

A CONCORDÂNCIA PLURAL VARIÁVEL NO SINTAGMA NOMINAL DO PORTUGUÊS  
REESTRUTURADO DA COMUNIDADE DE ALMOXARIFE, SÃO TOMÉ  
(Desenvolvimento das Regras de Concordância Variáveis no Processo de  
Transmissão-Aquisição Geracional)  
Vol. 1

por

**Carlos Filipe Guimarães Figueiredo**

Dissertação de Doutoramento em Linguística

2010



FACULDADE DE CIÊNCIAS SOCIAIS E HUMANAS

DEPARTAMENTO DE PORTUGUÊS

**UNIVERSIDADE DE MACAU**

A CONCORDÂNCIA PLURAL VARIÁVEL NO SINTAGMA NOMINAL DO PORTUGUÊS  
REESTRUTURADO DA COMUNIDADE DE ALMOXARIFE, SÃO TOMÉ  
(Desenvolvimento das Regras de Concordância Variáveis no Processo de  
Transmissão-Aquisição Geracional)  
Vol. 1

por

**Carlos Filipe Guimarães Figueiredo**

**Orientador:** Professor Doutor Alan Norman Baxter

Departamento de Português

Dissertação em Linguística para obtenção do grau de Doutor.

2010

Faculdade de Ciências Sociais e Humanas

UNIVERSIDADE DE MACAU

Copyright 2010 by Guimarães Figueiredo, Carlos Filipe  
University of Macau

Ao meu pai, que partiu em vésperas de eu me licenciar.

À minha mãe, exemplo ímpar de sacrifício e dedicação.

À Nani e ao Diogo, pela felicidade que trazem à minha existência.

Ao meu irmão Rui e aos meus filhos José Manuel e Ana Cátia.

À minha sogra Maria Ivone.



## **Agradecimentos**

Durante o período de tempo em que decorreu o presente estudo, fui acumulando dívidas, impossíveis de saldar. A primeira destas é para com todas as minhas fontes de informação, as quais permitiram que partilhasse das suas vidas e dos seus conhecimentos. De entre as fontes profissionais, tive o privilégio de cultivar amizade com algumas. Sem estabelecer prioridades, gostaria de referir o grande apreço e profundo sentimento de gratidão para com o meu orientador, Prof. Alan Norman Baxter, que, paciente e tolerantemente, dedicou horas sem fim a revisões, permitindo-me entender os objectivos e caminhos do rigor científico.

A segunda dívida de gratidão é para com a Universidade de Macau, que não só me concedeu a oportunidade para encetar um novo caminho na minha vida profissional mas também me proporcionou apoio e auxílio para levar a cabo a presente pesquisa. De entre o corpo profissional da Universidade, um agradecimento especial é para a Prof. Maria Antónia Espadinha, que permitiu que fosse acolhido no seio desta família e me deu estímulo antecipado, ajudando a lançar este projecto. Os agradecimentos estendem-se ao Hugo Cardoso, pelas sugestões e apoio na revisão de texto, bem como ao Aldino Dias, pelo incentivo e ajuda na revisão de texto. Quero também agradecer a todos aqueles que, de uma forma ou de outra, se interessaram pelo andamento da pesquisa. De entre estes, não posso deixar de citar a Paula Campos, a Sara Gonçalves, a Ana Paula Godinho, a Leonor Seabra, a Zhang Jing, o Yao Jing Ming e a Gao Lili, pela amizade e sentido de cooperação.

Mas as minhas dívidas intelectuais vão para além do corpo docente da Universidade de Macau. Esta pesquisa é uma excrescência do gosto pela Linguística, que nasceu durante o percurso de obtenção da minha licenciatura. O fascínio pelas variedades de português estava no gene, sem que me apercebesse, e palpitava de cada vez que, em Angola, ao abandonar a sala-de-aulas da escola primária, mudava o tipo de discurso para comunicar, de modo “genuíno”, com os colegas, fosse no pátio do recreio fosse nas brincadeiras de rua. As vicissitudes da vida determinaram que tivesse de interromper os estudos, quando aportei em Portugal, integrado no grosso dos “retornados” que buscava, na metrópole, alternativas de sobrevivência, finda que estava a aventura portuguesa da expansão colonial. Lembro-me, então, como me soava a estranho o idioma que

aprendera nas carteiras da escola primária, de cada vez que contactava com falantes das diversas regiões de Portugal. Retomando os estudos, já em fase em que começavam a despontar os primeiros cabelos brancos, o incentivo dos Profs. Malaca Casteleiro e Maria José Grosso foi determinante para que reavivasse em mim o interesse em tentar perceber a riqueza encerrada na “minha” variedade de português e entender que a mesma poderia dar contributo valioso para os estudos da aquisição, em particular, e da Linguística, em geral. Para ambos, deixo registada uma palavra de apreço muito especial.

Ao ser integrado no quadro profissional da Universidade de Macau, estava-me reservada a grata surpresa de poder dar continuidade a dois percursos que me completam pessoal e profissionalmente: exercer a carreira de docente, contactando com alunos que me ajudam a descobrir a profundidade de sentimentos e me animam para me envolver em projectos didácticos; desenvolver trabalhos de pesquisa sobre temas que me fascinam e que podem dar contributo valioso para os estudos em Linguística, para além de permitirem que contacte com especialistas que me possibilitam um cada vez maior crescimento intelectual.

A nível estreitamente pessoal, tenho ainda mais dívidas que nunca poderei saldar. Para com os meus pais, Céu e Amílcar, ímpares nos exemplos de rectidão humana e verdadeiros azimutes para me indicarem o trajecto da persistência, sacrifício e esforço. Para com o meu irmão Rui, sempre na vanguarda para me escudar nos momentos mais difíceis. Para com a minha sogra, Maria Ivone, nunca ausente com o seu estímulo e apoio. Para com o José Manuel e a Cátia, que me acolherem e aceitaram no seu seio como pai, ajudando-me a preencher a minha vida.

Para com a Nani, a minha dívida impagável deve-se ao sofrimento em silêncio, à paciência e à tolerância pelas minhas horas de “ausência”, dedicadas ao presente estudo. Quanto ao Diogo, a dívida tem a ver com a adolescência que não sentiu a minha presença diária, devido também ao tempo votado a este projecto. Vocês são, ambos, a magia da minha vida, o meu porto de acolhimento, o meu tudo.

## RESUMO

A pesquisa analisa a aquisição de marca de plural (PL) no sintagma nominal (SN) de uma variedade de português reestruturado, falada pela comunidade bilingue (crioulo e português) de Almojarife, São Tomé. Numa perspectiva da sociolinguística variacionista, são feitas comparações com variedades africanas de português que têm substratos do grupo níger-congo atlântico (o português vernacular de Angola, o português do Maputo, Moçambique, e o português de uma comunidade de descendentes de serviçais na ilha de São Tomé), o português cabo-verdiano, uma variedade brasileira de português rural e variedades brasileiras de português urbano. O estudo recorreu ao pacote GOLDVARB X, que permite identificar as variáveis linguísticas e extralinguísticas que influenciam a marcação plural nos constituintes do SN do português de Almojarife (PA), determinando a natureza da concordância plural variável (CPL-var) a nível comunitário. A análise estrutural da concordância no SN é sustentada por um quadro teórico heterogéneo da área da aquisição de primeiras línguas (L1's), aquisição bilingue e aquisição de L2's (ASL), enfocando, particularmente, o papel da Gramática Universal (GU) na ASL, bem como os preceitos da Morfologia Distribuída (MD) acerca da marcação variável nos elementos do SN. Comparando o português com o santomense, o crioulo de Cabo Verde (CCV) e certas línguas do grupo níger-congo atlântico, determina-se o impacto das características tipológicas das L1's tanto na interlíngua como na posterior língua nativizada que emerge em situação de transmissão linguística irregular, visando explicar a forma como ambos os sistemas podem ser modelados por aspectos linguísticos, psico-cognitivos e sociohistóricos. Uma das propostas do trabalho, baseada na Teoria dos Princípios e Parâmetros, é a hipótese de ocorrerem paralelismos na aquisição da marca de PL e desenvolvimento da CPL no SN das L1's de falantes monolíngues e bilingues, por um lado, e das L2's, por outro lado. Assume-se também que os falantes adultos de L2's acedem parcialmente à GU e que os parâmetros desactivados após o período crítico da aquisição não são totalmente reinstanciados. Os valores obtidos revelam que a CPL-var é motivada, essencialmente, por condicionantes do tipo estrutural, não se vislumbrando uma influência significativa das variáveis do tipo fonológico e semântico no desenho da mesma. A variável *classe gramatical* revelou que, em SN's de estrutura reduzida ocorre relação entre posição e



categoria gramatical: determinantes e primeira posição (posição pré-nuclear), nomes e segunda posição (núcleo) e adjetivos e terceira posição (posição pós-nuclear). Quanto às variáveis extralinguísticas, identificaram gramáticas em competição, com a dos falantes mais idosos evidenciando proximidade à gramática do substrato e a das gerações mais novas caminhando no sentido de uma ligeira aquisição das regras da gramática da língua-alvo (LA). A comparação dos resultados do PA com os das outras variedades africanas e brasileiras de português permitiu concluir que o português adquirido em situação de contacto por falantes de um crioulo português (CP) atlântico manifesta os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular, na qual a língua ancestral não é um CP atlântico.

## ABSTRACT

This study focuses on the acquisition of plural (PL) marking in the noun phrase of a variety of restructured Portuguese spoken by the bilingual (Creole – Portuguese) community of Almojarife (São Tomé). Within a variationist sociolinguist perspective, comparisons are made with African varieties of Portuguese which have Niger-Congo Atlantic substrates (Angola vernacular Portuguese, Maputo Portuguese in Mozambique, and the Portuguese of descendents of indentured Africans in São Tomé), the Portuguese of Cape Verde, as well as one rural and two urban varieties of Brazilian Portuguese. The study of variation in PA makes use of the GOLDVARB X package, which permits the identification of linguistic and extralinguistic variables influencing plural marking of noun phrase constituents, thereby determining the nature of variable plural agreement at a community level. The structural analysis of noun phrase plural agreement is rooted in a heterogeneous theoretical framework within the domain of first language (L1) acquisition, bilingual acquisition, and second language acquisition (SLA), with particular emphasis on the role of Universal Grammar (UG) in SLA, as well as the tenets of Distributed Morphology (DM) with regard to variable marking of NP constituents. Comparisons of Portuguese with Santomense, Cape Verdean Creole (CVC) and certain Niger-Congo Atlantic languages establish the impact of L1 and its typological characteristics on both the learners' interlanguage and the nativised language in a context of *imperfect language shift*, in order to explain how both systems can be shaped by linguistic, psycho-cognitive and sociohistorical aspects. One of the working hypotheses is that the acquisition of plural marking parallels the development of NP number agreement in monolingual/bilingual L1's and L2's, based on the Theory of Principles and Parameters. It is also assumed that adult L2 speakers access UG partially and that parameters deactivated after the critical period are not entirely reset.

The values obtained demonstrate that PLA-var is essentially motivated by structural constraints, since phonological and semantic variables appear not to intervene significantly in shaping it. The *grammatical class* variable revealed that, in simple NP's, there is a correlation between position and grammatical category: determiners and first (pre-head) position, nouns and second (head) position, adjectives and third (post-head) position. Extra-linguistic variables, on the other hand, unearthed competing grammars:

that of the older speakers approaching substrate models, and that of the younger generations moving towards partial acquisition of target-language (LA) grammatical rules. The systematic comparison of PA results with those for other African and Brazilian varieties of Portuguese led to the conclusion that the Portuguese acquired in contact situations by speakers of an Atlantic Portuguese-based Creole displays the same patterns of variation as the Portuguese acquired in an *imperfect language shift* setting for which the ancestral language is not an Atlantic Portuguese-based Creole.

## Índice

### Lista de abreviaturas

### Lista de tabelas

### Lista de figuras

<b>Introdução</b> .....	1
<b>Capítulo 1. O tema</b> .....	11
1.1. Português do Brasil e concordância nominal variável de número .....	11
1.2 Variedades africanas de português e concordância nominal variável de número .....	31
1.2.1. Variedades africanas de português com substrato do grupo níger-congo atlântico ...	31
1.2.1.1. Português dos tongas (Roça Monte Café, São Tomé) .....	31
1.2.1.2. Português vernáculo de Angola .....	38
1.2.1.3. Português de Moçambique .....	40
1.2.2. Variedades africanas de português com substrato crioulo .....	45
1.2.2.1. Português caboverdiano e seu substrato (crioulo de Cabo Verde) .....	45
1.2.2.2. Português de Almojarife (São Tomé) e seu substrato (santomense) .....	59
1.2.2.3. A comunidade de Almojarife .....	85
1.2.2.3.1. Características da fala de Almojarife .....	90
1.3. Conclusões parciais .....	99
<b>Capítulo 2. Fundamentação teórica</b> .....	103
2.1. Os pressupostos teórico-metodológicos sobre aquisições de L1, bilingue e L2 .....	106
2.1.1. Variação e sistematicidade na <i>interlíngua</i> : perspectiva generativista .....	108
2.1.2. Variação e sistematicidade na <i>interlíngua</i> : perspectiva psicolinguística .....	113
2.1.3. Variação e sistematicidade na <i>interlíngua</i> : perspectiva sociolinguística .....	114
2.1.4. Variação e sistematicidade na <i>interlíngua</i> : aplicação da perspectiva heterogénea ao estudo do PA .....	116
2.2. Abordagens teóricas sobre aquisição de L1 .....	123
2.2.1. Modelo generativista .....	124
2.2.2. A teorização em torno da GU .....	125
2.2.2.1. O Modelo de Princípios e Parâmetros ( <i>Parameter settings</i> ) .....	125
2.2.2.1.2. Teoria das Regências e Ligações .....	127
2.2.3. Papel da GU na aquisição da L1 .....	130
2.2.4. GU e estudos de aquisição .....	135
2.3. Abordagens teóricas sobre aquisição bilingue .....	137

2.3.1. Acesso ao léxico por parte do bilingue .....	138
2.3.2. Papel da GU na aquisição bilingue: a Hipótese da Diferenciação da Linguagem .....	141
2.4. Abordagens teóricas sobre aquisição de L2's .....	145
2.4.1. O modelo generativista e papel da GU na aquisição de L2's .....	146
2.4.1.1. GU e teorias em torno da aquisição de L2's .....	147
2.4.2. Aquisição de L2's e modelos psicocognitivos .....	152
2.4.2.1. Conexionismo ( <i>Connectionism</i> ) e modelos de processamento paralelo distribuído .....	153
2.4.2.2. Teoria da "Processabilidade" ( <i>Processability Theory</i> ) .....	154
2.4.2.3. Modelo dos Quatro Morfemas ( <i>The 4M Model</i> ) .....	160
2.4.3. Modelo da mudança evolucionista da língua .....	161
2.5. Variação: pressupostos teóricos .....	166
2.5.1. Variáveis extralinguísticas, variação e mudança .....	169
2.5.2. Variáveis linguísticas e contexto .....	172
2.5.3. Modelos matemáticos para reprodução do efeito global dos dados empíricos .....	174
2.6. Concordância de número PL a nível do SN .....	178
2.6.1. Posicionamento e flexão em número dos constituintes do SN: visões descritivistas .....	178
2.6.2. O quadro teórico da concordância .....	186
2.6.2.1. Concordância e núcleos funcionais .....	187
2.6.3. O SN no português europeu .....	195
2.6.4. Categorias funcionais e categorias lexicais .....	197
2.6.4.1. Aquisição do sintagma determinante .....	201
2.6.4.2. Aquisição de género e número .....	206
2.6.5. Aquisição do sintagma determinante em L2 .....	214
2.6.6. Aquisição da concordância de PL no sintagma determinante .....	218
2.7. Aquisição das categorias funcionais da L2: os casos do santomense e do PA .....	225
2.8. A refixação dos parâmetros .....	229
2.9. Concordância variável ao nível do SN .....	231
2.10. Fossilização .....	239
2.10.1. Fossilização e variação .....	239
2.10.2. Potenciais causas de fossilização .....	242
2.10.3. GU e fossilização .....	244
2.10.4. Fossilização no PA .....	246

2.10.4.1. PA: transferência e fossilização fonológica .....	246
2.10.4.2. PA: transferência e fossilização morfossintáctica .....	250
2.11. Aquisição massiva de L2's: línguas crioulas e teorias de aquisição de L2's .....	253
2.11.1. Universais linguísticos vs. transferência: a Hipótese da Interlíngua ( <i>Interlanguage Hypothesis</i> ) .....	253
2.11.1.1. Teoria da “Processabilidade” e flexão morfológica das línguas emergentes por contacto .....	254
2.11.1.2. Teoria da “Processabilidade” e processamento sintáctico das línguas emergentes por contacto .....	257
2.11.2. Hipótese do Desenvolvimento Moderado da Transferência ( <i>Developmentally Moderated Transfer Hypothesis</i> ) .....	260
2.11.3. Transmissão linguística irregular .....	263
2.11.3.1. Vectores para a emergência da transmissão linguística irregular .....	264
2.11.3.1.1. Sociohistória e transmissão linguística irregular .....	267
2.11.3.1.2. Transmissão linguística irregular: parâmetros linguísticos .....	269
2.11.3.1.3. Transmissão linguística irregular e gramaticalização: heterogénese GU/transferência .....	273
2.11.3.2. Variedades crioulizantes .....	279
2.11.3.3. Transmissão linguística irregular e variação: gramáticas em competição .....	282
2.12. Princípio da coesão estrutural e variação a nível do sintagma determinante .....	294
2.13. Conclusões parciais .....	295
<b>Capítulo 3. A metodologia</b> .....	299
3.1. Aplicação dos métodos quantitativos no estudo da aquisição de línguas .....	299
3.1.1. Recolha e transcrição dos dados do <i>corpus</i> .....	302
3.1.2. Tratamento dos dados .....	307
3.1.3. Os dados dos informantes .....	307
3.1.4. Caracterização dos dados .....	309
3.2. O suporte computacional .....	312
3.3. Descrição da variável dependente e das variáveis independentes .....	315
3.3.1. Variável dependente .....	316
3.3.2. Variáveis independentes .....	317
3.4. Codificação dos dados .....	319
3.5. Variável dependente constituída .....	321
3.6. Variáveis independentes constituídas .....	322

3.6.1. Variável independente <i>saliência fónica</i> e suas subvariáveis .....	322
3.6.1.1. <i>Saliência fónica</i> .....	322
3.6.1.2. <i>Tonicidade (Tonicidade dos itens lexicais singulares)</i> .....	330
3.6.2. Variável independente <i>marcas precedentes ao item analisado</i> .....	334
3.6.3. Variável independente <i>contexto fonológico posterior</i> .....	338
3.6.4. Variável independente <i>classe gramatical do item analisado</i> .....	343
3.6.4.1. Constituição e definição dos factores para a variável <i>classe gramatical</i> .....	345
3.6.5. Variável independente <i>posição do item analisado em relação ao núcleo do SN</i> .....	350
3.6.6. Variável independente <i>posição do item na cadeia do SN (Posição linear)</i> .....	354
3.6.7. Variável independente <i>marcação de género</i> .....	359
3.6.8. Variável independente <i>ordem do item na cadeia dos constituintes flexionáveis em número</i> .....	362
3.6.9. Variável independente <i>grau de concordância de número no SN</i> .....	364
3.6.10. Variável independente <i>traço semântico do SN (Animacidade)</i> .....	365
3.6.11. Variável independente <i>sexo</i> .....	368
3.6.12. Variável independente <i>idade</i> .....	371
3.6.13 Variável independente <i>estadia (Permanência fora da comunidade)</i> .....	372
3.6.14. Variável independente <i>escolaridade</i> .....	374
3.7. Codificação dos itens de acordo com os códigos atribuídos aos factores .....	377
3.8. Quantificação dos dados para selecção das variáveis independentes e dos factores a serem incorporados na análise final .....	378
3.8.1. Iteração entre grupos de factores .....	381
3.8.2. Variáveis independentes a serem incorporadas na análise final .....	388
<b>Capítulo 4.</b> Análise dos resultados: variável dependente e variáveis independentes do tipo estrutural .....	405
4.1. Resultados das variáveis constituídas e respectivas análises .....	407
4.1.1. Variável dependente .....	407
4.1.2. Variáveis independentes do tipo estrutural .....	409
4.1.2.1. Variável independente <i>posição em relação ao núcleo do SN</i> .....	410
4.1.2.1.1. Relação entre <i>classe gramatical</i> e <i>posição em relação núcleo do SN</i> .....	411
4.1.2.2. Variável independente <i>ordem do item na cadeia dos constituintes flexionáveis</i> .....	415
4.1.2.2.1. Relação entre <i>ordem do item na cadeia dos constituintes flexionáveis</i> e <i>posição linear</i> .....	417
4.1.2.3. Variável independente <i>classe gramatical</i> .....	420

4.1.2.4. Variável independente <i>marcas precedentes ao item analisado</i> .....	424
4.1.2.4.1. <i>Marcas precedentes</i> : influência dos factores individuais na marcação PL do elemento analisado .....	435
4.1.2.4.2. <i>Marcas precedentes</i> : contribuição dos factores amalgamados para a marcação PL elemento analisado .....	446
4.1.2.4.2.1. Factores amalgamados para estudo do efeito das <i>marcas precedentes</i> na marcação PL do elemento analisado: análise I .....	447
4.1.2.4.2.2. Factores amalgamados para estudo do efeito das <i>marcas precedentes</i> na marcação PL do elemento analisado: análise II .....	453
4.1.2.4.2.3. Factores amalgamados para estudo do efeito das <i>marcas precedentes</i> na marcação PL do elemento analisado: análise III .....	457
4.1.2.4.3. Relação entre <i>marcas precedentes</i> , <i>posição linear</i> e <i>classe gramatical</i> na marcação PL dos itens do SN .....	462
4.1.2.4.4. Análises alternativas considerando <i>marcas precedentes</i> , <i>posição linear</i> e <i>classe gramatical</i> .....	464
4.1.2.4.5. <i>Faixas etárias</i> e perfil de marcação motivado pelas <i>marcas precedentes</i> .....	470
4.1.2.4.6. <i>Marcas precedentes</i> representadas por marcador semântico .....	475
4.1.2.5. Variável independente <i>grau de concordância de número no SN</i> .....	483
4.1.2.6. Variável independente <i>posição do item na cadeia do SN (Posição linear)</i> .....	491
4.1.2.6.1. Relação entre <i>posição linear</i> e <i>classe gramatical</i> na marcação PL dos itens do SN .....	494
4.1.2.6.2. Relação entre <i>posição linear</i> e <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN .....	503
<b>Capítulo 5. Análise dos resultados: variáveis independentes do tipo semântico e fonológico; variáveis extralinguísticas</b> .....	537
5.1. Variável independente do tipo semântico .....	538
5.1.1. Variável independente <i>traço semântico do SN (Animacidade)</i> .....	538
5.2. Variáveis independentes do tipo fonológico .....	541
5.2.1. Variável independente <i>saliência fónica (Processos morfofonológicos de formação de PL)</i> .....	541
5.2.1.1. Relação entre <i>saliência fónica</i> e <i>escolaridade</i> na marcação PL do PA .....	548
5.2.2. Variável independente <i>tonicidade (Tonicidade dos itens lexicais singulares)</i> .....	553
5.2.2.1. Relação entre <i>tonicidade</i> e <i>escolaridade</i> na marcação PL do PA e do MRJ ....	555
5.2.2.2. Efeito da <i>saliência (saliência fónica x tonicidade)</i> na marcação PL dos itens do SN do PA .....	557
5.2.2.3. Efeito da <i>saliência (saliência fónica x tonicidade)</i> na marcação PL do PA e da NURC .....	561



5.2.2.4. Relação entre <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) e escolaridade (PA e NURC) .....	564
5.2.2.5. Efeito da <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) na marcação PL do PA e do PT .....	567
5.2.2.6. Relação entre <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) e <i>idade</i> na marcação PL dos itens do SN do PA e do PT .....	570
5.2.2.7. Efeito da <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) na marcação PL dos itens do SN: 3 variedades de português .....	576
5.2.3. Variável independente <i>contexto fonológico posterior</i> .....	581
5.2.3.1. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo os traços de fonte .....	584
5.2.3.2. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando as consoantes num único factor .....	588
5.2.3.3. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando as pausas num único factor .....	591
5.2.3.4. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo os traços de classe principal .....	593
5.2.3.5. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo os traços coronal e não-coronal .....	596
5.2.3.6. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo o modo de articulação: 6 factores .....	599
5.2.3.7. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo o modo de articulação: 7 factores .....	601
5.2.3.8. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo as zonas de articulação .....	605
5.2.3.9. Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN, englobando consoantes segundo os pontos de articulação .....	608
5.2.3.10. Efeito do <i>contexto fonológico posterior</i> na marcação PL da <i>saliência fónica</i> .....	612
5.2.3.10.1. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo os traços de fonte, para a marcação PL dos itens do SN .....	613
5.2.3.10.2. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo os traços de classe principal, para a marcação PL dos itens do SN .....	627
5.2.3.10.3. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo os traços coronal e não-coronal, para a marcação PL dos itens do SN .....	629
5.2.3.10.4. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo o modo de articulação, para a marcação PL dos itens do SN: 6 factores .....	631

5.2.3.10.5. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo o modo de articulação, para a marcação PL dos itens do SN: 7 factores .....	633
5.2.3.10.6. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo as zonas de articulação, para a marcação PL dos itens do SN .....	635
5.2.3.10.7. Efeito do <i>contexto fonológico posterior</i> na <i>saliência fónica</i> , englobando consoantes segundo os pontos de articulação, para a marcação PL dos itens do SN .....	637
5.2.3.11. Efeito das <i>marcas precedentes</i> na <i>saliência fónica</i> .....	640
5.3 Variáveis extralinguísticas .....	652
5.3.1. Variável independente <i>escolaridade</i> .....	653
5.3.2. Variável independente <i>idade</i> .....	663
5.3.3. Variável independente <i>estadia (Permanência fora da comunidade)</i> .....	671
5.3.4. Variável independente <i>sexo</i> .....	674
<b>Conclusões finais</b> .....	681
Referências bibliográficas .....	711
I. Bibliografia consultada e citada .....	711
II. Bibliografia pertinente consultada .....	753
III. Sítios pertinentes consultados .....	785
<i>Curriculum vitae</i> do autor .....	789



## Lista de abreviaturas

<b>ADJ</b>	Adjectivo
<b>Adj</b>	Adjunto
<b>ADV</b>	Advérbio
<b>AFX</b>	Afixo
<b>AHM</b>	Arquivo Histórico Militar (Lisboa)
<b>AP</b>	Complemento pré-nominal
<b>ART</b>	Artigo
<b>AISL</b>	Aquisição imperfeita de segundas línguas
<b>ASL</b>	Aquisição de segundas línguas
<b>AUX</b>	Verbo auxiliar
<b>CCV</b>	Crioulo de Cabo Verde
<b>cf.</b>	Conforme
<b>CGB</b>	Crioulo da Guiné Bissau
<b>CGEN-var</b>	Concordância variável de género
<b>CGG</b>	Crioulo do Golfo da Guiné
<b>Cod.</b>	Código
<b>COD's</b>	Construções com Objecto Duplo
<b>COMP</b>	Complementador; marcador completivo; marcador comparador
<b>Conc.</b>	Concordância
<b>CP</b>	Crioulo de base portuguesa
<b>CPLP</b>	Comunidade dos Países de Língua Portuguesa
<b>CPL-var</b>	Concordância plural variável
<b>CPrep's</b>	Construções com Preposições
<b>CVS's</b>	Construções com Verbos Seriais
<b>DEF</b>	Definido
<b>DEM</b>	Demonstrativo
<b>DE'S</b>	Descrições estruturais
<b>DET</b>	Determinante
<b>DLP's</b>	Dados linguísticos primários ou <i>triggers</i>
<b>DOC</b>	Construções com duplo objecto
<b>doc.</b>	Documento
<b>DOC.</b>	Documentalista

<b>EP</b>	Espanhol do Panamá
<b>EPR</b>	Espanhol de Porto Rico
<b>Esp</b>	Espanhol
<b>Estrutura-P</b>	Estrutura profunda
<b>Estrutura-S</b>	Estrutura de superfície
<b>ex.</b>	Exemplo
<b>FE-1</b>	Faixa etária 1 (20 – 40 anos)
<b>FE-2</b>	Faixa etária 2 (41 – 60 anos)
<b>FE-3</b>	Faixa etária 3 (+ 60 anos)
<b>Fem.</b>	Feminino
<b>FFFH</b>	<i>Failed Functional Features Hypothesis</i> (Falha do Restabelecimento dos Parâmetros)
<b>Fig.</b>	Figura
<b>fl.</b>	Folha
<b>FL</b>	Forma Lógica
<b>FLEX</b>	Flexão
<b>FL0</b>	Estado inicial do processo de aquisição de uma língua
<b>FLf</b>	Estado final do processo de aquisição de uma língua
<b>FTFA</b>	<i>Full Transfer Full Access Hypothesis</i> (Acesso Total aos Princípios e Parâmetros)
<b>FUT</b>	Futuro
<b>GEN</b>	Género
<b>GU</b>	Gramática Universal
<b>HEL-Ba</b>	Dialecto de Helvécia, Bahia (Brasil)
<b>HI</b>	Hipótese da Interlíngua
<b>I</b>	Sintagma entoacional
<b>i.e.</b>	Isto é
<b>IFX</b>	Infixo
<b>IL</b>	Interlíngua
<b>IMP</b>	Imperfectivo
<b>IND</b>	(Presente do) Indicativo
<b>INDF</b>	Indefinido
<b>ININT</b>	Ininteligível
<b>INF</b>	Informante

<b>INT</b>	Interrogativo(a); interrogação
<b>L</b>	Língua
<b>L1</b>	Língua primeira
<b>L2</b>	Língua segunda
<b>LA</b>	Língua-alvo
<b>LDB</b>	Lei de Directrizes e Bases da Educação
<b>LCV</b>	Língua Caboverdiana
<b>LE</b>	Língua estrangeira
<b>LF</b>	Língua-fonte
<b>LM</b>	Língua materna
<b>LO</b>	Língua de origem
<b>LOC</b>	Locativo
<b>Masc.</b>	Masculino
<b>MD</b>	Morfologia distribuída
<b>MLF</b>	<i>Matrix Language Frame Model</i>
<b>MLU</b>	<i>Mean Length of Utterance</i> (Comprimento Médio do Enunciado)
<b>MRJ</b>	Dialecto do Município do Rio de Janeiro (Brasil)
<b>MSIH</b>	<i>Missing Surface Inflection Hypothesis</i>
<b>MSpec</b>	Especificador de marcação
<b>N</b>	Nome; substantivo
<b>NEG</b>	Negação; negative
<b>NUM</b>	Numeral
<b>NURC</b>	Norma urbana culta (Brasil)
<b>OD</b>	Objecto directo
<b>OInd</b>	Objecto indirecto
<b>OP</b>	Oração pequena
<b>p.</b>	Página
<b>PA</b>	Português de Almojarife (São Tomé)
<b>PB</b>	Português do Brasil
<b>PCV</b>	Português de Cabo Verde
<b>p.e.</b>	Por exemplo
<b>PE</b>	Português europeu
<b>Pe.</b>	Padre

<b>PEH</b>	Princípio do Espelho
<b>PERF</b>	Perfectivo
<b>PF</b>	Forma Fonológica
<b>PEX</b>	Prefixo
<b>PIL</b>	Princípio da Integridade Lexical
<b>PL</b>	Plural
<b>PLd</b>	Português de Luanda (Angola)
<b>PM</b>	Português de Moçambique
<b>PMn</b>	Português de Minas Gerais (Brasil)
<b>PMp</b>	Português do Maputo (Moçambique)
<b>POP</b>	Grupo do português popular
<b>POS</b>	Possessivo
<b>pp.</b>	Páginas
<b>PP</b>	Pronome pessoal
<b>PPA</b>	Português de Porto Alegre, Brasil
<b>pr.</b>	Peso relativo
<b>proc.</b>	Processo
<b>PREP</b>	Preposição
<b>PRET</b>	Pretérito
<b>PRO</b>	Pronome
<b>PROG</b>	Progressivo
<b>PST</b>	Português de São Tomé
<b>PT</b>	Português dos tongas (São Tomé)
<b>Ptg</b>	Português
<b>PtgLA</b>	Português língua-alvo
<b>PtgL1</b>	Português língua primeira
<b>PtgL2</b>	Português língua segunda
<b>PtgLE</b>	Português língua estrangeira
<b>PtgLM</b>	Português língua materna
<b>PVA</b>	Português vernáculo de Angola
<b>PVB</b>	Português vernáculo do Brasil
<b>Q</b>	Quantificador
<b>REL</b>	Pronome relativo

<b>SADJ</b>	Sintagma adjectival
<b>SADV</b>	Sintagma adverbial
<b>SCOMPL</b>	Sintagma complementador
<b>SDET</b>	Sintagma determinante
<b>SFLEX</b>	Sintagma flexional
<b>SFX</b>	Sufixo
<b>Spec</b>	Especificador
<b>SPREP</b>	Sintagma preposicional
<b>SG</b>	Singular
<b>SN</b>	Sintagma nominal
<b>SNN</b>	Sintagma nominal nu
<b>SQ</b>	Sintagma quantificador
<b>SUJ</b>	Sujeito
<b>SV</b>	Sintagma verbal
<b>T</b>	Projecção máxima de informações de tempo
<b>Trad.</b>	Tradução; tradutor
<b>TMA</b>	Tempo, modo e aspecto
<b>UNI</b>	Grupo do português universitário





## Lista de tabelas

<b>Tabela</b>	<b>Título da tabela</b>	
Tabela 1.1	Informação encerrada nos SNN's do santomense .....	72
Tabela 2.1	Hierarquia hipotética da articulação dos processamentos na aquisição da interlíngua .....	158
Tabela 3.1	Articulação dos processamentos na aquisição da interlíngua inglesa .....	159
Tabela 4.1	Taxionomia dos factores motivadores da fossilização .....	246
Tabela 5.1	A assunção: similaridades estruturais configuram transferência .....	258
Tabela 6.1	Os factos: similaridades estruturais podem não representar transferência ....	259
Tabela 7.1	Inversão da ordem frásica na interlíngua alemã .....	259
Tabela 8.1	Informantes de Almoxarife: faixas etárias e seus aspectos sociais .....	309
Tabela 9.1	A variável dependente (análise atomística) .....	317
Tabela 10.1	Variáveis independentes (análise atomística) .....	319
Tabela 11.1	Factores constituídos para a variável dependente .....	321
Tabela 12.1	Factores constituídos para a variável independente <i>saliência fónica</i> .....	329
Tabela 12.2	Factores constituídos para a variável independente <i>tonicidade</i> .....	333
Tabela 12.3	Factores constituídos para a variável independente <i>marcas precedentes ao item analisado</i> .....	337
Tabela 12.4	Factores constituídos para a variável independente <i>contexto fonológico posterior</i> .....	342
Tabela 12.5	Factores constituídos para a variável independente <i>classe gramatical do item analisado</i> .....	349
Tabela 12.6	Factores constituídos para a variável independente <i>posição do item analisado em relação ao núcleo do SN</i> .....	354
Tabela 12.7	Factores constituídos para a variável independente <i>posição do item na cadeia do SN</i> .....	358
Tabela 12.8	Factores constituídos para a variável independente <i>marcação de género</i> ....	361
Tabela 12.9	Factores constituídos para a variável independente <i>ordem do item na cadeia dos constituintes flexionáveis em número</i> .....	363
Tabela 12.10	Factores constituídos para a variável independente <i>grau de concordância de número no SN</i> .....	364
Tabela 12.11	Factores constituídos para a variável independente <i>traço semântico do SN</i> .	368
Tabela 12.12	Factores constituídos para a variável independente <i>sexo</i> .....	370
Tabela 12.13	Factores constituídos para a variável independente <i>idade</i> .....	372
Tabela 12.14	Factores constituídos para a variável independente <i>estadia</i> .....	373
Tabela 12.15	Factores constituídos para a variável independente <i>escolaridade</i> .....	377

Tabela 13.1	Exemplo de cadeia de códigos resultante da codificação dos elementos do SN .....	378
Tabela 14.1	Escala hierárquica decrescente das variáveis independentes seleccionadas como válidas para análise da CPL-var no SN do PA .....	384
Tabela 15.1	<i>Variável dependente</i> : percentuais de marcas plural nos itens do SN do PA .	407
Tabela 15.2	Percentuais de SN's plenamente marcados no PA .....	407
Tabela 15.3	Percentuais de SN's plenamente marcados: 2 variedades de português .....	408
Tabela 15.4	Percentuais de marcas PL nos itens do SN: 7 variedades de português .....	409
Tabela 16.1	Efeito da <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN do PA .....	410
Tabela 16.2	Relação entre <i>classe gramatical</i> e <i>posição em relação ao núcleo do SN</i> : distribuição das classes gramaticais relativamente ao núcleo do SN .....	412
Tabela 17.1	Efeito da variável independente <i>ordem na cadeia dos constituintes flexionáveis</i> na marcação PL dos itens do SN .....	416
Tabela 17.2	Efeito da relação entre <i>posição linear</i> e <i>ordem do item na cadeia dos constituintes flexionáveis</i> na marcação PL dos itens do SN .....	418
Tabela 18.1	Efeito da <i>classe gramatical</i> na marcação PL dos itens do SN .....	420
Tabela 18.2	<i>Categoria gramatical</i> : efeito da <i>classe gramatical</i> na marcação PL no elemento analisado: 3 variedades de português .....	421
Tabela 19.1	Efeito das <i>marcas precedentes</i> na marcação PL dos itens do SN .....	426
Tabela 19.2	<i>Marcas precedentes</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – EPR e 5 variedades de português ....	436
Tabela 19.3	<i>Marcas precedentes</i> : configuração SS_ (item na terceira posição, duas marcas formais precedentes) – Informantes e grau de escolarização .....	443
Tabela 19.4	Efeito das <i>marcas precedentes</i> na marcação PL dos itens do SN do PA [Factores amalgamados: Análise I] .....	450
Tabela 19.5	<i>Marcas precedentes</i> : contribuição dos factores amalgamados para a marcação PL do elemento analisado – 3 variedades de português .....	452
Tabela 19.6	Efeito das <i>marcas precedentes</i> na marcação PL dos itens do SN do PA [Factores amalgamados: Análise II] .....	455
Tabela 19.7	<i>Marcas precedentes</i> : contribuição dos factores amalgamados para a marcação PL do elemento analisado – 2 variedades de português .....	455
Tabela 19.8	Efeito das <i>marcas precedentes</i> na marcação PL dos itens do SN do PA [Factores amalgamados: Análise III] .....	458
Tabela 19.9	<i>Marcas precedentes</i> : contribuição dos factores amalgamados para a marcação PL do elemento analisado – 3 variedades de português .....	460
Tabela 19.10	Efeito das <i>marcas precedentes</i> na marcação PL dos itens do SN do PA: <i>faixas etárias</i> .....	473

Tabela 20.1	Efeito do <i>grau de concordância de número no SN</i> na marcação PL dos itens do SN do PA .....	484
Tabela 20.2	Efeito do <i>grau de concordância de número no SN</i> no percentual de marcação PL dos itens do SN: comunidade e faixas etárias .....	488
Tabela 21.1	Grupo de factores <i>posição do item na cadeia do SN</i> e respectivos factores, para estudo da configuração do SN PL do PA. [Grupo de factores a testar para incorporação na análise de dados] .....	492
Tabela 21.2	Teste de significância estatística dentro do grupo de factores <i>posição do item na cadeia do SN</i> .....	492
Tabela 21.3	Efeito da <i>posição linear</i> na marcação PL dos itens do SN do PA .....	493
Tabela 21.4	<i>Posição linear</i> : contribuição dos factores individuais para a marcação PL do elemento analisado: 2 variedades de português .....	594
Tabela 21.5	Relação entre <i>posição linear</i> e <i>classe gramática</i> : distribuição das classes gramaticais pelas diversas posições do SN .....	595
Tabela 21.6	Efeito da relação entre <i>posição linear</i> e <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN do PA: total de ocorrências e percentuais de marcação .....	504
Tabela 21.7	Efeito da relação entre <i>posição linear</i> e <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN do PA .....	504
Tabela 21.8	Relação entre <i>posição linear</i> e <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN do PA: comunidade e faixas etárias .....	507
Tabela 21.9	Relação entre <i>posição linear</i> e <i>posição em relação ao núcleo do SN</i> na marcação PL dos itens do SN: 6 variedades de português .....	514
Tabela 22.1	Efeito do <i>traço semântico do SN</i> na marcação PL dos itens do SN do PA ...	538
Tabela 22.2	<i>Traço semântico do SN</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – 3 variedades de português .....	539
Tabela 23.1	Efeito da <i>saliência fónica</i> na marcação PL dos itens do SN do PA .....	541
Tabela 23.2	<i>Saliência fónica</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – 7 variedades de português .....	544
Tabela 23.3	Relação entre <i>saliência fónica</i> e <i>escolaridade</i> na marcação PL dos itens do SN do PA .....	550
Tabela 24.1	Efeito da <i>tonicidade</i> na marcação PL dos itens do SN do PA .....	553
Tabela 24.2	<i>Tonicidade</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – 3 variedades de português .....	554
Tabela 24.3	Relação entre <i>tonicidade</i> e <i>escolaridade</i> na marcação PL dos itens do SN do PA .....	556
Tabela 24.4	<i>Saliência</i> resultante do cruzamento entre <i>saliência fónica</i> e <i>tonicidade</i> (todos os dados): grupo de factores constituído consoante metodologia de Lopes .....	559

Tabela 24.5	Efeito da <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) na marcação PL dos itens do SN do PA: grupo de factores constituído consoante metodologia de Lopes .....	560
Tabela 24.6	<i>Saliência</i> ( <i>saliência fónica x tonicidade</i> ): contribuição dos factores individuais para a marcação PL no elemento analisado – 3 variedades de português .....	562
Tabela 24.7	Relação entre <i>saliência</i> e <i>escolaridade</i> na marcação PL dos itens do SN do PA: grupo de factores constituído consoante metodologia de Lopes .....	565
Tabela 24.8	<i>Saliência</i> resultante do cruzamento entre <i>saliência fónica</i> e <i>tonicidade</i> (todos os dados): grupo de factores constituído consoante metodologia de Baxter .....	568
Tabela 24.9	Efeito da <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) na marcação PL dos itens do SN do PA: grupo de factores constituído consoante metodologia de Baxter .....	568
Tabela 24.10	<i>Saliência</i> ( <i>saliência fónica x tonicidade</i> ): contribuição dos factores individuais para a marcação PL do elemento analisado – 2 variedades de português .....	569
Tabela 24.11	Relação entre <i>saliência</i> e <i>idade</i> na marcação PL dos itens do SN do PA: grupo de factores constituído consoante metodologia de Baxter .....	572
Tabela 24.12	<i>Saliência</i> resultante do cruzamento entre <i>saliência fónica</i> e <i>tonicidade</i> (todos os dados): grupo de factores constituído consoante metodologia de Jon-And .....	577
Tabela 24.13	Efeito da <i>saliência</i> ( <i>saliência fónica x tonicidade</i> ) na marcação PL dos itens do SN do PA: grupo de factores constituído consoante metodologia de Jon-And .....	577
Tabela 24.14	<i>Saliência</i> ( <i>saliência fónica x tonicidade</i> ): contribuição dos factores individuais para a marcação PL do elemento analisado: 4 variedades de português .....	579
Tabela 25.1	Efeito do <i>contexto fonológico posterior</i> na marcação PL dos itens do SN do PA .....	581
Tabela 25.2	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços de fonte) na marcação PL dos itens do SN .....	585
Tabela 25.3	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes num único factor) na marcação PL dos itens do SN .....	589
Tabela 25.4	Efeito do <i>contexto fonológico posterior</i> (englobando pausas num único factor), na marcação PL dos itens do SN .....	592
Tabela 25.5	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços de classe principal) na marcação PL dos itens do SN .....	594
Tabela 25.6	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços coronal e não-coronal) na marcação PL dos itens do SN .....	598
Tabela 25.7	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo o modo de articulação) na marcação PL dos itens do SN: 6 factores .....	600

Tabela 25.8	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo o modo de articulação) na marcação PL dos itens do SN: 7 factores .....	602
Tabela 25.9	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo as zonas de articulação) na marcação PL dos itens do SN .....	606
Tabela 25.10	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os pontos de articulação) na marcação PL dos itens do SN .....	611
Tabela 25.11	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços de fonte) na marcação PL da <i>saliência fónica</i> .....	613
Tabela 25.12	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços de classe principal) na marcação PL da <i>saliência fónica</i> .....	628
Tabela 25.13	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os traços coronal e não-coronal) na marcação PL da <i>saliência fónica</i> .....	630
Tabela 25.14	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo o modo de articulação) na marcação PL da <i>saliência fónica</i> : 6 factores .....	632
Tabela 25.15	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo o modo de articulação) na marcação PL da <i>saliência fónica</i> : 7 factores .....	634
Tabela 25.16	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo as zonas de articulação) na marcação PL da <i>saliência fónica</i> .....	635
Tabela 25.17	Efeito do <i>contexto fonológico posterior</i> (englobando consoantes segundo os pontos de articulação) na marcação PL da <i>saliência fónica</i> .....	637
Tabela 25.18	Efeito das <i>marcas precedentes</i> na marcação PL da <i>saliência fónica</i> .....	641
Tabela 25.19	Itens com singular em <i>-s</i> e <i>-r</i> (Marcação fonológico-geracional) .....	644
Tabela 26.1	Grupo de factores <i>escolaridade</i> e respectivos factores, para estudo da configuração do SN PL do PA: grupo de factores para teste de factores a incorporar na análise de dados .....	655
Tabela 26.2	Efeito da <i>escolaridade</i> na marcação PL dos itens do SN do PA .....	656
Tabela 26.3	<i>Escolaridade</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – 4 variedades de portugueses .....	658
Tabela 27.1	Efeito da variável <i>idade</i> na marcação PL dos itens do SN do PA .....	663
Tabela 27.2	SN's plenamente marcados no PA: valores por <i>idade</i> .....	664
Tabela 27.3	<i>Idade</i> : contribuição dos factores individuais para a marcação PL do elemento analisado – 6 variedades de portugueses .....	668
Tabela 28.1	Efeito da <i>estadia</i> na marcação PL dos itens do SN do PA .....	673
Tabela 29.1	Efeito da variável social <i>sexo</i> na marcação PL dos itens do SN do PA .....	674
Tabela 29.2	Marcação PL dos itens do SN do PA: valores por <i>idade</i> e <i>sexo</i> .....	675
Tabela 29.3	SN's plenamente marcados no PA: valores por <i>sexo</i> .....	676
Tabela 29.4	Variável social <i>sexo</i> : contribuição dos factores individuais para a marcação plural do elemento analisado – 5 variedades de portugueses .....	678



## Lista de figuras

<b>Figura</b>	<b>Legenda</b>	
Figura 1	Dialecto de HEL-Ba e PT: desenvolvimento da atribuição de PL .....	36
Figura 2	Marcação plural no CCV (Marcador visível em DET) .....	56
Figura 3	Marcação plural no CCV (Ausência de marcador em DET) .....	56
Figura 4	Estrutura sintáctico-discursiva dos sintagmas nominais nus no santomense ....	72
Figura 5	Estrutura sintáctica do SDET no santomense (marcação de número) .....	76
Figura 6	Índice demográfico de Santana e Almojarife .....	86
Figura 7	Percentagem das crianças com 5-17 anos que têm acesso à escolaridade .....	89
Figura 8	Evolução da população de São Tomé e Príncipe – 1940 a 2001 .....	90
Figura 9	Modelo de Princípios e Parâmetros .....	129
Figura 10	Modelo de aquisição das línguas.....	130
Figura 11	Estrutura X-barra .....	131
Figura 12	Estrutura alternativa gerada a partir da estrutura X-barra .....	132
Figura 13	Modelo de transmissão/aquisição de L1 .....	132
Figura 14	Incrementação estrutural da interlíngua .....	156
Figura 15	Processamento-S e Processamento Frásico .....	158
Figura 16	<i>Spell-out</i> em morfologia distribuída .....	188
Figura 17	Modelo da morfologia distribuída .....	189
Figura 18	Estrutura minimalista da concordância sintagmática: deslocação do objecto para receber caso .....	194
Figura 19	Estrutura do SDET: parte funcional e parte lexical .....	198
Figura 20	Estrutura do SDET: nó funcional SNUM .....	202
Figura 21	Representação funcional da estrutura do SN .....	202
Figura 22	Movimento de N para NUM .....	226
Figura 23	Sintaxe na OP para concordância .....	234
Figura 24	Sintaxe da concordância sujeito/verbo (PA e PB) .....	236
Figura 25	Sintaxe da concordância interna ao SN (PA e PB) .....	236
Figura 26	Sintaxe do possessivo núcleo para a concordância .....	237
Figura 27	Fossilização temporária e fossilização permanente .....	247
Figura 28	Modelo de aquisição em situação de contacto de línguas .....	283
Figura 29	Inglês: ausência de Mov-V .....	288
Figura 30	Português e francês: Mov-V-para-I .....	288



Figura 31	PA: Mov-V-para-I .....	288
Figura 32	CCV: Mov-V-para-I .....	288
Figura 33	Configurações de Mov-V para diferentes línguas .....	289
Figura 34	Estrutura do SN na FLO da aquisição .....	528
Figura 35	Romeno: movimento de N e consequente fixação de DET .....	529

## INTRODUÇÃO

“Não há nada mais perigoso que uma ideia, quando só temos uma.”  
(Paul Claude)

As línguas que emergem em situação de contacto linguístico têm fornecido material de valor inquestionável para o avanço dos estudos de um dos temas mais debatidos na actual linguística: a aquisição. O fascínio pelos vernáculos, que cedo marcou presença em estudos sobre línguas,<sup>1</sup> ganhou dimensão acrescida no Brasil, país no qual a língua portuguesa é o reflexo de uma situação sociolinguística bipolarizada (Lucchesi, 1994, 1996, 1998a, 1999, 2000a, 2001; Baxter & Lucchesi, 1997), dividida entre o padrão mais normatizado e o padrão que exhibe concordância variável, afectando estruturas das variedades de fala mais populares. No caso concreto da CPL-var registada entre os elementos que compõem o SN, a mesma é apontada como um fenómeno geral e categórico<sup>2</sup> dos mais marcantes do português do Brasil (PB), e que o distingue do português europeu (PE). O fenómeno, tem também sido amplamente estudado (p.e. Braga & Scherre, 1976; Braga, 1977; Scherre, 1978, 1988; Ponte, 1979; Nina, 1980; Guy 1981a; Fernandes, 1996; Carvalho, 1997a; Lopes 2001, P. Ribeiro, 2003; Tieppo, 2003; Baxter, 2009), originando um debate aceso acerca das motivações que o determinam. Sucintamente, a discussão em torno destas motivações remete, actualmente, para quatro hipóteses de condicionamentos:

- (i) A hipótese de influências das línguas de substrato africano, levantada por Guy (1981a) e que sugere que essa variação provirá de um ou mais substratos. Assim sendo, o cenário apontará para a possibilidade de o PB se estar a descrioulizar (Guy, 1981a, 1989; Houaiss, 1985; Holm, 1987) a partir de uma língua crioula histórica, influenciada por línguas africanas;
- (ii) A hipótese da transmissão linguística irregular (*imperfect language shift*),<sup>3</sup> postulada por Baxter & Lucchesi (1997), e que propõe que a mudança de língua e reestruturação se ficará a dever a um forte *input* de dados de PtgL2 (*transmissão irregular*) no processo de aquisição de L1, criando uma crioulação leve;
- (iii) A hipótese oposta às primeiras, isto é, a questão da deriva românica, avançada por Naro (1981), Scherre (1988) e Naro & Scherre (1993, 2000, 2007a), ambos

defendendo que as derivações direccionadas de fenómenos históricos internos à própria língua estarão na origem da variação;

- (iv) Por último, a hipótese mais ecléctica avançada por Holm (1998; 2004), que propõe explicar a formação das línguas reestruturadas com uma matriz sociolinguístico-histórica, assimilando as três perspectivas anteriormente apontadas: (1) a deriva; (2) a transmissão linguística irregular, com influências das L1's ancestrais (Winford, 2003a); (3) a influência ou influências de línguas pidgin ou línguas crioulas.

O comportamento do fenómeno da concordância nominal no SN do PB popular é marcado por redução das categorias morfológicas e pode ser conotado ao das línguas africanas (Guy, 1981a:301-302) do grupo níger-congo atlântico (Baxter, 2009:292), que o transmitiram também aos CP's atlânticos e variedades reestruturadas de português africano. Assim, é importante que se considere, neste trabalho, o funcionamento da regra da concordância nominal de número não só em outras línguas resultantes de um processo intenso de transmissão linguística irregular, tenham elas CP's atlânticos (português da comunidade de Almojarife, São Tomé, e português cabo-verdiano) ou línguas do grupo níger-congo atlântico (português vernáculo de Angola, português do Maputo e português dos Tongas, São Tomé), como substrato directo. Nesta conformidade, foram já levados a cabo estudos comparativos, numa perspectiva sociolinguística, sobre o português dos Tongas (PT) (Baxter 2004, 2009), o PA (Figueiredo, 2008, 2009a, 2009b), o português do Maputo (PMp) (Jon-And, 2008) e o português de Cabo Verde (PCV) (Jon-And, 2009). Nestes trabalhos, tem sido apontado um perfil de variação e funcionamento dos factores condicionantes estruturais semelhantes aos do PB. Portanto, de certa forma é corroborado o pressuposto de Guy (1981a), quanto aos mecanismos responsáveis pela CPL-var no SN, quando o português é adquirido por falantes de uma L1 do grupo níger-congo atlântico ou de um CP. De facto, os ancestrais falantes das variedades de português referidas poderão, durante a fase de aquisição do PtgL2, ter associado os determinantes do português aos classificadores pré-nominais que controlam a pluralização no SN das línguas do grupo níger-congo atlântico. Em resultado de tal, a CPL-var de tais variedades, a exemplo do

que acontece no português vernacular do Brasil (PVB), é fortemente condicionada pelo item colocado na posição imediatamente anterior ao núcleo do SN.

Face às questões expostas, a pesquisa que aqui se propõe levar a cabo pretende aprofundar e complementar os estudos já iniciados por nós sobre a CPL-var no SN da variedade reestruturada de português falada pelos almozarifanos, membros bilingues em santomense e português (reestruturado) de uma comunidade semi-isolada de São Tomé, que descende de antigos escravos resgatados em África e enviados para a ilha, alguns dos quais, possivelmente, remetidos depois para o Brasil. O PA, que tem como substrato um CP do Golfo da Guiné, o santomense, formou-se no bojo do contacto entre falantes deste e do português,<sup>4</sup> que determinou uma situação de transmissão linguística irregular no processo de nativização do PtgL2 falado pelos primeiros.

Um dos aspectos centrais no âmbito dos estudos sobre ASL e reestruturação de línguas emergentes em contexto de transmissão linguística irregular prende-se com a natureza altamente variável evidenciada em muitos dos elementos que compõem a cadeia frásica: utilização de determinantes, aplicação da concordância de género e número nos constituintes do SN, uso de preposições, construções relativas, categorias de Tempo, Modo e Aspecto (TMA), etc. Algumas das questões que determinaram a selecção da variação da concordância de número entre os elementos flexionáveis do SN no PA, como objecto primordial do nosso estudo, foram já avançadas em Figueiredo (2008, 2009a, 2009b). Contudo, continua por se precisar até que ponto certos factores linguísticos e extralinguísticos poderão determinar o actual estado de variação registado no PA. À medida que progredimos na análise dos fenómenos que determinarão a CPL-var no SN, mais convictos ficámos de que o processo de aquisição da marcação e desenvolvimento das regras da concordância PL reflecte uma intrincada correlação de forças não só entre características tipológicas das L1's e línguas-alvo em contacto mas também entre todas estas línguas e factores de aquisição universais, independentes das mesmas.

Desta forma, a nossa análise combina, essencialmente, vários aspectos da gramática generativa (Teoria dos Princípios e Parâmetros da Gramática Generativa), para explicação da aquisição da concordância nominal, e da sociolinguística quantitativa, para apuramento dos pesos relativos dos contextos linguísticos e extralinguísticos na

estruturação do SN e desenvolvimento das suas regras de concordância. Paralelamente, o peso quer de modelos psicolinguísticos e cognitivos, como o Modelo dos Quatro Morfemas (*The 4M Model* – Myers-Scotton & Jake, 2000a, 2000b; Myers-Scotton, 2002), o modelo da mudança evolucionista da língua (Givón, 1998; Croft, 2000; Mufwene, 2001; Clements, 2009) ou o modelo da incrementação dos processamentos linguísticos assente numa hierarquia universal inata (*The Processability Theory* – Pienemann, 1988, 2005b), quer os pressupostos da MD, avançados por Halle & Marantz (1993), Costa & Silva (2006b) ou Noyer (2006), foram também por nós tidos em consideração. O ecletismo que decidimos adoptar foi determinado pelo propósito de se observar até que ponto a variação inerente à fala dos almoxarifanos, um grupo de indivíduos que partilha a mesma L1, reflecte ou não aquisição de acordo com os princípios da GU, e se a variação actual resulta de reestruturação determinada por gramáticas em conflito. Por outro lado, o método utilizado na recolha de dados seguiu as premissas propostas por Labov (1972a, 1983) e Weinreich, Labov & Herzog (2006 [1968]), que apontam o discurso espontâneo de uma dada língua como o *corpus* ideal para os estudos e análises em linguística, visto contribuir de forma incontornável para ajudar a compreender cinco questões problemáticas e fundamentais acerca da variação registada na fala das comunidades: (i) a problemática da restrição (*constraints problems*), que estabelece os limites entre variação e mudança; (ii) a problemática da transição (*transition problem*), que determina a ocorrência de diferentes fases de variação até se consumir o processo de mudança; (iii) a problemática do encaixamento (*embedding problem*), que observa a relação estabelecida entre mudança e factores da estrutura sociolinguística; (iv) a problemática da avaliação (*evaluation problem*), que analisa a atitude dos falantes relativamente à variação; (v) a problemática da implementação (*actuation problem*), que se debruça sobre as potenciais causas conducentes à mudança linguística.

Paralelamente, a observação do desenvolvimento dos factores identificados como condicionadores da aquisição da regra de concordância no processo de transmissão geracional permitirá uma conclusão acerca do perfil da concordância e sua influência na variação registada a nível do SN na fala actual de Almojarife. A comparação com os perfis registados noutras comunidades ajudará também a coligir, por certo, elementos

que permitirão confirmar se o português adquirido em situação de contacto por falantes de um CP manifesta padrões de variação semelhantes aos do português resultante de uma situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Constatada esta premissa, poder-se-á trazer mais alguma luz a um dos dilemas que tem preocupado a actual crioulistica: terão as variedades parcialmente reestruturadas em situação de contacto sido desenvolvidas por influência de uma língua crioula (ou por línguas crioulas), ancorando a CPL-var no SN em mecanismos sintácticos dos substratos (Guy, 1981a:3001-302), ou, em contrapartida, ficaram a dever-se a processos internos à própria língua, independentes de tal influência, como a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a)?

Tendo como entendimento o pressuposto de que a ASL por adultos implica um acesso parcial à GU (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), debateremos algumas diferenças registadas a nível da aquisição de L1's e L2's, enfocando questões como a aquisição das categorias funcionais. A escolha de uma comunidade específica que produz variação relativamente uniforme na sua fala foi também determinada pelo propósito de se utilizarem os dados levantados como *corpus* de controlo, para ajudar a fornecer resposta à questão que Figueiredo (2008; 2009a) tem vindo a constatar, ou seja, que o português adquirido em situação de contacto por falantes de um CP manifesta padrões de variação semelhantes aos do português resultante de uma situação de transmissão linguística irregular em que a língua ancestral não é um CP. Assim julgamos estarem reunidas condições para melhor se identificarem as características linguísticas da comunidade, uma vez que este tipo de produção tem a vantagem de evitar artificialidades que adulteram os resultados dos dados submetidos a análise. Efectivamente, as limitações impostas à fala em tempo real impedem que, por exemplo, os falantes mais escolarizados possam assumir um controlo metalinguístico nas suas produções.

A análise do desempenho individual de cada um dos informantes foi feita com recurso ao pacote GOLDVARB X, que permitiu identificar não só as variáveis linguísticas e extralinguísticas que influenciam a marcação plural nos constituintes do SN da fala de Almojarife mas também que a variação registada a nível da concordância

dos itens flexionais é sistemática. A aplicação ou ausência das regras de concordância na sintaxe aparente do PAL1 poderá estar relacionada quer com a tipologia do santomense quer com a tipologia dos substratos africanos deste. Não obstante, se esta fosse a única justificação para a variação, não se entenderia o porquê de os falantes almorávidas aplicarem regularmente a regra em determinados contextos. Deste modo, tentaremos verificar não só se a variação em análise resulta da interacção complexa entre gramáticas da L1 e da LA mas também avaliar a forma como os princípios universais e psicolinguísticos subjazem à construção das referidas gramáticas.

O número de variedades de português africano e brasileiro a que recorreremos para estudos comparativos com o PA, o extenso quadro teórico que apoiou o nosso estudo e a metodologia pela qual optámos produziram, em agregado, uma quantidade elevada de tabelas e dados para análise. Desta forma o nosso estudo alcançou dimensão considerável, facto que, reconhecemos, o torna bastante volumoso. Depois de reflectirmos sobre este aspecto, optámos por não reduzir muito a dimensão inicial do trabalho, pois, ao fazê-lo, teríamos que retirar também achados que consideramos importantes para estudo da CPL-var numa perspectiva comparada. Desta forma, elaboraram-se cinco capítulos, cabendo ao primeiro destes estabelecer o pano de fundo sociohistórico da comunidade de Almorávida e apresentar também uma resenha dos contextos sociohistóricos das comunidades cujas línguas foram já observadas em estudos anteriores, os quais servirão para efectuarmos a nossa análise numa perspectiva comparada. Nesta parte do trabalho, dirigiremos a atenção para particularidades sociohistóricas que determinaram a aquisição do português em determinadas regiões do Brasil (Rio de Janeiro, Salvador, Helvécia e Rio de Contas) e de África (Angola, Moçambique, Cabo Verde e São Tomé). A identificação das condicionantes sociolinguísticas que poderão ter determinado o desenvolvimento da CPL-var no SN de todas essas variedades de português, com especial relevância para a da comunidade de Almorávida, objecto nuclear do nosso estudo, permitirá ver que tais condicionantes se tocam em muitas das suas especificidades, sugerindo transferência morfossintáctica, que se manteve depois por transmissão irregular. A fim de melhor se observarem estes fenómenos, o capítulo finaliza apresentando alguns dos traços gerais que caracterizam a fala de Almorávida e listando alguns paralelismos morfossintácticos, a nível do SN, que

os mesmos estabelecem com variedades de português (brasileiras e africanas), CP's atlânticos e línguas africanas, com especial incidência para as línguas do grupo níger-congo atlântico.

No capítulo 2 serão apresentados os pressupostos teóricos do nosso trabalho, elaborando-se não só um panorama acerca das teorias de aquisição de L1, bilingue e L2 mas também da forma como determinados estudos da actual Sociolinguística Variacionista têm procurado estabelecer a relação entre o comportamento linguístico dos falantes que integram uma dada comunidade de fala (língua-E) e a sua competência linguística enquanto usuários nativos de uma determinada língua natural (língua-I). Neste capítulo daremos ainda conta das perspectivas gerais sobre a variação na L2 e causas que poderão determinar fenómenos de fossilização. Estes aspectos implicam que se abordem também as questões em torno da morfossintaxe do SN no PE, PB e PA, observados numa perspectiva comparada. Apresentaremos ainda o quadro teórico da concordância a nível frásico e a nível do SN, abordando os pressupostos acerca da *interface* entre módulos gramaticais, com enfoque nas abordagens dos Princípios e Parâmetros da Gramática Generativa (Chomsky, 1957, 1965), Minimalismo (Chomsky, 1996 [1995]) e MD (Halle & Marantz, 1993; Costa & Silva, 2006b; Noyer, 2006). Sem deixar de levar em conta as diferenças registadas a nível da concordância plural no SN do PE e do PB, comentaremos as perspectivas acerca da aquisição das categorias funcionais em L1's e em L2's e condicionamentos que determinam a CPL-var no SN das segundas. Estes condicionamentos conduzem-nos, então, às questões em torno da transmissão linguística irregular e sua determinação na reestruturação de línguas, nomeadamente no desenho da configuração sintagmática do PA.

Quanto ao capítulo 3, será dedicado à descrição da metodologia aplicada na recolha, selecção, codificação e tratamento dos dados. Neste capítulo apresentaremos e descreveremos ainda o tipo de variáveis linguísticas e extralinguísticas a que recorreremos para postular hipóteses acerca das motivações que poderão determinar a CPL-var no SN do PA. No mesmo capítulo faz-se, ainda, uma descrição acerca do suporte estatístico computacional a que recorreremos para quantificar e analisar os dados.

Os resultados da análise serão apresentados em dois capítulos. No capítulo 4, tratámos os dados e analisámos os resultados das variáveis dependentes e dos grupos de



fatores linguísticos do tipo estrutural, comparando os nossos resultados com os de outros trabalhos sobre a CPL-var no SN de variedades brasileiras e africanas de português. No capítulo 5, recorreremos a idêntica metodologia para tratamento de dados e análise de resultados das variáveis linguísticas dos tipos semântico e fonológico, bem como das variáveis extralinguísticas.

Após o capítulo 5, serão apresentadas as conclusões do estudo. Determinou-se que a variação da concordância de número entre os elementos flexionáveis do SN é sistemática e orientada por contextos linguísticos específicos. Estes aspectos permitem prever em que tipo de estruturas é que os aprendentes do PtgL2 que tenham línguas crioulas como L1 revelam propensão para colocar, ou não, as marcas formais de PL. Os resultados permitiram concluir também que a aplicação da regra de concordância no PA se encontra em estágio de variação estável, fazendo antever uma situação de mudança linguística em progresso. Paralelamente, verificou-se que o português adquirido em situação de contacto por falantes de um CP manifesta os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a L1 ancestral não é um CP (Figueiredo, 2008; 2009a).

Dada a quantidade de páginas produzida, e para facilitar o manuseamento do trabalho, optámos por dividi-lo em duas partes, estando incluídos no primeiro volume a introdução e os três primeiros capítulos. Quanto ao segundo volume, dele fazem parte os capítulos 4 e 5, dedicados à análise dos resultados, bem como as conclusões do estudo e a lista bibliográfica consultada. No final do texto de cada um dos volumes, encontram-se inseridas as notas de fim de texto respeitantes ao mesmo.

Resta apenas referir que, durante o período que decorreu a presente investigação, a mesma permitiu ao autor produzir os seguintes trabalhos:

#### Participação em congressos:

Figueiredo, Carlos Filipe Guimarães. 2006. *O sintagma nominal no português reestruturado de Almojarife (São Tomé)*. Congresso Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLe). Universidade de Coimbra, Portugal, 26-28 de Junho.

—2008. *Encaixamento linguístico da configuração do sintagma nominal plural do português reestruturado da comunidade de Almojarife (São Tomé)*. Congresso Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLe). Universidade Nova de Lisboa, Portugal, 03-05 de Julho.

- 2009. *Concordância plural variável no sintagma nominal do português reestruturado de Almojarife: motivações morfofonológicas*. Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade de Colónia, Alemanha, 11-15 de Agosto.
- 2010. *Paralelismos morfossintáticos em variedades bantu, crioulos atlânticos de base portuguesa e variedades africanas e brasileiras de português: transferência ou trajecto universal de aquisição?* 10ème Colloque International de l'Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Llacan, Campus CNRS de Villejuif, Paris, France, 1-3 de Julho.

Publicações:

- Figueiredo, Carlos Filipe Guimarães. 2008. *A concordância variável no sintagma nominal plural do português reestruturado de Almojarife (São Tomé)*. *Papia – Revista Brasileira de Estudos Crioulos e Similares*, 18. 23-43.
- 2009. A configuração do SN do português reestruturado da comunidade de Almojarife – São Tomé [em linha]. *RCBPLE – Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, 1(1). 28-55. Disponível em:  
<http://rcblpe.sftw.umac.mo/doc/A%20Configura%C3%A7ao%20do%20SN%20Plural%20do%20Portugu%C3%AAs%20Reestruturado%20da%20Comunidade%20de%20Almojarife%20-20Sao%20Tom%C3%A9.pdf>.



## CAPÍTULO 1

*“O mais grave nos negros não-cristãos não é serem selvagens. Mas é estarem convertendo em selvagens os portugueses que são brancos e cristãos.”*

(In, Mia Couto, *O Outro Pé da Sereia*, p. 229)

### O tema

O debate em torno da CPL-var no SN do PB tem motivado estudos em sociolinguística variacionista, principalmente no Brasil (p.e. Braga & Scherre, 1976; Braga, 1977; Scherre, 1978; Ponte, 1979; Nina, 1980; Guy 1981a; Scherre 1988; Fernandes, 1996; Carvalho, 1997a; Lopes 2001, P. Ribeiro, 2003; Tieppo, 2003; Baxter, 2009). Alguns destes trabalhos têm procurado determinar se as motivações da CPL-var ancoram nas ancestrais línguas de substrato (Guy, 1981a:301-302), originando reestruturação devido a forte input de PtgL2 por transmissão linguística irregular (Baxter & Lucchesi, 1997), ou se, em contrapartida, têm origem em derivações direccionadas de fenómenos inerentes à própria língua portuguesa (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a). Estes aspectos determinaram que os estudos acerca da concordância variável, numa dimensão comparada, tivessem ganhado alento, analisando-se o fenómeno quer em textos do português medieval (Tarallo, 1996; Naro & Scherre, 2000) quer em variedades brasileiras de português (Lucchesi, 2000a; Lopes, 2001, P. Andrade, 2003, Baxter, 2009) quer ainda em variedades africanas de português (Baxter, 2004, 2009; Figueiredo, 2008, 2009a, 2009b; Jon-And, 2008, 2009).

Neste capítulo faremos uma resenha de alguns trabalhos elaborados para análise do fenómeno da CPL-var no SN em determinadas comunidades linguísticas, observando-se também o pano de fundo sociohistórico desta. Posteriormente, apresentaremos os traços mais marcantes que caracterizam a fala dos membros da comunidade de Almojarife, São Tomé, listando alguns paralelismos morfossintácticos que o PA estabelece, a nível do SN, com variedades de português (brasileiras e africanas), CP's atlânticos e línguas africanas do grupo níger-congo atlântico.

#### 1.1. Português do Brasil e concordância nominal variável de número

O PB apresenta diversidade acentuada relativamente ao PE, sobretudo a nível dos falares populares e rurais, aspecto que, desde cedo, chamou a atenção de dialectólogos e

linguistas. Os estudos científicos sobre o fenómeno, sobretudo na década de 70 e início dos anos 80 do século XX, permitiram concluir que tal diversidade reflecte, em si, a própria sociohistória do país. Alguns linguistas chamam a atenção para a inapropriação das gramáticas normativas no Brasil (Lucchesi, 1994, 1998a; Bagno, 2001, 2002), subjugadas às regras teórico-descritivas europeias: *“Fundamentam-se as regras da gramática normativa nas obras dos grandes escritores, em cuja linguagem as classes ilustradas põem o seu ideal de perfeição, porque nela é que se espelha o que o uso idiomático estabilizou e consagrou”* (Lima, 2002 [1957]:7). Como as produções escritas de carácter literário, tomadas como exemplo, não reflectem o registo oral espontâneo, Bagno radicalizaria o seu discurso, lançando aos pesquisadores o repto de determinarem se no Brasil se fala *“português ou brasileiro”* (Bagno, 2001:19).

De entre os fenómenos morfossintácticos do PB, as concordâncias variáveis nominal e verbal evidenciam distinção acentuada em relação às regras categóricas padronizadas. Dado que a variação apresenta diferenciados graus de aplicação, consoante os falantes sejam usuários da norma culta ou da norma popular, não surpreende que os linguistas a tomem como elemento privilegiado de análise, a fim de determinarem o modo como o fenómeno se encaixa na gramática e na sociedade, ou seja, quais as motivações que o determinam e a forma como ele é percebido socialmente. No concreto, os estudos acerca da concordância variável nominal de número não têm evidenciado consenso sobre as motivações que a impulsionam. Basicamente, alguns estudiosos, defensores da linha contactista, entendem que o desenvolvimento do PL no SN tem, no seu bojo, a intervenção de processos funcionalmente irrestritos de crioulização. Estes, por sua vez, levaram a que a variação na concordância do PB fosse determinada por fenómenos morfossintácticos indiciadores da ocorrência de reestruturações colaterais, impulsionadas pelo contacto massivo entre línguas que marcaram a sociohistória do Brasil (Baxter & Lucchesi, 1997; Lucchesi, 1998a, 2001), mas que teriam uma fonte comum: os ancestrais dialectos africanos falados pelos escravos africanos transplantados (Guy, 1981a, 1989, 2005; Holm, 1987, 1992; Baxter & Lucchesi, 1997, 1998; Lucchesi, 2000a; P. Andrade, 2003, Baxter, 2009).<sup>5</sup> Por seu lado, os defensores da linha derivista sugerem que a variação é consequência de o português estar a trilhar um trajecto próprio no Brasil, demarcando-se morfossintacticamente do PE, quer por incrementar o uso do

pronome na posição de sujeito e deixar de aplicá-lo no lugar do objecto quer por o seu sistema de concordância nominal e verbal estar a ser afectado por mudanças fonéticas naturais (Révah, 1959), determinadas, principalmente, pela deriva diacrónica (Naro & Lemle, 1976; Scherre, 1988).<sup>6</sup> Deste modo, o padrão de concordância variável do SN no PB terá sido fornecido por variedades do PE (Naro & Scherre, 1993, 2000, 2007a, 20007b). Contudo, no Brasil, o fenómeno apresentará uma extensão diferente, por ter sido intensificado e acelerado pelo contacto do português com as línguas indígenas e africanas, ou seja, em virtude de os falantes do PtgL2 terem exacerbado o modelo europeu (Naro, 1981; Naro & Scherre, 1993, 2000, 2007a). Ainda que assim seja, o facto de a aprendizagem do PE (modelo sem variação) por contacto produzir uma L2 com variação, é suficiente, por si só, para concluir que não faz sentido incidir atenção sobre a questão da deriva. Com efeito, e independentemente de este fenómeno ocorrer ou não no PE, o contacto irá originar sempre um produto final marcado por variação.

Por outro lado, é também importante ter em consideração que a noção de deriva latina foi formulada por filólogos numa época em que os fenómenos de contacto eram pouco compreendidos, o que levava a atenção dos estudiosos a centrar-se na língua-I (*I-language*).<sup>7</sup> Por conseguinte, sendo as línguas miscigenadas encaradas como deturpações da LA<sup>8</sup> envolvida no contacto, as questões extralinguísticas não eram consideradas essenciais para os estudos em diacronia. Este aspecto determinou que a fonética ganhasse grande protagonismo no final do século XIX e princípio do século XX, tendo o Positivismo dos neogramáticos introduzido a hipótese da existência de leis fonéticas de carácter universal e absoluto. Com estas leis aplicadas cegamente sobre os sons, as mudanças linguísticas eram explicadas de forma idêntica para todas as línguas, originando uma quase total ignorância acerca dos processos de mudança e aquisição de língua em estudos linguísticos (Baldinger, 1972 [1958]; Vänänen, 1981 [1957]; Williams, 1986 [1938]), até à sexta década do século XX. Contudo, e sendo inquestionável que todas as línguas são afectadas por mudanças de geração para geração, mesmo em contextos monolíngues, a verdade é que tais estudos não negam as interferências fonológicas e adopção de elementos lexicais dos substratos (Branco, 1984:7). É o que acontece, por exemplo, com o fenómeno da lenição (abrandamento) das consoantes oclusivas em posição intervocálica, que afectou e simplificou, na zona da

România Ocidental, o sistema de consoantes duplas do latim falado, muito por força do seu contacto com o substrato celta. Este fenómeno despoletou uma série de outros em cadeia, com vista a evitar situações de homonímia e homografia, e que, nalguns casos, chegou a culminar com o aparecimento de novas consoantes, como acontece, por exemplo, com a consoante /v/, inexistente no latim clássico (p.e. [-pp- > -p- > -b- > -ɸ- > -v-]). Outro fenómeno identificado de interferência substratal ocorreu com a conservação do [l] e do [n] intervocálicos em determinadas palavras no galego-português, em virtude do seu contacto com o moçárabe (p.e. *canito*; *male*) (Branco, 1984: 15). Quanto a incorporações lexicais dos substratos, poder-se-á apontar, por exemplo, o uso do sufixo celta *-occus* no latim, em combinação com um nome, geralmente um deus (p.e. *Indovellicus*) (Branco, 1984:13).

No entanto, cabe aqui questionar se os fenómenos de influências ou transposição se limitaram aos aspectos meramente fonológicos ou lexicais. Thomason & Kaufman (1988:37) identificam duas tendências gerais nos fenómenos determinados por contacto, cuja extensão depende dos factores prevaletentes neste: a transferência e o empréstimo. Para acontecer a primeira, é necessário que os falantes envolvidos no contacto linguístico sejam monolíngues. Estes, à medida que vão aprendendo a LA, transferem estruturas da língua nativa para a língua na qual procuram comunicar (Thomason & Kaufman, 1988:73-74), originando a emergência de uma língua que se distingue quer do substrato quer da LA. Posteriormente, e caso o contacto se mantenha, a nativização da nova língua, que geralmente ocorre entre três e seis gerações (Clements, 2009:34), é também marcada por reestruturação assente em empréstimos, uma vez que os falantes submetidos ao contacto se tornam bilingues (Thomason & Kaufman, 1988:73-74). No caso concreto da transição do latim para o português, apesar da tenaz oposição do povo indígena da Ibéria Central e Ibéria do Noroeste aos romanos, há registo de existirem, nestas regiões, durante o período entre 100 aC e 300, falantes bilingues em variedades celtas e latim (Tovar, 1977:124). A grande maioria destes seria analfabeta, o que minorava a recepção de *input* negativo acerca das produções da LA que evidenciavam variação. A intervenção dos bilingues neste extenso período de contacto linguístico terá sido determinante para a manutenção da morfologia pré-romana nos romances ibéricos central e ocidental. Nas variedades celtiberas, os nomes com vogal temática em *-o*

formam o nominativo PL em *-os*. Este traço entrou, por transferência, no sistema da segunda declinação dos nomes pré-romanos e latinos, com vogal temática em *-us* (*-us, -i, -o, -um, -o; -i, -orum, -is, -os, -is*). Posteriormente, e por empréstimo, os nomes temáticos em *-o* do português passaram a registar o PL actual em *-os* (o sufixo nominativo PL do latim) (Clements, 2009:34). Como se verifica, a variação que configura a mudança não estará apenas dependente de fenómenos inerentes à própria língua, como a deriva românica, já que pode ser também consequência de outros aspectos, como o contexto em que se dá a aquisição, o grau de similaridade estrutural das variedades em contacto, o tipo de falantes envolvidos neste e até a sua atitude comunicativa.

No caso concreto do Brasil, a visão teocêntrica do português quinhentista demonstra, logo desde o achamento, a manifestação de um preconceito que, durante séculos, discriminaria dialectos cujas tipologias fossem distintas da sua L1. Em *A Carta*, por exemplo, é notório o paralelismo que o cronista real Pêro Vaz de Caminha estabelece entre as práticas sociolinguísticas dos ameríndios e o paradigma europeu: “... *não pôde deles haver fala nem entendimento que aproveitasse*” (Seabra, 2000).<sup>9</sup> Por seu lado, situações como a implantação tardia do ensino no país concorrerem para que a unidade construída em Portugal para o seu idioma não ocorresse também no Brasil. Neste, a sociedade é altamente estratificada e regista um *continuum* linguístico bastante marcado, que vai desde as variedades vernáculas presentes nas comunidades rurais isoladas e quase analfabetas até às formas populares urbanas ou culta, usadas nos centros cosmopolitas.

Estudos de finais do século XIX (Coelho, 1880/1881, 1882, 1886) dão-nos conta de diferenças registadas entre o PB e o PE, sem, contudo, apontarem qualquer tipo de variação no segundo. Tais diferenças fazem perceber uma situação de bipolarização no PB, instalada até ao final do século XIX (Lucchesi, 1994, 1996, 1998a, 1999, 2000a, 2001, 2009a; Baxter & Lucchesi, 1997), altura em que ocorreu o súbito processo de industrialização e urbanização do país. A situação de bipolarização ter-se-á iniciado quando, no final do século XVI, o quadro habitacional do Brasil se começou a modificar drasticamente com a instalação dos engenhos de cana-de-açúcar no Nordeste do território. De facto, das diferenças linguísticas nos davam já conta quer Silva Neto (1986



[1957]), que chega a citar o Padre António Vieira por, no século XVII, fazer menção ao avesso e direito do português no território, o primeiro falado pelos naturais e africanos e o segundo pela elite que zelava pelos padrões europeus. Nesta fase, são também comuns os relatos de alguns viajantes estrangeiros que fazem referência aos falares dos índios e escravos negros, impregnados de características crioulas. Posteriormente, uma série de acontecimentos sociohistóricos viriam a afectar a hegemonia e representatividade dos falares populares no Brasil: declínio dos engenhos do açúcar e consequente deslocamento de parte das massas escravas do nordeste para trabalhar nas minas de esmeraldas e ouro do sudeste da colónia; penetração nos sertões brasileiros do surto das bandeiras, composto por grupos paulistas que englobavam indivíduos de condição social baixa, fossem eles brancos, mestiços ou índios, aí erigindo povoados que evoluiriam para grandes cidades e nas quais se movimentam também poetas, intelectuais e pequenos burgueses; novo deslocamento de escravos, desta vez para as roças de café que se iam espalhando pelos estados de São Paulo e do Rio de Janeiro, após esgotadas as reservas de minerais da zona de Minas Gerais; abandono das fazendas por parte dos libertos, após assinatura da Lei Áurea (1888) e extinção da escravatura, que se dirigem para os centros urbanos, exponenciando o crescimento vertiginoso destes e distendendo, diatopicamente, o contacto do PVB a contextos conservadores e europeizantes. Observada numa perspectiva sociolinguística, a entrada de escravos africanos no Brasil e posterior abolição da escravatura neste reflectiram-se extensamente no idioma do país, já que a distribuição dos falantes do PVB pelo país, traduzido na fixação e aculturação da população de origem africana um pouco por todo o território, acabou por impor não só as variantes vernáculas do português sobre a língua geral inicialmente falada pelas populações autóctones mas também por converter o modelo defectivo do PtgL2, falado por escravos, imigrantes pobres e colonos afastados dos contextos linguísticos urbanos, em L1 dos seus descendentes, fossem eles pedógamos ou mestiços. Contudo, ao verem dificultada a promoção da sua cidadania, os ex-escravos que aportaram aos centros cosmopolitas acabaram por buscar a ascensão social recorrendo a diversos tipos de estratégias, entre as quais se poderá destacar a abertura à pressão linguística, traduzida na introdução de características da norma culta no português popular. Ainda outro aspecto a ter em linha de conta, enquanto fenómeno originado pelo êxodo dos ex-

escravos para os centros cosmopolitas, foi o facto de a sociedade brasileira deixar de assentar grandemente em padrões de características rurais e transitar, em finais do século XIX, para uma sociedade do tipo urbano.

As mudanças verificadas na sociedade brasileira em finais do século XIX viriam também a ser incrementadas pelo rápido desenvolvimento de estruturas de diversos tipos, as quais afectariam, directa ou indirectamente, todo o seu sistema linguístico-cultural. O ensino básico, apesar de precário, massificou-se e a população menos elitista viu também facilitado o acesso aos meios de comunicação social. Quanto à rede viária, ramificou-se por todo o país e as grandes empresas capitalistas ligadas às áreas da mineração, extracção madeireira e produção de celulose expandiram-se pelo interior, incrementando a exploração dos recursos naturais do país. Submetidas à pressão deste jogo de forças, a norma popular e a norma culta foram estreitando laços, cedendo ambas nos seus padrões mais marcantes e caminhando para um patamar de equilíbrio das suas características linguísticas. A primeira deu passos no sentido de aproximar as suas particularidades às do padrão urbano culto e fazendo com que se diluíssem as marcas mais salientes das alterações resultantes do extenso contacto linguístico ocorrido nos séculos antecedentes. E a este processo de mudança não escaparam também os contextos linguísticos mais afastados e isolados, que viram a sua cultura ser desarticulada e o seu modo de vida radicalmente alterado. No entanto, os estudos empíricos levados a cabo nestas comunidades evidenciam que, apesar de terem perdido as suas marcas mais crioulizantes, as suas falas continuam a revelar características que permitem conotá-las aos processos de transmissão linguística irregular que determinaram a sua emergência.

Quanto à norma culta, sujeita à pressão quer dos falantes das classes imigrantes que se mudavam do campo para a cidade e se integravam na estrutura social quer dos defensores da demarcação em relação às normas europeizantes, via-se democratizada e sujeita a mudanças que faziam com que aproximasse também o seu registo do da norma urbano-popular que se ia implementando. A prova da penetração nos estratos sociais médios e altos de algumas das características resultantes da aquisição/nativização do português falado por ex-escravos e seus descendentes é dada, por exemplo, pela *“variação na concordância de número (tanto verbal, quanto nominal), que se teria originado no processo de transmissão linguística irregular ocorrido na base da*

*pirâmide das sociedades colonial e do Império, e se teria difundido tornando-se hoje um fenômeno que se verifica em todos os extratos da sociedade brasileira”* (Lucchesi, s.d.). Em consequência da pressão exercida, em simultâneo, sobre as normas popular e culta, atenuou-se o quadro marcadamente bipolarizado que havia imperado até então, caminhando-se para o actual panorama registado no país. Ainda assim, o ensino bastante tardio do português não conseguiu evitar que, no Brasil, emergisse uma situação de quase diglossia, cavando-se um fosso entre a língua da população miscigenada e de origem escrava, por um lado, e a das elites escolarizadas, por outro lado. Deste modo, a unidade construída em Portugal para o seu idioma não ocorreu no Brasil, onde a sociedade é altamente estratificada e regista um *continuum* linguístico bastante marcado, que vai desde as variedades vernáculas presentes nas comunidades rurais isoladas e quase analfabetas até à formas populares urbanas ou culta, patentes nos centros cosmopolitas. Como consequência das diferentes políticas de língua adoptadas, a norma do PE é diferente da do PB, e ambas sê-lo-ão também das outras que estão a germinar em todos os outros países de expressão portuguesa. Não obstante estes aspectos, os primeiros trabalhos sobre a concordância nominal de número no SN teimavam em apresentar uma visão tradicionalista deslocada da realidade do país, indicando-a como categórica (p.e. Lima, 2002 [1957]:305-306; Bechara, 1999 [1928]:544-545).<sup>10</sup>

Assim, a concordância variável de género e número no SN de variedades do PB continuou a ser apontada apenas em trabalhos de dialectólogos, como os de Amaral (1920), sobre o dialecto caipira, ou de Monteiro (1933) e Marroquim (1945 [1934]), sobre o português popular nordestino. Quanto aos linguistas, têm privilegiado a sociolinguística quantitativa como modelo de análise, uma vez que esta permite determinar o modo como o fenómeno se encaixa na gramática e na sociedade, ou seja, entender quais as motivações que o determinam e como ele é percebido socialmente. Mais recentemente, análises do mesmo tipo têm sido feitas em variedades africanas de português (Moreno & Tuzine, 1998; Baxter 2004, 2009; Figueiredo, 2008, 2009a, 2009b; Jon-And, 2008, 2009) e PtgLE em contexto de ensino (Godinho, 2005, 2008), abrindo o leque de possibilidades para o estudo do fenómeno da variação no SN numa perspectiva comparada.

No Brasil, o primeiro estudo sistemático sobre a CPL-var no SN no Brasil foi levado a cabo por Braga & Scherre (1976), que observaram o comportamento da concordância PL do PB falado na área urbana do Município do Rio de Janeiro (MRJ). Este trabalho teve como modelo base o estudo de Naro & Lemle (1976) sobre a concordância sujeito/verbo produzida por falantes semi-escolarizados, também da área urbana do Rio de Janeiro, e que frequentavam o curso de alfabetização de adultos. Na análise de Naro & Lemle (1976:240-241), é advogada que a mudança registada no PB é determinada pela deriva natural da língua, começando nos contextos menos salientes e atingindo, posteriormente, as formas mais salientes, por “imitação”. Por seu lado, Braga & Scherre (1976:474) concluíram que os itens de maior *saliência fônica* favorecem mais a inserção da pluralização, por oposição aos elementos terminados em *-s* no singular e que pluralizam com a inserção de *-es* final, que apresentam tendência para inibirem a marcação. A tendência para a não-marcação destes elementos deve-se, segundo as autoras, ao fenómeno da analogia,<sup>11</sup> comum em situações de aprendizagem que envolvem dados divergentes ou insuficientes, uma vez que os itens em questão apresentam já um *-s* final, que poderá ser entendido pelos falantes como uma marca de pluralização.

Seguidamente, Braga (1977) observou a concordância PL do português falado pelas classes média e baixa do Triângulo Mineiro (PMn), voltando a enfatizar que a *saliência fônica* exerce influência na marcação PL e que os itens terminados em *-s* a inibem, mas que estes comportamentos não são uniformes em todos os estratos sociais. Deste modo, Scherre (1978) avaliou a influência das variáveis sociais na marcação PL produzida por falantes de diversos níveis de escolarização do MRJ. Nesta análise, a linguista detectou que a *saliência fônica* propicia mais a concordância nos falantes escolarizados do que nos semi-escolarizados, apesar de ambos favorecerem positivamente a pluralização. Com base nas constatações de todos estes trabalhos, Scherre (1978:127) advogaria que a mudança emerge nos ambientes de menor *saliência fônica*, alastrando-se e alcançando, depois, os de maior *saliência*.

Ponte (1979:264-266), analisando o português de semi-analfabetos de Porto Alegre (PPA), Rio Grande do Sul, concluiu que o SN deste dialecto apresenta uma hierarquia de marcação PL idêntica à dos informantes semi-escolarizados do MRJ. Nina (1980), por

seu lado, observou os percentuais de pluralização realizados por informantes analfabetos da micro-região bragantina, Estado do Pará, tendo concluído que os itens na primeira posição da cadeia sintagmática são os que mais recebem marcas de PL. Constituindo a variável *grau de saliência fônica na oposição singular/plural*, Nina (1980:103) constatou que ela não é relevante para a marcação, mas que ocorre tendência para se pluralizarem os itens regulares e metafônicos. Por seu lado, os elementos que formam o PL com inserção de *-s* final e mudança silábica (p.e. *casal/casais*) apresentavam uma taxa de inibição total (100%).

Com recurso à variável *natureza flexional do elemento precedente*, Nina (1980:144-145) verificou também que os factores que mais favorecem a marcação são o numeral e os elementos sem marca formal de PL, mas com carga semântica de pluralização. Face aos resultados das variáveis linguísticas, a pesquisadora conclui que o português paranaense apresenta tendência para eliminar a concordância redundante, seja ela nominal seja ela verbal. Este aspecto vem ao encontro do pressuposto da hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que advoga que a inserção de PL na primeira oportunidade restringe a marcação nas seguintes, em virtude de esta se revelar redundante. Por outro lado, as variáveis sociais constituídas por Nina (1980:152) – *sexo e idade* –, levaram a autora a constatar que as mulheres usam mais as marcas e que a regra de concordância se encontra em processo de aquisição pelos escalões etários mais novos.

Naro (1981:76), retomando a questão da possível relação entre maior frequência de marcação em número e maior saliência fônica, propõe que o peso desta na marcação passe a ser observada sob duas perspectivas, levando em conta a diferença fônica, por um lado, e a tonicidade, por outro lado. No trabalho, Naro (1981:97) insiste que a deriva subjaz às mudanças linguísticas, pelo que a variação registada no PB será motivada por tal fenómeno. Em contrapartida, Guy (1981a:189), no seu estudo sobre o português dos falantes semi-escolarizados do Rio de Janeiro, apesar de seguir a linha directriz proposta por Naro para a configuração dos grupos de factores, conclui que a *saliência fônica* não exerce efeito relevante na marcação PL.

Dado que o Brasil foi o principal destino dos escravos resgatados na região afro-atlântica, Guy (1981a), para identificar os contextos linguísticos e sociais que podem

determinar o apagamento do *-s* final, seja ele marca de PL, ou não, teve em consideração o facto de ter ocorrido contacto entre línguas europeias e africanas. De acordo com esta visão, a inserção de marcas é fonologicamente determinada pelas sílabas tónicas e pelo contexto fonológico seguinte, quando este é realizado por pausa ou pelos traços [–vozeado] e [velar]. Na cadeia linear, o item que mais recebe a marcação é o que se encontra na primeira posição, com todas as outras posições revelando tendência para a inibição. Constituindo a variável *informação de plural precedente*, Guy (1981a:179) refere que, caso ocorra um numeral antes do item analisado, a tendência é para inibir a marcação no elemento seguinte. Paralelamente, e tendo como base os estudos de Braga (1977) e Scherre (1978), Guy (1981a:184) observou também o comportamento da *saliência fónica* das categorias morfológicas, estabelecendo dois factores para o efeito: *palavras terminadas em vogal* e *palavras terminadas em consoante*. A conclusão é que as primeiras inibem a inserção de marcas, enquanto as segundas a favorecem.

A fim de testar a hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), Guy (1981a:191) constituiu as variáveis *correlação entre marcas de PL no SN e marcas de PL no verbo* (factores *verbo seguinte com marca de plural* e *verbo seguinte sem marca de plural*) e *efeito do local da informação adicional de plural*. A primeira das variáveis não confirmou a hipótese, mas Guy (1981a:192) justificou o facto referindo que os falantes, por uma questão de estilo, pluralizam quer o SN quer o SV, caso pretendam usar mais a marca de número PL. O predomínio de determinada gramática, em detrimento de outra, pode ser também explicada através do Princípio da Coesão Estrutural (Haiman, 1983:782), que determina que o falante revela tendência para constituir mentalmente grupos que compartilham traços, a fim de empregar itens linguísticos em consonância com a gramática ou código que selecciona para poder interagir nas diferentes situações de fala (cf. Lucchesi, 2000a:141). Quanto à variável *efeito do local da informação adicional de plural*, confirmou a hipótese funcionalista. Assim, existe correlação entre gramática e posição linear, daí a marcação PL revelar tendência para ser inserida na primeira posição da cadeia sintagmática e apresentar inibição da marcação nos lugares à direita desta posição. Já no que concerne às variáveis sociais,<sup>12</sup> Guy (1981a:240) considera que a variável *sexo* é aquela que merece mais

credibilidade, com as mulheres aplicando mais as marcas, mas em valores bastante próximos dos apresentados pelos homens.

Estabelecendo paralelos entre a variação do português e dialectos hispano-americanos, Guy (1981b:179) reafirma que a posição linear constitui a chave para a inserção de marcas de PL e estabelece conexões entre posição e classe gramatical. Com base nas suas constatações deste e de outros trabalhos posteriores, Guy (1981a; 1989; 2005) situa o eixo da marcação PL nos determinantes, estabelecendo um paralelismo configuracional entre o PB e as línguas dos ancestrais substratos bantu e kwa falados pelos escravos transplantados para o Brasil e seus descendentes. Desta forma, é de considerar a possibilidade de o PB se encontrar em processo de descrioulização (Guy, 1981a, 1989; Houaiss, 1985; Holm, 1987), a partir de uma língua crioula histórica, influenciada por estruturas bantu e kwa (Guy, 1981a, 1989, 2005).

Scherre (1988) reanalisa a concordância nominal do MRJ, discordando de Guy (1981a, 1981b) no que concerne à correlação entre gramática e posição linear. Fazendo uma revisão sobre os trabalhos que haviam abordado a questão da CPL-var do SN, a autora analisa, atomística<sup>13</sup> e sintagmaticamente,<sup>14</sup> as estruturas do SN, estabelecendo ainda paralelos com o espanhol de Porto Rico (EPR). Constituindo um considerável número de hipóteses acerca da variação do fenómeno da concordância, Scherre (1988) testa-as, recorrendo às potencialidades concedidas pelo programa computacional VARBRUL. Para além de confirmar que as formas salientes são as mais marcadas, a autora conclui ainda a favor da intervenção do Princípio do Processamento Paralelo,<sup>15</sup> defendendo que a presença de marcas nos itens do SN conduz a mais marcas nos restantes elementos, isto é, que *“marcas levam a marcas e zeros levam a zeros”* (Scherre, 1988:208). Posteriormente, Silva & Scherre (1998 [1996]) e Scherre (1998a [1996]) confirmariam este pressuposto, mas P. Andrade (2003:22) chama a atenção para o facto de as autoras não tomarem em consideração, para as suas conclusões, os valores exibidos pelos factores *ausência de marcas fora do SN* e *zero formal na primeira posição*, isto é, *“justamente nas posições em que, no português do Brasil, preferencialmente se marca o plural”*.<sup>16</sup> Assim, será também de considerar que a marcação no interior do SN possa ser orientada pelo Princípio da Coesão Estrutural

(Haiman, 1983:782; Lucchesi, 2000a:141), e que leva o falante a constituir, mentalmente, grupos que compartilhem traços.

Até ao estudo de Scherre (1988), as pesquisas efectuadas em torno da marcação no SN vinham postulando que a primeira posição deste era o factor preferencial para a inserção da marca de PL, enquanto as restantes posições inibiam a marcação. Ao constituir a variável *relação dos elementos não nucleares em função do núcleo e posição dos elementos no SN*, Scherre (1988:167) considera este pressuposto falacioso, visto ter verificado que a ausência de marca de PL estabelece uma relação inequívoca com as posições antepostas ao elemento nuclear. No entanto, P. Andrade (2003) e Baxter (2004) observariam que o modelo linear pode funcionar, caso a estrutura do SN seja simples, isto é, composta por apenas dois elementos.

Outra questão em que há divergência entre Scherre (1988, 1998a [1996]) e Guy (1981b:179) tem a ver com a relação estabelecida entre posições lineares e categorias gramaticais na influência da marcação PL, uma vez que o segundo defendera que os determinantes surgem, por norma, na primeira posição do SN, os nomes na segunda posição e os adjectivos na terceira posição. Scherre discorda, referindo que os determinantes favorecem a pluralização, enquanto os substantivos e adjectivos a inibem. Por outro lado, Scherre (1988:152-153, 1998a [1996]:94-97) entende também que as posições lineares não podem ser vistas como exclusivas de determinadas categorias gramaticais, já que há uma diversidade de classes gramaticais nas primeira, segunda e terceira posições do SN, com qualquer uma delas revelando tendência para ser marcada, caso surja no primeiro lugar da cadeia sintagmática. Desta forma, e independentemente da classe gramatical do item analisado, o que se tem é que os elementos favorecem ou inibem a inserção de marcas em virtude de se encontrarem pospostos ou antepostos ao núcleo do SN. Segundo Scherre (1988:47), este comportamento é determinado pelo Princípio da Iconicidade, que advoga existir uma relação directa entre a dimensão linguística ou formal e a dimensão não-linguística ou conceptual (Haiman, 1983:782).

As variáveis sociais observadas por Scherre – *sexo, grau de escolarização, faixa etária e ambiente de origem do falante* –, permitiram concluir que a CPL-var do PB se encontra em fase de variação sociolinguística estável (Scherre, 1988:520), uma vez que a influência da faixa etária na inserção de marcas não é relevante. Por fim, estas mesmas



variáveis levaram ainda a autora a discordar dos resultados de Guy (1981a:301-302) acerca da possibilidade de o PB popular estar sujeito a um processo de descrioulização. Assim, em trabalhos levados a cabo posteriormente sobre a concordância sujeito/verbo no PE e no PB, Naro & Scherre (1993, 2000) mostram-se em desacordo com a hipótese de terem ocorrido processos de transmissão irregular do tipo crioulizante no português popular do Brasil. Em contrapartida, e sem desconsiderarem o contexto sociohistórico do país, Naro & Scherre, (1993:437) assumem uma posição semelhante à de Révah (1959), reafirmando que as derivações direccionadas de fenómenos internos ao português estarão na origem da variação. Os linguistas chegam, posteriormente, a recorrer a textos medievais do português, em particular do período anterior à estabilização das normas gramaticais, para sustentarem o seu ponto de vista, apesar de, no total das estruturas analisadas, a ausência de concordância apresentar apenas 1% de ocorrências (Naro & Scherre, 2000:250). Se, por si só, os dados escritos não reflectem o registo oral espontâneo, ou seja, não constituem o *corpus* ideal para a investigação científica em sociolinguística variacionista e quantitativa, o facto de os fenómenos registados configurarem apenas 1% de ocorrências não torna confiáveis os resultados estatísticos apresentados.

Recorrendo também a dados do PB antigo escrito (cartas e textos de teatro) e do Projecto Norma Urbana Culta de Salvador (NURC), Tarallo (1996) recusa a hipótese da crioulização e posterior descrioulização, uma vez que o que está a acontecer no Brasil é um afastamento em relação ao PE, e não uma aproximação a este. O entendimento de que o fenómeno seria uniforme em todo o território brasileiro, acaba também por descartar a possibilidade de ter ocorrido transmissão linguística irregular na formação sociohistórica do PB, já que esta pressupõe um *continuum* de níveis diferenciados de socialização/nativização de uma L2. Por outro lado, e a exemplo do que sucedeu com o estudo de Naro & Scherre (2000), os textos seleccionados por Tarallo (1996) não se aprestam para identificação dos fenómenos resultantes do contacto entre línguas, tal como o *corpus* da NURC, apesar de ser oral, não será o mais indicado para a observação das características da fala popular (Baxter & Lucchesi, 1997; Lucchesi, 2001).

Mussa (1991a, 1991b), por seu lado, apresenta uma visão superstratista do PB, já que entende que este não pode ser visto como um resultado das influências dos substratos

africanos. Segundo o autor, não há diferenças formais entre os processos de criouliização e de mudança natural, pois são ambos caracterizados por simplificação fonológica e estrutural (Mussa, 1991b:76). Assim, o que aconteceu no PB foi uma competição entre formas diferentes dos dialectos em contacto, optando os falantes pelas mais simples e/ou menos estigmatizadas. Como idêntica estratégia terá sido também seguida pelos falantes do português africano, o facto de este partilhar traços com o PB não passará, pois, de mera coincidência (Mussa, 1991a:237).

Já Lopes (2001:102), estudando atomisticamente a concordância nominal das normas culta e popular usadas na cidade de Salvador, encontra evidências de que o contacto entre línguas poderá estar na base da concordância variável: *“processos variados de aquisição do português ocorreram no decorrer da história desse país, envolvendo aquisição de Português L1 com dados divergentes ou insuficientes oriundos de contactos com falantes que aprenderam essa língua como L2”*. Adoptando a metodologia de Scherre (1988), Lopes (2001) constitui uma nova variável social, a *etnia*, justificada pelo facto de a cidade de Salvador possuir uma população maioritariamente negra. A variável envasa dois factores (*falantes com sobrenomes religiosos e falantes com sobrenomes não-religiosos*) e a constituição dos grupos foi viabilizada pelos estudos da área de genética da Universidade da Bahia (UFBA), que permitiram estabelecer a relação entre o sobrenome e a etnia através de análises do tipo sanguíneo das pessoas pesquisadas. Com recurso a esta variável, a autora pôde confirmar que a aquisição das regras de concordância por parte dos indivíduos negros e afrodescendentes é inferior à dos indivíduos não-afrodescendentes, um aspecto que surge também intrinsecamente relacionado com as oportunidades educacionais.

O quadro teórico que suportou os resultados de Lopes (2001) baseou-se no pressuposto cognitivo dos Quatro Morfemas (*The 4M Model*), de Myers-Scotton (1997, 2001) e Myers-Scotton & Jake (2000a; 2000b), a que nos referiremos mais detalhadamente no ponto 2.4.2.3. do presente trabalho. Paralelamente, o comportamento da CPL-var foi também observado em função dos pressupostos de Lightfoot (1999a), que entende que o organismo linguístico, apesar de ser biologicamente controlado, é desenvolvido a partir dos estímulos ambientais que intervêm no processo de aquisição. Desta forma, as conclusões indicam que a regra da concordância se dá a partir das

formas fonológicas mais salientes, enquanto a inibição se conota aos itens menos salientes. Os resultados evidenciam ainda que os elementos pré-nucleares são mais marcados do que os itens pós-nucleares, enquanto as categorias nucleares são mais pluralizadas quando ocorrem na primeira e na quinta posições. Aliás, a marcação na primeira posição é quase categórica, o que confirma a sequência de aquisição dos morfemas, com os morfemas sistêmicos estruturais (*early system morphems*) sendo primeiramente adquiridos em paralelo com os morfemas de conteúdo (*content morphems*) e, só posteriormente, se dando a aquisição dos morfemas de nível funcional sistêmico (*late system morphems*).<sup>17</sup>

No seu trabalho, Lopes (2001), contrariamente ao avançado por Guy (1981a), dá ainda conta que os numerais, devido ao seu elevado grau de saliência semântica, favorecem a inserção de marcas. Por outro lado, o favorecimento da marcação é também motivado pela pausa final, enquanto a inibição é favorecida pelas consoantes sonoras. Já relativamente às variáveis sociais, Lopes (2001) constata que a escolaridade exerce forte influência na aquisição das regras de concordância, a ponto de fazer desaparecer divergências na marcação registada entre sexos, uma vez que nas classes não-escolarizadas são as mulheres quem mais aplica a pluralização.

Estabelecendo ainda comparações com os estudos levados a cabo para o MRJ (Scherre, 1988), para a região sul (Fernandes, 1996) e para a fala universitária de João Pessoa, Paraíba (Carvalho, 1997a), Lopes (2001) conclui que o uso da concordância nominal de número nestes dialectos apresenta o mesmo padrão da cidade de Salvador. Ainda assim, os elementos nucleares do *corpus* de Salvador em segunda, terceira, ou quarta posições exibem uma marcação menos acentuada, justificando Lopes (2001:377) este aspecto com o facto de os dados linguísticos primários (DLP's) para aquisição do português dos falantes desta região serem mais divergentes do que os do MRJ, região sul e João Pessoa.

Dando conta dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil, Baxter & Lucchesi (1997) chamam a atenção para a pobreza a nível da concordância na fala de comunidades negras rurais que se encontram praticamente isoladas no Estado da Bahia. Como os estudos sobre comunidades rurais (Amaral, 1920; Veado, 1982; Wagner, 2001) e comunidades brancas de Mato Grosso,

interior da Bahia (Isensee, 1964; Callou, 1998), dão conta do mesmo tipo de fenómeno, Scherre (1994:38) entende que a “*variação na concordância de número no português falado do Brasil, longe de ser restrito a uma região ou classe social específica, é característico de toda comunidade de fala brasileira, apresentando diferenças mais de grau do que de princípio, ou seja, as diferenças são mais relativas à quantidade de marcas de plural e não aos contextos lingüísticos nos quais a variação ocorre*”. Deste modo, no Brasil não existirão características estruturais novas introduzidas pelo contacto entre línguas ou pela nativização do português entre os segmentos de falantes de outras línguas e seus descendentes, mas apenas uma ampliação de fenómenos lingüísticos regidos pelos mesmos factores estruturais que regiam o PE arcaico (Naro & Scherre, 2003:295). No entanto, esta visão afigurar-se-á simplista, uma vez que convive com a perspectiva formalista hegemónica e normativa de um sistema lingüístico que estaria na base de uma história lingüística única no Brasil (Lucchesi, 2008a:368), descartando a hipótese de existirem processos distintos e independentes de variação e mudança que concorreram, em paralelo, para a formação das duas grande normas do PB actual (Lucchesi, 2004a, 2007). Efectivamente, a mínima ou reduzida incorporação de morfologia flexional na estrutura morfológica e processos de recomposição da gramática do português popular brasileiro, ou seja, as suas estruturas mais marcadamente criouliantes, têm sido conotados à formação da norma popular brasileira, determinada, ao longo de toda a história do Brasil, pelos processos de transmissão lingüística irregular circunscritos, isto é, às variedades rurais mais refractárias à influência normatizadora dos círculos institucionais urbanos (Lucchesi, 2003:278).

O facto de os traços lingüísticos referidos se verificarem mais a nível do registo oral popular valida que se estude a fala das comunidades rurais isoladas formadas por descendentes de africanos que se fixaram no Brasil. Deste modo, P. Andrade (2003), decidiu analisar, atomística e sintagmaticamente, a concordância nominal de número no dialecto afro-brasileiro da comunidade de Helvécia, extremo sul do Estado da Bahia, a fim de verificar quais os elementos que resguardam as marcas das origens africanas do dialecto. Refira-se que, em comunidades rurais isoladas, como a de Helvécia (HEL-Ba), que fez parte da Colónia Leopoldina que perdurou de 1818 até à abolição da escravatura, é possível procurar determinar se a CPL-var entre os elementos do SN tem a ver com o

modo como se deu a aquisição entre os ancestrais falantes da comunidade. Assim, e partindo do pressuposto que a redução da morfologia flexional registada no PB resulta dos processos de transmissão linguística irregular decorrentes do amplo contacto entre línguas que ocorreu em solo brasileiro, este tipo de aquisição terá também sido regido pela ordem como os morfemas são adquiridos, isto é, pode ser conotado ao Modelo dos Quatro Morfemas (Myers-Scotton, 1997, 2001; Myers-Scotton & Jake, 2000a; 2000b), já avançado por Lopes (2001) no seu estudo sobre as normas culta e popular de Salvador.

A análise da CPL-var do SN de HEL-Ba permite concluir acerca de um quadro inicial de quase inexistência da regra de concordância entre os elementos que constituem a cadeia sintagmática, mas que o dialecto se encontra em processo de mudança em progresso no sentido de adquirir a regra. Este aspecto é confirmado com recurso à observação das variáveis sociais *faixa etária, sexo, escolaridade e estada fora da comunidade*. Quanto aos resultados da análise quantitativa, P. Andrade (2003:140) conclui que, relativamente à posição dos constituintes na cadeia sintagmática, o efeito mais significativo na aplicação da marcação é dado pela tendência geral verificada no PB de se inserir a marcação na primeira posição (Nina, 1980; Guy, 1981a; Scherre, 1988; Lopes, 2001). Este aspecto, segundo a autora, vem também confirmar a aplicação da hipótese funcionalista, já que há uma tendência para se expressar a pluralização na primeira oportunidade (Kiparsky, 1972:195). Tal fenómeno será ocasionado pela aquisição de um modelo imperfeito do português, determinado pelo processo de transmissão linguística irregular. E não havendo registo deste tipo de marcação no PE, a explicação só poderá ser conotada ao fenómeno do contacto entre línguas que aconteceu no Brasil (cf. Guy, 1981a).

P. Andrade (2003:141) confirma também a propensão de os elementos mais salientes serem mais marcados, uma vez que, havendo na composição do SN um elemento morfofonologicamente mais perceptível, ocorre igualmente tendência para se aplicar a regra da pluralização. A ocorrência deste tipo de fenómenos dar-se-á porque os contextos mais salientes concedem ao ouvinte uma margem para perceber os desvios gramaticais do falante, levando-o a evitar o uso de formas não-padrão.

Relativamente às *marcas precedentes ao elemento analisado*, evidenciam que o falante tem noção da pluralização semântica. Desta forma, os numerais apresentam-se como um factor determinante para a inserção de marcas, em virtude de serem morfemas de conteúdo semanticamente transparentes (P. Andrade, 2003:112), possibilitando a aquisição da regra em situações de contacto entre línguas. Outro factor determinante para a aquisição da regra é o número de elementos que compõem o SN, já que, quanto mais reduzidas forem as estruturas deste, mais facilmente o falante processa as regras de marcação. Já no que diz respeito às variáveis sociais, para além de permitirem ajuizar que a comunidade de Helvécia se encontra em processo de mudança (os jovens pluralizam mais os elementos do SN do que os idosos), revelam também que os homens aplicam mais a marcação PL, em virtude de terem um papel mais interventivo na vida social de HEL-Ba (P. Andrade, 2003:144).

Face aos resultados obtidos, P. Andrade (2003:95-96) confirma, na análise distribucional do PL, a correlação entre posição linear e classe gramatical proposta por Guy (1981b:179), dado o dialecto de HEL-Ba apresentar SN's de estrutura mais reduzida e regular do que os dos dados urbanos analisados por Scherre (1988), ou seja, configurando apenas dois elementos: DET+N. Paralelamente, P. Andrade (2003:146) corrobora em parte o pressuposto de Guy (1981a), no que concerne ao facto de a variação resultar do amplo e intenso contacto entre línguas que aconteceu durante a formação sociohistórica do país. Assim sendo, a variação teria como consequência uma criouliização leve, decorrente da mudança de língua e reestruturação determinada pelo forte *input* de PtgL2 irregular (Baxter & Lucchesi, 1997).

Tieppo (2003) realiza também estudos quantitativos, numa perspectiva sintagmática, sobre a variação na concordância de número da fala de uma comunidade afro-brasileira do município de Rio de Contas, localizada na Chapada da Diamantina, com vista a poder determinar o peso que o contacto entre línguas teve na formação histórica da presente realidade linguística do Brasil. A variável linguística *configuração sintagmática do SN* evidenciou que as marcas são preferencialmente inseridas quando o SN é formado apenas por DET+NUM (p.e. *os quatro*). A autora alega que estes casos não configuram, em rigor, verdadeiros casos de concordância, já que apenas é necessário incluir marcação no primeiro elemento. Ora, como a tendência geral é a de inserir marcação

apenas neste elemento, encontra-se justificado o facto de a marcação ser praticamente categórica neste tipo de configuração sintagmática. Aliás, este aspecto é depois confirmado pela estrutura DET+N, que desfavorece altamente a concordância (Tieppo, 2003:5). Outro aspecto confirmado por Tieppo (2003:4-5), e já verificado em estudos anteriores, é a tendência de os elementos modificadores (possessivos, adjectivos ou quantificadores) surgirem bastante mais marcados quando se encontram à esquerda do núcleo, por oposição às situações em que estão à direita deste.

Recorrendo à variável *número de constituintes flexionáveis do SN*, Tieppo (2003:5) corrobora P. Andrade (2003) quanto à tendência de a regra de concordância ser mais aplicada em constituintes com apenas um item flexionável, isto é, quando o processamento linguístico se revelar mais facilitado, por ocorrer em estruturas simplificadas. Refira-se que idêntico comportamento não é registado no PA, uma vez que os dialectos de HEL-Ba e Rio de Contas se encontram em situação de mudança no sentido da aquisição da regra de concordância, enquanto a fala de Almojarife, ao que tudo indica, apresenta estabilização da variação. Estes aspectos encontram-se discutidos noutras partes do presente trabalho, nomeadamente no ponto 4.1.2.6.2 e no capítulo 5, no qual se observa o comportamento das variáveis extralinguísticas. Relativamente às variáveis sociais *faixa etária* e *escolaridade*, Tieppo (2003:7) verifica que tanto as faixas etárias mais jovens como as mais escolarizadas realizam mais a concordância, confirmando uma situação típica de mudança aquisicional no sentido do aumento do uso da regra de concordância nominal de número na comunidade.

Tendo em conta que comunidades afro-brasileiras, como a de HEL-Ba, se constituíram a partir de uma diversidade étnica que originou contacto entre línguas, especialmente das europeias com as do grupo níger-congo atlântico, proporcionou-se uma situação de transmissão linguística irregular, determinante para a aquisição do português com uma simplificação da sua estrutura gramatical, isto é, retirando-lhe os elementos menos relevantes para a comunicação emergente. Considerando ainda que este tipo de contacto se espalhou por todo o país nos primeiros quatro séculos da sua existência, com o português culto sendo influenciado pelo português popular com a chegada de imigrantes estrangeiros a partir do século XIX, é possível inferir que o fenómeno da variação na concordância de número no PB não se restringe a uma

determinada região ou classe social específica, sendo característico de toda a comunidade de fala brasileira. Por outro lado, é possível verificar também que vários factores de ordem linguística e extralinguística interferiram para determinarem as actuais diferenças marcantes entre o PB e o PE. É o que acontece a nível da CPL-var entre os elementos do SN, tendo tais diferenças que ser entendidas como não aleatórias (Tieppo, 2003:7), isto é, como fenómenos sistemáticos de variação previsível, resultantes do quadro alargado de contacto entre línguas (P. Andrade, 2003:146) e que originou uma situação de transmissão linguística irregular de L2 para L1 no Brasil (Baxter & Lucchesi, 1997).

Como se referiu no início do presente capítulo, o comportamento do fenómeno da CPL-var no português popular brasileiro poderá ser conotado ao das línguas nativas do grupo níger-congo atlântico, enquanto substratos dos CP's atlânticos e substratos directos ou ancestrais de variedades africanas de português. Deste modo, passaremos, de seguida, a dar conta de trabalhos pioneiros sobre CPL-var que têm sido recentemente levados a cabo sobre a CPL-var no SN de algumas variedades africanas de português, incluindo o PA. Abordaremos igualmente o modo como se processa a regra da concordância nominal de número em alguns CP's atlânticos (CCV e santomense), bem como em substratos destes, pertencentes ao grupo níger-congo atlântico. Posteriormente, faremos o enquadramento sociohistórico e linguístico da comunidade de Almojarife, visto a sua fala constituir o principal objecto de estudo do nosso trabalho. Portanto, apontaremos também algumas características do PA, estabelecendo paralelismos morfossintácticos compartilhados, a nível da estrutura do SN, entre este dialecto e línguas africanas do grupo níger-congo atlântico, CP's atlânticos, variedades africanas de português e variedades brasileiras de português.

## **1.2. Variedades africanas de português e concordância nominal variável de número**

### **1.2.1. Variedades africanas de português com substrato do grupo níger-congo atlântico**

#### **1.2.1.1. Português dos tongas (Roça Monte Café, São Tomé)**

Os estudos quantitativos acerca da CPL-var no SN de variedades africanas iniciaram-se com a pesquisa atomística de Baxter (2004) sobre o português da comunidade semi-



isolada dos tongas, fixada na Roça Monte Café, São Tomé. Os falantes da comunidade são descendentes de africanos falantes de umbundu e kwa, “contratados” no final do século XIX e durante o século XX. Estes trabalhadores foram transportados para a ilha de São Tomé, a fim de prestarem serviço nas grandes roças de café e cacau, com contratos renováveis por períodos de três a cinco anos. Porém, antes do estudo de Baxter (2004), já Rougé (1992) havia chamado a atenção para a variedade de português falada pelos tongas, que apresenta traços dos seus substratos africanos (muitas vezes do tipo koiné), mas que não teria sido afectada pelo esparso contacto com as línguas de outros residentes permanentes ou temporários da ilha, nomeadamente o santomense,<sup>18</sup> o CCV ou as variedades linguísticas de Moçambique (zona de Quelimane).

Segundo Baxter (2004:104), o complexo mosaico sociolinguístico que determinou a emergência do PT pode ser sumariado da seguinte forma:

- (i) o modelo predominante de português utilizado pela massa trabalhadora seria o PtgL2, o qual terá existido num *continuum* de variedades de interlíngua, face à constante chegada de contratados que eram integrados na comunidade;
- (ii) as primeira e segunda gerações de crianças tiveram como L1 um koiné de base umbundu, que aprendiam antes do PtgL2;
- (iii) as crianças aprenderam português nos campos de trabalho e outros locais da roça, principalmente através do contacto com o PtgL2 dos trabalhadores e, esporadicamente, com os trabalhadores falantes de PtgL1 ou capatazes europeus.

Face a este contexto de transmissão linguística irregular, o PTL1 da geração mais idosa apresenta substancial e consistente reestruturação a nível do SN (representações de género e número), do SV (marcação de tempo e aspecto) e na realização das relações de caso ou construção das negações. Estes traços morfossintácticos são partilhados com os das gramáticas do CCV ou do santomense (Baxter, 2002), enquanto que, no plano do nivelamento e desenvolvimento das regras variáveis de concordância, o PT estabelece semelhanças com as variedades vernáculas do PB, de Angola, de Moçambique, de São Tomé (variedade popular e variedade de Almojarife), de Cabo Verde (Mindelo) e mesmo de Macau.

Para observar o desenvolvimento diacrónico da regra de concordância PL no SN do PT, Baxter (2004) seguiu as metodologias propostas por Scherre (1988) e Lopes (2001), constituindo cinco variáveis linguísticas e quatro extralinguísticas.<sup>19</sup> Tais variáveis permitiram que Baxter (2004) descartasse a possibilidade de derivações direccionadas de fenómenos internos ao PE estarem na origem da variação registada na regra de concordância PL do SN (cf. Naro, 1981; Naro & Scherre, 1993, 2000). Contrariamente, Baxter (2004:121) conota a variação à situação de contacto que envolveu componentes de PtgL2 influenciadas pelo umbundu e, simultaneamente, pela aquisição deste. Deste modo, a marcação revela tendência para ser inserida no elemento imediatamente à esquerda do SN, deixando, depois, de haver inserção de marcas nos itens seguintes. Como os falantes do PT eram confrontados com modelos difusos, terão optado por uma estrutura simplificada, isto é, isenta de marcação. Efectuou-se, desta forma, uma reconfiguração a nível de marcação no SN, a qual estabelece paralelismo com a do grupo níger-congo atlântico (balanta e wolof), substrato do CVV. Daí o autor entender que, no contexto sociohistórico específico de Monte Café, a L1 africana tenha sido, directa ou indirectamente, determinante no processo das transferências registadas na aquisição do PtgL1, via DLP's defectivos do PtgL2. Isto mesmo seria confirmado com uma observação sobre a marcação PL no desempenho dos falantes mais novos da comunidade, conforme a procedência dos pais: se nascidos em África ou na roça Monte Café. De facto, os falantes com pais nascidos na roça favorecem a inserção de pluralização, enquanto os falantes com pais nascidos em África a inibem (Baxter, 2009:283-284). O efeito, segundo Baxter (2009:284), radica no PtgL2 falado pelos pais dos informantes, o que, em última instância, vem corroborar Guy (1981a:300-301), quanto ao facto de o substrato africano ser uma potencial fonte da variação da concordância de número no SN, localizando a marcação na posição à esquerda do núcleo, tal como acontece no PA e PVB.

Os aspectos linguísticos e sociolinguísticos compartilhados pelas falas das comunidades da Roça Monte Café e de HEL-Ba levaram Baxter (2009) a produzir um estudo quantitativo, numa perspectiva comparada, sobre a CPL-var de ambos os dialectos. A metodologia adoptada seguiu uma perspectiva que combina teorias estruturais e aquisicionais, permitindo as primeiras que se observassem quer a

configuração estrutural da concordância quer o lugar da inserção do morfema PL, e as segundas que se percebessem as motivações que estarão na base da variação de número. As análises estruturais tiveram como base de partida as visões de autores que postulam a ocorrência de uma categoria funcional SNUM entre o DET e o SN (Picallo, 1991; Carstens, 1991; Ritter, 1991; Valois, 1991; Bernstein, 1993; Cinque, 1994; Montrul, 2004), entendimento este que determina que as projecções funcionais são motivadas pela semântica. Simultaneamente, foi também adoptado o pressuposto de que a noção de concordância de número é reflexo de uma relação sintáctica local (Chomsky, 2001b), isto é, que a concordância é computada entre os elementos do SDET, por acção de uma relação de C-comando a partir do DET, com vista a acontecer o mapeamento e consequente pareamento (*matching*) de traços. Além destas noções, foram também levados em consideração os pressupostos já avançados por Lopes (2001) e P. Andrade (2003) acerca da teoria cognitiva dos Quatro Morfemas (Myers-Scotton, 1997, 2001; Myers-Scotton & Jake, 2000a; 2000b), bem como a visão compartilhada por Costa & Silva (2006b) acerca dos princípios sobre morfologia distributiva e não distributiva. Estes, tendo como base os trabalhos de Galves (1993), Scherre (1994) e Embick & Noyer (2001), entendem que o PE e o PB manifestam distintas realizações do morfema associado à pluralidade. Como tal, os morfemas podem associar-se aos núcleos, independentemente dos processos de movimento, realizando-se de duas maneiras: (i) como morfemas dissociados, que não figuram na sintaxe, mas são inseridos após o *spell-out*, como um reflexo de configurações sintácticas; (ii) como morfemas *singleton*, que se associam aos elementos funcionais, responsáveis pela informação de número e referência.

Relativamente aos aspectos aquisicionais, a presença de PtgL2 nas duas línguas, determinada por situações de contacto que envolvem estágios intermédios de ASL em simultâneo com aquisição de L1 a partir do *input* de modelos de L2, isto é, fases de bilinguismo, fez com que se levassem em consideração as teorias sociolinguísticas que avaliam o contributo dos processos de ASL para análise da variação. Também foi tido como relevante o facto de a aquisição envolver mudança de alvo, dado que, em fases iniciais, a L2 das massas escravizadas teria várias fontes: (i) a L1 dos administradores; (ii) a L2 de outros escravos, baseada na L1 dos administradores; (iii) a L1 da população

escrava, baseada na L2 de outros escravos (cf. Smith, 2006). Paralelamente, foi ainda tida em conta a situação de desequilíbrio sociodemográfico entre contingentes escravizados (a grande maioria) e não escravizados (uma acentuada minoria), que levou o contacto com a LA a tornar-se esparso. Desta forma, ao configurar-se um cenário propício à mudança de língua, substancial material da L2 pôde ser não só incorporado na nova L1 da comunidade nas fases iniciais da aquisição mas também transmitido geracionalmente (Winford, 2003a:245-247), fossilizando-se algum desse material. Simultaneamente, a presença de bilinguismo na comunidade determinou a ocorrência de transferências de elementos das L1's originais para a nova L1 (cf. Thomason & Kaufman, 1988; Lefebvre, 2003; Montrul, 2006; Lefebvre, White & Jourdan, 2006b; Siegel, 2008).

Atendendo não só ao facto de diversos estudos sobre CPL-var do PB apresentarem a posição linear do item pluralizável como a chave para a inserção da flexão de PL mas tendo também em conta que Guy (1981b:179) advoga existir uma relação intrínseca entre posição linear do item e sua classe gramatical, fixando o eixo da marcação PL nos determinantes e originando uma configuração de marcação que terá respaldo no sistema linguístico bantu e kwa (Guy, 1981a, 1989, 2005), Baxter (2009) observa a forma como estas duas variáveis, em concomitância com a variável social idade, podem intervir na forma como se perfila e desenvolve aquisicionalmente a variação PL nas duas comunidades, uma vez que ambas apresentam configurações sintagmáticas de complexidade reduzida. Para tanto, Baxter (2009) tem ainda como primordial, para as suas análises, determinados factores que podem ter influência na forma da L2 dos adultos e da L1 irregular das crianças. Para a primeira, devem então ser considerados os seguintes aspectos: (i) simplificação, isto é, presença na L2 de formas reduzidas e regularizadas (Klein & Perdue, 1997; Siegel, 2004a; Mather, 2006), orientadas pelo tipo de *input* (p.e. frequência, saliência e transparência semântica ou pragmática) e pela GU; (ii) transferência e relexificação, ou seja, presença de estruturas da L1 do aprendente na sua L2 incipiente (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006); (iii) ASL imperfeita, que ocorre em virtude de o falante adulto não adquirir os traços das novas categorias funcionais (*Failed Functional Features Hypothesis* – FFFH) (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995;

Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto já ter sido ultrapassado o período crítico para a ASL. Quanto à L1 irregular das crianças, não podem ser descartados os recentes achados sobre aquisição da morfologia de PL no PB, nomeadamente do modo como as classes morfológicas fechadas (*closed class morphemes*) intervêm no processamento dos DLP's disponíveis para a aquisição de L1 e orientação do desenvolvimento da pluralização (Ferrari-Neto, 2003; Corrêa, Augusto & Ferrari-Neto, 2005; Ferrari-Neto, Corrêa & Augusto, 2005).

Os resultados de Baxter (2009) revelam que, em ambos os dialectos (PT e HEL-Ba), ocorre uma forte correspondência entre posição linear e classe gramatical, sendo a posição imediatamente pré-nuclear o ponto de inserção da pluralização. Nesta conformidade, Baxter (2009) entende que o perfil inicial da variação da comunidade bilingue de Monte Café deve ser essencialmente atribuído ao substrato africano. Quanto às variáveis extralinguísticas, ajudam a perceber que a variação aponta para mudança, no sentido da aquisição da morfologia padrão e aplicação das regras de concordância. Por seu lado, Figueiredo (2009a) aponta para o PA um perfil idêntico ao de HEL-Ba, com os itens imediatamente pré-nucleares a revelarem tendência para serem os primeiros a receberem a marcação mórfica de PL. Como tal, “*estes elementos têm um papel mais activo na construção estrutural da concordância, relegando os itens pós-nucleares para funções sintácticas de mera adjunção*” (Figueiredo, 2009a:47). Baxter (2009), com recurso ao mecanismo formal misto que propõe, indica que a distribuição da concordância PL no SN segue um padrão uniforme em HEL-Ba e no PT, e que, a avaliar pelos pressupostos de Figueiredo (2009a), será também idêntico ao do PA. Deste modo, para se perceber como se processa a ordem de entrada da marcação PL no SDET de HEL-Ba e no PT, Baxter (2009:293) avança com o seguinte esquema:

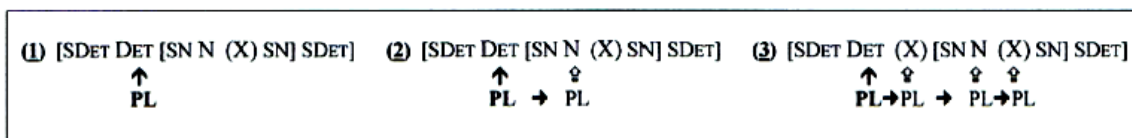


Fig. 1. Dialecto de HEL-Ba e PT: desenvolvimento da atribuição de PL.

A configuração mostra que o desenvolvimento da concordância PL se inicia com a introdução de um PL *singleton* na categoria funcional DET, que serve de âncora para o

controle da pluralização. Quando essa posição não é ocupada por um elemento DET, o SN torna-se alvo de PL, em virtude das suas características lexicais, fazendo “*lembrar o papel fulcral do PL no classificador nominal na tipologia níger-congo, e portanto no substrato níger-congo atlântico (balanta e wolof) do crioulo de Cabo Verde*” (Baxter, 2009:292). Portanto, a concordância torna-se parcial e é atribuída a partir da posição DET, isto é, a partir do elemento pré-nuclear adjacente. Posteriormente, o PL desenvolve-se, ainda com características de concordância parcial, através de morfemas dissociados, que são atribuídos aos elementos mais próximos do *fulcro*. A ordem de atribuição contempla os elementos do SDET, primeiro, e os itens à direita do núcleo, depois. Não obstante, o perfil da âncora pré-nuclear é mantido geracionalmente. De certa forma, é aqui estabelecido um paralelismo com os pressupostos advogados por Lucchesi (2000a) no seu estudo sobre a concordância variável de género (CGEN-var) da fala de HEL-Ba, e que, a propósito da concordância plural no SN, sustentam actuar uma correlação entre Princípio da Simplicidade, que leva a concordância a fixar-se inicialmente nas estruturas mais simples dos SN’s, constituídas por nome e determinante, e Princípio da Integração, que faz com que a marcação ocorra mais nos itens flexionáveis à esquerda do núcleo do SN (com incidência na posição adjacente) e menos nos elementos à sua direita.

Desta forma, se houve realmente um modelo inicial baseado no PE, será lícito concordar com o pressuposto teórico de que os africanos e os seus descendentes exageraram esse modelo (Naro, 1981; Naro & Scherre, 1993, 2000, 2007a). Contudo, a semelhança entre o papel da posição pré-nuclear para o sistema de marcação PL de HEL-Ba e Monte Café com CP’s, como o CCV (Castro & Pratas, 2006), contribui para que se entenda esta estrutura como resultante das características sociohistóricas compartilhadas pelos três dialectos. Portanto, parece bastante lógico assumir que o africano teria influenciado o desenvolvimento do PL, havendo ou não variação no modelo original (Baxter, 2009). Desta forma, tanto a hipótese contactista como a hipótese derivista são capazes de explicar a realização da CPL-var, propondo Baxter (2009:293) que ambas passem a figurar num modelo sociolinguístico único para o PB. Ainda assim, em pesquisas futuras sobre esta língua, deverá ser tido em conta que as

pressões exercidas na variação quer pelo contacto quer pela deriva terão sido determinadas pelas diferentes circunstâncias em que emergiu a variedade em análise.

### **1.2.1.2. Português vernáculo de Angola**

Não temos conhecimento de existirem estudos numa perspectiva sociolinguística quantitativa sobre o português vernáculo de Angola (PVA). Contudo, vários autores têm feito referência à concordância variável neste, estabelecendo paralelos morfossintáticos e morfofonológicos quer com os substratos africanos quer com o PVB (Marques, 1985 [1983]; Gärtner, 1989, 2003; Asher, 1994; Mello, 1997; Baxter, 1998; Chatelain, 2001; Mingas, 2000; Zau, 2001; Veiga, 2002a; Figueiredo, 2003; Inverno, 2004, 2005, 2008, 2009; Cabral, 2005a, 2005b). Ainda que o PVA seja língua de classes sociais específicas em Angola, estabelecidas nos centros urbanos costeiros, numa situação de nítida diglossia, as partilhas estruturais entre a variedade de português e os dialectos do substrato, nomeadamente a nível da CPL-var no SN, resultam de processos paralelos de formação no primeiro, ou seja, de reestruturação parcial, em consequência de situações de transmissão linguística irregular. Após a sua fixação no litoral do reino do Congo, não como colonizadores, mas como parceiros no tráfico de escravos e marfim, os portugueses rapidamente assumiram o controlo de Loando, em 1576, estabelecendo a colónia de Angola (Parreira, 1989; Ogot, 1999:278). O domínio foi assegurado por uma elite de afro-portugueses, resultado de uma política de miscigenação entre europeus e mulheres da nobreza angolana (Chatelain, 2001:14). Esta elite, sendo bilingue em português e línguas bantu, assumiu postos na função pública e servia de interlocutora entre os portugueses e os chefes indígenas, a fim de se garantir a paz com estes e se manterem activas as rotas que possibilitavam o resgate e o comércio de escravos (Venâncio, 1996:51).

Assim, e ao contrário do que sucedeu no Brasil, onde a variedade de português reestruturado acabou por se expandir por todo o território, gerou-se em Angola uma situação de diglossia estável, englobando uma elite afro-portuguesa bilingue, detentora de um código linguístico gramaticalmente elaborado (o PE), por um lado, e uma grande maioria indígena, falante de variedades vernaculares de português, por outro lado. Como o desenraizamento desta população não se realizou em grande escala a nível interno,

também não aconteceu uma situação que impedisse a inteligibilidade linguística entre africanos. Simultaneamente, como a aculturação destes à língua e cultura portuguesas foi bastante incrementada, sobretudo a partir do século XVIII (Carreira, 1979; Venâncio, 1996; Ogot, 1999), não se verificou uma estabilização generalizada das variedades básicas de português faladas no território, que pudesse estar na origem do PVA. Contudo ocorreram interferências mútuas entre o português e línguas bantu, sobretudo nas zonas residenciais periféricas às cidades litorais (os musseques). Deste modo, no século XIX Schuchardt (1888:249) referia-se já ao português “*mestiço*” de Luanda, um composto de “*palavras portuguesas acomodadas ao genio do bundo*”, enquanto Valman (1978:22), posteriormente, fazia também menção a um “*pequeno português*”, falado nos principais centros urbanos angolanos. Estes aspectos levam a crer que os autores davam conta de interferências, configurando, possivelmente, um processo de “*empréstimo estrutural moderado*” (Inverno, 2004:4), situação que requer, por considerável período de tempo, não só um acentuado bilinguismo por parte dos falantes das línguas que concedem os empréstimos mas também a manutenção das L1’s dos falantes envolvidos na situação de empréstimo (Thomason & Kaufman, 1988:37).

Dois outros aspectos concorreram para acentuar a situação de diglossia em Angola: (i) a ocupação do território limitada à orla marítima até meados do século XX (Gann & Duignan, 1981:9; Chatelain, 2001:5), que restringiu o uso do PE aos centros urbanos costeiros; (ii) o atenuar do processo de miscigenação com a posterior chegada de grande quantidade de mulheres provenientes da metrópole e que acompanharam os seus maridos no processo de colonização do interior, a partir da segunda metade do século XX. Nesta fase, também as políticas de aculturação levadas a cabo pelo Estado Novo visaram a implementação do PE, já que as oportunidades de ascensão social, nomeadamente através da colocação em cargos da administração pública, eram apenas concedidas aos que demonstravam “*saber ler, escrever e falar fluentemente português*” (Bender, 1980:216).<sup>20</sup> Contudo, uma vez que a maioria da população não tinha acesso à educação, o português era aprendido pelo contacto prático e funcional, o que fez com que as línguas do substrato fossem afectadas não só por empréstimo lexical massivo mas também pelo sistema fonológico e morfológico do português (Oliveira, 1990:69-89). Tal levou a que, posteriormente, ocorresse uma mudança para o português, já que este se



assumiu como uma das L1's dos falantes bilingues das classes menos escolarizadas que se fixaram nas zonas costeiras. Emergiu, deste modo, uma variedade reestruturada de português marcada por influências dos substratos. Este processo de mudança linguística, apesar de apontar para uma situação de formação comparável à do Brasil, é mais recente e restrito social e geograficamente, não estando, portanto, concluído (Inverno, 2004:5).

No que concerne à CPL-var no SN do PVA, tem sido indicada em diversos estudos a ausência de concordância entre os elementos não-nucleares do SN e o núcleo deste, um fenómeno que alguns autores apontam como sendo um dos traços morfossintáticos mais evidentes da variedade reestruturada de Angola (Cabral, 2005:73). Baxter (1998), a exemplo do que apontaria posteriormente para o PVB (Baxter, 2009), entende que o fenómeno resulta da convergência entre posição linear e classe gramatical, muito por influência do sistema de marcação PL dos substratos bantu (umbundu, kimbundu e kikongo). Assim, na aprendizagem do PtgL2, os nomes terão sido interpretados como elementos invariáveis, enquanto os elementos não nucleares do SN foram entendidos como afixos variáveis de marcação semântica pré-nuclear, a exemplo do que sucede nos dialectos do substrato. Daí que a marcação de número ocorra neste tipo de elementos, colocados à esquerda do núcleo do SN (Inverno, 2005:135), e não no núcleo (Marques, 1985 [1983]:218; Gärtner, 1989:159; Vilela, 1999:56; Inverno, 2004:5, 2005:134). Este aspecto acaba mesmo por determinar fenómenos de apócope sobre o *-s* final que não detém função de marcador de número PL em algumas palavras:

[1] PVA:  
*Quando acabar as férias\_* (Inverno, 2005:133)

### 1.2.1.3. Português de Moçambique

O contexto sociohistórico que levou à formação do português de Moçambique (PM) é em tudo comparável ao de Angola. Assim, estamos perante uma variedade não-nativa de uma língua ex-colonial, adquirida natural ou formalmente como L2 pela maioria das crianças falantes de L1's bantu, com especial destaque para o emakhuwa, ainda que este seja usado apenas por 26,3% dos utilizadores das línguas nativas (Instituto Nacional de Estatística de Moçambique, 2010). Quanto à presença dos portugueses em Moçambique, data do século XV (1498). Contudo, por razões económicas e culturais, a difusão do

português foi bastante incipiente durante os primeiros quatro séculos de colonização da região. De facto, quando os portugueses desembarcaram nesta ex-colónia, já os árabes se haviam aí estabelecido, controlando os principais centros urbanos. Assim, só em finais do século XIX/princípios do século XX é que os europeus conseguiram o controlo do território, nomeadamente através das chamadas campanhas de “pacificação” (1886-1918) (Gonçalves, 2004:230). Por outro lado, como a atenção da metrópole estava essencialmente dirigida para as riquezas do Brasil, o domínio total da região terá sido adiado em virtude de a mesma ser considerada periférica para o contexto do império, como bem o demonstra o facto de a sua administração ter sido feita a partir da Índia até 1752.

A exemplo do que aconteceu em Angola, os portugueses fixaram-se na zona costeira, sendo aí que ocorreu maior contacto entre o PE e as línguas indígenas, enquanto no interior a população continuava a usar as suas L1's. Este aspecto originou uma situação de diglossia, que se mantém até aos dias de hoje, já que nas zonas rurais o português é apenas utilizado em escolas, onde o seu ensino é compulsivo (Gonçalves, 2004:233). Aliás, a tardia implementação de uma política educacional consistente<sup>21</sup> terá também acentuado a situação periférica da ex-colónia. As consequências foram idênticas às de Angola, criando-se uma população bilingue minoritária e sendo as línguas do substrato, faladas pela maioria da população (98%) (Firmino, 1995:221), afectadas lexical, fonológica e morfologicamente com elementos do português. Este aspecto impediu a emergência de um pidgin e, posteriormente, com a independência do país (1975), a comunidade de falantes do português cresceu consideravelmente, devido a razões socioeconómicas e ao incremento da rede de educação (Gonçalves, 2004:232). Assim, o número de falantes de PtgL2 aumentou cerca de 15% (Firmino, 1995:224; Gonçalves, 2004:232), acentuando a variação linguística que caminha no sentido da mudança, uma vez que o número daqueles que usam o PtgL1 reestruturado marcado por influências dos substratos cresceu mais de 50% (Firmino, 1995:224; Gonçalves, 2004:232). Como se verifica, este processo de variação linguística ainda não estabilizado (Carvalho, 1991:62; Gonçalves, 2004:236), tal como em Angola, é recente e restrito social e geograficamente, mas apontando para uma situação de formação que pode ser comparada à do Brasil.

Em trabalho sobre a concordância variável no SV do PML2, numa perspectiva generativista, Gonçalves (2004) analisa o modo como factores linguísticos e extralinguísticos podem bloquear o restabelecimento de novos valores paramétricos (cf. Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashes, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), tornando o processo de mudança da L2 menos gradual que o da L1 e também menos dramático que o da crioulização. O estudo evidencia ainda como o conhecimento prévio da gramática da L1 bantu dificulta o processo de aquisição do PtgL2, necessitando este, para valores de determinados parâmetros da LA, a exposição a DLP's extra. Pelo facto de Moçambique constituir uma sociedade pós-colonial, estes DLP's extra poderão não se encontrar disponibilizados, principalmente devido à preponderância quantitativa de falantes do PtgL2. Em consequência de tal, a instanciação em cadeia de novos/errados parâmetros, que origina a variação, torna-se cada vez mais frequente no *input*, eliminando a possibilidade de um reajustamento paramétrico e fazendo com que os novos/errados parâmetros sejam tomados como modelos para uma nova reinstanciação paramétrica do PML1. Consoma-se, desta maneira, a mudança que acentua a distância tipológica entre a LA e a capacidade gramatical (ou língua-I) atingida pelos falantes do PM (Gonçalves, 2000:255). Em resumo: alguns traços das L1's bantu são retidos no PtgL2 e sobrevivem na gramática do PM, tal como algumas propriedades gramaticais da LA se perdem no *continuum* de aquisição (Gonçalves, 2004:229), originando uma “*moçambicalidade gramatical*” (Gonçalves, 2000:255).

Tendo não só em conta que a sociohistória do PM é bastante recente mas também considerando que a gramática do PE não foi plenamente abandonada por todos os falantes moçambicanos, os núcleos comunitários operam com mais do que uma gramática. Ocorre, desta forma, uma espécie de diglossia não só no plano intercomunitário mas até a nível “interno” dos próprios falantes (cf. Lightfoot, 1999b:2). Tais espécies de diglossia tornam difícil precisar qual das gramáticas gera determinadas produções linguísticas, seja no plano da realização de grupo seja a nível individual. Este aspecto faz com que a frequência de traços não-padrão no discurso dos falantes do PM não seja uniforme para todas as comunidades nem mesmo para todos os membros de uma determinada comunidade, originando diferentes variedades comunitárias e

subvariedades nestas, sendo tais traços disponibilizados através do *continuum* dialectal. Contudo, e de um modo mais geral e uniformizado, poder-se-á dizer que o *continuum* dialectal do PM apresenta uma extensão que vai desde a variedade não-padrão falada pelos indivíduos menos escolarizados até à variedade padrão das classes superiores e com mais elevado nível de escolaridade.

A existência deste *continuum* em Moçambique levanta questões metodológicas, quando se toma determinada subvariedade como objecto empírico para o estudo da gramática do PM. No caso concreto da subvariedade de português falada no Maputo (PMp), é na área da sintaxe (léxico-sintaxe, morfossintaxe e sintaxe propriamente dita) que ela mais se diferencia do PE (Carvalho, 1991; Gonçalves, 1997). No que concerne aos fenómenos sintácticos e morfossintácticos que apresentam desvios mais sistemáticos relativamente ao modelo da LA, podem apontar-se “*as regras de utilização dos artigos e pronomes, os mecanismos de encaixe de orações subordinadas, e as regras de concordância verbal e nominal*” (Gonçalves *et alii*, 1998:36). Ainda no caso concreto da variação a nível da concordância nominal, estes autores referem que o nome se apresenta no singular, em ocorrências que exigem a flexão no PL. Por outro lado, a pluralização apenas aparece inserida morfológicamente nos artigos e pronomes (exemplo [2]), ou é determinada pela presença de um numeral não singular (exemplo [3]) (Gonçalves *et alii*, 1998:107-108):

[2] PMp:  
*Eu não mostro **os meus dente**\_* (Gonçalves *et alii*, 1998:108)

[3] PMp:  
*O lobolo aumentou, eram **quatro conto**\_* (Gonçalves *et alii*, 1998:108)

Em adição, também se verifica a falta de concordância em número do núcleo nominal com adjectivos (exemplo [4]), pronomes (exemplo [5]) e artigos (exemplo [6]) (Gonçalves *et alii*, 1998:108):

[4] PMp:  
*... com **uns preços um bocado baixo**\_* (Gonçalves *et alii*, 1998:111)

[5] PMp:  
***Meus neto**\_ são dezasseis neto\_* (Gonçalves *et alii*, 1998:110)

- [6] PMp:  
*Os vizinho\_ hãode vir.* (Gonçalves *et alii*, 1998:110)

Tudo indica que o fenómeno da CPL-var no SN resulta da convergência entre posição linear e classe gramatical, evidenciando a forma específica como as propriedades gramaticais particulares das L1's dos falantes em aquisição do PtgL2 podem interferir no processamento do *input*, levando à posterior instanciação de novos/errados parâmetros (Gonçalves, 2004:235).

Moreno & Tuzine (1998) elaboram também um estudo sobre o PMp, mas numa perspectiva sociolinguística laboviana que visa analisar a relação existente entre a variação linguística e os factores sociais. Trabalhando unicamente com valores percentuais,<sup>22</sup> os autores observaram a interferência de quatro variáveis sociais (*escolaridade, idade, profissão e local de residência*) no léxico-sintaxe (pronomes pessoais reflexivos) e morfossintaxe (concordância verbal e concordância nominal) do PMp. Apesar da dimensão reduzida do *corpus* analisado, os resultados sugerem que os fenómenos da variação do PMp não podem ser explicados recorrendo-se unicamente a uma ou duas variáveis isoladamente, já que a variação é motivada por vários factores actuando em conjunto. Ainda assim, tudo parece indicar que o nível de escolaridade e o tipo de profissão são as variáveis sociais mais relevantes para a ocorrência da variação registada no discurso dos informantes. Não obstante, Moreno & Tuzine (1998:86) alertam para o facto de os fenómenos de variação terem sido observados a partir de dados “*recenseados numa pequena amostra e, por isso mesmo, só futuros estudos poderão confirmar ou não as presentes conclusões*”.

Assim, seria Jon-And (2008) quem se debruçaria novamente sobre o PMp numa perspectiva quantitativa, observando a CPL-var no SN produzido por 18 informantes de ambos os sexos, distribuídos por três faixas etárias (20-40 anos; 41-60 anos e +60 anos).<sup>23</sup> Curiosamente, a autora levou em conta a variável social *idade*, que os resultados de Moreno & Tuzine (1998:84) haviam considerado não-significativa para a motivação da variação, em virtude de o português dominar cada vez mais o registo oral da vida quotidiana no Maputo, principalmente entre as gerações menos idosas. Como Jon-And (2008) não constituiu as variáveis *escolaridade e profissão* para o seu estudo, ficam por

confirmar os achados de Moreno & Tuzine (1998) no que concerne à maior influência destas variáveis na variação registada no PMp.

Para além das variáveis extralinguísticas *idade* e *sexo*, Jon-And (2008) observou ainda o peso exercido pelas seguintes variáveis linguísticas na CPL-var no SN: *saliência fónica, posição em relação ao núcleo, contexto fonológico seguinte* (excluída por ter pouco efeito), *marcas precedentes* e *classe gramatical* (ambas excluídas por revelarem interferência com outras variáveis). As discussões detalhadas em torno deste estudo serão apresentadas nos capítulos 4 e 5, dedicados à análise e comparação dos nossos resultados com os de outras pesquisas. Não obstante, é possível adiantar que as conclusões da pesquisa apontam para que as variáveis fonológicas detenham efeitos semelhantes aos dos registados para o PT e para o Brasil, enquanto a variável linguística *posição em relação ao núcleo do SN* se apresenta como aquela que mais motiva a variação no PMp, “*possivelmente por causa de influência de línguas bantu, L1s dos informantes maputenses*” (Jon-And, 2008:7).

## 1.2.2. Variedades africanas de português com substrato crioulo

### 1.2.2.1. Português cabo-verdiano e seu substrato (crioulo de Cabo Verde)

O CVV e o santomense têm em comum o facto de possuírem, como substrato, línguas do grupo níger-congo atlântico, estabelecendo paralelismos formais no que concerne ao sistema de marcação PL e que podem ser conotados às L1's dos ancestrais aprendentes da interlíngua. De facto, ambos os crioulos podem inserir a marcação PL apenas no elemento pré-nominal, permanecendo inalterado o radical dos restantes itens do SN:

[7] CCV:  
*Góra, nu ta kume uns banana.* (Pereira, Arim & Carvalho, 2006:12)  
Agora 1PL ASP comer IND banana  
“Agora, nós vamos comer umas bananas.”

[8] FORRO:  
*Ê sêbe tudu inen kwa se.* (Hagemeijer, 2007:122)  
3SG saber tudo 3PL coisa MSPEC  
“Ele sabe todas estas coisas.”

Este sistema de marcação estrutural transitou quer para o PCV (exemplo [9]) quer para o PA (exemplo [10]), permitindo considerar a probabilidade de as distintas variedades de português adquiridas em situação de contacto por falantes de diferentes CP's atlânticos manifestarem idênticos padrões de variação. Por outro lado, o paralelismo deste tipo de marcação com o do PB argumenta a favor do pressuposto da influência das ancestrais línguas africanas no sistema de pluralização deste (cf. Guy, 1981a:301-302):

[9] PCV:  
*As coisa foram subindo.* (Jon-And, 2009:1)  
“As coisas foram subindo”

[10] PA:  
*Vem do pé todas dia paqui* [FRANCM2]  
“Vem a pé todos os dias para aqui”

Formado algumas décadas depois do início da ocupação do arquipélago de Cabo Verde por parte dos portugueses e escravos resgatados na costa africana (actuais regiões do Senegal e Guiné-Bissau) e transplantados pelos primeiros, o CCV tem na sua génese condições sociolinguísticas especiais, já que a comunidade relativamente estável do arquipélago era constituída por falantes de diferentes L1's do grupo níger-congo atlântico. Muitos destes falantes já falariam um pidgin rudimentar de base lexical portuguesa, que deixaria de ser visto como L2 de comunicação emergencial para assumir o estatuto de língua de comunicação interétnica na comunidade. Sendo esta transplantada, apesar de ser geográfica e culturalmente heterogénea, a sua língua interétnica evoluiu para um crioulo, uma vez que os falantes deixaram de usar os seus idiomas de origem, salvo em situações de ambiente familiar (Pereira, Arim & Carvalho, 2006:3). Contudo, das dez ilhas que constituem o arquipélago, apenas nove foram povoadas, mas em épocas diferentes.<sup>24</sup> O tempo que mediou entre a primeira e a última fase do povoamento do arquipélago e o modo como o mesmo aconteceu nas diferentes ilhas foram determinantes para que ocorresse uma variação dialectal acentuada, nomeadamente entre as ilhas de Santiago e de S. Vicente (Pereira, Arim & Carvalho, 2006:3).

Desde a sua formação até aos nossos dias, o CCV tem mantido um contacto quase exclusivo com a língua portuguesa. Privados das relações com as suas terras de origem, os africanos e seus descendentes acabaram por deixar de usar as suas L1's. Desta forma, em Cabo Verde existem apenas duas línguas nacionais: o CCV e o português, que ganhou o estatuto de língua oficial, embora se preveja para breve o estatuto de “língua de ensino e da administração” para o crioulo. Estigmatizado durante séculos, o CCV, apesar de ser a L1 de todos os cabo-verdianos, não era estudado nas escolas nem usado em contextos oficiais. A política de sucessivos governos portugueses, que visava erradicar este dialecto e substituí-lo pelo português, chegou mesmo ao ponto de proibir o seu uso em locais públicos. Em consequência de tal, o português e o CCV passaram a coabitar em estado de afastamento linguístico, com o segundo sendo remetido para o espaço familiar, enquanto o primeiro se tornava a língua elitista do ensino e da vida pública, único idioma de acesso à literatura e aos bens sociais. Estas barreiras artificiais fizeram com que alguns falantes do CCV tentassem torná-lo o mais parecido possível com o português que sabiam. Como muitos deles não eram escolarizados, emergiram duas variedades regionais no CCV: a das ilhas do Barlavento, isto é, do Norte (Santo Antão, São Vicente, São Nicolau, Boavista, Sal, Santa Luzia), mais próxima do português actual; e a das Ilhas do Sotavento, ou seja, do Sul (Santiago, Maio, Fogo, Brava), classificada como acrolecto. O conjunto destes aspectos conduz, pois, a um *continuum* de variedades, umas mais fundas (menos assimiladas) e outras mais leves (mais assimiladas). Assim, mesmo que olhemos apenas para a variedade de Barlavento, verificam-se algumas diferenças nos registos dos falantes do interior, dotados de fraca ou nenhuma escolarização, que falam um crioulo fundo, e dos falantes das zonas urbanas, mais escolarizados, que utilizam um crioulo leve. Por outro lado, o *continuum* resultante da presença generalizada do português, entre variedades basilectais e acrolectais do CCV, tem levado a que todas elas se encontrem em processo de descrioulização (Pereira, Arim & Carvalho, 2006:3).

Müller (1994a:87), analisando a aquisição do alemão e do francês como L1's em crianças bilingues, defende que a mesma sofre um processo de maturação, razão pela qual as gramáticas não apresentam marcação de género ou número no estado inicial da aquisição. Portanto, estas noções gramaticais só são disponibilizadas pela GU num



estádio posterior de desenvolvimento gramatical, pelo que Müller (1994a:62) parte do pressuposto que as gramáticas adultas possuem pelo menos uma projecção funcional acima do SN (Fig. 34), no qual se encontram alojados os traços flexionais dos nominais que não estão disponibilizados na fase inicial da aquisição. Esta será a situação da fase inicial de aquisição dos pidgins que estiveram na base do CCV e do santomense, já que os indivíduos em idade adulta, quando submetidos à aquisição de uma L2 emergente em situação de contacto massivo, não assimilam o sistema de flexões da LA, passando a sua interlíngua a registar ausência ou redução destas. Posteriormente, os falantes responsáveis pela estabilização estrutural, face à continuada ausência de realização do DET nos DLP's do *input* que recebem, e conforme sugerido pela FFFH, poderão não reinstanciar os traços perdidos pelas gerações antepassadas, pelo que continuam a apagar, em muitas das suas realizações, as marcas formais de plural, como sucede no PCV e no PA.

Em português, alguns nomes genéricos (p.e. *O homem é um ser inteligente. / A coisa agora está preta.*) também podem ser usados no singular, com referência PL. Contudo, e ao contrário do que sucede nos crioulos referidos, nos quais os exemplos [7] e [8] não permitem outras alternativas de pluralização, os genéricos portugueses admitem várias realizações: (i) recorrendo ao singular, antecedido de artigo definido (exemplo [11a]); (ii) utilizando o PL, antecedido de artigo definido (exemplo [11b]); (iii) fazendo recurso ao singular, antecedido de artigo indefinido (exemplo [11c]); (iv) aplicando o singular e omitindo o artigo (exemplo [11d]); (v) servindo-se do PL e elidindo o artigo (exemplo [11e]):

- [11] PTG:
- a. *A cobra é peçonhenta.*
  - b. *As cobras são peçonhentas.*
  - c. *Uma cobra é um bicho peçonhento.*
  - d. *Cobra é bicho peçonhento.*
  - e. *Cobras são bichos peçonhentos.*

Para os casos em que a marca formal de PL não é realizada, Scherre (2001:98-99) considera estar-se perante representações abstractas de pluralização, em que a carga semântica de pluralidade é percebida pelos falantes a partir do contexto, uma situação que acontece também nos crioulos acroletais de Cabo Verde. Daí que, no

português, se flexione apenas o adjunto adnominal com determinados etnónimos (p.e. *os bantu*), estabelecendo-se aqui um sistema de pluralização estrutural absolutamente idêntico ao dos exemplos [7], [9] e [10]. Contudo, este tipo de marcação não representa a regra geral do português, ou seja, não constitui sistematização, como acontece no CCV, PCV e PA, nos quais as motivações que a determinam são distintas das do português, abrangendo outras categorias de nomes.

Depreende-se então que os crioulos se apresentam como sistemas linguísticos que apresentam similaridades a nível de traços linguísticos, em virtude de as formas captadas pelos aprendentes de L2 envolvidos no contacto dependerem de aspectos compartilhados por estes: idade da aquisição, frequência no *input* disponível, saliência fônica, compreensão, etc. Tal comprovação pode ser observada recorrendo quer à flexão em número quer à configuração estrutural do SN de outros crioulos. Assim, como característica comum aos CP's da Alta Guiné, observa-se que, no SN, os artigos (exemplo [15]), demonstrativos (exemplos [13] e [16]), possessivos (exemplo [12]) e quantificadores (exemplos [14] e [17]) precedem o nome, funcionando como determinantes tanto no CCV como no crioulo da Guiné-Bissau (CGB).

- [12] CCV: *Nos kàza e grànd* (Cardoso, 1989:30)  
 1PL casa é grande  
 “A nossa casa é grande.”
- [13] CCV: *kel minina* (Cardoso, 1989:21)  
 DEM menina  
 “Aquela menina.”
- [14] CCV: El deve ten *algun dñer.* (Cardoso, 1989:35)  
 3SG deve tem Q dinheiro  
 “Ele deve ter algum dinheiro.”
- [15] CGB: *N tené un libru.* (Peck Jr., 1988:33)  
 1SG tem INDF livro  
 “Eu tenho um livro.”

[16] CGB: *es gazya li* (Kihm, 1994:136)  
 DEM rapariga aqui  
 “Esta rapariga aqui.”

[17] CCGB: *Parsi-n kuma no ka ten ja mafe* (Kihm, 1994:237)  
 Parece-1SG COMP 1PL NEG tem Q *mafe*  
 “Parece-me que já não temos mais *mafe*.”

Por seu lado, o artigo indefinido deriva do português “um”, mas é invariável, tanto no CCV (exemplo [18]) como no CGB (exemplo [19]):<sup>25</sup>

[18] CCV: *un mnina* (Cardoso, 1989:21)  
 INDF menina  
 “Uma menina/umas meninas.”

[19] CGB: *um karta* (Kihm, 1994:54)  
 INDF rapaz  
 “Uma carta/umas cartas.”

Por conseguinte, nestes crioulos, a indicação de número do referente é dada pelo contexto, pelo que o determinante possui apenas traço semântico de singular (exemplos [18] e [19]). Contudo, os acrolectos, variedades mais próximas da LA, apresentam já algumas flexões, seguindo regras muito próprias para codificar o género e o número. Deste modo, o número PL só é obrigatoriamente marcado num dos elementos da expressão nominal, em geral o primeiro (o referente). Assim, caso o nome seja antecedido de um determinante artigo (exemplo [7]), um demonstrativo (exemplo [20]) ou um possessivo (exemplo [21b]), é sobre este elemento que vai incidir a pluralização, estabelecendo-se aqui, um paralelismo evidente com a forma de marcar a pluralização não só do PA mas também das outras variedades de português reestruturado de África e do Brasil, que temos vindo a observar neste trabalho:

[20] CCV – VARIEDADE DE SANTIAGO: *Kes mininu fémia* (Pereira, 2006:35)  
 DEM meninofêmea  
 “Aqueles meninas.”

- [21] CCV – VARIEDADE DE S. NICOLAU:
- a. **ña** kàza (Cardoso, 1989:30)  
 1SG casa  
 “A minha casa.”
- b. **ñas** kàza (Cardoso, 1989:30)  
 1PL casa  
 “As minhas casas.”

Se o nome não for antecedido de nenhum determinante PL ou que indique uma quantidade PL, será ele a receber a marca de número (exemplo [22]):

- [22] CCV – VARIEDADE DE SANTIAGO:
- Fidjus** di Maria sta duenti (Pereira, 2006:35)  
 “Os filhos da Maria estão doentes.”

Todavia, a marcação de número também pode ser feita com recurso ao morfema *-s*, depois de palavras acabadas em vogal (exemplo [22]), *-is*, depois de palavras terminadas em consoante (exemplo [23a]), ou *-sis*, depois de palavras finalizadas em vogal nasal (exemplo [23b]):

- [23] CCV – VARIEDADE DE SANTIAGO:
- a. kudjer > kudjeris  
 (Pereira, Arim & Carvalho, 2006:12)  
 “a colher” > “as colheres”
- b. armun > armuns/armunsis  
 (Pereira, Arim & Carvalho, 2006:12)  
 “o irmão” > “os irmãos”

Note-se, ainda, que os falantes das variedades acrolectais usam também a forma singular para construir frases como as do exemplo [24], em que os nomes não carecem de marca morfológica de PL, visto fazerem parte de um conhecimento geral compartilhado pelos falantes, ou seja, actuarem como indicadores de uma dada identidade PL:

- [24] CCV – VARIEDADE DE SANTIAGO:
- Es tene **odju** reagladu na kabésa. (Pereira, 2006:47)  
 Eles tem olho arregalado na cabeça.  
 “Eles estão espantados.”

Formas singulares essas que podem também ser usadas, quando respeito a entidades plurais já referidas anteriormente no discurso, isto é, previamente identificadas como tal:

- [25] CVV – VARIEDADE DE SANTIAGO:  
*Mininu dja txiga.* (Pereira, 2006:93)  
 “As crianças/os meninos/as meninas [já referidos antes] chegaram”

Por fim, importa referir que algumas variedades descrioulizantes do CGB podem apresentar também flexões de número, marcadas no núcleo do SN (exemplo [26]):

- [26] CGB:  
*N tené libru(s).* (Baptista, Mello & Suzuki, 2007:73)  
 1SG tem livro(s)  
 “Eu tenho livros.”

As diferentes formas de marcar o PL nos CP’s atlânticos evidenciam que ocorre um trajecto de inserção de marcas que começa por apresentar paralelismos com o dos substratos, mas que se vai reestruturando, posteriormente, em direcção ao sistema de pluralização da LA. Por outro lado, quer as diferentes maneiras de marcar o número quer a forma como se configura o SN poderão levar-nos a entender que a simplicidade da estrutura destas línguas é, afinal, aparente. Isto mesmo pode ser confirmado por construções praticadas pelos falantes nativos do crioulo de Santiago ou do santomense, que, a exemplo do que sucede com os falantes nativos do português, podem mudar a ordem básica dos constituintes da sua língua (SVO), bastando que recorram, para tal, à alteração de traços prosódicos a fim de apresentar uma frase deslocada à esquerda (exemplos [27] e [28]):

- [27] CCV – VARIEDADE DE SANTIAGO:  
*Es sok, bo traz-êl pra li*  
 (Pereira, Arim & Carvalho, 2006:22)  
 DEM saco 2SG traz-3SG para cá  
 “Esse saco, trá-lo para cá”

- [28] FORRO:  
*Inen poto se, mosu fis’ e.* (Hagemeijer, 2007:19)  
 3PL porta MSPECrapaz fechar 3SG  
 “As portas, os rapazes fecharam-nas.”<sup>26</sup>

Analisando a concordância de número internamente ao SDET do CCV, Castro & Pratas (2006) indicam que a mesma pode ocorrer de acordo com as seguintes possibilidades:

- (i) Se existir um DET no SDET, somente este elemento apresenta o morfema *-s*, marca de PL, uma situação que encontra paralelo evidente no PA:

[29] CCV:  
*Uns / alguns / kes / nhas livru bonito*  
 (Castro & Pratas, 2006:12)  
 um-PL / algum-PL / este-PL / meu-PL livro bonito  
 “uns / alguns / estes / os meus livros bonitos”

[30] PA:  
 a. *Quero vi passia, fica **uns** dia aqui* [OSVALH1]  
 “Quero vir passear, ficar uns dias aqui”  
 b. *fico lá **alguns** tempo* [MMDEUSM3]  
 ?“Fico lá alguns tempos”  
 c. *mataram **esses** home* [MANOH3]  
 “mataram esses homens”  
 d. *fazê **meus** serviço* [SALVH3]  
 “fazer os meus serviços”

- (ii) Caso o DET seja representado por uma categoria vazia, a marcação de PL é inserida no primeiro elemento do SDET, seja ele um nome (exemplo [31a]) seja ele um adjetivo (exemplo [31b]). Estas construções também encontram paralelismos nas estruturas do PA (exemplos [32a] e [32b]):

[31] CCV:  
 a. *livrus bonito* (Castro & Pratas, 2006:12)  
 livro-PL bonito  
 “livros bonitos”  
 b. *purmerus livru* (Castro & Pratas, 2006:12)  
 primeiro-PL livro  
 “primeiros livros”

[32] PA:  
 a. *Eu faço **trabalhos** privado.* [CASTEH1]  
 “Eu faço trabalhos privados”  
 b. *Temo que utilizá arma, e com **boas** barbatana* [OSVALH1]  
 “Temos que utilizar arma, e com boas barbatanas”

(iii) As ocorrências com quantificadores universais apresentam, por seu lado, três possibilidades. Caso exista um DET (demonstrativo ou possessivo), a pluralização é realizada neste (exemplo [33a]); se o quantificador universal estiver localizado depois do nome, a marca de PL é inserida neste (exemplo [33b]); por fim, quando o quantificador universal é colocado inicialmente e sem DET's, não é inserida a marca de PL –s em nenhum dos elementos do SDET (exemplo [33c]). Construções similares são também detectadas no PA, uma vez mais (exemplos [34a], [34b] e [34c]):

- [33] CCV:
- a. *tudu kes / nhas mininu*  
(Castro & Pratas, 2006:13)  
todo este-PL / meu-PL menino  
“todos estes meninos / todos os meus meninos”
  - b. *mininus tudu* (Castro & Pratas, 2006:13)  
menino-PL Q  
“os meninos todos”
  - c. *tudu mininu* (Castro & Pratas, 2006:13)  
Q menino  
“todos os meninos”

- [34] PA:
- a. *fez tudo **essas** coesa* [MANEVEM2]  
“fez todas essas coisas”
  - b. *já cozeu **folhas** toda na panela* [ANTOM1]  
“já cozeu as folhas todas na panela”
  - c. *nós aqui, tudo criança daqui pa brincadera pode falá diolecto*  
[MINISM3]  
“nós aqui, todas as crianças daqui podem falar dialecto nas brincadeiras”

A propósito da construção em que o quantificador universal é colocado inicialmente sem DET's e que leva a que nenhuma marca de PL –s seja inserida nos outros elementos do SDET (exemplo [34c]), veja-se como o informante [CASTEH1], do estrato mais novo e mais escolarizado de Almojarife (8º ano de escolaridade, Tabela 8.1), começa por realizar uma estrutura idêntica à do crioulo, mas tem percepção da variação e tenta corrigi-la, de imediato, de acordo com a regra da concordância do português:

- [35] PA:  
*isso e todo documento... os documentos que eu tenho* [CASTEH1]

Soares & Grosjean (1984:385), abordando as questões inerentes ao acesso ao léxico por parte dos falantes bilíngues, defendem que, que estes têm de aceder a duas redes linguísticas distintas, nas quais se encontram os armazenamentos lexicais de cada uma das suas línguas, ou seja, a forma (fonologia e propriedades sintáticas) e o significado das palavras, mas que se encontram interconectadas no que concerne às representações conceptuais. Assim, estes falantes são capazes de falar numa só língua, mas buscam primeiro as palavras no léxico menos activado, apresentando interferências latentes deste na língua em que estão a comunicar, ou seja, o falante usa a língua com códigos alternados e com recurso a léxico emprestado (Paradis, 1981, 1985, 2001 [1989]).

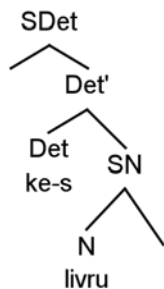
Por seu lado, Castro & Pratas (2006), seguindo as propostas avançadas em termos de MD (Halle & Marantz, 1993, 1994; Noyer, 1997; Embick, 1997; Embick & Noyer, 2001), que abordaremos no ponto 2.6.2.1 do presente trabalho, entendem que não existe Léxico,<sup>27</sup> pelo que as funções que lhe são atribuídas são distribuídas entre outras componentes da gramática.<sup>28</sup> Deste modo, e conforme esquemas das Figs. 16 e 17, a Sintaxe manipula apenas as categorias abstractas dispostas nas estruturas não lineares. Depois da Sintaxe, a Inserção Vocabular (*Vocabulary Insertion*) insere as formas fonológicas – itens de vocabulário – nas estruturas sintáticas abstractas. Por seu lado, a Ordem Linear (*Linear Order*) é uma propriedade unicamente da representação fonológica, enquanto a Linearização (*Linearization*) ocorre concomitantemente com a Inserção Vocabular. Contudo, previamente a esta, a propriedade da Ordem Linear (adjacência e relações precedentes) não pode ser definida numa estrutura sintáctica.

Tal como procederia Baxter (2009) para o PT e o dialecto de HEL-Ba, Castro & Pratas (2006) levaram igualmente em conta as assunções de Costa & Silva (2006b) acerca das diferenças na concordância de número internamente ao SDET no PE (concordância plena) e no PB (número marcado quer nos artigos definidos ou indefinidos e demonstrativos, núcleos DET, quer em todos os elementos pré-nominais). Partindo do pressuposto que o PL é um morfema *singleton* (Costa & Silva, 2006b), Castro & Pratas (2006) assumem que somente o DET carrega a marca de PL, uma vez que é neste elemento que ancora a informação de número (Enç, 1991; Ferrari-Neto, Corrêa & Augusto, 2005; Castro & Ferrari-Neto; 2007), sendo redundante a marca morfofonológica de PL no nome, uma vez que o que conta é a sua expressão em DET

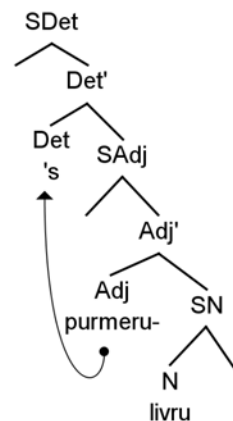


(Castro & Ferrari-Neto, 2007:75). Ora, como vimos nos exemplos que fornecemos, esta é uma situação igualmente válida para o PA. Nesta conformidade, seja no CCV seja no PA, o DET é uma extensão da projecção da expressão nominal (Abney, 1987), sendo aí realizado o morfema de PL, em virtude de assumir a posição de núcleo do SDET. Por seu lado, os itens como *purmeru* (exemplo [31b]), no CVV, e “boa” (exemplo [32b]), no PA, têm que ser observados como complementos pré-nominais (AP’s), e não como adjuntos ou especificadores (Cinque, 1994) que carregam complementos do SN (Abney, 1987; Castro, 2004).

Segundo Costa & Silva (2006b), em línguas do tipo do CCV, o marcador de PL emerge pós-sintacticamente como um morfema do tipo *singleton*, visível no núcleo DET através de um sufixo. Caso haja um elemento visível em DET (Fig. 2), o sufixo é atribuído aos elementos patentes nos exemplos [29] e [33a], para o CCV, e exemplos [30 (todos)] e [34a], para o PA. Se não existir elemento visível em DET (Fig. 3), o sufixo “baixa” e é atribuído ao núcleo complemento (exemplos [31a], [31b] e [33b], para o CCV, e exemplos [32a], [32b] e [34b], para o PA. Internamente ao SDET, a marca PL “baixa” da posição vazia DET e é atribuída ao nome nu.



**Fig. 2.** Marcação plural no CCV  
(Marcador visível em DET).



**Fig. 3.** Marcação plural no CCV  
(Ausência de marcador em DET).

Contudo, este tipo de concordância morfológica reduzida não é válida para o exemplo [33c], do CCV, nem para o exemplo [34c], do PA, uma vez que nenhum dos elementos do SDET recebe marcação. Tal deve-se ao facto de em estruturas do tipo Q+N, o quantificador “*tudu*” deter a função de modificador do SDET e, conseqüentemente, a

sequência ser interpretada como massiva, não necessitando nenhum dos itens de receber marca morfológica de PL (Maling, 1976:716; Vicente, 2006). Contrariamente, nos exemplos [33a] e [33b], do CCV, e exemplos [34a] e [34b], do PA, a marca de PL é inserida porque, em ambos os casos, as sequências são interpretadas como definidas. Assim, na primeira sequência (Q+DET+N), Q é núcleo do SQ, enquanto o SDET é o seu complemento, sendo válida a premissa de que o marcador de PL é atribuído como sufixo do tipo *singleton* ao núcleo DET (Costa & Silva, 2006b). Quanto à segunda situação, verifica-se uma estrutura do tipo N+Q que requer algum tipo de especificação pós-quantificador, apesar de a interpretação ser definida. Neste caso, Q é o núcleo do SQ, um adjunto à esquerda do SX.

Partindo do princípio que a GU actua na aquisição de uma L2 através da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006) e que durante o processo de aquisição o falante vai testando sucessivas possibilidades gramaticais de aproximação à LA, criando ele próprio uma interlíngua que se vai reformulando à medida que adquire novos dados (Corder, 1967; Leiria, 1996), Cardoso (2005) aponta algumas possíveis interferências do crioulo de Santiago na aquisição do PtgL2 falado em Cabo Verde. Tendo em conta que, neste arquipélago, a aquisição do PtgL2 ocorre essencialmente a nível formal, isto é, na escola, configura-se uma situação que propicia a ocorrência de interferência forte ou moderada na interlíngua, muitas vezes responsável por “*ruptura tipológica*” (Mota, 1996:519), já que “*as regras de concordância no crioulo distanciam-se do português como o céu da terra, tanto em número como em género*” Silva (1998:114). Assim, a nível morfossintático, o PtgL2 que tem o CCV como substrato apresenta interferências caracterizadas pela substituição das categorias gramaticais funcionais ou de determinadas regras gramaticais, por outras semelhantes da L1 (Cardoso, 2005:13). No caso da pluralização, estes aspectos são evidentes nas realizações linguísticas do PCV que evidenciam ausência de realização do artigo definido (exemplos 36a. e 36b.), isto é nomes nus, ou de marcação de número (exemplos 37a. e 37b.), apresentando transposições estruturais do CCV, em tudo semelhantes às realizadas no PA, e que abordaremos já no próximo ponto deste trabalho:

[36] PCV:  
 a. *O sobrinho não deu corda* (Cardoso, 2005:14)

↓  
 ∅

CCV:  
 b. *Xibinhu ka da kórda* (Cardoso, 2005:14)  
 sobrinho ASP dar corda  
 “o sobrinho não deu a corda”

[37] PCV:  
 a. *O sobrinho disse lobo para amarar o macaco pé.* (Cardoso, 2005:14)

↓  
 ∅

CCV:  
 b. *Xibinhu fla lobo pa mara makaku pé* (Cardoso, 2005:14)

Sobrinho falar lobo para amarrar macaco pé  
 “O sobrinho disse ao lobo para amarrar os pés do macaco”

Aplicando a mesma metodologia que utilizara para analisar a CPL-var no SN do PMp, Jon-And (2009) observa a influência de cinco variáveis linguísticas – *saliência fônica, posição em relação ao núcleo, posição linear, categoria gramatical e animacidade* – e duas variáveis extralinguísticas – *escolaridade e idade de início de aquisição (age of onset) do português* – para determinar as motivações que subjazem à CPL-var nos SN’s produzidos por falantes do PtgL2 do Mindelo, que têm o crioulo acrolectal de São Vicente/Santo Antão (Barlavento) como L1. Ambas as variáveis sociais foram rejeitadas como não significantes para a marcação, o que levanta a possibilidade de o facto de a LA ser apenas aprendida em contexto formal não ser suficiente para eliminar o fosso tipológico entre o PE e o PCV (Cardoso, 2005:11).

Relativamente às variáveis linguísticas *categoria gramatical e posição em relação ao núcleo* cruzada com a *posição linear*, ambas apresentam um efeito forte para a marcação PL, possivelmente por interferência do substrato (Jon-And, 2009:7), confirmando-se as substituições de categorias ou regras gramaticais da LA por outras semelhantes da L1 (Cardoso, 2005:13). Estes efeitos são semelhantes no PCV e outras variedades africanas e variedades brasileiras de português (Jon-And, 2009:7), fazendo pressupor, uma vez mais, que a instanciação de novos/errados parâmetros e sua frequência no *input*, que apaga a possibilidade de ocorrer reajustamento paramétrico (Gonçalves, 2004:229), leva a que o PCVL1 apresente uma nova reinstanciação paramétrica, com a variação

caminhando no sentido da mudança. Assim, o fenómeno que resulta da convergência entre posição linear e classe gramatical poderá, também aqui, ser associado ao sistema de marcação PL dos substratos africanos do CCV (cf. Baxter, 1998, 2009), com alguns traços das L1's sendo retidos neste e sobrevivendo na gramática da variedade reestruturada (cf. Winford, 2003a:245-247), ao mesmo tempo que algumas propriedades gramaticais da LA se esvanecem no *continuum* de aquisição (Gonçalves, 2004:229). Das similaridades registadas entre as gramáticas do PCV, adquirido por falantes de um CP, e as gramáticas do PVB, PT, PVA e PM, adquiridos por falantes cujos substratos não são CP's, parece resultar evidente que não ocorrem divergências substanciais no que concerne aos padrões de variação (Figueiredo, 2008:37), já que todos estes dialectos têm na sua génese uma situação sociohistórica específica, isto é, de transmissão linguística irregular.

#### **1.2.2.2. Português de Almojarife (São Tomé) e seu substrato (santomense)**

O PA reestruturado, língua falada por uma comunidade semi-isolada de São Tomé, tem, no bojo da sua formação e evolução, uma extensa situação de contacto entre línguas, que se iniciou logo com a colonização do Golfo da Guiné, em 1485. Nesta fase, ocorreram as primeiras e esparsas tentativas de fixação de colonos na ilha, que aí estabeleceram rudimentares engenhos do açúcar (Negreiros, 1895:33), recorrendo à mão-de-obra de escravos resgatados no Benim e no reino do Congo (Brásio, 1952-1:54). Muitos destes escravos seriam já falantes de um proto-crioulo (Hagemeijer, 2009), que ascendeu a língua de comunicação interétnica na ilha. Esta evoluiria para crioulo, uma vez que as comunidades geográfica e culturalmente deslocadas pelos colonos constituíram uma sociedade nova de falantes de distintas L1's do grupo níger-congo atlântico, que abandonaram as mesmas (excepto em situações de ambiente familiar), visto necessitarem de comunicar entre si e com os seus amos.<sup>29</sup>

No início do século XVI, São Tomé era já um importante entreposto no tráfico negreiro (Ferronha, 1989), iniciando-se o período da colonização massiva da ilha (1493 a 1515). Nesta altura, o número de colonos fora já ultrapassado pelo dos escravos, originando uma situação de parco contacto linguístico entre ambos, e que fazia com que os segundos tivessem cada vez menos acesso às formas da LA. Todavia, o número de

mestiços e falantes bilingues exponenciou-se quando D. João II de Portugal, por decreto régio, atribuiu aos colonos o direito de tomarem uma mulher escrava por companheira (Brásio, 1952-1:331-332). Ao conceder liberdade às mulheres africanas tomadas pelos convictos, bem como aos seus descendentes (Brásio, 1952-1:331-332), primeiro (1515), e às concubinas (Brásio, 1952-1:376), depois (1517), o rei D. Manuel I permitiu que a franja da população liberta assimilasse ainda mais a cultura e traços linguísticos dos colonos, constituindo a comunidade de forros (escravos que recebiam carta de alforria). Quanto aos mestiços nascidos destas relações, bilingues em português e crioulo, ganharam projecção na vida política, social e económica da colónia, sendo decisivos para o processo de complexificação e distensão léxico-gramatical do santomense (Ferraz, 1979:17). Após conflitos vários com escravos revoltosos e a tomada do arquipélago pelos holandeses (1641), os colonos abandonaram temporariamente este e o santomense sistematizou-se na forma próxima do actual basilecto (Ferraz, 1979:19). Presentemente, este crioulo é uma das línguas de comunicação generalizada em São Tomé, sendo falado por 72,4% da população (Instituto Nacional de Estatística de São Tomé e Príncipe, 2003/2005). Não obstante deter um léxico em grande parte de origem portuguesa, as suas características morfossintáticas estabelecem paralelismos com as L1's africanas faladas pelos ancestrais escravos transplantados (Hagemeijer, 1999, 2009). O seu substrato linguístico será maioritariamente kwa (falantes oriundos do Benin) e bantu (falantes provenientes da Guiné, Gabão e Angola) (Ferraz, 1975, 1979). Recentes pesquisas têm evidenciado que o proto-crioulo do Golfo da Guiné, que representa a raiz dos quatro crioulos desta região (Hagemeijer, 1999, 2009), evidencia traços que o relacionam quer com o edóide quer com os idiomas bantóides da área H, especialmente com as variedades kongo (Lorenzino, 1998; Hagemeijer, 1999, 2007, 2009; Schang, 2000; Tomás *et alii*, 2002; Güldemann & Hagemeijer, 2006).

No caso concreto da estrutura sintagmática do santomense, uma das formas de marcar o PL recorre ao determinante definido *inen* (exemplo [38b]), homófono do pronome pessoal, 3ª pessoa PL, possivelmente proveniente do pronome do edo,<sup>30</sup> língua do grupo níger-congo atlântico, falada na Nigéria, e colocado em posição pré-nominal para funcionar como marcador de flexão em número, isto é, como elemento que determina a

inserção da marcação PL em nomes de traço [+humano]. Quanto aos restantes itens do SN, mantêm-se invariáveis, sendo iguais para o singular e para o PL:

- [38] FORRO: (Hagemeijer, 2000:118)
- a. *ome plejidu*  
 homem preguiçoso  
 “o homem preguiçoso”
- b. ***inen*** *ome plejidu*  
 3PL homempreguiçoso  
 “os homens preguiçosos”

Estruturalmente, o sistema de marcação PL no SN do santomense estabelece paralelismos quer com línguas do grupo níger-congo (exemplo [39]) quer com construções do PA (exemplo [40]):

- [39] IORUBÁ:  
***àwon*** *omo*<sup>31</sup>  
 3PL criança  
 “as crianças”

- [40] PA:  
*já trabalhê esses lugar tudo* [LUISH2]  
 “já trabalhei nesses lugares todos”

A transposição da estrutura frásica do grupo níger-congo atlântico por falantes nativos pós-adolescentes, em situação de aquisição do PtgL2, determinou, no santomense, o uso de construções sintagmáticas em que a pluralização é inserida apenas no primeiro elemento pré-nominal. Aliás, são várias as similaridades entre o santomense e o seu substrato níger-congo, a nível do sistema morfossintáctico (Ferraz, 1975; 1979). Uma delas manifesta-se na presença de pronomes disjuntivos em frases preposicionadas, tanto nas construções interrogativas como nas relativas (exemplos [41] e [42]):

- [41] FORRO: (Ferraz, 1975:36)
- Nãji su bi de?*  
 Onde você vem de?  
 “De onde vem você?”

- [42] ZULU: (Ferraz, 1975:36)
- Umuzi engavela kuwo*  
 Aldeia eu~vim de  
 “A aldeia de onde eu vim”

No processo de aquisição do português por falantes do santomense L1 ocorreu transferência do sistema estrutural de pluralização deste para a L2. Posteriormente, o mesmo sistema transitaria geracionalmente (cf. Winford, 2003a:245-247) no processo de nativização do PtgL2, com base nos DPL's defectivos transmitidos em cadeia, e que determinaram uma nova reinstanciação paramétrica no PtgL1 (PA). Deste modo, na situação de transmissão linguística irregular originada, os ancestrais falantes do PA marcaram a pluralização com recurso a morfemas do *singleton* na categoria funcional DET, que passou a servir de âncora para o controlo da pluralização. Este tipo de marcação está também presente noutras variedades africanas e brasileiras de português, tenham elas CP's atlânticos ou línguas do grupo níger-congo atlântico como substrato directo ou ancestral:

- [43] PLd:  
*vigia as criança* (Mingas, 2000:67)
- [44] PMp:  
*Há muitas dificuldade nas escola.* (Gonçalves & Stroud, 1998:110)
- [45] PT:  
*issos coesa tudo* (Baxter, 2004:98)  
“essas coisas todas”
- [46] PCV:  
*as coisa foram subindo* (Jon-And, 2009:1)
- [47] MRJ:  
*meus filho casado* (Scherre, 1988:356)

Ainda a propósito da marcação PL no PA, refira-se que detectámos também possíveis resquícios do recurso ao uso do pronome pessoal, 3ª pessoa PL, da LA (exemplo [48]), um fenómeno que abordaremos mais em pormenor no ponto 2.11.3.1.3 do presente trabalho:

- [48] PA:  
*outra pessoa foi tirô eles rôpa, viu arma aqui [ZECAH1]*  
“outra pessoa foi buscar a(s) roupa(s) e viu a arma lá”  
(?) “outra pessoa foi buscar a(s) roupa(s) deles e viu a arma lá”

O tipo de marcação PL estrutural do santomense determina, igualmente, que os nomes e os adjectivos permaneçam invariáveis, baseando-se nas formas masculinas singular do português (terminadas em *-o*), quando esta língua as marca em género (exemplo [49]):

- [49] FORRO:  
*gatu* (Ferraz, 1979:60)  
 “o gato/a gata”

Este fenómeno demarca-se do sistema da LA, visto que o PE implica concordância entre o género do núcleo do SN e os seus determinantes e qualificadores (adjectivos):

- [50] PE:  
*eu tenho um irmão simpático* > *eu tenho uma irmã simpática*

No kimbundu, por exemplo, quase todos os nomes que referem a ideia de sexo são epicenos, salvo raras excepções, como acontece com os termos *tata* (pai) > *mama* (mãe), *diala* (homem) > *muhetu* (mulher) ou *dikolombulu* (galo) > *sanji* (galinha).<sup>32</sup> Verifica-se então que, para as diversas classes de nomes representadas por afixos que fazem a diferenciação entre singular e PL, “as distinções sexuais não são importantes” (Welmers, 1973:159). Deste modo, caso não se realize um dos nomes atrás mencionados, a distinção masculino/feminino é fornecida quer pelos termos *diala* e *muhetu*, que acompanham o nome no caso de ele deter o traço [+humano] (exemplos [51a] e [51b]) ou [+animado] (exemplos [52a] e [52b]). Por vezes, com substantivos do tipo [+animado] e [+humano] referentes a entidades que ainda não atingiram a idade adulta podem ocorrer também os vocábulos *ndumbi* (macho) e *mukaji* (fêmea). Contudo, importa reter que o nome núcleo do SN se mantém invariável:

- [51] KIMBUNDU:  
 a. *mona adiala* (Zau, 2001:24)  
     filho PREP+homem  
     filho de homem  
     “filho”  
 b. *mona amuhetu* (Zau, 2001:24)  
     filho PREP+mulher  
     filho de mulher  
     “filha”



- [52] KIMBUNDU:
- a. *mbua adiala ami* (Zau, 2001:23)  
 cão PREP+homem 1SG  
 cão de homem meu  
 “o meu cão”
- b. *mbua amuhetu ami* (Zau, 2001:23)  
 cão PREP+mulher 1SG  
 cão de mulher meu  
 “a minha cadela”

Como se verifica, a marcação de género restringe-se aos elementos nucleares, já que os outros constituintes do SN se mantêm inalterados. Quanto aos CP’s da Alta Guiné e do Golfo da Guiné, é também a partir do núcleo do SN que se recompõe a morfologia do género. Deste modo, configura-se um claro paralelo no plano da morfossintaxe de género entre estes CP’s e as línguas bantu, com as construções do santomense a revelarem que o valor referencial do sexo ocorre, por norma, com os nomes adjectivados ‘*ome* (homem) e ‘*muala* (mulher), os quais acompanham o nome, mantendo-se este invariável nas formas masculina e singular:<sup>33</sup>. Formam-se, assim, nomes compostos que, no PE, têm forma simples:

- [53] FORRO:
- a. *bué ‘ome ame* (Ferraz, 1979:60)  
 boi homem 1SG  
 “o meu boi”
- b. *bué ‘muala ame* (Ferraz, 1979:60)  
 boi mulher 1SG  
 “a minha vaca”

Lucchesi (2000a), para a variedade de HEL-Ba, e Figueiredo (2009a), para o PA, concluíram que a marcação de género se dá a partir do núcleo nominal em ambos os dialectos, o qual surge como pólo irradiador da marcação mórfica de género. Não obstante, ambas as variedades de português apresentam não-realização de matéria gramatical da LA, com erosão da morfologia flexional do nome afectando mais o mecanismo morfossintáctico da concordância do que a flexão mórfica do género dos núcleos nominais. Desta forma, ao desaparecer o mecanismo morfossintáctico de função meramente gramatical para a indicação do género, conservou-se a flexão deste nos substantivos dos seres animados recorrendo-se ao valor referencial do sexo através de nomes adjectivados do tipo *macho/fêmea* ou *homem/mulher*, a exemplo do que sucede

nos substratos, sejam eles bantu sejam eles CP's, como o santomense ou o CCV (exemplo [54]). Assim, é possível encontrar no PA (exemplos [55a] e [55b]) realizações cujas marcações de género são idênticas não só às do seu substrato (o santomense) e do CCV mas também às de algumas línguas do grupo níger-congo atlântico:

[54] CCV:  
*Mininu fêmia é spértu* (Pereira, 2006:34)  
“As meninas são espertas”

[55] PA:  
a. *mia mãe ñõ pariu homem macho.* [MANEVEM2]  
“A minha mãe não deu à luz/teve rapazes”  
b. *Dois filho mulher qu'ê tem* [FRANCM2]  
“As duas filhas que eu tenho”

O mesmo tipo de variação na marcação de género é ainda detectável tanto no PVA (exemplo [56]), reflectindo transposições literais do substrato kimbundu, como nas falas rurais de Piracicaba, São Paulo, resultantes do contacto da língua portuguesa com dialectos do grupo ameríndio tupi-guarani e africanos do grupo níger-congo atlântico. Em Piracicaba (exemplo [57]), tal como sucede em Almojarife (exemplo [55a]), o uso dos substantivos adjectivados chega mesmo a marcar redundantemente a distinção de sexo nos nomes com traço [+animado]:

[56] PVA:  
*O meu avô de mulher come banana.* (Zau, 2001:24)  
“A minha avó come banana”.

[57] PIRACICABA:  
*Tenho mais fê em menina mulher do que em menino homem.*  
(Rodrigues, 1974:52)

Apesar de este tipo de marcação de género não estar presente no PE, não deve ser desconsiderada a possibilidade de representar uma estratégia natural aplicada em situações de dificuldade comunicativa intercultural, visto que as línguas kwa, também pertencentes ao grupo níger-congo, recorrem, frequentemente, ao género gramatical. Esta é, aliás, uma situação que também acontece no santomense:

- [58] FORRO:  
 a. *soglu* (Ferraz, 1979:60)  
     “o sogro”  
 b. *sogla* (Ferraz, 1979:60)  
     “a sogra”

Os dois tipos de marcação de género existentes no santomense (recurso à marcação com morfema sufixal e às construções que apelam ao valor referencial de sexo) poderão, também, estar na origem da hesitação na aplicação da regra de concordância em género que se verifica no PA:

- [59] PA:  
*faço **trabalhos caseira** até à tarde* [CARMOM1]

A exemplo do que sucede no PA, outras variedades de português, sejam africanas sejam brasileiras, também apresentam CGE-var, um fenómeno já apontado em alguns estudos (p.e. Baxter, 1998):

- [60] PT:  
*agora que tem **muitos capoeira** tá chamare cobra* [Isabel Simba – M2]<sup>34</sup>
- [61] PVA:  
***no mesmo barriga*** (Mendes, 1985:149)
- [62] PMp:  
*Eu tenho **minhas netos*** (Gonçalves & Stroud, 1998:107)
- [63] HEL-Ba:  
*ôpero **no perna e tudo*** (Lucchesi, 2000a:207)

Estes tipos de marcação poderão ter implicação directa no sistema de pluralização de todas estas variedades de português. No caso da marcação do tipo sufixal, e como debateremos no ponto 3.6.7, dedicado à constituição da variável independente *marcação de género*, constituída para constatar se existe uma correlação entre a marcação de género e a marcação PL em cada constituinte do SN, o morfema de género, por ser mais directamente conectado à estrutura mórfica do item em análise, visto estar estreitamente ligado ao radical deste, é aplicado anteriormente ao da marcação de PL (Müller, 1994a; Franseschina, 2002, 2005; Corrêa, Name & Netto, 2004; Montrul, 2004; White *et alii*,

2004;). Deste modo, é de esperar que a marcação de número PL, por ser posterior à do género, possa ser condicionada por esta, quando ela ocorre. Por seu lado, a marcação de género com recurso ao valor referencial de sexo terá determinado que a variável independente *marcação de género* não fosse considerada relevante pela ferramenta computacional VARBRUL para a questão da CPL-var no SN.

Continuando a observar aspectos relacionados com a estrutura do SN do santomense, importa referir que, em trabalho que aborda os nomes nus e o domínio nominal no santomense, Alexandre & Hagemeyer (2007:42) defendem que este idioma apresenta um elevado grau de sensibilidade em relação à identificação dos nomes, em virtude de a relevância destes se limitar ao domínio estrutural do SN. Nos casos da marcação de género e número, a sensibilidade aos princípios semânticos determina que a correcta interpretação das frases requeira não apenas a interpretação do significado lógico-semântico mas também a interpretação do significado informal (Vallduví, 1993:12-13). Nesta conformidade, os traços da animacidade do nome jogam um papel importante na marcação de número e de género, com a inserção de marcas revelando apetência pelos nomes com traço [+humano], visto que estes são interpretados como elementos individuais, enquanto os nomes com traço [-humano] são tidos como mais genéricos (Alexandre & Hagemeyer, 2007:41). Com a marcação de género dependendo da animacidade, os nomes humanos (exemplos [58a] e [58b]) ou extensões do humano, como as profissões (p.e. *kuzinhêru* > *kuzinhêra*), apresentam distinção lexical de género, ao passo que os nomes não humanos necessitam de informação adicional de género (exemplo [53a] e [53b]) (Alexandre & Hagemeyer, 2007:42, nota 7).

Relativamente à estrutura do SN, constata-se, a exemplo do que sucede nos outros CP's da Alta Guiné, a ocorrência do determinante indefinido invariável *ũa*<sup>35</sup> antes do nome núcleo (exemplo [64]) para marcar o número singular, sendo a ausência de flexão atribuída à influência dos substratos africanos (Ferraz, 1975:36):

[64] FORRO:  
*ũa ja* (Alexandre & Hagemeyer, 2007:46)  
“um dia”

Tendo em conta as questões da animacidade em relação à marcação de género e número, o determinante lexical *ũa* transporta especificações de número e definitude

(Alexandre & Hagemeyer, 2007:40), introduzindo informação nova nos contextos sintáticos (Lucchesi, 1993) e tendo como oposto o quantificador *ũa dôsu* (“alguns”):

[65] FORRO:  
*ũa dôsu ome* (Alexandre & Hagemeyer, 2007:40)  
 um dois homem  
 “alguns homens”

Quanto ao sistema de pluralização (exemplo [38b]), Alexandre & Hagemeyer (2007:42) entendem estar-se na presença de um determinante definido marcador de PL, que se terá desenvolvido a partir do âmbito do domínio dos demonstrativos, já que a literatura não é consensual acerca da origem etimológica do mesmo. Considerando-se ainda a relação entre animacidade e marcação de número, verifica-se que a sequência *inen+N* apenas se realiza com nomes de traço [+humano], expressando um significado definido que diz respeito a um grupo não-individualizado familiar ao ouvinte e ao falante. Assim e a exemplo do que sucede com *ũa*, para o singular, o determinante definido lexical *inen* carrega, para o PL, as especificações de número [+plural] e [+definitude] (Alexandre & Hagemeyer, 2007:42). Este aspecto aponta para uma conexão entre o SN definido singular e ausência de artigo definido, estabelecendo semelhanças quer com o CCV (Baptista, 2007) quer com HEL-Ba (Baxter & Lopes, 2009). Contudo, como o santomense, ao contrário dos Crioulos da Alta Guiné, de HEL-Ba e do português, não exhibe marcadores de PL em forma de sufixo do N, isto é, não possui marca de número visível, o seu oposto singular é o SN nu (SNN) (exemplo [66]), que abrange um vasto campo de funções semânticas distribuídas por todas as posições sintáticas, sejam elas predicativas sejam argumentativas (Alexandre & Hagemeyer, 2007:37). Como tal, os determinantes nulos devem receber o mesmo tratamento dos determinantes lexicais, podendo o SNN ser totalmente integrado no mesmo tipo de estrutura nominal que incorpora os determinantes lexicais (ponto 1.3.2.2.2, dedicado à estrutura do SN no forro).

[66] FORRO:  
*Maji vimpema ka<sup>36</sup> fe ome mali* (Alexandre & Hagemeyer, 2007:50)  
 Mas vinho de palma ASP faz homem mal  
 “Mas o vinho de palma é mau para os homens/um homem.”

Outra forma de pluralizar os nomes faz-se com recurso à reduplicação (Ferraz, 1979:58-61), que é também usada para marcar a intensificação:

[67] FORRO:  
*Ne anka se godo-godo* (Ferraz, 1979:58)  
 3PL caranguejo MSPEC gordo gordo  
 “Aqueles caranguejos muito gordos”

Os itens utilizados como pluralizadores na reduplicação podem: (i) intensificar o PL (exemplo [68]); (ii) fornecer o significado de “*cada um dos dois*” (exemplo [69]):

[68] FORRO:  
*E pega ka se kota mō-mō-mō-mō-mō d~e* (Ferraz, 1979:59)  
 3SG pegou caranguejo MSPEC cortou mão-mão-mão de-3SG  
 “Ele pegou naquele caranguejo e cortou-lhe todas as patas”

[69] FORRO:  
*ome-ome o?* (Ferraz, 1979:59)  
 Homem-homem INT  
 “São ambos rapazes?”

Um dos tipos de reduplicação do santomense elimina a terminação do participio passado no primeiro elemento do composto:

[70] FORRO:  
*Ine sa zunta-zuntadu* (Ferraz, 1979:60)  
 3PL está junto-junto  
 “Eles estão todos juntos (misturados)”

Lorenzino (1998:119) considera a estratégia da reduplicação uma fonte produtiva deste tipo de construções nos crioulos. Ao que tudo indica, estar-se-á, uma vez mais, perante uma estratégia natural que visa facilitar situações de comunicação. No PA são também comuns as reduplicações (exemplo [71]), mas, tal como no português (exemplo [72]), representam estratégia de marcação do grau superlativo dos adjectivos. Daí que, na caracterização dos nossos dados (ponto 3.1.4), tenhamos optado por considerar apenas o primeiro item da reduplicação para efeitos de codificação, a fim de não levar à análise um número desproporcionado de instâncias com o mesmo tipo de intenção semântica, que acabariam por viciar os resultados finais (cf. Tagliamonte, 2006:94):

[71] PA:  
*como esses cacau **alto alto*** [CLOTIM2]  
“como esses cacaeiros muito altos”

[72] PE:  
*Este vestido é **lindo, lindo**.*

No PA, há ainda outras estruturas que evidenciam paralelismo com as do seu substrato crioulo, e que têm influência na CPL-var do SN. No caso do uso do determinante artigo definido, por exemplo, ocorre uma variação do tipo ternário, já que o mesmo tanto pode surgir sob a forma do determinante do PE (*o, a, os, as*, exemplo [75]) como ser omitido (nomes nus, exemplo [73] e [74]), ou ainda aparecer representado pelo pronome pessoal terceira pessoa do PL *eles*, (exemplo [48]):

[73] PA:  
*faço trabalhos caseira até tarde* [CARMOM1]  
“faço os trabalhos caseiros até tarde”

[74] PA:  
*E amigo meu tinha uma mulher* [OSVALH1]  
“E o meu amigo tinha uma mulher”

[75] PA:  
*É os meus pai que me deu essa ideia.* [OSVALH1]  
“Foram os meus pais que me deram essa ideia”

Estas questões podem ter influência directa na codificação dos nossos dados, já que estruturas simplificadas do tipo DET+N, do PE, podem surgir no PA sem o determinante, implicando que o nome surja isolado na cadeia sintagmática. Deste modo, dados que seriam codificados e levados à análise, foram retirados destas porque configuram realizações sintagmáticas com um único item gramatical.

As realizações com nomes nus do PA (exemplos [73] e [74]) têm respaldo no seu substrato, o santomense. Neste, o uso deste tipo de nomes apresenta uma distribuição e interpretação livre a nível gramatical, mas que é restringida pela componente extra-gramatical. Assim, os nomes nus podem ser licenciados somente na *interface* entre o nível gramatical e extra-gramatical, que relaciona o discurso à pragmática (Alexandre & Hagemeyer, 2007:49). A interpretação de definitude atribuída aos nomes nus deve-se ao

facto de os mesmos terem sido previamente licenciados no discurso, sendo já referenciados pelo ouvinte. Se, pelo contrário, os nomes nus constituem informação nova para o ouvinte, então são entendidos como plurais indefinidos. Como as línguas que têm nomes nus para realizarem funções específicas não apresentam determinantes visíveis para preencherem essas mesmas funções (Crisma, 1997:236), o pressuposto adequa-se ao santomense no que concerne ao indefinido singular *ũa*, uma vez que este dialecto não exhibe SNN's com leitura singular indefinida. Por outro lado, o mesmo pressuposto é também válido para a questão do definido PL *inen*, já que este apenas selecciona nomes humanos, que são preferencialmente interpretados como singulares.

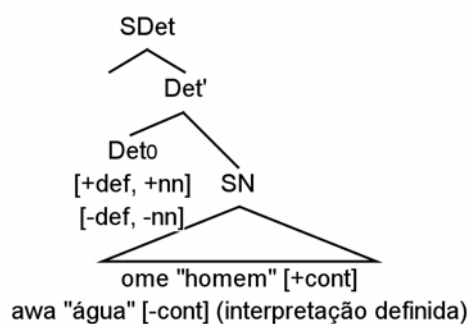
Dado que o sistema de determinantes do santomense usa recursos mínimos para duas funções específicas (*ũa/inen*), as restantes são preenchidas pelos determinantes nulos. Assim, os definidos nulos contemplam, normalmente, o singular (a interpretação preferencial dos SNN's sujeito, que encerram informação previamente licenciada). Quanto aos indefinidos nulos, por norma acompanham o PL (SNN's lexicalmente governados e que contêm informação nova). Contudo, como a informação de número ancora na definitude, a *interface* entre os níveis gramatical e extra-gramatical é crucial para interpretar correctamente as funções dos determinantes nulos. Portanto, os SNN's em posições lexicalmente governadas (V, S) podem ser entendidos como singulares/plurais definidos e plurais indefinidos, mas nunca como singulares indefinidos, que têm de ser precedidos por *ũa* em qualquer posição sintáctica. Por seu lado, os SNN's sujeito não carregam informação nova, pelo que não constituem um passo novo na troca entre definitude e indefinitude, contendo apenas traços do singular, tal como sucede no PA (exemplo [74]), em que o informante já havia feito uso de um ancorador, a fim de licenciar a informação (*ela vivia perto de mô amigo Adento* [OSVALH1] > “ela vivia perto do meu amigo Adérito”). A Tabela 1.1 fornece o panorama sobre a informação contida na estrutura dos SNN's do santomense:



**Tabela 1.1.** Informação encerrada nos SNN's do santomense (Alexandre & Hagemeijer, 2007:56).

Informação nova (indefinido)		Informação licenciada (definido)	
Sujeito	Objecto	Ancorador	Ancorado
SGL	<i>ũa ome/ũa kabla</i> “um homem /uma cabra”	<i>Ome se/Kabla se</i> <i>Ome/Kabla</i>	<i>Ome/Kabla</i> <i>Ome/Kabla</i>
PL	-	<i>Inen ome se/Inen kabla se</i> <i>Inen ome/Kabla</i>	<i>Ome/Kabla</i> <i>Ome/Kabla</i>

Dado que os SNN's exibem um leque variado de leituras que podem ocorrer em posições argumentativas ou predicativas, deverão ser integrados na estrutura que requer um nível de interpretação extra-agramatical proposta na Fig. 5. Esta integração é motivada pelo suporte empírico de que este tipo de SN's projecta sempre um SDET, nível onde ocorre a *interface* entre gramática e discurso pragmático, fundamental para a ancoragem da interpretação (Fig. 4). Note-se que, no santomense, os nomes são sempre interpretados como massivos, requerendo um sistema classificador para a derivação das interpretações contáveis (Alexandre & Hagemeijer, 2007:51). Esta particularidade estabelece parecenças tipológicas com a língua chinesa, que Chierchia (1998) classifica como [-predicativa, +argumentativa].



**Fig. 4.** Estrutura sintático-discursiva dos sintagmas nominais nus no santomense (Alexandre & Hagemeijer, 2007:51).

Já no que concerne ao dialecto de HEL-Ba, as similaridades estendem-se muito para além da mera interpretação dos massivos e contáveis, abarcando questões como o desfavorecimento do uso do artigo quer com nomes de referência única ou específica quer com SN's de número singular, ou, em alternativa, o favorecimento da aplicação do artigo definido em SN's nos quais há co-ocorrência deste e modificadores pós-nominais, como o adjectivo, o possessivo ou a oração relativa (Baxter & Lopes, 2009).

Paralelamente, Baxter & Lopes (2009:328-329) notam ainda que o uso do artigo definido pleno em HEL-Ba só é consistentemente favorecido em SN's com função sintáctica de adjuntos adnominais, ao passo que é desfavorecido caso os SN's desempenhem qualquer outra função sintáctica. Estes resultados estabelecem um paralelismo evidente com idênticos achados dos autores para a fala de Salvador (Baxter & Lopes, 2004b) e, no que concerne especificamente aos SN's com funções de sujeito e OD, ocorrem também similaridades entre HEL-Ba e o santomense, já que este não apresenta assimetria na distribuição dos SN's sem determinante de referência definida que figuram como argumentos internos do verbo (Alexandre & Hagemeijer, 2007:57).

As realizações que configuram o uso ternário do determinante no PA têm respaldo, todas elas, em formas do santomense, pelo que importa fazer uma abordagem à estrutura do SN, neste crioulo. Observando a colocação do adjectivo cadeia sintagmática (exemplos [38a] e [38b]), verifica-se que o mesmo só pode ocorrer em posição pós-nominal, tal como sucede nas línguas do ramo bantu (exemplos [76a] e [76b]). No português, o adjectivo não detém esta rigidez formal (exemplos [77a] e [77b]):

- [76] KIMBUNDU:
- |    |  |  |   |
|----|--|--|---|
| a. | <i>diala (di~ala)</i><br>AFXSG+homem<br>“o homem branco”             | <i>dia (di~a)</i><br>AFXSG+PREP(de)        | <i>mundele (mu~nde)</i> <sup>37</sup><br>AFXSG+branco |
| b. | <i>mala (ma~ala &gt; mala)</i><br>AFXPL+homem<br>“os homens brancos” | <i>ma (ma~a &gt; ma)</i><br>AFXPL+PREP(de) | <i>mindele (mi~nde)</i><br>AFXPL+branco               |
- [77] PTG:
- |    |                    |
|----|--------------------|
| a. | <i>casa bonita</i> |
| b. | <i>bonita casa</i> |

A concordância entre os elementos que constituem o SN é controlada nas línguas bantu pelo afixo pré-nominal e concordantes aglutinados em posição prefixal aos restantes itens. A similaridade com o sistema de marcação PL do santomense e do PA, nas situações em que há recurso ao item *inen*, no primeiro, e de marcação PL no elemento pré-nominal no PA, com recurso a morfema do tipo *singleton*, determina que o radical das palavras se mantenha inalterado nos restantes elementos que compõem a cadeia sintagmática. No PA a esmagadora maioria dos adjectivos surge também preferencialmente na posição pós-nominal, conforme se verá no ponto 4.1.2.1.1 do

presente trabalho. Em SN's de estrutura reduzida (2 ou 3 elementos), que configuram a grande maioria das produções sintagmáticas dos almozarifanos, estabelece-se, então, um claro paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjetivos e terceira posição (cf. Guy, 1981b).

A não-marcação de PL sufixal nos elementos do SN do santomense implica que determinados itens gramaticais, como os possessivos, também se mantenham inalterados. Esta categoria gramatical ocorre também em posição pós-nominal (exemplos [53a] e [53b]), um aspecto que encontra paralelo nos substratos níger-congo atlânticos (exemplo [78]) e noutros CP's atlânticos (exemplo [79]):

[78] IORUBÁ:  
*omo mi*<sup>38</sup>  
 filho 1SG  
 “o meu filho”

[79] CCV:  
*na ja bráb di mi* (Holm, 1988/1989:274)  
 na ilha Brava de 1SG  
 “na minha ilha Brava”

A construção é também detectada no PA (exemplo [74]) e outras variedades africanas de português (exemplo [80]), originando frases agramaticais no que concerne à posição linear do possessivo:

[80] PVA:  
*eu tenho horário meu* (Inverno, 2005:140)  
 “eu tenho o meu horário”

Idêntica situação acontece com a pluralização e colocação dos determinantes demonstrativos, que apenas possui uma forma no santomense (*se*) e se situa sempre em posição pós-nominal (exemplos [67] e [68]). Lucchesi (1993) atribui a *se* a função de definido, servindo para enfatizar o nome. Como tal, não será um elemento nuclear no santomense. Todavia, esta não é a visão de Alexandre & Hagemeyer (2007:44), que entendem que *se* não detém estas características no santomense. De facto, não só o traço definido é preenchido pelos SNN's como *se* pode ainda co-ocorrer com os marcadores definido *ũa* e indefinido *inen* (Alexandre & Hagemeyer, 2007:44), especialmente com o segundo (exemplo [81]). Assim sendo, os autores propõem que o elemento gramatical *se*

seja observado como um especificador de marcação (MSPEC), que é núcleo do SSPEC (Fig. 5).

- [81] FORRO:  
*So inen ja se so sun Fenanji na ska*  
Então 3PL *dia* MSPECFOC senhor Fernandes NEG ASP  
*texila vin fa.* (Alexandre & Hagemeijer, 2007:41)  
Tira vinho NEG  
“Então, estes dias o senhor Fernandes não está a tirar vinho de palma.”

Uma das características estruturais rígidas de *se* é a de apenas poder ocorrer de modo adjectival, isto é, requerer obrigatoriamente a presença do núcleo nominal (exemplo [82]), o que não sucede com o sistema do português, reforçando o facto de estarmos na presença de um MSPEC.

- [82] FORRO:  
*Ka se sa bwa* (Ferraz, 1979:74)  
Coisa MSPECestá boa  
“isto/aquilo é bom”

A relação entre núcleo e MSPEC leva Alexandre & Hagemeijer (2007:45) a considerarem *se* um clítico com forte adjacência nominal visível, enquanto o marcador de PL *inen* e o quantificador *ũa* provêm do léxico com um traço de número inerente, comportando-se nominalmente e permitindo lacunas (nomes nus). De facto, a definitude, enquanto instanciação formal que permite a identificação, conecta a frase ao nível do discurso e pode ser desenvolvida antes do número (Wood, 2003:127), enquanto este transita directamente a partir do seu próprio traço, quando os determinantes são visíveis. A ligação entre definitude e número no santomense activa a leitura do número singular, sendo necessário recorrer ao marcador de PL *inen* para transformar esse singular em PL. Por sua vez, o MSPEC tem a função de reforçar o traço de definitude.

Por outro lado, o alto grau de referencialidade que resulta da co-ocorrência entre o marcador de definitude e o MSPEC é bastante comum com nomes de traço [-humano] (exemplo [81]), estabelecendo o contraste com as ocorrências em que apenas intervêm o marcador de definitude e o nome, aplicadas aos nomes de traço [+humano] (exemplo [38b]). A análise das questões inerentes ao SDET incidindo na sintaxe e semântica dos modificadores que expressam definitude, número, especificidade e interacção entre

todos terá, então, que pressupor que o traço de número está directamente associado à definitude, embora esta possa cair fora do nível gramatical quando se trata de SNN's. Considerando que o traço de número pode ser mapeado a nível do DET0 (Longobardi, 1994:621) e que não ocorre marcação PL por intermédio de afixos no santomense, Alexandre & Hagemeijer (2007:43) entendem que não é requerida uma categoria funcional separada para os SNN's, propondo uma estrutura sintáctica que dispensa uma projecção independente de número e integra ambos os DET's, lexicalizados e ausentes, num único sistema. Desta forma, para os exemplos [83a] e [83b], ter-se-á a estrutura do SDET representada na Fig. 5, a qual, em virtude da sua especificidade, difere das propostas por Castro & Pratas (2006) para os dois tipos de SDET existentes no CCV (Figs. 2 e 3):

- [83] FORRO:
- a. *ũa mina se* (Alexandre & Hagemeijer, 2007:47)  
 IND menina MSPEC  
 “uma menina em questão”
- b. *inen mina se* (Alexandre & Hagemeijer, 2007:47)  
 3PL menina MSPEC  
 “as meninas em questão”

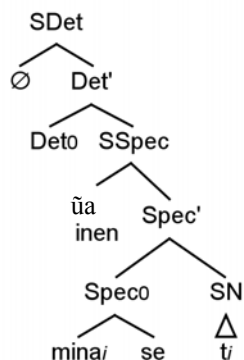


Fig. 5. Estrutura sintáctica do SDET no santomense (marcação de número)  
 (Alexandre & Hagemeijer, 2007:47).

De acordo com a configuração da Fig. 5, o MSPEC *se* é núcleo do SSPEC, um nó que apenas projecta quando ocorre evidência lexical. O nome é adjunto à esquerda do clítico *se*, que é invariável e detém leitura de singular associada a si. Bernstein (1993:217) entende que os demonstrativos e os determinantes definidos são núcleo do SDET, daí excluïrem-se mutuamente (p.e. *\*o este carro / \*este o carro*). Este pressuposto torna

feliz o entendimento de que *ũa* e *inen* são núcleos do SDET no santomense. Enquanto determinantes, ambos C-comandam o seu complemento, SN e SSPEC, se projectados. Como mais nenhum movimento do nome é necessário, a árvore sintáctica proposta configura a generalidade da estrutura dos SDET's no santomense.

Relativamente ao uso do possessivo, obedece ao mesmo tipo de colocação verificada para o MSPEC (pós-nominal), recorrendo os falantes ao pronome pessoal objecto para realizarem tal categoria (exemplo [84]), uma situação que, ao que tudo indica, se aplica ao núcleo do SDET *inen*, homónimo do pronome pessoal, 3ª pessoa PL:

[84] FORRO:  
*Livlu mu* (Ferraz, 1979:81)  
 livro PP/POS  
 “o meu livro”

Outra questão que poderá ter interferência na observação da CPL-var do SN do PA relaciona-se com o uso das partículas locativas que marcam obliquamente o nome e, por norma, não marcam presença no santomense. Tal aspecto reflecte o reduzido sistema preposicional dos crioulos:

[85] FORRO:  
*Amaña nga ba vende* (Ferraz, 1979:74)  
 Amanhã 1SG-FUT ir loja  
 “Amanhã eu vou (irei) à loja.”

Se atentarmos às línguas do grupo níger-congo atlântico, verificamos que o seu sistema preposicional é mais reduzido que o do português. No kimbundu, por exemplo, verifica-se que os complementos circunstanciais de lugar são introduzidos por apenas três preposições:

- (i) *Ku*, locativo referente a direcção, lugar de destino distante, interioridade e proveniência. Estabelece correspondência com as preposições *a*, *em*, *para* e *de*.
- (ii) *Mu*, com o sentido de interioridade. Corresponde à preposição *em* e às locuções prepositivas *dentro de*, *no interior de*, etc.
- (iii) *Bu*, que indica sobreposição, à superfície de. Condiz com a preposição *sobre* e as locuções prepositivas *em cima de*, *acima de*, etc.

As preposições podem ser pré-afixadas aos nomes ou pronomes, formando então SN's com valor espacial:

- [86] KIMBUNDU:  
*mwene wamuya kubata* (Mingas, 2000:75)  
 “mwene u~amuya **ku**~(di)bata”  
 3SG 3SG+PROG+ir LOC+casa  
 “ele vai para casa”
- [87] KIMBUNDU:  
*mwene wala monzo* (Mingas, 2000:76)  
 “mwene u~ala **mu**~inzu”  
 3SG 3SG+IMP+estar LOC+quarto  
 “ele está no quarto”
- [88] KIMBUNDU:  
*Mwene wambata anyi bu mutwe?* (Mingas, 2000:76)  
 “mwene u~ambata anyi **bu** mutwe?”  
 3SG 3SG+IMP+levar o que é LOC cabeça  
 “o que é que ele leva na cabeça?”

O PE, por seu lado, possui um sistema de preposições mais extenso para expressar as funções espaciais (*a, ante, após, até, contra, de, desde, em, entre, para, perante, por, sob, sobre, trás*), podendo algumas destas preposições surgir contraídas com os definidos, os indefinidos (artigos e pronomes), os pronomes pessoais e os pronomes demonstrativos. Deste modo, os falantes de L1's bantu em situação de aquisição do PtgL2 não conseguem estabelecer as funções fornecidas pelas diversas preposições do PE e acabam por produzir ocorrências que se revelam agramaticais neste (Mingas, 2000:76). A este propósito, refira-se que a preposição “em”, correspondente à forma *ku*, é a mais produtiva e, conseqüentemente, aquela que mais desvios revela:

- [89] PANG:  
*Pessoas que estavam lá na trás...* (Cardoso, 1980:56)  
 “As pessoas que estavam lá atrás”
- [90] PANG:  
*Passavam e olhavam nele...* Cardoso, 1980:35)  
 “Passavam e olhavam para ele...”
- [91] PANG:  
*farristas vieram na procura dos bens deixados* (Cardoso, 1980:55)  
 “os farristas vieram à procura dos bens deixados”

[92] PANG:  
*Pra quê queres saber? É **na** tua conta?* (Cardoso, 1980:58)  
“Para que é que queres saber? É da tua conta?”

[93] PANG:  
*se queres gozar, vai gozar **na** tua mãe.* (Cardoso, 1980:60)  
“se quiseres gozar, vai gozar com a tua mãe.”

No PA, a não identificação das funções das preposições do PE é também uma evidência, a ponto de elas poderem ser adoptadas para exprimir situações que se demarcam da referência espacial (exemplos [95], [96] e [97]) ou serem aplicadas de forma redundante (exemplo [97]):

[94] PA:  
*daqui é fui **para até** hospital* [OSVALH1]  
“daqui, eu fui ao (para o) hospital”

[95] PA:  
*o quê que ocê faz de manhã até **na** hora que ocê vai dormir **de** noite?*  
[OSVALH1]  
“o que é que você faz desde a manhã até à hora em que você vai dormir à noite”

[96] PA:  
*o irmão também que stava **na** xofer na Praia Conchas* [CELESH3]  
“o irmão que também era chofer/motorista na Praia das Conchas”

[97] PA:  
*agora quase toda criança naci já **com a** falá português* [CELESH3]  
“agora quase todas as crianças nascem já a falar português”

A indeterminação das funções e uso das preposições do PE é também comum nas outras variedades africanas de português:

[98] PM:  
*Os alunos, aqueles que ser mandado **na**queelas escolas profissionais.*  
(Gonçalves & Stroud, 1998:127)  
“Os alunos, aqueles que são mandados para aquelas escolas profissionais.”

[99] PM:  
*... para os portugueses ir também **no** país deles.*  
(Gonçalves & Stroud, 1998:127)  
“... para os portugueses irem também para o país deles.”



- [100] PM:  
*Tem que viver com comida que ser oferecido **com** outros países.*  
 (Gonçalves & Stroud, 1998:127)  
 “Têm que viver com a comida que é oferecida pelos outros países.”
- [101] PM:  
*Eu é que é mais velho em relação **dos** outros.*  
 (Gonçalves & Stroud, 1998:127)  
 “Eu é que sou o mais velho em relação aos outros.”
- [102] PT:  
*põe dentro de garrafa, vai jogar **no** mar* [Amélia do Nascimento – M1]  
 “põe-se dentro da garrafa e vai jogar-se (atirar-se) ao mar”
- [103] PT:  
*Pa entregá **nas** crianças* [Joana Nunes – M1]  
 “Para entregar às crianças”
- [104] PT:  
*porque avó falô com ele só **com** kimbundu* [Amélia do Nascimento – M1]  
 “porque os avós falaram com ele só em kimbundu”

Abordando a concordância variável de género no SN de HEL-Ba, Lucchesi (2000a:98) chama a atenção para as estruturas desviantes deste, nomeadamente no que concerne à “*redução/eliminação do emprego de nexos gramaticais*” que levam este dialecto a apresentar não só “*o uso de em por a e de de por para/a*” (exemplos [105] e [106]) mas também ausência de preposições em determinadas realizações que requerem o seu uso (exemplos [107] e [108]):

- [105] HEL-BA:  
*fui lá **em** procura de um remédio* (Lucchesi, 2000a:98)  
 “fui lá à procura de um remédio”
- [106] HEL-BA:  
*fica **de** sempre* (Lucchesi, 2000a:98)  
 “fica para sempre”
- [107] HEL-BA:  
*porque eu nunca gostei **nenhuma*** (Lucchesi, 2000a:98)  
 “porque eu nunca gostei de nenhuma”
- [108] HEL-BA:  
*mas nós nunca liguemo festa não* (Lucchesi, 2000a:98)  
 “mas nós nunca ligámos à festa, não”

Estas estruturas apresentam “*um paralelo com crioulos de base portuguesa da África (a exemplo do crioulo de S. Tomé)*” (Lucchesi, 2000a:97). Efectivamente, o santomense caracteriza-se por possuir um reduzido sistema de preposições, ocorrendo frequentemente estruturas que apresentam ausência das mesmas (exemplo [85]). Sendo a redução no uso das preposições um dos fenómenos de não-realização gramatical que caracteriza também os processos de mudança resultantes da transmissão linguística irregular, a ausência destes itens gramaticais na estrutura das frases do santomense (exemplo [85]) e do PVB (exemplos [107] e [108]) estabelece também paralelismos não só com o PA mas também com outras variedades africanas de português:

[109] PA:  
*É, bêra strada.* [CLOTIM2]  
 “É, à beira da estrada.”

[110] PANG:  
*Sim, trabalhar aí sês casa* (Inverno, 2005:140)  
 “Sim, trabalho aí na vossa casa”

[111] PM:  
*Pega cadeira!* (Gonçalves & Stroud, 1998:41)  
 “Pega na/numa cadeira”

[112] PT:  
*a missora começô transmitir* [Pedro Fuíca – H3]  
 “a emissora começou a transmitir”

Ainda no que concerne ao emprego das preposições, Lucchesi (2000a:97) menciona igualmente o uso da forma “ni” por “em” (exemplos [113] e [114]), que se revela quase categórico na fala de Mato Grosso, Bahia. Segundo o autor, dado tratar-se de um item vocabular isolado, “*torna-se ainda mais aceitável a hipótese de um empréstimo lexical, com origem em outros dialectos rurais brasileiros que desenvolveram a forma a partir do contacto entre línguas*” (Lucchesi, 2000a:97-98).

[113] HEL-BA:  
*nunca trabalhei **ni** roça* (Lucchesi, 2000a:97)

[114] HEL-BA:  
*eu já fui **ni** Rio de Contas* (Lucchesi, 2000a:97)

O mesmo locativo regista-se também no santomense, que pode, opcionalmente, recorrer a três tipos de prefixos nominais, com função locativa:

(i) O morfema *ni*, ou sua variante *n-*, antes de vogais:

- [115] FORRO:  
 a. ***ni*** *ke* (Ferraz, 1979:74)  
     em casa  
     “em/na casa”  
 b. ***n-*** *ome* (Ferraz, 1979:74)  
     em meio  
     “no meio”

(ii) A consoante hormogânica nasal *m*<sup>39</sup>:

- [116] FORRO:  
***m-*** *panela* (Ferraz, 1979:75)  
 em panela  
 “na panela”

(iii) A vogal morfofonémica *î*<sup>40</sup>:

- [117] FORRO:  
***î-*** *liba* (Ferraz, 1979:75)  
 em cima  
 “em cima”

Idêntico tipo de realização ocorre em línguas do grupo níger-congo atlântico, como o iorubá (exemplo [118]) ou o kimbundu (exemplo [119]), substituindo o advérbio de lugar que marca obliquamente o nome:

- [118] IORUBÁ:  
***Ni*** *ile* > ***nile*** (Ferraz, 1979:75)  
 PREP casa  
 “em/na casa”

- [119] KIMBUNDU:  
*Ngala* ***ninzala*** (Zau, 2001:28)  
 1SG+estar PREP+fome  
 “eu estou com fome/eu tenho fome”

O mesmo item gramatical marca também presença tanto no PA, que tem um CP atlântico como substrato, como no PT, que tem línguas do grupo níger-congo como substratos:

- [120] PA:  
*porque eu stive lá ni empresa trabalhar* [MANOH3]  
 “porque eu estive lá na empresa a trabalhar”
- [121] PA:  
*com chicote ni costas pa vi pa o pé trabalho* [MANOH3]  
 “com chicote nas costas para vir a pé para o trabalho”
- [122] PT:  
*quando levantei encontrei arroz ni dezoito dobra.*<sup>41</sup>  
 [Amélia do Nascimento – M1]  
 “quando me levantei encontrei o arroz a dezoito dobras”
- [123] PT:  
*eu tenho um tio lá ni sede também* [Amélia do Nascimento – M1]  
 “eu também tenho um tio lá na sede”

Baxter (2004:109) refere que o locativo invariável “ni” substitui, no PT, componentes flexionados do português como <preposição+artigo+número>, representados pelas contracções prepositivas espaciais “nos/nas”, com sentido locativo ou direccional. A forma “ni” ocorre com SN’s sem artigo definido do português, pelo que a dupla função locativo-direccional atribuída pelos falantes do PT a este locativo se ficará a dever a fenómenos de transferência resultantes de, nas línguas do substrato africano, as preposições compartilharem as referidas funções locativo-direccionais. Entretanto, não temos conhecimento da existência deste tipo de realização nem no PANG nem no PM, podendo a mesma ter origem no santomense, por transferência, daí transitando geracionalmente, via DLP’s, para o português de São Tomé e suas variantes, como o PA. Posteriormente, devido ao contacto linguístico entre os falantes destes e os utilizadores do PT, a mesma ter-se-á generalizado na ilha, comprovando-se “*a enorme complexidade da questão do contacto entre línguas em geral*” (Lucchesi, 2000a:98). Baxter (2004:100) advoga também que os valores de “ni” no PT, embora sendo entendidos como conectados ao crioulo de S. Tomé, não terão sido adquiridos pelos tongas via santomense nos estádios primordiais da actividade laboral nas roças, mas sim posteriormente, via português de São Tomé (e PA). Apesar de a construção locativa “ni” não ocorrer no PANG e no PM, note-se a sua aproximação fonológica às forma “no/na”, bastante generalizadas nestas variedades de português (especialmente a forma “na”). Ainda segundo Baxter (2004:123), tanto a dupla função locativo-direccional da

preposição portuguesa “em” como o facto de a consoante /d/ inicial ocorrer como /nd/ em línguas bantu, como o umbundu, terão igualmente concorrido para que a preposição “de” [di] passasse a ser pronunciada como /ni/.

Em HEL-Ba há evidências de que o locativo actua como demonstrativo, que pode funcionar como artigo (Baxter, Lucchesi & Guimarães, 1997:13-14). Segundo Cerqueira (1993), os locativos funcionam como demonstrativos na fase inicial da aquisição do PBL1, pelo que os dados do PA poderão reflectir uma fase transitória em que o artigo se desenvolveu “*em estruturas que previamente dependiam do locativo pós-nominal como marcador das propriedades referenciais do substantivo. Nessa fase transicional, esse papel pretérito do locativo começa a recuar enquanto se desenvolve a estrutura pré-nominal*” (Baxter & Lopes, 2009:326-327). Estes aspectos têm repercussão directa na estrutura do SN do PA, já que as formas “ni” e “di” impedem que se incluam, nos dados a serem quantificados e codificados na presente análise, SN’s que no PE configuram a cadeia sintagmática PREP+DET+N, mas que na fala de Almojarife interpretámos com a sequência PREP+N, isto é, SNN’s. Como tal, este tipo de SN’s não foi codificado e quantificado por nós, em virtude de representar estruturas em que acontece o apagamento da categoria DET (exemplos [120] e [121]). Opostamente, o uso desviado de algumas preposições do PE, com especial incidência para a forma contraída “na”, não só pode determinar concordância variável de género entre o núcleo do SN e o DET como originar também SN’s simples no PA (PREP+DET+N), codificados e quantificados na presente análise. Note-se que este tipo de SN’s poderá ser inexistente no PE, caso configure estruturas do tipo PREP+N ou com nome simples (exemplo [96]).

Resta referir que, embora o aparecimento da forma “ni” no dialecto de Mato Grosso possa também ser equacionado como um empréstimo lexical proveniente de outras variedades rurais brasileiras, a verdade é que não pode ser descartada a possibilidade de o seu surgimento nestas ter a ver com uma transplantação da responsabilidade dos escravos provenientes de S. Tomé, após terem aqui permanecido durante algum tempo, prestando trabalho nos engenhos de açúcar (Holm, 1988/1989, 1992; Baxter, 1992, 1997, 2003).

### **1.2.2.3. A comunidade de Almojarife**

Após a reconquista de São Tomé aos holandeses, em 1644, os portugueses depararam-se com uma realidade diferente da anterior, reflectida na “africanização” do arquipélago. A ilha encontrava-se praticamente dividida em duas partes (Henriques, 1989:276), criando-se uma situação de parco contacto linguístico entre os membros estabelecidos nas regiões interiores (escravos africanos e seus descendentes, a maioria falante do santomense e alguns bilingues em forro e PtgL2 ou PtgL1 nativizado) e litorais (colonizadores, falantes do PE, e seus descendentes libertos, bilingues em santomense e PtgL1). Ganhando autonomia ao refugiarem-se no mato, os africanos deixavam de estar inseridos no espaço linguístico-cultural dos colonos. Paralelamente, a conjuntura económica não favorecia a exploração açucareira (Henriques, 1989:278), evitando nova instalação massiva de colonos no arquipélago e acentuando a situação de afastamento linguístico entre estes e a população sediada nas zonas rurais.

Face a este panorama, a melhoria económica só seria retomada no século XIX, com a introdução das culturas do café (1800) e do cacau (1822) nas ilhas. No entanto, vivendo em situação de auto-suficiência no interior das ilhas, a maioria dos libertos ou descendentes de africanos sedeados em São Tomé não se mostrava predisposta a trabalhar para os colonos, fazendo com que os proprietários das roças tivessem de recorrer à mão-de-obra “contratada”, provinda inicialmente de Angola e Moçambique (finais do século XIX), mas também de Cabo Verde, em período posterior (meados do século XX). Apesar de a abolição da escravatura ter acontecido no império português a 25 de Fevereiro de 1869, os trabalhadores das roças eram forçados a trabalhar em condições desumanas. Esta situação levou a que a maioria dos libertos continuasse a não procurar emprego nas roças (Ribeiro, 1875:12), pelo que, nestas, apenas permaneciam os “contratados”, constituindo comunidades cada vez mais isoladas. Por seu lado, os libertos, tal como sucedeu no Brasil, buscavam outras alternativas de sobrevivência, estabelecendo-se a sua grande maioria na periferia dos centros urbanos. Outros, porém, permaneceram em pequenas comunidades rurais semi-isoladas, compostas por núcleos familiares, como terá acontecido com os forros que se instalaram em Almojarife.

Ainda que a falta de documentação sobre esta pequena comunidade não permita determinar com exactidão a sua origem, o facto de os seus falantes serem bilingues em

santomense e português reestruturado indicia que ela não terá emergido aquando da desactivação dos engenhos do açúcar e retirada dos colonos após tomada do arquipélago pelos holandeses, no século XVII, e que determinou também a fixação de vários grupos comunitários no interior da ilha, recorrendo, essencialmente, à agricultura de sobrevivência. Assim, o grupo de forros que se estabeleceu na ampla Baía de Almojarife seria constituído por libertos santomenses que, em movimentos deambulatórios iniciados em finais do século XIX/princípios do século XX, buscavam pontos de fixação afastados das roças, a fim de não serem compulsivamente integrados na força braçal destas. Uma vez instalada na baía, a comunidade procurou o seu sustento na agricultura de pequena escala e na pesca artesanal, uma actividade até então exercida quase exclusivamente pelos angolares (Esquadrão de Polícia Militar N° 2222, 1962:1-2).

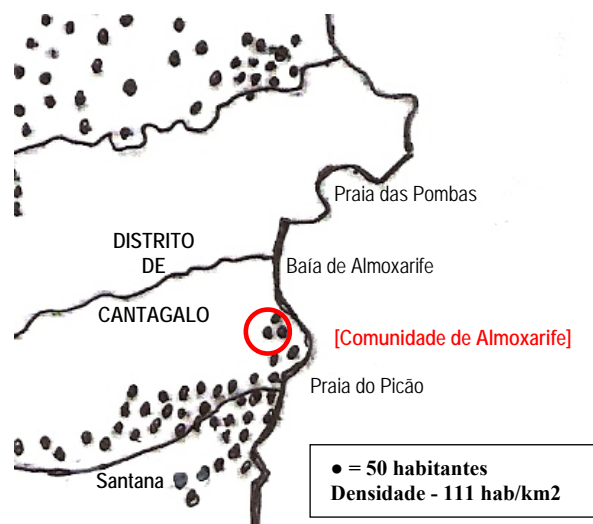


Fig. 6. Índice demográfico de Santana e Almojarife (Tenreiro, 1961:164).

Situada na zona Este da ilha de São Tomé, caracterizada como litoral baixo, a Baía de Almojarife é confinada a norte pela Praia das Pombas e a sul pela Praia do Picão. Pertence ao Distrito de Cantagalo, que tem por capital a povoação de Santana. De acordo com o Censo de 1960 (Tenreiro, 1961:163), o distrito albergava cerca de 13258 habitantes, espalhados por uma área de 119 km<sup>2</sup>. A densidade populacional rondava os 111 habitantes por km<sup>2</sup> (Fig. 6). Quanto à capital, possuía, à data, 7702 habitantes (4353 do sexo masculino e 3349 do sexo feminino), enquanto a pequena comunidade de

Almoxarife era constituída por cerca de 150 forros, que se concentraram na parte sul da baía (Tenreiro, 1961:164).

A situação de afastamento linguístico relativamente ao PE acentuou-se para os almoxarifanos, ajudando a fixar as características linguísticas defectivas na língua nativizada, em virtude de os DLP's transmitidos geracionalmente preservarem a gramática dos falantes das gerações mais velhas, marcada por acentuada não-aquisição de morfologia relativamente à LA. O isolamento da comunidade e conseqüente pouco contacto com os falantes do PE prolongaram-se ainda por bastante tempo, fosse pelo receio dos membros da comunidade em serem obrigados a regressar às roças fosse pela política do Estado português, que somente em 1926 aprovaria, por decreto emitido pelo Estado Novo, o *Estatuto Orgânico das Missões Católicas Portuguesas de África e Timor*. Este, por seu lado, atribuía à Igreja Católica o papel exclusivo na evangelização e “civilização” das populações indígenas não assimiladas dos territórios coloniais portugueses. Quanto à rádio, que poderia permitir às comunidades isoladas um contacto mais próximo com o PE, seria introduzido em São Tomé em 1925, mas somente em 1949 é que foi fundada a Rádio Clube de São Tomé, com uma maior capacidade de alcance e que se encontra ainda em funcionamento.<sup>42</sup> Relativamente à televisão, só seria introduzida na ilha alguns anos depois da independência, já em 1982. Contudo, privados de rede eléctrica em Almoxarife, os membros da comunidade apenas tinham acesso a ela caso se deslocassem a Santana, uma situação que se mantém até aos dias de hoje.

No final da década de 60 do século XX (apenas sete anos antes da independência) surgem as primeiras referências sobre a subida do nível de educação da população em relatórios oficiais, que não é acompanhado pelas oportunidades de emprego (CTI de São Tomé, 1968a:1) e faz com que alguns forros das novas gerações voltem a procurar trabalho nas roças (CTI de São Tomé, 1968b:1-2). Como se verifica, apesar de o ensino e o exército português, que integra elementos nativos, permitirem às gerações mais novas um contacto estreito com o PE, facultando ao PtgL1 nativizado vectores de expansão estrutural em direcção à LA e minorando os processos de transferência, o contacto com o português dos “contratados”, marcado por características das línguas bantu, não cessa para os falantes forros.

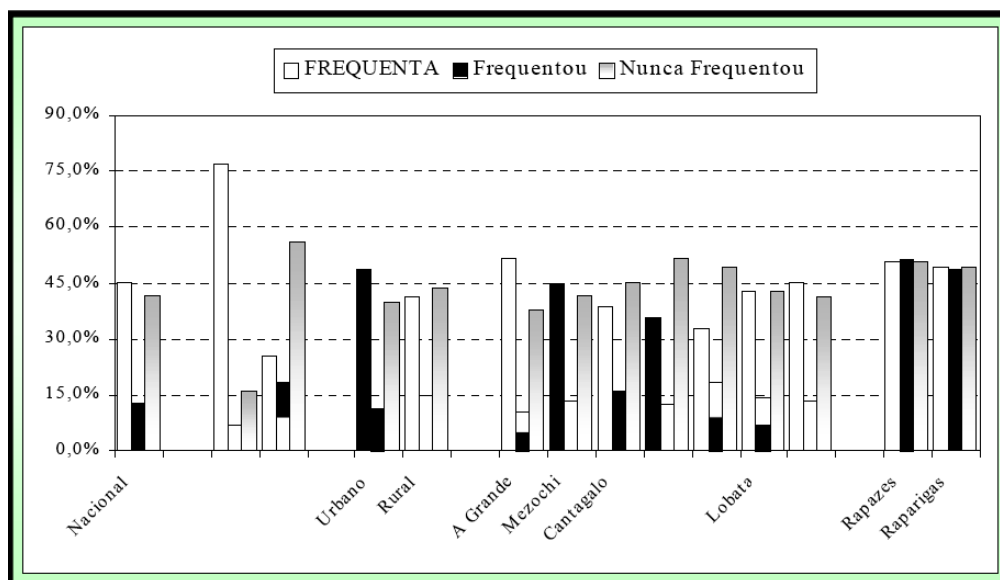


Nos anos 70 do século XX, o preço do cacau cai nos mercados internacionais e a economia de plantação da ilha entra em recessão, acontecendo o êxodo de muitos cabo-verdianos, que regressam a casa ou partem para a metrópole (CTI de São Tomé, 1970:1, 1972a:1, 1972b:1). Actualmente, e no que diz respeito aos estrangeiros residentes no Distrito de Cantagalo, representam cerca de 10% da população deste, sendo também o grupo cabo-verdiano aquele que continua a deter o maior efectivo de residentes estrangeiros no arquipélago (57%) (Instituto Nacional de Estatística de São Tomé, 2003a:24-26). Na década de 70, a ilha possui já uma eficiente cobertura escolar a nível do território, coadjuvada por “... *uma boa rede de estradas em que a pavimentação de asfalto ultrapassou já os 220 Km...*” (Unidade da Companhia de Artilharia 3376, 1973a:3), bem como a existência de “... *boas rodovias que normalmente servem as sedes das propriedades agrícolas*” (Unidade da Companhia de Artilharia 3376, 1973a:3), aspecto que terá contribuído para que, de certa forma, o português reestruturado falado pelos núcleos urbanos e pelas comunidades rurais mais isoladas se nivelasse, eliminando-se bipolarizações entre a fala de ambas, sobretudo a nível das gerações mais novas e mais escolarizadas da ilha. Nesta fase, o contacto mais estreito com o PE por parte dos falantes bilingues de São Tomé, em geral, e da comunidade de Almojarife, em particular, vai permitir a incrementação de empréstimos de traços lexicais e estruturais da LA (cf. Thomason & Kaufman, 1988; Lefebvre, 2003; Montrul, 2006; Lefebvre, White & Jourdan, 2006b; Siegel, 2008), colocando gramáticas em competição e originando a CPL-var no SN produzido, actualmente, pelos almojarifanos.

Em 12 de Julho de 1975, com a independência de São Tomé e Príncipe, fechava-se um capítulo de meio século de vida do arquipélago como colónia, marcado por períodos de auge e decadência económica, com consequências profundas na construção da identificação sociolinguística de um povo descendente de antigos escravos, transplantados e desenraizados geográfica e culturalmente das suas origens. Com a independência, foi instaurado um regime monopartidário, que procedeu à nacionalização das roças, levando cerca de 4000 portugueses a abandonarem o arquipélago. O período que se segue à independência é marcado por grande conflitualidade interna, sucedendo-se as tentativas de golpes de estado. Com as questões socioeconómicas jogadas ao esquecimento, as estruturas do país foram-se degradando, sendo as comunidades rurais

votadas ao abandono e tornando-se cada vez mais isoladas, face à desarticulação dos sistemas educacional, rodoviário e de comunicação.

No que concerne à questão da educação no Distrito de Cantagalo, presentemente cerca de  $\frac{1}{4}$  da sua população jovem encontra-se em fase de idade escolar obrigatória (6-12 anos),<sup>43</sup> não existindo praticamente assimetria entre ambos os sexos. Contudo, os valores que se referem à frequência escolar das crianças com idade compreendida entre os 5 e os 17 anos evidenciam que o número de analfabetos do Distrito de Cantagalo (46%), nesta faixa etária, é superior ao dos que frequentaram (15%) ou frequentam (39%) a escola (Fig. 7). Este aspecto, não só revela as dificuldades que as crianças dos meios rurais isolados ou semi-isolados, nos quais se inclui a comunidade de Almoxarife, sentem para conseguir frequentar presentemente a escola como tem também evidentes reflexos na questão da estabilização da variação registada a nível da concordância do SN, visto o elevado número de analfabetos não ter contacto com as normas do PE, que poderiam ajudar na aquisição/expansão da regra de concordância.

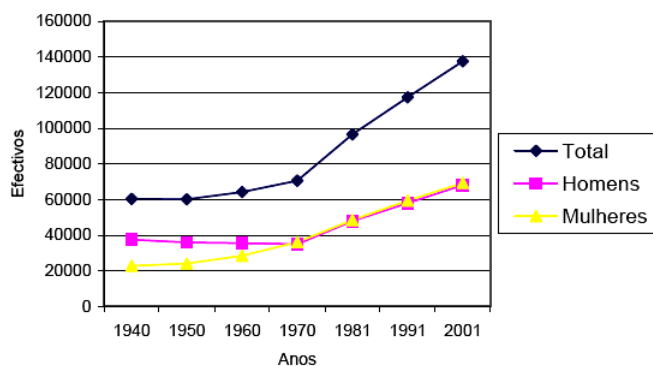


**Fig. 7.** Percentagem das crianças com 5-17 anos que têm acesso à escolaridade. (Fonte: Instituto Nacional de Estatística de São Tomé e Príncipe, 2003b:23).

A desarticulação do sistema educacional, traduzido no corte do contacto entre o português popular e o PE, tem como consequência, segundo o Censo de 2001, o facto de 98,9 % da população actual do arquipélago entender e usar o português, mas apenas a classe política e a alta sociedade, ou seja, aquelas que fazem grande parte dos seus

estudos em Portugal, utilizarem a variedade padrão (Instituto Nacional de Estatística de São Tomé e Príncipe, 2003/2005).

Dados do Censo de 1970 (Unidade da Companhia de Artilharia 3376, 1973b) mostravam que a população de Santana, contrariando a tendência geral de crescimento populacional verificada no país, decrescera ligeiramente (7496 indivíduos), levando a um ajustamento na distribuição entre sexos (3794 homens e 3702 mulheres), que se mantém até à actualidade. Este aspecto indicia que os membros mais jovens abandonam a região e não regressam, fazendo com que a fala almoxarifana continue a conservar as suas características mais crioulizantes. Quanto ao esbatimento da diferença que existia no efectivo de homens e mulheres, estará relacionado, sobretudo, com o facto de, na região, serem os primeiros quem mais migra para os centros urbanos<sup>44</sup> ou emigra para o estrangeiro. No entanto, a nível geral do país, é possível constatar que, entre 1940 e 1970, a população masculina foi sempre superior à feminina, tendo o desequilíbrio origem na população estrangeira levada para São Tomé sob contrato forçado para trabalhar nas roças, nas quais predominava a mão-de-obra masculina. Contudo, finalizada a entrada de “contratados” e aumentando a emigração dos homens em busca de melhores condições de trabalho e vida no estrangeiro, inverteu-se a situação, com o número de habitantes do sexo feminino<sup>45</sup> ultrapassando o do sexo masculino (Fig. 8):



**Fig. 8.** Evolução da população de São Tomé e Príncipe – 1940 a 2001.  
(Fonte: Instituto Nacional de Estatística de São Tomé e Príncipe, 2003a:9).

### 1.2.2.3.1. Características da fala de Almoxarife

A situação que determinou o aparecimento da comunidade de Almoxarife tem, obviamente, implicações directas no português usado por esta. A economia da comunidade nunca teve qualquer relevância para o país, já que os seus membros se

debatem com as dificuldades resultantes do semi-isolamento e dos seus poucos recursos de subsistência, provenientes da pesca e da agricultura (milho, batata-doce, mandioca ou quiabo) de mera subsistência. Ainda assim, a comunidade, principalmente por intermédio das mulheres, transacciona uma escassa parte destes recursos com os poucos comerciantes locais ou em mercados populares de reduzida escala. Portanto, a fala de Almojarife, ainda que de forma pouco significativa, não deixa de estar submetida a um *continuum* linguístico, aspecto que, por exemplo, não sucede tão frequentemente com as comunidades mais isoladas de algumas roças.

O acesso entre Almojarife e Santana é feito através de uma estrada de terra batida, mas nunca existiu uma rede de transportes que permitisse aos membros da comunidade utilizarem-na para se deslocarem até à capital do distrito. Como o percurso tinha que ser feito a pé, na época colonial eram os próprios comerciantes que, utilizando as suas viaturas, se deslocavam à baía para comerciarem com os almojarifanos. Este aspecto terá contribuído bastante para o semi-isolamento da comunidade antes da independência do país. Não obstante, vencido o percurso entre Almojarife e Santana, era possível aos membros da comunidade deslocarem-se com facilidade para outros pontos da ilha, já que, estando incluída no leque das vilas rurais mais importantes da colónia, com a categoria de povoação administrativa, Santana era servida por uma estrada com pavimento em asfalto e que a ligava a São Tomé, Angolares e Caué (Unidade da Companhia de Artilharia 3376, 1873b:6). Contudo, após a independência, o isolamento da comunidade acentuou-se consideravelmente, visto que os colonos abandonaram a ilha, as rodovias se degradaram, a ponto de se tornarem intransitáveis, e se extinguiu a rede de transportes públicos que permitia aos seus usuários deslocarem-se regularmente para outros locais. Disso mesmo nos dá conta o excerto do diálogo entre o entrevistador e o informante [OSVALH1], que deixa transparecer a inexistência de contactos entre os membros da comunidade e os habitantes de Santana:

DOC: *Você custuma i a Santana?*

INF: *Vô umas vezes devido quê zê essa prigiça de i a pé e volta a pé.*

DOC: *Hum, mas você vai assim em Santana mais pra fazê o quê?*

INF: *Quê zê é vô mais a Santana assisti programa.*

DOC: *Hum.*

INF: *Só!*

Como se verifica, o *continuum* linguístico da comunidade almoxarifana tinha como consequência, até um passado bastante recente, a manutenção da situação de afastamento linguístico relativamente ao PE, contribuindo para conservar as características defectivas do PtgL2. Os falantes da faixa etária mais idosa (+65 anos) afirmaram todos que tiveram o santomense como L1, embora alguns deles tenham também aprendido o PtgL1 através do contacto com o PtgL2/PtgL1 dos pais, impregnado de características crioulizantes.

No que concerne à geração intermédia de Almojarife (41-65 anos), apenas a informante [MANEVEM2] (Tabela 8.1) aprendeu o português e o santomense simultaneamente, já que todos os outros falantes tiveram o segundo como L1. Alguns dos falantes desta faixa etária trabalharam sazonalmente nas roças, fazendo com que o PA entrasse em contacto com o português adstrato dos “contratados”, marcado por características das línguas bantu. Este aspecto terá também contribuído para que a situação de pouco contacto com o PE se mantivesse, já que a comunicação entre os capatazes e trabalhadores era feita, muitas vezes, com os primeiros recorrendo a formas precárias do português.

O *continuum* linguístico da comunidade almoxarifana teve como consequência, até um passado bastante recente, a manutenção de características defectivas em relação à LA. Estas constituíram, posteriormente, os modelos disponíveis para o processo de nativização do PAL1, adquirido em simultâneo com o santomense pela geração mais nova da comunidade (20-40 anos), mas com o primeiro revelando-se como a língua dominante e o segundo sendo estigmatizado no contexto de aquisição bilingue:

DOC: *Qual qual foi a língua que você aprendeu primeiro?*

INF: *Português.*

DOC: *Não aprendeu dialecto em casa não?*

INF: *Ê també aprendi. [Põe lá pra mim, Sã!]*

DOC: *Sua vó falava em casa com ocê como?*

INF: *Falava dialecto também.*

DOC: *Falava mais dialecto ou português em casa, sua avó?*

INF: *Hum, mas assim quan gen fala dialecto assim senhora não gostava.*

DOC: *Ahã tinha que falar português?*

INF: *Hum.*

DOC: *A... a criança que fala dialecto com mais velhos é falta de respeito?*

INF: *Não.*

DOC: *Hum, pode falar?*

INF: *Pode falá.*

DOC: *Mas por que que sua avó não gostava?*

INF: *Minha avó diz gen era muita criança ainda quê pa tá usa dialecto e assim quem gen pa um meio de pissoa vai falar assim mesmo.*

DOC: *Hum.*

INF: *Calha feio.* [ANTOM1]

No ponto 1.1 do presente abordou-se a questão dos fenómenos induzidos por contacto: a transferência e o empréstimo (Thomason & Kaufman, 1988:73-74). O santomense e o português dos “contratados” compartilhavam algumas afinidades relativamente ao sistema de marcação PL, já que ambas as línguas recorriam a um marcador pré-nominal, ou seja, um marcador inserido apenas antes do núcleo do SN, mantendo-se invariáveis todos os restantes elementos da cadeia sintagmática. Este sistema de marcação ancora no dos substratos do grupo níger-congo atlântico e resulta de transferência, já que os falantes responsáveis por esta eram monolíngues e, à medida que elaboravam a interlíngua, transferiam estruturas da língua nativa para a língua na qual procuravam comunicar (Thomason & Kaufman, 1988:73-74). Posteriormente, os falantes bilingues do português nativizado levaram a cabo empréstimos a partir da LA (Thomason & Kaufman, 1988:73-74), incorporando os determinantes do português no seu sistema de marcação e diminuindo a frequência de utilização de formas que revelavam apego às estruturas do santomense, como acontece com o uso do pronome pessoal terceira pessoa do PL (exemplo [48]) ou de nomes nus (exemplos [73] e [74]). No entanto, apesar de a faixa etária mais nova de Almoxarife ter tido um contacto mais próximo com formas da LA, sobretudo por intermédio do ensino, os factores linguísticos e extralinguísticos que intervieram na instanciação de novos/errados parâmetros (Gonçalves, 2004:235) não possibilitaram que ocorresse o restabelecimento de novos valores paramétricos (cf. Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), suficientes para diluir completamente os traços já adquiridos. Desta forma, o PA reestruturado revela algumas diferenças em relação ao PE, seja a nível fonológico (p.e. fenómenos de lambdacismo ou de abaixamento na realização de vogais altas do PE) seja no plano da morfossintaxe (p.e. não realização de determinadas preposições ou variação na concordância nominal de número e na concordância verbal), ocorrências típicas das situações de transmissão linguística irregular e que, no caso concreto da fala de Almoxarife, respaldam nas duas

gramáticas em competição: crioula do substrato santomense e portuguesa, da LA (PE elitista e do ensino).

Os fenómenos que afectam o PA apresentam uma incidência mais significativa entre os informantes mais idosos, fazendo prever que a frequência da variação seria mais elevada na fase em que a comunidade foi constituída, o que confirmaremos empiricamente nos capítulos 4 e 5, dedicados à análise das variáveis constituídas. Ainda assim, é possível detectar no PA estruturas variáveis decorrentes de processos de não-aquisição de matéria gramatical que acontecem na sequência do contacto entre línguas. Listamos, seguidamente, alguns dos traços que caracterizam o dialecto de Almoxarife:

#### ► Estrutura do SN

- (i) Uso variável do artigo definido (exemplos [240] e [243]), que pode determinar a realização de nomes nus:

[124] PA:  
*depois eu ia tê com meu amigo* [OSVALH1]  
“depois eu ia ter com o meu amigo”

[125] PA:  
*faço trabalhos caseira até tarde* [CARMOM1]  
“faço os trabalhos caseiros até tarde”

- (ii) Variação na colocação do possessivo (exemplo ([68]);

[126] PA:  
*E amigo meu tinha uma mulher* [OSVALH1]  
“E o meu amigo tinha uma mulher”

- (iii) Concordância variável de número entre os elementos do SN, com tendência para a marcação ser feita com recurso a um PL *singleton*, visível na categoria DET (Castro & Pratas, 2006), dispensando, depois, a marcação no nome núcleo (Castro & Ferrari-Neto, 2007:75) e em outros elementos do SN (Kiparsky, 1972:195), por se revelar redundante:

[127] PA:  
*já trabalhê esses lugar tudo* [LUISH2]  
“já trabalhei nesses lugares todos”

- (iv) Variação na flexão de género dos nomes referentes a seres animados, com recurso ao valor referencial do sexo usando nomes adjetivados:

[128] PA:  
*Dois filho mulher qu'ê tem* [FRANCM2]  
“As duas filhas que eu tenho”

- (v) Concordância variável de género entre os elementos do SN:

[129] PA:  
*faço trabalhos caseira até tarde* [CARMOM1]  
“faço os trabalhos caseiros até tarde”

### ► Morfologia verbal

- (vi) Uso de formas do presente para expressar acções e estados passados:

[130] PA:  
*Ê os meus pai que me deu essa ideia.* [OSVALH1]  
“Foram os meus pais que me deram essa ideia”

- (vii) Uso do modo indicativo para expressar acções e estados que requerem modo conjuntivo (exemplo [131]) e vice-versa (exemplo [132]):

[131] PA:  
*eu ñ viu nenhum governo que tá judá nós* [FRANCM2]  
“eu não vi nenhum governo que nos esteja a ajudar”

[132] PA:  
*às vez ponha quinze* [MMDEUSM3]  
“às vezes põe quinze”

- (viii) Uso de formas aparentemente finitas em contextos que requerem formas do infinitivo (exemplo [133]), e vice-versa (exemplo [134]):

[133] PA:  
*pode compra livro depois mais tarde* [CELESH3]  
“pode comprar o livro depois, mais tarde”

[134] PA:  
*Viver mal, viver num país tá muito mal* [MMDEUSM33]  
“Vivemos mal, vivemos num país que está muito mal”

- (ix) Não distinção entre posse e existência, com recurso ao verbo “ter” para expressar a segunda:



[135] PA:  
*aqui ni praia Mocharife nã tem nada nã tem luz, nã tem nada*  
[OSVALH1]  
“aqui na praia de Almocharife não há nada, não há luz, não há nada”

(x) Redução de concordância verbal, que atinge todas as pessoas do verbo:

[136] PA:  
*eu nã viu nenhum governo que tá judá nós* [FRANCM2]  
“eu não vi nenhum governo que nos esteja a ajudar”

[137] PA:  
*nós compra coesa bucado barato* [FRANCM2]  
“nós comprávamos as coisas um bocado baratas”

[138] PA:  
*É os meus pai que me deu essa ideia.* [OSVALH1]  
“Foram os meus pais que me deram essa ideia”

(xi) Variação relativamente à presença de verbo copulativo (exemplo [139]), que pode também ser substituído por verbo paracopulativo (exemplo [140]):

[139] PA:  
*depois dôtör disse aqui só pa fazer operação* [SALVH3]  
“depois o doutor disse-me que ali era só para fazer operações”

[140] PA:  
*Nã, tá sorteido.* [FRANCM2]  
“Não, sou solteira.”

### ► Relações sintagmáticas e oracionais

(xii) Variação no emprego das preposições (exemplo [141]), que pode passar pela omissão destas (exemplo [142]), pela aplicação de formas redundantes (exemplo [143]) ou pelo uso do item “ni”, em substituição de “em” (exemplo [144]):

[141] PA:  
*o quê você faz de manhã até na hora que ocê vai dormir de noite?*  
[OSVALH1]  
“o que é que você faz desde manhã até à hora em que você vai dormir à noite”

[142] PA:  
*É, bêra strada* [CLOTIM2]  
“É, à beira da estrada”

[143] PA:  
*daqui ê fui para até hospital* [OSVALH1]  
“daqui, eu fui ao (para o) hospital”

[144] PA:  
*porque eu estive lá ni empresa trabalhar* [MANOH3]  
“porque eu estive lá na empresa a trabalhar”

(xiii) Variação no emprego de complementadores:

[145] PA:  
*há momento pessoa levanta matabicha, nã almoça até jantá* [ANTOM1]  
“há momentos em que a pessoa se levanta, matabicha,<sup>46</sup> não almoça até ao jantar”

[146] PA:  
*depois dôtor disse aqui só pa fazer operação* [SALVH3]  
“depois o doutor disse-me que aqui era só para fazer operações”

(xiv) Não distinção entre formas rectas e átonas do pronome pessoal, com tendência para uso das primeiras em detrimento das segundas:

[147] PA:  
*eu tem entregá el coesa del* [OSVALH1]  
“eu tinha de lhe entregar as coisas dele”

[148] PA:  
*cercaram ele* [CARMOM1]  
“cercaram-no”

(xv) Não distinção entre as formas rectas e tónicas do pronome pessoal, usando-se as segundas pelas primeiras:

[149] PA:  
*mim depois saída de outro comício* [CELESH3]  
“eu depois saí de outro comício”

(xvi) Não distinção entre as formas do OD e OInd pronominais, substituindo-se as primeiras pelas segundas:

[150] PA:  
*nã lhe vi* [ABILH2]  
“não o vi”

(xvii) Variação na colocação do pronome pessoal:

[151] PA:  
*o seôro se dá bem co esse seu filho* [ABILH2]  
“o senhor dava-se bem com esse seu filho”

A variação e/ou não-realização de partículas gramaticais (artigos, morfemas flexionais, preposições e complementadores) são marcas distintivas do PA reestruturado, enquanto dialecto que emergiu por transmissão linguística irregular. A sociohistória da comunidade de Almojarife, por seu lado, concede evidências que permitem relacionar, empiricamente, tais fenómenos quer com o processo massivo, defectivo e não-normatizado de aquisição do português, marcado por acentuada alteração nos seus parâmetros sintácticos e levado a cabo por escravos africanos geográfica e culturalmente transplantados para o entreposto de São Tomé (p.e. não realização de preposições), quer com o processo defectivo e não-normatizado de nativização do PtgL2, marcado por fenómenos de reestruturação morfossintáctica mais leves (p.e. CPL-var no SN, com a marcação a incidir apenas na categoria DET). Inicialmente, devido à situação de contacto restrito relativamente ao PE, o PA não teve substancial acesso aos modelos deste, o que determinou que o seu processo de reestruturação ocorresse sem referências normatizadoras. Com campo aberto para a fixação de estruturas desviantes, o PA incorporou variação de diferentes tipos, que foi transmitida de geração em geração. Mais tardiamente, com o nascimento de um sistema de rodovias e de uma rede de transportes públicos que interligava meios rurais e urbanos, a criação e expansão de um sistema de educação pública abrangente, a implantação da rádio na ilha e a incorporação de nativos nos quadros do exército colonial e da administração pública local, alguns falantes do PA passaram a ter contacto mais estreito com a LA. Este facto possibilitou que a variação se diluísse, mas o processo em direcção à aquisição das regras da norma foi interrompido após a independência do país, já que o sistema socioeconómico deste se desestruturou, lançando as comunidades do interior para um isolamento que condiciona o *continuum* linguístico. Como tal, a variação registada na comunidade encontra-se presentemente estabilizada (capítulo 5, análise das variáveis extralinguísticas), pelo que os actuais traços do dialecto de Almojarife terão que ser atribuídos não só ao processo de transmissão linguística irregular que levou ao seu aparecimento mas também ao extenso período de semi-isolamento em que a comunidade se encontrou (e encontra) mergulhada.

Os fenómenos de variação que se detectam no PA não encontram similaridades significativas no PE, mas estabelecem paralelos claros com outras variedades africanas e brasileiras de português que emergiram por contacto, nomeadamente no que concerne à variação que atinge os sistemas de concordância verbal e de concordância nominal de número. Esta particularidade permite-nos, desde já, questionar a hipótese de estes fenómenos ancorarem em mudanças fonéticas inerentes ao próprio português, como tem sido defendido pela corrente derivista a propósito das motivações históricas das características que diferenciam o PB do PE (Naro, 1981; Naro & Scherre, 1993, 2000, 2007a, 2007b). Por outro lado, os paralelismos referidos permitem assumir que o contacto entre línguas terá sido crucial para a emergência e fixação dos fenómenos em questão quer no PA quer no PB. Assim, uma análise variacionista e comparativa sobre a CPL-var sistemática no SN do PA, enquanto dialecto falado por descendentes dos antigos escravos transplantados para o entreposto de São Tomé na rota do tráfico afro-negreiro para o Brasil, permitirá visualizar até que ponto se pode estabelecer uma corrente que conecte tal fenómeno às línguas africanas (cf. Guy, 1981a, 1989, 2005).

### **1.3. Conclusões parciais**

No presente capítulo lançámos um olhar sobre alguns trabalhos acerca da CPL-var no SN em variedades brasileiras e africanas de português, emergentes em situação de transmissão linguística irregular. A CPL-var no SN de tais variedades reflecte aspectos da própria sociohistória que determinou o aparecimento das comunidades que comunicam com recurso a elas. No caso do Brasil, este tipo de variação tem sido conotado, por um lado, a mudanças fonéticas naturais internas ao português, que foram determinadas pela deriva diacrónica (Naro & Lemle, 1976; Naro, 1981; Naro & Scherre, 1993, 2000, 2007a, 2007b), e, por outro lado, a fenómenos morfossintácticos de reestruturações colaterais, resultantes do contacto massivo entre línguas ocorrido no país (Baxter & Lucchesi, 1997, Lucchesi, 1998a, 2001; Baxter, 2009), mas que radicam nos ancestrais dialectos africanos falados pelos escravos africanos transplantados e seus descendentes (Guy, 1981a, 1989, 2005; Baxter & Lucchesi, 1997, 1998; Lucchesi, 2001; P. Andrade, 2003; Baxter, 2009). No caso específico da comunidade de HEL-Ba, Baxter (2009:293) observou a CPL-var no SN combinando teorias estruturais e aquisicionais,

tendo concluído que: (i) a variação é motivada por questões como o desequilíbrio demográfico entre contingentes escravizados e não-escravizados, a mudança de LA na comunidade, a presença de bilinguismo nesta, ou as situações de pouco acesso às formas linguísticas da LA; (ii) o desenvolvimento da concordância PL inicia-se com a introdução de um PL *singleton* visível na categoria DET, para, posteriormente, ganhar características de concordância parcial, através de morfemas dissociados, que são atribuídos aos elementos do SDET mais próximos do núcleo, primeiro, e aos itens à direita deste, depois. Para Baxter (2009:292), a variação da concordância deverá ser conectada à situação de contacto que envolveu componentes de PtgL2 influenciado pelas línguas bantu, cujas estruturas de marcação PL se assemelham às das línguas atlânticas da tipologia níger-congo atlântica (balanta e wolof), substratos do CVV. Nestas, a pluralização é controlada por um classificador nominal que estabelece paralelismos estruturais com a categoria DET das variedades africanas e brasileiras de português, reestruturadas por transmissão linguística irregular, e que tenham, como substratos directos, quer as próprias línguas africanas do grupo níger-congo atlântico (PM, PVA e PT) quer os CP's atlânticos dos quais estas são também substratos (PA e PCV). Para reforçar este pressuposto, apresentámos exemplos de paralelismos estabelecidos a nível do SN entre línguas africanas atlântico-ocidentais, CP's atlânticos, variedades africanas de português com substrato do grupo níger-congo atlântico, variedades africanas com substrato crioulo e variedades brasileiras de português. Quanto à manutenção geracional da CPL-var no SN nas referidas variedades reestruturadas, é determinada pela preponderância quantitativa de falantes de PtgL2, que esbate a possibilidade de um reajustamento paramétrico e leva a que os novos/errados parâmetros sejam tomados como modelos para uma nova reinstanciação paramétrica no PtgL1 nativizado (Gonçalves, 2004:235).

A elaboração do capítulo abordou ainda a descrição sociohistórica da comunidade de Almojarife e levantou as marcas pertinentes do português reestruturado falado pelos almojarifanos. As marcas apontam para a possibilidade de o seu entendimento ter que ser verificado a partir de constatações que ancoram em modelos teóricos distintos, sejam eles da aquisição, sociolinguísticos ou psicolinguísticos. Portanto, o suporte para análise dos nossos dados terá que ser o mais abrangente possível, pelo que, no próximo capítulo,

daremos conta dos quadros teóricos que, no nosso entender, poderão ajudar a determinar se a aquisição da pluralização nos constituintes do SN, e seu posterior desenvolvimento estrutural, ancoram em princípios regidos pelos universais linguísticos ou, se, contrariamente, configuram quer casos de transferência que se fossilizam quer empréstimos provenientes da LA.



## CAPÍTULO 2

*“Ab uno disce omnes.”*

(In, Virgílio, *Eneida*, II, p.65)

### Fundamentação teórica

Algumas das variedades de português que emergiram em situação de transmissão linguística irregular mais radical no Brasil exibem traços de criouliização leve (Baxter, 1992, 1997, 2009; Baxter & Lucchesi, 1997; Lucchesi, 2000a; Lucchesi & Baxter, 2009). No presente capítulo abordaremos as perspectivas teóricas que permitirão verificar se as estruturas criouliizantes de tais variedades apresentam idêntico perfil ao de outras variedades de português, especialmente as que possuem um crioulo como substrato, como acontece com o PA. Portanto, o estudo da CPL-var no seu SN é fundamental para se perceber a forma como os factores que regulam a variação e a mudança actuam em conjunto com o processo de aquisição da L1 para determinarem o modo como o conhecimento linguístico se orienta na mente quer dos falantes almorarifanos quer dos falantes das referidas variedades de PB. Assim, e independentemente de estarmos verificando se a aquisição dos constituintes do SN e seu consequente desenvolvimento estrutural ancoram em princípios regidos pelos universais linguísticos ou, em contrapartida, configuram casos de transferência, mais ou menos fossilizada, estaremos também assumindo que a variação e mudança decorrentes no PA, a existirem, não serão unicamente reguladas pelas suas estruturas gramaticais mas também por estruturas inatas comuns ao ser humano e formalizadas pelo conceito da existência de uma GU (Chomsky, 1957, 1965, 1981a, 1981b, 1986a, 1996 [1995]).

Para Labov (1982:18), a Sociolinguística Variacionista deve centrar atenção na forma como os factores sociais determinam os padrões linguísticos usados colectivamente numa certa comunidade, elaborando os estágios do desenvolvimento histórico da sua fala. Desta forma, a mudança linguística pode ser predita estabelecendo-se a relação entre a variação linguística analisada num momento específico do devir da fala dessa mesma comunidade e o sentido em que a mudança em curso vem operando. Portanto, as variações reflectem estágios não-aleatórios do processo de mudança em progresso, tendo na sua origem a conjunção de fenómenos linguísticos e sociais que vão edificando, progressiva e estruturalmente, o todo linguístico-comunitário.



Assim, caberá à Sociolinguística o papel de observar os fenómenos linguísticos numa perspectiva sociohistórica, relacionando subjectivamente, à estrutura social e ideológica da comunidade, os dados empíricos que reflectem variação e indiciam uma mudança em curso. Não obstante, a não elaboração de uma teoria da estrutura em termos sociolinguísticos tem criado obstáculos ao encaixamento dos processos de variação e mudança na estrutura linguística, uma vez que apenas um reduzido número de estudos (p.e. Tarallo, 1987, 1991; Kato & Tarallo, 1989; Kato, 1999; Ramos, 1999; Duarte, 1999; Kroch, 1989b, 2001, 2004; Kroch & Taylor, 1997; Roberts, 2007) tem procurado estabelecer uma relação entre o comportamento linguístico dos falantes que integram uma dada comunidade de fala (língua-E) e a sua competência linguística enquanto usuários nativos de uma determinada língua natural (língua-I).

No presente capítulo procuraremos dar conta dos princípios teóricos que orientarão a metodologia do nosso estudo, a qual adoptará como pedra basilar de fundamentação a conjugação dos dois modelos referidos, que serão complementados por outros quando o entendermos necessário. Ainda que os modelos concernentes à língua-I e língua-E apresentem divergências na sua forma de observar as línguas, estas divergências não deixam também de se conjugarem *“para o entendimento da questão da mudança, tanto no plano dos princípios teóricos fundamentais, quanto no plano da metodologia e dos critérios da verificação empírica”* (Lucchesi & Ribeiro, 2009:126). Assim sendo, no presente estudo procuraremos, de alguma forma, integrar a dimensão psico-biológica do fenómeno da CPL-var registada no SN do PA na dimensão sociohistórica da comunidade almoxarifana. Com tal metodologia, buscamos entender até que ponto os mecanismos da faculdade humana da linguagem (cf. Chomsky, 1986a, 1996 [1995]) intervêm também nas relações sociais da comunidade, conjugando-se com estas no processo que determina o fenómeno da variação registado a nível da concordância de número entre os elementos do SN do português falado em Almojarife.

Ao aplicarmos os princípios teóricos e a metodologia mencionados ao estudo da CPL-var no SN do PA não temos a pretensão de fornecer uma solução para colmatar a problemática que emerge quando se pretende conjugar dois modelos tão distintos. No entanto, acreditamos que é possível recorrer também a elementos da teoria da gramática para uma análise dos padrões colectivos de comportamento linguístico da comunidade

almoxarifana, contribuindo, com tal, para se observar como se processa a aquisição e regularização de fenómenos variáveis em línguas emergentes em contexto massivo de mono e bilinguismo, sejam elas L2's sejam elas L1's. De facto, analisando os processos em que as gramáticas concorrentes na comunidade de Almoarifife afectam o desempenho linguístico dos seus utilizadores, isto é, o modo como as distintas gramáticas geracionais determinam alternâncias de códigos que se encaixam no conhecimento linguístico dos almoarififanos para elaborarem estruturas gramaticais coesas, é possível não só compreender o desenvolvimento da CPL-var do PA, em geral, mas também fornecer mais alguns dados para complementar o entendimento acerca do *puzzle* que é a realidade actual dos falares rurais e populares do Brasil.

Face ao exposto, iremos focar aspectos em torno das abordagens sobre a teoria da gramática e o estudo dos padrões colectivos de comportamento linguístico, nomeadamente acerca de alguns trabalhos que buscam conjugar os modelos discutidos sob o ponto de vista da Sociolinguística (cf. Weinreich, Labov & Herzog, 2006 [1968]) e da Gramática Generativa, a fim de explicarem, com recurso a modelos matemáticos para a reprodução do efeito global dos dados empíricos, como se dão as mudanças linguísticas a nível comunitário.

Paralelamente, e visto o nosso estudo ter como objectivo principal observar o comportamento da concordância de número a nível do SN, abordaremos a questão da flexão em número nos constituintes que integram o SN do PE, verificando-se que é aparente o conceito de que este idioma insere marcas de número PL em todos os elementos do SN. Para tanto, importará também caracterizar o funcionamento do SN em português, referindo as dificuldades que o sistema de flexão e concordância em número podem colocar na aquisição do PtgL2. Para esta caracterização, iremos descrever, sucintamente, o quadro teórico que orientará as nossas observações, sobretudo as propostas em torno da aquisição de L1, aquisição bilingue e ASL em situações sociohistóricas específicas.

Trataremos igualmente a questão particular da aquisição da categoria funcional SDET, na qual se alojam os traços de género (Koehn, 1994; Müller, 1994a; Ferrari-Neto, Corrêa & Augusto, 2005) e de número (Enç, 1991; Koehn, 1994; Müller, 1994a; Longobardi, 1994; Ferrari-Neto, Corrêa & Augusto, 2005; Castro & Pratas, 2006; Costa & Silva,

2006b; Alexandre & Hagemeyer, 2007; Castro & Ferrari-Neto, 2007; Baxter, 2009). Neste aspecto, será também fundamental perceber como a concorrência entre gramáticas pode afectar a morfologia de número a nível do SDET em situações de transmissão linguística irregular. De facto, esta configura a realidade sociolinguística que determinou a emergência do PA e de determinadas variedades reestruturadas do PB, caracterizadas por uma não-realização de morfologia flexional, que pode ir de grau mais leve a grau mais acentuado. Dado que a manutenção geracional da não-realização poderá estar directamente conotada à não-reinstanciação de determinados parâmetros, que determina fossilização, importará também fazer referência ao possível papel desta no processo de reestruturação do português falado em Almocharife.

## 2.1. Os pressupostos teórico-metodológicos sobre aquisições de L1, bilingue e L2

As variações da fase precoce da aquisição de uma L1 em situação que não configura transmissão linguística irregular são geralmente regularizadas bastante cedo. Contrariamente, tanto a L2, adquirida sobretudo após o período crítico, como a L1 adquirida em contexto de transmissão irregular são sistemas linguísticos que mantêm as variações por mais tempo, podendo algumas fossilizar-se. No PA convivem diferentes gramáticas geracionais e até individuais, com determinadas realizações do mesmo falante revelando apego à reduzida morfologia do santomense (exemplo [152a]), enquanto outras apresentam características da elaborada morfologia do PE (exemplo [152b]):

[152] PA:

- a. *Foi o tempo do Frente, os homem de Frente.*<sup>47</sup> [MAURIH2]  
“Foi no tempo da Frente, dos homens da Frente”
- b. *resolvê poblema dos homens que tão cá* [MAURIH2]  
“resolver o problema dos homens que estão cá”

A variação na marcação de número tem sido conotada quer a mudanças internas ao próprio português, no caso, a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a), quer ao sistema de pluralização dos substratos africanos (Guy, 1981a:301-302) do grupo níger-congo atlântico (Baxter, 2009:292), que faz incidir o papel fulcral PL no classificador pré-nominal. Contudo, uma explicação da variação linguística com base exclusiva em comparações entre a L1 e a L2 é bastante

reduzida (Baxter, 2004; Godinho, 2005), já que existem paralelismos entre a aquisição da L1 e da L2, com os aprendentes a atravessarem diversos estádios na aplicação das regras da língua em elaboração, até atingirem um patamar final de aquisição (Cerqueira, 1994; Lléo, 1997, 1998, 2001; López-Ornat, 1997; Faria, Freitas & Miguel, 2001; Name & Corrêa, 2003; Plag, 2008a, 2008b).

Na ASL, o aprendente elabora uma interlíngua (Selinker, 1972:216), marcada por etapas que se vão aproximando da LA (*approximative system*) (Nemser, 1971:119). Este sistema apresenta interferências da L1<sup>48</sup> a nível morfossintático, fonológico, lexical e semântico (Thomason & Kaufman, 1988:73-74), originando, por vezes, estruturas com marcas da L1 e da LA, e que demonstram que o falante detém conhecimentos da segunda, mas resiste às suas regras gramaticais. Contudo, o facto de detentores de diferentes L1's revelarem idênticos estágios de desenvolvimento na aquisição de uma L2 levanta a hipótese de ambas as aquisições serem também determinadas por mecanismos cognitivos internos e independentes do substrato (Romaine, 2003:411; Godinho, 2005:55). Para Plag (2008a:115), o processo de interlíngua ocorre unicamente a nível individual, enquanto a emergência de uma nova língua em contexto de contacto massivo pressupõe a aquisição e uso colectivo de uma língua-E, por um lado, e de uma língua-I, por outro lado. Daí que a noção de interlíngua e da sua sistematicidade constituam, hoje em dia, pontos fulcrais dos estudos em ASL (p.e. Bayley & Preston, 1996; Siegel, 2004a, 2004b, 2006, 2008; Franseschina, 2005; Pienemann, 2000, 2005a, 2005b; Bruhn de Garavito, & Atoche, 2006; Sprouse, 2006).

Alguns autores têm apontado paralelismos entre a aquisição da L1 e da L2, que afectam quer as crianças quer os adultos detentores de diferentes L1's (Bruhn de Garavito & White, 2002; Montrul, 2004; White *et alii*, 2004). As etapas de variação no processo de aquisição de línguas são influenciadas por fenómenos vários, que vão desde a acção dos mecanismos cognitivos internos relacionados com a L1 e a L2 (Montrul, 2004), até às regras não conectadas a esta (Pienemann, 1998, 2000, 2005a, 2005b; Pienemann & Håkansson, 2007; Plag, 2008a, 2008b, 2008c, 2008d), sem deixar ainda de passar pela própria competência linguística subjacente a cada falante. No caso concreto da ASL massiva, são determinantes, para a maior ou menor variação registada na interlíngua, os contextos sociohistóricos em que esta se dá (Winford, 2003a:355),

fundamentais para a motivação dos falantes em relação à aprendizagem da LA. Por conseguinte, uma correcta análise da variação e da sistematicidade da interlíngua e das línguas nativizadas terá que levar em conta pressupostos relacionados com as perspectivas generativista, sociolinguística e psicolinguística (Ellis, 1994; Godinho, 2005; Baxter, 2009; Silva, 2010).

### **2.1.1. Variação e sistematicidade na *interlíngua*: perspectiva generativista**

O estudo dos fenómenos da *interlíngua* devem distinguir entre o conhecimento da língua que o falante detém, realmente, e aquilo que ele produz em situação de comunicação (Yip, 1995:6, *apud* Godinho, 2005:56, nota 66). Em termos generativistas (Chomsky, 1965), a competência não será a variação detectada na *performance*, mas antes, o conhecimento interiorizado que o falante detém da L2. Desta forma, o patamar avançado de aprendizagem é a condição de excelência para os estudos em gramática generativa, não podendo a discussão em torno da sintaxe complexa demarcar-se das especificidades inerentes à intervenção da GU. Esta activa quer os princípios universais das línguas quer os parâmetros específicos de cada idioma, permitindo que o aprendente construa a gramática da LA com base em *input* restrito. Este aspecto determina que a competência subjacente à *performance* apresente paralelos entre L1 e L2, interessando apurar o conhecimento que o falante detém de um determinado sistema linguístico e não o modo como o utiliza (Lightbown & Spada, 2003 [1993]). Ainda assim, alguns generativistas (p.e. Bruhn de Garavito & White, 2002) não se têm limitado a comparar juízos de gramaticalidade entre falantes de L1's e L2's, uma vez que se debruçam também sobre dados da *performance* (Godinho, 2005:57).

Assim, e no que concerne concretamente às abordagens que procuram estabelecer uma relação entre a mudança linguística e a teoria da gramática generativa, é assumido que a gramática da criança é construída com base na conjugação estabelecida entre os dados auditivos (os DLP's) e os universais linguísticos biologicamente determinados para aquisição das línguas (a GU). A GU demarca o que pode variar entre línguas (os parâmetros) e aquilo que é constante (os princípios). Os primeiros constituem um conjunto de opções definido pela própria GU, que estabelecem as diferenças entre dois sistemas gramaticais. Por seu lado, os segundos configuram os mecanismos mentais

predeterminados dos falantes que, ao interagirem com os parâmetros, permitem analisar as estruturas que poderão estar a violar um dado sistema linguístico.

Apesar de os princípios serem imutáveis, os DLP's podem ser afectados por mudanças culturais, sociais e pragmáticas, o que os leva a variarem de geração em geração. Deste modo, o *input* linguístico que determina a aquisição da linguagem conduzirá, inevitavelmente, a mudanças, vistas como o resultado de um processo abduativo (lei + resultado => caso) (Lightfoot, 1979; Lucchesi & Ribeiro, 2009).

Trabalhos diversos sobre as diferenças de parâmetros entre línguas têm também tentado descrever o parâmetro e suas consequências, bem como as implicações tipológicas e experiências decorrentes das distintas selecções de valores paramétricos. Especificamente no que diz respeito às diferenças paramétricas registadas em diferentes estágios de uma mesma língua, é fundamental explicar não apenas qual o parâmetro que regista diferença e os efeitos originados por esta mas também o porquê de um novo valor se sobrepor geracionalmente a outro (Adams, 1987; Vance, 1989; Roberts, 2007).

Em alguns estudos sobre variação e mudança linguística tem sido assumido que a mudança resulta, essencialmente, da competição entre formas conservadoras e outras que se vão inovando ao serem influenciadas por factores externos de diversos tipos. Portanto, trata-se de um fenómeno que evolui difusa e gradualmente no tempo, em virtude de as gramáticas serem geracionalmente construídas, uma vez que as crianças não se podem abstrair das mudanças que afectam os DLP's a que estão expostas (Lightfoot, 1979:391).

No entanto, tem também sido assumido que a mudança é algo que acontece de modo abrupto, envolvendo apenas a transição de uma geração para outra (Kroch, 1989b). Este aspecto resultará do facto de nenhum sistema linguístico se poder encontrar em estágio intermediário entre dois valores paramétricos. Com efeito, e de acordo com os princípios da Lei do Meio Excluído (*Law of the Excluded Middle* – Roberts, 2007:296), os valores paramétricos respeitantes a entidades como verbos, nomes, tempos ou fonemas têm de ser tratados como entidades formais de natureza discreta, visto os seus traços possuírem apenas um de dois valores possíveis.

Estes pressupostos configuram uma visão generativista sobre parâmetros, aquisição e mudança, sendo ainda coadjuvados por mais dois princípios fundamentais: (i) a

mudança catastrófica ocorrida em determinado momento histórico altera, irreversivelmente, determinados valores paramétricos; (ii) a mudança é um fenómeno que diz respeito apenas à língua-I, não sendo determinada por questões conectadas à língua-E (p.e. o ambiente social, cultural ou histórico em que o aprendente se encontra inserido). À primeira vista, parece que os três pressupostos não encontram pontos de conciliação com a visão da Sociolinguística Variacionista, mas alguns autores entendem que tal contradição é meramente aparente (Kroch, 2001; Roberts, 2007), conquanto se conciliem *“as ideias de sistemas algorítmicos homogêneos, discretos, que são propriedades dos indivíduos, com a mudança gradual da linguagem num dado momento histórico em uma dada comunidade de fala, formalizada analiticamente, pelos sociolinguistas, por algoritmos heterogêneos, graduais e probabilísticos”* (Lucchesi & Ribeiro, 2009:133).

Segundo Roberts (2007), os efeitos das mudanças discretas e instantâneas a nível estrutural encontram-se inseridas nos dados históricos, pelo que é possível observá-las conciliando e analisando a frequência de realizações entre formas inovadoras e conservadoras com o trajecto temporal da mudança linguística. Por outras palavras: estudando-se, em determinado sistema linguístico, o efeito gradual exercido por factores independentes uns dos outros, ou seja, factores sociais (idade, escolaridade, variação idiolectal, etc.) e factores de natureza gramatical (difusão lexical, microparâmetros, macroparâmetros, etc.), obtém-se uma visão acerca do modo como ocorrem as mudanças no mesmo. No que concerne à difusão lexical, tem sido confirmado que mudanças fonológicas (Labov, 1994) e mudanças de traços formais do léxico substantivo (Lightfoot, 1991a) se difundem gradualmente, afectando o sistema lexical de determinados idiomas. Contudo, no segundo caso, os processos de reanálise podem afectar somente estruturas específicas, incidindo unicamente sobre determinados parâmetros do léxico funcional (mudanças paramétricas), e não sobre a totalidade da gramática (Borer, 1984). Nestas situações, e dado que as categorias funcionais são compostas por traços formais, estamos perante casos de mudança gradual, que se opõem aos casos de mudança abrupta, em que é afectado todo o léxico substantivo (Roberts, 1993, 2007).

Quando uma série de mudanças discretas atinge, gradualmente, o conjunto de traços de categorias funcionais de um determinado dialecto contemporâneo, é necessário identificar, a partir da noção de microparâmetros, quais as unidades mínimas da língua histórica que registaram variação paramétrica (Kayne, 2000; Ladgeway, 2000). Efectivamente, variações sincrónicas mínimas registadas em traços que alteram os seus valores no conjunto poderão apontar para uma mudança diacrónica gradual nas comunidades (Roberts, 2007:300-305). Contudo, as mudanças graduais assentes em opções de natureza sociolinguística, que levam determinada forma a tornar-se predominante, poderão não constituir mudança paramétrica. Este aspecto torna-se elucidativo em trabalhos com recurso ao modelo minimalista da gramática generativa, os quais demonstram que a opcionalidade formal numa mesma gramática pode constituir apenas uma possibilidade técnica, sem consequências semântico-pragmáticas (Biberauer & Roberts, 2005; Biberauer & Richards, 2006; Roberts, 2007).

Face à complexidade que a variação nos usos de uma determinada comunidade pode encerrar, Kroch (1989b:200) alertou para a hipótese do Efeito de Taxa Constante (*Constant Rate Effect*), em que a mudança gradual é consequência de gramáticas em competição, que põem em contraste formas conservadoras e inovadoras vernaculares, levando uma delas a impor-se diacronicamente em determinados contextos linguísticos, por melhor se adequar a eles. Tal imposição não deve, então, “*ser considerada um fato da gramática em si mesma, e sim um fato sobre a forma como os indivíduos com competência em mais de uma gramática acessam as opções disponíveis e sobre como os fatores sociais atuam nessas escolhas*” (Lucchesi & Ribeiro, 2009:136).

Todavia, as abordagens sobre mudanças paramétricas e mudanças no léxico substantivo não fornecem resposta para se analisar a forma como a mudança afecta massivamente determinadas comunidades, já que a observam unicamente a nível do sujeito. Ainda assim, a proposta generativa de que a mudança linguística é direccionada pela aquisição da linguagem permite entender o problema da implementação (*actuation problem*) (Weinreich, Labov & Herzog. 2006 [1968]), isto é, observar os estágios em que as duas formas de uma mesma língua têm a sua maior intervenção. Por seu lado, a noção de gramáticas em competição e do Efeito de Taxa Constante (Kroch, 1989b) ajuda a explicar questões relacionadas com a problemática da transição (*transition*



*problem*) (Weinreich, Labov & Herzog. 2006 [1968]), ou seja, como a nova gramática se difunde entre os indivíduos de uma comunidade de fala bilingue. De facto, forças externas actuando sobre as gramáticas em competição produzem variação a nível social, atestada na diglossia sintáctica (p.e. nível formal e nível coloquial), enquanto o bilinguismo origina registos marcados por alternância de códigos (*code-switching*) (ponto 2.3.1 do presente trabalho) ou mistura de códigos (*code-mixing*).

Note-se que a competição entre gramáticas, por si só, não fornece resposta cabal para se entender a mudança linguística na sua totalidade, já que não explica as questões conectadas ao início da mudança (*constraints problems*) nem à problemática da sua conclusão (*actuation problem*), vistas na perspectiva da Teoria da Variação (Weinreich, Labov & Herzog. 2006 [1968]). Assim, o conceito de diferenciação ordenada (*orderly differentiation*), fundamental para se entender a questão da mudança linguística, (Weinreich, Labov & Herzog. 2006 [1968]:151), sobretudo a problemática da transição (*transition problem*) que determina mudanças sintácticas, é explicado por Roberts (2007:320) com base em mudanças a nível paramétrico e que ocorrem em três estágios distintos: no primeiro acontece a remarcação do valor de um determinado parâmetro, que gera nova opção formal e muda a proporção de indivíduos bilingues na comunidade; no segundo, o valor paramétrico difunde-se na comunidade, colocando gramáticas em competição, tornando opaca a opcionalidade formal respeitante a um único sistema e alterando, gradualmente, a proporção de falantes em que as duas formas coexistem; por fim, a opção formal conservadora extingue-se no terceiro estágio.

Para explicarem a difusão gradual da mudança, alguns generativistas entendem-na como uma consequência de “erros” no processo de aquisição da L1 ou L2 e recorrem à noção de parâmetros discretos, procurando, com tal, definir onde ocorre a variação nos sistemas linguísticos em observação (Kroch, 2004). Se os referidos “erros” forem compatíveis com a GU, dá-se a mudança no valor do parâmetro e, seguidamente, a alteração na frequência de uso da forma, que a levam a propagar-se na comunidade. Visão contrária têm outros autores, que admitem um processo inverso na alteração da frequência de uso de uma certa forma linguística (Adams, 1987; Vance, 1989). Assim, esta será determinada por estratégias discursivas que alteram frequências de uso e tornam opaco um determinado parâmetro, conduzido à alteração no seu valor.

Seja de uma forma seja de outra, os estudos necessitam de recorrer à quantificação para poderem analisar as tendências gerais do fenómeno sintáctico em observação. Por outro lado, como as línguas variam parametricamente, o recurso à sintaxe comparada fornece indicações precisas sobre as diferenças registadas no desenvolvimento histórico de um determinado idioma, que podem ser vistas como o resultado da remarcação diacrónica de parâmetros.

### **2.1.2. Variação e sistematicidade na *interlíngua*: perspectiva psicolinguística**

A perspectiva psicolinguística debruça-se sobre os mecanismos internos que conduzem à variação detectável na *performance* (Godinho, 2005:57). Para justificar a variação, Selinker (1972:216) propõe a existência de uma estrutura psico-cerebral que é accionada quando o aprendiz começa a ASL, iniciando-se também a construção da *interlíngua*. Segundo esta perspectiva, a *interlíngua* e a L1 são sistemas unos e independentes, pelo que as produções de ambas representam o uso natural de uma língua (Adjemian, 1976:307). Assim, a realização linguística tem categorias distintas, visto assentar em diferentes fases de produção de uma língua (Levelt, 1989:147), que pode ser consciente, ou não. Para os psicolinguistas, a variação reside nas realizações de superfície das entradas marcadamente abstractas do léxico mental, ou seja, nos morfemas – os lemas – que encerram em si parte da informação semântica, sintáctica e até fonológica do item seleccionado para a produção (Biase & Kawaguchi, 2002; Håkansson, Pienemann & Sayehli, 2002; Pienemann, 2000, 2005a, 2005b; Plag, 2008a, 2008b). A realização de alguns dos morfemas é accionada de forma indirecta, visto deterem uma funcionalidade dependente do elemento directamente seleccionado, como acontece com as preposições que regem determinados verbos. Quanto à informação encerrada nos lemas, diz respeito aos três subsistemas da estrutura lexical (Ellis, 1994:130-131), listados por Godinho (2005:58):

- “ - *A estrutura lexical/conceptual que combina a informação semântica e pragmática disponível universalmente;*
- *A estrutura “predicado-argumentos”, que especifica as propriedades dos verbos em diferentes sub-categorias e a forma como são codificados*

*gramaticalmente os argumentos, isto é, quantos argumentos pode ter um verbo e qual o papel temático que cada um deles recebe;*  
*- Padrões de realização morfológica que explicam os mecanismos de superfície para ordem das palavras, caso, concordância, marcas de tempo e aspecto, etc.”.*

Estes subsistemas interagem, de modo progressivo, em etapas que se assemelham nas L1's e L2's (Myers-Scotton, 2002:23), indo desde as intenções de produção linguística a nível conceptual (tomada de decisões, por vezes inconsciente, e que levam em conta quer aspectos quer consequências sociais e psicolinguísticas), até à efectiva realização da mesma (Levelt, 1989:1-28, *apud* Godinho, 2005:58-59).

### **2.1.3. Variação e sistematicidade na *interlíngua*: perspectiva sociolinguística**

A perspectiva sociolinguística de análise da variação e da sistematicidade da *interlíngua* relaciona a variação linguística e a sociohistória dos seus falantes, ou seja, estabelece a relação entre a competência linguística e a *performance* (Dickerson, 1975; Beebe, 1980, 1988; Bayley, 1994, 1996; Preston, 1996a, 1996b; Chambers, Trudgill & Schilling-Estes, 2003 [2002]). Embora a nossa análise incida sobre dados da produção real, obtidos com recurso a entrevistas, é de considerar que os informantes poderão, de modo mais ou menos espontâneo, ter recorrido também à introspecção. Assim, e para termos em conta aqueles que serão os padrões colectivos de comportamento linguístico da comunidade de Almojarife, entendemos que os actos de fala dos seus membros, tal como sucede com os de qualquer outra comunidade, “*são regidos por factores relativos:*

- (i) *ao conhecimento intuitivo, que faz do indivíduo um membro da comunidade de fala (ou seja, um falante);*
- (ii) *a certas características sociais desse indivíduo (tais como sexo, idade, nível de escolaridade, etc.);*
- (iii) *aos juízos de valor desse indivíduo sobre as escolhas disponíveis no repertório linguístico da comunidade;*
- (iv) *à dimensão ideológica do ato de interação verbal, que remete às relações de poder, estratégias de convencimento, nível de interesse na conversação, etc.;*
- (v) *às circunstâncias específicas nas quais se realiza o ato, relativamente ao conhecimento compartilhado entre os participantes, à intervenção eventual de participantes adventícios ao ato em questão, etc.”* (Lucchesi & Ribeiro, 2009:129).

Para se entender a CPL-var no SN do PA, é fundamental estabelecer, com recurso a métodos quantitativos, a diferença entre realizações pontuais e sistemáticas. Como tal, integraremos no nosso estudo factores que possam reflectir, probabilisticamente, com recurso a variáveis que arrolaremos, as tendências de uso de formas linguísticas e sua proporção de aplicação nos distintos contextos linguísticos. Esta metodologia permitirá elaborar predições comunitárias, avaliando até que ponto as frequências de uso da CPL-var no SN estão relacionadas com uma situação de mudança linguística em curso.

Contudo, o modelo apenas dá conta dos aspectos relacionados com a língua-E, não facultando uma compreensão sobre o conhecimento internalizado na mente dos seus membros, a nível individual, ou seja, da sua língua-I. Nestas circunstâncias, se não observarmos também os pressupostos teóricos acerca da forma como o conhecimento linguístico internalizado determina o comportamento linguístico dos membros da comunidade de Almocharife, não nos é possível entender, abrangentemente, os aspectos que determinam a CPL-var no interior SN do PA nem quais os elementos estruturais que são efectivamente afectados por esta. Portanto, o recurso à teoria da gramática, visto na perspectiva da problemática das restrições (*constraint problems*) à luz da Teoria da Variação (Weinreich, Labov & Herzog, 2006 [1968]), será fundamental para estabelecermos os limites da variação na estrutura do SN do PA, observando não só onde estará situado o ponto de partida da implementação de um possível processo de mudança em curso (cf. Tarallo, 1991) mas definindo também os potenciais processos que determinam esta.

Paralelamente, a estrutura gramatical, ao intervir no processo de difusão da mudança, permite entender a problemática do encaixamento (*embedding problem*) (Weinreich, Labov & Herzog, 2006 [1968]), uma vez que determina quais os contextos específicos que favorecem certa variante e se a sua frequência de uso tem implicação directa no desenvolvimento diacrónico de uma outra variável linguística. Do exposto, verifica-se que a resposta para os fenómenos da variação e da sistematicidade não pode assentar apenas num dos modelos referidos (Zuelenger, 1989:66, *apud* Godinho, 2005:60). Nesta conformidade, Ellis (1994:120) propõe uma conjugação dos quadros teóricos sociolinguístico e psicolinguístico, para dar conta dos referidos fenómenos. Porém, e a exemplo da metodologia aplicada em determinados estudos sobre aquisição de L2/LE

(p.e. Godinho, 2005; Silva, 2010), estas perspectivas podem ser complementadas com outras abordagens, levando-se em consideração os princípios respeitantes à GU, já que este modelo, conforme mencionado no ponto 2.1.1, se revela robusto para uma análise sobre a morfologia concordante, sobretudo quando se trata de estudar padrões colectivos de comportamento linguístico afectados por mudanças linguísticas decorrentes de processos históricos específicos, como é o caso da transmissão linguística irregular.

Apesar de tudo, importa considerar que a Gramática Generativa formaliza as suas constatações, em termos de língua-I, recorrendo a unidades invariáveis e regras categóricas, ao passo que a Sociolinguística Variacionista faz uso de unidades e regras variáveis para análise dos padrões colectivos de comportamento linguístico. Além disso, ao procurar explicar o modo como a língua funciona e entendendo a língua-I como o resultado do desenvolvimento da faculdade inata da linguagem, o generativismo situa as suas observações sobre os fenómenos linguísticos num plano estritamente psico-biológico. Já a Sociolinguística Variacionista, ao lançar a sua atenção sobre a relação entre variação e mudança e operar com esquemas probabilísticos de análise, coloca os seus estudos sobre tais fenómenos num plano sociohistórico. Nesta perspectiva, o nosso modelo de análise será heterogéneo, observando o funcionamento da gramática em concomitância com factores sociais, culturais ou ideológicos envolvidos nos processos históricos de mudança linguística que afectaram, ou estarão a afectar, o PA. Assim, será também possível não só alargar o conjunto de respostas para entendimento das questões inerentes à CPL-var no SN do PA mas estabelecer igualmente quais os fenómenos que concorrem para a implementação (*actuation problem* – Weinreich, Labov & Herzog, 2006 [1968]) de determinadas variantes em detrimento de outras (cf. Lucchesi & Ribeiro, 2009:130) na comunidade de Almojarife. E, ao conjugarmos aspectos da teoria da gramática e análise sociohistórica da mudança linguística, estaremos, em última instância, a procurar contribuir para diminuir a distância teórico-metodológica que separa, actualmente, estas duas perspectivas de estudo.

#### **2.1.4. Variação e sistematicidade na *interlíngua*: aplicação da perspectiva heterogénea ao estudo do PA**

Como a fala de Almojarife encerra variação, adoptámos a metodologia da perspectiva sociolinguística para recolha de dados em contextos heterogéneos (Labov,

1972a). O tratamento dos dados seguiu ainda as propostas labovianas para análise quantitativa na variação linguística, também aplicadas por Dickerson (1974) e Dickerson (1976), para detecção dos contextos linguísticos em que acontece a variação, em primeira mão, e por Preston (1996b), para apurar a direcção em que acontece a mudança linguística, em segunda etapa (cf. Godinho, 2005:61).

A investigação de dados por Labov (1966, 1972a), na perspectiva da sociolinguística quantitativa, permitiu identificar quais as influências internas (p.e. a saliência fónica ou o contexto fonológico em que determinado som se realiza) e externas à língua (p.e. a classe social, a idade, o nível de escolaridade) que actuam sobre as variáveis fonológicas. Seguindo esta metodologia, poderemos, por exemplo, confirmar se a inserção ou apagamento do –s morféxico, marca de PL, nos itens do SN do PA terão subjacentes motivações fonológicas do mesmo tipo das que foram observadas por Labov (cf. Godinho, 2005:61).

A variação nos dados de um determinado falante pode ser livre ou sistemática, com a primeira reflectindo formas que resultam de motivações em competição conotadas aos sistemas de regras da L1 e da L2 (Ellis, 1994:144; Romaine, 2003:411, *apud* Godinho, 2005:62). De acordo com o primeiro autor, a variação livre ocorre por curto período de tempo, restringindo-se aos estados precoces do desenvolvimento linguístico. O falante começa por utilizar um traço de forma não sistemática, mas à medida que a interlíngua se desenvolve, a forma é integrada nesta e passa a ser aplicada em todos os contextos. Portanto, a variação livre vai desaparecendo, à medida que os falantes desenvolvem a sua interlíngua em direcção à LA (Preston, 1966:250, *apud* Godinho, 2005:62).

Como temos vindo a evidenciar, subjazendo à transição de variação livre para variação constante actuam factores situacionais, linguísticos e psicolinguísticos, o que determina que a observação das questões relacionadas com a variação livre e a variação condicionada não podem ser entendidas sem se lançar mão de um modelo teórico adequadamente formulado para tratar a variação e a mudança. O fenómeno da mudança linguística, sociolinguisticamente observado (cf. Weinreich, Labov & Herzog, 2006 [1968]), é também uma problemática da Gramática Generativa, apesar de esta o entender como um processo de aquisição da linguagem com restrições impostas pela GU. Assim, a questão da implementação (*actuation problem*) (Weinreich, Labov & Herzog, 2006

[1968]), que tenta determinar o porquê de acontecer mudança linguística, “*encontra uma resposta no programa gerativista em termos de cadeia de aquisição*”:

- (i) *a aquisição resulta de um processo seletivo e abduativo, de interação entre estruturas cognitivas e tipos de experiência selecionados;*
- (ii) *os dados da experiência relevantes para a aquisição são constituídos de elementos robustos e estruturalmente simples;*
- (iii) *a criança é um aprendiz grau-0 – sensível somente a dados de sentenças raiz e do complementador da sentença encaixada;*
- (iv) *a morfologia desempenha papel importante na definição do valor de um parâmetro;*
- (v) *parâmetros são marcados nos estágios iniciais da aquisição”* (Lucchesi & Ribeiro, 2009:140).

Nesta perspectiva, a reanálise abduativa dos DLP’s, resultante de processos morfofonológicos, ambiguidade estrutural, DLP’s inovadores, factores extralinguísticos ou contacto entre línguas, potencia a mudança sintáctica (Lightfoot, 1991a; Kroch, 2004; Roberts, 2007).

A problemática do encaixamento (*embedding problem*) (Weinreich, Labov & Herzog, 2006 [1968]) é observada sob dois ângulos em termos sociolinguísticos: (i) o modo como se encaixa na estrutura linguística; (ii) a forma como se encaixa na estrutura social. Relativamente à primeira visão, estudos de sintaxe diacrónica têm apresentado três possibilidades, estando a primeira conectada à noção de valor paramétrico (positivo ou negativo) e sua correlação com as hipóteses estruturais (Lightfoot, 1979; 1991a; Rizzi, 1982; Roberts, 2007). Nesta perspectiva, a mudança linguística não afectará a totalidade de um sistema linguístico (Weinreich, Labov & Herzog, 2006 [1968]:185), visto que as propriedades que podem variar entre línguas não são autónomas, antes se organizam em grupos estáveis.

No caso do parâmetro do sujeito nulo, por exemplo, Rizzi (1982) especifica propriedades que diferenciam o PE, o italiano e o espanhol de outras línguas, como o inglês ou o francês. As conclusões demonstram que, uma vez perdida a propriedade correspondente ao valor do parâmetro, as outras propriedades relacionadas com ele se apagam também.

Lightfoot (1979, 1991a), por seu lado, analisa a evolução de verbos plenos do inglês antigo, para verbos auxiliares modais no inglês médio. O autor conclui que a reanálise terá sido impulsionada por um conjunto de mudanças gramaticais distintas

Quanto a Roberts (2007), debruça-se sobre os aspectos que levaram o inglês a divergir, entre 1100 e 1700, quer das línguas românicas quer de outras línguas germânicas. A mudança de uma língua de um tipo para outra de outro tipo ficará a dever-se a uma mudança no seu sistema funcional, cujo efeito dominó activará mudanças paramétricas em cascata. Efectivamente, tais mudanças, ao afectarem os traços de um determinado núcleo funcional, têm como consequência uma extensa reorganização nos sistemas conotados ao mesmo.

No que concerne à problemática do encaixamento na estrutura social (*embedding problem*) (Weinreich, Labov & Herzog, 2006 [1968]), a mesma pode levar a mudanças linguísticas na fala dos adultos, que alteram os DLP's para a aquisição da L1 de uma geração para a outra e afectam algoritmos generativos. Nesta conformidade, a mudança é activada pela alteração nas frequências de uso de uma determinada forma, como acontece na variação que afecta a concordância entre os elementos do SN no PtgL2 adquirido pelos almoxarifanos, e que alteram os dados robustos para aquisição do PA como língua materna pelas gerações precedentes. Este tipo de alterações sociohistóricas não é ocasionado pela gramática em si, mas sim pela mudança de hábitos linguísticos associada a mudanças comportamentais a nível de relações sociais, culturais ou ideológicas. Contudo, poderão ter impacto na gramática ao determinarem também mudanças paramétricas, como o comprova o caso da substituição do pronome pessoal canónico *vós* por *vocês* no PE, que adoptou a forma verbal da 3ª pessoa do plural, reduzindo o paradigma flexional do verbo (p.e. *vós ides* > *vocês vão/eles vão*).

Kroch (2001, 2004), ao propor que se observem os efeitos gerados pelas gramáticas em competição numa determinada comunidade de fala, “*permite conciliar a concepção de uma mudança gramatical abrupta com uma gradualidade que resulta da distribuição social dos usos linguísticos*” (Lucchesi & Ribeiro, 2009:143-144). Efectivamente, a possibilidade de os falantes almoxarifanos poderem optar por estruturas disponíveis nas gramáticas em concorrência na sua comunidade, sobretudo por esta ter na sua génese



uma situação de contacto entre línguas, foi fundamental para a distribuição dos seus usos linguísticos.

Face a estes aspectos, resulta que, para além da adopção de um apropriado modelo teórico, a observação da variação e mudança nos sistemas linguísticos não é possível se não se recorrer também a um dispositivo estatístico adequado para avaliação da mesma. Desta forma, a variação sistemática pode ser analisada recorrendo a variáveis dependentes e independentes, com as primeiras representando o conjunto das variantes linguísticas (os itens em variação) e as segundas dizendo respeito aos prováveis grupos de factores que condicionam as variáveis dependentes. No nosso caso, a variável dependente em análise representa a ausência ou inserção de marca morfológica de número PL nos itens do SN. Por seu lado, as variáveis independentes identificam o conjunto de factores linguísticos e extralinguísticos que beneficiam ou inibem o uso dessa mesma marca. Sujeitando à análise quantitativa os factores prováveis de exercerem influência sobre a variação, se se verificar que os mesmos não têm efeito sobre as variáveis dependentes, então não está configurada uma situação de variação sistemática (cf .Preston, 1996:250, *apud* Godinho, 2005:62).

A CPL-var no SN do PA configura variação sistemática, visto ser um fenómeno abrangente em termos de estratos etários ou de sexo. Mas, esta sistematicidade pode também ser determinada por outros contextos, que não os meramente sociais, como se pode constatar em exemplos do nosso *corpus*, mais concretamente os que dizem respeito à variação que afecta a inserção de PL no nome “trabalhadores”:

[153] PA:  
*não tem possibilidade de pagar os trabalhador antão Banco Mundial*  
[ABILH2]

[154] PA:  
*não tem possibilidade de pagar os trabalhadores principalmente o Stado*  
[ABILH2]

[155] PA:  
*os trabalhadô stão co sacrificá no mato pa dá produção* [CELESH3]

[156] PA:  
*cinquenta metros quadrado daquela terra enquadrá os trabalhadores.*[CASTEH1]

[157] PA:  
*para recolher alguns trabalhadores, os empregados pra levar*  
[CASTEH1]

Em contextos linguísticos idênticos, os falantes aplicam marcação PL diferente, pelo que a explicação para esta terá que ser fornecida em função de motivações sociais, como os anos de escolaridade, e/ou psicolinguísticas, como nos casos em que o falante monitoriza o *output*, após ter prestado atenção às realizações linguísticas do seu interlocutor (Wolfson, 1976:203). Portanto, subjazendo à variação estarão a actuar factores de natureza distinta, isoladamente ou em concomitância (Ellis, 1994:120), pelo que a sua observação terá que ser efectuada de modo heterogéneo. Deste modo, a codificação dos nossos dados, que adoptou a perspectiva variacionista amplamente utilizada em estudos brasileiros, permitiu que constatássemos também coincidências com a matriz proposta por Ellis (1994). Observar a variação numa perspectiva interna à própria língua significa, pois, levar em conta fenómenos que são idênticos em aquisição de L1 e L2 e que se podem conotar às teorias sobre a marcação e aos universais linguísticos. Será o caso, por exemplo, de ocorrer marcação apenas no elemento que é entendido como semanticamente relevante (p.e. *os home branco daqui ficaram* [CELESH3]), um fenómeno que terá respaldo nas línguas africanas (cf. Guy, 1981a:301-302) do grupo níger-congo atlântico (Baxter, 2009:292). Ainda a propósito deste tipo de marcação, Kiparsky (1972:195) advogou a hipótese funcionalista de que acontece uma tendência para os falantes reterem a informação relevante na estrutura de superfície (Condições de Distintividade), evitando o apagamento do morfema de PL apenas nas situações em que ele acarreta a informação semântica da pluralidade (Young, 1993:78).

Outras ocorrências em que se regista variação sistemática poderão estar ligadas às tendências universais detectadas em fonética, como acontece com a não-realização de consoantes finais (Romaine, 2003:413, *apud* Godinho, 2005:63), caso do apagamento do /r/ em final de palavra, um fenómeno vulgar no PB e também no PA (exemplo [155]). Lembremos, a este propósito, que determinadas línguas do grupo níger-congo atlântico, como o kimbundu, para além de não registarem a ocorrência da consoante /r/, apresentam, regra geral, uma vogal em final de palavra. Por outro lado, estes dialectos apenas apresentam vogais foneticamente salientes, não realizando as vogais /a/ [- alt], /e/ [+ alt] e /o/ [+ alt] (Mingas, 2000:36). Fenómenos deste tipo poderão ter tido

influência directa na realização de vogais foneticamente salientes e estruturas do tipo dissilábico CVCV registadas no santomense, apresentando estas uma tendência para a simplificação, quando comparadas com as do português (Ferraz, 1979:26). Assim sendo, tanto o falante de uma língua africana do grupo níger-congo atlântico em situação de aquisição do PtgL2 como a criança em contexto de aquisição do PtgL1 nativizado poderão revelar tendência para perceberem as formas mais salientes, acabando, no caso da pluralização, por inserir marcas nos itens de acordo com o advogado no Princípio da Percepção da Saliência (*Principle of Perceptual Saliency*), que postula que quanto mais distantes forem as formas fonéticas, maior é a probabilidade de elas surgirem marcadas (Wolfram, 1989:189).

De acordo com as Condições de Distintividade, é então de prever que, no caso do PA, a inserção de /s/, marca de número PL, não venha a ocorrer em nomes que sejam antecidos por elementos que já contenham a informação de PL, seja ela semântica (p.e. numerais) seja ela estrutural; em contrapartida, se a informação de PL ainda não tiver ocorrido, isto é, se o nome surgir como primeiro elemento do SN, então é provável que ele seja pluralizado (cf. Godinho, 2005:64). Paralelamente, também é de esperar que os elementos que apresentem maior saliência fónica na oposição singular/plural (p.e. *ovo* > *ovos* ou *patrão* > *patrões*) sejam mais marcados do que os itens em que tal saliência não seja tão óbvia (p.e. *coisa* > *coisas* ou *metro* > *metros*), como sucede quer nas variedades brasileiras (Scherre, 1988; Lopes, 2001; P. Andrade, 2003) quer nas variedades africanas de português (Baxter, 2004, 2009; Figueiredo, 2008, 2009b; Jon-And, 2008, 2009). Contudo, este tipo de marcação tem também sido conotado aos princípios universais (Romaine, 2003:413). Sobre a possibilidade de determinada estrutura da L2 representar ou não uma transferência da L1, voltaremos a pronunciar-nos no ponto 2.4.2.2 do presente trabalho, nomeadamente quando abordarmos as questões concernentes à Teoria da “Processabilidade” (Pienemann, 1998, 2005b).

A observação diacrónica da variação e da sistematicidade na fala da comunidade, verificadas através do chamado tempo aparente, ajuda a detectar possíveis mudanças em curso na mesma. Segundo Ellis (1994:144, *apud* Godinho, 2005:64), o aprendente de uma L2 começa por utilizar, de modo categórico, o mesmo traço para todos os contextos. É o que sucede, por exemplo, com o informante que produziu o exemplo

[155], que nunca pluraliza o nome “trabalhador” no *corpus* que possuímos, como pudemos constatar. Contudo, na geração seguinte vamos encontrar formas marcadas e não-marcadas (exemplos [153] e [154]), o que denota que a variação se torna livre, à medida que os membros mais jovens e mais escolarizados da comunidade vão aprendendo novas formas do mesmo nome. O efeito exercido por determinados factores sociolinguísticos leva ainda a que, posteriormente, a terceira geração evidencie uma maior aquisição da regra, passando a controlar a marcação de modo consciente, sistemático e correcto, conforme o demonstram os exemplos [156] e [157], produzidos pelo informante [CASTEH1] (8º ano de escolaridade, cf. Tabela 8.1), para o qual não foram detectadas formas do nome “trabalhador” sem marca PL, quando esta é exigida.

Trabalhos levados a cabo por Dickerson (1975, *apud* Godinho, 2005:64) ou Plag (2008a, 2008b) demonstram que a ASL é um processo de aprendizagem contínua, apresentando etapas flutuantes e transitórias, mas caracterizadas por sistematicidade. Factores linguísticos e extralinguísticos actuam, desenvolvendo a gramática do falante em direcção ao sistema de concordância da LA (Romaine, 2003:416), de acordo com o observável nos exemplos [153] e [154]. Portanto, quando há consciência da regra de concordância, entram em jogo fenómenos psicolinguísticos, podendo o falante monitorizar o uso das formas. No caso concreto dos informantes do PA, embora se tenha tentado afastar pressões do tipo social que pudessem interferir no seu registo espontâneo, nunca é de descartar a possibilidade de, no decurso das entrevistas, eles terem adoptado diferentes estratégias de comunicação, consoante a pragmática o exigisse. Porém, consideramos que, no cômputo geral, os discursos se adequam às exigências de uma correcta análise em sociolinguística quantitativa para o grau de sistematicidade registado na fala da comunidade (cf. Godinho, 2005:65).

## **2.2. Abordagens teóricas sobre aquisição de L1**

A aquisição de L1 pressupõe, em primeira instância, a aquisição da L1 por falantes monolingues. Contudo, o próprio conceito de monolingüismo puro é visto, à luz da teorização actual, como algo fictício, já que qualquer falante pode modelar e modular o seu registo, adaptando-o às exigências pragmáticas (Berg & Schade, 1992:418). Desta forma, o monolingue recorre a “falas” distintas, caso, por exemplo, esteja a comunicar

com uma criança ou se encontre no tribunal, perante um juiz. Noutras situações, pode mesmo misturar formas ou elaborar paráfrases, usando, para tanto, palavras ou registos linguísticos bastante distintos entre si. No presente trabalho, independentemente da polémica em torno do conceito de monolingüismo, interessa mais procurar entender como se processa não só a aquisição da L1 da criança que está igualmente a tentar comunicar com o adulto mas também o modo como essa aquisição se desenvolve até ao estágio em que ela, depois de atingir a idade adulta, vai fazer uso da recursividade, isto é, vai utilizar o *input* restrito e finito que recebeu para produzir *outputs* infinitos e conscientemente adaptados às diferentes situações de uso com que se vier a confrontar (Hornstein & Lightfoot, 1981:202).

### **2.2.1 Modelo generativista**

Ao elaborar a teoria do inatismo sobre a aquisição da linguagem, Chomsky (1965, 1972) defende que esta capacidade se desencadeia de forma involuntária em todas as crianças, concluindo-se em espaço de tempo breve,<sup>49</sup> mas envolvendo uma grande complexidade de operações. Apesar de existirem diferenças entre as línguas aprendidas, o processo de aquisição actua de forma semelhante, permitindo que todas as crianças, a partir de dados limitados, e com o suporte da GU, não só construam a gramática da língua mas activem também a capacidade da recursividade, isto é, a faculdade da linguagem. Assim sendo, a recursividade torna-se candidata a característica nuclear da linguagem humana, assentando os seus pressupostos quer na Teoria dos Princípios e Parâmetros (Chomsky & Lasnik, 1999 [1993]) quer na Teoria das Regências e Ligações (Chomsky, 1981b) quer ainda em preceitos avançados em posteriores trabalhos de Chomsky (1986a, 1986b). No caso concreto da GU, actua rapidamente, permitindo a aquisição da competência linguística em espaço de tempo muito curto. Tal só é possível porque ela está dotada de propriedades universais formais (referem-se às condições abstractas a que uma gramática obedece para combinar unidades linguísticas), substantivas (apontam para as categorias universalmente presentes nas línguas, isto é, de substância visível e necessárias para se analisar a língua, como acontece com os traços distintivos, nomes, verbos, sílabas singular/plural, masculino/feminino, etc.), implicativas (determinam a existência de uma correlação entre propriedades universais,

como acontece, por exemplo, nas línguas do tipo SVO, que necessitam obrigatoriamente de recorrer ao uso de preposições), ou pertencentes a determinada área da gramática (por exemplo, os universais fonológicos, semânticos, etc.). Como estas categorias universais se manifestam de modo diferenciado nas diversas línguas, a GU tem papel primordial na determinação (e explicação) das propriedades da gramática do ser humano, enquanto a Gramática Generativa tem como principal objectivo o estudo da natureza e propriedades da GU. Tendo em conta estes aspectos, percebe-se o motivo que leva os investigadores que trabalham dentro destes pressupostos a optarem, como objecto de estudo, por descrições linguísticas de gramáticas.

### **2.2.2. A teorização em torno da GU**

#### **2.2.2.1. O Modelo de Princípios e Parâmetros (*Parameter Settings*)**

Estando defendidas as propriedades universais da linguagem humana, Chomsky (1981a), avançou com a Teoria dos Princípios e Parâmetros para tentar descrever a diversidade das línguas. O modelo, composto por princípios extremamente gerais, visava, inicialmente, explicar as diferenças sintácticas entre as línguas, numa dimensão puramente sincrónica. Deste modo, defende-se que a gramática se organiza em módulos simples e autónomos, mas que abrangem diferentes domínios da linguagem, em virtude de manterem uma rede de interações com as outras componentes desta. Consequentemente, o modelo possibilita que se cruzem estudos puramente gramaticais sobre a estrutura das línguas quer com observações psicolinguísticas sobre aquisição e desenvolvimento da linguagem quer com estudos de gramática comparativa. Com base nos pressupostos deste modelo, a GU contém dois tipos de princípios:

- (i) os princípios rígidos e invariáveis (os universais substantivos e formais), como o Princípio de Projecção, o Princípio de Dependência Estrutural ou o princípio de que todas as línguas possuem SN sujeito e SV predicado (Chomsky, 1965: 27-30);
- (ii) os princípios abertos ou parâmetros, não especificados, e que se vão configurar para um valor num sistema binário (Chomsky, 1981a, 1986b).

Os parâmetros funcionam como “comutadores linguísticos”, cujo valor final é alcançado durante o processo de aquisição através da fixação numa das posições

possíveis, com base na informação colhida no meio linguístico ambiente. Desta forma, a gramática final da língua adquirida é caracterizada por dois aspectos: (i) a aprendizagem lexical da língua, sendo as propriedades fonológicas, sintáticas e semânticas guiadas pelo “dicionário mental”; (ii) os parâmetros da GU, com valores que lhes são atribuídos pelas características da língua.

Os parâmetros pertencem ao âmbito da variação, que tanto pode ser lexical, isto é, aprendida de forma explícita, como do domínio da aprendizagem não-explícita. Fornecem as opções possíveis à GU, bastante limitadas em virtude de serem finitas e binárias: o “parâmetro da direccionalidade” e o “parâmetro do sujeito nulo”. O “parâmetro da direccionalidade” é fixado muito cedo e permite que as crianças não errem ao juntar verbo e objecto na estrutura da produção. Como tal, é o parâmetro que permite estudar as configurações das línguas OV ou VO. Por seu lado, o “parâmetro do sujeito nulo” concede à criança a possibilidade de detectar se a língua que é falada à sua volta admite ou não construções com sujeito nulo.

A ligação entre todos os parâmetros concede uma “gramática nuclear” (*Core Grammar*) à criança, composta por uma complexa rede de conexões entre os princípios rígidos (universais) e os parâmetros, acabando, assim, por determinar as propriedades específicas de cada língua. Por conseguinte, a escolha dos parâmetros não é orientada conscientemente pela criança, correspondendo antes à aquisição da linguagem a partir dos enunciados que recebe da comunidade (os DLP’s), aquisição essa sujeita a um percurso de maturação da GU, que se inicia num estado em que os parâmetros não estão ainda fixados e termina num patamar com parâmetros fixados através de um dispositivo computacional. Em resumo: a GU, para além de definir as propriedades universais da linguagem humana, vai também fixar os parâmetros em determinados valores, originando a gramática particular de cada língua, isto é, a variação entre línguas.

Assim, na perspectiva do Modelo de Princípios e Parâmetros, a mudança sintáctica não é mais do que uma mudança no valor de parâmetro, Efectivamente, no estado inicial da aquisição (FL0),<sup>50</sup> o parâmetro é fixado a partir de dados simples e facilmente acessíveis, isto é, a partir dos DLP’s que a criança tem ao seu alcance para determinar a posição a conceder na referida fixação. No caso dos princípios rígidos, por exemplo aquele que determina que todas as línguas possuem SN sujeito e SV predicado, o

“parâmetro do sujeito nulo”, que vem por defeito com o valor positivo (a língua tem sujeito nulo), possui duas opções: (i) a língua tem sujeito nulo; (ii) a língua não tem sujeito nulo. Estas opções vão possibilitar que a criança escolha entre a realização fonética obrigatória do sujeito (valor *não-pro-drop*) ou realização fonética opcional do sujeito (valor *pro-drop*), fixando aquela que corresponde ao paradigma da sua língua específica. Nas situações que configuram princípios abertos, em que as crianças acedem a frases gramaticais pouco frequentes ou “estranhas”, a fixação pode não corresponder à dos parâmetros efectivos da GU. Ainda assim, em certas circunstâncias, a não ocorrência de determinadas estruturas nos DLP’s pode levar, durante o processo de maturação da GU, à fixação de parâmetros em determinada posição, sem necessidade de se recorrer a correcções ou instruções explícitas (Chomsky, 1981a).<sup>51</sup>

#### **2.2.2.1.2. Teoria das Regências e Ligações**

À luz da Teoria das Regências e Ligações (Chomsky, 1981a, 1986a), cada língua é um procedimento generativo, responsável pela produção de *descrições estruturais* (DE’s) compostas de propriedades semânticas e fonológicas. Deste modo, os parâmetros encontram-se localizados no léxico, no qual, por sua vez, estão alojadas as categoriais lexicais e funcionais das línguas. As primeiras representam os morfemas de conteúdo (categorias como nomes e verbos) e têm papel activo no quadro temático da produção linguística, atribuindo o valor semântico a esta. Como tal, configuram-se como núcleos da sua própria projecção lexical: o SN, o sintagma verbal (SV), etc. Estas categorias encerram em si as propriedades responsáveis pela definição das categoriais lexicais das palavras (p.e. [+V] e [+N]).

Por seu lado, as categoriais funcionais dizem respeito à morfologia flexional que contém a informação gramatical de pessoa, número, concordância, caso, negação, tempo ou aspecto. Portanto, para além de possuírem características semânticas e de flexão, as categoriais funcionais controlam também as formas morfofonológicas que correspondem à realização destas (p.e. *-s*, para marca de flexão em número PL, e *-a*, para marca de flexão em género feminino). Quanto à informação gramatical contida nas categoriais funcionais, é codificada de acordo com o tipo a que diz respeito, por exemplo: FLEX



(categoria funcional da flexão), CONC (concordância) ou T (projecção máxima de informações de tempo) (Godinho, 2005:99).

O dispositivo computacional que determina a fixação dos parâmetros extrai as categorias funcionais do léxico e transforma-as em representações sintácticas abstractas. Deste maneira, na hierarquização da estrutura frásica são as categorias funcionais que dominam as categorias lexicais (o SN é dominado pelo SDET, enquanto o SV é dominado por todas as categorias acima da categoria ASPECTO). Quando o dispositivo computacional extrai as categorias funcionais do léxico, acciona operações que as levam da área dos itens lexicais para o campo dos elementos funcionais (operações de “movimento”), a fim de entrarem em relação próxima com as outras categorias e permitirem que os seus traços interpretativos efectuem o mapeamento dos correspondentes traços não-interpretativos, originando o pareamento entre eles (operação de “pareamento”),<sup>52</sup> isto é, a concordância entre os elementos que constituem a frase (Montrul, 2004:8; Godinho, 2005:100).

Na Teoria das Regências e Ligações, os parâmetros são propriedades abstractas da GU e não estabelecem correspondências biunívocas com propriedades simples (as que são directamente observáveis) do sistema linguístico. Assim, a fixação do parâmetro é estabelecida pela interacção com outros princípios da GU e determina, por sua vez, feixes complexos de propriedades concretas em diversos pontos da gramática. Desta forma, o dispositivo computacional e as características formais das línguas constituem universais, o que não acontece com a edificação das frases, dada a sua dependência do léxico (Godinho, 2005:100).

Pode concluir-se que, em termos gerais, o pressuposto básico da teoria generativa é o de que a língua possui léxico e gramática (o dispositivo computacional). Ao segundo, que é inato, compete gerar as expressões ao nível da representação linguística (as DE's), sendo o léxico, por seu lado, responsável pela caracterização dos itens lexicais que surgem nas DE's (Chomsky, 1996 [1995]:59, *apud* Godinho, 2005:100). Como o léxico (as palavras) não é inato, tem que ser aprendido, necessitando a criança de adquiri-lo, tal como acontece com os parâmetros da sua L1 (a morfologia), já que ambos representam o conjunto das propriedades sintáctico-semânticas abstractas das categorias funcionais (Montrul, 2004:10, *apud* Godinho, 2005:100). Por sua vez, a aquisição de uma L2

significa a aquisição e fixação dos parâmetros da LA, já que os princípios são universais, não necessitando de ser aprendidos (Godinho, 2005:101).

Desta forma, a sintaxe trabalha mais com o léxico, levando a que o falante interprete determinados enunciados, por exemplo “*gostar de*” e “*amar*”, como sinónimos. Contudo, a sintaxe também estabelece relações com a componente interpretativa das palavras (a semântica), a prosódia (a fonologia) e a estrutura informacional (p.e. topicalizações, em que o primeiro discurso determina a escolha do enunciado posterior). Para dar conta da maneira como as diferentes componentes gramaticais interagem entre si, foi concebido o Modelo-T (modelo derivacional – Fig. 9), bastante rico em fontes possíveis de variação, e que tem o léxico como ponto de partida de toda a derivação, sendo as outras componente gramaticais alimentadas por ele (Chomsky & Lasnik, 1999 [1993]). Assim, as DE’s não serão mais do que representações em diferentes níveis da estrutura da língua (Estrutura-P, Estrutura-S, Forma Lógica (FL) e Forma Fonológica (PF)):

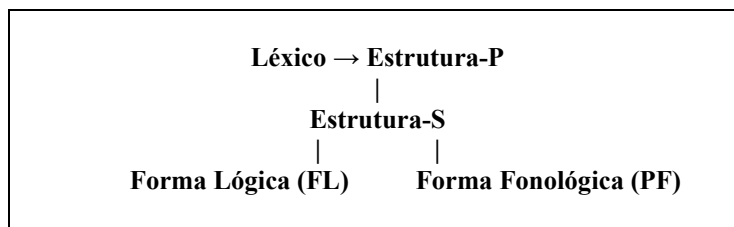


Fig. 9. Modelo de Princípios e Parâmetros (a partir de Chomsky, 1996 [1995:22]).

De acordo com o esquema, a Estrutura-P representa a projecção directa da estrutura argumental, isto é, a estrutura profunda, enquanto a Estrutura-S diz respeito ao nível sobre o qual incidem as regras de interpretação fonológica e de interpretação semântica (a *performane*), ou seja, a estrutura de superfície. O que se passa entre a Estrutura-P e a Estrutura-S é a aplicação das regras sintácticas (essencialmente as regras de movimento), sujeitas a variação paramétrica. Por seu lado, a estrutura de superfície, enquanto sistema de *performance*, alimenta duas vias: (i) a componente fonológica (o módulo articulatorio-perceptual); (ii) a componente lógica (o módulo conceptual-intencional). Como a prosódia pode interferir na interpretação do enunciado, a PF constrói os seus constituintes a partir da sintaxe, mas precisa também dos seus próprios constituintes para elaborar a hierarquia prosódica, uma vez que a sintaxe não lhe

concede todos os elementos de que necessita (Nespor & Vogel, 2007:32). Já em termos semânticos, ocorrem também movimentos não explícitos (subjacências) entre a Estrutura-S e o módulo FL, a fim de se evitar a ambiguidade (Huang, 1982:63; May, 1985:97; Nespor & Vogel, 2007:265). Assim, e apesar de o Modelo-T apresentar vários instrumentos de análise e dos domínios de variação paramétrica, esta não é satisfatoriamente explicada por ele. Paralelamente, também a relação da sintaxe com outros domínios gramaticais, como a morfologia ou a escolha da ordem das palavras com vista a condicionar a estrutura informacional, não se encontra plenamente contemplada no Modelo de Princípios e Parâmetros. Voltaremos a esta questão no ponto 2.6.2.1 do presente trabalho, no qual faremos alusão aos modelos desenvolvidos a partir do Modelo de Princípios e Parâmetros, nomeadamente o Programa Minimalista (Chomsky, 1996 [1995]).

### 2.2.3. Papel da GU na aquisição da L1

Ao partir do princípio que a aquisição da linguagem se processa interna e rapidamente, a linguística generativa entende que a GU é o órgão biológico (o dispositivo computacional) que não só caracteriza o estado inicial da faculdade da linguagem (FL0) mas orienta também todo o processo de aquisição de uma L1 até ao seu estado final (FLf), isto é, até à fixação da gramática do adulto. Deste modo, o Dispositivo de Aquisição das Línguas (*Language Acquisition Device*) inicia-se no estado zero (FL0), completando-se no estágio estável do conhecimento da língua (FLf) (Chomsky, 1965), o que nos permite elaborar a seguinte estrutura:

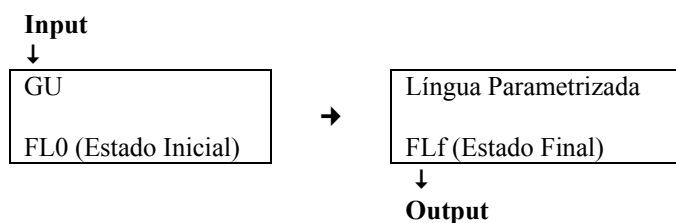
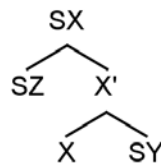


Fig. 10. Modelo de aquisição das línguas.

No dispositivo computacional encontra-se o estado inicial da aquisição (FL0), dotado de princípios rígidos e uma série finita de parâmetros, pelo que:

*“todas as funções a operar vão depender de elementos funcionais e propriedades gerais do léxico. A selecção  $\Sigma$  de opções determina uma língua. Assim, a aquisição de uma língua será o fixing de  $\Sigma$ , i.e., a Gramática da Língua descreve um  $\Sigma$ . As representações nada mais vão ser do que motivadas a condições, por propriedades de interface. As expressões linguísticas vão ser as realizações óptimas das condições de interface. Estas expressões linguísticas serão o objecto formal que satisfaça as condições de interface, da melhor maneira”* (Chomsky, 1996 [1995]:168).

Para representar o “programa” do dispositivo computacional responsável pelas modificações da estrutura projectada a partir do léxico, Chomsky (1996 [1995]:172) propôs a Teoria X-barras, que advoga que são os núcleos projectados que vão ter a principal intervenção nas relações básicas estabelecidas a nível do sistema linguístico, sendo estas relações tipicamente locais (Fig. 11):



**Fig. 11.** Estrutura X-barras.

As relações fundamentais estabelecidas entre os núcleos são então:

- (i) A relação entre SZ e X, isto é, a relação entre Especificador (SPEC) e Núcleo, representando X as restantes categorias;
- (ii) A relação de irmandade a nível temático entre núcleo e complemento, ou seja, entre X e SY.

Como as regras são recursivas, podem combinar-se e gerar um número infinito de estruturas sintáticas. Assim, é possível não só formarem-se estruturas genéricas, tipo SX (Fig. 11), como elaborarem-se também estruturas com múltiplos complementos, árvores sintáticas mais pequenas que omitam partes opcionais, ou estruturas com nós adicionais de SX's ou X's de vários tipos. O diagrama da Fig. 12 ilustra um dos possíveis modos como as regras se combinam para formar uma árvore sintática de estrutura genérica SX:



Fig. 12. Estrutura alternativa gerada a partir da estrutura X-barra (a partir de Carnie, 2007 [2002]:118).

Note-se que o complemento (COMP) contendo X' pode ser distinto de um adjunto (ADJT) contendo X', visto o complemento ter X (núcleo) como irmão, enquanto o adjunto possui X' como irmã.

Independentemente das questões relativas à forma de representar o “programa” do dispositivo computacional responsável pelas modificações da estrutura projectada a partir do léxico, é importante reter que a GU actua por fases de desenvolvimento, que se assemelham às dos domínio anatómico, fisiológico ou mental (Chomsky, 1965). No entanto, a aquisição da língua é um processo impreciso, daí nada garantir que os diferentes aprendentes, a partir dos mesmos DLP's, estruturam internamente a mesma gramática (Kroch, 2004:1). De acordo com o modelo generativista de aquisição de L1's, a transmissão/aquisição linguística ocorre segundo o seguinte esquema:

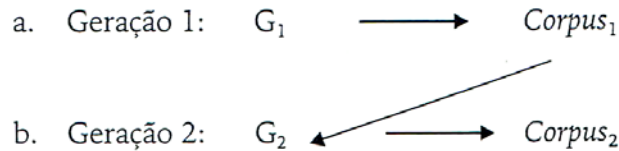


Fig. 13. Modelo de transmissão/aquisição de L1 (Lucchesi & Ribeiro, 2009:144)

Como a criança (G<sub>2</sub>) só tem acesso a DLP's a partir do ambiente linguístico em que está inserida, isto é, à língua-E (*Corpus*<sub>1</sub>) e não ao conjunto de parâmetros valorados para determinadas propriedades morfossintáticas que configura a língua-I (G<sub>1</sub>), vai elaborar sistemas gramaticais hipotéticos e intermédios (*Corpus*<sub>2</sub>), traduzidos em desvios e variação, até atingir um FLf, mais ou menos rígido (a gramática do adulto), e que lhe permitirá gerar as produções linguísticas (*output*) adequadas às situações de

comunicação (Montrul, 2004:4). Ainda assim, o desenvolvimento da linguagem da criança guia-se por uma ordem que vai do genérico para o específico, começando na fase pré-linguística, mas tendo uma progressão bastante rápida. Depois de ultrapassadas as fases dos sons controlados por estímulos<sup>53</sup> e do balbúcio, começa a articular as primeiras palavras<sup>54</sup>, geralmente monossilábicas e com a forma CV, por volta do primeiro ano. O desenvolvimento gramatical é nulo, mas as palavras demonstram que a criança já percebeu que os sons se relacionam com os significados. A partir dos dezoito meses emerge a fase das duas palavras, que revelam não só a existência de relações semânticas e sintácticas mas também entoação sem pausas. Estas proto-frases caracterizam-se pelo recurso a nomes, que constroem frases do tipo SO ou para demonstrar a relação sujeito-lugar. A criança ainda não utiliza os pronomes nem recorre a marcas sintácticas e morfológicas, mas algumas das palavras apresentam flexões, isto é, denotam ter-se iniciado a aquisição gramatical. Após começar a encadear mais do que duas palavras para transmitir mensagens principais, a criança rapidamente inicia a construção de frases e, com dois anos e meio de idade, em situação de desenvolvimento normal, produz já enunciados semelhantes aos dos adultos nativos, revelando direccionamento para os usos da LA, traduzido em domínio flexional nas construções gramaticais e aplicação das estruturas básicas da L1 (Cerqueira, 1994; Faria, Freitas & Miguel, 1998, 2001; Freitas & Miguel, 1998; Costa & Santos, 2004 [2003]; Ferrari-Neto, Corrêa & Augusto, 2005; Castro & Ferarri-Neto, 2007). À medida que vão adquirindo mais linguagem, complexificam as construções gramaticais e, cerca de dezoito meses depois, a variação é esporádica e os enunciados apresentam uma gramática já bastante próxima da dos adultos. Nesta fase, as realizações linguísticas das crianças já não são imitações da fala dos adultos, mostrando que está completado o percurso da aquisição que possibilitará à criança passar a funcionar como um “*telégrafo para o infinito*” (Fromkin & Rodman, 1993 [1974]:357).

No início da aquisição da L1, a GU disponibiliza todos os elementos e possibilidades gramaticais que possui, mas com a exposição ao *input* específico de determinada língua, que leva a criança a testar, em fases sucessivas, as hipótese de utilização do mesmo, ela apreende as regras que a levam a usar a língua adequadamente. A variação registada, que produz erros próprios do processo de desenvolvimento da linguagem, afecta as

partes não-substantivas do léxico e as propriedades gerais dos itens lexicais, sendo, por norma, ultrapassada relativamente cedo. Pelo contrário, na variação linguística referente aos aprendentes de uma L2, os erros mantêm-se durante tempo prolongado e podem mesmo nunca vir a ser corrigidos, fossilizando-se. Deste modo, as línguas não são mais do que subconjuntos particulares da GU, que governa todos os seus estágios de aquisição (Godinho, 2005:97), consoante as necessidades específicas em que aquelas são utilizadas.

Partindo destes princípios, os estudos generativistas têm centrado a sua atenção em dois objectivos: por um lado, perceber e avançar explicações acerca das propriedades universais da faculdade da linguagem, isto é, determinar os princípios abstractos que concedam justificação para as representações fonológicas e sintáctico-semânticas da linguagem; por outro lado, entender e explicar como se processa a variação sistemática entre línguas registada durante o processo de desenvolvimento da aquisição. Assim, os linguistas avançaram com quatro hipóteses, listadas por Godinho (2005:102):

1. A teoria continuísta (*Continuity Hypothesis*), que entende que as gramáticas das crianças e dos adultos são em tudo semelhantes (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996);
2. A teoria não-continuísta, que advoga que a gramática dos adultos é acentuadamente diferente da das crianças (Slobin, 1973, 1985);
3. A teoria da maturação (*Maturation Hypothesis*), que entende que os princípios da GU actuam logo a partir dos primeiros enunciados da criança, a que se seguem, mais tardiamente e de modo descontínuo, outros aspectos operativos (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997).
4. A teoria da aquisição gradual, que se aproxima da anterior, mas postulando que a construção gramatical se dá faseadamente, a partir da aquisição de novos itens lexicais e morfológicos (Montrul, 2004).

Com base nestes pressupostos, muitos têm sido os estudos levados a cabo para se perceber como se processa a aquisição das diferentes L1's e/ou L2's, em contextos de monolingüismo, bilingüismo ou plurilingüismo. No caso concreto da aquisição do

PtgL1, poder-se-ão apontar como relevantes, entre outros, os trabalhos de Cerqueira (1994), Faria, Freitas & Miguel (1998, 2001), Costa & Santos (2004 [2003]), Ferrarri-Neto, Corrêa & Augusto (2005) e Castro & Ferrarri-Neto (2007), enquanto que, para o PtgL2/PtgLE, podem ser indicados as pesquisas de Fransceschina (2002), Godinho (2005, 2008), Matos (2008), Baxter (2009) e Silva (2010).

O modelo abstracto de parâmetro permite não só que a Gramática Comparativa reduza o feixe de diferenças superficiais entre dois sistemas linguísticos a um único parâmetro mas também que a Gramática Histórica passe a conceber mudanças em termos da alteração de valor de um ou mais parâmetros em determinado sistema linguístico. Nesta perspectiva, o nosso trabalho segue o pressuposto de que a GU está subjacente a todo e qualquer contexto de aquisição de L1. Contudo, a própria tradição racionalista entende que o desenvolvimento das capacidades linguísticas, visto apenas sob a perspectiva de *“uma pseudogramática generativa reduzida a uma mera análise das expressões em árvores, representa uma posição de efeitos absolutamente nefastos e revela uma incompreensão total tanto dos fundamentos básicos da gramática generativa como do processo de aquisição e desenvolvimento da linguagem da criança”* (Raposo, 1992:36-37). Portanto, factores sociais, como o meio ambiente, a idade da aquisição, a estabilidade emocional e até a saúde física do falante, têm também influência na parametrização, permitindo aos estudos variacionistas fazerem previsões acerca do realinhamento das propriedades sintácticas dos dialectos sujeitos a contacto linguístico (Kato & Tarallo, 1989:37).

#### **2.2.4. GU e estudos de aquisição**

Tanto a Teoria dos Princípios e Parâmetros como a Teoria das Regências e Ligações constituem a base actual dos estudos sobre aquisição, já que fornecem um quadro teórico bastante fundamentado para o entendimento do modo como esta se processa. De acordo com estes princípios, os dados do *input*, mesmo que não constituam uma aprendizagem explícita, interagem com os princípios (as propriedades) e parâmetros da GU, “engatilhando” nos primeiros para activar os segundos. Portanto, a gramática é fixada, permitindo que a língua reflecta a interacção entre o mecanismo genético humano e o ambiente sociolinguístico aquando da aquisição (Lightfoot, 1999a:51). Como o produto



final – a competência linguística –, representa a conjugação do processamento dos DLP's<sup>55</sup> com o inatismo do indivíduo para a faculdade da linguagem, as características do *input* que a criança recebe são de importância vital para a construção do FLf (Godinho, 2005:91).

A fixação da gramática complexa é uma realidade na fase adulta, independentemente da pobreza dos estímulos, podendo estes ser condicionados pela pressão de factores extralinguísticos aos quais os falantes estão sujeitos: estatuto social e profissional, idade, escolaridade, contacto com outras línguas, etc. Deste modo, a GU é, ao mesmo tempo, flexível (acomoda a variação registada entre as diferentes línguas) e rígida (segue um rumo específico em direcção ao conhecimento final), permitindo que se adicionem às propriedades universais inerentes a ela própria quer as propriedades específicas das diferentes línguas quer as condições sociolinguísticas em que ocorre o *input*. Com base neste pressuposto, estudos recentes têm evidenciado que a GU condiciona a aquisição da linguagem tanto nos contextos monolíngues de aquisição da L1 como nas situações de bilinguismo e de ASL por adultos (White, 1989, 2003; Bruhn de Garavito, 1994; Meisel, 1994a; Hawkins & Chan, 1997; Franceschina, 2002, 2003; Montrul, 2004; Godinho, 2005, 2008; Silva, 2010).

Tendo como ponto de partida os pressupostos teóricos abordados, ao analisarmos a aquisição e desenvolvimento geracional de um aspecto particular da morfossintaxe do português, no caso concreto a CPL-var no SN de uma variante reestruturada resultante de transmissão linguística irregular, dentro da perspectiva da GU, estaremos não só a procurar descrever tal fenómeno mas também a dar o nosso contributo para validação de uma teoria linguística que leva em conta quer os princípios universais quer os aspectos particulares da aquisição das línguas, tal como tem sido feito em outros trabalhos que abordam particularidades específicas de aquisição de L2/LE (Godinho, 2005; Silva, 2010). Paralelamente, a conjugação de análises teóricas com metodologias quantitativas permitirá entender melhor o fenómeno da referida variação, já que cada uma delas nos concede instrumentos importantes para se concluir melhor acerca do trajecto e desenvolvimento diacrónico das características estruturais no interior do SN do PA, com vista a configurar uma possível mudança em curso.

### 2.3. Abordagens teóricas sobre aquisição bilingue

Os falantes de Almojarife são bilingues em português e santomense, pressupondo-se que tenham estado em situação de maior ou menor exposição aos dois dialectos como L1's, ainda que, por vezes, esta evidência seja negada por preconceito sociolinguístico. Independentemente da discussão estabelecida em torno do próprio conceito de bilinguismo, restringi-lo-emos, no âmbito do presente trabalho, a duas perspectivas: (i) considerar o indivíduo bilingue como um falante com uma modelagem única e específica, entendido como um todo (Grosjean, 1982, 1985); (ii) observar o bilingue levando em linha de consideração o papel desempenhado pela aquisição da sua L2. Assim sendo, o bilinguismo acaba por reflectir quer as situações de domínio de duas ou mais línguas quer a adopção de diferentes registos das mesmas, de acordo com o exigido pela pragmática (Dufour & Kroll, 1995:136).

O facto de um cada vez maior número de sociedades tender para o multiculturalismo levou a Neuropsicologia, a Psicologia Cognitiva e a Psicolinguística a procurem entender não só a forma como as línguas do bilingue estão representadas na sua rede cognitiva mas também o modo como ele acede à informação lexical (Oliveira, 2002:87). Os achados revelam que os bilingues desenvolvem estratégias de processamento da informação de acordo com o contexto em que se dá a aquisição de ambas as línguas e que os factores linguísticos (p.e. as diferentes estruturas das diversas línguas) e extralinguísticos (p.e. a idade e modo de aquisição ou a ordem de aprendizagem das línguas) intervêm directamente no modo como as duas línguas são adquiridas pelos bilingues (Hamers & Blanc, 2000 [1989]:115). Assim, as produções linguísticas de cada uma das línguas do bilingue representam *“uma matriz de traços distintivos, incluindo especificações semânticas, e vão ser conectadas a um sistema comum de representações conceptuais. A mensagem, antes de ser codificada e decodificada em cada língua, não faz parte da competência linguística implícita mas de um sistema conceptual.”* (Oliveira, 2002:89). Estes aspectos parecem indiciar que, no bilingue, existem duas representações lexicais específicas, uma para cada língua, enquanto a representação conceptual é comum a ambos os idiomas (Potter *et alii*, 1984:28). Contudo, vários autores discordam deste pressuposto, alegando que, tal como acontece na ASL, uma das línguas (a L2) é mediada pela outra (L1). Assim, uma das línguas do bilingue exerce dominância sobre a

outra (Kroll & Scholl, 1992:196), levando os aprendentes em situação inicial de aquisição das L1's a produzirem erros de morfologia flexional em ambas. A este propósito, lembremos que o presente trabalho tem igualmente como pressuposto a hipótese de que a L1 é o estado inicial para aquisição de uma L2, isto é, de que a L1 exerce dominância sobre a LA. Todavia, e como referiremos no ponto 2.4.1.1, não descartamos também a possibilidade de o estágio final de sistematicidade revelado pelos falantes de Almojarife aliar outro tipo de interferências às provenientes da L1, sejam elas de carácter cognitivo (maturidade, quantidade e qualidade do *input*, frequência de uso da LA, conhecimento de outras línguas, etc.) sejam do tipo extralinguístico (motivação, idade, sexo, escolaridade, aspectos culturais, etc.).

Paradis (1987:117) entende que o bilingue pode funcionar como um nativo nas duas línguas, uma vez que os itens lexicais armazenados contêm quer a forma quer o significado das palavras. Deste modo, nos dois armazenamentos lexicais estarão também presentes quer a forma (fonologia e propriedades sintácticas) quer os significados das palavras, enquanto o armazenamento conceptual, que é comum, contém as representações mentais das coisas observadas e dos eventos experienciados. Deste modo, o modelo entende que, aquando do uso da palavra, compete aos itens lexicais activar toda a panóplia de traços conceptuais que permitem obter uma representação mental do seu referente (Hagège, 1996:112). Este aspecto determina uma perfeita coordenação das duas línguas no bilingue, fazendo-o actuar com um nativo em ambas e levando-o a distinguir as características léxico-semânticas encerradas em termos equivalentes em tradução, mas que activam representações mentais distintas (Paradis, 1987:166; Oliveira, 2002:91).

### **2.3.1. Acesso ao léxico por parte do bilingue**

Alguns estudiosos entendem que cada palavra contém diversos tipos de informação a que correspondem várias representações na rede cognitiva do falante. Contudo, estas estão todas conectadas (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982), o que levanta a questão acerca do modo como o bilingue acede ao léxico. Experiências com recurso ao *priming*<sup>56</sup> revelam que o *priming* semântico acontece entre as línguas do bilingue, em virtude de elas acederem a uma representação conceptual

comum. Por outro lado, o *priming* de repetição não ocorre, evidenciando que uma palavra e a sua tradução não compartilham representação mental. Para os autores que aplicaram este tipo de metodologia (Grainger & Beauvillain, 1988:59; Jin & Fischler, 1990:1153), a situação só é possível porque, a cada uma das línguas, correspondem entradas lexicais independentes, isto é, porque os sistemas lexicais do bilingue estão separados.

Fenómenos característicos dos falantes bilingues, como a alternância de códigos,<sup>57</sup> revelam a capacidade dos bilingues em alternar as suas línguas nas situações normais de fala (Romaine, 1989; Muysken, 1995; Myers-Scotton, 1997). No entanto, esta capacidade de alternância não é arbitrária (Pfaff, 1979; Clyne, 1980; Poplack, 1980b), sendo governada por regras de restrições semântico-estruturais, resultantes das interconexões estabelecidas entre as duas gramáticas do falante (Pfaff, 1979:307). Por conseguinte, a alternância de códigos incide unicamente sobre a Estrutura-S, não actuando sobre o significado das palavras e das frases, uma vez que este é armazenado a nível metalinguístico, ou seja, na Estrutura-P (Clyne, 1980:403). Este aspecto leva também o bilingue a buscar primeiro as palavras no léxico menos activado e, só posteriormente, no da língua de base (Soares & Grosjean, 1984:385, exemplo [35]).

Para Green (1986:213), as línguas do bilingue estão sujeitas a três níveis de activação: (i) nível da selecção, que leva a língua seleccionada a controlar o *output* da fala; (ii) nível da activação, no qual a língua activa colabora no processamento da fala, trabalhando paralelamente à língua seleccionada, mas sem aceder ao canal de fala em uso; (iii) nível da latência, correspondente ao da língua em repouso, ou seja, a língua que se encontra armazenada na memória a longo termo, mas sem intervenção activa no processamento imediato da linguagem. Deste modo, no momento em que o falante vai seleccionar uma das línguas para a usar numa determinada situação, as outras são activadas ou colocadas em suspenso. Estes aspectos permitem ao bilingue seleccionar mais do que uma língua para uso (Green, 1986:219), pelo que as palavras escolhidas durante o processo de fala são da língua seleccionada, mas a partir da língua activa, ou mesmo da língua latente. Por sua vez, a activação da língua implica que se accionem não apenas os seus itens lexicais mas também todo um conjunto de lexemas e unidades sublexicais (p.e. sílabas e fonemas). Assim sendo, a tipologia morfológica de qualquer

língua, em virtude de implicar o envolvimento da estrutura interna das palavras na produção linguística, tem influência determinante na alternância de código. A noção de equivalência entre gramáticas das duas línguas facilita o uso das mesmas pelo bilingue, levando a que ocorram não só alternâncias de código mas também interferências lexicais na sua fala. Este aspecto é tanto mais compreensível se considerarmos que existe equivalência não apenas de categorias (itens lexicais, estruturas fráscas, traços morfossintáticos) mas igualmente de relações entre categorias sintagmáticas (ordem das palavras na frase) ou paradigmáticas (oposições equivalentes). Realizações dos falantes do PA evidenciam isso mesmo, como acontece no caso das construções relativas ou relativas cortadoras (exemplo [160]), em que ocorrerá interferência/transferência da conjunção usada no seu substrato, o santomense (exemplos [158] e [159]), dada a semelhança fônica entre as formas do substrato e da LA ([ku] e [kə]):

[158] FORRO:  
*floli se ku sa i* (Ferraz, 1979:74)  
 flor MSPEC REL estar aqui  
 “Esta flor que está aqui”

[159] FORRO:  
*om ku Zõ sa ka fla n-e* (Ferraz, 1979:71)  
 homemREL João estar ASP falar em-ele  
 “o homem de que/de quem o João está a falar”  
 “o homem que o João está a falar”  
 “o homem que o João está a falar nele/dele”

[160] PA:  
*dia co ele tem, ele dá, mas dia que ele ãõ tem, ele ãõ tem mesmo*<sup>58</sup>  
 [SALVH3]  
 “nos dias (em) que ele tem, ele dá, mas nos dias (em) que ele não tem, ele não tem mesmo”

No santomense pode também ser observado que a conjunção *ku* cumpre ainda outras funções gramaticais, como acontece nas construções integrantes, (exemplo [161]):

[161] FORRO:  
*m-pesa kwe ka sobe* (Ferraz, 1979:65)  
 eu-pensar INTEGRANTE (KWE < KUE) KA chover  
 “eu penso que vai/irá chover”

Idêntica situação ocorre no PA, com a mesma conjunção estendendo a sua função a situações de subordinação:

[162] PA:  
*Mas eu ño tenho canoa, co aparado ño tem dinheiro pa fazê mas canoa*  
[SALVH3]  
“Mas eu ño tenho canoa, se fico parado ño tenho dinheiro para fazer  
mais canoas”

[163] PA:  
*é de vez quando co gente quero co vem fazer, manda-mo co ê faz*  
[SALVH3]  
“é de vez em quando que as pessoas querem que eu faça, mandam que eu  
faça”

Deste modo, alternâncias de códigos e interferências dependem de um conjunto de princípios linguísticos abstractos, patentes em diferentes comunidades linguísticas e, ao que tudo indica, baseados em aspectos cognitivos (Myers-Scotton, 1997:187). Portanto, estes fenómenos, apesar de emergirem na produção linguística, têm o léxico como base. E tal só é possível porque as estruturas gramaticais estão contidas em lemas (as entradas abstractas no léxico mental do falante) e incluem os aspectos fonológicos de um elemento de informação lexical. Por outras palavras: compete ao nível conceptual, o mais abstracto da produção da língua, a selecção dos lemas que vão ser activados, pelo que, neste tipo de fenómenos, a decisão sobre a língua que vai conceder os lemas a serem activados se encontra feita *a priori*. Conclui-se então que, seja na fala monolíngue seja na fala bilingue, a estrutura dos enunciados vai estar dependente das condições pragmáticas e sociopragmáticas em que as referidas falas acontecem.

### **2.3.2. Papel da GU na aquisição bilingue: a *Hipótese da Diferenciação da Linguagem***

Embora as duas línguas do bilingue sejam adquiridas de modo idêntico, na realidade não ocorre simetria entre ambas. De facto, cada uma das línguas está conectada a determinados indivíduos ou actividades do quotidiano, pelo que, com o desenvolvimento da criança, uma delas acaba por ser preferencial e passar a exercer dominância sobre a outra. Não obstante ainda não se ter determinado como e quando se dá a dominância de uma língua sobre a outra na aquisição bilingue, é aceite que a frase é inicialmente

construída de modo parcial na língua dominante e, só posteriormente, traduzida e articulada na língua dominada. O fenómeno da dominância é também frequente em falantes políglotas, como sucede com os aprendentes de PtgLE que tenham previamente adquirido o inglês como L2. De facto, é comum ouvir-se estes falantes realizarem frases do tipo “*Ele introduziu-me ao amigo dele*”, na qual acontece uma clara predominância do verbo inglês “*to introduce*” sobre o seu congénere português “apresentar”. Como se verifica, o que ocorre aqui é todo um conjunto de estratégias de compreensão baseado na língua dominante e não na língua em uso (Bates *et alii*, 1982:274).

Por norma, as crianças monolíngues aprendem a L1 com os pais. Contudo, no caso das crianças bilingues, a sua exposição às duas línguas é diferenciada, já que o contacto pode ser mais intenso com um dos pais, caso, por exemplo, o outro tenha que se ausentar de casa para trabalhar. Estudos feitos com crianças em início de aquisição bilingue, entre dois e três anos de idade, cujos pais detêm e falam diferentes L1’s, têm evidenciado que este tipo de aquisição bilingue é complexo, uma vez que as referidas crianças apresentam uma mistura de sons e palavras das suas duas línguas nas mesmas produções orais, mesmo que estejam a comunicar com falantes que usem apenas uma língua. Uma das justificações para as crianças produzirem este tipo de misturas assenta no facto de elas não deterem vocabulário suficiente numa das línguas, ou mesmo nas duas, para se expressarem completamente numa delas (Genesee, 1989:176). Portanto, a estratégia para tornar a comunicação eficaz é recorrer ao vocabulário que elas conhecem, mas que pertence à outra língua (Vihman, 1985:316).

Face a estes aspectos, e tendo em conta que a GU modela a aquisição em situações nas quais a criança é exposta simultaneamente a duas línguas, Genesee (1989:172) e Meisel (1994a:93) propõem a Hipótese da Diferenciação da Linguagem, que assenta na Teoria da Maturação (*Maturation Hypothesis*) de aquisição de L1’s (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997) e defende que a GU orienta os parâmetros das duas gramáticas (Pfaff, 1979:315), que actuam logo a partir dos primeiros enunciados da criança (Köppe, 1994b:225), desenvolvendo-se, depois, de modo paralelo, mas interagindo (e mesmo influenciando-se) mutuamente (Sridhar & Sridhar, 1980:414). Assim, e de acordo com a hipótese maturacional, a fase FL0 apenas envolve o módulo pragmático da linguagem, não

detendo o módulo do conhecimento gramático, já que o desenvolvimento deste está sujeito a maturação, isto é, à aquisição das categorias funcionais.

Em estudos circunscritos ao quadro dos Princípios e Parâmetros (Chomsky, 1981a), levados a cabo no âmbito do projecto DUFDE (*Deustsh und Französisch – Doppelter Erstspracherwerb*),<sup>59</sup> ou seja, com crianças em aquisição bilingue simultânea de francês e alemão como L1's, Köppe (1994b) estudou a ordenação das palavras em estágio precoce de desenvolvimento da gramática infantil, com vista a detectar a fase em que a criança tem acesso ao movimento SN e respectiva elevação do sujeito. Baseando-se no Princípio do Comprimento Médio dos Enunciados (*Mean Length of Utterance*) e atentando ao momento em que a criança passa a combinar verbos com SN's, Köppe (1994b) buscou determinar se o movimento constitui um princípio da GU disponibilizado na fase FL0, ou, pelo contrário, se encontra sujeito a maturação.

A conclusão de Köppe (1994b:233) é que os princípios da GU estão presentes desde sempre nas duas línguas do bilingue, passando a operar a partir do momento em que as categorias funcionais emergem. A evidência empírica para este pressuposto é fornecida através do modo como as crianças conseguem analisar precocemente as propriedades lexicais de diferentes tipos de verbos, o que lhes permite projectar as diferenças estruturais a nível do SV. O pressuposto de Köppe (1994b), claramente continuísta, leva a entender que a gramática das crianças bilingues não só é igual à das monolingues como também o é em relação à dos adultos, uma vez que tem todos os princípios disponíveis desde o começo da elaboração gramatical (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996).

Também no âmbito do projecto DUFDE, Meisel (1994a) investigou o modo como se processa a aquisição de categorias funcionais e quais as consequências desta para o desenvolvimento das gramáticas em questão. Para tanto, Meisel (1994a:92) partiu do pressuposto que a gramática dos pequenos falantes tem como base inicial as categorias lexicais, sendo as funcionais implementadas gradualmente. Deste modo, o autor opina a favor da Teoria da Maturação (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994), que se contrapõe à Teoria Continuísta (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996), defensora de que “*a linguagem*



*primitiva da criança já é constrangida por princípios da GU e pelo valor não-marcado dos parâmetros” (Kato, 1995:69).*

Observando a aquisição de finitude, concordância e tempo nas gramáticas das crianças, Meisel (1994a:127) conclui que estas não possuem categorias funcionais na fase FL0, as quais vão apenas emergir faseadamente no tempo, acabando por determinar o desenvolvimento das categorias gramaticais associadas à flexão (concordância, tempo e finitude). Assim sendo, a criança não tem inicialmente acesso à gramática, pelo que recorre a princípios semântico-pragmáticos para organizar a sua fala. Por volta dos 2;0 anos, a maturação neurológica da criança disponibiliza-lhe a GU, passando a sua gramática a operar, nesta fase, dentro da Hipótese Continuista (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996). Por outras palavras: Meisel (1994a) não acredita que exista qualquer dependência entre o processo que operava antes dos 2;0 anos e aquele que passa a agir a partir da maturação, ou seja, os princípios semântico-pragmáticos não estão sujeitos a processo de gramaticalização.

As observações de Meisel (1994a:128) levaram-no a concluir que, na gramática infantil, a concordância é a primeira categoria gramatical que permite distinguir as formas verbais finitas das infinitas. Deste modo, as marcas de tempo só aparecem após a criança apresentar todas as marcas de concordância pessoal e grande maioria das concordâncias de número. Estes aspectos fizeram Meisel (1994a:128) concluir também que o processo de maturação se fica a dever à instanciação de traços, já que a finitude é definida inicialmente nos moldes da concordância e, só depois, em termos de tempo.

Com base nos achados de Köppe (1994b) e nos postulados da Hipótese da Diferenciação da Linguagem (Genesee, 1989; Meisel, 1994a), é lícito inferir que, dentro do mesmo espaço de tempo em que as crianças monolingues se tornam competentes numa língua, as bilingues aprendem duas e utilizam-nas hábil e apropriadamente em meios socialmente distintos. Portanto, monolingues e bilingues apresentam progressos similares no desenvolvimento da linguagem, aproximadamente nas mesmas idades. As diferenças na aquisição da linguagem a nível individual reflectem apenas que algumas crianças, sejam elas monolingues sejam elas bilingues, adquirem as suas primeiras palavras ou usam estruturas orais complexas mais precocemente do que outras.

Ainda no âmbito do projecto DUFDE, Müller (1994b) analisa a questão do restabelecimento dos parâmetros, a que nos referiremos mais detalhadamente no ponto 2.8 do presente trabalho. Observando a ordem dos constituintes em orações encaixadas do alemão e do francês, a autora entende que a GU contém uma restrição para o accionamento paramétrico. Nesta conformidade, Müller (1994b:268) postula que os parâmetros não podem ser reaccionados durante o desenvolvimento linguístico, indo ao encontro dos pressupostos avançados pelos defensores de que, apesar de os parâmetros não poderem ser restabelecidos, a GU continua a intervir no processo de ASL (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashes, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003).

Face ao exposto sobre aquisição bilingue, o nosso trabalho segue também o pressuposto de que a GU não está apenas subjacente a todo e qualquer contexto monolingue de aquisição de L1, já que estende a sua acção aos domínios da aquisição bilingue ou plurilingue. Contudo, teremos também em conta que, neste tipo de aquisição, o contexto sociohistórico é determinante para que uma das línguas acabe por se tornar preferencial e passe a exercer dominância sobre a outra. Na comunidade de Almojarife, a primeira fase da aquisição, por monolingues, conduz ao aparecimento do santomense como língua preferencial de comunicação interétnica, restringindo o uso das L1's do grupo níger-congo atlântico a contextos meramente familiares. Posteriormente, a aquisição do português em situação de transmissão linguística irregular vai determinar, em contexto de bilinguismo, a emergência do PA, marcado por elevada variação, dada a defectividade dos DLP's. Posteriormente, a variação dilui-se ligeiramente em direcção ao sistema da LA, como bem o revela a aquisição da regra de concordância pelas gerações mais novas e mais escolarizadas de Almojarife, ou seja, pelas gerações expostas a maior contacto com a LA.

#### **2.4. Abordagens teóricas sobre aquisição de L2's**

Presentemente, as teorias sobre a intervenção da GU na ASL apoiam-se em diferentes perspectivas: (i) acesso total à GU, isto é, a versão forte sobre o que é transferido (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006); (ii) acesso parcial à GU, ou seja, a versão fraca sobre o que é transferido (Lefebvre, 1986;

1998; Jake & Myers-Scotton, 1998; DeGraff, 1999a, 1999b; Siegel, 2004a; 2004b); (iii) influência da L1 nos estágios iniciais de aquisição da L2, entendendo-se aqui uma nova perspectiva sobre a abordagem contrastiva entre L1 e L2, não limitada à questão da transferência de hábitos (Lightbown & Spada, 2010 [2006]).

#### **2.4.1. Modelo generativista e papel da GU na aquisição de L2's**

Como mencionámos nos pontos 2.2.3 e 2.3.2, a aquisição de L1's (contexto monolíngue ou bilingue) processa-se interna e rapidamente, ainda que as crianças produzam variação no processo de desenvolvimento da linguagem durante as fases precoces da aprendizagem, o que não acontece na ASL. No que concerne a esta, actualmente bastantes linguistas defendem que a melhor forma de a entender passa pela perspectiva da GU. Porém, a ASL, depois de ultrapassado o período crítico do desenvolvimento, terá que ser explicada com recurso a outro quadro teórico, pois os mecanismos da GU poderão já não se encontrar disponíveis nesta fase (Clahsen & Muysken, 1986; Muysken, 2001). Efectivamente, se se levar em conta que a faculdade da linguagem é biologicamente determinada e que está presente na aquisição de uma L1, é também de considerar que esses princípios tenham um papel interventivo na ASL (Flynn, 1989:92), ainda que o contexto desta se afaste radicalmente daquele que determinou a emergência da gramática da LA. Assim, ambos os tipos de falantes passam por vários estágios de construção da gramática, acabando a aquisição das regras reguladas pela GU por reflectir as competências linguísticas nesses mesmos estágios. Assim, as diferentes gramáticas que a interlíngua em maturação conhece, (IL1, IL2, IL3...), desde a fase FL0 até ao patamar FLf, traduzem-se na competência final (ILn) (cf. esquema de Godinho, 2005:98), a qual só difícil e esporadicamente alcançará o mesmo nível de *performance* dos falantes nativos da LA.

Na ASL, as opiniões quanto ao que será o seu estágio inicial divergem em duas correntes: uns defendem que é a própria GU, tal como acontece na aquisição da L1 (Flynn, 1989; White, 1989, 1990); outros entendem que é a L1, mantendo-se a GU activa enquanto a interlíngua se desenvolve, a fim de substituir a L1 quando esta revela lacunas ao nível da assimilação do *input* ou do fornecimento de representações (Montrul, 2004). Por seu lado, os que compartilham a visão de que o estágio inicial da ASL é a GU

acham que esta possui um dispositivo computacional inato de princípios universais e invariáveis (cf. Chomsky, 1996 [1995]), comuns a todos os idiomas, por forma a regerem a forma e o funcionamento da gramática, imprimindo as diferenças estruturais entre as línguas. Tal é possível porque o léxico é dotado de parâmetros que actuam sobre as categorias lexicais e funcionais específicas das línguas. As primeiras (categorias lexicais), como o seu próprio nome indica, detêm propriedades que definem a categoria lexical da palavra, fornecendo o conteúdo semântico às produções linguísticas; por seu lado, as segundas (os morfemas funcionais) possuem características universais semânticas e de flexão, realizadas através de formas morfofonológicas, e interagem com a morfologia flexional para concederem a informação gramatical (género, número, concordância, caso, negação, tempo e aspecto). Dispondo destes dados, e tendo em conta que as categorias lexicais são dominadas pelas funcionais, isto é, que as diferentes unidades linguísticas estabelecem relações hierárquicas entre si, compete ao dispositivo computacional filtrar estas, com vista a transformá-las em representações sintácticas. Assim, ao contrário do que sucede com o dispositivo computacional ou com as categorias funcionais, a construção da frase não é universal, pois depende do léxico da língua (ponto 2.2.2.1.2 do presente trabalho).

Face à teoria de que o estágio inicial da ASL é a GU, conclui-se que o dispositivo computacional, pelo facto de ser inato, não necessita de ser aprendido. O mesmo não acontece com o léxico particular (as palavras e a morfologia) das diferentes línguas, que precisa de ser aprendido durante o processo de aquisição. Ao adquirir a L1, a criança procede também à aquisição do léxico e das categorias funcionais desta (os parâmetros que regem o conjunto de propriedades sintácticas e semânticas abstractas da L1). Na ASL, ocorre idêntico processo, pois é necessário adquirir o léxico e os parâmetros da LA.

#### **2.4.1.1. GU e teorias em torno da aquisição de L2's**

Tendo em conta que, no estado inicial da ASL, o aprendente já possui a gramática da GU, com princípios e parâmetros estabelecidos, das quatro teorias que foram mencionadas acerca da disponibilidade da GU na aquisição de L1's (ponto 2.2.3) apenas a Hipótese Continuista (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996)

e a Hipótese da Maturação (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997) estão implicadas na ASL (Godinho, 2005:102). De acordo com a Hipótese Continuista, a gramática das crianças é igual à dos adultos, tendo todos os princípios disponíveis desde o início da elaboração gramatical. Deste modo, e dado que a criança possui competência gramatical total na língua a que se encontra exposta, se os parâmetros não forem fixados, é porque se verificam dificuldades na aquisição. Caso não ocorram estas, as crianças produzem frases gramaticais, conduzidas pelos parâmetros filtrados pela GU.

Por seu lado, a Hipótese da Maturação entende que os princípios e parâmetros estão geneticamente programados para serem fixados de modo gradual, isto é, em diferentes estágios de maturação. Como tal, apenas determinadas categorias gramaticais operam na fase inicial da elaboração de frases, actuando outras em período mais tardio, e segundo uma ordem específica. Assim, interessa perceber em que fase da maturação ocorre a transformação da GU em gramática da língua a ser aprendida.

A Teoria da Maturação advoga ainda que existe uma diferença acentuada na ASL entre crianças e adultos. Assim, enquanto nas primeiras ocorre uma fase de maturação semelhante à da aquisição da L1, nos segundos essa fase já se encontra concluída por terem efectuado a aquisição da L1 (Yip, Rutherford & Clahsen, 1995:25). De facto, estudos sobre aprendentes da mesma L2, mas detentores de distintas L1's, apontaram para diferentes potencialidades de ASL (Schachter, 1989; Franceschina, 2002; White *et alii*, 2004). Estes resultados permitiram também que se avançasse com três hipotéticas premissas direccionando a ASL, podendo as mesmas partilhar aspectos comuns ou distinguir-se nas suas etapas até ao estágio final (Godinho, 2005:103-104): (i) a GU actua a partir do estado inicial da aquisição; (ii) a GU actua a partir do estado inicial da aquisição, mas assentando na L1; (iii) registam-se importantes transferências da L1 para a L2. Actualmente, é consensual considerar que a GU actua através da L1 na ASL, divergindo as teorias apenas acerca do momento da operacionalidade: estado inicial ou intervenção apenas quando se verificar necessidade de tal?

Considerando que a GU opera por intermédio da L1 na ASL, compreende-se o porquê de a interlíngua deter características formais da L1 e da L2. Este aspecto, por outro lado, vem confirmar a existência de diferenças e semelhanças nos processos de aquisição de

uma L1 e de uma L2. A nível das diferenças, poder-se-ão destacar (Godinho, 2005:104-105): o facto de o aprendiz da L2 já ter desenvolvido um sistema linguístico (a L1); o facto de os adolescentes e adultos possuírem maturidade cognitiva; o facto de o estágio final da L2 poder revelar variações (erros) fossilizadas; e ainda o facto de os aprendentes da LA nunca atingirem o nível de proficiência dos falantes nativos desta. Quanto às semelhanças entre aquisição de uma L1 e ASL, podem apontar-se as seguintes (Godinho, 2005:105): ambas têm como objectivo a aquisição de um sistema complexo, a partir de *input* reduzido; os aprendentes das duas cometem o mesmo tipo de variação, fazendo pressupor a existência de um sistema linguístico interiorizado de regras sistemáticas (os princípios e parâmetros da GU), que vai governar as construções linguísticas. Face a este quadro de possibilidades, os generativistas, no que concerne à ASL, têm procurado explicar como se desenvolve a competência linguística do aprendiz, seja em situação de aprendizagem formal seja em contexto natural de aquisição.

Assim, e ainda relativamente à questão da disponibilidade/indisponibilidade da GU na aquisição de L2's, são quatro as posições assumidas actualmente pelos linguistas (cf. Godinho, 2005:106-111; Figueiredo, 2009a:31-32; Silva, 2010:40-44):

- (1) A GU actua na ASL, mas apenas de modo parcial, porque, apesar de os princípios se manterem acessíveis, os parâmetros não podem ser restabelecidos (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashes, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003). Uma evidência para sustentar esta hipótese é fornecida fonologicamente, por exemplo, pela perda de capacidade discriminatória a nível dos padrões trocaicos e jâmbicos que ocorre em idade pós-adolescência.
- (2) A GU actua de forma igual (total) na aquisição da L1 e da L2, sendo ambos os processos bastante similares. Desta forma, os aprendentes da L2 têm acesso total aos princípios e parâmetros desde o início da aquisição e até que esta se complete (Cook, 1994; White, 1989, 2003; Bruhn de Garavito & White, 2002; Montrul, 2004; Cook & Newson, 2007 [1988]). As competências em L1 e L2 são paralelas (duas configurações na GU), só assim se compreendendo o porquê de os falantes avançados de uma L2 deterem representações mentais complexas desta e que se

podem diferenciar das que são elaboradas pelos falantes da LA. Portanto, a relação entre a GU e a aquisição da L1 traduz-se na competência em L1 (Gramática *n*), enquanto a relação entre GU e L2 produz a competência em L2 (IL *n*). Porém, o estágio final de aquisição, e conseqüente variação da interlíngua, estão dependentes da maior ou menor pobreza dos estímulos durante a aprendizagem. Defendem os seguidores do acesso total à GU (*Full Transfer/Full Access Model – FTFA*) que, mesmo em situação de pobreza de estímulos muito acentuada, a ASL não cessa. Assim, e como esta não pode ser atribuída ao *input* que provém do meio envolvente, a ASL tem que ser vista como uma propriedade genética do ser humano. Como se vê, a perspectiva do acesso total à GU engaja na Teoria Continuista da disponibilidade da GU na aquisição de L1's (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996), levando a entender que existe continuidade no desenvolvimento linguístico que configura a ASL.

- (3) A GU actua na L2 através da L1,<sup>60</sup> pois é hoje considerado que a maioria dos aprendentes de uma L2 parte do princípio que esta é igual à L1. Argumentam os autores da corrente do acesso indirecto à GU que o ponto de partida para a ASL é a gramática da L1 e que as propriedades computacionais desta se transferem em bloco para as fases iniciais da ASL, excepto as matrizes fonéticas e os itens léxico-morfológicos (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006). Assim, numa fase inicial da ASL, os aprendentes cometem erros de morfologia flexional, mas, com a aquisição intermédia de morfologia flexional relevante e de itens lexicais, é possível refixar os parâmetros e atingir um estágio final de aquisição. Portanto, a posterior exposição ao crescendo de *input* reconstrói, sucessivamente, a interlíngua. Erros em fase tardia são justificados pela FFFH (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto as categorias funcionais deixarem de estar disponibilizadas depois do período crítico do desenvolvimento (Lenneberg, 1967; Chomsky, 1972; Larsen-Freeman & Long, 1991; Ellis, 1994). Contudo, é tido como certo que a construção gramatical é orientada pela GU, que mantém sempre os seus princípios activos, permitindo que os aprendentes adquiram valores

paramétricos a partir dos valores de parametrização da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006).<sup>61</sup>

- (4) A GU deixou de estar disponível para actuar na ASL, pois já desempenhou o seu papel na aquisição da L1 ou aquisição bilingue e impede os aprendentes da L2 de acederem aos princípios universais ou de estabelecerem os parâmetros (Clahsen & Muysken, 1986; Muysken, 2001). Neste aspecto, destaque-se a posição radical de Meisel (1997, *apud* Godinho, 2005:106), que corrobora Selinker (1972:216) ao considerar que a interlíngua não é um sistema natural nem sistemático, uma vez que os aprendentes de L2's são incapazes de estabelecer correctamente a dependência estrutural da sintaxe complexa. Assim sendo, a L2 não tem qualquer relação com a faculdade da linguagem, sendo adquirida com recurso a capacidades cerebrais distintas desta. O pressuposto assenta nos princípios advogados pela perspectiva psicolinguística de interlíngua (Selinker, 1972), já que a competência em L2 é bastante distinta da competência em L1, podendo a LA ser adquirida por indução, o que não acontece com a L1.

O Programa Minimalista (Chomsky, 1996 [1995]), a que faremos referência no ponto 2.6.2.1, advoga que os traços não interpretáveis na fase FL0 da aquisição, como acontece com os traços de género e número, são apenas usados para activar operações sintácticas. Como tal, podem ser apagados, caso a concordância se dê com traços valorados a partir de outros elementos nucleares da estrutura frásica. Este aspecto pode determinar uma competição de gramáticas na interlíngua, caso se entenda que a L1 é o estado inicial para a ASL e ambas as línguas possuam sistemas distintos de marcação em género e número, que é exactamente o que sucede a nível do santomense e do português. Assim, a concordância entre os constituintes frásicos está dependente de aspectos como a Condição de Visibilidade,<sup>62</sup> que leva a que os elementos componham o seu valor paramétrico a partir do valor paramétrico do núcleo, produzindo-se a concordância entre todos os itens de acordo com o valor patente neste. Seguidamente, os elementos são projectados para a FL, levando os seus radicais terminais definidos para um determinado valor, no caso o valor do núcleo. Ora, aqui levanta-se um problema a nível da construção sintáctica na interlíngua, já que o padrão de marcação pode ser diferente na L1 e na L2,



levando o falante a enfrentar dificuldades na fixação dos valores na interlíngua. Daí que, tal como advogam os defensores da hipótese do restabelecimento de parâmetros (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), os aprendentes de L2's, em idade pós-crítica elaborem construções em que é evidente a simbiose entre sintaxe da L1 e itens lexicais da LA (formas morfofonológicas da L2 em especificações de traços da L1) (DeGraff, 1999b:482).

O facto de as construções divergirem nas produções da L1 e da L2 revela que elas são condicionadas por princípios universais, o que mostra a importância da GU na ASL. Este aspecto leva também a que se direcione atenção aos traços funcionais, já que se estes exibirem proximidade na L1 e LA, ocorre também um desempenho mais aproximado entre aprendentes da L2 e falantes nativos da LA (Hawkins & Chan, 1997:220). Por outro lado, reside também nos traços funcionais a justificação para o facto de os aprendentes de uma L2, ainda que continuamente expostos a dados da LA, não conseguirem adquirir as mesmas representações sintácticas dos falantes nativos desta. Deste modo, apesar de os falantes da L2 alcançarem competência no que respeita às propriedades sintácticas distribucionais dos itens funcionais, poderão não as reproduzir de modo sistemático, porque não conseguem fazer corresponder as características formais às formas morfofonológicas (Godinho, 2005:109). Neste caso, estamos perante uma situação que alguns autores (Prévost & White, 1999:219) consideram como superficial (*Missing Surface Inflection Hypothesis* – MSIH).

Como se vê, qualquer dos modelos proposto acerca do acesso à GU na ASL pode ser válido para diferentes aprendentes, ou mesmo para distintos aspectos da língua em relação ao mesmo aprendente (Cook, 1994; Cook & Newson, 2007 [1988]). Assim, o problema do investigador será escolher, entre todos eles, os que melhor resposta dêem às situações com que se depara no processo de análise.

#### **2.4.2. Aquisição de L2's e modelos psicocognitivos**

Os seguidores do cognitivismo entendem que o conhecimento linguístico é indissociável das capacidades cognitivas do aprendente (Godinho, 2005:81). Assim, a ASL representa uma construção de sistemas do conhecimento, direccionada para a compreensão e uso. Dado que a capacidade de retenção do conhecimento é limitada, o

aprendente necessita da experiência e da prática para ir, progressiva e automaticamente, utilizando os conhecimentos que vai interiorizando (McLaughlin, 1987; McLaughlin & Heredia, 1996, *apud* Godinho, 2005:81).

#### **2.4.2.1. Conexionismo (*Connectionism*) e modelos de processamento paralelo distribuído**

Alguns pesquisadores entendem a capacidade de aquisição como o resultado de uma série de conexões que o cérebro estabelece a nível das suas redes neurofisiológicas. Estas ligações ocorrem em módulos informativos e tornam-se mais fortes ou mais fracas, caso haja ou não activação das conexões mentais e neurológicas que se interligam, levando-as a excitarem-se ou inibirem-se umas às outras, consoante os elementos linguísticos ocorram conjunta e frequentemente, ou não. Como não existe um módulo neurológico específico para a aprendizagem da linguagem, esta deve ser demarcada do conjunto de regras abstractas que lhe é atribuído, pois a aquisição não passa de um mero processamento cerebral do *input* (Gass & Selinker, 2001:216, *apud* Godinho, 2005:82).

Segundo o modelo, os aprendentes de uma língua são sensíveis ao *input* das suas formas regulares, delas extraindo e testando padrões probabilísticos de aplicação. Por conseguinte, os pressupostos do Conexionismo não lidam exclusivamente com a construção do conhecimento linguístico (a competência) mas debruçam-se também sobre a forma como se acessa a este conhecimento e se aplica o mesmo (a *performance*). Para observar a forma como se estabelecem as conexões no processo de aquisição, Rumelhart & McClelland (1986) testaram hipóteses com recurso a modelos computacionais programados para reproduzirem as redes neurofisiológicas do cérebro humano e suas activações, à medida que a aprendizagem vai prosseguindo. Trabalhando com crianças, os pesquisadores observaram a aquisição das formas verbais regulares e irregulares do pretérito no inglês, conseguindo que o programa indicasse o modo como as conexões iam, progressivamente, sendo activadas. Estas observações seriam continuadas por Pinker (1991:487), o qual concluiu que somente os verbos irregulares são recuperados a partir de associações de memória, já que os verbos regulares são produzidos em resultado do processamento das regras de sufixação.

Quanto à ASL, Sokolik & Smith (1992:58), com recurso a métodos conexionistas, investigaram a atribuição do género nos nomes em francês, concluindo que a diferença

registada entre aprendentes do francês L1 e L2 é explicada com a troca de duas variáveis no modelo: (i) nas situações em que o computador atribuiu um estado zero de conexão aos aprendentes de L1, assumiu também que os aprendentes de L2 cumpriam a mesma tarefa, aplicando padrões conectivos pré-existentes, os quais interferiam na aquisição; (ii) o modelo considerou ainda que a ASL apresentava um grau de aprendizagem mais baixo. Como se pode constatar nas considerações em (i), ocorre um paralelismo com os pressupostos generativistas de que o ponto de partida para a ASL é a gramática da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006).

Por seu lado, Ellis & Schmidt (1997), partindo dos pressupostos advogados por Rumelhart & McClelland (1986) e Pinker (1991), analisaram a aquisição da morfologia PL em adultos, verificando que a mesma resulta da associação de determinados mecanismos a nível cerebral. Nesta conformidade, as estruturas complexas são construídas a partir de processamentos simples, ou seja, envolvem automatizações, tal como fora advogado em trabalhos psicolinguísticos prévios (Schiffrin & Schneider, 1977; Schneider & Schiffrin 1977; McLaughlin, 1987, 1990).

Contudo, o modelo conexionista apresenta desvantagens, já que se torna questionável determinar até que ponto se podem isolar variáveis que interagem em contextos naturais e, conseqüentemente, até que ponto os resultados obtidos desta forma espelham o que acontece na vida real. Por outro lado, o que se observa com recurso a esta metodologia tende a ser demasiado específico e localizado, daí resultando o perigo de se ignorar o modo como os diferentes aspectos da aquisição interagem. Assim, o conhecimento sobre os eventos, conceitos e linguagem surge representado de modo bastante difuso no sistema cognitivo. Este aspecto, associado ao facto de a aprendizagem ser apresentada como um sistema cumulativo, dificulta ainda a explicação acerca do modo como ocorre a aquisição das categorias funcionais marcadas por sintaxe complexa.

#### **2.4.2.2. Teoria da “Processabilidade” (*Processability Theory*)**

A morfologia flexional pressupõe a distinção entre três categorias: flexão inerente, flexão contextual e flexão morfofonémica. A primeira não é requerida pela sintaxe, mas tem relevância sintáctica (Booji, 1995:2). É o caso da flexão da categoria de número dos nomes, do grau comparativo e superlativo dos adjectivos e do tempo e aspecto dos

verbos, em que a presença de traços morfossintáticos num determinado lexema afecta a sua referência. Quanto à flexão contextual, é orientada pela sintaxe e abarca a pessoa e número dos marcadores verbais que concordam com o sujeito e/ou objectos, marcadores de concordância dos adjectivos e marcadores estruturais de caso dos nomes (Booji, 1995:2). Portanto, os traços morfossintáticos expressam-se no lexema como consequência de uma parte da construção sintáctica, não afectando a sua referência. Por fim, a flexão morfofonémica não expressa traços de significado ou traços morfossintáticos funcionais, possibilitando apenas que os traços morfossintáticos possam ser realizados (Kihm, 2010:1). É o que acontece com a flexão de género e número, cujos exponentes não podem ser afixados ao radical verbal sem a inserção de uma vogal temática, mas que não expressa qualquer tipo de traço: (p.e. cant-a-s/\*cant-s).

Assentando em modelos psicolinguísticos que visam explicar o modo como a produção do discurso se desenvolve, a Teoria da “Processabilidade” (Processability Theory) defende que o desenvolvimento morfossintático da interlíngua segue um trajecto universal perfeitamente definido (Pienemann, 1998, 2005b). Observando a aquisição da ordem das palavras e de estruturas gramaticais no alemão L2, Pienemann (1998) defende que estas obedecem a um processamento rígido. Assim, as novas estruturas só são aprendidas caso as etapas de aquisição prévia estejam concluídas, isto é, que os processamentos linguísticos arquitectados pelo processador da língua estão dependentes de uma hierarquia universal inata, não necessitando o aprendente de utilizar estratégias ou habilidades para produzir as realizações da L2. Desta forma, as predições sobre o desenvolvimento da ASL podem ser testadas empiricamente. Tomando-se como exemplo a frase “O filho deu um gato à mãe”, poder-se-á elaborar o diagrama da Fig. 14, que ilustra o modo como se processa a aquisição dos elementos frásicos.

A produção do primeiro SN da frase só é possível após ser gerada a mensagem pré-verbal na unidade que Levelt (1989:23) classificou de conceptualizador. Esta transmite o conceito “FILHO” ao codificador gramatical, responsável pela produção da pertinente estrutura gramatical, ou seja, o SN. O processamento de parte da estrutura frásica é feito com itens lexicais resgatados do léxico, pelo que a selecção do lema “FILHO” determina a categoria N, que fornece o DLP em torno do qual é construído o SN.

CONCEPTUALIZADOR

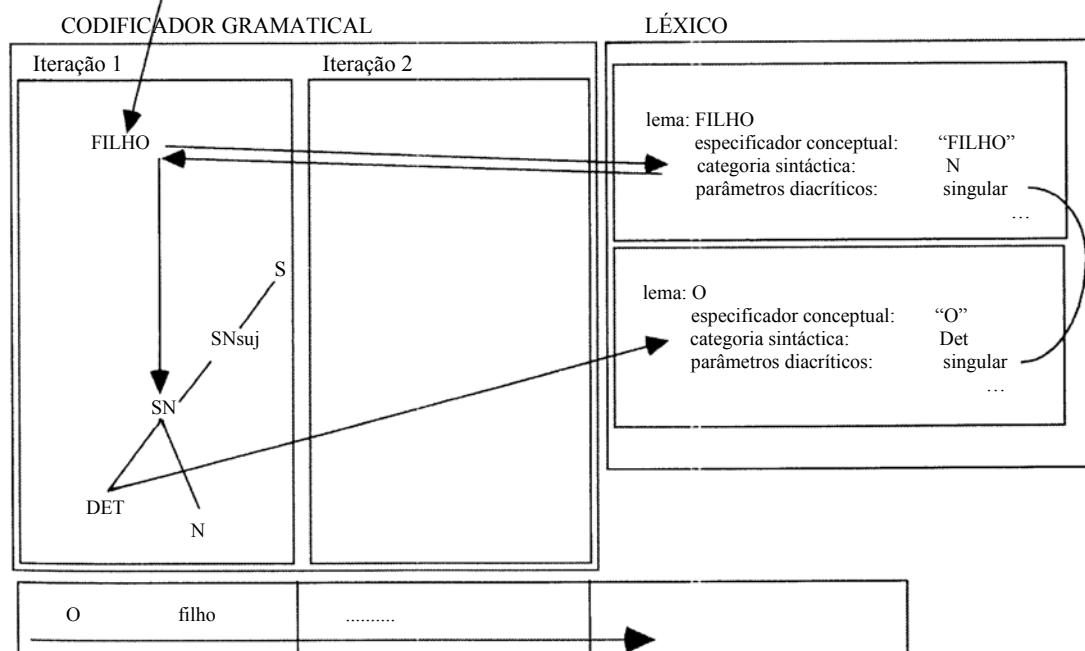
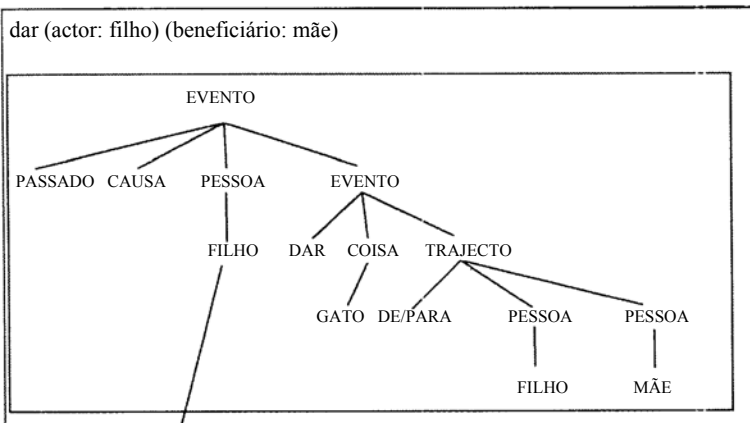


Fig. 14. Incrementação estrutural da interlíngua (Pienemann, 1998:68)

O diagrama revela que o processo de construção frásica implica, necessariamente, as seguintes tarefas (cf. Levelt, 1989:236):

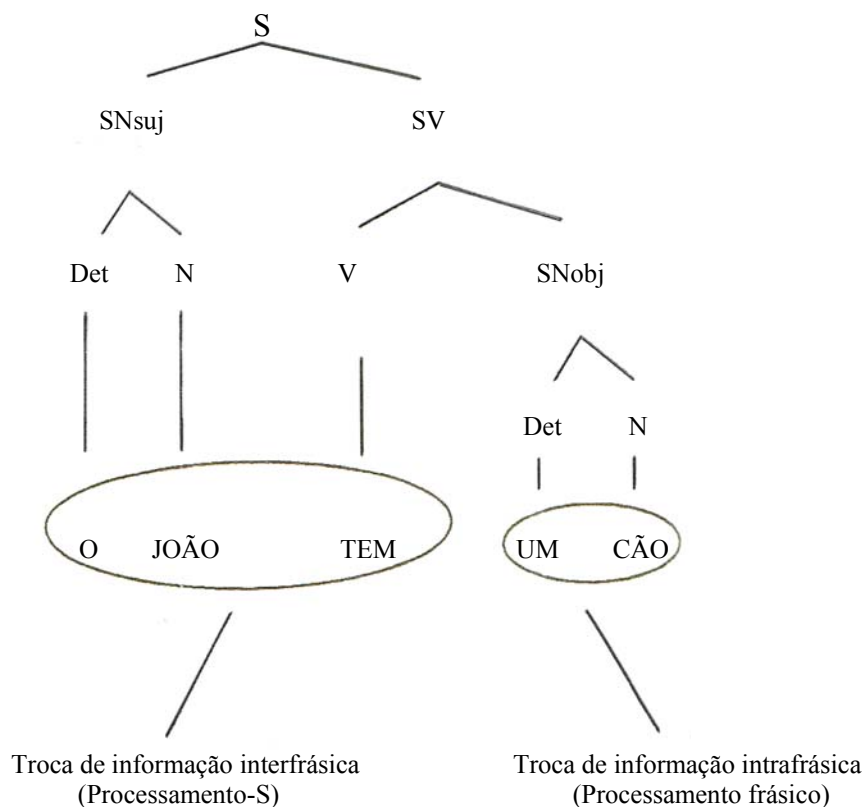
- fornecimento da estrutura sintáctica na qual poderão ser inseridos os itens lexicais;
- busca dos traços morfossintácticos que estabeleçam o pareamento com a estrutura conceptual (p.e. [+ definido]);
- selecção do traço [singular] para o lema “FILHO”;
- pareamento dos traços de itens que possam ser inseridos sob o nó do DET do SN com os do lema seleccionado que detém a função de núcleo;

- resgatar do léxico os lemas pertinentes (p.e. o determinante “O”);
- por fim, transferir a estrutura resultante para unidades de processamento posterior, como o codificador fonológico e o articulador.

O exemplo abordado ilustra uma das características fundamentais da produção do discurso: a incrementação. Esta leva a estrutura a ser gradualmente construída com o processamento da conceptualização, operando os processamentos subsequentes conjuntamente com *outputs* prévios ainda incompletos, que obrigam a disponibilizar na memória, a curto termo, *outputs* intermédios incompletos. Verifica-se então que produção linguística envolve bastantes processamentos paralelos, consideráveis constatações memorizadas a curto termo e a disponibilidade para lidar com todo o tipo de estruturas rotineiras de processamento específico (Pienemann, 1998:7), tais como:

- acesso ao lema;
- processamento categorial;
- processamento frásico;
- processamento oracional (processamento-S);
- processamento da oração subordinada, se for caso de tal.

Como o processamento frásico é responsável pela troca da informação gramatical no interior da frase, os traços morfossintáticos dos constituintes são pareados durante a sua realização (Fig. 15). Note-se que a troca de informação intra e interfrásica seria impossível se não ocorressem, antes, o processamento do acesso ao lema – fornece o material lexical e os seus respectivos traços diacríticos – e o processamento da categoria – concede a informação da categoria sintáctica que permite a elaboração da estrutura posterior –. Por seu lado, o processamento-S implica o acordo entre o sujeito e o verbo. Este só é possível caso se respeite a estrutura sintáctica da direita, isto é, caso o sujeito estabeleça pareamento com a informação gramatical pertinente (a terceira pessoa do singular). Verifica-se então que o processamento-S é aquele que permite a troca gramatical entre os SV e SN sujeito:



**Fig. 15.** Processamento-S e Processamento Frásico (Pienemann, 2000:101).

Seguidamente, na Tabela 2.1 apresenta-se uma hipotética hierarquia do desenvolvimento dos estágios de ASL em inglês:

**Tabela 2.1.** Hierarquia hipotética de articulação dos processamentos na aquisição da interlíngua (Pienemann, 2000:108).

	T1	T2	T3	T4	T5
<b>Processamento-S'</b> (S embebido)	-	-	-	-	+
<b>Processamento-S</b>	-	Simplificado	Simplificado	Troca de informação interfrásica	Troca de informação interfrásica
<b>Processamento frásico</b>	-	-	Troca de informação frásica	Troca de informação frásica	Troca de informação frásica
<b>Processamento categorial (categoria lexical)</b>	-	Morfemas lexicais	Morfemas lexicais	Morfemas lexicais	Morfemas lexicais
<b>Acesso (palavra / lema)</b>	+	+	+	+	+

Na tabela acima verificam-se 5 momentos de aquisição (T1 a T5), em que ocorrem outros tantos estágios diferenciados de desenvolvimento da interlíngua. As evidências

empíricas têm revelado que um processamento-S simplificado já se encontra disponível em estágios bastante precoces da aquisição. Em T1, o aprendiz apenas consegue produzir realizações holofrásticas, isto é, frases compostas por um único lema de categoria indefinida e resgatado do léxico. No estágio seguinte, o falante resgata morfemas lexicais e constrói frases bastante simples, tipo NVN (SVO) ou NNV (SVO ou SOV), mas que demonstram que os referidos morfemas já exibem uma categoria específica nesta fase da aquisição. No período T3, ocorre apenas concordância interna no SN, o que deixa perceber que acontece troca de informação intrafrásica, não havendo ainda concordância interfrásica, isto é, concordância entre sujeito e verbo. Na fase T4 verifica-se o processamento-S completamente desenvolvido, o que não acontece no PA. Por fim, no estágio T5 já é possível ao falante intercalar frases.<sup>1</sup>

Na Tabela 3.1 é possível observar a articulação dos processamentos que correspondem aos estágios de desenvolvimento da sintaxe mencionados anteriormente:

**Tabela 3.1.** Articulação dos processamentos na aquisição da interlíngua inglesa (Plag, 2008a:124).

<b>Estágio</b>	<b>Articulação do processamento</b>	<b>Processamento da L2</b>	<b>Morfologia</b>	<b>Sintaxe</b>
<b>T1</b>	Palavra / lema	Palavras	Formas invariáveis	Constituinte único
<b>T2</b>	Processamento categorial	Morfemas lexicais; Pronomes possessivos	Plural dos nomes	Ordem canónica
<b>T3</b>	Processamento frásico	Troca de informação intrafrásica	Concordância no SN	Frases iniciadas com ADV / “do”; Topicalização
<b>T4</b>	Processamento-S / ordem da frase	Troca de informação interfrásica	Neg+ V	Inversão “yes/no”; Inversão da cópula
<b>T5</b>	Processamento-S / ordem da frase	Troca de informação interfrásica	Concordância SV (3SG –s)	“Do” / Auxiliar 2ª posição
<b>T6</b>	Processamento das orações subordinadas	Orações principais e subordinadas	—	Cancelamento da inversão

A pluralização dos nomes acontece em T2, enquanto a concordância a nível do SN ocorre na fase T3, ou seja, quando se inicia a troca de informação gramatical intrafrásica. Por seu lado, a concordância sujeito-verbo apenas se verifica no estágio T5, isto é, numa fase em que a troca de informação interfrásica já está consolidada. Observando a distinção entre flexão inerente e flexão contextual, verifica-se que a falta



de flexão é uma característica da fase T1, mas que se pode prolongar pelos estágios seguintes, dependendo do tipo de flexão que se pretende observar. A flexão inerente instancia-se a partir do estágio T2 (p.e. pluralização dos nomes nus), enquanto a aplicação da concordância requer a articulação de processamentos posteriores, acontecendo unicamente em estágios mais avançados.<sup>64</sup>

Diversos estudiosos têm testado os princípios da Teoria da “Processabilidade” em distintas L2’s com falantes de diferentes L1’s, confirmando o processamento e as hierarquias do desenvolvimento morfossintático (Kawaguchi, 2005; Mansouri, 2005; Pienemann, 2005a; Pienemann, Biase & Kawaguchi, 2005; Pienemann *et alii*, 2005b; Zhang, 2005). Este aspecto dá consistência aos seus princípios, validando a sua aplicação nos estudos que visam observar a forma como surge e se desenvolve a morfossintaxe flexional de todo o tipo de línguas, incluindo as adquiridas em contexto massivo ou emergentes em situação de transmissão linguística irregular.

#### **2.4.2.3. O Modelo dos Quatro Morfemas (*The 4M Model*)**

Apesar de, presentemente, as teorias sobre ASL versarem mais sobre a essencialidade da GU na ASL, o Modelo dos Quatro Morfemas, baseada no modelo psicolinguístico de Levelt (1989:182) e avançado por Myers-Scotton & Jake (2000a, 2000b) e Myers-Scotton (2002), tenta descrever a produção normal e espontânea dos adultos. Exaustivamente descrita e analisada por Lopes (2001:48-59) e Godinho (2005:84-86), esta teoria procura descrever o modo como se organiza o léxico mental e a forma como se acede a ele para se conseguir a produção linguística. Em termos de ASL, o modelo prevê que os morfemas de conteúdo da LA sejam adquiridos antes dos morfemas de sistema. Por outro lado, tanto os morfemas de sistema semanticamente mais relevantes como os de realização morfológica na L1 que se assemelhem aos seus correspondentes na LA são igualmente adquiridos antes dos morfemas menos salientes, ou que ocorram nas estruturas de superfície apenas para preencher requisitos estruturais da L2 (Wei, 1996:91).

Lopes (2001), P. Andrade (2003) e Baxter (2004) chamam a atenção para as demonstrações de Myers-Scotton (2001) acerca do modo como a CPL-var no SN envolve o recurso aos morfemas sistémicos estruturais (*early system morphems*) e aos

morfemas de nível funcional sistémico (*late system morphems*), os quais originam a redução de estruturas gramaticais durante a fase da aquisição (cf. nota de fim de texto 17). De facto, no caso de algumas línguas crioulas, e tal como analisaremos mais detalhadamente no ponto 4.1.2.6.2 do presente trabalho, alguns elementos do português, caso dos artigos definidos, poderão, de acordo com o sistema de pluralização das línguas africanas do grupo níger-congo atlântico, ter sido fonologicamente interpretados como prefixos classificadores dos nomes e, depois, transpostos para as estruturas morfosintáticas do PtgL2 nos seus estágios iniciais de aquisição em situação de transmissão linguística irregular.

### **2.4.3. Modelo da mudança evolucionista da língua**

De acordo com os defensores deste modelo (Givón, 1998; Croft, 2000; Mufwene, 2001; Clements, 2009), a evolução da língua estabelece analogias com a evolução das espécies, constituindo-se de traços lexicais e estruturais (análogos aos genes em biologia) e sendo os conjuntos desses traços (*feature pools*) usados para codificar a comunicação humana (relações comunicativas) no seio das comunidades de falantes, ou seja, entre elementos que interagem rotineiramente num certo grupo, recorrendo, para tal, a uma determinada língua. Neste aspecto, e à imagem daquilo que é defendido pelos conexionistas, não existe um módulo neurológico específico para a aprendizagem da linguagem, pelo que esta não se configura como um conjunto de regras abstractas (Gass & Selinker, 2001:216), sendo antes determinada pelo comportamento de seres humanos individuais, enquanto elementos integrantes de grupos utilizadores de um idioma específico (Clements, 2009:210).

Para Givón (1984:41-42), à linguagem humana competem duas funções primordiais: (i) representar o conhecimento; (ii) transmitir o conhecimento representado. Partindo destes pressupostos, o autor propõe que a linguagem seja vista em termos de aproximação à evolução-adaptativa manifestada biologicamente, uma vez que a estrutura e a função da primeira, a exemplo do que sucede com as componentes da segunda (anatomia e fisiologia), se complementam e não podem ser observadas em separado. Por seu lado, o processo que decorre entre o estado FL0 e o estado FLf da aprendizagem da língua é consequência de uma incrementação gradual da aquisição.

Assim sendo, a língua é um sistema dinâmico e em constante mutação, que tem a finalidade de representar o conhecimento face ao propósito de comunicar. Em tal sistema coexistem, então, variadas estruturas em determinado momento da sua etapa evolutiva (Pienemann, 1998:7).

Não obstante, as distinções não se estabelecem apenas em termos de linguagem, já que são também extensíveis aos sistemas de comunicação. Efectivamente, em termos de propriedades estruturais, funcionais e cognitivas, tais sistemas evidenciam diferenças entre o estágios de comunicação pré-gramatical – a interacção é feita com recurso a léxico codificado e combinatórias de regras rudimentares – e de comunicação gramatical – a interacção é estabelecida com recurso a léxico e regras gramaticais elaborados –. O desenvolvimento que permite a evolução do código pré-gramatical ao gramatical só é possível com base nas repetições que acontecem no discurso, sendo este essencial para a compreensão da gramática (Hopper, 1987, 1988, 1998).

Os padrões de reconhecimento e o papel que desempenham na forma, função e mudanças da língua, envolvem a noção de protótipo porque, com recurso a ele, é possível perceber quer a frequência das ocorrências no uso da língua quer o modo como as estruturas e classes emergem nesta. Estudos estatísticos sobre a estrutura de orações absolutas no espanhol (López Meirama, 1997) evidenciam que cerca de 85% dos sujeitos em posição inicial são altos na hierarquia de definitude (SN's realizados como nomes próprios, pronomes pessoais ou nomes com um determinante deítico), enquanto 90% dos sujeitos não-iniciais são baixos na hierarquia (SN's sem determinante). Por outro lado, observações sobre a ordem dos constituintes no castelhano medieval mostram que os casos de sujeito pospostos são baixos na hierarquia, uma vez que só ocorrem quando o objecto é um demonstrativo ou um relativo (Suárez Fernández, 2008:293, nota 86). Portanto, no caso do espanhol, o tópico protótipo é alto em definitude, tal como a ordem protótipo do sujeito é alta pré-verbalmente.<sup>65</sup>

Croft (2000) e Mufwene (2001) constroem a noção de língua como espécie assente em propriedades comportamentais, da mesma forma que as espécies são definidas em biologia. A aproximação é feita sob a perspectiva genética das populações, considerando-se os aspectos sociohistóricos que determinam variação como a ecologia das línguas. No entender de Croft (2000:13-15), tal como as espécies irmãs isoladas

(*sibling species*), que se reproduzem separadamente, mas não se distinguem estruturalmente, e as espécies politípicas (*polytyp species*), que são estruturalmente distintas entre si, também uma língua detém propriedades estruturais essenciais e imutáveis que a identificam, distinguindo-a ou aproximando-a de outras. Nesta conformidade, independentemente da possibilidade de as espécies linguísticas se reproduzirem isoladamente, não são definidas por traços estruturais abstractos, mas antes por propriedades fundamentais compartilhadas pelos membros da sociedade que a utilizam.

A classificação taxionómica das espécies é baseada em semelhanças e diferenças dos traços estruturais das línguas, enquanto a classificação filogenética corresponde ao quadro teórico que caracteriza as populações. Desta forma, a associação filogenética de duas populações acontece em termos históricos, uma vez que a mesma só é possível caso algum traço inovador venha a ser compartilhado por ambas (Croft, 2005:15). Os traços que distinguem ou aproximam duas línguas podem ser do tipo cultural e lexical (p.e. o hindu e o urdu são dialectos idênticos da mesma língua, mas os seus falantes distinguem-se por questões religiosas), do tipo cultural, social e político (p.e. os falantes do mandarim e cantonês compartilham o mesmo sistema político, cultural, social e de escrita, mas falam diferentes dialectos da mesma língua) ou até de identidade nacional e política (p.e. o alemão e o holandês falados na fronteira dos dois países são mutuamente inteligíveis pelos habitantes raianos, apesar de serem assumidos como dialectos de línguas distintas, isto é, como variedades do alemão *standard* e do holandês *standard*).

No caso concreto do PE e do PA, duas línguas politípicas da família romance, as diferenças são de vários níveis, mas mais na oralidade do que na escrita. Fonologicamente, o sistema de sílabas átonas é mais reduzido no segundo do que no primeiro, enquanto as consoantes apicais [t], [d], [s] e [z], antes das vogais [+ altas] [ə] e [i], são palatalizadas em [k], [g], [j] e [ç] no PA (Ferraz, 1979:22). Por outro lado, na morfossintaxe dos almorarifanos acontece variação na concordância de género e número no SN, enquanto os pronomes pessoais sujeito podem ser usados com função de objecto (exemplo [148]). Também a colocação dos clíticos varia nas duas línguas, uma vez que, no PA, podem ser colocados pré-verbalmente (exemplo [151]), enquanto no PE o são pós-verbalmente. Outra das distinções entre as duas línguas verifica-se a nível do

sistema de conjugação verbal, mais elaborada no PE do que no PA (exemplos [136], [137] e [138]). Note-se que as diferenças se estendem ainda ao plano lexical (exemplo [145], nota de fim de texto 46). No entanto, apesar das muitas distinções entre as duas variedades, ambas são consideradas como elementos integrantes de uma única língua, em virtude de Portugal e São Tomé e Príncipe deterem traços culturais e sociais estreitos. Por outro lado, se se levar em consideração que as características apontadas para o PA são também válidas para determinadas variedades do PB, então estas configurarão casos de espécies irmãs isoladas (*sibling species*) relativamente à fala de Almoxarife.

Assim, a interação comunicativa depende não apenas das similaridades estruturais das variedades faladas mas também do comportamento social dos falantes. Este aspecto determina que a definição de língua assente mais em critérios sociais do que linguísticos (Chambers & Trudgill, 1980:3). Nesta perspectiva, se a espécie biológica é definida como um conjunto de indivíduos isolados que se reproduzem, por um lado, e se a espécie linguística é caracterizada como um conjunto de indivíduos isolados que comunicam entre si, por outro lado, então as noções biológicas de raça e *deme* geográficas corresponderão às redes sociais de um determinado grupo de indivíduos isolados que comunicam entre si, e não tanto com indivíduos externos à sua rede social (Clements, 2009:8).

Para Mufwene (2001:151-152), a língua é como um parasita, ou seja, um género de espécie simbiótica, que necessita de um hospedeiro (o falante) para com ele poder coexistir. Em adição, a transmissão linguística tem relevância secundária no seu modo vertical, ou seja, de pais para descendentes, mas primordial na sua forma horizontal. Esta transmissão assemelha-se à das plantas consideradas superiores, ou seja, à que ocorre de forma poliplóide, levando as línguas a poderem ter uma origem múltipla, marcada por estágios iniciais incipientes na sua génese, como acontece com os pidgins, que geralmente se desenvolvem de modo dinâmico e recorrente em apenas duas ou três gerações. Paralelamente, este desenvolvimento é acompanhado de mudanças genéticas e epigenéticas, conduzindo a uma extensa reestruturação em todos os níveis do genoma, e que originam repadronização, novos padrões de expressão genética, invasão intergenómica e evolução coordenada. Um exemplo de que a transmissão vertical não é

relevante para a evolução linguística é concedido pelo pidgin falado pelos africanos transplantados para o sudoeste de Portugal nos finais do século XV e princípios do século XVI, que não evoluiu para crioulo (Clements, 2009:43-48). Daí inferir-se também que, muitas vezes, não são as crianças a fonte das inovações que activam a variação conducente à mudança. Por conseguinte, a emergência da variação em L1 e L2 ancora em factores diversos e envolvendo vários tipos de desencadeamento (*bootstrappings*).<sup>66</sup>

Tendo em conta que os sistemas complexos emergem de modo simples e que, pouco a pouco, vão desenvolvendo complexidades (Ellis & Schmidt, 1997), conclui-se que as estruturas sintácticas se processam com base no léxico. Trabalhos sobre aquisição verbal (p.e. Bresnan, 1982, 2001; Langacker, 1987; Goldberg, 1995, 1999) têm evidenciado a forte relação entre significado, por um lado, e estrutura argumentativa dos verbos e estrutura sintáctica, por outro lado. A relação entre aquisição vocabular e aquisição gramatical é tão evidente que, segundo os defensores do modelo evolucionista, a segunda não será viável de modo independente, antes configura um desenvolvimento do léxico, sendo este possível apenas através da frequência e rotina de uso dos verbos em questão. Por conseguinte, a gramática emerge de modo primário durante a aquisição, sendo afectada pela combinação do *input* linguístico, pelas exigências das situações comunicativas e por constrangimentos cognitivos.

No caso específico da ASL, estes aspectos aliam-se aos sistemas de marcação universal para orientarem o tipo de desenvolvimento e forma que determinado pidgin ou crioulo vai evidenciar (Clements, 2009:26). Estudos levados a cabo sobre a escolha das formas copulativas em CP's e crioulos de base lexical espanhola revelam que os falantes dão primazia às estruturas do tipo CV (p.e. a forma *sã* em vez de *é*, no maquista), em virtude da sua força lexical (Clements, 2009:22). Neste caso, a saliência perceptual impõe-se à frequência na LA, fazendo com que uma estrutura não-marcada universalmente, por ser mais facilmente processável (Thomason & Kaufman, 1988:331), seja responsável pela emergência da estrutura silábica de uma variedade que nasce por contacto linguístico. Por outro lado, as relações de dominância entre os grupos de falantes de diferentes línguas em contacto podem também ser determinantes para que traços relativamente marcados, como a flexão morfológica, se instalem nas variedades

submetidas a contacto. Estes aspectos conduzem-nos, assim, às questões das transferências e empréstimos (Thomason & Kaufman, 1988:73-74) por nós debatidas nos pontos 1.1 e 1.3.2.3.1 do presente trabalho, já que ambas se reflectem na actual estrutura sintagmática do PA.

Após os considerandos tecidos em torno dos pressupostos teórico-metodológicos sobre aquisição de L1, aquisição bilingue e ASL, passaremos agora a dirigir a nossa atenção para os aspectos que determinam a concordância PL entre os elementos do SN no português, antes de abordarmos as questões em torno da aquisição massiva de L2's e, conseqüentemente, do português pelos falantes da comunidade de Almojarife.

## **2.5. Variação: pressupostos teóricos**

A Sociolinguística visa o estudo dos padrões de comportamento linguístico observáveis dentro de uma comunidade de fala, formalizando-os analiticamente com recurso a um sistema heterogéneo, constituído por unidades e regras variáveis (Lucchesi & Ribeiro, 2009:127; Lucchesi & Araújo, 2010). Partindo-se destes princípios, tem-se buscado uma resposta satisfatória para a questão da mudança linguística, enquanto fenómeno que alia determinados pressupostos incontornáveis (Labov, 1972a, 1974, 1982, 1994; Weinreich, Labov & Herzog, 2006 [1968]): (i) os sistemas linguísticos são heterogéneos e plurais, a exemplo do que acontece com as comunidades que os usam; (ii); não há mudança sem variação, mas pode ocorrer variação sem mudança; (iii) conseqüentemente, os padrões linguísticos registadas nos distintos momentos da variação reflectem o estágio e percurso da mudança em curso. Nesta conformidade, os elementos de uma determinada comunidade, apesar de conjugarem, entre si, normas e atitudes face ao uso da fala comunitária, não recorrem todos às mesmas formas linguísticas. Estas, no entanto, não deixam de compartilhar traços, os quais, por seu lado, marcam diferença relativamente aos traços utilizados noutras comunidades.

A exemplo do que sucede com qualquer outra língua, o PA regista também formas que correlacionam aspectos linguísticos e sociais usados heterogeneamente pelos seus falantes, isto é, pelos membros que compõem a pequena comunidade de Almojarife, São Tomé. Deste modo, determinados usos linguísticos produzidos pelos almojarifanos, ainda que distintos entre si, apresentam equivalência semântica a nível de vocabulário,

da sintaxe e morfossintaxe, do subsistema fonético-fonológico e no campo pragmático-discursivo (cf. Mollica, 2003a:9). A presença e ausência de marcas de concordância nominal em idêntico tipo de construções (exemplos [152a] e [152b]) ou de fonemas que podem ter diferentes realizações fonéticas que se alternam no mesmo contexto linguístico (exemplos [211] e [212]) ilustram bem a variabilidade linguística patente em Almojarife, sendo tais alternâncias sistemáticas, as variantes linguísticas, determinadas pela pressão que factores estruturais (as variáveis linguísticas) quanto sociais (as variáveis extralinguísticas) exercem, em simultâneo, sobre os usos da língua, aumentando ou diminuindo a sua frequência de ocorrência e impedindo que os mesmos sejam aleatórios (Tarallo, 1986:8). Consequentemente, estes usos em co-ocorrência e concorrência revelam a capacidade que a heterogeneidade linguística possui para não produzir alterações significativas nas mensagens que o locutores pretendem transmitir aos alocutários. Não obstante, os condicionalismos que actuam sobre as variantes em competição poderão determinar que uma delas apresente um maior uso em determinados contextos, podendo este ser analisado e previsto com recurso a metodologia estatística cientificamente apropriada.

Ao seleccionarmos como objecto de estudo a análise da variação registada num idioma que tem na génese da sua emergência uma situação de contacto entre línguas, estamos não só a considerar como socialmente relevante a linguagem da pequena comunidade de Almojarife mas também a atribuir à sua variação um papel que pode ajudar a dar contributo valioso para os estudos sociolinguísticos, nomeadamente no que concerne às questões relacionadas com a mudança linguística e factores que a determinam. Para tanto, procuraremos fazer a descrição estatisticamente fundamentada da CPL-var no SN do PA, isto é, tentaremos analisar, apreender e sistematizar como este fenómeno se caracteriza de acordo com as propriedades da língua, por um lado, e se compromete, a nível do seu *status* positivo ou negativo, a nível do sistema linguístico da comunidade, por outro lado. Para procurarmos entender este grau de comprometimento, calcularemos a influência que cada factor, interno ou externo ao sistema linguístico de Almojarife, possui na realização da CPL-var no SN do PA. Estes aspectos permitirão detectar se as variantes em competição na realização da pluralização a nível do SN configuram um caso de mudança em processo (verifica-se uma tendência de



predominância de uma variante sobre a outra) ou se, pelo contrário, manifestam uma situação de variação estável (o quadro da variação tende ainda a manter-se por longo período de tempo) ou de recuo na inovação. Portanto, para se concluir sincronicamente acerca do processo de mudança em diacronia que estará, ou não, actuando sobre o fenómeno que nos propomos analisar, será necessário avaliar o grau de estabilidade ou mutação patente nos usos alternativos (*evaluation problem*, cf. Labov, 1972a, 1982; Weinreich, Labov & Herzog, 2006 [1968]), detectando-se, com exactidão, quais as variáveis internas à língua (p.e. fono-morfossintácticas ou lexicais), inerentes ao indivíduo (p.e. idade ou sexo), de tipo social (p.e. escolarização) ou meramente contextuais (p.e. tensão discursiva), que estarão afectando sistematicamente a variante, no sentido de a difundir e implementar entre os falantes almoxarifanos (*actuation problem*, cf. Labov, 1972a, 1982; Weinreich, Labov & Herzog, 2006 [1968]). Ainda assim, é também importante realçar que a complexidade dos condicionamentos actuantes sobre os usos que determinam a variação contínua da CPL-var no SN do PA não possibilita que se prevejam todos os tipos de agentes que estabelecem correlação com as suas variantes linguísticas, ou seja, não é possível demarcar nitidamente as fronteiras em que a referida variação ocorre.

No PA, a estigmatização linguística e o grau de monitoramento maniqueísta certo/errado exercido sobre alguns dos seus falantes complementaram-se no sentido de parametrizarem determinadas formas linguísticas, que tinham sido adquiridas informalmente, ao encontro do padrão da variedade de português mais culta falada na ilha. No entanto, concorrendo com estas são também realizadas algumas variantes motivadas, que constituem inovações relativamente ao PE. Estas, por seu lado, não deixam também de ser realizações coesas, configurando padrões estruturais e estilísticos que estão na base da caracterização do idioma usado pela comunidade almoxarifana. Portanto, estamos perante variações que, de certa forma, escaparam à monitorização e são usadas regularmente num eixo diatópico e diastrático marcado por grau considerável de isolamento geográfico e social. Desta forma, as variações acabaram por se implementar com padrões cujas tendências são preditivas e mensuráveis, já que não deixaram de se estruturar de acordo com as propriedades sistémicas das línguas, que

impõe também restrições categóricas às variantes, impedindo que as mesmas se tornem aleatórias (Tarallo, 1986:8).

No presente trabalho, a Teoria da Variação abordada instrumentaliza a análise sociolinguística e quantitativa do tipo laboviano, já que a mesma se tem revelado coerente e metodologicamente eficaz em diversos estudos anteriores que pretenderam efectuar descrições linguísticas numa perspectiva idêntica à que adoptámos. Todavia, e como temos vindo a sustentar no presente capítulo, não deixaremos de ter igualmente em consideração outros pressupostos teóricos, conquanto os mesmos dêem contributo importante para a análise que pretendemos levar a cabo.

No caso concreto dos factores estruturais, por exemplo, dever-se-á ter em conta o seu encaixe na estrutura relevante (cf. Naro, 2003a:16), o qual poderá originar a falta de concordância de número quando acontece maior ou menor diferença fónica entre as formas singular e PL. Face a questões como esta, o desafio em Teoria da Variação será determinar o quanto cada um deste tipo de oposições concorre para a realização de uma ou de outra variante em competição. Os dados orais do nosso *corpus* exibem categorias morfológicas em que o fenómeno da concordância PL ocorre com ou sem marca formal em diferentes posições na cadeia sintagmática, dificultando a verificação do efeito das ditas categorias de um modo isolado, isto é, sem se recorrer, em simultâneo, à observação do seu posicionamento na estrutura frásica. Este aspecto alerta para a impossibilidade de se medir o efeito de uma determinada categoria na CPL-var do SN do PA sem levar em consideração o efeito de outras categorias que, com ela, concorrem para a motivação do fenómeno em observação.

### **2.5.1. Variáveis extralinguísticas, variação e mudança**

Agentes externos à língua, como marcadores regionais de comunidades geograficamente circunscritas ou formas usadas especificamente por determinados estratos sociais, detêm um papel importante no controlo da variação, impedindo que esta seja caótica e aleatória, por um lado, para se tornar regular, sistemática e previsível, por outro lado. Contudo, estas variáveis não operam isoladamente, antes co-actúan com as variáveis linguísticas e também com outras variáveis extralinguísticas. De entre as segundas, variáveis como a idade do falante, seu nível socioeconómico ou de

escolarização, ou ainda acesso que ele tem aos media, poderão determinar a ocorrência de distintos padrões linguísticos (popular e culto) no seio da comunidade em que está inserido, como sucede em Almoхарife.

Embora sejam expressivos os estudos que correlacionam diferentes variáveis extralinguísticas com fenómenos de uso na fala de diferentes comunidades, o debate acerca dos seus efeitos na padronização do registo apresentado está longe de se encontrar esgotado. Observando o efeito que factores não linguísticos exercem sobre formas estigmatizadas do inglês vernacular de comunidades negras urbanas, Labov (1972b) concluiu que tais formas resultavam de preconceitos relacionados com as pressões étnica, social e escolar. Por seu lado, Sankoff, Kemp & Cedergren (1978), debruçando-se sobre estruturas alternativas do francês de Montréal, Canadá, com vista a determinar o maior ou menor grau de marcação social nas mesmas, verificaram que este era influenciado pela escolarização e estatuto profissional dos falantes. A importância das variáveis sociais na determinação de formas linguísticas usadas pelos mesmos falantes seria ainda coadjuvada e confirmada em outros estudos (p.e. Laberge, 1977; Clermont & Cedergren, 1979; Kemp, 1979, 1981).

Para além de observarem o efeito que factores sociais exercem sobre formas linguísticas, os estudos labovianos permitiram também que se fizessem inferências sobre o desenvolvimento diacrónico da língua, a partir dos resultados das análises sincrónicas, já que estas reflectem em tempo real, em um determinado ponto da estrutura da gramática da fala da comunidade, o processo da mudança em curso projectado no tempo aparente. O pressuposto da projecção do tempo real no tempo aparente é explicado por Chambers & Trudgill (1980:165) com recurso à variável social idade, alegando os autores que, por exemplo, a fala dos indivíduos de 40 anos de idade não só reflecte directamente, e no tempo actual, a produção dos indivíduos de 20 anos há 20 anos mas também pode ser comparada com a dos falantes de 20 anos da actualidade. Caso ocorram discrepâncias entre os usos das faixas etárias com 20 e 40 anos, essas discrepâncias reflectem o progresso registado nas realizações durante o período de tempo que separa as duas faixas etárias. Ainda assim, é preciso considerar que os sistemas linguísticos não são sempre perfeitamente estáveis nem totalmente predizíveis quanto às tendências de mudanças futuras, já que estas podem ser afectadas por

processos inexistentes no momento do estudo, mas que venham a ocorrer posteriormente (Labov, 1981:177).

Ainda no que concerne à variável social idade, podem registar-se duas situações (Chambers & Trudgill, 1980:91-93): (i) se as faixas etárias intermédias apresentarem o maior uso das formas inovadoras, o padrão curvilíneo denuncia estar-se perante uma situação de variação estável; (ii) se a aplicação da inovação incidir mais fortemente nos estratos etários mais novos, o padrão inclinado evidencia uma mudança em progresso. Todavia, estas inferências terão que ser sustentadas pelos resultados de outras variáveis sociais, como por exemplo, a escolarização.

Assim, no que às variáveis escolarização e/ou actividades de letramento diz respeito, normalmente associadas aos indivíduos das classes mais altas, tem sido confirmada a influência de ambas no sentido da aquisição da norma culta (p.e. Bortoni-Ricardo, 1994; Silva & Scherre, 1998 [1996]; Godinho, 2005; Silva, 2010), configurando-se situações de variação estável. Porém, alguns estudos têm revelado que o emprego de determinadas estruturas linguísticas por parte de falantes das classes média ou baixa, apesar de pertencerem a grupos mais escolarizados, pode apresentar flutuações, não se moldando ao sistema de ensino (Silva & Scherre, 1998 [1996]). Este aspecto revela não apenas que se está perante falantes que tendem a impulsionar processos de mudança mas também que, afinal, o efeito de indicadores sociais sobre o perfil sociolinguístico dos falantes é bastante mais complexo do que à primeira vista pode parecer (Mollica, 2003b:29). De facto, numa sociedade linguisticamente complexa, como a de São Tomé, serão imensos os indicadores sociais com influência nas formas de uso, competindo ao investigador seleccionar aqueles que mais impacto poderão ter sobre a variação que se pretende observar.

Nas sociedades mais industrializadas, a classe social parece ser o factor mais marcado linguisticamente (Chambers, 2002 [1995]), podendo os seus índices ser avaliados em função de aspectos como a ocupação, a educação ou o poder económico. Contrariamente, círculos sociais fechados, localizados e economicamente mais desfavorecidos, como acontece com a comunidade de Almojarife, tendem a incrementar relações familiares, de vizinhança ou de bairro. De qualquer forma, seja nuns seja nos outros, há tendência para se constituírem estereótipos, pelo que os índices de classe

social passam tanto por factores objectivos como por factores subjectivos. Paiva & Scherre (1999), em estudos pioneiros sobre usos do PB (PEUL – Programa de Estudos sobre os Usos da Língua), tentaram determinar a correlação estabelecida entre variáveis não convencionais, como a origem social, os bens materiais e os bens culturais, por um lado, e os usos linguísticos prestigiados e não prestigiados, por outro lado. Não obstante, os resultados levantaram a possibilidade de tais categorias não poderem ser medidas com recurso a critérios linguísticos, já que a língua é apenas mais um no conjunto dos elementos que constituem o património pessoal do indivíduo, enquanto membro integrante de uma determinada comunidade.

Do exposto é possível concluir que ocorrem incertezas quanto à possibilidade de certos factores extralinguísticos representarem uma opção válida como identificadores e determinadores do estatuto social de determinado indivíduo ou comunidade. Paralelamente, dúvidas parecem não existir no que concerne a outros factores, enquanto limitadores ou impeditivos da ascensão social do indivíduo e, conseqüentemente, de uma maior aquisição dos usos linguísticos de maior prestígio. Este é exactamente o caso da pequena comunidade de Almojarife, submetida a uma situação de isolamento cada vez mais acentuado, a qual determina que aspectos como as condições miseráveis de habitação, o acesso precário à saúde e escolarização ou a fome impeçam, aos seus membros, o acesso ao estatuto de cidadãos plenos.

### **2.5.2. Variáveis linguísticas e contexto**

Os trabalhos pioneiros de Labov (1972a, 1972b) sobre a centralização de ditongos em Martha's Vineyard, costa ocidental dos Estados Unidos, e sobre o uso do /r/ em Nova Iorque, com recurso a dados colectados em situações reais de comunicação, concederam primazia à observação do modo como condicionamentos externos à língua interferiam na produção dos referidos fenómenos. Centralizando-se atenção nos factores sociais, com recurso a métodos de análise quantitativa, buscava-se não só uma teorização oposta à preconizada pelos princípios generativistas mas procurava-se evidenciar também que os usos linguísticos das comunidades são heterogéneos. Nos estudos referidos, notava-se também uma apetência pela selecção de fenómenos morfofonológicos como objecto de observação. Este aspecto justificava-se, já que, em colectas de dados reais, os sons e

fonemas se aprestam como entidades que garantem um número de realizações mais elevado do que o de qualquer outra entidade linguística. Paralelamente, os estudos com recurso a variáveis de natureza fonológica vinham já de longa data, sendo uma constante em trabalhos estruturalistas, embora sem recurso a métodos quantitativos.

Tendo-se evidenciado eficientes para a observação da variação sistemática e quantificável dos fenómenos fonológicos, os estudos labovianos deram o mote para a análise, na mesma linha, da variação registada na sintaxe, na semântica e até no discurso, áreas onde, apesar de tudo, os fenómenos não são tão recorrentes como na fonologia. Além do mais, a busca de formas sintácticas alternantes esbarrava, segundo alguns autores, na impossibilidade de se manter idêntico significado em ambas as realizações. Lavandera (1984) propõe, então, que se tratem as variáveis linguísticas levando em conta o enfraquecimento da equivalência semântica e dando primazia à “comparabilidade funcional”. Portanto, frases que transmitem a mesma intenção comunicativa, ainda que distintas (p.e. *Está calor aqui dentro.* > *Por favor, ligue o ar condicionado.*) devem ser tratadas como variantes de uma mesma variável. Contudo, outros estudiosos dos fenómenos da variação (Weiner & Labov, 1983) entendem que as condições para análise das variantes sintácticas se devem restringir ao valor referencial, como sucede nas alternâncias das vozes activa e passiva, em que a noção de verdade não sofre alteração.

Seja num caso seja no outro, o estudo variacionista ver-se-á sempre confrontado com diferenças conectadas aos valores semânticos, aos contextos discursivos e ao tipo de interacção entre locutor e alocutário. Todavia, tais diferenças poderão ser controladas com recurso a factores que o estudioso entende estarem correlacionados com o fenómeno em análise. Portanto, é possível determinar os referentes portadores de informação nova ou velha, testando o peso que exercem no uso das formas comunitárias e aquilatando a sua influência na mudança em curso. Ainda assim, tal só poderá acontecer determinando se o traço discursivo ou semântico a observar constitui uma das variáveis admitida como escolha alternativa num mesmo contexto, isto é, a variação.

A questão da escolha contextual ganhou projecção em análises variacionistas, sendo determinante para o estudos das variantes sintácticas, o que significa dizer que estas não podem ser observadas a partir de produções linguísticas que configuram realizações

isoladas (Klein-Andreu, 1983; Bentivoglio, 1987; Kroch, 1989a). Estruturas em que acontece o deslocamento de um SN para o início da frase ou ocorrem realizações com marcadores discursivos (p.e. *sabe?*, *assim*, etc.) dificultam a decisão sobre o que constituirá, de facto, o conjunto de variantes a analisar. Nesta conformidade, estudos recentes sobre variação não se limitam a seleccionar fenómenos gramaticais de contornos perfeitamente marcados, antes observam domínios abrangentes. É o que sucede com variantes que se manifestam através de distintas realizações linguísticas, (p.e. *nós* > *a gente* > *a malta* > *o pessoal*). Nestas situações, não se estabelece “*apenas uma escolha entre duas alternativas, mas uma escala de possibilidades para aquilo que poderíamos chamar de um domínio funcional, e aí estudar a variação. Neste caso, o significado referencial seria preservado, mas a questão do contexto exigiria um cuidado especial, ao levar em conta a situação discursiva*” (V. Silva, 2003:71).

Em virtude de enxertarem um no outro, significado e contexto determinaram que estudos variacionistas do PB refinassem a observação das variantes sintáticas e discursivas, levando a avanços consideráveis nas áreas da linguística textual e linguística funcional, as quais seleccionam fenómenos de âmbito e estrutura superiores ao das sentenças: (i) o *status* informacional dos itens lexicais (Braga, 1984; Berlinck, 1989) e oracionais (Paiva, 1991); (ii) aspectos da coesão textual (Omena, 1978, V. Silva, 1988), relacionados a fenómenos de nível morfossintático e oracional; (iii) relação semântica de contraste, aplicada a fenómenos morfossintáticos (Mollica, 1984; V. Silva, 1988), oracional (Oliveira & Braga, 1997) e discursiva (Gryner, 1990). Paralelamente, o interesse pelos fenómenos não-fonológicos internos à língua ganhou novo fôlego à luz da Teoria da Variação, até porque estes constituem um desafio acrescido, quando se trata de os correlacionar às variáveis sociais. Quanto aos resultados, têm evidenciado que este tipo de fenómenos não está tão sujeito à influência das variáveis sociais, alertando, consequentemente, para o facto de a sua variação advir mais da pressão que é exercida pelos níveis sintático, semântico e discursivo-pragmático (Braga, 2003).

### **2.5.3. Modelos matemáticos para reprodução do efeito global dos dados empíricos**

Nos estudos pioneiros sobre CPL-var no SN do PB (Braga, 1977; Scherre, 1978), postulou-se que a maior distinção fónica singular/plural favorecia a inserção de marca

formal de PL, um aspecto que, no entanto, era contrariado pelos valores estatísticos enviesados, já que o percentual de marcação era bem mais elevado nas formas em que a referida oposição era simples, como acontecia com os artigos ou demonstrativos (p.e. *o>os* ou *este>estes*). Assim, só observando a forma como as distinções fónicas co-actuaem com a posição linear de cada elemento flexionável do SN é que foi possível perceber o modo como ambos os factores intervêm na configuração da regra variável de marca de PL. De facto, as altas frequências reveladas pelas oposições simples estavam a inflacionar os resultados em virtude de se encontrarem concentradas na primeira posição da cadeia linear e não por razões que pudessem ser conectadas à sua estrutura morfológica. Removendo os dados das ocorrências nesta posição foi então possível confirmar que, de facto, a maior distinção fónica singular/plural favorece a inserção de marca formal de PL e que, conseqüentemente, *“as frequências brutas, embora concretas e intuitivamente bastante ‘reais’, podem ser falaciosas, porque seu cálculo não leva em conta as inter-relações existentes entre as categorias que atuam numa regra variável”* (Naro, 2003a:19).

As tentativas pioneiras de prover um modelo matemático consistente com os dados linguísticos iniciaram-se com o modelo aditivo proposto por Labov (1969), no seu clássico estudo sobre a variação patente no uso da cópula e do verbo auxiliar *“to be”* por parte de falantes do inglês registado nos *“ghettos”* urbanos de diversas comunidades negras. No modelo laboviano, a frequência geral da marca em observação procurava representar a soma dos factores contextuais, isto é, a co-actuação dos factores oposição simples, posição, número total de categorias relevantes e média global de aplicação da variante sob estudo.

Posteriormente, Naro (1981) testaria o modelo para analisar a concordância variável na fala de adultos cariocas alfabetizados. Neste estudo, a média global de aplicação da variante servia como ponto de referência para os diversos factores, cada um dos quais registando maior ou menor efeito na regra da concordância, ou seja, aumentando ou diminuindo a frequência da variante em determinado contexto. Não obstante, o modelo representa apenas uma soma de números, que pode ultrapassar os 100% ou ir abaixo dos 0%, aspecto que inviabilizaria a sua aplicação na elaboração de uma hipótese que



definissem, com exactidão, a força da actuação conjunta de categorias presentes num dado contexto, de modo a reproduzir o efeito global da variação.

Antes de Naro (1981) ter testado o modelo, Cerdergren & Sankoff (1974), partindo do pressuposto que a *performance* é modelada pela conjugação de factores generativos e extralinguísticos, desenvolveram um modelo matemático a partir das propostas de Labov (1969), que visava testar, em registos do espanhol, francês e inglês, as frequências preditas pelas regras variáveis. O modelo, denominado de aplicativo de multiplicação, propunha que a observação da variação se efectuasse a partir de probabilidades, e não de frequências. Assim, a variação representa o produto da actuação simultânea de factores independentes, e não a soma destes. No entanto, o modelo apenas se revelaria satisfatório para observar a actuação conjunta de dois factores altamente desfavorecedores (o modelo multiplicativo de aplicação das regras linguísticas), já que, para um estudo apropriado da co-actuação dos factores favorecedores, seria necessário recorrer a um modelo multiplicativo de não-aplicação das regras linguísticas.

Com vista a ultrapassar este impasse, Rousseau & Sankoff (1978) elaboraram um modelo logístico que conjuga propriedades dos três modelos anteriores (aditivo, multiplicativo de aplicação e multiplicativo de não-aplicação) e se destinava a observar, estatisticamente, o modo como as categorias estabelecem inter-relações em áreas distintas da linguística, como sucede na Biologia. Por outro lado, o modelo apresentava a vantagem de se encontrarem também respostas satisfatórias a partir da desvinculação dos factores entre si. Este aspecto revela-se importante porque a variação em direcção à mudança ocorre em virtude de, num determinado momento do tempo, uma das duas formas alternativas (a inovadora vernacular) passar a ser favorecida em detrimento da outra (a conservadora). Estudos empíricos de determinados fenómenos linguísticos (p.e. Silva, 1982; Kroch, 1989a) têm evidenciado que o processo de substituição conhece três etapas distintas, seguindo o trajecto de curva em  $\sim$ : (i) o estágio inicial de mudança, em que uma das ocorrências em uso se sobrepõe lentamente à outra; (ii) o estágio intermédio de mudança, em que a ocorrência que se sobrepõe evidencia um aceleração notório de uso; (iii) o estágio final de mudança, no qual a aceleração se retrai, até se consumir a mudança, traduzida no desaparecimento da forma em competição menos usada. Como o modelo logístico permite recorrer a probabilidades do

*input* (a média geral de aplicação da regra), é possível medir a tendência de presença da variante, isolando-a do efeito dos factores. Portanto, a metodologia da Teoria da Variação, adoptada no âmbito do presente trabalho, com recurso ao modelo de regressão logarítima múltipla VARBRUL (ponto 3.2 do presente trabalho), constitui uma ferramenta poderosa e fiável quando se trata de observar e descrever o dinamismo da variação sincrónica de determinados fenómenos linguísticos, enquanto retrato dos processos de mudança no tempo real. No entanto, as limitações da ferramenta poderão esbarrar nas do próprio pesquisador, caso este não detecte os factores relevantes, não levante e codifique acertadamente os dados empíricos e não interprete correctamente, à luz do quadro teórico disponível, os resultados obtidos.

Face ao exposto, temos que a caracterização de um processo de variação estável ou de mudança em curso independe dos resultados isolados de cada variável em observação. Nesta conformidade, para caracterizarmos a actual situação linguística do PA, no que concerne à CPL-var no SN, procuraremos ancorar a coerência argumentativa na representação que elaboraremos do processo como um todo, a partir das evidências empíricas que nos forem fornecidas pelos resultados das distintas variáveis (cf. Lucchesi, 2004b; Lucchesi & Araújo, 2010), as quais irão ser constituídas e pormenorizadamente descritas no capítulo 3. Deste modo, acreditamos vivamente que, superadas as dificuldades evidenciadas neste ponto do nosso trabalho, iremos também dar o nosso contributo para o avanço da Teoria Variacionista, já que a discussão em torno dos resultados alcançados com recurso ao tratamentos dos dados empíricos produzidos pelos informantes de Almojarife permitirá que aprofundemos igualmente questões de natureza teórica. Para tanto, não nos limitaremos a fazer dos factores internos um ponto exclusivo do nosso debate nem relacionaremos a sua variação restritamente às questões de nível sintáctico, semântico e discursivo-pragmático, uma vez que procuraremos identificar também até que ponto os factores sociais do tipo clássico (p.e. idade, sexo do informante ou anos de escolaridade) se aliam a estes para determinarem o uso da CPL-var no SN da fala de Almojarife, ou seja, para moldar o fenómeno que constitui o objecto principal da análise sociolinguística desenvolvida no âmbito do presente estudo.

## 2.6. Concordância de número PL a nível do SN

### 2.6.1. Posicionamento e flexão em número dos constituintes do SN: visões descritivistas

Para elaboração do ponto 2.6 e seus subpontos, tomámos como base de orientação o capítulo 3 do trabalho de Godinho (2005:112-165) e leituras sugeridas no mesmo (p.e. Pérez-Pereira, 1989; Cerqueira, 1994; López-Ornat, 1997; Lléo, 1997; 1998, 2001; Faria, Freitas & Miguel, 1998, 2001; Marrero & Aguirre, 2003; Name & Corrêa, 2003), tendo ambos sido complementados quer com dados e conclusões de outras leituras quer com inferências nossas. Relativamente às generalidades acerca da flexão em número do português, as mesmas estão contempladas em qualquer gramática prescritiva (p.e. Cunha & Cintra, 1997 [1984]:181-189, para os nomes; Cunha & Cintra, 1997 [1984]:252-253, para os adjectivos; Bechara, 1999 [1928]:117-131, para os nomes; Bechara, 1999 [1928]:145-146, para os adjectivos). Godinho (2005:112-119) lista alguns exemplos para confirmar que, no próprio português padrão, a marcação do PL com o sufixo *-s* acontece de forma aparentemente sistemática. A listagem, apesar de exaustiva, não é completa. Assim, a propósito da flexão em número de nomes e adjectivos, julgamos ser importante chamar a atenção para mais algumas especificidades da norma padronizada não contempladas na referida análise. Deste modo, nem todos os substantivos são passíveis de marcação PL em português, como acontece com os nomes que terminam em *-s* átono (exemplo [164a]). De igual forma, também determinados monossílabos com igual terminação se mantêm invariáveis (exemplo [164b]),<sup>67</sup> o mesmo não acontecendo com termos que finalizam em *-s* tónico (exemplo [164c]):

- [164] PTG:
- |    |                 |   |                    |
|----|-----------------|---|--------------------|
| a. | <i>o atlas</i>  | > | <i>os atlas</i>    |
| b. | <i>o cais</i>   | > | <i>os cais</i>     |
| c. | <i>o ananás</i> | > | <i>os ananases</i> |

Analisando o efeito do *contexto fonológico posterior* sobre a *saliência fónica* dos itens terminados em *-s* no singular, Braga & Scherre (1976), Lopes (2001) e Baxter (2004) referem que estes são responsáveis pela inibição da pluralização no elemento seguinte, já que os falantes fazem uma interpretação analógica do /s/ final com a marca de PL, deixando de aplicar o PL depois, por este se revelar redundante (Kiparsky,

1972:195). Contudo, este é um fenómeno que não ocorre no PA, como procuraremos demonstrar no ponto 5.2.3.10.1 do presente trabalho.

No português padrão, alguns nomes, classificados como não-contáveis, não possuem a noção de separação em partes diversas para uma determinada matéria ou para uma ideia abstracta. Recebem a denominação de *singularia tantum* porque, apesar de possuírem forma PL, só se usam, normalmente, no singular. Estão incluídos neste critério: (i) os nomes que pertencem a categorias semânticas que exprimem a ideia de matéria, como os metais (exemplo [165a]); (ii) os nomes que transmitem noções abstractas, como os pontos cardeais (exemplo [165b]), as qualidades do ser humano (exemplo [165c]), <sup>68</sup> as ciências (exemplo [165d]), as artes (exemplo [165e]) ou as religiões (exemplo [165f]); (iii) os nomes que, traduzindo a ideia de pluralização, fornecem também o conceito de não poderem ser contados, ou seja, os nomes massivos (exemplo [165g]):<sup>69</sup>

[165]	PTG:			
a.	<i>o ouro</i>	>	–	
b.	<i>o norte</i>	>	–	
c.	<i>a gentileza</i>	>	–	
d.	<i>a história</i>	>	–	
e.	<i>a pintura</i>	>	–	
f.	<i>o budismo</i>	>	–	
g.	<i>a relva</i>	>	–	

Os não-contáveis incluem ainda alguns nomes colectivos que, na forma singular, encerram a noção de pluralidade, fazendo referência a uma colecção ou conjunto de objectos. Entre os colectivos há duas categorias: os universais, que não são contáveis e só se pluralizam se houver necessidade de fazer referência à classe (exemplo [166a]); e os particulares, que se contam e são passíveis de pluralização (exemplo [166b]):<sup>70</sup>

[166]	PTG:			
a.	<i>O casario de Macau.</i>	>	<i>O casario de Macau e da Taipa.</i>	
b.	<i>A turma B do segundo ano.</i>	>	<i>As turmas B e C do segundo ano.</i>	

Alguns nomes são usados no singular com o valor de PL genérico, visto encerrarem representações abstractas de pluralização (Pereira, 1984:46), como acontece com as palavras *homem* ou *coisa*:

- [167] PTG:  
 a. *O homem é um ser inteligente.*  
 b. *A coisa agora está preta.*

Não obstante, e como referimos no ponto 2.6.1, Scherre (2001:98-99) adverte para a possibilidade de a representação abstracta de pluralização traduzir realizações com carga semântica de pluralidade que é percebida pelos falantes a partir do contexto, originando-se, assim, a não-realização da marca formal de PL. Sobre este tipo de nomes nos pronunciaremos, de novo, no ponto 4.1.2.4.6.

Outra situação de nomes usados no singular com o valor de PL genérico ocorre com os substantivos etnónimos (referentes a povos, tribos, castas ou agrupamentos em que é prevalente o conceito de etnia), apesar de admitirem a forma PL:

- [168] PTG:  
*o bantu* > *os bantu*

Como apontámos no ponto 1.2.1.1, a pluralização destes nomes apresenta uma estrutura idêntica à das variedades africanas e brasileiras de português reestruturado. Contudo, no caso do português, não estamos perante uma marcação que define o padrão PL sistemático da língua, ao contrário do que sucede nas referidas variedades.

Continuando a observar formas de marcação plural no português, alguns empréstimos lexicais acolhidos por este idioma conservam também o PL original:

- [169] PTG:  
*o curriculum* > *os curricula*

Determinados nomes podem possuir duas ou três formas de PL, como acontece com muitos dos que terminam em *-ão*:

- [170] PTG:  
 a. *o vulcão* > *os vulcões* > *os vulcões* > –  
 b. *o sacristão* > *os sacristãos* > – > *os sacristães*  
 c. *o guardião* > – > *os guardiões* > *os guardiães*  
 d. *o aldeão* > *os aldeãos* > *os aldeões* > *os aldeães*

A análise da influência da variável independente *saliência fónica* na marcação PL no SN do MRJ revelou que os itens finalizados em *-ão* tendem a inibir a pluralização,

tendo Scherre (1988:124) atribuído esta à incerteza provocada nos falantes pelas distintas formas de se pluralizarem os referidos elementos. Estas questões encontram-se debatidas no ponto 5.2.1 do presente trabalho.

A pluralização de alguns nomes diminutivos do português, com recurso ao sufixo –*zinho*, apresenta grau de dificuldade acrescido para os falantes em aquisição do PtgL2, já que estes itens flexionam em número colocando os dois elementos no PL, mas suprimindo o –s do nome grau normal:<sup>71</sup>

[171] PTG:  
*o cãozinho* > *os cãesinhos*

Em português existem também nomes que, por norma, se utilizam unicamente na forma PL, isto é, os *pluralia tantum*:

[172] PTG:  
*as alvíssaras* > –

Fazemos ainda referência a alguns nomes, por norma concretos, cuja flexão em número acarreta uma alteração semântica do valor da palavra:

[173] PTG:  
*a costa* (orla marítima) > *as costas* (parte do corpo humano)

Com substantivos próprios<sup>72</sup> que configuram nomes e apelidos é possível a pluralização quando se deseja fazer referência a um conjunto de indivíduos que possuem um mesmo antropónimo (exemplo [174a]), ou se quer mencionar, não um indivíduo, mas vários que detêm características semelhantes à da individualidade referida (exemplo [174b]):

[174] PTG:  
a. *Os Antónios constituíram uma Associação.*  
b. *Ainda há muitas **Teresas de Calcutá** por esse mundo fora.*

Os apelidos, por seu lado, podem também ocorrer no PL, caso façam referência ao conjunto de indivíduos de uma determinada família (exemplo [175a]), ou se pretenda reportar a obras ou objectos produzidos pelo detentor do apelido que os elaborou ou manufacturou (exemplo [175b]):

[175] PTG:

- a. *Os **Freires** têm muitas casas em Leiria.*
- b. *O Museu de Arte Moderna tem muitos **Picassos** em exposição.*

De particular relevância para o nosso estudo são os nomes metafônicos, visto agregarem dupla pluralização, isto é, um PL morfémico (inserção de *-s* final) e um PL fônico (abertura vocálica), como acontece com os nomes do seguinte tipo: *jogo>jogos*, *ovo>ovos* ou *olho>olhos*. De facto, é comumente aceite que os falantes em situação de ASL revelam tendência para fazer incidir um maior peso de marcação PL sobre os itens fonicamente salientes. Estes aspectos encontram-se debatidos no capítulo 5, no qual procederemos à análise das variáveis fonológicas do PA.

No que respeita à flexão em número dos compostos morfológicos, remetemos a informação pertinente para Villalva (2003a:928-929). No entanto, convém referir que a flexão em género e número não se limita aos nomes, já que outros elementos que constituem o SN, como os indefinidos, os quantificadores ou os adjectivos, podem também apresentar este tipo de flexões.

Como no português é exigida a concordância entre os constituintes que compõem o SN (nomes, determinantes, quantificadores e adjectivos), os falantes de L1's africanas do grupo níger-congo atlântico ou do santomense, em aquisição do português, são confrontados com dificuldades na aprendizagem da regra, o que vai originar variação na interlíngua. No que concerne à concordância entre o núcleo e todos os constituintes do SN no PE, verifica-se que os determinantes e os quantificadores não colocam problemas maiores, reservando-se ao adjectivo algumas questões mais pertinentes. Godinho (2005:114-115) remete para Casteleiro (1981) as explicações detalhadas sobre a sintaxe do adjectivo, mas menciona as três situações em que se dá a ocorrência do adjectivo na frase: (i) adjectivos que ocorrem antes e depois do nome com alteração de sentido; (ii) adjectivos que só ocorrem antes do nome; (iii) adjectivos que só ocorrem depois do nome. Contudo, a propósito da primeira situação, importa também referir que a flutuação da posição do adjectivo, pré ou pós-nominal, nem sempre acarreta alteração semântica. De facto, adjectivos que não encerram um traço mais informativo específico (avaliativo), são colocados antes ou depois do nome apenas por mera questão estilística ou com a intenção de focalizar:

- [176] PTG:
- a. *A situação **actual** do Primeiro-Ministro é periclitante.*
  - b. *A **actual** situação do Primeiro-Ministro é periclitante.*

Sobre as outras duas situações referidas, constata-se que os adjectivos unicamente marcados por um traço mais qualitativo (conotativo) ocorrem antes do substantivo, como acontece com o adjectivo *mero* (Godinho, 2005:115). Contudo, neste aspecto discordamos com o outro exemplo avançado, ou seja, o do adjectivo *simples*,<sup>73</sup> já que este pode também ocorrer em posição pós-nominal:<sup>74</sup>

- [177] PTG:
- a. *Só gostava de iogurtes **simples**.*
  - b. *Uma **simples** testemunha bastou para o condenar.*

Por seu lado, os adjectivos de característica informativa, ou seja, denotativa, (por exemplo, os que se referem a nacionalidades ou cores), só podem ser posicionados pós-nominalmente (exemplos [178a] e [178b]), excepto em ocorrências literárias, já que estas subvertem regras sintácticas por questões estéticas:

- [178] PTG:
- a. *Uma mulher **chinesa** / \*Uma **chinesa** mulher.*
  - b. *Uma estante **metálica** / \*Uma **metálica** estante.*

Note-se que alguns destes adjectivos, se colocados pré-nominalmente, mudam de traço sintáctico (exemplo [179a]), passando a adjectivos substantivados que exigem a anteposição de um determinante. Classificados como modificadores internos, estes adjectivos possuem um traço formal rígido que obriga a que ocorram depois do nome. Caso tal não aconteça, deixam de ser itens subordinados para se tornarem o elemento central do SN (exemplo [179b]):

- [179] PTG:
- a. *O céu **azul**.*
  - b. *O **azul** do céu.*

Já as situações que envolvem o uso de dois ou mais adjectivos no SN implicam um posicionamento ordenado, no qual apenas um poderá ser pré-verbal. Contudo, caso ocorram dois adjectivos de traço informativo, ambos terão de ser situados depois do nome (exemplo [180a]). De facto, a colocação de qualquer dos adjectivos em posição



pré-nominal torna a frase agramatical (exemplos [180b] e [180c]). Quanto à ordem entre os adjetivos, é hierarquicamente determinada pelo maior ou menor grau informativo que encerram em si, surgindo em primeiro lugar o mais informativo. Assim, se a ordem hierárquica dos adjetivos for alterada, a frase torna-se agramatical (exemplo [180d]):

- [180] PTG:
- a. *O comboio **eléctrico amarelo**.*
  - b. *\*O **eléctrico** comboio **amarelo**.*
  - c. *\*O **amarelo** comboio **eléctrico**.*
  - d. *\*O comboio **amarelo eléctrico**.*

Nos casos em que houver um adjetivo de traço mais denotativo (p.e. *castanha*) e outro de traço mais conotativo (p.e. *pesada*), o segundo poderá ocorrer em posição pós (exemplo [181a]) ou pré-nominal (exemplo [181b]). Se optarmos pela primeira hipótese, o adjetivo de traço mais conotativo terá que obedecer à hierarquia formal, isto é ser colocado depois do adjetivo de traço mais informativo, pois o contrário torna a frase agramatical (exemplo [181c]). A colocação do adjetivo mais denotativo em posição pré-nominal produz também uma frase agramatical (exemplo [181d]):

- [181] PTG:
- a. *Uma pasta castanha **pesada**.*
  - b. *Uma **pesada** pasta castanha.*
  - c. *\*Uma pasta **pesada** castanha.*
  - d. *\*Uma **castanha** pasta **pesada**.*

Situações há em que se verifica também a presença de um adjetivo de traço informativo (p.e. *enjoativo*) e outro que não altera o traço semântico com a deslocação para a posição pré-nominal (p.e. *intenso*). Nestes casos, o adjetivo que não altera o traço pode ser usado antes (exemplo [182a]) ou depois do nome (exemplo [182b]). No segundo caso, e também por razões hierárquicas, ocupa o segundo lugar da estrutura em virtude de possuir um traço conotativo, o que não acontece com o adjetivo de traço mais informativo. A deslocação deste adjetivo para a posição pré-nominal (exemplo [182c]) ou pós-adjectivo que não altera o traço semântico (exemplo [182d]), torna as frases agramaticais:

- [182] PTG:
- a. *A sopa tem um **intenso** cheiro enjoativo.*
  - b. *A sopa tem um cheiro enjoativo **intenso**.*
  - c. *\*A sopa tem um **enjoativo** cheiro intenso.*
  - d. *\*A sopa tem um cheiro intenso **enjoativo**.*

O ordenamento de um SN em que ocorra a existência de um adjectivo modificador interno que não altera o traço semântico se colocado antes do nome (p.e. *política*) e um outro adjectivo que também não altera o traço semântico com a deslocação para a posição pré-nominal, mas que não funciona como modificador interno (p.e. *actual*), é também hierarquicamente definido. Assim, o primeiro adjectivo terá que ocorrer sempre pós-nominalmente (exemplo [183a]), já que o seu deslocamento para antes do nome origina uma frase agramatical (exemplo [183b]). Como se vê ainda no exemplo [183a], caso os dois adjectivos ocorram depois do nome, o ordenamento hierárquico impõe que o adjectivo modificador interno ocupe a primeira posição, visto deter também um traço mais informativo. Pelo facto de o adjectivo que não funciona como modificador interno poder ser deslocado para a posição pré-nominal, a construção frásica do exemplo [183c] é gramatical:

- [183] PTG:
- a. *A situação **política** actual do país é estável.*
  - b. *\*A **política** situação actual do país é estável.*
  - c. *A **actual** situação política do país é estável.*

As questões que se prendem com o posicionamento do adjectivo no português poderão ter implicações na aquisição dos mesmos por parte de falantes das línguas do grupo níger-congo atlântico ou do santomense, já que, nestes idiomas, a posição do adjectivo detém rigidez formal, ocorrendo sempre em posição pós-nominal (ponto 1.3.2.2 do presente trabalho).

Após termos apontado algumas especificidades do SN do português, numa abordagem descritivista, importa também observar quer o funcionamento do fenómeno da concordância de número PL na referida língua quer o modo como se dá a aquisição das suas categorias funcionais. Este aspecto impõe-se, dado que a correlação entre a variação e a caracterização das estruturas sintagmáticas, que seguidamente iremos

abordar, ajudará a formular hipóteses (capítulo 3), relativamente às questões que estarão subjacentes à aquisição dessas mesmas estruturas.

### 2.6.2. O quadro teórico da concordância

A concordância de número detém regras bastante rígidas no português padrão,<sup>75</sup> podendo o seu sistema de flexão verbal, por exemplo, fazer com que o sujeito pronominal não seja realizado (Cunha & Sintra (1997 [1984]:494), identificando-se apenas pelos traços de pessoa e número da concordância (Raposo, 1992:43).<sup>76</sup> Segundo a visão generativista, mais concretamente de acordo com os pressupostos do Princípio da Projecção Alargado da Teoria X-barras (Grimshaw, 2005 [1991]), existe uma categoria SN na posição de sujeito das frases, pelo que, em português, este SN pode ter um nome como núcleo ou um pronome foneticamente realizado ou não, e que constitui uma categoria vazia sujeito ou *pro* (Campos & Xavier, 1991:20). Portanto, a concordância é um fenómeno de *interface* por excelência, uma vez que traduz uma manifestação morfológica que é consequência de uma relação sintáctica (Costa & Silva, 2006a:1) Este aspecto determina também que as manifestações de concordância entre diferentes elementos sintácticos sejam distintas nas diferentes línguas (Béjar, 2003; Sigurðsson, 2004), já que algumas destas não aplicam uma das propriedades mais salientes deste fenómeno: a do traço da redundância, que repete informação já expressa em determinadas categorias (Kiparsky, 1972:195). É o que acontece, por exemplo, no santomense, em que a marcação PL a nível do SN pode ocorrer apenas no elemento pré-nominal, mantendo-se inalterado o radical de todos os restantes elementos da cadeia sintagmática. Partindo do pressuposto que a ASL se dá a partir da gramática da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), as produções dos falantes do PA vão apresentar competição entre as gramáticas das línguas em contacto (santomense e PE). Desta forma, a CPL-var no SN do PA é bastante semelhante à de outras línguas que passaram por transmissão irregular, como o PT ou o PVB, nas quais todo o tipo de flexões continua a ser consideravelmente reduzido. Face a estes aspectos, torna-se legítimo questionar sobre o quanto da concordância pode ser estabelecido na gramática, ela mesma, e o quanto da concordância deve ser visto como

*interface*, isto é, como mero reflexo morfológico de uma relação sintáctica, para o qual a própria sintaxe não tem explicações a oferecer.

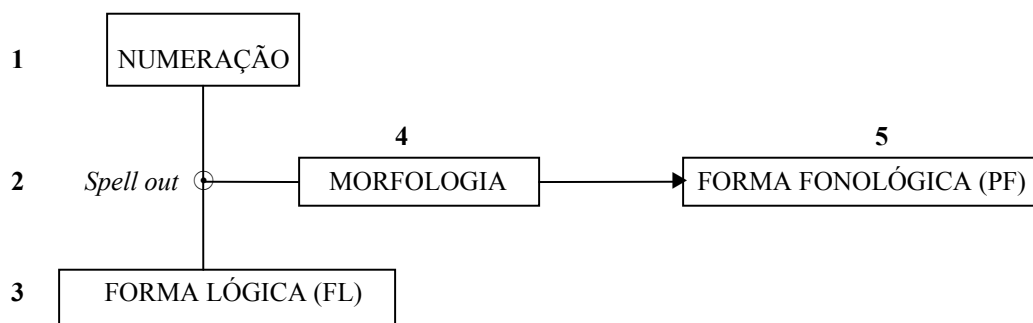
### **2.6.2.1. Concordância e núcleos funcionais**

Segundo a corrente tradicional lexicalista, as relações de concordância só se estabelecem depois de se juntarem todos os elementos que compõem a frase. Nesta conformidade, a formação de palavras ocorre no léxico, já que as regras morfológicas são de natureza distinta das sintáticas, actuando em elementos primitivos diferentes dos das segundas. Portanto, a morfologia opera sobre o radical e os morfemas para produzir palavras, enquanto a sintaxe incide sobre as palavras já flexionadas para introduzi-las nas frases. Esta visão configura o Princípio da Integridade Lexical (PIL), preconizado por Harris (1982:26), que assegura que a sintaxe opera a nível das partes da palavra, impedindo que, por exemplo, os afixos sejam desagregados desta com recurso a regras sintáticas.

Contudo, para os não-lexicalistas, parte da formação das palavras não é feita na morfologia, mas sim na própria sintaxe. Baker (1985), observando o sistema de concordância no chamorro, língua malaio-polinésia das ilhas Marianas, avançou com o Princípio do Espelho (PEH), que estabelece um estreito paralelismo entre morfologia e sintaxe, mas não especifica os princípios gerais segundo os quais os afixos são concatenados na estrutura sintáctica do chamorro. Posteriormente, Baker (1988) estudaria também o falar dos índios hopi, Arizona, língua que não possui determinantes e na qual o verbo e o seu complemento aparecem incorporados um no outro, formando uma única palavra.<sup>77</sup> O facto de os determinantes não serem incorporados, permite perceber que este tipo de operação obedece a uma condição sintáctica, e não morfológica. Deste modo, a sintaxe opera a nível de palavras e morfemas, verificando-se que uma palavra complexa pode ser formada através de regras sintáticas e, mais especificamente, com movimento nuclear, através da incorporação de um radical lexical num morfema. Por outro lado, o processo deixa também entender quais são as restrições às incorporações, isto é, qual é o tipo de nomes que pode ser incorporado. Como os movimentos sintáticos operam de baixo para cima, não há hipótese de ocorrer incorporação dos nomes sujeito.

Este tipo de abordagem permite observar o comportamento de elementos morfológicos e sintáticos, em especial a questão do PEH (Baker, 1985), que estabelece que a derivação morfológica reflecte a derivação sintáctica, e vice-versa. Se a estrutura morfológica de uma palavra complexa resulta de um movimento nuclear, o PEH deixa perceber que *“a ordem dos morfemas numa palavra complexa reflecte a incorporação sintáctica de núcleos que correspondem a esses morfemas”*<sup>78</sup> (Baker, 2002:326). Note-se que a abordagem abrange também os efeitos do PIL, uma vez que o resultado do processo de incorporação é ainda uma categoria a nível da palavra. Este aspecto demarca o modelo da incorporação das outras abordagens sintáticas que interagem com a morfologia, nas quais é assumido que uma palavra complexa é formada de palavras lexicais distintas na sintaxe, e que estas só formam palavras a nível fonológico. Neste caso, os efeitos do PIL terão que receber explicações independentes.

Por seu lado, Halle & Marantz (1993) avançam com a proposta da MD. Essencialmente não-lexicalista, o modelo é consensual, visto não retirar protagonismo à componente lexical (Fig. 16).



**Fig. 16.** *Spell-out* em morfologia distribuída (adaptado de Cook, 1996:320-326).

De acordo com esta proposta, há um *input* para a componente morfológica, que lê o que a sintaxe produziu. Por conseguinte, os processos que não são resolvidos pela sintaxe, são-no pela morfologia, em virtude de esta ser sensível às questões da adjacência. Para Bobaljik (1995:243), o modelo demonstra que a afixação não é feita sintacticamente e, como a operação sintáctica está a ser sensível a questões morfológicas, então o processo também não é morfológico. Nesta perspectiva, a sintaxe não trabalha com palavras, mas sim com formas abstractas (traços). Estas questões levam ainda a concluir que a morfologia actua como filtro da sintaxe, já que cada

morfema corresponde a um encaixe (*slot*) sintáctico, mas cuja afixação não está dependente dos movimentos.

As propostas da MD propõem, duas linhas orientadoras-base (Embick, 2000:192):

- (i) As palavras e subpartes que as compõem são constituintes sintácticos mais pequenos, mas que não diferem, no essencial, dos constituintes maiores, como as orações ou frases. Assim, embora cada constituinte tenha propriedades *sui generis*, todas as suas estruturas internas são elaboradas por uma única componente gramatical: a sintaxe.
- (ii) O Léxico não existe na MD, estando as funções que determinadas correntes linguísticas lhe atribuem distribuídas entre outras componentes da gramática.

Segundo Noyer (2006, *apud* Silva, 2010:19/20 ), a estrutura gramatical em MD pressupõe uma divisão em duas partes (Fig. 17): (i) a *informação listada* (uma lista morfémica, um vocabulário e uma enciclopédia), que é depositada em alguns repositórios; (ii) um *dispositivo generativista*, subdividido em operação sintáctica, propriamente dita, e diversos mecanismos pós-sintácticos (abaixamentos, ornamentações, empobrecimentos, deslocamentos locais e linearização).

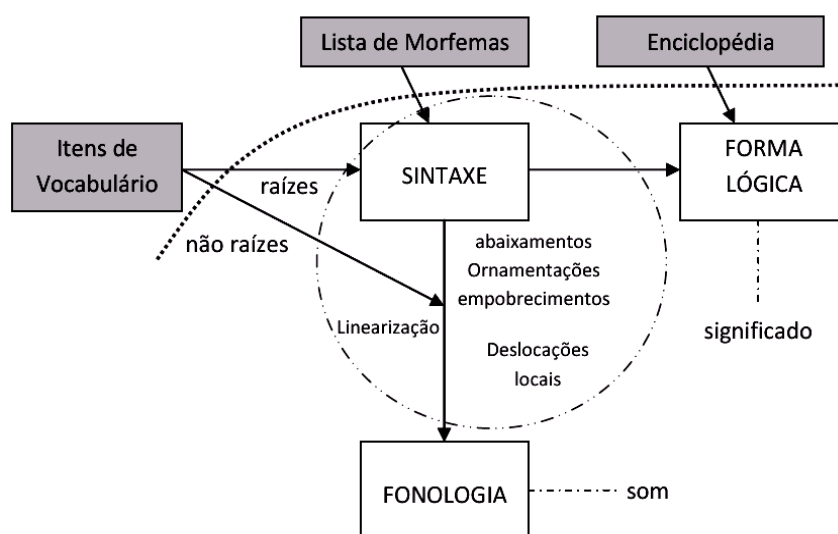


Fig. 17. Modelo da morfologia distribuída (Noyer, 2006:734, *apud* Silva, 2010:20).

O *dispositivo generativista* elabora as expressões linguísticas que provêm dos morfemas seleccionados pela lista morfémica, interpretando fonológica e semanticamente estas expressões com informação concedida pelo vocabulário e pela enciclopédia. Por seu lado, os morfemas contidos na lista morfémica, que constituem os terminais da estrutura morfossintáctica e não contêm traços fonológicos, são de dois tipos: radicais (morfemas-L) e categorias funcionais de sintaxe (morfemas-F). Os primeiros representam itens de classe aberta de categorias indeterminadas, cujas funções são determinadas pelos seus contextos sintácticos (Noyer, 2006:734). Quanto aos segundos, dizem respeito a categorias como o tempo, categorias vazias (CV), CONC ou DET. Portanto, os itens vocabulares (radical das palavras, morfemas de PL ou morfemas de tempo) estabelecem a relação entre representantes fonológicos e morfemas, inserindo-os de acordo as exigências contextuais. Por fim, as entradas da enciclopédia relacionam interpretações com expressões linguísticas estruturadas, que podem ser palavras ou frases.

Halle & Marantz (1993, *apud* Silva, 2010: 21/22) estabelecem, então, três propriedades em MD:

- (i) A Inserção Tardia (*Late Insertion*), que especifica que todas as expressões fonológicas devem ser inseridas somente depois da sintaxe. Deste modo, as categorias sintácticas são abstractas e não possuem conteúdo fonológico, daí resultando que, antes da inserção, elas detenham apenas traços morfossintácticos. Quanto ao momento da inserção do vocabulário, não tem gerado consenso entre os teorizadores. Embick (2000:193) entende que os representantes de morfemas radicais são inseridos anteriormente ou, talvez, durante a sintaxe propriamente dita, enquanto os outros elementos morfemas são-no depois desta. Por um lado, os pressupostos da MD partilham afinidades com os de outras teorias, como a “*Extended Word*” ou o “*Paradigm*” (Anderson, 1982), enquanto, por outro lado, estabelecem semelhanças com os conceitos tradicionais da Morfologia de Base Lexema-Morfema (Beard, 1995);
- (ii) A Subespecificação (*Underspecification*) de itens vocabulares, que determina que as expressões fonológicas não precisam de ser totalmente especificadas para as

posições sintáticas nas quais irão ser inseridas. Este conceito, bastante controverso, rejeita os pressupostos do recurso a um mecanismo de filtragem (Lumsden, 1987; Wunderlich, 1996; Halle, 2000);

- (iii) A Estrutura Sintáctica Hierarquizada *Top-down* (*Hierarchical Structure All the Way Down*), isto é, a elaboração sintáctica dos elementos da sintaxe e da morfologia que se combinam e entram no mesmo tipo de estrutura de constituintes (árvores binárias). A estrutura hierarquizada manifesta-se em todos os níveis da análise estrutural, não ocorrendo morfemas e mantendo-se uma separação restrita entre as regras morfofonológicas e as que atendem aos alomorfes, que podem alterar as propriedades fonológicas dos elementos. Como tal, a inserção do vocabulário limita-se a fornecer elementos. É o que sucede, por exemplo, com a derivação, em inglês, que afecta o termo *foot*, que pressupõe a inserção do item [fut] numa posição de radical. A especificação do contexto de morfema PL implica, desta forma, que se insira um elemento zero na posição canónica de morfema PL. Posteriormente, acontece o reajustamento da forma derivante [i], sendo esta derivação gerada num contexto especial de PL (Noyer, 2006:735).

As propriedades referidas permitem estabelecer o modelo da Fig. 16, em que a sintaxe (1) combina os traços morfossintáticos, cabendo ao *spell-out* (2) a função de conectar a componente morfológica (4), na qual acontecem as operações morfológicas ou de reajustamento à PF (5), ocorrendo as operações fonológicas. Deste modo, encontra-se descrita a teoria que representa, de forma global, o modo como a gramática funciona, no sentido de levar um determinado conceito mental a entrar num *interface* pós-sintáctico conceptual, do qual se processará uma instanciação resultante da construção final na PF. Por seu lado, a FL (3) é apenas um nível de representação que exhibe as relações estruturais relacionadas com o significado das expressões, pelo que não expressa nem representa qualquer significado.

Como se verá no ponto 2.9 do presente trabalho, os pressupostos da MD dão contributo valioso para o entendimento acerca da CPL-var originada pela assimetria de determinados itens gramaticais em posições pré e pós-nucleares, tanto no PB (Costa & Silva, 2006b) como no PA. De facto, tais pressupostos consideram que a estrutura



interna das palavras acontece a partir quer das operações sintáticas, especialmente em sequência do movimento de núcleo (ou Mov-X°), quer de outros mecanismos pós-sintáticos específicos. Portanto, o *merge*<sup>79</sup> morfológico permitirá que os morfemas operem relações de adjacência linear a outros constituintes, daí resultando uma relação de adjunção relativamente aos seus núcleos. Paralelamente, operações de empobrecimento (*impoverishment*) são também de considerar nesta fase, uma vez que as mesmas, ao apagarem os valores de determinados traços morfossintáticos antes da inserção do vocabulário, levam à neutralização sistemática de determinadas categorias (Noyer, 2006:736).

Os estudos da concordância quer entre sujeito e verbo quer a nível dos elementos do SN levaram também à necessidade de se observar a dependência entre concordância (nomeadamente a concordância ao nível do SN) e outros fenómenos sintáticos. Em algumas línguas, como o árabe, a concordância pode ser reflexo das configurações, isto é, resultar de correlações entre as marcas sintáticas e as configurações. De facto, este idioma permite que se altere a ordem dos constituintes, criando-se frases do tipo SV ou VS. Contudo, a configuração entre os constituintes determina que, na estrutura SV, o verbo concorde em número com o sujeito pré-verbal, enquanto na configuração VS essa marca se fixa unicamente no sujeito pós-verbal, permanecendo o verbo no singular (Brustad, 2000:363). Ora, se a concordância é manifestação de algo sintático, mas surge codificada na parte funcional, então esta reflecte aquilo que se passa na sintaxe. Daí existirem línguas que permitem que o verbo concorde com o objecto, como acontece nas estruturas clitizadas do francês (exemplos [184a] e [184b]), e que confirmam a ocorrência de dependência entre as marcas de concordância e a configuração específica:

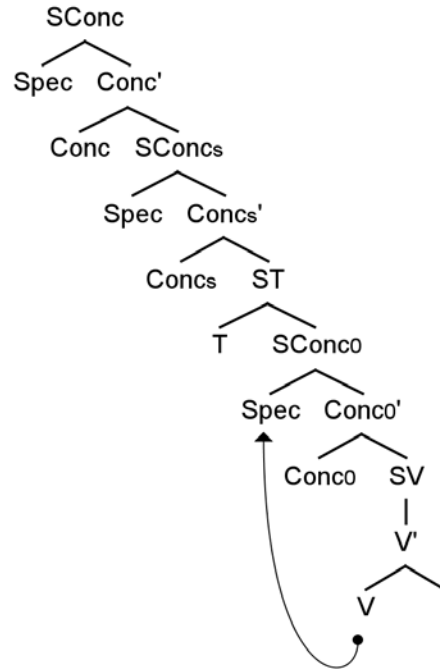
- [184] FRANCÊS:
- a. *J'ai ouvert **les fenê**tres.*  
“Eu abri as janelas”
  - b. *Je **les ai ou**vertes.*  
“Eu abri-as”.

No árabe é ainda possível constatar que a concordância pode ter maior ou menor visibilidade, dependendo da configuração estrutural da realização. Deste modo, a falta de visibilidade morfológica não significa, necessariamente, ausência de processo sintático.

Tendo em conta o pressuposto de Pollock (1989:386), que propõe que o verbo é formado na sintaxe,<sup>80</sup> muitos linguistas trabalharam a questão da concordância partindo do princípio que esta é uma manifestação directa da categoria sintáctica CONC. Porém, levando em consideração que em algumas línguas há uma concordância visível com o objecto, e assumindo que o Caso é uniformemente atribuído na configuração nuclear SPEC, Chomsky (1993), seguindo Kayne (1989), propõe a distinção entre CONC-S(ujeito) e CONC-O(bjecto). Assim, estes dois núcleos são responsáveis pelo licenciamento do Caso, sendo a concordância estabelecida quando um SX e um núcleo inserem uma relação nuclear de SPEC no domínio dessas categorias.<sup>81</sup>

Este tipo de visão sobre a concordância levantou um conjunto de problemas, sendo alguns destes abordados pelo próprio Chomsky (1996 [1995]), quando avançou com os pressupostos que estão na base do Programa Minimalista. Como o espírito do programa é limpar tudo o que se revelava redundante no Modelo de Princípios e Parâmetros, o autor propõe que se reduzam as operações de “movimento” dos constituintes (*move*) e de concatenação da estrutura (*merge*) ao mínimo indispensável (*Derivation by Face*), isto é, que se considere apenas o que é interpretável e que interessa para as *interfaces*, descartando-se aquilo que não é interpretável para estas. Observada sob a perspectiva minimalista, a sintaxe é, então, a aplicação conjunta das duas operações, *merge* e *move*, com a primeira a combinar dois elementos sintácticos para produzir um outro elemento sintáctico, e a segunda a deslocar visivelmente (*overtly*)<sup>82</sup> um item sintáctico de uma posição para outra (Fig. 18).

A proposta de alteração do Modelo-T não só permite ao objecto deslocar-se para receber o caso como também deixa o sujeito ocupar mais do que uma posição, evidenciado uma estrutura dissociada entre este e o caso nominativo. O modelo prevê que, algures na aplicação das regras entre o léxico e a FL, há um ponto, o *spell out*, em que a derivação sintáctica é enviada para a PF. Uma das consequências inerente a este pressuposto é a de que, independentemente da língua em questão, os movimentos são visíveis, apenas variando o momento em que ocorre o *spell out*. Como tal, a interpretação das produções linguísticas, a FL, é igual para todas as línguas.



**Fig. 18.** Estrutura minimalista da concordância sintagmática (Chomsky, 1996 [1995]:173): deslocação do objecto para receber caso.

Por seu lado, a noção de economia e derivações avançada pelo Programa Minimalista determina que uma operação sintáctica só acontece no caso de ser motivada. Numa língua natural existem dez categorias de mecanismos gramaticais minimalistas<sup>83</sup> que são automaticamente postas em acção. Se estes traços estiverem ausentes ou forem insuficientes no *input*, as crianças accionam os princípios cognitivos inatos para os desenvolverem na sua linguagem. Com vista a definir-se o que é ou não é um movimento motivado, foi, então, criado um sistema de *traços-phi*, associado a qualquer categoria sintáctica. Estes podem ter uma força forte (o elemento sintáctico tem movimento) ou fraca (o elemento sintáctico não tem movimento), indo as forças determinar o factor de variação entre as línguas. Na estrutura minimalista da concordância sintagmática (Fig. 18), CONC-S e CONC-O são mnemónicas informais para distinguir os dois *papéis-phi* de CONC. Assim, este representa um conjunto de *papéis-phi* (género, número, pessoa), ou seja, os sistemas comuns de CONC para sujeitos e objectos, uma vez que CONC-S e CONC-O podem representar selecções distintas de níveis de acordo. Por outro lado, os *traços-phi* associados às categorias linguísticas são ainda de dois tipos: interpretáveis, que têm relevância na FL (p.e. o número); e não-interpretáveis (p.e. o género e o caso).

Na visão de que as categorias vêm equipadas com traços, a sintaxe serve basicamente para eliminar, através das operações de movimento, aqueles que não são interpretáveis, ocorrendo o apagamento quando traços semelhantes se encontram. Como tal, o falante selecciona como gramatical aquilo que é mais económico, esvaziando a componente sintáctica e colocando-a no léxico. Nesta perspectiva, o movimento existe para seleccionar as operações que não interessam para as *interfaces*, pelo que constitui uma propriedade mais económica (preferida) da língua do que a inserção de material lexical. Assim, as *interfaces* ganham lógica e tudo o que não cabe no modelo mínimo cai na sua área. Este aspecto determina que alguns movimentos, como o movimento- $\bar{A}$ , que tem por base inúmeras motivações (p.e. questões de escopo ou discursivas), afectem qualquer tipo de sintagma, enquanto outros, como o movimento-A, motivado pelo caso e cujo alvo é a posição SPEC, afectem apenas os SN's argumentais.

A partir dos pressupostos minimalistas, foi possível procurar entender a estrutura das *interfaces*. Deste modo, não só a morfologia, que não é contemplada no Modelo-T, pôde ser tratada de forma mais clara, como também se abriu caminho para que, posteriormente, se observasse a *interface* entre sintaxe, morfologia e semântica.

### **2.6.3. O SN no português europeu**

Qualquer construção linguística configura um sintagma, ou grupo, cujo elemento central, o núcleo, é realizado com recurso a uma categoria sintáctica (nuclear), vulgarmente designada por “parte do discurso”, “classe de palavras” ou “classe morfosintáctica” (Brito, 2003a:326). No caso do português, a estrutura interna do SN inclui obrigatoriamente um núcleo e, opcionalmente, outros dois tipos de constituintes: complementos e especificadores. Se as palavras inseridas no SN configuram unidades pertencentes ao reduzido grupo de palavras ou unidades morfológicas da língua, cujo significado remete para as noções mais abstractas desta, estamos na presença de categorias funcionais. Opostamente, se as unidades incluídas no SN fazem parte do extenso e renovável inventário do vocabulário ou léxico da língua, cujos significados remetem para entidades, situações, propriedades ou relações entre entidades, estamos perante categorias lexicais, como acontece com os nomes, os adjectivos, as preposições, os verbos e os advérbios. Estas categorias, por seu lado, constituem o núcleo das respectivas categorias sintagmáticas em que estão incluídas: sintagma nominal (SN),

sintagma adjetival (SADJ), sintagma preposicional (SPREP), sintagma verbal (SV) e sintagma adverbial (SADV). Deste modo, as categorias sintagmáticas são projecções endocêntricas dos seus próprios núcleos, já que os segundos são da mesma natureza das primeiras.

Face ao exposto, e observando a estrutura geral do SN, verifica-se que ele representa uma categoria sintagmática que é projecção de um nome, podendo ser composto apenas por esse nome (exemplo [185a])<sup>84</sup> ou por um pronome (exemplo [185b]),<sup>85</sup> os quais são, simultaneamente, núcleo do referido SN. Contudo, este pode ainda ser realizado por um nome e por outros elementos opcionais, tais como determinantes (exemplo [185c]), quantificadores (exemplo [185d]), expressões quantitativas (exemplo [185e]) e modificadores, dividindo-se estes em SADJ's (exemplo [185f]), SPREP's (exemplo [185g]), sintagmas oracionais (exemplo [185h]) ou epítetos (exemplo [185i]) (Brito, 2003a:328-329):

[185] PE:

- a. <sup>SN</sup>[**Deus**] *perdoa os que se arrependem.*
- b. <sup>SN</sup>[**Ninguém**] *quis sobremesa.*
- c. <sup>SN</sup>[**Esses livros**] *estão à venda.*
- d. <sup>SN</sup>[**Muitos livros**] *ficaram por vender.*
- e. <sup>SN</sup>[**Uma porção de livros**] *ficou por vender.*
- f. *O Pedro é* <sup>SN</sup>[*um homem grande*].
- g. *O Pedro tem* <sup>SN</sup>[*muitas camisas de algodão*].
- h. *Perdi* <sup>SN</sup>[*o isqueiro que comprei ontem*].
- i. <sup>SN</sup>[*O João, aquele malandro*], *tem faltado às aulas.*

Os exemplos evidenciam que núcleo e complementos constituem a estrutura lexical da categoria SN, enquanto determinantes e quantificadores configuram a sua estrutura funcional. Por seu lado, os modificadores recobrem constituintes com variados valores semânticos, surgindo alguns mais ligados à estrutura lexical, enquanto outros se conectam mais à estrutura funcional. Em algumas realizações, o núcleo do SN é também aparentemente preenchido por especificadores (exemplo [186a]), mas, na verdade, está-se perante SN's em que o núcleo é uma categoria vazia, uma vez que tais especificadores se podem conjugar com outros especificadores e com diversos tipos de complementos (exemplo [186b]):

- [186] PE:
- a. *Todos gostaram do espectáculo.*
  - b. *Todos aqueles não gostaram do espectáculo.*

Sendo a forma do SN marcada por mecanismos de concordância, os exemplos patentes em [185] e [186] evidenciam também que o núcleo do SN determina a concordância dos especificadores, dos determinantes (artigos, demonstrativos e possessivos) e de alguns complementos (SADJ's e epítetos). Apesar de haver critérios que permitem a identificação dos nomes, deve ter-se em consideração que a concordância interna do SN é morfológica e semanticamente susceptível ao tipo de nome que pode servir de núcleo (nomes animados humanos, nomes animados não humanos, nomes não animados, nomes comuns, nomes próprios, nomes contáveis, nomes não contáveis ou massivos, nomes concretos e nomes abstractos). Estes aspectos determinam, então, as condições em que se estabelecem as relações de complementação a nível das estruturas lexical e funcional do SN. De facto, “*os nomes relacionáveis lexicalmente com outras categorias são quase sempre predicados semânticos, com lugares ou argumentos que recebem uma dada interpretação, que pode ser descrita através de relações temáticas*” (Brito, 2003a:331).

#### **2.6.4. Categorias funcionais e categorias lexicais**

A teoria linguística tem procurado estabelecer paralelismos entre a sintaxe das frases e a sintaxe dos SN's, já que este tipo de observações pode ajudar a entender quer aquilo que motiva os diferentes padrões de ordem das palavras nos SN's das diferentes línguas quer a relação estabelecida entre a sintaxe e a morfologia flexional dentro do SN (Montrul, 2004:12, apud Godinho, 2005:131). Tendo em conta que núcleo e complementos constituem a estrutura lexical da categoria SN, então os determinantes e as formas de quantificação (quantificadores, expressões quantitativas e expressões partitivas) formam a sua estrutura funcional, localizada à esquerda do nome. Assim sendo, na perspectiva da gramática generativa actual, o SN tradicional deixou de ser entendido como tal, para, de acordo com o princípio da endocentricidade, passar a ser visto como um S<sub>DET</sub>, a projecção máxima de um DET (Fig. 19). Esta hipótese postula, então, que os determinantes e formas de quantificação, que restringem a extensão do núcleo nominal, devem ser considerados o núcleo da categoria funcional S<sub>DET</sub> ou

sintagma quantificador (SQ), os quais são responsáveis pela selecção dos SN's como complementos. Desta forma, N é o núcleo da parte lexical ou categoria (Brito, 2003a:345-349, Fig. 19), cabendo ao SDET a função de atribuir valores referenciais do tipo [+definido] ou [-definido] ao SN complemento. Portanto, SDET e sintagma complementador (SCOMPL) estabelecem uma relação de paralelismo no SV.

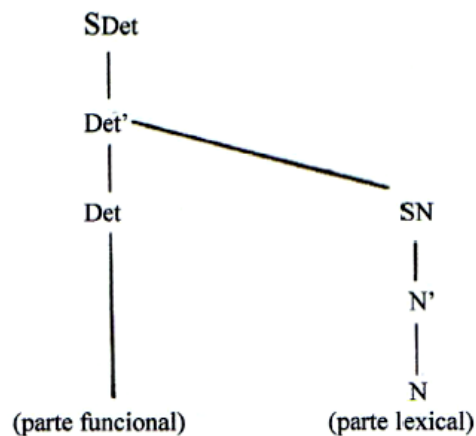


Fig. 19. Estrutura do SDET: parte funcional e parte lexical (Brito, 2003a:345, *apud* Godinho, 2005:132).

Na estrutura proposta, o núcleo da categoria SDET é preenchido por artigos ou demonstrativos. Assim, os determinantes (artigos e, por norma, possessivos, demonstrativos e quantificadores) precedem o nome núcleo do SN, enquanto os adjectivos, complementos e adjuntos aparecem, geralmente, depois deste (cf. Godinho, 2005:132). Ocorre, desta forma, uma relação entre posição linear e categoria gramatical, que influencia a marcação PL (Guy, 1981b:179), aspecto com o qual Scherre (1988, 1998a [1996]) entra em desacordo. Esta questão, à qual fizemos já referência no ponto 1.2 do presente trabalho, será por nós amplamente debatida no capítulo 5, especialmente no ponto 4.1.2.6, dedicado à observação do comportamento da variável *posição do item na cadeia do sintagma nominal* e seu cruzamento com outras variáveis independentes.

No que concerne aos artigos, precedem sempre os nomes e são de dois tipos: definidos e indefinidos. A distinção entre ambos é feita com recurso aos respectivos valores semânticos e, ao contrário dos segundos, os definidos nunca podem surgir isolados, quer por elipse do N (p.e. \**Eu comi o.* / *Eu comi um.*) quer devido ao movimento do N e do complemento (p.e. \**Manual de botânica, eu comprei o.* / *Manual de botânica, eu comprei um.*). Ainda relativamente aos definidos, podem surgir em

realizações com elipse nominal, caso sejam seguidos de um complemento (p.e. *Eu comprei os chinelos azuis e tu os amarelos.*), um modificador adjetival (p.e. *O Pedro comprou os verdes.*) ou uma oração relativa (p.e. *O João comprou os que são castanhos.*). Estes artigos surgem também em distribuição complementar relativamente aos demonstrativos, tornando agramaticais as frases que não respeitam este preceito (p.e. *\*Eu comprei o este jornal de hoje.*). Por sua vez, os indefinidos não acompanham nomes próprios (p.e. *\*Ontem, eu falei com um António.*) nem podem ser colocados depois de *todos* ou *ambos* (p.e. *\*Todos uns alunos foram à visita de estudo./ \*Ambos uns alunos faltaram às aulas hoje.*).

Os demonstrativos, ainda que não co-ocorram com definidos, podem realizar-se em construções de elipse nominal (p.e. *Prefiro estes aqui.*). Nas realizações de tipo déitico, os demonstrativos *este*, *esse* e *aquele* conjugam-se com localizadores para funcionarem como “filtros de vizinhança” (Lopes, 1972:141), mas apenas em distribuição complementar (p.e. *\*Este aqui livro é do Pedro.*). Os demonstrativos podem ainda possuir valor anafórico, remetendo para o discurso anterior (p.e. *Eu comprei livros de matemática e botânica. Aqueles foram mais caros do que estes.*) e ser realizados pré (exemplo [185c]) ou pós-nominalmente com exclamativas (exemplo [187a]) ou relativas apositivas (exemplo [187b]), mas sempre em distribuição complementar com os artigos, sob pena de a construção se tornar agramatical (exemplo [187c]):

[187] PE:

- a. *Que calor este!*
- b. *Comprei um carro ontem, carro esse que está na minha garagem.*
- c. *\*Os esses livros estão à venda.*

Ainda assim, a distribuição frásica dos demonstrativos justifica a aplicação da estrutura frásica patente na Fig. 19 (Brito, 2003a:349).

Relativamente aos possessivos, exprimem valores temáticos e de determinação, possuindo marcas de género e número denunciadoras de uma dupla natureza lexical-funcional. Portanto, podem co-ocorrer na cadeia pré-nominal com artigos definidos (exemplo [188a]), demonstrativos (exemplo [188b]) e quantificadores (exemplo [188c]). Os possessivos podem pospor-se ao nome, se este não for antecedido de artigo (exemplo [188d]) ou se realizar com artigo indefinido (exemplo [188e]). A localização pós-



nominal pode ocorrer com cardinais (exemplo [188f]) ou estruturas não argumentais, envolvendo vocativos (exemplo [188g]) ou posições predicativas (exemplo [188h]):

[188] PE:

- a. ***O meu irmão*** chega amanhã a Macau.
- b. ***Aquela tua amiga*** é muito simpática.
- c. ***Estes meus muito amigos*** partem hoje para Lisboa.
- d. Estava à espera de ***notícias tuas*** há mais de uma semana.
- e. Recebi ***uma carta tua***, que me deixou muito contente.
- f. Hoje vi dois ***empregados seus***.
- g. ***Virgem nossa***, que estupidez!
- h. ***Aquele ali*** é um ***empregado teu***?

Nos exemplos até agora exibidos, temos focalizado atenção nas estruturas em que o SDET é constituído por uma sequência de palavras em torno de um nome. Apesar de as sequências referencialmente autónomas (p.e. os pronomes pessoais ou os demonstrativos *isto, isso, aquilo*) poderem constituir, por si sós, SDET's (p.e. *Vocês fizeram serão?* / *Isto tem muito que se lhe diga.*), abster-nos-emos de as observar e comentar, já que caem fora do âmbito do presente estudo. Pelos mesmos motivos, não observaremos as elipses nominais com determinantes nem as realizações com nomes próprios.

Relativamente à quantificação dos nomes, pode ser expressa com recurso aos plurais simples (p.e. *Hoje comprei chocolates para ti.*), aos quantificadores propriamente ditos (p.e. *Hoje comprei muitos livros na Feira dos Livros.*) ou às expressões quantitativas (p.e. *Hoje comprei uma data de livros na Feira dos Livros.*). Sobre os tipos de elementos abrangidos pelos quantificadores e expressões quantitativas, bem como as suas projecções endocêntricas em estruturas sintácticas, remetemos as explicações pertinentes para Brito (2003a:356-365), que abrangem também as expressões partitivas (p.e. *Alguns dos livros que comprei, são raríssimos.*).

No que aos modificadores nominais (restritivos ou apositivos) diz respeito, centraremos a nossa atenção nos adjectivos. Estes podem ocorrer antes (exemplo [189a]) ou depois do nome (exemplo [185f]), mas os posicionamentos pré-nominais contemplam situações bastante restritas. Contudo, como se observou no ponto 2.6.1 do presente trabalho, há adjectivos que possuem um traço formal rígido, pelo que só podem ser realizados antes (exemplos [189b] e [189c]) ou depois do nome, (exemplos [189d] e [189e]), consoante o caso. Os que podem ocorrer em posições indistintas, mudam

normalmente de valor semântico, conquanto se encontrem antes ou depois do nome. Nestes casos, a posição pós-nominal está associada a significados mais objectivos ou denotativos (exemplo [185f]), enquanto a localização pré-nominal surge mais ligada a interpretações subjectivas ou conotativas (exemplo [189a]). Não obstante, há adjectivos cuja alteração de posição relativamente ao nome não acarreta mudança de valor semântico (exemplos [176a] e [176b]), estando o posicionamento dependente de uma mera questão estilística ou da intenção de focalizar a informação.

[189] PE:

- a. *O Pedro é um **grande** homem.*
- b. *Aquilo não passou de um **mero** mal-entendido.*
- c. *\*Aquilo não passou de um mal-entendido **mero**.*
- d. *O vestido **azul** da Maria é muito bonito.*
- e. *\*O **azul** vestido da Maria é muito bonito.*

Note-se também que há nomes não relacionáveis lexicalmente com outras categorias e que podem surgir seguidos de SPREP's (os complementos determinativos). Estes, por seu lado, podem restringir o significado do SN (p.e. *O manual de botânica.*), tal como acontece nas ocorrências com relativas restritivas (p.e. *O manual que trata de botânica.*) ou com determinantes (p.e. *Esse manual.*). Outros SPREP's, contudo, podem exprimir relações semânticas mais ténues, a ponto de originarem realizações ambíguas (p.e. *Gostei muito do quadro do Manuel da Glória.*).

#### 2.6.4.1. Aquisição do sintagma determinante

Godinho (2005, 132-138) refere pormenorizadamente o modo como ocorre a aquisição da categoria funcional SDET. Dado que esta domina o SN, a projecção máxima do nome, mas que não contém o determinante, é também uma categoria que descreve uma expressão nominal referencial, aparecendo como categoria funcional ligada à categoria lexical SN. Desta forma, o N pode ocorrer sem DET, originando SNN's, mas que não invalidam a hipótese de existir projecção do SDET (Baxter. Lucchesi & Guimarães, 1997:8, nota 14, *apud* Godinho, 2005:132). Meisel (1994a:92) constatou que, na fase FL0 da aquisição bilingue, não ocorrem categorias funcionais. Estas acabam apenas por emergir faseadamente no tempo, para determinarem o desenvolvimento das categorias gramaticais associadas à flexão (concordância, tempo e finitude). Por

consequente, o SDET, apesar de ser foneticamente vazio, continua a encerrar em si as propriedades referenciais de especificação do SN (Godinho, 2005:132). Note-se que, sem o primeiro, o segundo estaria também impedido de actuar como um argumento do verbo da frase, um aspecto que vai determinar que o núcleo DET detenha a função de subordinador, ou seja, uma função semelhante à desempenhada pelos complementadores, enquanto núcleos das frases encaixadas (Godinho, 2005:133).

Como a flexão é um processamento sintáctico, não pode ocorrer dentro do SN, como o comprova o facto de, empiricamente, acontecer marcação variável de género e número no núcleo do SN (Baxter, Lucchesi & Guimarães, 1997:9, *apud* Godinho, 2005:133, nota 173). De facto, é na categoria NUM, pertencente ao nó funcional SNUM, (categoria intermédia entre o SN e o SDET), que estão comportados os traços de número [+/- plural], a fim de serem mapeados e resgatados pelo SDET. (Picallo, 1991; Carstens, 1991; Bernstein, 1993; Cinque, 1994; Koehn, 1994; Müller, 1994a; Montrul, 2004; Godinho, 2005) (Fig. 21):

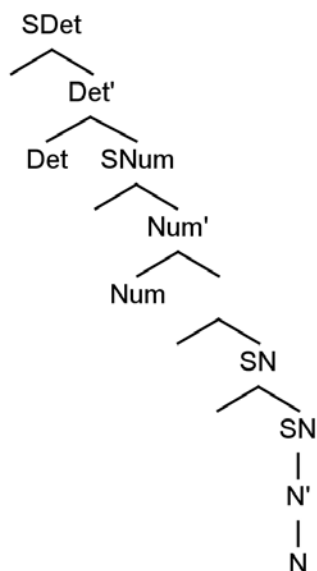


Fig. 20. Estrutura do SDET: nó funcional SNUM.

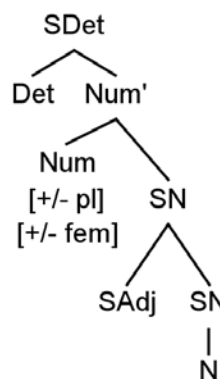


Fig. 21. Representação funcional da estrutura do SN (Montrul, 2004:35, *apud* Godinho, 2005:134).

No português, os traços funcionais de número atraem os nomes, que sobem de N para NUM (cf. Fig. 22), originando a ordem frásica típica do português (N > ADJ) (Bruhn de Garavito & White:2002:154, *apud* Godinho, 2005:134). Ainda a ter em consideração é o pressuposto de que a categoria SNUM contém, igualmente, a propriedade de

concordância em género (Ferrari-Neto, Corrêa & Augusto, 2005:1077), assumindo, assim, as funções de categoria CONC (ou AGREE) (Cerqueira, 1994; Godinho, 2005:134).

As aquisições de L1 e L2 exibem aspectos comuns e são orientadas por operações particulares. No caso concreto da aquisição do SDET na L1, as categorias funcionais apenas detêm função gramatical, subordinando a si as categorias lexicais e circunscrevendo os parâmetros de variação aos seus traços (cf. Godinho, 2005:135). No caso concreto dos falantes de línguas do grupo níger-congo atlântico ou do santomense em aquisição do PtgL2, eles não só terão de adquirir os aspectos morfossintácticos da LA (aprender que os nomes da LA não são marcados com afixos temáticos) mas também perceber que a marcação de género está sujeita a marcação PL e concordância em género e número, normalmente com recurso a sufixos que se dissociam pelo nome, determinantes e adjectivos (cf. Godinho, 2005:136).

A aquisição precoce da categoria funcional SDET no PtgL1 revela variação na concordância em género e número nas suas fases iniciais (Faria, Freitas & Miguel, 1998, 2001; Name & Corrêa, 2003, Costa & Santos, 2004 [2003]; Corrêa, Augusto & Ferrari-Neto, 2005; Ferrari-Neto, Corrêa & Augusto, 2005; Castro & Ferrari-Neto, 2007), uma situação ocorre também na aquisição bilingue (Barreña, 1997; Lléo, 2001; Meisel, 1994a). As etapas de aquisição são discriminadas por Godinho (2005:137), iniciando-se com a produção de nomes, categorias que exibem também os traços de número em primeiro lugar (Faria, Freitas & Miguel, 2001:55, *apud* Godinho, 2005:137). Por volta dos catorze meses, os pequenos falantes produzem também sílabas pré-nominais, entendidas como proto-determinantes ou proto-morfemas (*fillers*) (Faria, Freitas & Miguel, 1998; Name & Corrêa, 2003; López-Ornat, 1997; Lléo, 1997, 1998, 2001), visto assumirem a função de marcadores de lugares morfológicos. A partir dos dezoito meses, as crianças começam a elaborar SN's compostos por um elemento vocálico, do tipo DET, e um N marcado em género. No final desta etapa da aquisição, inicia-se a inserção correcta dos adjectivos na ordem frásica, cuja concordância em género com o N é feita quase sem variação, especialmente quando este é do tipo inanimado. O uso de flexões e concordância revela que se iniciou a fase da aquisição gramatical. Por outro lado, a realização dos diferentes tipos de DET nas fases precoces da aquisição da L1 fornece indicativos importantes não só para a discussão acerca da natureza das gramáticas

precoces mas também para ajudar a determinar as etapas da aquisição em que as categorias funcionais são disponibilizadas (Godinho, 2005:137).

A ausência de SDET's na etapa FL0 da aquisição tem sido apontada por diversos estudiosos (Koehn, 1994; Müller, 1994a; López-Ornat, 1997), pelo que, na transição desta fase para a fase gramatical, “*as crianças fazem corresponder ‘pistas’ fonoprosódicas a funções pragmático-semânticas*” (Godinho, 2005:138). Portanto, a criança não consegue discriminar determinados sons, entendendo formas distintas (DET+N) como sequências fonoprosódicas únicas (Koehn, 1994:49; DeGraff, 1999b:482; Faria Freitas & Miguel, 2001:53). A finalidade pragmático-semântica destas realizações permite que a criança percepcione as regularidades linguísticas, extraia as suas representações morfológicas e produza, depois, DET's que concordam em género e número com o N (cf. Godinho, 2005:138).

Na aquisição do espanhol L1, a importância da fonologia na aquisição morfossintáctica é evidenciada pelas vogais pré-nominais proto-morfémicas, algumas das quais apresentando regularidade linguística (López-Ornat, 1997, *apud* Godinho, 2005:138). Deste modo, a elaboração gramatical segue um trajecto que transita da produção fonoprosódica para a realização sintáctica. Estes aspectos vêm contrariar os pressupostos de uma observação da aquisição da linguagem numa perspectiva meramente sintáctica (Montrul, 2004:42, *apud* Godinho, 2005:138), relevando a importância da fonologia e da prosódia na aquisição.

No caso concreto do PEL1, Faria, Freitas & Miguel (1998, *apud* Godinho, 2005:138-139) apontam a realização de proto-morfemas (*fillers*) na posição do DET, confirmando que as categorias funcionais estão presentes desde as etapas primordiais da aquisição. Para além destes proto-determinantes, Faria, Freitas & Miguel (1998, 2001, *apud* Godinho, 2005:139, notas 185 e 186) mencionam ainda a inserção de morfemas de género e número nos nomes, sob a forma de fricativa morfológica em coda de sílaba, que é realizada anteriormente às outras fricativas em coda de sílaba.<sup>86</sup> Como se verifica, tudo aponta para que a categoria SDET esteja disponibilizada desde a fase FL0 da aquisição, ainda que não haja controlo sobre as formas morfofonológicas dos elementos que a compõem. Nesta conformidade os proto-morfemas atestam que as posições estruturais já se encontram contempladas, embora as operações de mapeamento de

traços só venham a acontecer mais tardiamente (Faria, Freitas & Miguel, 1998:119, *apud* Godinho, 2005:139), por norma a partir dos 2;0 anos, isto é, quando a maturação neurológica da criança lhe disponibiliza a GU (Meisel, 1994a:127), passando a sua gramática a operar a partir desta fase dentro da Hipótese Continuista da aquisição (Pinker, 1984; Hyams, 1987; Kato, 1995; Rice & Wexler, 1996). Estes aspectos são confirmados, nas mesmas pesquisas, com o facto de a gramática das crianças que adquirem o PE se apresentar mais próxima da gramática dos adultos do que a gramática de crianças que adquirem línguas germânicas ou outras línguas românicas (Godinho, 2005:139).

Também em estudos sobre aquisição de artigos no espanhol e no alemão é apontado o aparecimento de proto-morfemas na posição do DET, preenchendo e assumindo, posteriormente, a posição sintáctica destas categorias (Lléo, 1998; Montrul, 2004, *apud* Godinho, 2005: 139-140), o que sugere continuidade.

Os estudos de Name & Corrêa (2003), Ferrari-Neto (2003), Corrêa, Augusto & Ferrari-Neto (2005) e Ferrari-Neto, Corrêa & Augusto (2005), realizados com crianças em processo de aquisição do PB, evidenciam que elas são, igualmente, sensíveis aos determinantes, ainda que estes não apresentem um padrão fónico regular (Godinho, 2005:140). As conclusões dos estudos não confirmam, desta forma, os pressupostos de que o uso de determinantes não é frequente na produção inicial das crianças (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993-1994; Meisel, 1994a, 1997). De facto, as crianças com cerca de catorze meses de idade conseguem já identificar os determinantes como uma classe fechada e mapeá-los com o DET, evidenciando que, nesta fase da aquisição do PB, os requisitos para identificação do género, número e pessoa, se encontram já contemplados. (Name & Corrêa, 2003:25, *apud* Godinho, 2005:140).

Estes pressupostos são compartilhados por Lléo (1997, 1998, 2001), que entende a projecção do SDET numa fase mais precoce como um requisito solicitado pelo contexto, configurando morfemas funcionais relacionados com os artigos e confirmando a hipótese de maturação. Por conseguinte, o acesso a estas categorias não é inato, sendo a aquisição da categoria SDET em PtgL1 governada por aspectos fonológicos e prosódicos. Daí as crianças portuguesas adquirem esta categoria antes das crianças em

aquisição do alemão ou inglês L1, já que estas estão submetidas a um *input* fonológico e prosódico distinto do português (Godinho, 2005: 141-142).

#### **2.6.4.2. Aquisição de género e número**

A ultrageneralização das regras de concordância de género nas fases primordiais da aquisição da L1 evidencia que a aplicação da regra de concordância entre DET e N já se encontra disponibilizada nestas, apesar de não ser aplicada correctamente (Godinho, 2005:142). Ao procurar determinar o momento preciso em que as crianças adquirem o mecanismo sintáctico para a concordância do género no PB, Name & Corrêa (2003) constataram que a variação morfofonológica ocorre na categoria DET. Esta particularidade evidencia que as crianças possuem um dispositivo inato para processar as propriedades formais desta categoria (Godinho, 2005:142), aspecto que seria confirmado por Ferrari-Neto, Corrêa & Augusto (2005:1077) em pesquisa sobre o modo como as crianças com cerca de dois anos de idade percebem a informação morfofonológica relativa ao número e processam a concordância de número no âmbito do SDET no PBL1. No estudo, os autores concluem que tanto o género como o número são processados no âmbito do SDET, uma vez que é *“a informação de género expressa no DET e o processamento da concordância de género expressa no SDET que possibilitam à criança identificar o género de palavras novas (nomes com traço semântico inanimados, nos quais o género é intrínseco, sem motivação semântica, de um ponto de vista sincrónico)”*. Como tal, os autores entendem que o traço de género pode ser mapeado a nível do DET0, uma situação idêntica à do mapeamento de traço de número (Longobardi, 1994:621). Contudo, essa não é a visão de Lucchesi (2000a), para o falar de HEL-Ba, e Figueiredo (2009a), para o PA, que entendem que a inserção da marca de género ocorre no nome núcleo do SN destas variedades de português, expandindo-se a partir deste para se fixar nos elementos à sua esquerda, primeiramente, e só depois nos itens à sua direita, embora de forma marginal. Também Godinho (2008, 2009), em estudos sobre aquisição da concordância de género e número do PtgLE por adultos falantes de cantonense L1, em três momentos distintos, refere que os nomes núcleo do SN são os elementos que exibem maior índice de marcas, sendo responsáveis pela fixação dos traços da flexão morfológica de género no primeiro momento da aquisição. Posteriormente, os traços

disseminam-se pelos outros itens, adjacentes ao núcleo (cf. Lucchesi, 2000a; Figueiredo, 2009a). Segundo a autora, estes aspectos permitem predizer que a concordância se vai espalhando pelos diferentes elementos do SN, fazendo prever que a aquisição da concordância neste é um sistema que sofre maturação gramatical. Desta forma, os princípios da GU actuam a partir dos primeiros momentos da aquisição, a que se seguem, mais tardiamente e de modo descontínuo, outros aspectos operativos (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997), sendo o desenvolvimento sintáctico conduzido pela aprendizagem de novos itens lexicais e morfológicos (Montrul, 2004:96). Este aspecto implica que, de certa maneira, as novas estruturas só são aprendidas caso as etapas de aquisição prévia estejam completadas (Pienemann, 1998, 2005b), conforme o confirmam, aliás, os momentos dois e três da aquisição observados por Godinho (2008, 2009). Por outro lado, os achados permitem também concluir que os adultos falantes de cantonense L1 em aquisição do PtgLE cometem erros de concordância em virtude de os traços parametrizados para aquisição do género e número deixarem de estar disponibilizados depois do período crítico do desenvolvimento.

Para entendimento de aspectos concernentes à concordância de número, Godinho (2005, 143-144) aponta a importância dos estudos de Marrero & Aguirre (2003), Pérez-Pereira (1989) e Bruhn de Garavito (1994). No primeiro são abordadas as etapas de desenvolvimento do PL no espanhol, em que as crianças começam por não fazer a distinção entre formas singular/plural, para, posteriormente, aplicarem o PL unicamente na classe DET e, depois, o inserirem nos outros elementos do SDET, ainda que com variação na aplicação da regra de concordância. Para Montrul (2004, *apud* Godinho, 2005:144), esta aquisição tardia dos alomorfes de PL é determinada por factores fonológicos, enquanto a aquisição do género é mais precoce, em virtude de esta categoria ser motivada por factores meramente linguísticos. Por seu lado, Pérez-Pereira (1989, *apud* Godinho, 2005:143) debruçou-se sobre a aquisição dos alomorfes /s/ e /es/ na mesma língua, constatando dois aspectos: (i) ultrageneralização do primeiro alomorfe, sendo também aplicado nas palavras que requerem o uso da vogal epentética, sem que esta seja realizada; (ii) tendência para manter no singular os itens terminados em -s.



Quanto ao estudo de Bruhn de Garavito (1994), analisa a forma como aprendentes de L1 francesa adquirem o PL no espanhol L2. Os estádios de aquisição são idênticos aos apontados por Pérez-Pereira (1989), já que os falantes começam por não realizar a marcação de PL, adquirindo depois o alomorfe de PL /s/ e, apenas posteriormente, o marcador de pluralização /es/. A conclusão é que “*a aquisição tardia do alomorfe /-es/ não é um problema morfológico, mas sim de natureza fonológica e diz respeito à aquisição da epêntese e das regras de silabação do espanhol*” (Godinho, 2005:144).

Ferrari-Neto, Corrêa & Augusto (2005), no seu estudo sobre a forma como as crianças com cerca de dois anos de idade percebem a informação morfofonológica relativa ao número e processam a sua concordância no âmbito do SDET no PB, levaram em conta os princípios minimalistas de que as relações sintácticas entre elementos seleccionados do léxico se realizam por meio do pareamento de traços formais mais e menos interpretáveis (Chomsky, 1996 [1995], 2001b). Portanto, é esse pareamento que permite ao dispositivo computacional universal operar sobre traços do léxico de origem semântica. A existência de traços não-interpretáveis correlatos a traços semanticamente interpretáveis (p.e. o número) permite que o número conceitual encontre uma expressão gramatical, ou seja, tome parte em operações sintácticas (*merge* e AGREE). Neste contexto, os autores entendem que a concordância é o pareamento de traços interpretáveis com traços não-interpretáveis, numa configuração de C-comando. No processo de pareamento, os traços não-interpretáveis são valorados ou eliminados, tendo o resultado dessa valoração uma expressão morfológica, naquilo que tradicionalmente se designa como concordância de número, género e pessoa, em línguas morfológicamente mais elaboradas, como as românicas. No caso específico da concordância de número, o traço interpretável encontra-se no nome (Chomsky, 1996 [1995]) ou numa projecção funcional independente (Ritter, 1991:61).

Em ambas as situações, o número interpretável tende a expressar-se numa desinência flexional do nome, ganhando visibilidade na *interface* fonética. Quanto aos restantes elementos do SDET (determinantes, possessivos, adjectivos, etc.), apresentam traços não-interpretáveis de número, estabelecendo, desta forma, uma relação de concordância com o nome ou com a projecção funcional NUM.<sup>87</sup> Na sequência de tal relação, o SDET assume o traço de número definido no nome (ou no SNUM). Como em algumas línguas

os elementos portadores dos traços não-interpretáveis adquirem um afixo flexional correspondente ao valor do traço de número, aquando da sua codificação morfológica, a informação relativa ao número nesses elementos torna-se visível na *interface* fonética. Deste modo, co-existem dois sistemas no PB: (i) no modelo padronizado, tanto a expressão morfológica do traço interpretável de número na desinência do nome como a expressão morfológica do traço valorado de número no determinante, e demais elementos que entram em concordância com o nome, encontram-se visíveis na *interface* fonética (p.e. *A-s casa-s branca-s*); (ii) no modelo não-padrão, a informação de número tende a ser visível na *interface* fonética apenas no DET (ou no elemento mais à esquerda no SDET), originando concordância variável (p.e. *A-s casa branca*). Por conseguinte, quando a expressão do número é feita apenas no DET, a identificação dessa informação, por parte da criança, durante os enunciados da língua em aquisição, envolve o processamento da concordância no SDET.

Ferrari-Neto, Corrêa & Augusto (2005:1081) chamam ainda a atenção para o facto de o fonema /s/ interferir nos estímulos testados. Todavia, o estudo não clarifica se este efeito pode ser tomado como indicativo de uma fase de desenvolvimento linguístico na qual a criança sofre interferência desse tipo de condicionamento silábico, e que a leva a processar singulares quando deveriam ser inseridos os morfemas de número PL. Ainda assim, os resultados do estudo permitem verificar que as crianças com cerca de dois anos de idade estão sensíveis à expressão gramatical de número no PB, visto que, nesta fase, passam a processar diferentemente as condições gramatical e não-gramatical (cf. Meisel, 1994a:127). A não-diferença entre as condições gramatical padrão e não-padrão indica ainda que a informação de número é extraída da categoria DET, o que implica processamento da concordância no SDET, consoante o modelo de língua a adquirir. Face a estes resultados, os autores atribuem o processamento da concordância a um dispositivo computacional universal, que se encontra disponível aquando dos primeiros contactos da criança com a língua, sendo esta, simultaneamente, agente e paciente da identificação dos aspectos específicos no idioma da comunidade.

Mais recentemente, Castro & Ferrari-Neto (2007) procuram observar se as crianças em aquisição do PEL1 e do PBL1 processam a concordância de número de forma idêntica, isto é no âmbito do SDET. Para o estudo, os autores tomaram em consideração

três perspectivas com assunções diferentes acerca da localização do traço interpretável de número no SDET: (i) em termos minimalistas, o número interpretável situa-se em DET e o não-interpretável em N (Magalhães, 2004); (ii) a expressão morfológica de número expressa-se na categoria funcional SNUM e operação de AGREE (Brito, 1996); (iii) na perspectiva da MD, o morfema de PL *-s* é do tipo *singleton* no PB e do tipo dissociado no PE (Costa & Silva, 2006b). Note-se que “*esta última proposta acentua o carácter de micro-variação paramétrica relativamente à expressão morfológica de número entre duas variedades, assumindo que o D é a posição relevante para a interpretação semântica (em que se inclui a referência de pluralidade) de um DP*” (Castro & Ferrari-Neto, 2007:67). Assim, é esperado que a referência PL se recupere a partir de uma marca morfofonológica unicamente em DET.

Lembre-se que, para o PB, Ferrari-Neto, Corrêa & Augusto (2005) verificaram que as crianças de cerca de dois anos percebem informação morfofonológica relativa a número a partir da categoria DET, processando a concordância de número no âmbito do SDET, numa fase em que, não necessariamente, a morfologia de número se encontra estabelecida na produção (Ferrari-Neto, 2003; Simões, 2004). Dos mesmos resultados se inferiu que a posição estrutural DET já está activa nesta fase inicial do processo de aquisição da linguagem em PB e que a mesma desempenha um papel fundamental na identificação de elementos de uma classe aberta (Höhle *et alii*, 2002:108) e, conseqüentemente, na aquisição de novos nomes (Corrêa, 2001, 2006).

No que diz respeito ao PE, alguns estudos baseados em dados de produção apontam também para a hipótese de as crianças serem sensíveis tanto à posição estrutural sintáctica como à presença de marcas de acordo morfológico de número, uma vez que produzem proto-determinantes em estágios precoces do processo de aquisição da língua (Freitas & Miguel, 1998; Faria, Freitas & Miguel, 1998, 2001; Soares, 1998; Costa & Santos, 2004 [2003]; Castro, 2007; Castro & Ferrari-Neto, 2007). Todavia, Faria, Freitas & Miguel (2001) e Costa & Santos (2004 [2003]) argumentam que os dados disponíveis não concedem evidência convincente para que o proto-determinante não seja visto como um proto-morfema prosódico. Desta forma, o estudo de Castro & Ferrari-Neto (2007) visa testar qual é a posição relevante do morfema para a interpretação da pluralidade por

crianças em processo de aquisição do português, confrontando as duas variantes, PB e PE: (i) se nas duas posições (DET e N); (ii) ou se apenas em uma (DET ou N).

No caso das crianças em aquisição do PE, os resultados apontam para um efeito significativo de *expressão morfológica do número*, com uma quantidade maior de respostas PL para as condições elicitadoras de resposta PL. Contrariamente, e tal como fora revelado no estudo de Ferrari-Neto, Corrêa & Augusto (2005), as crianças em processo de aquisição do PB sugeriram uma interferência da marca –s na marcação PL nos nomes terminados em /s/ no singular. Desta forma, as crianças portuguesas tratam de forma indiferenciada as duas condições do conjunto de estímulos elicitadores de PL, assim como as duas condições do conjunto de estímulos não gramaticais (para PL). Opostamente, as crianças brasileiras entendem os dois estímulos de forma diferenciada, podendo inferir-se que a posição relevante para a percepção da marca de pluralidade é DET, e não N. Daí resulta que as crianças portuguesas identificam a informação morfológica relativa ao número da sua variedade, como identificam também a posição DET como sendo a do morfema de PL para a interpretação da pluralidade. Desta forma, a *performance* destas crianças está mais próxima da gramática-alvo, levando os autores do estudo a relacionarem a constatação com o facto de o sistema de marcação de número em PE ser mais estável. Por outro lado, como as crianças brasileiras são confrontadas com um *input* que obedece a padrões inconsistentes, revelador de várias gramáticas em competição, apresentam um sistema morfológico de marcação de número no SDET que se estabiliza mais tardiamente (Castro & Ferrari-Neto, 2007:74-75).

A comparação dos resultados obtidos para o PE e o PB vem também corroborar os pressupostos de Costa & Silva (2006b), confirmando-se não só que DET é a posição relevante para a interpretação da referência PL de um SDET mas também que a marca morfofonológica de PL no nome é redundante (Castro & Ferrari-Neto, 2007:75), uma vez que também as crianças portuguesas recuperam a referência PL a partir de uma marca morfofonológica apenas em DET, ainda que essa não seja a expressão gramatical da sua variedade. O exposto permite concluir que as crianças com cerca de dois anos de idade são sensíveis à expressão morfológica de número da sua língua, tanto em PE como em PB, e reconhecem que a informação crucial de número é extraída de determinantes nas duas variedades. Assim, ao que tudo parece indicar, a categoria DET estará activa no

processo de aquisição desde uma fase bastante precoce, podendo as crianças valer-se, exclusivamente, da informação morfofonológica relativa ao número para atribuição de referência PL de um SDET flexionado em número, sendo a informação de número no nome entendida como redundante (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75). Assim, o que conta é a expressão da pluralidade em DET, independentemente da variedade de português. Este aspecto torna ainda feliz a hipótese de Costa & Silva (2006b) acerca da variação na expressão morfofonológica de PL em PE e PB.

Numa outra perspectiva, o facto de as crianças colocarem correctamente o adjectivo, desde o início da aquisição, argumenta, igualmente, a favor de elas conhecerem precocemente a sintaxe do SDET, apesar de realizarem variação no uso da concordância. Esta particularidade sugere que o SNUM já está activo desde o início da aquisição, uma vez que a operação de elevação do N é aplicada de forma correcta. Paralelamente, a questão da aquisição do parâmetro *N-drop*, comum no espanhol e no português, vem também reforçar a hipótese acerca do conhecimento precoce da sintaxe do SDET. Estabelecendo uma relação entre estes dois aspectos, Snyder, Senghas & Inman (2001:172) concluem que há relação entre concordância e *N-drop*.

Ao que tudo indica, as crianças detêm os rudimentos do género e número desde a fase mais precoce das aquisição, mas apenas dominam essas categorias mais tardiamente. Este aspecto ficar-se-á a dever às distintas características morfológicas e prosódicas das diferentes línguas, como o comprova o facto de os determinantes surgirem primeiramente no espanhol e no português e, só depois, no francês, alemão e inglês (Godinho, 2005:145). Ainda assim, é preciso ter em conta que a aquisição de género ocorre primeiro do que a de número (Müller, 1994a; Franceschina, 2002, 2005; Corrêa, Name & Ferrari-Netto, 2004; Montrul, 2004; White *et alii*, 2004) um aspecto que Lucchesi (2000a) sustenta, ao analisar sintagmaticamente a inserção de marcas de género no dialecto de HEL-Ba. Deste modo, a captação da regra da concordância de número, por ser funcional, implica, quase necessariamente, a aquisição da regra da concordância de género, que é lexical. Paralelamente, os resultados do nosso estudo (ponto 3.6.7) vêm mostrar que, no plano mórfico, o oposto não constitui condição necessária, ou seja, que a aquisição da concordância de género não implica, necessariamente, a posterior aquisição da concordância de número em SN's de estrutura

reduzida. Koehn (1994:34), em estudos sobre aquisição bilingue das línguas alemã e francesa como L1's, apresenta duas abordagens que podem dar conta do fenómeno da aquisição de morfologia de género e número no SN:

- (i) Item e Processo, que considera que itens e afixos estão separados do radical. De acordo com este princípio, ocorre um processo que une ambos os elementos e faz com que a aquisição das marcas morfológicas seja entendida como um aprendizado de formas (alomorfes) e regras;
- (ii) Modelo de Esquemas (*Schema Model*), que defende que não existem apenas relações entre formas básicas e derivadas em termos de regras de afixação, mas também interligações a nível fonológico, que podem representar generalizações. Deste modo, os afixos não serão armazenados independentemente dos radicais com os quais podem ocorrer, levando a que a estrutura interna das formas morfológicamente complexas seja reconhecida por comparação com outras formas, com recurso a conexões lexicais.

Para Koehn (1994:35), a aquisição de marcas morfológicas é explicada de acordo com o Modelo de Esquemas, estando a gramática sujeita a um processo de maturação (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997) e que passa pelas seguintes etapas: (i) o desenvolvimento do conceito semântico subjacente para número (um ou mais do que um); (ii) o reconhecimento de que género e número são sistematicamente codificados em categorias sintáticas específicas; (iii) a aquisição das realizações morfológicas apropriadas desses traços (e de fenómenos de concordância com eles relacionados).

Assim sendo, o falante em processo de aquisição lida inicialmente com a palavra como um todo (Koehn, 1994:49; DeGraff, 1999b:482; Faria Freitas & Miguel, 2001:53), pelo que a aplicação da pluralização não depende da aplicação da marca de género, prevalecendo o Modelo de Esquemas como base para novas formações, o que leva a que a aquisição de marcas se processe com recurso à indução. Por conseguinte, se se considerar que o ponto de partida para a ASL é a gramática da L1 e que as propriedades computacionais da L1 se transferem em bloco para as fases iniciais da ASL, excepto as matrizes fonéticas e os itens léxico-morfológicos (Schwartz & Sprouse, 1996; Sprouse

& Vance, 1999; Siegel, 2006; Sprouse, 2006), a aquisição das marcas dar-se-á por comparação e analogia, acabando as estruturas sintagmáticas do PA por reflectirem um sistema de pluralização idêntico ao do santomense e, conseqüentemente, dos ancestrais substratos deste, isto é, das línguas do grupo níger-congo atlântico. Paralelamente, como as distinções sexuais não são importantes para estas línguas, a marcação de género pode limitar-se ao núcleo do SN (ponto 1.4.2 do presente trabalho), não ocorrendo concordância em género em virtude de os outros constituintes da cadeia sintagmática permanecerem inalterados. Conclui-se, então que a correlação entre concordância de género e concordância de número não é categórica a nível atomístico, indicando não só que “*o encadeamento/aquisicional género => número se verifica de forma mais efectiva no plano sintagmático do que no mórfico*” (Lucchesi, 2000a:265) mas também que a não-concordância de natureza criouliizante tende a ser simultânea. Como simultânea tende a ser também a aplicação dos mecanismos de natureza descriuolizante, confirmada pelos dados de Lucchesi (2000a) para a fala de HEL-Ba, indiciando a intervenção do Princípio da Coesão Estrutural (Haiman, 1983:782) nesta.

Conforme observado por Godinho (2005:145), os fenómenos a que fizemos referência na aquisição da L1 estabelecem paralelismos com o processo de ASL, especialmente no que concerne à aquisição do PtgL2. A diferença reside no facto de, normalmente, a variação desaparecer bastante cedo na aquisição de L1, mas prolongar-se na ASL, chegando a fossilizar-se na interlíngua. Tal ficará a dever-se à incapacidade de aquisição dos traços virtuais não especificados das categorias funcionais (FFFH), já que estas deixam de estar disponibilizadas depois do período crítico do desenvolvimento.

### **2.6.5. Aquisição do sintagma determinante em L2**

A variação registada no processo de ASL é importante para se observar a problemática em torno da refixação dos parâmetros e para análise de aspectos metafísicos relacionados com a aquisição das categorias funcionais na L2 (Godinho, 2005:145). Partindo da aceitação, em linguística, “*que a morfologia explícita conduz à aquisição da sintaxe*” (Montrul, 2004:71, *apud* Godinho, 2005:146), passaremos a analisar a questão dos dados que poderão fornecer evidência sobre a forma como se dá a

aquisição das categorias funcionais da LA, tendo em conta que esta possui traços que não existem na L1 dos aprendentes. Considerando que a aquisição das categorias funcionais pressupõe a aquisição quer de traços quer de formas morfofonológicas, é lícito questionar se os falantes de uma L2 são capazes de adquirir plenamente os traços que não existem na sua L1 e, conseqüentemente, se a variação registada resulta da impossibilidade de adquirir as categorias funcionais marcadas por tais traços (Godinho, 2005:146).

Tendo em conta os pressupostos sobre ASL que avançámos nos pontos 2.3 e 2.4.1.1 do presente trabalho, passemos a observar as diferenças registadas a nível de traços no NUM de algumas línguas que possuem determinantes e uma categoria funcional SDET. Em inglês, a ocorrência de movimento não-visível do N não só determina a ordem ADJ > N como origina concordância fraca em género e número entre determinantes, nomes e adjectivos (Godinho, 2005:147). Deste modo, o SDET em inglês vai apresentar diferenças substanciais em relação ao SDET em espanhol ou português, sendo a mais relevante registada a nível da força dos traços de NUM, mais fracos na primeira língua e mais forte nas outras duas (Carstens, 1991; Ritter, 1991; Valois, 1991; Bernstein, 1993; Montrul, 2004).

Normalmente, os falantes em processo de ASL são foneticamente mais elaborados do que as crianças em situação de aquisição da L1. Este aspecto leva os primeiros a não produzirem proto-determinantes, sendo esta uma das diferenças mais sintomáticas a nível das duas aquisições (Godinho, 2005:147). Contudo, e exceptuando esta particularidade, alguns autores entendem que, em tudo o mais, os dois processos são bastante semelhantes. Produções de aprendentes de espanhol L2 apresentam realização de *N-drop* em contextos idênticos aos da L1 (SADJ, SPREP e SCOMPL), implicando que, *“independentemente de não haver concordância e a omissão de determinantes ocorrer na fase em que ocorrem os proto-determinantes, os restantes aspectos têm um desenvolvimento semelhante nos dois tipos de aquisição”* (Godinho, 2005:147).

A variação registada a nível da concordância em género e número na ASL tem sido constatada por alguns autores (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) e entendida como dificuldade na aquisição dos traços que não existem na L1, já que as categorias deixam de estar



disponibilizadas depois do período crítico do desenvolvimento (FFFH – Godinho, 2005:147). Este aspecto é evidente nos estudos sobre aquisição do género no espanhol L2 por parte de falantes de francês L1 (dois idiomas detentores de morfologia de concordância), não se tendo detectado variação significativa na aquisição dos traços do SDET (Bruhn de Garavito & White, 2002:167, *apud* Godinho, 2005:148). Contudo, no que diz respeito à questão da concordância de género, constatou-se uma ultrageneralização da forma masculina em palavras que requerem o género feminino, uma particularidade já notada em outros estudos (Tarone, Frauenfelder & Selinker, 1976; Cain, Weber-Olsen & Smith, 1987; Finneman, 1992; Dewaele & Véronique, 2001).

Dos estudos referidos interessa ainda destacar duas particularidades idênticas em aquisição de L1 e ASL: *“a concordância em género é primeiro estabelecida com o determinante e depois com o adjetivo. O número de erros também é menor com determinantes do que com outros modificadores”* (Godinho, 2005:148). Este segundo aspecto foi também observado em outros estudos sobre aquisição do género (Lucchesi, 2000a; Figueiredo, 2009a), fazendo pressupor que o género masculino parece encerrar o valor de *“default”*, daí a opção pela sua aplicação, quando não existe a certeza sobre o género do nome. Contudo, este detalhe não seria confirmado por McCowen & Alvord (2006), cujo estudo não evidencia diferenças significativas na realização das formas masculina e feminina, um pressuposto vem ao encontro dos achados de Lucchesi (2000a) sobre a aquisição do género pelos falantes de HEL-Ba. De facto, os falantes submetidos a observação exibiram um elevado grau de variação afectando ambas as formas, sendo algumas delas regularizadas posteriormente. Segundo McCowen & Alvord (2006:168), a tendência para o uso da forma masculina só acontece quando o falante não tem a certeza acerca do género do nome. Deste modo, este tipo de variação não pode ser entendido como uma generalização, visto registar-se, de forma esporádica, a nível individual. Por outro lado, a regularização da variação evidencia que os traços podem ser adquiridos numa L2 depois de ultrapassado o período crítico da aquisição, conforme preconizado pelos defensores da FTFA (Cook, 1994; White, 1989, 2003; Bruhn de Garavito & White, 2002; Montrul, 2004; Cook & Newson, 2007 [1988]). Por conseguinte, as dificuldades reveladas na aquisição não são da responsabilidade da L1,

resultando a variação no género do facto de os falantes em processo de ASL não conseguirem fazer a correspondência do traço de género à sua forma morfofonológica (Bruhn de Garavito & White, 2002, *apud* Godinho, 2005:149).

Estudos sobre aquisição do género em espanhol L2 por falantes de inglês L1 (língua que não detém género) e francês L1 (idioma com género) mostram que os primeiros apresentam maior dificuldade na aquisição deste traço. (White *et alii*, 2004, *apud* Godinho, 2005:149). No que concerne à aquisição do número, acontece nivelamento final entre ambos os tipos de falantes, apesar de os ingleses apresentarem um grau de variação ligeiramente mais elevado nas fases iniciais da aquisição, mas que é regularizada posteriormente. Por seu lado, a aquisição de concordância do masculino é a mais problemática e, novamente, são os franceses a apresentar a melhor *performance*. Estas observações permitem concluir não só que a L1 interfere no estado inicial da ASL (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006) mas também que os erros iniciais de morfologia flexional podem ser ultrapassados com a aquisição da morfologia flexional e itens flexionais, refixando-se, assim, os parâmetros. Todavia, o estudo acaba concluindo que, no geral, a interferência da L1 não é relevante na ASL, ao contrário do que sucede com o grau de proficiência, que é determinante neste tipo de aquisição (White *et alii*, 2004, *apud* Godinho, 2005:150).

Os aspectos expostos permitem perceber que o espanhol e o português têm projecções de SDET desde a fase inicial da aquisição e também que, apesar de ocorrer variação na concordância em género e número durante determinados estádios de aquisição da L1 e ASL, ela tem tendência a regularizar-se, fazendo pressupor acesso total à GU em ASL (Cook, 1994; White, 1989, 2003; Bruhn de Garavito & White, 2002; Montrul, 2004; Cook & Newson, 2007 [1988]). Por seu lado, outros autores (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) propõem que a gramática da interlíngua se restrinja às propriedades formais da L1, o que impossibilita a aquisição de novas propriedades, em virtude de elas não estarem instanciadas nas L1's. Como tal, a aquisição dos traços não será problemática se as duas línguas os possuírem, uma vez “*estão activados na representação da L1 e, por isso, disponíveis na gramática da IL*” (Godinho, 2005:150).

Os estudos abordados evidenciam igualmente que, apesar de se registar variação nas etapas iniciais da aquisição, as categorias sintáticas já se encontram universalmente presentes desde a FL0, manifestando-se, precocemente, através da produção de proto-determinantes. No entanto, ainda que os determinantes sejam adquiridos bastante cedo, a concordância em género e número a nível do SDET só acontece posteriormente em ambas as aquisições, podendo persistir mais tempo na ASL, a ponto de se fossilizar em alguns falantes (cf. Godinho, 2005:151). Deste modo, a CPL-var no SN do PA poderá ser estar intimamente ligada ao facto de o santomense e as línguas do grupo níger-congo atlântico, a exemplo do que Godinho (2005:385) constatou relativamente aos falantes de cantonenseL1 em aquisição de PtgL2/LE, não possuírem traços de género e número.

#### **2.6.6. Aquisição da concordância de PL no sintagma determinante**

A aquisição da concordância de PL no SDET pressupõe a detecção da categoria número e a realização dos processos morfossintáticos inerentes à mesma, levando as regras da formação das palavras (*Word Formation Rules*) a actuarem em duas direcções (Aronoff, 1976; Spencer, 1991; Cerqueira, 1994:119):<sup>88</sup>

*“(a) As regras são responsáveis por garantir a formação das palavras da língua «estabelecendo as operações que produzem os resultados desejados»;*

*(b) As regras operam em sentido inverso, não para formar palavras mas para analisar as suas partes constituintes” (Godinho, 2005:152).*

Estes aspectos implicam que as flexões só poderão ser alojadas se a palavra for analisada em função dos traços morfológicos das diferentes categorias morfossintáticas accionadas. Deste modo, Faria, Freitas & Miguel (2001:51-54, *apud* Godinho, 2005:153-154 e nota 204), ao debruçarem-se sobre as projecções funcionais do SN do PE, verificaram que o morfema PL realizado com a fricativa /s/, em final de palavra, acontece, geralmente, em sílabas não acentuadas, em virtude de o acento principal em PE gerar, por norma, palavras paroxítonas. Por outro lado, como as sílabas tónicas protegem a informação fonológica dos processos fonéticos de apagamento, (Delgado-Martins, 1996:319), as fricativas morfológicas /s/, marca de PL, deveriam ser os últimos itens a emergir, visto ocorrerem em posição final, ou seja, em sílabas átonas. Todavia, e

conforme apontado, listado e comentado por Godinho (2005:153-155), a propósito do momento em que sobrevêm as fricativas que determinam a pluralização, a *saliência fónica* não é relevante para a inserção das mesmas. Face a este achado, que corrobora o postulado de Cerqueira (1994) acerca da disponibilidade inata de um dispositivo para recepção e descodificação do *input* fonológico, verifica-se que “*a aquisição «precoce» das fricativas em codas morfológicas, em posição final de palavra, em sílabas não acentuadas, deve-se ao seu estatuto morfológico no sistema da língua (português europeu) e à posição (final) que ocupam na palavra, e que as crianças são capazes de as isolar e interpretá-las como um constituinte separado – um morfema – ao nível da palavra*” (Godinho, 2005:154-155). Assim, apesar de as crianças revelarem incapacidade de isolar e interpretar separadamente os morfemas a nível das palavras nas fases precoces da aquisição bilingue (cf. projecto DUFDE, Koehn, 1994:49) e ASL (cf. DeGraff, 1999b:482), conseguem perceber as propriedades semânticas e morfossintáticas dos elementos lexicais.

As realizações que consubstanciam o nosso *corpus*, a exemplo do que sucede com os dados produzidos por falantes de PtgL1 (Guy, 1981a; Cerqueira, 1994; Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Baxter, 2004, 2009; Jon-And, 2008, 2009) e PtgL2/LE (Godinho, 2005), não só espelham realizações decorrentes do contacto entre línguas mas evidenciam também que, a exemplo do que sucede com as produções registadas na aquisição de L1 e bilingue, a marcação de PL pode revelar variação, não incidindo em todos os constituintes do SN (ou do SDET) que a requerem. Não obstante, o facto de ocorrerem marcas de PL, mesmo que esporádicas, faz pressupor que os falantes em questão identificam a fricativa /s/ em final de palavra como um marcador do valor PL. Assim, estamos perante dois padrões de marcação, semelhantes aos identificados por Godinho (2005:156) nas produções de falantes do cantonense L1 em aquisição do PtgL2/LE: (i) o sistema de marcação PL respeita as regras de aplicação da flexão em número do português e esta é inserida em todos os elementos da frase que a requerem (exemplo [190]); (ii) o sistema de marcação PL não respeita as regras de aplicação da flexão em número do português e esta não é inserida na totalidade dos elementos da frase que a requerem (exemplo [191]):

[190] PA:  
*os homens holandeses vieram* [CASTEH1]

[191] PA:  
*As quatro roda pa cima* [OSVALH1]

Considerando que os quantificadores, determinantes e modificadores adjetivos concordam com o núcleo sintagmático, então estes elementos deverão ser portadores da marca morfológica PL patente no mesmo (exemplo [190]). Todavia, o exemplo [191] contraria este entendimento, evidenciando que este tipo de traço configura uma categoria vazia no núcleo. Deste modo, e de acordo com a Hipótese do SDET (Picallo, 1991; Carstens, 1991; Ritter, 1991; Cerqueira, 1994; Valois, 1991; Bernstein, 1993; Cinque, 1994; Koehn, 1994; Müller, 1994a; Montrul, 2004), que postula a existência de categorias funcionais entre o SDET e o SN, então é lícito inferir que o traço de número [+/- plural] do núcleo nominal é mapeado em NUM na sintaxe (visível ou invisível), consoante a sua força morfológica, a qual pode determinar a elevação do N, a partir do seu posicionamento original (Cerqueira, 1994:121, *apud* Godinho, 2005:157).

Segundo o Princípio da Projecção Alargado da Teoria X-barras (Rizzi, 1982; Jaeggli, 1984; Burzio, 1986; Jaeggli & Safir, 1989; Grimshaw, 2005 [1991]), N, um item lexical, e DET, um elemento funcional, compartilham os mesmos traços de categoria sintáctica [-V, +N], apesar de, funcionalmente, pertencerem a classes distintas. A distinção entre ambos, a nível de traços funcionais, é justificada por em termos de desdobramento (Cerqueira, 1994:121-125; Godinho, 2005:157-165), ou seja, é assegurada por um traço funcional, no caso o traço {F}, que se vai desdobrar em dois: {F0}, para as categorias lexicais; e {F1}, para as categorias funcionais. Assim, o valor {F} de um nó não consubstancia parte da sua análise categorial, visto não possuir o mesmo valor dos traços categoriais na formação de projecção alargada nem ser um traço binário que estabeleça interacção com outros traços categoriais (Godinho, 2005:157-158). Desta forma, as categorias funcionais são essencialmente relacionais, já que estabelecem relações com categorias lexicais, delas se constituindo a sua parte funcional, como sucede com o SDET, relativamente ao N. (Cerqueira, 1994:121; Godinho, 2005:158).

Tendo em conta a questão do desdobramento do traço funcional  $\{F\}$ , proposta por Cerqueira (1994:121), as categorias derivadas de DET e N, listadas por Godinho (2005:158), serão as seguintes:

(1)	DET	$[-V, +N]$	$\{F1\}$	L0
	DET'	$[-V, +N]$	$\{F1\}$	L1
	SDET	$[-V, +N]$	$\{F1\}$	L2
(2)	N	$[-V, +N]$	$\{F0\}$	L0
	N'	$[-V, +N]$	$\{F0\}$	L1
	SN	$[-V, +N]$	$\{F0\}$	L2

Por seu lado, as projecções não-alargadas (ou de núcleo perfeito) são estabelecidas de acordo com os seguintes parâmetros (Godinho, 2005:158-159):

#### PROJECCÃO DE NÚCLEO NÃO-ALARGADO

$x$  é núcleo perfeito de  $y$ , e  $y$  é uma projecção perfeita de  $x$ , caso:

- (i)  $y$  domine  $x$ ;
- (ii)  $y$  e  $x$  compartilhem de todos os traços categoriais  $[-V, +N]$ ;
- (iii) todos os nós entre  $x$  e  $y$  tenham os mesmos traços categoriais;
- (iv) o valor de S de  $y$  seja igual ao valor de S de  $x$ .

Segundo estes princípios, SDET é projecção de DET e DET', visto compartilhar traços categoriais e flexionais com estes. Porém, SDET não é projecção de N, N' ou SN, visto não compartilharem traços flexionais. Por seu lado, SN é projecção de N' e N.

Quanto às projecções alargadas, são definidas de acordo com os seguintes princípios (Godinho, 2005:159):

#### PROJECCÃO DE NÚCLEO ALARGADO E PROJECCÃO ALARGADA

$x$  é núcleo alargado de  $y$ , e  $y$  é uma projecção alargada de  $x$ , caso:

- (v)  $y$  domine  $x$ ;
- (vi)  $y$  e  $x$  compartilhem de todos os traços categoriais;
- (vii) todos os nós entre  $x$  e  $y$  tenham os mesmos traços categoriais;
- (viii) o valor de S de  $y$  seja maior que o valor de S de  $x$ , onde  $n$  actue entre  $x$  e  $y$ , se  $y$  dominar  $x$  e  $n$ . Por outro lado,  $n$  domine  $x$ , mas  $x$  não domine  $y$ .

De acordo com estes pressupostos, é possível combinar projecções de núcleo perfeito com outras maiores, o que leva o SDET a deter cinco núcleos alargados: DET', DET, SN, N' e N. Simultaneamente, SDET é projecção alargada de DET, DET', SDET, N, N' e SN, o que determina que um sintagma seja, ao mesmo tempo, COMPL de um núcleo e parte da mesma projecção alargada, caso observe os requisitos exigidos pelo Princípio da Projecção Alargada (Cerqueira, 1994:122; Godinho, 2005:159).

Se dos núcleos de uma projecção alargada é projectada uma dada informação, todos os seus traços deverão concordar com a projecção SN > SDET, manifestando-se caso a morfologia registe os seus valores (Grimshaw, 2005 [1991]:114). Todavia, nos casos em que não ocorre concordância entre DET e N a manifestação não é visível, já que esta se torna também alargada, devido à projecção (Godinho, 2005:159-160).

Além destas determinações, a Teoria da Projecção Alargada garante igualmente que a projecção COMPL de um núcleo seja parte integrante dessa projecção, visto compartilharem traços funcionais e categoriais. Portanto, se número é traço de N, sê-lo-á igualmente de DET, o que leva a projecção de SDET e DET a incluírem traços do núcleo alargado N e SN. Perante tal, a concordância é um fenómeno derivado que assegura a correcta elaboração das estruturas nominais (Cerqueira, 1994:123; Godinho, 2005:160).

Sendo a Teoria da Projecção Alargada uma das componentes da Faculdade da Linguagem, o pareamento do traço morfológico de número entre todos os elementos do SN, que permite elaborar a concordância visível, só será possível caso o falante receba a informação acerca do traço *-s* (Godinho, 2005:160). Contudo, tal pode não suceder, conforme exemplos [190] e [191]. Nestes casos, a não-realização fonética da marca PL licenciada por {s} obedecerá a uma operação de identificação semântico-funcional, idêntica à que afecta a categoria sintáctica *pro* Cerqueira (1994:125, *apud* Godinho, 2005:161). De facto, a ocorrência de sujeitos pronominais nulos em línguas como o português depende da elaborada morfologia do seu sistema flexional de concordância, o qual permite a recuperação dos traços gramaticais do sujeito a partir do conteúdo morfológico das terminações verbais (Raposo, 1992:477-478). Deste modo, a projecção alargada vai não apenas assegurar os licenciamentos formal e funcional mas também permitir que se formulem os pressupostos que determinam o licenciamento de {s} nulo, os quais, por seu lado, permitem que se compreenda o porquê de não acontecer

licenciamento de {s} nulo quando o SN é configurado unicamente por N (Godinho, 2005:161):

LICENCIAMENTO DE {S} NULO

{s} pode ser nulo, caso:

- (ix) seja parte integrante de uma projecção alargada;
- (x) a sua presença seja notificada por algum item da projecção.

Levando em consideração quer os princípios do licenciamento de {s} nulo quer o facto de a concordância ser um fenómeno derivado, Godinho (2005:161) estabelece os considerandos finais sobre as condições de base que determinam a segunda: (i) advir da relação entre SPEC e núcleo; (ii) resultar de projecção. No segundo caso, caberá ao SDET, enquanto projecção máxima de N e SN, projectar os traços destes dois núcleos alargados, permitindo que eles estabeleçam o pareamento de traços entre si. Por sua vez, a marca PL em DET vai permitir que a mesma não seja visível em N, ou seja, vai tornar possível o licenciamento de {s} nulo neste.

Tendo em conta tais considerandos, determinadas realizações dos falantes de Almojarife (exemplo [192]) permitem constatar que as mesmas encerram representações sintácticas distintas, identificadas consoante o contexto em que são produzidas:

[192] PA:

*lodava as conta com el poque a... a roça piquena.* [CELESH3]

Nesta realização, o SN “*a roça piquena*” permite duas configurações estruturais (cf. Godinho, 2005:162):

- (i) SDET composto pelo artigo “a”, o nome “roça” e o adjectivo pós-nominal “piquena” (pequena);
- (ii) Uma construção predicativa com apagamento de cópula, “*a roça era piquena*”, na qual o adjectivo é parte integrante do SV.

Partindo do princípio que, na aquisição de L1 e ASL, N e DET+N são produzidos numa mesma fase (Montrul, 2004:42-43; Godinho, 2005:161-162), é lógico inferir que



os falantes almorarifanos dominem a categoria funcional SDET, apesar de nem sempre a realizarem. Este entendimento poderá ser extensivo às questões que determinam a ausência visível, na sintaxe, da marcação PL em todos os itens do SDET (exemplo [191]). Assim, o facto de os falantes de Almorarife não marcarem o número em todos os elementos do SN não implica, necessariamente, que estejam a violar princípios gramaticais, uma vez que essas sequências se assemelham às que são realizadas na FLf de falantes nativos da LA (cf. Godinho, 2005:162). Para a autora, os pressupostos sobre a elaboração da concordância PL, por um lado, e acerca da forma como se manifesta a ausência de marcas PL, por outro lado, argumentam a favor da proposta que o ponto de partida para a ASL é a gramática da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), pelo que a variação resulta de um problema na correspondência de traços, e não da impossibilidade de aquisição destes, seja por falha no restabelecimento dos parâmetros (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) seja pela ausência dos mesmos traços na L1 (Leung, 2003:206, *apud* Godinho, 2005:162). Portanto, com a aquisição intermédia de morfologia flexional relevante e de itens lexicais, é possível refixar os parâmetros e atingir um estágio final de aquisição.

Não obstante, segundo Hawkins (1993, 1998, 2001), os traços [-interpretáveis] não estão presentes no processo de aquisição, em virtude de serem os mais difíceis de adquirir. Por seu lado, Franceschina (2002, *apud* Godinho, 2005:163) sustenta que a aquisição dos traços [-interpretáveis] de número e género não oferece dificuldades na ASL, caso os falantes os possuam nas suas L1's. Como este tipo de traços é opcional nas línguas, os falantes nativos daquelas que os não possuem experimentam dificuldades na sua aquisição, quando em processo de ASL's nas quais eles se encontram instanciados. É o que acontece com os falantes da generalidade dos idiomas do grupo níger-congo atlântico ou do santomense, submetidos à aquisição do Ptg, uma língua parametrizada, já que as suas L1's não detêm concordância do tipo sufixal em número e género. Deste modo, é lícito pressupor que a CPL-var registada no SN dos informantes do PA poderá ser motivada pelo facto de o traço de número não estar instanciado nas L1's dos seus antepassados (cf. Godinho, 2005:164). Estudos sobre a concordância PL no SN do PB (Braga & Scherre, 1976; Braga, 1977; Scherre, 1978; Ponte, 1979; Nina, 1980; Guy,

1981a; Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Tieppo, 2003; Baxter, 2009) e de variedades africanas de português (Moreno & Tuzine, 1998; Baxter, 2004, 2009; Figueiredo, 2008, 2009a, 2009b; Jon-And, 2008, 2009) têm revelado que a presença ou ausência de marcas não é aleatória (Tarallo, 1986:8), já que os contextos linguísticos e/ou extralinguísticos que as motivam permitem que se prevejam as tendências da marcação (ponto 2.5 do presente trabalho). Alguns destes trabalhos sobre o PB, como os de Guy, (1981a; *apud* Godinho, 2005:164), Lopes (2001, *apud* Godinho, 2005:164), P. Andrade (2003) ou Baxter (2009), apontam a hipótese de a variação resultar de um processo de transmissão linguística irregular, que afectou a aquisição da língua portuguesa por parte de ameríndios e escravos africanos deslocados geográfica e culturalmente pelos colonizadores. Quantos às variedades africanas, exceptuando o trabalho de Moreno & Tuzine (1998), que não aborda a questão da aquisição, todos os outros referem que a reestruturação terá sido condicionada por factores que ancoram na aquisição do PtgL2 e PtgL1 nativizado, em situação de contacto que envolveu influências das línguas nativas do grupo níger-congo atlântico ou línguas crioulas. Alguns destes trabalhos têm também corroborado a hipótese levantada por Guy (1981a:3001-302) de que o ancestral substrato africano será a potencial base que origina a CPL-var no PB, cuja marcação, tal como nas variedades referidas, revela tendência a ser inserida no elemento imediatamente à esquerda do núcleo do SN.

### **2.7. Aquisição das categorias funcionais da L2: os casos do santomense e do PA**

Os radicais dos nomes não possuem flexão em género e número no santomense nem nos seus substratos. Nestes, os classificadores nominais, apensos inicialmente, regem a concordância PL. Nas línguas bantu, por exemplo, todos os elementos que compõem o SN e se relacionam com o nome recebem marcação temática, mantendo-se invariáveis os radicais dos constituintes da cadeia sintagmática. Estes aspectos determinam que os falantes destas L1's, em situação de aquisição do PtgL2, tenham que fixar novas propriedades sintácticas na interlíngua, a fim de elaborarem enunciados com acordo no SN, já que as regras da LA determinam que a concordância no PE se processe com recurso a morfemas sufixais e dissociados ao longo da estrutura sintagmática.

A possibilidade de ocorrer conflito na aquisição dos traços que determinam a marcação em género e número na interlíngua leva-nos a observar como se processa a concordância no SN do PE, semelhante à que ocorre no domínio da frase. Sem nos determos muito em pressupostos generativistas, interessará reter que a concordância ao nível do SN dependente, entre outras, das seguintes condições: (i) relações entre SPEC e núcleo; (ii) categorias funcionais relacionadas com a concordância; (iii) movimento do núcleo internamente ao SN; (iv) mecanismos pronominais. A partir da Teoria X-Barra, obtivemos os diagramas das Figs. 19 e 20, que determinam que a categoria funcional NUM, localizada entre o DET e o SN, é aquela onde se situam os traços de número (Picallo, 1991; Carstens, 1991; Ritter, 1991; Valois, 1991; Bernstein, 1993; Cinque, 1994; Montrul, 2004). Deste modo, se tivermos em conta o posicionamento rígido de determinados adjectivos do português, como é o caso dos adjectivos relacionais ou que digam respeito a formas e cores (exemplos [193a] e [193b]), o nome movimenta-se para permitir a concordância, podendo este movimento traduzir uma propriedade económica da língua, à qual os falantes recorrem para evitarem a inserção de mais elementos lexicais na estrutura frásica. Contudo, como a posição DET se encontra preenchida, é necessário criar mais derivações nas árvores sintagmáticas para se inserir o nome. Deste modo, para o exemplo [193a] teríamos, então, a configuração da Fig. 22:

- [193] PTG:  
 a. *A bola vermelha.*  
 b. *\*A vermelha bola.*

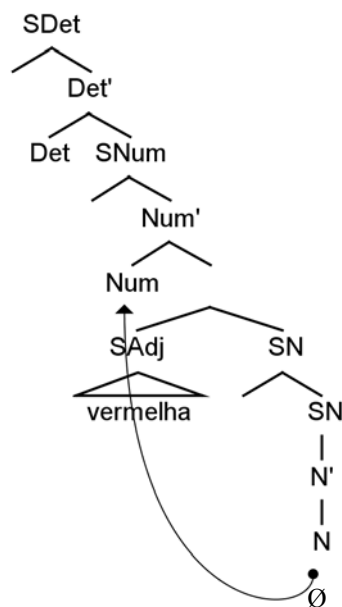


Fig. 22. Movimento de N para NUM.

Como o ADJ detém uma posição fixa, é o N que sofre elevação, transitando da área dos itens lexicais para a zona dos elementos funcionais, ou seja, eleva-se da posição de N para a de NUM, assumindo a função de núcleo do SN. A posição de especificador vai então torná-lo morfologicamente visível e, ao entrar em relação próxima com os outros itens do SN, permite que os seus traços [+interpretáveis] de número procedam ao mapeamento dos correspondentes traços [-interpretáveis], originando a concordância sintagmática. No entanto, convém referir que o género possui um traço bastante lexical, revelando-se indiferente na especificação da entidade, enquanto o número possui um traço mais funcional, tendo importância para a construção da referência. Assim sendo, a aquisição de género ocorre antes da aquisição de número no plano da concordância dentro do SN, (Müller, 1994a; Franceschina, 2002, 2005; Corrêa, Name & Ferrari-Netto, 2004; Montrul, 2004; White *et alii*, 2004).<sup>89</sup>

Como as gramáticas do santomense e seus substratos do grupo níger-congo atlântico não apresentam sistema de marcação idêntico ao do PE, a categoria N, ao elevar-se para a posição nuclear, não vai exibir morfema de PL, originando a concordância a nível meramente temático. Tendo ainda em conta que a competência em L2 é igualmente condicionada pela ecologia das línguas, deve dar-se especial atenção ao contexto em que esta ocorre: endolingue ou exolingue. O santomense, substrato do PA, desenvolveu-se em local onde um grupo dominante (os colonos portugueses), que veio de fora com o objectivo de formar uma sociedade nova, estabeleceu contacto com comunidades geográfica e culturalmente deslocadas por esse mesmo grupo dominante. Inicialmente, a população deslocada, constituída por poucos elementos, era submetida a um *input* regular da LA, pelo que apenas fazia uso das suas línguas de origem em situações restritas, isto é, em ambiente familiar. Contudo, esta população variada sobrepôs-se em número ao do grupo dominante em determinado momento da constituição histórica da nova sociedade. Assim, apesar de associada cultural e economicamente a este, a maioria dos membros da sociedade dominada, face a um contacto cada vez mais distanciado em relação aos detentores da LA, via os estímulos linguísticos desta diminuir acentuadamente. Estes aspectos determinaram, então, que o uso das L1's se mantivesse, ao mesmo tempo que se desenvolvia o santomense. Paralelamente, como o sistema das línguas do grupo níger-congo atlântico faladas por eles não possui a categoria DET,

impossibilitava-os de fazerem a correspondência das características formais às formas morfofonológicas (cf. DeGraff, 1999b; Prévost & White, 1999 – MSIH). Deste modo, o DET surgia fonologicamente interpretado como um afixo semântico apenas no início do N (Baxter, 2004:120),<sup>90</sup> e não como uma categoria funcional que determina a construção a nível sintáctico, originando a instanciação de novos/errados parâmetros (Gonçalves, 2004:235) na L2, daí resultando itens gramaticais relexificados a partir de estruturas do substrato e lexemas derivados do superstrato. Posteriormente, a instanciação em cadeia dos novos/errados parâmetros nos DLP's determinou uma situação de transmissão linguística irregular e inviabilizou a possibilidade de um reajustamento paramétrico, fazendo com que os novos/errados parâmetros fossem tomados como modelo para uma nova reinstanciação paramétrica do PAL1. Consumava-se, deste modo, a variação no sentido da mudança que, por um lado, acentua a distância tipológica entre o PE e a capacidade gramatical (ou língua-I) atingida pelos falantes do PA e, por outro lado, faz com que alguns traços do santomense e das ancestrais línguas do substrato níger-congo atlântico sejam retidos no PtgL2 e sobrevivam na gramática do PAL1. Em sequência de tal, algumas propriedades gramaticais do PE perderam-se também no *continuum* de aquisição, originando a *especificidade gramatical* do PA.

Assim sendo, as variações que se detectam no DET do PA terão origem em dificuldades de percepção do tipo fonológico, visando a construção mental de concordância temática, e não tanto do tipo estrutural, já que estas efectuam a concordância a nível sintáctico. Porém, como a aprendizagem da L2 pode acontecer por indução (Selinker, 1972:216), é também de considerar que, a nível mais avançado da aquisição do PA, sobretudo em falantes escolarizados, ocorra a percepção de que se está perante itens lexicais distintos, passando a concordância a operar sintacticamente. Como o valor semântico expresso pelos afixos flexionáveis não é suficientemente robusto para determinar a sua sobrevivência nos estágios mais avançados do desenvolvimento da interlíngua (Plag, 2008a:125), ocorre, posteriormente, uma refixação paramétrica e conseqüente substituição do sistema de pluralização do santomense e dos seus substratos níger-congo atlânticos pelo sistema de pluralização sufixal do PE, levado a cabo pelos falantes bilingues e mais escolarizados de Almoxarife.

## 2.8. A refixação dos parâmetros

No ponto 2.4.1.1 do presente trabalho, referimos que a corrente linguística que defende o acesso indirecto à GU entende que o ponto de partida para a ASL é a gramática da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006). Por conseguinte, as propriedades computacionais da L1 transferem-se em bloco durante as fases iniciais da ASL, com excepção das matrizes fonéticas e dos itens léxico-morfológicos. Nesta conformidade, a construção gramatical é orientada pela GU, que mantém sempre os seus princípios activos, fazendo com que, no processo de ASL, os aprendentes adquiram valores paramétricos a partir dos valores de parametrização da L1 e cometam erros de morfologia flexional.

O Programa Minimalista (Chomsky, 1996 [1995]) advoga que os *traços-phi* não-interpretáveis, entre os quais se encontram os traços de género e número, têm como finalidade activarem as operações sintácticas. Os traços não verificados são valorados e, caso necessário, apagados através da concordância com traços valorados a partir de outro núcleo da estrutura frásica. Deste modo, os *traços -phi* [*ugen*] e [*unum*] vão mapear o conjunto de traços valorados para o mesmo valor binário do traço em questão, que terá de partir de um elemento nuclear, no caso a categoria N, projectada em NUM. No PE, acontece uma relação de especificador em que o núcleo é morfologicamente visível, permitindo que os elementos em questão possam copiar ou compor o seu valor a partir deste núcleo, sendo verificados, posteriormente, por forma a produzir-se a concordância em todos os elementos de núcleo e seus complementos. Os morfemas de género e número são, então, do tipo dissociado, espalhando-se por todos os contextos do mesmo caso dentro do SN. Finalizadas as operações sintácticas, os elementos são remetidos para a FL, com radicais-terminais para um determinado valor. Como temos vindo a referir, no caso do género, ele vai deter características bastante lexicais (revela-se indiferente para especificar a entidade), enquanto o número é mais funcional, construindo a referência.

Sem assumirmos que a variação é provocada por mecanismos de transferência directa da L1, não poderemos, todavia, deixar de levar em consideração que os aprendentes adultos de uma L2 são proficientes, à partida, na sua L1. E no estado inicial da ASL, tal pode determinar que os aprendentes adquiram valores paramétricos a partir dos valores

de parametrização da L1. Como o PE é uma língua que detém valor positivo para ambos os *traços-phi* de género e número ao nível da sintaxe, este aspecto pode ser determinante na fase inicial da aquisição do PtgL2 quer por falantes do santomense quer por falantes dos seus substratos níger-congo atlânticos, já que estes aprendentes detêm L1's caracterizadas por configuração negativa para ambos os *traços-phi* referidos. De facto, tais idiomas possuem um reduzido sistema flexional, pouco elaborado em traços virtuais não especificados (as categorias funcionais). Portanto, não só estão privados de flexão sufixal de género e número como também marcam estas categorias de forma distinta do PE, recorrendo a classificadores pré-nominais que configuram morfemas do tipo *singleton*. Paralelamente, e no caso específico do género, este pode ser assinalado na morfossintaxe dos nomes com o traço semântico [+animado], já que, nas mencionadas línguas, os falantes expressam também o género recorrendo, por vezes, a nomes adjetivados colocados depois do nome, como acontece com os termos *mulher* ou *fêmea* (género feminino) e *homem* ou *macho* (género masculino) (ponto 1.4.2 do presente trabalho).

Como se depreende, esta opcionalidade de fixação de valores poderá levantar problemas na aquisição do PtgL2, levando à instanciação de novos/errados parâmetros (Gonçalves, 2004:235), aspecto este que seria minorado ou eliminado, caso os falantes partissem para a aquisição com valores positivos (cf. Franceschina, 2002:100), à semelhança do que acontece na LA. Por outro lado, e embora o modelo de acesso indirecto à GU determine também que, com a aquisição intermédia de morfologia flexional relevante e de itens lexicais se consegue ultrapassar os erros de morfologia flexional característicos das fases iniciais da aquisição (Schwartz & Sprouse, 1996; Sprouse, 2006), ou seja, se refixem os parâmetros, a verdade é que a FFFH entende que as categorias funcionais, nas quais se encontra incluído o número, deixam de estar disponibilizadas depois do período crítico do desenvolvimento, impedindo a refixação para tais categorias. Em consequência deste aspecto, acontece a ausência em cadeia das referidas categorias funcionais no *input*, o que impossibilita o posterior reajustamento paramétrico de acordo com os modelos da LA, fossilizando-se determinadas estruturas transferidas das ancestrais L1's, e que são responsáveis pela variação registada no PA. Portanto, será importante que se observem também quer as questões inerentes à

concordância variável ao nível do SN do PA quer as condições que poderão determinar as fossilizações e, conseqüentemente, consumarem a mudança da variação.

## 2.9. Concordância variável ao nível do SN

No que diz respeito à concordância variável, o PB tem-se revelado como um caso paradigmático para o debate acerca da visão completamente configuracional da concordância. Ao contrário do PE, que apresenta um padrão uniformizado de concordância verbal para a relação sujeito/verbo e que estabelece paralelos com o da concordância interna ao SDET, o PB exhibe dois padrões de concordância verbal e apenas um para a concordância interna ao SDET, inviabilizando um paralelo de acordo com o que sucede no PE. Nos casos de concordância verbal, verificam-se, então, três situações-base no PE: (i) realizações com ordem frásica padrão (SVO) exibindo concordância entre todos os constituintes frásicos (exemplo [194a]): (ii) aplicação de concordância quando ocorre inversão do sujeito na ordem da frase (exemplo [194b]); (iii) uso de concordância quando se levam a cabo realizações com verbos inacusativos (exemplo [194d]).

[194] PE:

- a. *Os carros são bonitos.*
- b. *Comeram, os meninos, o doce.*
- c. *\*Comeu, os meninos, o doce.*
- d. *Ontem chegaram dez navios ao porto de Lisboa.*
- e. *\*Ontem chegou dez navios ao porto de Lisboa.*

No que diz respeito à concordância interna ao SDET, o PE não levanta problemas em termos descritivos, já que apresenta concordância entre todos os elementos frásicos que a exigem:

[195] PE:

- a. *Os carros.*
- b. *Os carros bonitos.*
- c. *Os primeiros carros.*
- d. *Os meus carros.*

No caso do PA, a similaridade com realizações do PB é bastante evidente, sobretudo no que respeita à concordância interna ao SDET. Assim sendo, os pressupostos avançados para o PB são igualmente válidos para a fala de Almojarife. Concretamente



sobre a concordância verbal, o PA, a exemplo do PB, exhibe variação na relação sujeito/verbo, que pode ser de dois tipos. Desta forma, com a ordem frásica padrão (SVO), a variação pode realizar-se das seguintes maneiras: (i) ocorrendo concordância entre verbo e apenas um dos elementos do SN sujeito, por norma o item pré-nuclear (o determinante), isto é, aquele que revela tendência para receber a marca de PL (exemplo [196]); (ii) ausência de concordância entre verbo e elemento pré-nuclear (exemplo [197]).

[196] PA:  
*Os dirigente não são homens assim* [CASTEH1]  
“Os dirigentes não são homens assim”

[197] PA:  
*Depois as piquena não tá respondê* [OSVALH1]  
“Depois, as pequenas não responderam”

Por outro lado, acontecendo inversão do sujeito na ordem da frase, isto é, construções do tipo VS,<sup>91</sup> ocorre quase sempre falta de concordância entre estes dois elementos (exemplo [198]):

[198] PA:  
*Então táqui os home que veio preso* [MANOH3]  
“Então estavam aqui os homens que vieram presos”

Deste modo, temos dois tipos de concordância variável para o sintagma verbal, o que não acontece com a concordância variável interna ao SDET, que é idêntica nas construções em ambas as realizações SV e VS. Ainda assim, a marca de PL pode incidir sobre outros itens do SN, que não o elemento pré-nuclear. Não obstante, como a marcação é quase categórica no item pré-nuclear em todas as gerações, é consensual que se tenha esta como o modelo para efeitos de análise, isto é, se admita como paradigmática a concordância interna ao SDET apontada nas ocorrências do tipo da registada no exemplo [191]. De facto, o que sucede com as marcações que divergem desta é que, por norma, denotam gramáticas em competição no mesmo indivíduo, fruto de causas várias, que podem ser de carácter linguístico e/ou social. De facto, se olharmos para os exemplos que se listam seguidamente, constatamos que a variação é toda realizada pelo mesmo falante, e pode ir desde a marcação PL que incide apenas no

determinante (exemplo [199a]) até à marcação que afecta ambos os itens (exemplo [199c]), passando ainda pela marcação registada somente no nome (exemplo [199b]). Assim, nunca é de descartar a possibilidade de estas gramáticas em competição poderem dificultar uma análise em termos de se estabelecer uma visão configuracional uniforme da concordância interna ao SDET, tanto no PB como no PA:

- [199] PA:
- a. *Os dirigente ão são homens assim* [CASTEH1]  
“os dirigentes não são homens assim”
  - b. *Eis a razões que eu já ão stô cá.* [CASTEH1]  
“eis as razões pelas quais eu já não estou cá”
  - c. *As crianças já estão adormecido* [CASTEH1]  
“As crianças já estavam a dormir”

Como se constata, o que se passa internamente a nível do SN não é o mesmo que se passa no âmbito da frase, já que não é possível efectuar-se uma análise uniforme de todos os fenómenos de concordância na segunda. Nesta conformidade, e ao contrário do que sucede com o PE, não existe paralelo entre concordância verbal e concordância interna ao SDET. Ainda assim, o alto peso de pluralização evidenciado pelos determinantes pré-nucleares permite considerar que a posição DET é marcada com a morfologia de número, embora as ocorrências com possessivos do tipo das apontadas por Scherre (1988:163) (ausência de marcação na primeira posição e inserção de marca PL no possessivo em segunda posição) parecem retirar alguma robustez a este pressuposto no PB. Contudo, como veremos no ponto 4.1.2.6.1, esta particularidade não encontra paralelo no PA, já que, ao contrário do que sucede no Brasil, a expletividade do definido não é tão sistemática na fala almojarifana como no PB.

Relativamente à concordância entre SN e S<sub>ADJ</sub>, se tomarmos como exemplo uma oração pequena (OP) (exemplo [200]) e levarmos em conta que a CONC emerge entre SUJ e V até ao momento do *spell-out*, teremos a configuração sintáctica da Fig. 23 para representar os predicativos adjectivais e respectivo movimento sintáctico para CONC, proposta por Chomsky (1996 [1995]:175) para este tipo de orações:

[200] PE:  
*O João é alto.*

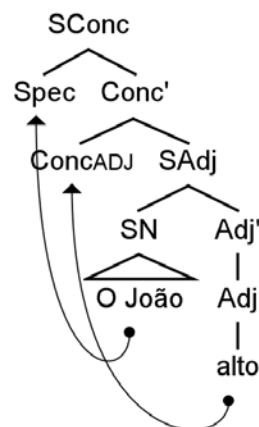


Fig. 23. Sintaxe na OP para concordância  
 (a partir de Chomsky, 1995:175).

Na configuração proposta, CONCADJ diz respeito à mnemónica de um conjunto de *traços-phi* relacionados com o adjectivo (Silva, 2010:26). O SN sobe para SPEC, enquanto o ADJ ascende a CONCADJ, gerando a estrutura interna entre SN-ADJ nas construções predicativas, ou seja, a configuração da OP. Na estrutura S-, o complemento de SV [SN *o João*] sobe para receber caso nominativo e CONC verbal, originando a estrutura *O João é alto*. De acordo com a estrutura minimalista da concordância (Fig. 18), o SN *o João* estabelece três tipos de relações: (i) Uma relação de Caso com [T CONC-S], daí resultando a relação verbal [[T CONC-S] V]; (ii) Uma relação de concordância com CONC-S, originando a complexidade verbal de *ser*; (iii) Uma relação de concordância com CONC da Fig. 23, estabelecendo a complexidade adjectival. Note-se, então, dois aspectos neste tipo de construções: (i) a OP impõe como requisito que o SN esteja fora do SAdj; (ii) a sintaxe na OP para a concordância estabelece paralelismo com o movimento de N para NUM (Fig. 22).

Outro aspecto a ter em consideração na CPL-var do PB e do PA é a assimetria entre posições pré e pós-nucleares. Para compreenderem esta, os teóricos têm recorrido também à MD, que advoga que o *output* da sintaxe é o *input* da morfologia (esta lê aquela). Deste modo, cada morfema contém em si a informação sobre os contextos em que pode ser inserido. Embick & Noyer (2001) desenvolveram uma análise sobre a marca de definitude no dinamarquês e no sueco, aplicando as assunções da MD, conjuntamente com as operações sintácticas que envolvem não só os movimentos do núcleo de N para DET mas também aspectos relacionados com o caso. Como se sabe, em

línguas como o latim, a classe de palavras que recebe o caso é determinada pelo nome, espalhando-se a morfologia por todos os contextos:

[201] LATIM:  
*Video illam cicadam pulchram.*

Nestas situações, estamos perante morfemas dissociados do mesmo tipo dos que ocorrem no PE, aos quais se opõem outros, que não se espalham pela cadeia sintagmática (os morfemas *singleton*), já que estes últimos aparecem apenas numa única posição, como acontece quer com os morfemas do pretérito perfeito quer com os que se fixam em determinada posição do SN, originando a variação (Costa & Silva, 2006b:44). Este aspecto permite prever o seguinte: se há variação morfológica, então também é possível existir especificação na variação entre morfemas dissociados e morfemas *singleton*. De facto, reside nestes a primeira fonte de variação que distingue o PB e o PA do PE.

Outra fonte de variação radica, ao que tudo indica, na visibilidade da relação entre especificador e núcleo. Alguns idiomas, como o francês, apresentam concordância entre o pronome objecto e o particípio passado (exemplo [202]), o que não acontece no PE (exemplos [203a] e [203b]):

[202] FRANCÊS:  
*Je les avais ouvertes.*

[203] PE:  
a. *Eu tinha-as aberto.*  
b. *\*Eu tinha-as abertas.*

Ora, este tipo de contraste permite dizer que, quando duas categorias estão na posição especificador/núcleo, a relação paramétrica deixa perceber se a concordância é morfológica visível. No PE e PB/PA, a concordância é visível nas construções com sujeito pré-verbal, (exemplos [194a] e [196]), o que não acontece nas realizações de Almojarife (exemplo [198]) e do Brasil com sujeito pós-verbal, já que, nestas línguas, o sujeito não desencadeia a concordância.

No que concerne à posição da estrutura sintáctica na qual é inserido o morfema de número, é necessário ter em conta que este tipo de marcação é funcional, enquanto a do género é lexical (Müller, 1994a; White *et alii*, 2004; Corrêa, Name & Ferrari-Netto,

2004). Como o número constrói referência, deve aparecer na posição de DET, seja no PE seja no PA ou no PB. Contudo, a concordância do PE é caracterizada por morfemas dissociados, enquanto no PA e PB pode ser determinada por morfemas do tipo *singleton*. Já no que concerne à concordância verbal, quando o sujeito é pré-verbal faz-se apelo à relação SPEC/verbo para justificar a inserção de marcas neste. Logo, é possível predizer que só se espera concordância quando há relação SPEC/núcleo, mas apenas algumas línguas optam pela visibilidade morfológica, como acontece com o árabe (ponto 2.6.2.1 do presente trabalho). Por conseguinte, a ocorrência da concordância no PA e no PB é estabelecida de acordo com as configurações das Figs. 23 (concordância entre sujeito e verbo) e 24 (concordância interna ao SN, exemplo [199a]):

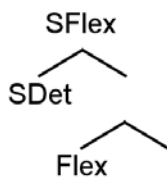


Fig. 24. Sintaxe da concordância sujeito/verbo (PA e PB).

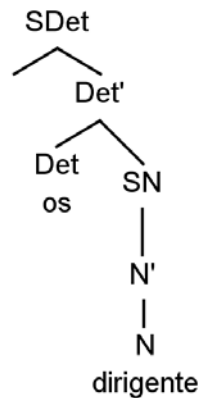


Fig. 25. Sintaxe da concordância interna ao SN (PA e PB).

Relativamente aos possessivos, que no PB evidenciam tendência para receberem a marcação PL quando estão em posição pré-nuclear, mas em segundo lugar da cadeia sintagmática (Scherre, 1988:163), a situação é justificada em termos da relação de definitude que estes itens gramaticais estabelecem com o nome. De facto, ocorre correlação entre posição e definitude, já que o estatuto dos possessivos é diferente quando se encontra em posição pré ou pós nominal. Assim, caso esteja colocado antes do nome (exemplo [204a]), o possessivo é forte portador de definitude e assume a função de núcleo, o que não sucede quando se encontra depois do substantivo (exemplo [204b]), visto deter o estatuto de sintagma nesta posição (Castro & Costa, 2003:108):

- [204] PE:
- a. *O meu livro.*  
↳ núcleo
  - b. *Um livro meu.*  
↳ sintagma

Paralelamente, existem fortes restrições de adjacência entre o definido e o possessivo, já que não admitem a interferência de nenhum elemento entre eles (exemplo [205a]), o que não acontece quando o segundo se encontra em posição pós-nominal (exemplo [205b]):

- [205] PE:
- a. *\*O inteiramente meu livro.*
  - b. *Um livro inteiramente meu.*

Não obstante, essa restrição desaparece no PA e em determinadas regiões do Brasil, visto o definido assumir uma função expletiva, isto é, ser dispensado em grande parte das realizações (exemplo [206]), sem que se perca a noção de definitude:

- [206] PA:  
*arramê mias bagagem, pegué vim* [CELESH3]  
“arrumei a minha bagagem, peguei nela e vim”

Como o número é inserido na categoria que permite a ancoragem com a posição de referência, verificamos que esta é garantida pelo possessivo, que recebe, assim, o morfema de PL. Estabelece-se, então, a configuração sintáctica seguinte:

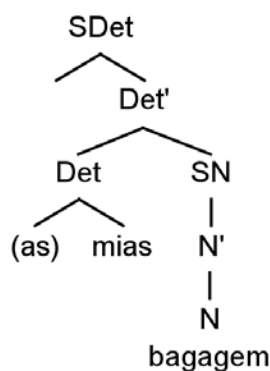


Fig. 26. Sintaxe do possessivo núcleo para a concordância.

Finalmente, a correlação entre posição e definitude é também advogável para os adjectivos que ocorrem em posição pré ou pós-nominal, justificando a assimetria

revelada pelos mesmos na concordância PL. Portanto, se estiverem posicionados antes do nome (exemplo [208a]), detêm a categoria de núcleo e revelam tendência para receberem a marca de número. Caso se encontrem localizados depois do substantivo (exemplo [208b]), possuem categoria de sintagma e não recebem marca de concordância:

[208] PA:

- a. *nós temo que utilizá arma e com boas barbatana* [OSVALH1]  
↳ núcleo  
“nós temos que utilizar armas e boas barbatanas”
- b. *eu faço trabalhos privado*. [CASTEH1]  
↳ sintagma  
“eu faço trabalhos privados”

Estas situações encontram paralelo em determinadas línguas, como o inglês, na qual a posição dos adjetivos é determinada pelo uso (ou não) de complemento. Assim, os adjetivos detêm categoria nuclear se ocorrerem antes do nome, em virtude de não possuírem complemento (exemplo [209a]), ou, em contrapartida, são categorias lexicais, em virtude de possuírem um complemento e se posicionarem pós-nominalmente (exemplo [209b]):

[209] INGLÊS:

- a. *a proud man*  
↳ núcleo
- b. *a man proud of his son*  
↳ sintagma
- c. *\*a proud of his son man*

Tendo em conta que os itens lexicais pertencentes às categorias DET, ADJ e N detêm radicais que são “partes de um conjunto” desses mesmos itens, as quais são seleccionadas a fim de serem manipuladas pelo dispositivo computacional-sintáctico, tais radicais vão determinar a elaboração de derivações sintácticas, já que são providas de distintos traços disponibilizados pela GU para as diferentes línguas (Chomsky, 1996 [1995]). Por outro lado, as relações de concordância entre sujeitos e predicativos do PE levam os primeiros a comportarem-se de acordo com o esperado pela Teoria da Verificação de Traços (Chomsky, 1999a). Contudo, a comparação de tais traços, partindo das estruturas da L1, não será a mais apropriada para a descrição de uma L2 em

qualquer dos seus estágios de evolução, já que algumas das línguas analisadas neste estudo (p.e o português e o iorubá ou o santomense) exibem diferenças radicais nos traços de alguns dos seus itens, que vão determinar variação na interlíngua.

## **2.10. Fossilização**

Partindo do pressuposto que a ASL assenta na gramática da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), então o processo é caracterizado por uma tentativa de descoberta das regras (Chomsky, 1969) que, inicialmente, são confrontadas com as da L1. Este aspecto implica que o aprendiz interiorize mais rapidamente as particularidades linguísticas da L2 tidas como mais simples, em detrimento das consideradas complexas. Como o indivíduo em situação de ASL não entende plenamente (ou não tem tempo de assimilar) o sistema de flexões da LA, a sua interlíngua vai registar ausência ou redução destas, situação que se pode traduzir em cristalizações que acabam configurando variação sistemática. Deste modo, enquanto a criança em fase precoce de aquisição inata da L1 produz variação linguística, elaborando erros naturais no processo de desenvolvimento da linguagem, mas que, geralmente, ultrapassa de forma fácil, no caso da ASL, a variação linguística pode manter-se durante bastante tempo e, caso se torne permanente e estável, acaba por configurar uma fossilização. Assim sendo, a fossilização é um fenómeno exclusivamente inerente à ASL e aquisição de LE's, já que não ocorre na aprendizagem da L1 (Selinker, 1972:211).

### **2.10.1. Fossilização e variação**

Os estudos sobre fossilização que buscam revelar as causas que subjazem à mesma têm, quanto a nós, versado muito sobre descrições do fenómeno, mas adiantado pouco acerca dos diferentes tipos e causas que originam as variações sistemáticas ou permanentes. A este propósito, Selinker (1996:226) chama a atenção para o facto de, em diversos trabalhos, fenómenos distintos, como a recaída (*backsliding*),<sup>92</sup> a desaceleração no processo da aprendizagem da L2 ou a baixa proficiência, serem conotados à fossilização, já que não são objecto de uma observação em diacronia. Para Long (2003:490), a estabilização de uma determinada variação, enquanto fenómeno que



encerra todas as características da fossilização, excepto a imutabilidade, representa o primeiro passo no sentido de se prever que a fossilização está em vias de ocorrer. Neste sentido, a estabilização tem particular interesse no estudo da ASL, já que permite entender não só a natureza e causas da fossilização, ela mesma, mas também conhecer quais são as formas particulares da interlíngua que são potencialmente fossilizáveis. E no seguimento deste raciocínio, é ainda possível prever se determinada variação sistemática pode ou não vir a configurar mudança. Deste modo, tanto a variação prolongada como a fossilização que se manifestam na interlíngua impedem que o indivíduo alcance uma competência gramatical que se aproxime da dos falantes da LA, mesmo em contexto de aquisição formal que envolva um estudo intensivo e um processo de memorização bastante consciente. Além de afectar a produção fonológica, em primeira instância, a variação na L2, sobretudo em período precoce da sua aquisição, vai também manifestar-se acentuadamente a nível morfossintáctico, em segunda instância. Em situações de aquisição massiva, se o *input* da LA for prolongadamente esporádico, as motivações empobrecidas, isto é, a instanciação em cadeia de novos errados/parâmetros (Gonçalves, 2004:235), acabam por ter reflexo tanto na variação registada na organização frásica dos constituintes como na produção da flexão e concordância da L2, levando à cristalização diacrónica de formas variáveis.

Para Corder (1973:194), a ASL pressupõe três estágios de aprendizagem: o pré-sistemático, o sistemático e o pós-sistemático. No ponto 5.3.2 do nosso trabalho (variável extralinguística *idade*), é possível verificar que a geração mais idosa de Almojarife apresenta uma *performance* que permite inseri-la no estágio pré-sistemático, já que os falantes apenas revelam uma vaga consciência acerca da ordem sistemática dos itens da LA, não sendo capazes de corrigir erros. Por outro lado, poucos membros da geração menos idosa da comunidade terão alcançado o estágio pós-sistemático da aquisição, uma vez que são raros os indícios de se estar perante discursos consistentes, em que os falantes sejam capazes de corrigir os seus próprios erros autonomamente. Deste modo, o quadro geral da fala de Almojarife configurará o estágio intermédio da aquisição, o sistemático, correspondente igualmente à fase 3 de ASL proposta pela Teoria da “Processabilidade” (Plag, 2008a – Tabela 3.1), dado que a grande maioria dos seus utilizadores, apesar de revelar já alguma interiorização das regras da língua,

continua a produzir variação sistemática, nomeadamente a nível da concordância no SN. Este aspecto indica que os falantes não corrigirão a variação se não captarem ou obtiverem um *feedback* negativo (cf. Corder, 1973:196), o que leva a considerar a possibilidade de a variação registada actualmente no SN do PA se poder vir a fossilizar.

Schumann (1978:34), em estudos sobre pidginização, foi o primeiro a chamar a atenção para a possibilidade de os factores sociais conduzirem a fossilizações. Para o investigador, as situações de fossilização estão intrinsecamente ligadas ao modo como o aprendente se acultura ao grupo detentor da LA (*The Acculturation Model*). Defendendo que os primeiros estados de ASL são comparáveis aos processo de formação das línguas pidgin, o autor entende que, no caso de haver uma grande distância psicocultural entre o aprendente e o grupo falante da LA, a pidginização precoce da L2 torna-se persistente e fossiliza. No quadro de transmissão linguística irregular experienciado pela comunidade de Almojarife, o distanciamento em relação à LA foi bastante marcante em determinado momento da emersão e nivelamento da L2, fazendo com esta preenchesse apenas requisitos de comunicação interétnica. Assim sendo, o contexto criado mostrou-se propício para a fixação de formas transferidas das L1's, candidatando-as à fossilização. Ainda assim, é preciso ter em conta que a proposta de Schuman (1978) não leva em conta dois aspectos fundamentais: (i) os factores que determinam a aculturação do indivíduo não são estáticos, mas variáveis e dinâmicos, consoante as experiências que o aprendente vai atravessando; (ii) os aprendentes não estão apenas submetidos às condições sociais, já que se podem tornar sujeitos destas e construir o próprio contexto de aprendizagem.

Para Hale (1988:32), a fossilização resulta do facto de o falante não conseguir erradicar determinados parâmetros da L1, que acabam intervindo na ASL.<sup>93</sup> De certa maneira, o pressuposto seria posteriormente retomado por Schwartz & Sprouse (1996), quando estes determinaram que a GU actua na L2 através da L1. Como tal, a fossilização resulta da dificuldade que o aprendente tem em refixar os parâmetros na L2, conforme postulado pela FFFH. Por seu lado, Selinker & Lakshmanan (1993:215) entendem que a fossilização ocorre quando determinados factores da aquisição de uma L2 actuam de forma múltipla (*Multiple Effects Principle* – MEP), ou seja, funcionam em continuidade, levando à estabilização de certas formas. Entre os ditos factores podem

apontar-se a simetria de estruturas, a afectividade, o princípio da distância mínima, a interferência e, sobretudo a transferência, tida como o factor central da ASL para possibilitar a fossilização. A conjugação destes factores levaram a que acontecesse a transmissão em cadeia de novos/errados parâmetros (cf. Gonçalves, 2004:235) durante a aquisição do PAL1, pelo que terá de ser considerada a hipótese de ter ocorrido reinstanciação paramétrica neste, possibilitando a fossilização de determinadas formas do substrato na fala dos almoxarifanos.

Ainda segundo Selinker & Lakshmanan (1993:215), a fossilização apresenta duas formas: (i) a fraca, que determina que a transferência de uma língua é um co-factor privilegiado na ocorrência de efeitos múltiplos; (ii) a forte, que postula que a transferência de uma língua é um co-factor necessário na ocorrência de efeitos múltiplos. Apesar de estes dois tipos de factores serem fundamentais para ocorrer a fossilização, os autores admitem ainda que esta pode igualmente acontecer sem transferência, conforme o haviam demonstrado Bean & Gergen (1990:219) em estudos sobre variação registada na *performance* individual de interlínguas com fossilizações.

### **2.10.2. Potenciais causas da fossilização**

Várias têm sido as causas apontadas como potenciais fontes da fossilização, de entre elas podendo destacar-se as que resultam de fenómenos de esquiva (*avoidance*)<sup>94</sup> intencional por parte do falante (Nakuma, 1988), a ultrageneralização de determinadas regras Ellis (1994)<sup>95</sup> ou da falha na interacção entre o conhecimento interior e os estímulos externos para determinar o desenvolvimento das gramáticas mentais (Franceschina, 2005).<sup>96</sup> No caso da aquisição do PtgL2, são comuns, por exemplo, os erros de ultrageneralização, e consequentes fossilizações, que actuam sobre o sistema verbal, com os falantes regulando o paradigma de conjugação dos verbos da segunda (terminados em *-er* no infinitivo) e terceira classes (terminados em *-ir* no infinitivo) de acordo com o paradigma dos verbos de primeira classe (terminados em *-ar* no infinitivo) ou aplicando, continuamente, a forma da terceira pessoa singular a todas as outras:

[209] PA:  
*nós é pobre miserável aqui Praia Moxarife* [FRANCM2]

Neste caso, estamos perante transferências intralinguísticas, que fazem com que o falante ignore sistematicamente as regras das restrições, incluindo as que incidem sobre itens lexicais ou de qualquer outro tipo linguístico, fossilizando-as em diacronia (Ellis, 1994:284). Segundo Brown (1994:367), os falantes que incorrem na sobregeneralização acreditam que estão a usar a gramática correcta, em virtude de terem interiorizado regras sistemáticas. Deste modo, o fenómeno da sobregeneralização ocorre de forma inconsciente, necessitando quase sempre de *feedback* negativo para não ser fossilizado.

Tendo ainda em conta as distintas teorias sobre fossilização, poder-se-á dizer que, genericamente, o conceito aponta para as seguintes determinações: (i) a fossilização pode aparecer em determinados níveis da interlíngua; (ii) a fossilização pode ocorrer não só em diferentes estádios de aprendizagem da LA mas também em variados estratos etários; (iii) a fossilização pode acontecer a nível estrutural ou a nível da competência; (iv) a fossilização manifesta-se, geralmente, através de formas que se desviam das normas da LA; (v) há graus suaves e fortes de fossilização. Do exposto, conclui-se que a fossilização representa um estádio inevitável na ASL e com influências nesta.

O conceito de transferência tem sido observado como uma questão central quer em estudos de ASL (Chomsky, 1969) quer em estudos sobre fossilização (Ellis, 1994), tendo o primeiro autor levantado a possibilidade de ser um processo consciente, através do qual o aprendente constrói a interlíngua, testando hipóteses sobre a L2. Nesta perspectiva, e tendo Selinker (1972:229) advogado que algumas das regras registadas na interlíngua resultam de transferências da L1, então esta constitui a base para a aprendizagem da nova língua (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006). Desta forma, os erros registados na L2 têm origem nas diferenças entre gramáticas da LA e L1, sendo as transferências as causas da fossilização.<sup>97</sup> No entanto, e de acordo com a HI (Plag, 2008a, 2008b, 2008c, 2008d), é importante estabelecer a diferença entre traços conectados a questões universais do desenvolvimento da interlíngua e os que configuram transferências, já que algumas formas registadas na interlíngua poderão não passar de falsas transferências (Tabela 6.1). De facto, para estas acontecerem é necessário que o processador da interlíngua tenha disponível um dispositivo que accione o processamento da estrutura da L1 passível de transferência, isto é, que a estrutura da L1 a transferir encontre condições para ser

processada no sistema de desenvolvimento da L2<sup>98</sup> (Andersen, 1983c:182; Plag, 2008b:314).

Alguns estudos sobre ASL apontam a intervenção de duas outras causas para ocorrer fossilização de formas incorrectas na interlíngua: a falta de instrução formal na LA e, conseqüentemente, de *feedback* negativo sobre a variação produzida (Vigil & Oller, 1976:287; Corder, 1981b:72); a idade em que se dá a ASL (Brown, 1980; Larsen-Freeman & Long, 1991; Selinker & Lakshamanan, 1993; Lightbown & Spada, (2003 [1993])). Sobre o primeiro aspecto, pronunciar-nos-emos no ponto 5.3.1, dedicado à análise da variável social *escolaridade*. Quanto à questão da idade, Selinker & Lakshamanan (1993:207) estabelecem uma associação entre fossilização e escalão etário adulto, reservando a noção de desenvolvimento para a idade infantil. Por seu lado, Lightbown & Spada (2003 [1993]:70) entendem que a aprendizagem de uma L2 por adultos se revela mais eficaz nos estágios iniciais de aquisição e que estes podem ser capazes de comunicar com sucesso, apesar de apresentarem diferenças a nível de sotaque e produzirem formas que, por vezes, não se adequam ao discurso. Estes aspectos levaram ainda os autores a sugerir que, depois do período crítico, a aquisição não será inata, demarcando-se do modo como a mesma acontece na infância, seja em L1 seja em L2.

### **2.10.3. GU e fossilização**

Para White (1989, 1990, 1993, 2000, 2003), defensora de que a GU actua de forma igual na aquisição de L1 e L2, o falante em estado inicial de ASL já possui a gramática da L1, com princípios e parâmetros estabelecidos. Como o aprendente já conhece um sistema linguístico, vai ocorrer transferência do mesmo para aquele que se encontra em estágio inicial de aquisição, já que a gramática a que o falante recorre para testar as hipóteses é a da L1, que vai actuar como um filtro (Krashen, 1982, 1985, 1988), impedindo que certas formas do *input* da LA sejam captadas e apreendidas. Desta forma, os aprendentes são impedidos de atingir, na gramática da L2, a mesma proficiência que têm na gramática da L1, pelo que, a não percepção de estruturas relevantes da LA conduz à fossilização. Por seu lado, a fixação de determinados parâmetros, em

detrimento de outros, deve-se à variação inter e intrapessoal, o que caracteriza a interlíngua do aprendente (Cook, 1993:186).

Para Flynn (1989:104), a ASL é facilitada quando ocorre um ajustamento entre os parâmetros da LA e da L1. Para que tal aconteça, os aprendentes em ASL necessitam de ser expostos a *input* prolongado e correcto, caso contrário os parâmetros da L2 não se ajustam aos já estabelecidos na L1, instanciando-se novos/errados parâmetros na interlíngua (Gonçalves, 2004:235). Assim sendo, os falantes em processo de aquisição precisam de estabelecer novos valores para os parâmetros a serem adquiridos, o que compromete a apreensão e, no caso de esta se retardar, induz em fossilização. Contudo, para se determinar se uma forma linguística configura fossilização, Selinker (1972:223) propõe que ela terá que permanecer estabilizada pelo menos durante cinco anos. Por seu lado, Han (2004:14) questiona o período, uma vez que, primeiro, seria necessário precisar o tempo que um determinado indivíduo necessita para adquirir um factor particular na L2. De facto, segundo a autora, múltiplos são os factores que devem ser considerados como motivadores da fossilização, estando alguns deles ainda por determinar nos estudos efectuados até à data.

De tudo o que foi exposto acerca das possíveis causas motivadoras da fossilização, conclui-se que alguns factores se mostram determinantes para que certas formas linguísticas fossilizem, como é o caso do meio ambiente no qual o falante em aquisição se encontra inserido ou as questões relacionadas com aspectos cognitivos, neurobiológicos e sócio-afectivos. E olhando individualmente para estes, enquanto elementos de um todo, salta à evidência que apenas dois tipos de factores concorrem para que se dê a fossilização: os factores internos e os factores externos. Estes aspectos permitem então elaborar a seguinte tabela taxionómica, para, de forma sumariada, não só se simplificar toda a explanação acerca das fossilizações mas servir também como proposta metodológica para o nosso trabalho, quando tivermos de abordar questões relacionadas com a problemática da fossilização:

**Tabela 4.1.** Taxionomia dos factores motivadores da fossilização (Xueping, 2008:131).

Externo	Meio ambiente		p.e.: ausência de instrução; défice de <i>input</i>
Interno	Cognitivo	Representação do conhecimento	p.e.: influência da L1; défice de acesso à GU
		Processamento do conhecimento	p.e.: défice de atenção
		Psicológico	p.e.: esquiva; simplificação
	Neuro-biológico		p.e.: idade; défice de talento
	Sócio-afectivo		p.e.: défice de aculturação

Outro aspecto que consideraremos tem a ver com o facto de o conceito de estabilização estar muito próximo do de fossilização, mas ambos os fenómenos configurarem realidades distintas. A este propósito, Han (1998:13) e Selinker & Han (2001:278) esclarecem que a estabilização e a fossilização configuram um *continuum*, no qual a primeira é parte integrante do processo da segunda e, por tal, não podem ser equiparadas. Segundo a perspectiva de Long (2003:521), a estabilização é o primeiro sinal de que poderá ocorrer uma fossilização. Contudo, os fenómenos distinguem-se em virtude de a primeira ser permeável e reversível, enquanto a segunda é hermética e permanente. Portanto, a forma estabilizada pode ser corrigida, mas a fossilização é imutável.

#### **2.10.4. Fossilização no PA**

##### **2.10.4.1. PA: transferência e fossilização fonológica**

A interlíngua dos falantes em ASL pode registar dois tipos de fossilização: (i) fossilização individual, determinada pela persistência na elaboração de certa forma durante o desenvolvimento da interlíngua do indivíduo; (ii) fossilização de grupo, motivada pelo nível de uso de uma forma durante o desenvolvimento diacrónico da língua de determinada comunidade. As segundas têm como ponto de partida alguns aspectos relacionados com as primeiras, uma vez que são originadas pela competência linguística fossilizada, primeiramente, ao nível de desenvolvimento das competências fonológica, gramatical, lexical e pragmática de indivíduos aprendentes de uma L2. Como tal, revelam-se na interlíngua de falantes de L2's que têm um contacto, de certa

forma prolongado, com a LA, atingindo um determinado nível de *performance* sem, contudo, deixarem de exhibir, repetidamente, certas variações, difundindo-as, depois, pelo grupo em que estão inseridos (Selinker, 1979:212).

Não obstante, note-se que as fossilizações poderão ser temporárias ou permanentes (Selinker, 1979:218) e ocorrerem a todos os níveis, desde o fonológico ao pragmático, sem deixar de passar pelos morfológico, sintáctico e semântico. As fossilizações temporárias, também designadas por estabilizações, revelam que a interlíngua fossilizada consiste de patamares de aprendizagem (Fig. 27), nos quais o desenvolvimento de determinados traços da LA é fixado ou inibido por períodos de tempo mais ou menos longos (Sims, 1989:70).<sup>99</sup> Como tal, as fossilizações deste tipo podem ser desfossilizadas, o que não acontece com as fossilizações permanentes, que resultam de uma conjugação de fenómenos conotados a aspectos sociais, psicológicos e interactivos e impedem que o estágio final de determinada forma da interlíngua alcance o mesmo estágio patente na LA:

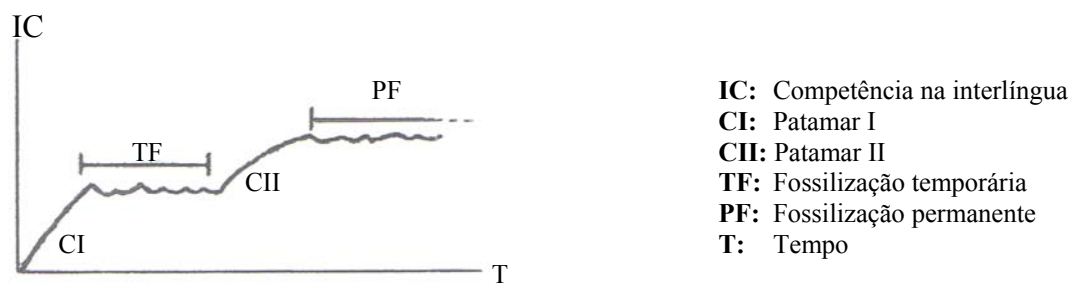


Fig. 27. Fossilização temporária e fossilização permanente (Xueping, 2008:131).

No presente trabalho abster-nos-emos de abordar todo o género de fossilizações, já que as variáveis linguísticas a que recorremos para estudo do fenómeno da CPL-var do PA são do tipo marcadamente fonético-fonológico e morfossintáctico. Assim, incidiremos a nossa atenção sobre as fossilizações directamente relacionadas com estas variáveis.

O “fenómeno de Joseph Conrad”, ou seja, a fossilização fonológica, é tido como inevitável na ASL após a puberdade (Scovel, 1969:252).<sup>100</sup> As diferenças fonológicas entre interlíngua e LA resultam da aquisição incorrecta da pronúncia de algumas formas desta, que são afectadas pelo sistema fonético da L1, constituindo, possivelmente, a maior diferença entre ambas. Como determinadas línguas do grupo níger-congo



atlântico, caso do ramo bantu, não possuem a consoante /r/, é difícil aos falantes adultos destas articularem-na correctamente quando em processo de aquisição do PtgL2. Observando os exemplos [210] e [211], em especial o nome “doutor”, constatamos que a ausência da consoante /r/ nas línguas bantu provoca uma dificuldade de articulação nos falantes adultos em aquisição do PtgL2, levando à realização sistemática de fenómenos de apócope e lambdacismo na interlíngua (o santomense), que se conservaram diacronicamente, configurando fossilizações no PA, possivelmente temporárias (exemplo [212]):

[210] FORRO:  
*dotolo da e zesõ*  
(Ferraz, 1979:66)  
“o doutor deu-lhe injeções”

[211] PA:  
*dotôlo deu lemede* [PRINCEM3]  
“o doutor deu um remédio”

[212] PA:  
*tenho bastante filho, outra filha professora, outra veio **doutora***  
[MAURIH2]

Para Tarone (1978, 1995), à fossilização fonológica subjazem causas fisiológicas e psicológicas. Deste modo, as dificuldades articulatórias poderão assentar no facto de as línguas bantu e o português recorrerem a distintos pontos articulatórios do tracto bucal para produzirem os seus sons, levando os falantes do PtgL2 a enfrentarem dificuldades na articulação de sons que não existem nas suas L1's (cf. Ellis & Beaton, 1995:161). Por seu lado, as justificações em termos psicológicos apontam para o fenómeno da formação de hábitos, que leva a percepção e a produção da fala a serem permanentemente influenciadas pelo sistema fonológico da L1, a ponto de o aprendente enfrentar sérias dificuldades em perceber ou elaborar uma nova fonologia (Tarone, 1978, 1995). Este aspecto pode gerar uma falta de empatia entre o aprendente e os falantes da LA, fazendo com que os primeiros só adoptem alguns padrões de pronúncia quando se sentem particularmente identificados com determinado grupo (Labov, 1972a:31). Debruçando-se ainda sobre a estrutura silábica de certas línguas (incluindo o português), Tarone (1978, 1995) defende também que o fenómeno da fossilização pode assentar em três

factores: a transferência, os processos relativos à aquisição da L1 e os fenómenos de carácter universal. A hipótese da transferência fonológica postula que o aprendente usa a estrutura silábica da L1 para tentar comunicar na L2, um aspecto que é facilmente constatado no santomense e que revela resquícios no PA, conforme analisado em detalhe no ponto 5.2.3.10.1 do presente trabalho. Assim, se a L1 do falante contém sílabas de acordo com a estrutura CV, a tendência do aprendente é transformar a estrutura silábica da LA em estruturas do mesmo tipo das que ocorrem na L1, até porque os falantes dão primazia às estruturas do tipo CV em virtude da sua força lexical (Clements, 2009:22). Neste caso, e como referimos no ponto 2.4.3, a saliência perceptual impõe-se à frequência na LA, em virtude de a estrutura CV ser mais facilmente processável (Thomason & Kaufman, 1988:331).

A hipótese sobre os processos relativos à aquisição da L1 entende que o aprendente da L2 tende a fazer aquilo que o aprendente da L1 faz com a estrutura silábica. Deste modo, os defensores desta linha, assentando pressupostos nos princípios advogados pela perspectiva psicolinguística de interlíngua (Selinker, 1972) de que a GU deixou de estar disponível para actuar na aquisição da L2 (Clahsen & Muysken, 1986; Meisel, 1997; Muysken, 2001), postulam que os processos que actuam na interlíngua são independentes daqueles que incidem sobre a L1 e o que ambos os falantes fazem é apenas simplificarem as sílabas difíceis (Oller, 1974:114). Porém, Tarone (1978, 1995) discorda desta hipótese, defendendo que os falantes em ASL usam estratégias para pronunciarem os sons difíceis. Assim, no caso de ocorrerem encontros consonantais, em vez de proceder a uma simplificação, o falante da L2 insere vogais na estrutura consonantal (epênteses ou paragoges), a fim de retomar a estrutura universal CV (Tarone, 1978:328). Portanto, estes fenómenos de inserção configuram estratégias próprias da fonologia dos aprendentes em ASL, levando a que a estrutura silábica da interlíngua seja diferente daquela que caracteriza a LA. Casos de inserção epentética ou paragógica são comuns no forro e no PA, como se pode também constatar no ponto 5.2.3.10.1 do nosso trabalho.

As fossilizações, em particular as fonológicas, têm, então, sido explicadas em termos de factores diversos: idade, resistência à aculturação, pressão comunicativa, ausência de aprendizagem formal, natureza do *input* (negativo ou positivo), perda de capacidade

discriminatória a nível dos padrões trocaicos e jâmbicos em idade pós-adolescência, falta de empatia em relação aos falantes da LA, formação psicológica de hábito, qualidade de voz para expressar emoções,<sup>101</sup> etc.. Contudo, e ao que tudo parece indicar, residirá na transferência a responsabilidade maior pelas fossilizações.

A fossilização fonológica é definitiva na maioria dos casos, permanecendo em indivíduos que adquirem um grande domínio de vocabulário e sintaxe da LA (Scovel, 1969:252). Se a distância entre os sistemas fonológicos da L1 e da LA for acentuada, maior será a tendência para os aprendentes da L2 construírem os sons que não existem na L1, por paronímia fónica com aqueles que fazem parte desta (exemplos [160], [162] e [163]). A incapacidade dos falantes da L2 em pronunciarem correctamente os sons da LA conduz, assim, à fossilização fonológica (Levelt, 1989:286). Por outro lado, como o conhecimento da L1 é o ponto de partida para a aquisição da L2 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), a produção fonológica desta vai seguir o mesmo processo de articulação fónica da L1. Deste modo, só dificilmente se poderá desassociar a fossilização fonológica do fenómeno da transferência, sendo a sua cristalização determinada pelo facto de os falantes não serem capazes de refixar, na L2, os parâmetros desactivados das matrizes fonéticas da LA (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clahsen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) nem conseguirem erradicar os novos/errados parâmetros transferidos das L1's (Hale, 1988:32; Gonçalves, 2004:235). Neste caso, os substratos do grupo níger-congo atlântico são a fonte de determinadas formas que transitaram para a interlíngua (santomense) e que acabaram fossilizadas no PA, devido à reinstanciação em cadeia de novos/errados parâmetros (cf. Gonçalves, 2004:235).

#### **2.10.4.2. PA: transferência e fossilização morfossintáctica**

As fossilizações morfológicas ocorrem em maior grau quando a LA possui uma gramática elaborada em morfemas, por oposição a uma L1 cuja gramática detém um reduzido sistema flexional. Isto é exactamente o que sucede quando comparamos as gramáticas do PE, por um lado, e do santomense e seus substratos do grupo níger-congo atlântico, por outro lado. O sistema de artigos portugueses e a flexão sufixal em género e número não encontram correspondência nestes dialectos. Como tal, podem não ser

fonologicamente interpretados pelos falantes em processo de aquisição do PtgL2, originando a sua ausência na interlíngua (Baxter, 2004:120; Kihm, 2003:347), como acontece a nível da estrutura do SN, que não exhibe determinantes no santomense e apresenta variação dos mesmos no PA, sendo esta originada pela conjugação de fenómenos de transferência e empréstimo.

Por outro lado, todas as línguas têm as suas próprias regras sintáticas. Uma das manifestações típicas de fossilização sintática registada em falantes do PtgL2 que têm línguas do grupo níger-congo atlântico como L1's prende-se com o uso do tempo e modo verbais, já para não falar no sistema de marcação do tempo pretérito e do participio passado. Como o sistema verbal das suas L1's não detém idênticas características, os falantes não conseguem fazer uma distinção clara das diferentes formas da LA, e acabam por recorrer às suas formas cognitivas para marcarem o tempo verbal, originando-se, assim, as fossilizações.

Alguns autores (Pienemann & Johnson; 1987; Pienemann, 2000; Plag, 2008a, 2008b, 2008c, 2008d) opinam que a ASL atravessa seis estágios incontornáveis de desenvolvimento da interlíngua (Tabela 3.1), iniciando-se com um desenvolvimento gramatical nulo, que leva à articulação de palavras simples, as quais, por sua vez, podem também representar fórmulas, tal como acontece na fase em que as crianças em aquisição de L1 produzem frases holofrásticas. No estágio seguinte a interlíngua começa a apresentar frases simples, mas completas, se bem que o falante ainda não faça a distinção entre os elementos que as compõem. Portanto, as palavras são ordenadas em função dos seus significados ou focos de informação nelas contidos. Tal como acontece com a produção das crianças em processo de aquisição de uma L1 na fase pós-holofrástica, as palavras, articuladas em função de relações semânticas, não apresentam marcas sintáticas e morfológicas nem flexão contextual. Plag (2008b:314) refere que a maioria das línguas crioulas não ultrapassa este estágio de desenvolvimento, podendo algumas revelar ligeira transição para o terceiro estágio (apresentam indícios primários de flexão contextual), mas sem conseguirem ultrapassá-lo. No caso do santomense, um crioulo do tipo basilectal, é evidente a estagnação de alguns dos seus falantes no segundo patamar do desenvolvimento da interlíngua, dado evidenciarem escassas produções de flexão inerente e quase total ausência de flexão contextual. Ainda assim,

os registos do nosso *corpus* revelam que a maioria dos falantes de Almojarife realiza flexão contextual em maior ou menor grau. Portanto, o terceiro estágio é aquele que melhor caracteriza a produção da generalidade da comunidade (Tabela 3.1), já que se nota nos seus membros a capacidade de conseguirem identificar o início e o fim da construção frásica, a ponto de trocarem a ordem dos elementos que podem ocorrer nestas duas posições (exemplo [213]) ou estabelecerem concordância, embora variável, a nível dos elementos do SN.

[213] PA:  
*lá pas 14 que nós conseguimos trazê isso pa casa* [OSVALH1]

Paralelamente, a flexão contextual registada em itens incorporados na estrutura do SN denota que a aquisição da pré-sintaxe foi alcançada. Não obstante, e apesar de alguns falantes revelarem algum domínio deste tipo de flexão e aplicarem as estruturas básicas da LA, nenhum deles evidencia uma aquisição completa das regras gramaticais que caracterizam o quarto estágio de desenvolvimento da interlíngua em direcção à LA. Como tal, são incapazes de, por exemplo, ultrapassar na totalidade a variação a nível da estrutura do SN ou caracterizar elementos particulares que compõem a cadeia frásica, movendo-os do interior desta para a posição inicial ou final.

Como a aquisição dos estágios é fixa, não é possível aos falantes abolir qualquer das etapas referidas. Este aspecto pressupõe que a capacidade de realizar as tarefas de um determinado estágio implica a competência de efectuar as do estágio anterior. Ora, é neste ponto que reside a diferença entre as aquisições de L1 e de L2, já que os falantes desta, em determinada fase da sua ASL, produzem construções típicas de diferentes estágios, comprovando-se que fossilizaram, suave ou definitivamente, formas de fases anteriores. Este aspecto evidencia ainda que a fossilização acontece por etapas (Fig. 27), já que, na aquisição de L1, se alguma destas for superada, o falante não apresenta variação característica dessa mesma etapa no patamar seguinte de aquisição, o que não sucede no PA, no qual o desenquadramento distribucional entre flexão inerente e flexão contextual transita de estágio para estágio.

Contudo, deve ser sempre levado em linha de conta que a aquisição dos diferentes estágios varia de indivíduo para indivíduo (Fernandes-Boëchat, 1989:212), comprovando-se que ela assenta em factores diversos, que não apenas os

morfossintáticos. Como tal, as diferenças individuais também actuam a nível da fossilização, com alguns aprendentes a fossilizarem formas de um determinado estágio e outros a fossilizarem formas de um patamar distinto. Assim sendo, a observação das fossilizações pressupõe que a busca das suas motivações tenha que ser vista não só levando em conta as condições internas e externas que condicionam a ASL mas também considerando o tempo durante o qual determinada variação permanece estabilizada.

Seguidamente, passaremos a tecer considerações em torno da aquisição massiva de L2's, a fim de se observar até que ponto a variação, que começa por ser individual, se instala a nível da comunidade, determinando alterações no padrão da fala desta.

### **2.11. Aquisição massiva de L2's: línguas crioulas e teorias de aquisição de L2's**

Pesquisas recentes em ASL têm permitido uma nova visão acerca das propriedades das línguas crioulas e natureza da crioulização (Kouwenberg & Patrick, 2003; Lefebvre, White & Jourdan, 2006a; Siegel, 2008). Com base nestes estudos e nos pressupostos da Teoria da “Processabilidade” (Pienemann, 1998, 2005b), Plag (2008a, 2008b) formulou a Hipótese da Interlíngua (*Interlanguage Hypothesis* – HI), que advoga que os crioulos representam interlínguas convencionais em estágios primários de aquisição. A HI tem respaldo nos estudos de Andersen (1980, 1983a) e observa os processos de ASL que ocorrem nos estágios iniciais do contacto (cf. Kouwenberg & Patrick, 2003), a fim de avançar pressupostos acerca da forma como se processa a flexão morfossintáctica dos crioulos.

#### **2.11.1. Universais linguísticos vs. transferência: a Hipótese da Interlíngua (*Interlanguage Hypothesis*)**

Dado que podem ocorrer transferências sem ASL (p.e. situações de diglossia ou aquisição de L1 em contexto bilingue) e processarem-se, durante a crioulização, estruturas que não resultam de transferência, a HI resgata o pressuposto de que os primeiros estados de ASL são comparáveis aos processos de formação das línguas pidgins (Schumann, 1978). Neste aspecto, os crioulos configuram interlínguas convencionais em estágios primários de aquisição, sendo os patamares avançados da aquisição explicados em termos de factores da língua-E, como, por exemplo, o tipo de acesso ao superstrato. Assim, estes factores terão pouco a ver com a simplicidade

gramatical dos crioulos, que será uma consequência da fase inicial de emersão das línguas nascidas por contacto.

Embora os crioulos apresentem morfologia reduzida, alguns deles possuem flexões a nível nominal (ponto 1.2.2.2), verbal e adjectival. Estas verificam-se unicamente a nível dos morfemas presos que expressam categorias gramaticais, uma vez que os morfemas gramaticais livres, como os marcadores de TMA, não configuram “morfologia flexional” no sentido clássico. Exemplos de ambos os tipos de morfologias podem ser consultados em alguns trabalhos sobre crioulos (p.e. Hudson, 1983; Holm, 1988/1989; Stolz, 1989; Kouwenberg, 1994, Bakker, 2003; Baptista, 2003; Veenstra, 2003). Luís (2008) providencia também algumas evidências de sufixos que englobam noções de tempo e aspecto em conjugações verbais de alguns crioulos indo-portugueses.

#### **2.11.1.1. Teoria da “Processabilidade” e flexão morfológica das línguas emergentes por contacto**

Os processos de flexão, independentemente de serem inerentes, contextuais ou morfofonémicos, fazem uso dos seguintes elementos formais: afixação, clitização, reduplicação, metafonía e supletismo. Os pidgins e crioulos exibem ausência de flexão quase total, em virtude de apresentarem um desenquadramento distribucional entre flexão inerente e flexão contextual. A distinção entre flexão contextual e flexão inerente dá-se a nível da informação que é necessária ou não para tornar a comunicação eficiente. A aplicação da concordância ou de casos estruturais envolve troca de informação intra e interfrásica, ao passo que a flexão inerente não requer troca de informação entre constituintes distintos. A concordância sujeito-verbo e o uso do caso sujeito, por exemplo, requerem um processamento-S, enquanto a atribuição dos casos objecto ou genitivo exigem troca de informação frásica entre núcleo e complemento (Fig. 15). Por seu lado, a concordância de número e/ou de género obrigam à troca de informação intrafrásica.

Aplicando estes processamentos aos crioulos, constata-se que eles possuem, quase exclusivamente, casos típicos de flexão inerente (p.e. pluralização de nomes e/ou flexão verbal de tempo e aspecto), ou seja, estruturas que praticamente não requerem troca de informação entre os seus constituintes, revelando tendência para a *interface* léxico-sintaxe mínima e transparente (Kihm, 2003).<sup>102</sup> Assim, os casos de flexão contextual são

raríssimos, enquanto situações de flexão morfofonémica não acontecem de todo (Crowley, 2008; Plag, 2008a). Partindo ainda do pressuposto que os crioulos configuram casos típicos de ASL, então é lógico predizer que estamos perante interlínguas, conforme o confirma a questão da morfologia flexional. Em adição, também se verifica que o valor semântico dos afixos flexionais não é, por si só, um garante da sua sobrevivência geracional (Plag, 2008a:125-126), como o demonstra, por exemplo, o facto de os afixos marcadamente inerentes de tempo/aspecto dos crioulos indo-portugueses precederem a aplicação de flexão contextual, como acontece com a concordância ou a atribuição de caso. Assim sendo, a preservação de marcadores flexionais aponta mais para uma questão de processamento do que para uma questão de transparência semântica ou relevância comunicativa. E nesta conformidade, poder-se-á afirmar que os falantes responsáveis pela emergência dos crioulos recorrem exactamente aos mesmos processos mentais dos outros falantes em situação de ASL. Estes aspectos levam então a concluir que a tipologia da flexão crioula emerge naturalmente, ou seja, em consequência de constrangimentos universais que actuam quer sobre o processamento linguístico quer sobre a aquisição de línguas, exibindo, em sequência de tal, os estágios típicos de desenvolvimento das interlínguas.

O facto de os falantes dos crioulos estabilizarem as suas interlínguas em estágios primários do desenvolvimento (Plag, 2008a:125-126) tem a ver com dois aspectos primordiais: a motivação e o *input*. Na grande maioria dos casos de crioulição, o acesso ao superstrato era bastante limitado, pelo que não havia exposição à LA que permitisse um desenvolvimento da interlíngua para estágios mais avançados de ASL. De facto, é tido como consensual que todo e qualquer conhecimento linguístico é indissociável das capacidades cognitivas do aprendente, pelo que a ASL representa uma construção de sistemas do conhecimento, direccionada para a compreensão e uso. Dado que a capacidade de retenção do conhecimento é limitada, o falante em situação de aprendizagem necessita da experiência e da prática rotineira para ir, progressiva e automaticamente, utilizando os conhecimentos que vai interiorizando (McLaughlin, 1987; McLaughlin & Heredia, 1996; Clements, 2009). Simultaneamente, as condições sociohistóricas em que ocorriam as nativizações das línguas fazem pressupor que os responsáveis pela crioulição não se sentiriam profundamente motivados para



aprenderem na perfeição as línguas dos seus opressores, antes olhariam para elas como meros veículos que serviam a comunicação interétnica (Hymes, 1971a; Silverstein, 1972; Baker, 1994; Baker, 2001; Smith, 2006). Neste aspecto, a hipótese da filtragem afectiva (*Affective Filter Hypothesis*), proposta por Krashen (1982, 1985, 1988) não deixa de ser aqui relevante, uma vez que advoga que, caso as condições psicológicas do aprendente estejam afectadas por aspectos subjacentes ao afecto (motivação, atitude, autoconfiança e ansiedade), ele revela tendência para filtrar o *input* e não processar a aquisição. Assim, se por um lado este tipo de motivação determinou a aquisição de material lexical e algumas propriedades estruturais do superstrato, por outro lado dispensou a aquisição de formas complexas e mais avançadas, como acontece com a atribuição de casos ou a aplicação da concordância morfológica. Portanto, é de esperar que, mesmo em estágios mais avançados da criouliização, se continuem a evidenciar manifestações características de estágios primários de ASL. Este aspecto permite ainda prever que o melhor acesso ao superstrato, a elevada motivação para aprender a LA ou o prolongado contacto com o superstrato, em situações que diluem o afastamento linguístico relativamente, conduzem à aplicação de estruturas crioulas mais próximas dos estágios avançados das interlínguas (Plag, 2008a:129).

Em virtude da variabilidade do desenvolvimento das L2's, é possível observar diferentes graus de traços flexionais em diferentes crioulos, dependendo do tipo de línguas envolvidas no contacto e das situações sociohistóricas em que este acontece. Caso as línguas em contacto sejam morfológicamente elaboradas e apresentem traços próximos, então as hipóteses de sobrevivência da flexão aumentam consideravelmente (Bakker, 2003; Clements, 2009). Esta particularidade pode ser interpretada como transferência em termos de Teoria da "Processabilidade", uma vez que algumas estruturas da L1 poderiam processar-se na LA e, como tal, transitam para a interlíngua, originando padrões não-uniformes na língua utilizada por grupos de aprendentes detentores de L1's tipologicamente distintas (Pienemann & Håkansson, 2007:492). O pressuposto de que os crioulos, enquanto produtos emergentes a partir de pidgins radicalmente reduzidos, evitam os traços desnecessários para a comunicação (McWhorter, 2006:180) e emergem como línguas mais simples do que as línguas não-crioulas (McWhorter, 2000:93), tem sido criticado em vários domínios pela forma

obscura e vaga como foi avançado (Siegel, 2004a; Klein, 2006a, 2006b; Kihm, 2008; Plag, 2008a). Contrariamente, a HI defende que a não-aquisição de categorias flexionais é uma consequência previsível dos constrangimentos que actuam sobre o processamento da língua nas fases primárias da ASL. Assim, a ausência de flexão nestas fases acontece porque, simplesmente, os falantes das interlínguas não detêm os recursos para o processamento da flexão. Quando o desenvolvimento desta se inicia, o ponto de partida é a flexão inerente, que não requer processamentos complicados. Posteriormente, nos estágios mais avançados de desenvolvimento da interlíngua, flexões de outro tipo podem ocorrer por influência de factores externos, especialmente por indução da LA em situações de interacção linguístico-cultural estreita (Winford, 2003a:313).

#### **2.11.1.2. Teoria da “Processabilidade” e processamento sintáctico das línguas emergentes por contacto**

No que concerne aos estudos sobre crioulos, ainda permanecem relativamente obscuros quais os mecanismos responsáveis pela natureza não-marcada de algumas das suas estruturas linguísticas. Outro dos desafios em análises crioulisticas prende-se com a tentativa de se clarificar, com exactidão, se determinada estrutura resulta de uma transferência do substrato ou se, pelo contrário, é regida por tendências universais. Aliás, em algumas situações parece ocorrer mesmo uma convergências de ambos os fenómenos, pelo que a HI pode ajudar a clarificar se as similaridades entre interlínguas e crioulos, por um lado, e entre desenvolvimento da interlíngua e pidginização/crioulização, por outro lado, não são meramente acidentais. De facto, os paralelismos entre a formação de crioulos e a ASL dão-se a nível do desenvolvimento de ambos, da natureza da variação registada nas interlínguas e nos crioulos, do papel da interacção e do factor idade.

Um fenómeno que estabelece paralelismo evidente entre ASL e pidginização/crioulização é o do grau da variação (*variability in “attainment”*), um termo usado em pesquisas de ASL para descrever o patamar de aproximação da variedade linguística do aprendente em relação à LA. Embora o termo *attainment* não seja adequado para os contextos de pidginização/crioulização (Baker, 1994:72), não deixa de estabelecer correspondência com a designação ‘aproximação ao superstrato’ (*proximity to the superstrate*), utilizada no campo das pesquisas crioulisticas.

Efectivamente, ambas as variações dependem da natureza da interacção entre os falantes envolvidos no contacto, já que esta é caracterizada pelos diferentes graus do limitado acesso ao superstrato, da limitada participação nas instituições e práticas reguladas pela língua lexificadora e baixa (ou nula) motivação relativamente à LA. Como tal, o *input* e a interacção, bem como a atitude e a motivação, têm influência notória no grau de variação registado quer na ASL quer na pidginização/crioulização.

A intervenção do factor idade estabelece também um paralelismo evidente entre interlínguas e crioulos. Bastantes estudos têm demonstrado que, na ASL, os aprendentes adultos não só são menos bem sucedidos do que as crianças (Lenneberg 1967; Chomsky, 1972; Brown, 1980; Larsen-Freeman & Long, 1991; Ellis, 1994; Beck, 1998) como se revelam também mais sensíveis a outras questões que interferem no grau de variação, como as sociológicas, as socioculturais e as psicológicas (Hyltenstamm & Abrahamsson, 2003:586). Nos contextos iniciais de línguas em contacto, adultos em condições socioculturais e psicológicas desfavorecidas configuram o público aprendente, deixando antever, à partida, um baixo grau de aproximação à LA.

Os paralelismos referidos permitem entender que algumas características universais da interlíngua podem ser utilizadas como ferramentas heurísticas para o estudo da formação dos crioulos, numa perspectiva comparada. Relativamente à questão das transferências em ASL, é geralmente aceite que podem ocorrer em todos os subsistemas, com especial incidência nos aspectos fonológicos. Alguns estudos crioulisticos têm entendido que similaridades estruturais entre L1 e crioulos representam evidências de transferências. Com base neste pressuposto, é possível estabelecer a assunção patente na tabela seguinte:

**Tabela 5.1.** A assunção: similaridades estruturais configuram transferência (Plag, 2008b:311).

	<b>Traço</b>	<b>L1 / substrato</b>	<b>LA / lexificadora</b>	<b>Interlíngua / Crioulo</b>
<i>a.</i>	traço x	+	-	+
<i>b.</i>	traço x	-	-	-

Em *a.* verifica-se que uma determinada estrutura está presente tanto na L1/substrato como na interlíngua, mas ausente na LA/superstrato, pelo que é forte a probabilidade de se estar perante uma transferência. Opostamente, como a estrutura *b.* não marca presença nem na L1/substrato nem na LA/superstrato, então também não pode ocorrer na

interlíngua. Contudo, para se confirmar se uma determinada estrutura representa transferência, é fundamental concluir que a mesma não emerge universalmente no desenvolvimento da L2, isto é, de forma independente, apesar de o mesmo tipo de estrutura existir na L1 do aprendente (Plag, 2008b:311). Assim, se tomarmos em consideração as questões inerentes aos universais de ASL, o cenário representado na tabela anterior torna-se mais complexo, conforme demonstrado na Tabela 6.1:

**Tabela 6.1.** Os factos: similaridades estruturais podem não representar transferência (cf. Plag, 2008b:311).

	<b>Traço</b>	<b>L1 / substrato</b>	<b>LA / lexificadora</b>	<b>Interlíngua / Crioulo</b>
<i>a.</i>	traço x	+	–	+
<i>b.</i>	traço x	–	–	–
<i>c.</i>	traço x	–	–	+ (!)
<i>d.</i>	traço x	+	+	– (!)

Como se constata agora, emergem novas estruturas (*c.*), enquanto outras, que são compartilhadas por ambas as línguas, desaparecem (*d.*). A exemplificação deste tipo de estruturas pode ser encontrada nos estudos sobre aquisição do alemão L2 por falantes de sueco L1, dois idiomas que compartilham o traço da inversão da ordem frásica em orações principais (V2nd) (Håkansson, Pienemann & Sayehli, 2002), de acordo com o representado na Tabela 7.1:

**Tabela 7.1.** Inversão da ordem frásica na interlíngua alemã (Plag, 2008b:311).

<b>L1: sueco</b>	<b>LA: alemão</b>	<b>Estágios da interlíngua</b>
V2nd	V2nd	1. SVO
*ADV SVO	*ADV SVO	2. ADV SVO
		3. ADV VSO

Nas duas línguas, a colocação do advérbio em primeira posição sem inversão é agramatical, uma similaridade que não é compartilhada pela interlíngua no seu estágio 2 de desenvolvimento, já que se mantém a ordem frásica SVO do estágio 1 da aquisição, configurando-se o esquema *c.* da Tabela 6.1 e não se confirmando a representação *b.* da Tabela 5.1. Note-se também que, apesar de o alemão e o sueco possuírem a estrutura V2nd, esta só emerge em sequência do desenvolvimento da interlíngua, ou seja, no estágio 3 (Håkansson, Pienemann & Sayehli, 2002:259). A questão da ordem frásica

levanta ainda outro problema: dado que a interlíngua emerge com a ordem SVO, inexistente na L1 e na LA, então a ocorrência deste tipo de traço não pode ser vista como uma transferência, caso a L1 possua a ordem SVO e entre em contacto, por exemplo, com uma língua do tipo VSO, ou seja, se configurem os traços patenteados em *a*. (Tabelas 5.1 e 6.1) e que apontariam para a evidência de transferência.

### **2.11.2. Hipótese do Desenvolvimento Moderado da Transferência (*Developmentally Moderated Transfer Hypothesis*)**

Face às questões apresentadas, é importante determinar quais os traços conectados a questões universais do desenvolvimento da interlíngua e quais os que resultam de transferência. Neste aspecto, a Teoria da “Processabilidade” (Pienemann, 1998, 2005b) e a HI (Plag, 2008a, 2008b) permitem não só elaborar assunções e predições mas também testá-las empiricamente (Figs. 13 e 14; Tabela 2.1, nota de fim de texto 63 e Tabela 3.1). De acordo com a hipótese de que a GU actua na L2 através da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), a FL0 da interlíngua tem que ser observado como a FLf da L1, o que leva a predizer que ocorrem transferências, principalmente na fase inicial da aquisição da L2. Contrastando com esta posição, Pienemann *et alii* (2005a, 2005b) e Pienemann & Håkansson (2007) avançaram com a Hipótese do Desenvolvimento Moderado da Transferência (*Developmentally Moderated Transfer Hypothesis*), que advoga que os aprendentes de uma L2 não iniciam a aquisição a partir da sua L1, mas sim activando o sistema de processamento e articulando os processos disponibilizados pela L2. De acordo com este pressuposto, qualquer estrutura da L1 só é transferida no caso de encontrar condições para ser processada no sistema de desenvolvimento da L2, isto é, caso o processador da interlíngua tenha disponível um processo que accione o processamento da estrutura da L1 passível de transferência. Como se depreende, a Hipótese do Desenvolvimento Moderado da Transferência não exclui a possibilidade de ocorrerem transferências, mas defende que elas são constringidas pela “processabilidade” da L2, ou seja, que elas se dão em tempo distinto daquele que é avançado pela sugestão de que a GU actua na L2 através da L1. Este aspecto pode ser de novo empiricamente observado nas construções que envolvem V2nd produzidas por falantes de sueco L1 em aquisição do alemão L2 (Tabela 7.1), já que os mesmos apenas produzem este tipo de estruturas no estágio 3 do desenvolvimento da

interlíngua (Håkansson, Pienemann & Sayehli, 2002:259). Este fenómeno apenas ocorre depois de os falantes terem ultrapassado dois estágios de desenvolvimento da interlíngua que apresentam estruturas sem qualquer correspondência na L1 e na LA, apontando para a possibilidade de o efeito apenas emergir quando o sistema de processamento é capaz de reorganizar a gramática da interlíngua com recurso aos apropriados conhecimentos da L1. Desta forma, não deixam de ser aqui pertinentes quer o princípio advogado por Montrul (2004:365) de que o desenvolvimento sintáctico é conduzido pela aprendizagem de novos itens lexicais e morfológicos quer determinados pressupostos da Hipótese da Maturação, que advogam que certos aspectos operativos da aquisição se dão de modo descontínuo, isto é, em período mais tardio do que aquele que seria previsto acontecerem (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997).

Note-se que, em virtude de a transferência ocorrer apenas se o processamento o permitir, isto é, em estágios hierárquicos apropriados do desenvolvimento da interlíngua, pode também acontecer no estágio inicial, conquanto as estruturas transferidas sejam processáveis neste. É o que sucede, por exemplo, com a ordem básica SVO, quando ela se regista também na L1. De facto, após um breve estágio inicial em que formulam frases holofrásticas, os aprendentes de L2's passam a produzir, predominantemente, construções com a ordem canónica, isto é, SVO ou SOV, independentemente de a L1 e/ou a LA possuírem ou não estas mesmas ordens (Håkansson, Pienemann & Sayehli, 2002:259). A explicação para este comportamento inicial dos aprendentes é explicada por Pienemann, Biase e Kawaguchi (2005:229) em termos de Gramática Lexical Funcional, já que ocorre uma associação fixada entre a estrutura argumental, a estrutura funcional e a estrutura constituinte, designada pelos autores de “sequência não-marcada”.<sup>103</sup>

De acordo com esta teoria, o mapeamento directo não requer nenhum processador específico da língua nem qualquer armazenamento, mas permite que o aprendente produza sequências SOV ou SVO. Deste modo, e ao contrário do que seria de esperar de acordo com o predito pela hipótese de que a GU actua na L2 através da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), os aprendentes ingleses de japonês L2, por exemplo, começam por produzir a sequência SOV (Biase &

Kawaguchi, 2002:293), em vez de transferirem a sequência nativa SVO. Este aspecto, se observado na perspectiva da transferência vs. desenvolvimento universal, evidencia que nem a sequência SOV nem a ordem SVO produzidas por aprendentes de L2's podem ser vistas como transferências da L1, já que ambas são realizadas nos estágios iniciais da aquisição.

Observando-se agora a ordem básica de SN's com sujeitos e objectos não pronominalizados de diferentes crioulos, verifica-se que todos eles compartilham a estrutura canónica da ASL, apesar de os substratos e as línguas lexificadoras apresentarem, entre si, similaridades e/ou diferenças:

- [214] a. CRIOULO DO HAITI:  
 Crioulo do Haiti: S AUX VO  
 Gbe: S AUX VO / OV  
 Francês: S AUX VO  
 (Lefebvre & Brousseau, 2002)
- b. SURINAMÊS ou SRANAM:  
 Surinamês: S AUX VO  
 Gbe: S AUX VO / OV  
 Inglês: S AUX VO; X S AUX V  
 (Bruyn, 2002)
- c. PALANQUERO:  
 Palanquero: S AUX VO  
 Kikongo: S AUX VO  
 Espanhol: S AUX VO  
 (Bentley, 1967 [1887]; Laman, 1936; Schwegler, 1991)
- d. NEGERHOLLANDS:  
 Negerhollands: S AUX VO  
 Kwa: S AUX VO  
 Holandês: V2nd, X AUX SOV  
 (Muysken, 2001)
- e. CRIOULO DE BASE HOLANDESA DE BERBICE:  
 Crioulo de Berbice: S AUX VO  
 Ijo Oriental: SOV  
 Holandês: V2nd, X AUX SOV  
 (Kouwenberg, 1992, 1994)

Os exemplos evidenciam paralelismos estreitos entre estruturas dos estágios iniciais da interlíngua e estruturas crioulas, cujas emergências não podem ser totalmente

explicadas em termos de transferência. Assim, e de acordo com a HI, é possível prever que algumas línguas emergentes por contacto apresentem também a ordem SOV, como de facto acontece:

[215] PIDGIN DE NAGALAND:  
*Kikatemla modu kha-yas-ile*  
Kikatemla vinho comer-PROG-PRET  
“O Kikatemla estava a beber vinho” (Bhattacharjya, 2007:240)

De um modo geral, a HI, em congruência com a Teoria da “Processabilidade”, sustenta que, nos crioulos, a ordem das palavras reflecte uma sequência não-marcada, não havendo evidência em favor da transferência (Plag, 2008b:320).<sup>104</sup> Relembre-se também que, de acordo com os princípios da mudança evolucionista da língua, a saliência perceptual se pode impor à frequência na LA, fazendo com que uma estrutura não-marcada universalmente, por ser mais facilmente processável (Thomason & Kaufman, 1988:331), seja adoptada pelos falantes da L2 (Clements, 2009:22).

Tendo ainda em conta que a emergência da flexão contextual se instancia a partir do nível 4 (Tabela 3.1) da hierarquia da “processabilidade” (Plag, 2008a, 2008b), pode inferir-se que os crioulos não alcançam este patamar de desenvolvimento sintáctico. A causa tem então origem em factores de natureza psicolinguística, nomeadamente nos traços universais que motivam a ASL, responsáveis pela emergência de muitas das estruturas crioulas, geralmente consideradas não-marcadas.

Não obstante o contributo da HI para ajudar a clarificar o porquê de a flexão contextual se encontrar ausente em pidgins e crioulos, ficam em aberto questões para as quais a hipótese não fornece respostas satisfatórias: (i) Porque é que algumas flexões inerentes são preferidas em detrimento de outras? (ii) Porque é que a flexão morfofonémica é também aplicada em pidgins e crioulos, contrariando o pressuposto de que não deveria ocorrer nestas línguas; (iii) Porque é que, nestas mesmas línguas, a flexão inerente apenas recorre a um subconjunto de significados formais disponíveis, desfavorecendo a afixação e excluindo os casos de metafonia e supletismo?<sup>105</sup>

### **2.11.3. Transmissão linguística irregular**

Os conceitos de transmissão linguística irregular e de pidginização/crioulização reportam-se a processos históricos de contacto massivo e prolongado entre línguas, nos



quais intervêm um idioma que é tomado como LA pelos falantes dos demais dialectos. Todavia, o primeiro conceito é mais abrangente do que o segundo, visto envolver processos de natureza linguística e sociohistórica que conduzem ao emergir quer de uma nova língua (pidgin ou crioulo) quer de uma nova variedade linguística, que se tornam predominantes na situação de contacto. Portanto, a criouliização pressupõe uma situação de reduzido acesso à LA, que vai determinar o desencadeamento da reestruturação linguística e, conseqüentemente, o aparecimento de uma nova língua. Esta é caracterizada por acentuada não-aquisição de morfologia flexional, o que a torna bastante diferente das que lhe concederam os modelos iniciais para aquisição e conseqüente nativização. Por seu lado, a transmissão linguística irregular configura um *continuum* de níveis diferenciados de socialização/nativização das L2's adquiridas em situações sociohistóricas específicas. Desta forma, a ecologia destas línguas tem como conseqüência resultados distintos entre si, caracterizados por não-aquisição de morfologia flexional, a qual pode ir de grau mais leve a grau mais acentuado.

#### **2.11.3.1. Vectores para a emergência da transmissão linguística irregular**

As massas populacionais detentoras de diferentes L1's, quando constituindo uma comunidade, têm que adquirir unicamente por contacto, para efeitos de comunicação funcional, uma L2 de emergência comunicativa. Esta interlíngua transmite, lexicalmente, aquilo que outras línguas expressam de modo gramatical, sendo caracterizada por vocabulário reduzido mas utilitário, não-realização de segmentos com pouco peso fonético e uma acentuada redução gramatical, uma vez que os falantes apenas incorporam no novo dialecto os itens que preenchem funções gramaticais básicas. Por norma, a redução fica a dever-se a três factores principais: (i) pouco contacto com os detentores da LA, principalmente porque estes controlam o poder político e se configuram como um grupo minoritário de difícil acesso; (ii) o facto de a esmagadora maioria dos falantes em aquisição ser adulta, ou seja, iniciar a aprendizagem da L2 após o período crítico da aprendizagem; (iii) a ausência de aplicação de uma norma que oriente formalmente o processo de construção da interlíngua, já que esta responde apenas a questões de urgência comunicativa.

De modo geral, o que antes se definiu foi a emergência de um pidgin, isto é, de uma L2 a partir de *input* reduzido (modelos/estímulos fragmentados e deficientes) e que não é L1 de nenhum dos diferentes grupos em contacto. O distanciamento em relação à LA determina, desde logo, que o pidgin resulte essencialmente dos processos universais de aquisição da L2 (simplificação morfológica ou nivelção) e transferências da L1 (Schumann, 1978; Gilbert, 1983; Bickerton, 1984a, 1988), pelo que não pode ser explicado apenas em termos de simplificação da LA (Ferguson, 1971). Nesta fase do seu desenvolvimento, o pidgin não desempenha funções interactivas (coesão social), expressivas (ideias abstractas), poéticas e metalinguísticas isto é, algumas das funções que caracterizam as línguas que não emergem por contacto massivo. Portanto, acaba por funcionar de forma pragmática/pré-sintáctica, com base em discursos universais independentes da sintaxe (Givón, 1979, 1984), podendo ser definido, convencionalmente, como *um simplificado sistema linguístico não-nativo, lexical e estruturalmente reduzido, instável, com padrões variáveis e inconsistentes (devido à influência das diferentes línguas nativas em contacto), criado e desenvolvido rapidamente para ser usado como meio de comunicação emergencial (L2) em situação restrita de contacto funcional entre adultos que são falantes nativos de línguas mutuamente ininteligíveis*. Esta caracterização não é meramente linguística, reflectindo uma resposta social necessária e adequada à realidade imposta pelo contacto. Por outro lado, o pidgin impossibilita o bilinguismo, por razões sociohistóricas e políticas.

Contudo, se a situação de contacto se mantiver, a interlíngua de emergência evolui naturalmente, tornando-se o dialecto que preenche as funções sociais na rede de interacção linguística, ao mesmo tempo que vai ganhando estatuto de L1 para os descendentes dos falantes do pidgin e das outras línguas de substrato que vão engrossando a comunidade (Cornips & Hulk, 2006). A limitação estrutural da língua começa a revelar-se insuficiente para dar resposta às novas exigências funcionais da comunidade, pelo que esta leva a efeito a expansão do pidgin (Mühlhäusler, 1986; Siegel, 2008), gramaticalizando os itens lexicais importados da LA e incorporando novos dispositivos gramaticais da língua do substrato (Thomason & Kaufman, 1988:73-74). Da troca entre expansão funcional e expansão gramatical emerge o crioulo, enquanto código linguístico que ganha identificação e referências próprias para passar a

preencher as necessidades sociolinguísticas da comunidade. Assim, e de um modo geral, a noção de crioulização surge relacionada com a emergência/desenvolvimento de uma L1, com *input* restrito, isto é, a partir de estímulos/modelos defeituosos de variedades de uma L2. Por seu lado, os efeitos do processo de aquisição variam de acordo com a natureza da L2, enquanto elemento que fornece os DLP's que actuam como estímulos para a aquisição da L1. Nesta fase, a língua nativizada detém ainda duas funções/características: a de L1 para os falantes nascidos no seio da comunidade; e a de L2 para os membros transplantados para a comunidade (DeCamp, 1971:16). Temos então crianças e adultos intervindo no processo de crioulização, com as primeiras emprestando a capacidade da competência linguística, enquanto os segundos a desdobram no processo de uso (Kroch, 1989a; Sprouse & Vance, 1999; Siegel, 2004b). Na comunidade processam-se então dois tipos de interacção linguística: a interacção na horizontal, entre todos os membros da comunidade; e a interacção na vertical, entre os falantes dos distintos substratos e seus descendentes directos, detentores da L1 nativizada. Quanto aos contactos com a LA, embora esporádicos, não cessam, ocorrendo, essencialmente, entre trabalhadores braçais e seus capatazes e/ou entre serviçais e seus amos. Assim, no momento em que a interlíngua da comunidade é objecto de expansão estrutural com a introdução de novas formas oriundas do substrato, estas entram em conflito com formas da LA.

Segundo Kihm (1980:12), para que uma língua seja considerada um crioulo, deve obedecer ao triplo “paradoxo crioulo”:

1. crioulos de bases diferentes assemelham-se mais entre si do que com a LA;
2. abstraído das diferenças lexicais, crioulos de bases diferentes estão mais próximos uns dos outros do que as LA's estão entre si;
3. crioulos da mesma base lexical não são, devido a isso, mais parecidos entre si do que os crioulos de bases diferentes.

Para alguns linguistas, estas similaridades são simplesmente conjunturais e decorrem da mera presença ou influência das línguas de substrato no processo de formação dos crioulos. Algumas propriedades são discutidas por Muysken (1981a; 1988) e vistas como resultando de valores não-marcados de parâmetros (Bickerton, 1988, 1999;

Roberts, 1997, 1999). Nesta etapa, já não faz então sentido continuar a observar a interlíngua em termos de pidginização ou criouliização, uma vez que o processo de transmissão da mesma varia em função de factores vários, provocando alterações significativas na sua estrutura. Por outro lado, é ainda de ter em consideração que a interlíngua poderá também entrar em contacto com outras línguas, como aconteceu no Brasil, quando o PtgL2 falado pelas comunidades escravas transplantadas passou a contactar com os dialectos dos grupos ameríndios ou oriundos de outras partes de África. Assim, o panorama que nos é oferecido é o de uma língua falada detentora de uma gramática reestruturada, mas que continua a incorporar, por empréstimo, traços das línguas com as quais passou a ter contacto (cf. Thomason & Kaufman, 1988; Lefebvre, 2003; Montrul, 2006; Lefebvre, White & Jourdan, 2006b; Siegel, 2008).

Nesta fase do novo contacto linguístico, os modelos linguísticos emergentes podem ser tidos como vigentes pelos descendentes dos membros da comunidade, que estabelecem uma norma linguística a partir dos mesmos, isto é, ao invés de facultarem a emergência de um novo sistema linguístico distinto da LA (um crioulo), apenas dão origem a uma nova variedade desta. Contudo, este novo código linguístico não deixa de exibir as características defectivas de que foi alvo durante o seu processo de socialização/nativização. Nesta conformidade os dialectos emergentes por transmissão linguística irregular, tal como os pidgins e crioulos, devem ser vistos como fenómenos que conjugam factores linguísticos e sociohistóricos na sua emergência.

#### **2.11.3.1.1. Sociohistória e transmissão linguística irregular**

Sociohistoricamente, o transplante de massas populacionais escravizadas entre os sécs. XVI e XIX por colonizadores europeus para novos contextos geográficos ou geográfico-culturais configurou-se como um dos factores para a emergência de línguas pidgins ou crioulas. A constante necessidade de mão-de-obra para implementação e expansão de empreendimentos agrícolas ou exploração mineira, sobretudo no continente americano, foi preponderante no processo de escravização de africanos, em virtude de as populações ameríndias, pelos seus hábitos culturais e de recolectagem, resistirem ao processo de aculturação e não possuírem características físicas para suportar esforço físico intenso.

No processo de transplantação para o *Novo Mundo*, os escravos resgatados no continente africano eram inicialmente conduzidos para o entreposto de São Tomé, onde permaneciam durante algum tempo a fim de prestarem serviço nos engenhos da cana-de-açúcar dos colonos, entretanto aí fixados. Detentores de diferenciadas L1's, os escravos viam-se na necessidade de aprender o PtgL2, enquanto código linguístico que respondia às necessidades emergentes de comunicação não só entre os membros da comunidade escrava (cf. Hymes, 1971a; Silverstein, 1972; Baker, 1994; Baker, 2001; Smith, 2006) mas também entre estes e os seus capatazes e senhores. Posteriormente, foi-se construindo um cenário complexo de interacções sociolinguísticas na ilha, que acabaram por deter papel preponderante, em maior ou menor grau, no processo de socialização/nativização da L2: constante chegada de novos escravos resgatados de vários pontos do continente africano, alguns já falantes de um PtgL2 defectivo; aumento da população nativa descendente dos escravos, que, conjuntamente com os membros transplantados, conduziu ao desequilíbrio populacional entre colonos e escravos, com os segundos constituindo a grande maioria da massa demográfica fixada na ilha; envio de cerca de duas mil crianças judias com cerca de oito anos de idade, de ambos os sexos, e oriundas de Castela, retiradas pelo rei João II de Portugal aos seus pais, a fim de pressionar os judeus residentes em Portugal a optarem pela religião cristã; incremento da população mestiça, já que o mesmo soberano havia permitido que os colonos portugueses tomassem uma escrava negra por companheira, população essa que, pelo seu bilinguismo e condição de homens livres, terá contribuído para um estreitamento das distâncias na rede de interacção sociolinguística entre colonos e escravos; êxodo dos colonos portugueses, em finais do século XVI e princípios do século XVII, que abandonaram o arquipélago por falta de segurança e regressaram a Portugal ou se fixaram no Brasil, levando o santomense a deixar de contactar com a LA e a estabilizar-se como basilecto (Ferraz, 1979:19); ocupação do arquipélago pelos holandeses em meados do século XVII e contacto da população residente com uma nova realidade linguística; chegada de novos fluxos de africanos a partir do século XIX, contratados em Angola e Moçambique, alguns falando PtgL2 defectivo marcado por características bantu, e que emprestaram mão-de-obra “contratada” às roças do cacau e do café; fixação de trabalhadores provindos de Cabo Verde, nos anos 50 do século XX, falantes do CCV;

criação de um sistema de ensino e aplicação de uma norma que orientava o processo de aquisição linguística; incorporação de elementos nativos nos quadros profissionais da administração e do corpo militar coloniais; retirada definitiva dos portugueses após a Revolução dos Cravos, em 1974; integração de professores nativos nos novos quadros do ensino, detentores de um código linguístico cujo registo se aproxima do modelo do português popular falado na ilha. Como se verifica, a conjuntura sociohistórica e linguística do arquipélago proporcionou que, diacronicamente, o crioulo se estabilizasse como L1 sob a forma de basilecto, ou seja, se elaborasse um código linguístico que favoreceu e fixou a transferência das estruturas dos substratos. Posteriormente, o contacto do santomense com o português determinou a emergência, em situação de transmissão linguística irregular, do PA reestruturado, marcado por variação que caminhava no sentido da aquisição das regras da LA. Como tal, as diferenças nos registos da comunidade foram-se acentuando geracionalmente em direcção à língua de maior prestígio, registando-se um quase apagamento de determinadas características estruturais dos substratos, directo ou ancestrais.

Os estudos de Baker (1982) sobre os crioulos das ilhas Maurício e Reunião revelaram que o maior ou menor acesso aos modelos da LA é determinante para o tipo de língua que emerge do contacto linguístico, o que viria a fundamentar o princípio de que o estudo sobre línguas em contacto não pode ser demarcado da noção de crioulição variável (Bickerton, 1984a). Nestes estudos, a crioulição do tipo clássico configurará um dos extremos do processo de aquisição de uma LA, em que o processo de reestruturação, face ao escasso acesso à LA, gera uma língua distinta daquelas que lhe forneceram os modelos para a aquisição/nativização. Por seu lado, o processo de transmissão linguística irregular configura um *continuum* de diversificados níveis de socialização/nativização de uma L2 adquirida de modo imperfeito ou de aquisição defectiva de uma L1 por grupos populacionais em contextos sociohistóricos específicos, como sucedeu em São Tomé.

#### **2.11.3.1.2. Transmissão linguística irregular: parâmetros linguísticos**

A transmissão linguística irregular tem como alicerce as situações de comunicação social defectivas em que emerge ou se desenvolve um código linguístico de urgência,

que assenta, por sua vez, em termos retirados do léxico da LA. O manancial dos itens lexicais presente na variedade simplificada da LA é mais ou menos extenso, apresentando formas que vão desde soluções sintáctico-gramaticais transferidas da L1 até aspectos mais efectivos da língua lexificadora. Deste modo, os processos de pidginização/crioulização levados a efeitos por comunidades compostas por falantes adultos de línguas tipologicamente distintas geram elementos da LA que contêm reduções e simplificações, mas também itens relexificados com base nas L1's. Portanto, os pidgins e crioulos vão exibir formas completamente inovadoras, isto é, inexistentes na LA (Ellis, 1985; Gass, 1988; Winford, 2003a; Siegel, 2006, 2008).

O fenómeno da redução visa fornecer soluções que tornam mais simples algumas formas gramaticais das outras línguas envolvidas no contacto e pode ser justificado por questões diversas. Em termos de uma abordagem construtivista (*constructivist approach*), a emergência de determinados itens do novo código linguístico ficar-se-á a dever não ao facto de os falantes terem a L2 como alvo, mas sim à necessidade de estabelecerem uma comunicação meramente interétnica (Hymes, 1971a; Silverstein, 1972; Baker, 1994; Baker, 2001; Smith, 2006), recorrendo, para tal, aos meios cognitivos que já se encontram disponibilizados (Plag, 2008a, 2008b). A constante integração de novos membros na comunidade pode conduzir a mudanças semelhantes às registadas pelos fluxos de imigrantes, uma vez que novas variedades são introduzidas na referida comunidade (Horvarth, 1985:84). Em trabalho levado a cabo no âmbito do European Science Foundation Project,<sup>106</sup> Klein & Perdue (1992) observaram, diacronicamente, a ASL por contacto entre grupos de imigrantes, tendo constatado que tanto os morfemas livres como os morfemas presos de função meramente gramatical se encontravam ausentes das variedades básicas faladas pelos ditos imigrantes (*foreigner talk*). Outro dos factores que intervém no processo de redução das interlínguas em emergência tem a ver com questões do tipo psicolinguístico, já que as limitações perceptuais por parte de falantes adultos de línguas tipologicamente distintas da LA podem levar estes a não estabelecerem a segmentação entre diferentes sequências morfológicas, que acabam sendo tratadas como um todo indivisível (DeGraff, 1999b:482), deixando-se, conseqüentemente, de se estabelecer a correspondência entre as características formais e as formas morfofonológicas de determinados itens da LA

(Prévost & White, 1999 – MSIH). Seja por um motivo seja pelo outro, certo é que a redução do acervo das funções desempenhadas pela comunicação verbal e a não-aquisição de matéria gramatical incidem, com maior acuidade, sobre as estruturas mais abstractas da língua, já que estas apresentam uma carga semântica menos transparente, logo menos funcional em situações de comunicação de emergência.

Lefebvre (1986; 1998) e Jake & Myers-Scotton (1998) avançaram com princípios advogando que a criouliização é um processo mental e sociolinguístico. Os resultados de Lefebvre evidenciam que acontece relexificação na ASL, confirmando-se que a mesma necessita de monolingüismo para poder ocorrer (Thomason & Kaufman, 1988:73-74). Outro aspecto a ter em conta é que, nas hipóteses de relexificação, o inventário de núcleos funcionais e suas propriedades permanece invariável na sua transmissão da língua de substrato para o crioulo. Contudo, a elaboração de novas categorias implica desenvolvimento da sintaxe complexa e da morfologia gramatical de uma língua em contacto na qual elas não existiam anteriormente, isto é, a emergência, por exemplo, de um sistema de marcação de TMA onde antes eram apenas usados advérbios. Klein & Perdue (1992:87) salientam o facto de, na linguagem dos imigrantes, os elementos funcionais não serem adquiridos da mesma forma que os itens lexicais. Deste modo, categorias funcionais como determinantes ou marcadores de TMA, apesar de terem origem nas categorias lexicais do superstrato e entrarem como categorias lexicais no crioulo incipiente, são reanalisadas como categorias funcionais com a não-aquisição da flexão, isto é, como marcadores pré-verbais. Na base deste desenvolvimento gramatical está a mudança linguística interna, envolvendo processos de reanálise e extensão, mas também a influência de outras línguas, especialmente dos idiomas do substrato (Siegel, 2004b:362).

O prolongamento da situação de contacto leva o código linguístico de emergência a ter de responder ao incremento da rede de relações sociais das comunidades. Deste modo, ele é adoptado como veículo de comunicação entre os membros da franja social dominada, que vai abandonando, a pouco e pouco, o uso das L1's, principalmente nas situações de falantes geográfica e culturalmente transplantados, como sucedeu em São Tomé. A cimentação do uso do código de emergência e a necessidade de este dar resposta às novas solicitações é determinante para a recomposição da sua estrutura



gramatical, que leva ao surgimento da nova língua, qualitativamente distinta dos dialectos inicialmente envolvidos no contexto do contacto. É o que sucede, por exemplo, com o *tok pisin*, da Nova Guiné, ou o *kriol*, de Belize, antigas Honduras Britânicas, vistos como pidgins expandidos, em virtude de se terem desenvolvido gramatical e funcionalmente, mas sem se tornarem na língua nativa da maioria dos seus utilizadores.

Este aspecto vem demonstrar que a reestruturação gramatical não está unicamente dependente da nativização da língua, conforme defendido pela linha bickertoniana. De facto, os contextos de crioulização determinam a aquisição de L1's em moldes distintos dos da aquisição de L1's que não conhecem uma situação de transmissão linguística irregular, uma vez que a variação orientada pela fixação dos parâmetros da GU está dependente dos DLP's que actuam como gatilhos (*triggers*) para fixação das regras gramaticais. Quando a sociohistória propicia a crioulização, os DLP's transmitidos pelos falantes adultos às gerações seguintes são bastante variáveis e defectivos, o que vai fazer com que a gramática dos mais novos se aproxime da dos mais idosos, apesar de não ser totalmente idêntica. Como se viu no ponto 2.11.2 do presente trabalho (Hipótese do Desenvolvimento Moderado da Transferência), os crioulos não ultrapassam, por norma, o estágio 3 do desenvolvimento da interlíngua (Plag, 2008b:325), o que compromete a morfologia flexional da L2, ficando esta marcada pelas reduções provenientes das fases iniciais da aquisição. Deste modo, também a aquisição da L1, sustentada por este tipo de DLP's em cadeia, vai apresentar defectividade, assente, por exemplo, em morfosintaxe do género analítico, que visa compensar a falta de informação advinda da ausência de morfemas de género e número. Nesta conformidade, factores externos (*E-factors*) de natureza sociohistórica constituem-se também como uma das causas das mudanças que ocorrem nas línguas (Meillet, 1906; Schuchardt, 1980 [1909]). E na crioulização, estes factores associam-se aos recursos linguísticos da faculdade humana da linguagem (factores internos > *I-factors*) para construir um novo sistema linguístico, definido por Rizzi (1999:465) como a “invenção gramatical”.

É tido como ponto assente que existem muitos traços em comum entre pidgins e crioulos, uma vez que ambos perdem, em maior ou menor grau, estruturas gramaticais, consoante o seu processo sociohistórico de formação. Este aspecto é determinante na fase inicial da transmissão linguística irregular, visto que vai, igualmente, determinar o

grau de reestruturação gramatical, se bem que os factores sociolinguísticos e psicocognitivos dêem também o seu contributo para esta. Posteriormente, a LA determina os vectores de expansão estrutural da nova variedade linguística, já que a progressão é feita em direcção a ela, com todas as envolventes que isso acarreta em termos de pressão sociolinguística. Nesta fase, compreende-se que os processos de transferência de formas dos antigos substratos sejam menores, até porque os falantes já são bilingues. Contudo, estruturas fossilizadas não deixarão de acompanhar, em diacronia, as consecutivas fases de expansão estrutural, ainda que o peso da reestruturação se faça sentir com maior intensidade não só em virtude da pressão sociolinguística mas também por causas bionaturais conotadas à aquisição de L1's (faculdade da linguagem).

#### **2.11.3.1.3. Transmissão linguística irregular e gramaticalização: heterogénese GU/transferência**

Genericamente, a gramaticalização pode ser definida como a mudança operada a nível dos itens e construções lexicais, levando a que percam, total ou parcialmente, o seu significado lexical, a fim de, em determinados contextos linguísticos, passarem a deter funções gramaticais que se encontravam ausentes do código linguístico vigente. Uma vez gramaticalizados, esses elementos continuam a desenvolver novas funções gramaticais (Hopper & Traugott, 2003:16). Assim, nomes e verbos portadores de significado lexical podem transformar-se, ao longo do tempo, em formas auxiliares, marcadores de caso e flexão ou conectores frásicos. Contudo, o processo de transição deste tipo de palavras principais em palavras gramaticais não é imediato, já que têm de passar por um gradual processo de esvaziamento lexical, isto é, transitar de palavras principais a palavras acessórias e, só depois, fixarem-se como palavras gramaticais.

Em termos generativistas, Roberts (1993:384-389) define a gramaticalização como a transposição de um item de categoria lexical para categoria funcional, acompanhada de não-aquisição do seu conteúdo funcional. Assim, ocorre também uma recategorização sintáctica do item ou forma gramaticalizada, levando a que o mesmo perca o seu significado referencial conotado às situações interlinguísticas (universo exterior da língua) e passe a representar um significado gramatical ligado às questões intralinguísticas (universo interior da língua). Com a recategorização sintáctica, o item

ou a forma gramaticalizada adquire “*propriedades funcionais na sentença, sofre alterações morfológicas, fonológicas e semânticas, deixa de ser uma forma livre, estágio em que pode até mesmo desaparecer, como consequência de uma cristalização extrema. Esse trajecto se dá no tempo real quanto no tempo aparente. Num sentido mais amplo, a gramaticalização é a codificação de categorias cognitivas em formas lingüísticas, aí incluídas a percepção do mundo pelas diferentes culturas, o processamento da informação, etc.*” (Castilho, 1997a:31-32). Um exemplo bastante conhecido de gramaticalização é o processo pelo qual a forma lexical inglesa “*let us go*” foi reduzida para a construção simples “*let’s*”, a qual, posteriormente, se cristalizou em “*lets*”, como acontece, por exemplo, na frase “*lets you and me study together*”. Como se pode verificar, a forma perdeu o seu significado lexical de “permitir; deixar”, adquirindo a categoria de auxiliar, ao mesmo tempo que o pronome “*us*” se reduziu a um sufixo, primeiro, e a um mecanismo puramente fonémico, depois. Com base na hipótese da unidirecção (*Unidirectionality Hypothesis*), Hopper & Traugot (2003:7) propõem o seguinte modelo para definir os diferentes estágios da trajectória da gramaticalização:

item de conteúdo > item gramatical > clítico > afixo flexional

Os contextos sociohistóricos que são favoráveis à ocorrência de processos de pidginização/crioulização são-no, conseqüentemente, também propícios para que aconteçam reduções que potenciam posteriores gramaticalizações, uma vez que em situação de ASL não só a GU, que mantém os seus princípios acessíveis, actua na L2 através da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), permitindo transferências a partir desta (Thomason & Kaufman, 1988:73-74), como também os falantes adultos revelam falhas no restabelecimento dos parâmetros (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto as categorias funcionais deixarem de estar disponibilizadas depois do período crítico do desenvolvimento.

Em fases posteriores do desenvolvimento, e conseqüente nativização do código linguístico vigente na comunidade, as reduções deixam de preencher os requisitos comunicativos impostos pelo processo de socialização, pelo que as novas gerações, normalmente bilingues, buscam nos itens do acervo lexical disponível (alguns dos quais

comprometidos por transferências resultantes, por exemplo, das limitações perceptuais dos adultos da primeira geração da ASL) aqueles que possibilitam papéis funcionais intralinguísticos, por forma a suprirem as necessidades interlinguísticas que emergem. Assim sendo, verifica-se que, na recomposição gramatical durante o processo de socialização/nativização de uma língua, intervêm dois fenómenos:

- (i) o da ampliação das funções dos itens gramaticais já existentes na fala da comunidade;
- (ii) o recurso aos itens lexicais para se expressarem relações gramaticais inexistentes, determinante no processo de gramaticalização da língua em emergência.

O sistema de marcação PL do santomense constitui um exemplo típico de gramaticalização conotado à influência do substrato, sendo efectuado com recurso ao item *inen*, do substrato edo, que é colocado em posição pré-nominal para funcionar como marcador de flexão em número (exemplo [38b]). Este tipo de estrutura, para além de evidenciar paralelismo com o sistema linguístico kwa (exemplo [39]), divide igualmente características com construções similares de outros CP's atlânticos:

[216] ANGOLAR:  
*ené*                      *omé*    (Ferraz, 1979:61)  
3PL    homem  
“os homens”

Para Holm (1988/1989:29), a evidência deste tipo de transferência é tão óbvia que nenhum argumento foi ainda proposto no sentido de atribuir tal gramaticalização à influência dos universais linguísticos. No caso do PA, detectamos reminiscências deste tipo de gramaticalização, agora com recurso ao pronome pessoal do português, 3ª pessoa do PL “eles” (exemplo [48]), revelando, mais uma vez que, para a produção linguística dos falantes bilingues de Almojarife, apesar de a língua activa (o PA) colaborar no processamento da fala, as estruturas da língua em repouso (o santomense) se mantêm latentes (Green, 1986:219), provocando interferências na primeira.

Contudo, a propósito da relação entre animacidade e marcação de número, Alexandre & Hagemeyer (2007:42) chamam a atenção para, no santomense, a sequência *inen+N*

apenas se realizar com nomes de traço [+humano], expressando um significado definido que diz respeito a um grupo não-individualizado familiar ao ouvinte e ao falante (ponto 1.3.2.2.1 do presente trabalho). Nesta conformidade, o pronome pessoal *eles* poderá ser entendido, em alternativa, como uma forma desviada do objecto indirecto da frase (“*outra pessoa foi, tirou-lhes a roupa e viu a arma lá*”), já que, como foi visto a propósito das relações sintagmáticas e oracionais patentes na fala de Almojarife, é frequente os almojarifanos não fazerem a distinção entre formas rectas e átonas do pronome pessoal, com tendência para uso das primeiras em detrimento das segundas (ponto 1.3.2.3.1 – exemplos [147] e [148]). Contudo, e a confirmar-se esta possibilidade, a realização em observação apresentar-se-ia descontextualizada na situação em que foi produzida, tornando-se completamente desprovida de conteúdo lógico-semântico.

Outra possibilidade seria a de o falante estar a usar um possessivo e não um pronome pessoal (“*outra pessoa foi, tirou a roupa deles e viu a arma lá*”). Porém, importa referir que o mesmo informante faz um uso correcto do possessivo em outras construções (exemplo [217]), pelo que não é provável que se esteja na presença deste item gramatical:

[217] PA:  
*puxemo laço assim... rabo dele fica assim... [ZECAH1]*  
“puxámos o laço assim... o rabo dele ficou assim...”

Paralelamente, e como se verá no capítulo 4, dedicado à análise das variáveis independentes do tipo estrutural, as formas das L1’s africanas exercem, conjuntamente com as do crioulo, importante influência no ponto de entrada do item com morfema de PL que origina a configuração do SN PL do PA. Deste modo, o que se verifica é uma tendência para a entrada da marcação ocorrer categoricamente na posição imediatamente anterior ao núcleo do SN, o DET, levando ao consequente apagamento das marcas de PL nos outros itens da cadeia sintagmática (*tirô eles rôpa*), em virtude de elas se revelarem redundantes (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75). Veja-se ainda que a forma “*tirou a roupa deles*” implicaria o uso de um nome massivo, categoria morfológica afectada pela redução gramatical que se verifica nos estágios iniciais de ASL por adultos e que, no caso da totalidade dos dados do nosso *corpus*, iria constituir a única ocorrência registada deste tipo de elemento gramatical.

Os exemplos [38b] e [48] evidenciam também como itens e formas do substrato se conjugam com os do superstrato no processo de socialização/nativização das L1's que emergem a partir de DLP's fornecidos por L2's defectivas, estando tal conjugação na base de uma heterogénese orientada por princípios universais e transferências dos substratos, tudo com vista a possibilitar uma comunicação efectiva. Assim, e embora o item gramaticalizado sofra uma alteração a nível meramente tradutológico, respondendo à pressão social que leva a que algumas formas sejam progressivamente abandonadas e substituídas por outras da LA (Thomason & Kaufman, 1988; Lefebvre, 2003; Montrul, 2006; Lefebvre, White & Jourdan, 2006b; Siegel, 2008), a verdade é que o sistema estrutural de marcação PL, por partilhar um traço de marcação estrutural com o do santomense, se mantém conforme o padrão das ancestrais L1's, quiçá por fossilização geracional. Isto mesmo parece comprová-lo o facto de a construção do exemplo [48] ter sido produzida por um falante da geração menos idosa de Almocharife e com um nível de escolarização que classificámos de médio na variável independente *escolaridade*.

Estas questões levam então a considerar que a reestruturação da variedade que emerge por transmissão linguística irregular ocorre em dois planos distintos, mas que confluem para a sua elaboração: por um lado, e a um nível mais profundo, actuando os dispositivos da faculdade da linguagem regulando estruturas das línguas do substrato e da LA que, conforme as diferentes situações sociohistóricas, vão conceder os distintos DLP's para accionarem a nativização da língua; e por outro lado, a um patamar mais superficial, acontecendo a escolha dos elementos lexicais que vão desempenhar as funções gramaticais requeridas pelas referidas estruturas.

Centrando a atenção no primeiro aspecto, o *bioprograma* de Bickerton (1981, 1983, 1984a, 1988, 1999), partindo da perspectiva avançada por Borer (1984) de que os parâmetros sintácticos são fixados a nível do léxico, apresenta uma distinção entre o que acontece na aquisição das L1's em situação que não configura transmissão linguística irregular e o que sucede no processo de nativização dos crioulos. Assim, enquanto às crianças da primeira é facultado todo um manancial de itens gramaticais que lhes vai permitir responderem às solicitações definidas pela sintaxe inata, às crianças do segundo não são disponibilizados tais itens gramaticais, o que as leva a terem de os criar a partir

dos elementos lexicais de que dispõem, os quais são, assim, alvo de erosão na sua função inicial: a de conceder o significado.

A gramaticalização da forma lexical inglesa "*let us go*" em "*lets*" revela que este é um fenómeno que actua também ao nível da evolução histórica das línguas que não emergem por contacto. Com efeito, todas as línguas são afectadas por diversos fenómenos no seu processo de desenvolvimento natural, que não só mostram que elas são entidades dinâmicas que apresentam variações ao longo do tempo como evidenciam também que os falantes não reagem todos do mesmo modo face às mudanças que decorrem nos seus idiomas. Se uma determinada forma linguística parece estar a ser substituída por outra, seja ao longo do tempo seja numa determinada dimensão socioeconómica ou demográfica da comunidade, então poderemos estar face a um processo de mudança.<sup>107</sup> Um dos fenómenos de gramaticalização que afectou o português e continua activo nas línguas adquiridas por contacto tem a ver com a gramaticalização do demonstrativo, que fez com que este passasse a deter as funções de artigo. Na passagem do latim para o português, o demonstrativo latino *ille/illa/illud*, com função deíctica/fórica ou enfática (positivo) (Andrade, 1999:161), forma acusativa, originou o artigo definido. No *papia kristang* também não existe este artigo, podendo a sua função ser desempenhada pelos demonstrativos *isi* ou *isti* (este) e *akeli* (aquele) (Hancock, 1969; Baxter, 1988). Por seu lado, o fanakalo<sup>108</sup>, um pidgin usado como língua franca, principalmente nas minas de ouro, diamante e carvão das zonas industriais mineiras da África do Sul, recorre ao item zulu *lo*, que faz as funções de demonstrativo, de artigo definido e de artigo indefinido (Ferraz, 1979:74). O uso deste último demonstrativo estabelece um claro paralelismo com o do MSPEC *se*, do santomense, que chegou a ser apontado como detendo função de artigo definido (Ferraz, 1979; Lucchesi, 1993):

[218] FORRO:  
*kanwa se* (Ferraz, 1979:74)  
canoa MSPEC  
“esta/aquela/a canoa”

A posterior socialização da língua e o prolongado contacto do PAL1 com o PE iriam determinar a aquisição do artigo definido da LA, originando-se formas híbridas do demonstrativo/artigo, conforme se pode constatar no exemplo seguinte:

[219] PA:  
*táqui*                    *issos*                    *home* [ZECAH1]  
está+LOC            DEM                    homem  
“estão aqui esses/os homens”

A diferença que ocorre a nível da gramaticalização entre as línguas que emergem em situação de contacto massivo e as que emergem sem ser por contacto massivo tem a ver com o ritmo de evolução e extensão do processo de variação, muito mais rápido no caso dos idiomas submetidos a contacto horizontal, um aspecto que alguns defensores da mudança evolucionista atribuem à característica poliplóide destas línguas (Mufwene, 2001:151-152). De facto, enquanto a gramaticalização de construções como as que referimos no inglês ou no português têm de superar estruturas fixas e normatizadas, a fim de se fixar a mudança, no caso das línguas emergentes por contacto, a erosão estrutural vai possibilitar uma reestruturação gramatical extensa e profunda em período relativamente curto, isto é, no tempo de uma ou duas gerações. Deste modo, e tal como defende Roberts (1997:23), o carácter criouliizante de uma determinada língua não se define em consequência das suas características estruturais, mas tão só em função do processo sociohistórico que determina a sua emergência. Aliás, nem de outra forma poderia ser, uma vez que, segundo o autor, a aquisição das línguas crioulas é determinada pela faculdade humana da linguagem, sendo as suas características estruturais e funcionais geneticamente codificadas da mesma forma que acontece com todas as outras línguas transmitidas geracionalmente.

### **2.11.3.2. Variedades criouliizantes**

O pressuposto de que os parâmetros sintácticos são fixados a nível do léxico (Borer, 1984; Bickerton, 1999) levantou também a hipótese de as línguas crioulas virem geneticamente programadas com parâmetros não marcados, princípio que Roberts (1997:12-13) considerou válido, mesmo nos casos em que a LA lexificadora é determinada por parâmetros marcados. Segundo o autor, os parâmetros dependem dos valores dos traços abstractos dos núcleos funcionais, acontecendo a variação paramétrica



em virtude de cada núcleo poder activar movimentos de diversos tipos. Como se viu a propósito da fixação do valor de parâmetros no processo de aquisição (ponto 2.2.2.1 do presente trabalho), os movimentos são determinados pelos valores «+», originando, por sua vez, a sintaxe complexa. Como as circunstâncias que subjazem à génese das línguas crioulas originam não-aquisição de morfologia, isto é, sintaxe simplificada, os valores não marcados de parâmetros serão os que se apresentam por defeito nos DLP's que vão funcionar como estímulo para as primeiras gerações de falantes do crioulo. Nesta conformidade, os referidos DLP's vão apresentar-se defectivos em relação às propriedades responsáveis pelo desencadeamento dos parâmetros de valor «+», fazendo com que as línguas crioulas sejam marcadas, por exemplo, por ausência de movimento do verbo, pela realização do sujeito referencial ou pela ordem SVO.

Não obstante, estas características não são exclusivas das línguas crioulas, já que, por exemplo, o inglês também não tem movimento do verbo, não realiza sujeitos referenciais nulos e obedece à ordem SVO. Deste modo, não se pode determinar o carácter crioulo de uma língua apenas em função da sua estrutura linguística, havendo necessidade de recorrer, para tal, à sua sociohistória particular de aquisição em contacto massivo. Como esta é determinante no processo de não-aquisição de morfologia e de outros elementos da LA, a fase inicial da transmissão linguística irregular vai aliar estes fenómenos à reestruturação gramatical em processo. Neste ponto, é importante referir que, na transição do sistema de marcação fonológico proposto por Jakobson (1966 [1932]) para o de Chomsky (1981b), o paradigma dos traços marcados perdeu o seu carácter restritivamente binário, sendo possível a sua multivalorização (Battistella, 1996:132). Neste aspecto, o pressuposto bickertoniano vai determinar que não existirá limite *a priori* para o número e tipo de instanciações não-defeito. Como tal, se o traço [+vozeado] contrasta unicamente com o [-vozeado], o mesmo não sucede, por exemplo, com o caso nominativo (um parâmetro não-marcado), que contrasta não-unicamente com todos os casos marcados, mas que contam como uma unidade no que concerne à marcação. Por outro lado, se levarmos em conta, por exemplo, que as construções anafóricas representam formas marcadas (Chomsky, 1981c) em virtude de caírem fora do escopo do SN (p.e. [*os livros*] uns dos outros), este aspecto, por si só, não nos revela qual será a construção não-marcada que expressa a mesma ideia. As possibilidades são

imensas, podendo recorrer-se a uma frase que seja semântica, mas não sintacticamente uma anáfora recíproca (exemplo [220]), aplicar-se uma forma que explicita a questão (exemplo [221]), ou utilizar-se uma construção ambígua (exemplo [222]):

[220] PTG:  
*Eles leram os livros um do outro.*

[221] PTG:  
*Cada um leu o(s) livro(s) do(s) outro(s).*

[222] CGB:  
*Kada un lei libru di utru.* [Kihm, 2008:414]  
“Cada um leu o livro do outro”

O último exemplo levanta uma questão: como em situação de socialização/nativização de uma L2 adquirida de modo imperfeito as crianças se encontram expostas a DLP's que podem não reflectir os parâmetros instanciados por defeito pelo bioprograma não-adulterado, as opções marcadas são limitadas em número, o que as leva a estabelecerem intrinsecamente uma relação finita com a opção não-marcada. Neste contraste de  $n_{marc} = 1$  vs.  $\{marc\}$ ,  $\{marc\}$ , apenas um conjunto limitado de formas permite originar o parâmetro  $n_{marc}$ , e vice-versa, isto é, pode determinar um “*sistema binário ampliado*” (Kihm, 2008:414). Este aspecto irá então ser fundamental na definição das características das línguas que emergem por transmissão linguística irregular, enquanto fenómenos altamente variáveis que conjugam factores psicolinguísticos e sociohistóricos na sua génese. Como tal, é possível avançar, desde já, com algumas particularidades que distinguem este tipo de línguas:

- (i) Não-aquisição ou variação a nível da morfologia flexional e de determinados itens gramaticais;
- (ii) Gramaticalização de itens lexicais para compensar lacunas do código linguístico;
- (iii) Alteração dos valores dos parâmetros sintácticos em função de valores não-marcados por defeito, já que estes não implicam a realização de movimentos aparentes (*covert movements*)<sup>109</sup> na cadeia frásica.

### 2.11.3.3. Transmissão linguística irregular e variação: gramáticas em competição

A emergência de pidgins e crioulos deve-se, em grande parte, à conjugação de duas situações: pouco acesso à LA no período inicial do contacto massivo entre línguas tipologicamente distintas e, posteriormente, durante a fase de socialização/nativização da interlíngua, uma fraca influência exercida na reestruturação gramatical desta pela LA. Em sequência de tal, os fenómenos que afectam a morfossintaxe dos crioulos são, regra geral, mais radicais e profundos do que aqueles que se registam nos dialectos resultantes de transmissão linguística irregular. Por norma, a afectação traduz-se em não-aquisição elevada – ou total – não só da morfologia flexional a nível do nome e do verbo mas também das regras de concordância a eles associadas, que praticamente não transitaram para determinados crioulos, como acontece com o santomense. Este dialecto, a exemplo de outros CP's atlânticos, apresenta, então, uma acentuada alteração nos seus parâmetros sintácticos, reflectida em gramaticalização e/ou transferências das línguas dos substratos do grupo níger-congo atlântico, como o demonstra o desenvolvimento de um sistema de partículas de tempo assente na reestruturação de advérbios e determinadas formas verbais, que se gramaticalizaram como marcadores das categorias gramaticais do verbo. Por seu lado, o valor marcado de parâmetro de sujeito nulo da LA foi alterado, adquirindo a função de valor não-marcado no santomense e nos outros CP's atlânticos. Como tal, uma das características que distingue estes CP's da sua LA é a obrigatoriedade de se realizar o sujeito referencial, excepto em determinados casos, como acontece com os sujeitos expletivos:

[223] CCV:  
*Sta faze calor oji.* (Baptista, 2002:254)  
PROG fazer calor hoje  
Hoje está calor.”

Mas se o santomense, enquanto basilecto, configura uma situação de afectação morfossintáctica profunda, já outras línguas – ou variedades de uma mesma língua – tidas como crioulas, ou sejam, os acrolectos e os mesolectos, apresentam diferenciados graus de reestruturação gramatical, como acontece com os crioulos do Barlavento e do Sotavento cabo-verdianos. Tal diferenciação é consequência do grau de intensidade do contacto entre línguas, seja na fase inicial deste seja no *continuum* pós-crioulo, que leva,

consequentemente, a um maior ou menor índice de reestruturação. Este, por sua vez, pode mesmo traduzir-se em descrioulização, isto é, na substituição de estruturas marcadamente crioulas por formas da LA.

Já as situações de transmissão linguística irregular configuram fenómenos de reestruturação morfossintáctica mais leves, centrando-se a variação mais a nível das formas de uso dos elementos gramaticais e não tanto no plano das alterações da estrutura linguística. Assim, na fase inicial do contacto, ao não atingir o grau de erosão morfológica registada no caso dos pidgns, a variação em situações de transmissão linguística irregular verifica-se na aplicação dos morfemas flexionais e uso dos elementos gramaticais. Este aspecto torna nubloso o tipo de estímulos responsáveis pela marcação dos valores dos parâmetros sintácticos, que não podem ser fixados na ausência de DLP's processáveis (Kihm, 2008:415), originando um acentuado grau de variação a nível do uso colectivo da língua, sem que se registre, nesta fase da emergência da nova língua, uma elevada quantidade de dispositivos originais relativamente à gramática da LA. Paralelamente, as lacunas são suprimidas com dispositivos da LA (Thomason & Kaufman, 1988; Lefebvre, 2003; Montrul, 2006; Lefebvre, White & Jourdan, 2006b; Siegel, 2008), originando-se um sistema de variação binária entre presença e ausência de mecanismos gramaticais da LA, uma vez que convivem na fala da comunidade duas ou mais gramáticas geracionais, isto é, diferentes valores de parâmetros sintácticos.

No caso da aquisição do PtgL1 por parte dos descendentes almorarifanos, deu-se com base em DLP's compostos por dados fragmentários e defectivos do PtgL2 adquirido pelos seu progenitores, configurando uma situação de transmissão linguística irregular que pode ser esquematizada da seguinte forma:

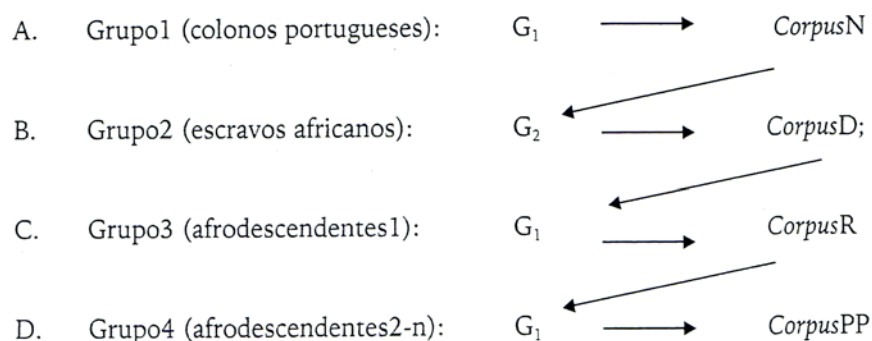


Fig. 28. Inglês: Modelo de aquisição para situação de contacto entre línguas (Lucchesi & Ribeiro, 2009:145).

Em S. Tomé, a relação entre os colonos portugueses (Grupo1) e os escravos adultos africanos (Grupo2) determinou a aquisição defectiva de uma L2 pelos segundos, com base nos dados fornecidos pelos primeiros (*CorpusN*). Esta variante é caracterizada quer por simplificações a nível dos mecanismos gramaticais mais abstractos e sem valor referencial quer por especificidades decorrentes de transferências e reanálises orientadas pelas gramáticas das L1 africanas (*CorpusD*). Posteriormente, a partir deste modelo de PtgL2 defectivo, os afrodescendentes do Grupo2 adquiriram o Ptg L1 característico de Almojarife. Este incorpora tanto as simplificações do *CorpusD* como transferências do substrato, que materializam eventuais fossilizações, sejam elas temporárias sejam permanentes. Por fim, ocorreu uma situação de transmissão linguística geracional entre os segmentos afrodescendentes (Grupo3 e Grupo4), com base no Ptg nativizado e reestruturado pelos afrodescendentes1 (*CorpusR*). Assim, a fala produzida pelos afrodescendentes2-n (*CorpusPP*) tendeu e tende a diminuir as interferências iniciais resultantes do contacto entre línguas, ao incorporar, continuamente, estruturas gramaticais do português falado tanto pelos outrora colonos portugueses como pelas actuais classes elitistas santomenses, que fizeram ou continuam fazendo a sua formação académica em Portugal.

Como se constata, aconteceu em São Tomé (e Almojarife) uma relação de implicação entre os diversos níveis. Daí resulta que a maior ou menor reestruturação na gramática do Grupo3 (afrodescendentes de primeira geração) será directamente proporcional à maior ou menor simplificação/contaminação produzida no *input* que configura o *CorpusD*. A partir de então, “o processo de transmissão da língua da segunda geração de afrodescendentes em diante se assemelhará mais a um processo de descrioulização” (Lucchesi & Ribeiro, 2009:146). Contudo, tal processo descrioulização não terá encontrado via livre para se poder consumir plenamente. Com efeito, a relação A-B-C repetiu-se ciclicamente na ilha, já que a chegada de africanos foi ininterrupta, primeiro traduzida nas levas de escravos em trânsito para outros continentes, com alguns emprestando temporariamente mão-de-obra aos engenhos do açúcar, e, posteriormente, reflectida nos contingentes de contratados para trabalharem nas roças do cacau e do café, muitos deles falando um PtgL2 de características idênticas ao do *CorpusD*.

Verifica-se então que, nos casos de transmissão linguística irregular, se observam duas etapas fundamentais que concorrem para a reestruturação da língua emergente: (i) o período inicial do contacto, marcado por variação mais, ou menos intensa, no uso dos mecanismos gramaticais, já que, nesta fase, a ausência de DLP's processáveis impede a fixação de parâmetros sintácticos; (ii) as fases seguintes de reestruturação linguística, com variação nas frequências de uso dos parâmetros sintácticos, fixados a partir do parâmetro da não-marcação fornecido por defeito às crianças e das pistas detectadas por estas nos DLP's defectivos, que as forçam a modificar o parâmetro por defeito em determinada direcção (Lightfoot, 1999a:439). Assim sendo, durante a primeira fase ocorre uma competição de formas gramaticais de todo o tipo, com os falantes a seleccionarem as que melhor respondem às necessidades de comunicação, cujo uso passa a ser generalizado. Seguidamente, as formas adoptadas acabam por se configurar como estímulos nublosos, isto é, como DLP's de parâmetros indeterminados (Kegl, Senghas & Coppola, 1999:2006), que facilitam diferenciados valores para a fixação dos parâmetros sintácticos que acontece nas fases posteriores de aquisição. Nestas, o nivelamento das formas nuelas pode mesmo levar à diluição diacrónica de algumas protoformas (Mufwene, 1999, 2001; Siegel, 2008), já que estas não conseguem resistir à pressão exercida quer pelos processos de gramaticalização determinados pela heterogeneidade da fala da comunidade quer pelas formas gramaticais que se cimentam por pressão da LA.

Nas variedades reestruturadas por transmissão linguística irregular, caso do PA – e variedades populares do PB –, são comuns, então, os seguintes quadros gramaticais:

- (i) Não-realização de dispositivos gramaticais marcadamente abstractos e de uso restrito na LA. É o que sucede, por exemplo, quer com o apagamento de marcas morfológicas de número, que afectam a concordância a nível do SN (sobretudo do SDET – exemplos [227a] e [227b]) e do SV (exemplo [224]), quer com a supressão do modo conjuntivo (exemplo [225]):

[224] PA:  
*vinte cinco de Abril nós ficou aqui muito abalado.* [SALVH3]  
“em vinte e cinco de Abril nós ficámos aqui muito abalados”

[225] PA:  
... *talvés esse semana... essa semana ele ão veio*. [MANOH3]  
“talvez nessa semana... nessa semana ele não tenha vindo”

(ii) Recomposição da estrutura gramatical da L1, ao encontro da LA, reduzindo-se o grau de variação de determinadas formas, como sucedeu com a marcação de género. Esta terá sido reintroduzida no PA (Figueiredo, 2009a) e na grande maioria das variedades populares do PB, mantendo-se a sua variação em maior escala apenas nas comunidades que conheceram um profundo processo de transmissão linguística irregular. A este propósito, veremos no ponto 4.1.2 do presente trabalho que a variável independente marcação de género não foi considerada relevante para observação do fenómeno da CPL-var que ocorre no SN do PA. Por outro lado, veja-se também como, em alguns casos, a pressão exercida pela LA conduz à eliminação da variação de determinadas formas, como sucede em alguns registos do falante [CASTEH1], detentor do 8º ano de escolaridade, apontados nos exemplos [156] e [157].

(iii) Conservação da variação no esquema presença/ausência do mecanismo gramatical da LA, como acontece a nível da concordância nominal de número (exemplos [190] e [191]) ou das marcas de pessoa-número do SV (exemplo [264]):

[226] PA:  
*nós fomos saber qual é lotes que nós vão receber*. [ABILH2]  
“nós fomos saber quais eram os lotes que nós íamos receber”

(iv) Manutenção do tipo ternário no uso da variação gramatical, isto é, de formas da LA (exemplo [227a]), da variante proveniente do processo de reestruturação da gramática (exemplo [227b]) e da variante zero (exemplo [227c]):

[227] PA:  
a. *todos os dias si ê nô i, ê perdo dia* [ANTOM1]  
“todos os dias que eu não for, eu perco o dia”  
b. *Correu com todos branco na roça*. [MINISM3]  
“Correu com todos os brancos da roça”  
c. *são coesa acontecida* [CASTEH1]  
“são coisas (acontecidas)/ são coisas que acontecem”

- (v) Alteração em várias estruturas típicas da LA, como é o caso das construções relativas (exemplo [160]), ou que envolvem partículas gramaticais como o artigo (exemplos [124] e [125]), as preposições (exemplos [141] a [144] e os clíticos (exemplos [147] a [151]);
- (vi) Alteração nas frequências de uso em relação à marcação de determinados parâmetros sintáticos, como sucede com a redução na aplicação do sujeito nulo (Duarte, 1993) (exemplo [224]), ou a não-realização de movimento do verbo (Morais, 1993; Rossi, 1993) (exemplo [228]), já que a ordem das palavras na estrutura frásica não tem consequências para a interpretação. As construções que configuram estes dois tipos de estruturas podem ser observados sob a óptica de uma expansão da gramática crioulezante, já que a marcação de valores consistentes para esses padrões sintáticos é fortemente generalizada no PE:

[228] PA:

*Depois eu sempre passávamo pá fila das pissoa* [OSVALH1]

A propósito desta realização, importa dizer que a concatenação da estrutura (*merge*) é o requisito mínimo para ocorrer sintaxe, pelo que não representa um parâmetro marcado. Em contrapartida, o movimento é a operação marcada, o que faz sentido do ponto de vista minimalista de uma economia da derivação e também se considerarmos a deslocação como uma “imperfeição” (*imperfection*) do *design* da língua (Chomsky, 1998).<sup>110</sup> Como a flexão implica comportamento sintático, surge envolvida no movimento do verbo. Desta forma, a não-aquisição de morfologia flexional determinada pela transmissão linguística irregular pode originar variação devido à não-movimentação do verbo para domínios funcionais mais altos.<sup>111</sup> A ausência de movimento aparente do verbo de “V-para-I” sucede também na língua inglesa (Pollock, 1989), que revela concatenação, mas cuja pobreza flexional origina a ordem ADV>V>O, registada nos exemplos [229a] e [229c], obtendo-se o diagrama da Fig. 29. No PE, contudo, é possível “movimento do verbo de V-para-I”, devido à elaborada flexão gramatical da língua (exemplo [230]), situação que acontece também no francês (exemplo [231]), podendo realizar-se a ordem V>ADV>O em ambas as línguas (Fig. 30).<sup>112</sup>



- [229] INGLÊS:  
 a. *John often kisses Mary.*  
 b. \**John kisses often Mary.*  
 c. *John has often kissed Mary.*

[230] PE:  
*O João beija frequentemente a Maria.*

[231] FRANCÊS:  
*Jean embrasse souvent Marie.*

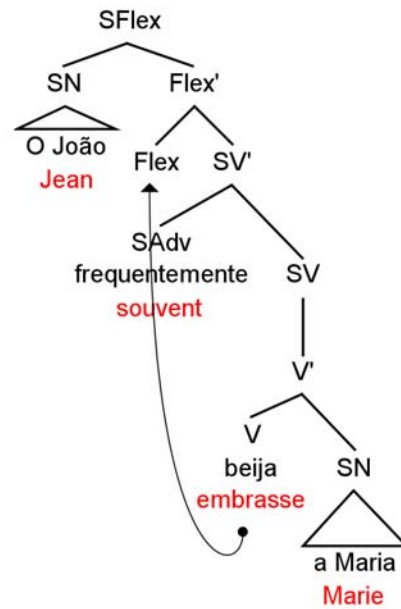
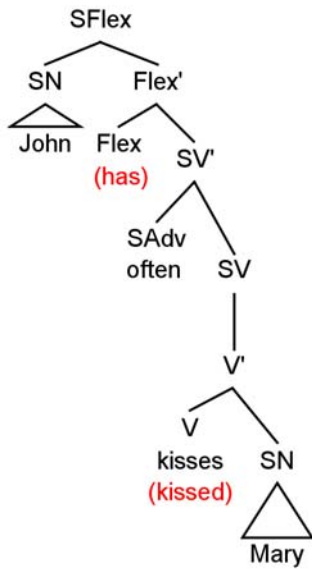


Fig. 29. Inglês: ausência de Mov-V.  
 (a partir de Silva, 2010:15)

Fig. 30. Português e francês: Mov-V-para-I.  
 (a partir de Silva, 2010:15)

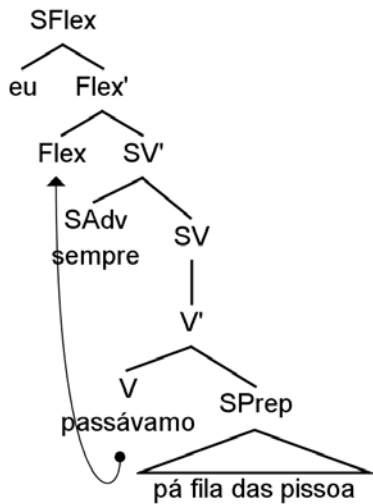


Fig. 31. PA: Mov-V-para-I.

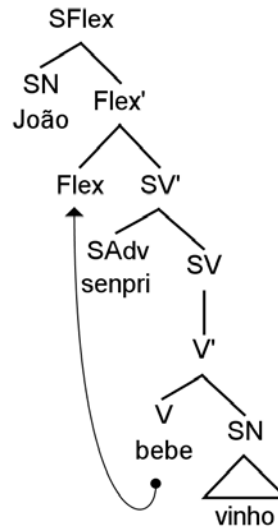


Fig. 32. CCV: Mov-V-para-I.

Todavia, note-se que a discussão em torno da questão do movimento do verbo não tem sido consensual na literatura actual. Analisando a posição de alguns advérbios do crioulo de base francesa do Haiti, entre os quais o advérbio “sempre”, DeGraff (1997:68-70) apresenta paralelismos entre esta língua e o inglês, e que a fazem divergir do idioma lexificador. A conclusão é que, no haitiano, tal como no inglês e ao contrário do francês, em virtude de ser uma língua de flexão “pobre”, o verbo apresenta ausência de “movimento de V-para-I”, isto é, instancia-se o parâmetro não-marcado que estabelece a força dos traços do verbo. Assim, neste tipo de línguas, o advérbio parece constituir uma barreira intransponível para a subida do verbo para a flexão, levando este a ganhar traços numa posição mais baixa da estrutura, isto é, na posição original.

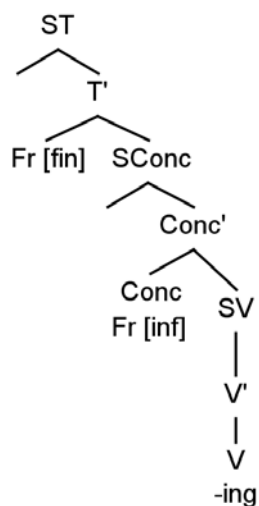


Fig. 33. Configurações de Mov-V para diferentes línguas (Silva, 2010:15).

Quanto à questão do movimento do verbo em francês, pode ser de dois tipos: (i) um movimento mais curto, para os casos em que vai aparecer no infinitivo (p.e. *ne pas manger*); (ii) movimento até ao núcleo do sintagma de tempo, para ganhar os traços de flexão relativos a T (p.e. *Jean embrasse souvent Marie*) (Fig. 30). Em português, o que distingue a colocação do ADV<sup>113</sup> é que sofre um movimento curto, semelhante ao movimento que ocorre nos verbos franceses no infinitivo (exemplo [233]). Assim, o movimento do verbo não é opcional em português, devido às restrições impostas por determinados advérbios, como os advérbios de negação, que surgem sempre antes do verbo. Por outro lado, este idioma regista também advérbios “baixos”, que surgem ligados ao sintagma. A não aquisição destas regras pelos falantes do PtgL2 vai

determinar conflito entre gramáticas, que origina a variação na colocação dos advérbios (exemplo [228]).

O estabelecimento de paralelismos com os estudos do movimento do verbo no inglês levados a cabo por Pollock (1989) e a tendência de DeGraff (1997) em generalizar as conclusões a todos os crioulos no que concerne à questão de estes representarem línguas não-marcadas tem que ser observada com alguma reserva, já que o primeiro não analisou o comportamento verbal no PE, que apresenta complexidade relativamente à posição do advérbio.<sup>114</sup> Efectivamente, apesar de os verbos principais se movimentarem nesta língua, não há evidência consistente de que se desloque sempre para FLEX (Costa, 1996:27), como o demonstram os exemplos seguintes com o advérbio “sempre”:

[232] PE:  
*O João vai sempre com a Maria ao cinema.*

[233] PE:  
– *O João sempre vai com a Maria ao cinema hoje?*  
– *Sim, sim. Ele sempre vai com ela.*

Como se pode verificar, embora o advérbio de frequência surja pós-verbalmente no exemplo [232], também pode ser colocado pré-verbalmente em interrogativas dubitativas, e em respostas às mesmas perguntas, se estas esclarecerem assertivamente a dúvida inicial (exemplos [233]). Verifica-se então que a posição do advérbio “sempre” em relação ao verbo não é regida por questões sintácticas, mas sim semânticas. De facto, no primeiro exemplo a sua semântica associa-se ao tempo verbal, enquanto no segundo caso se relaciona com o aspecto do verbo. Idênticas situações acontecem no CCV, no qual a posição do advérbio *senpri* (sempre) também varia, podendo ocorrer em vários lugares da cadeia frásica (Baptista, 2002:128):

[234] CCV:  
a. *Senpri João bebe vinhu.*  
Sempre João bebe vinho  
“O João bebe sempre vinho.”  
b. *João senpri bebe vinhu.*  
c. *João bebe senpri vinhu.*  
d. *João bebe vinhu senpri.*

Este aspecto vem evidenciar que o CCV exhibe marcação exactamente onde o postulado de Roberts (1997:12-13) aponta para que tal não aconteça.<sup>115</sup> Por outro lado, os exemplos [234a] a [234d] do CCV, bem como outros do PE (exemplos [235a] a [235i]) e do francês (exemplos [236a] a [236g]), demonstram que as posições do verbo e do advérbio não podem ser vistas nestas línguas como mero resultado de uma deslocação do primeiro para uma posição acima do segundo, a fim de preencher os requisitos de mapeamento:

[235] PE:

- a. *Frequentemente o João beija a Maria.*
- b. *O João frequentemente beija a Maria.*
- c. *O João beija frequentemente a Maria.*
- d. *O João beija a Maria frequentemente.*
- e. *Frequentemente o João tem beijado a Maria.*
- f. *O João frequentemente tem beijado a Maria.*
- g. *O João tem frequentemente beijado a Maria.*
- h. *O João tem beijado frequentemente a Maria.*
- i. *O João tem beijado a Maria frequentemente.*

[236] FRANCÊS:

- a. *Jean embrasse souvent Marie.*
- b. *Jean souvent embrasse Marie.*
- c. *Souvent Jean a embrassé Marie.*
- d. *Jean souvent a embrassé Marie.*
- e. *Jean a souvent embrassé Marie.*
- f. *Jean a embrassé souvent Marie.*
- g. *Jean a embrassé Marie souvent.*

(Kihm, 2008:423)

Tendo ainda em conta a questão dos traços do verbo, verificámos que este tanto pode aparecer marcado (*João bebe senpri vinhu*) como não-marcado (*João senpri bebe vinhu*) no CVV, em clara consonância com aquilo que sucede na sua LA. Ora, estes aspectos são também válidos para o PA (exemplos [228] e [237], estabelecendo-se, assim, um claro paralelismo entre alguns CP's, variedades reestruturadas de português e o próprio PE no que concerne à questão dos traços do verbo e seu "movimento de V-para-I":

[237] PA:

*Por acaso assistia sempre, todos os dia eu assistia televisão.*

[OSVALH1]

Considerando não só que a posição relativa dos marcadores de TMA dos crioulos, dos verbos e dos advérbios dos crioulos, línguas reestruturadas e línguas lexificadoras dependem da semântica dos advérbios e não de mecanismos sintácticos cegos (Kihm, 2008:422), a presença ou ausência do movimento do verbo para ajudar a sustentar que os crioulos são línguas sintacticamente não marcadas (Roberts, 1997:12-13) têm que ser vistas com alguma reserva, já que assentam em bases pouco consistentes. Por outro lado, a complexidade acerca do posicionamento do advérbio – e da possibilidade ou não de realização do sujeito referencial – mostra que a questão também não pode ser vista em termos de traços de marcação, uma vez que não é possível efectuar tal observação a uma escala meramente binária. Assim sendo, contempla-se o pressuposto de Muysken (1981b:415) de que línguas marcadas serão aquelas cuja interpretação conduz a uma falta de paralelismo na complexidade entre a sintaxe e a FL, que é exactamente o que sucede no caso do CCV, no PA, no PE, no crioulo do Haiti ou no francês. O que deixa em alternativa a questão de a complexidade entre a sintaxe e a FL representar o parâmetro por defeito, o qual é aplicado pela criança, a não ser que esta obtenha evidências para não o instanciar. Sob esta perspectiva, certas particularidades das línguas crioulas, como a colocação do advérbio na estrutura frásica, o uso de partículas de negação, o sistema de partículas pré-verbais de TMA, a ordem SVO, ou a ausência referencial de sujeitos nulos, têm que ser vistas, não como características de línguas não-marcadas, mas sim como instanciações inatas de parâmetros por defeito dessas mesmas propriedades, que podem originar uma série indefinida de instanciações não-defeito a adquirir pela criança, quando exposta a DLP's processáveis (Kihm, 2008:432). A este propósito, lembre-se que as formas não marcadas são as que se aplicam à língua como um todo, enquanto as formas marcadas são as que se aplicam a morfemas específicos.

No que respeita à afectação da morfologia de número no SDET, e consequente variação de concordância entre os seus elementos gramaticais, a criança mapeia dados a partir de várias possibilidades que são construídas na fala almoxarifana. Tomando como exemplo os SDET's que envolvem determinante+nome, a L1 nativizada a partir da L2 defectiva e fragmentária realiza, pelo menos nas primeiras fases da aquisição, as seguintes estruturas, ainda que o *input* não conceda evidências para a elaboração das mesmas:

1. DET<sub>[+PL]</sub> N<sub>[+PL]</sub> (exemplo [154]);
2. DET<sub>[+PL]</sub> N<sub>[-PL]</sub> (exemplo [227b]);
3. DET<sub>[-PL]</sub> N<sub>[+PL]</sub> (exemplo [199b]);
4. DET<sub>[-PL]</sub> N<sub>[-PL]</sub> (exemplo [227c]).

Posteriormente, os usos que não são consistentes com a LA poderão ir desaparecendo, uma vez que a *performance* se adequa às situações sociolinguísticas avaliadas como positivas, sendo estas determinadas, sobretudo, pela variante prestigiada. No caso concreto da comunidade de Almojarife, o prolongado isolamento restringiu o contacto com as formas prestigiadas do PE, mantendo geracionalmente a concorrência de gramáticas distintas e preservando o leque de formas flutuantes que constituem os DLP's para a aquisição/transmissão entre os níveis B-C-D (Fig. 28). Consequentemente, na fala almojarifana, as estruturas da gramática do português elitista continuam a conviver em concorrência com as estruturas típicas das situações de contacto e transmissão linguística irregular, isto é, com as formas resultantes dos processos de simplificação, transferência e reanálise. Por outras palavras: a variação registada no PA resulta da competição entre gramáticas e deve ser entendida como processos de alternância de código (*code switching processes*) (cf. Lucchesi & Ribeiro, 2009:149).

Este aspecto revela-se de primordial importância para entender a variação registada no SN da comunidade bilingue de Almojarife, uma vez que a alternância de códigos a nível do processamento inter e intrafrásico (Fig. 15) é regulada por mecanismos inerentes às situações de interacção verbal (ponto 2.3.1 do presente trabalho). De facto, alguns estudos têm mostrado que a alternância de códigos no processamento interfrásico é determinada por factores sociais e discursivos, enquanto a que atinge o processamento intrafrásico é condicionada por princípios sintácticos (Belazi, Rubin & Toribio, 1994:221). Nesta conformidade, a alternância de códigos, sobretudo a registada no interior da mesma sentença, terá implicação directa sobre o modo como o falante estrutura o seu conhecimento linguístico na altura de interagir verbalmente.

No caso concreto da alternância de códigos no interior da mesma frase, parece ocorrer “*uma tendência para predominância de uma gramática numa determinada porção da estrutura oracional e a predominância da outra gramática nos demais nódulos da*

*oração*” (Lucchesi & Ribeiro, 2009:149). A este propósito, convém lembrar igualmente que as transferências não se processam de modo caótico, antes necessitam de encontrar condições para serem processadas no sistema de desenvolvimento da L2 (Andersen, 1983c:182; Plag, 2008b:314) (ponto 2.10.2 do presente trabalho). Este aspecto alia-se aos anteriores no sentido de impedir que a fala almorarifana realize flexão contextual plena, quedando-se, como referimos no ponto 2.10.4.2, por produções que revelam apenas alguma flexão inerente e a posicionam entre os estágios T2 e T3 do desenvolvimento linguístico sugerido por Plag (2008b – Tabelas 2.1 e 2.3), os quais são caracterizados por realização de variação acentuada, mas estruturada. Assim, no plano variável do PA, será de considerar a tendência para a co-ocorrência de estruturas provenientes de uma mesma gramática numa mesma porção da estrutura da sentença nos processos de alternância de código (cf. Lucchesi & Ribeiro, 2009:149), configurando-se aqui as propostas avançadas pelo Princípio da Coesão Estrutural (Haiman, 1983:782).

## **2.12. Princípio da coesão estrutural e variação a nível do sintagma determinante**

Os casos de alternância de códigos são originados pela competição entre gramáticas alojadas na mente do falante, elaborando-se sentenças a partir das regras de uma das gramáticas e que se combinam com estruturas geradas pela outra gramática. Estas formas, apesar de reflectirem variação nos mais variados níveis da sentença, são orientadas e restringidas pelo Princípio da Coesão Estrutural (Haiman, 1983:782). Este determina, então, que as estruturas de uma mesma gramática apresentem propensão maior para co-ocorrerem nos nódulos mais coesos, isto é, naqueles que são internos a um único constituinte (p.e. número, género, pessoa), e não tanto nos nódulos mais periféricos, ou seja, nos que estabelecem uma relação de especificador/núcleo (p.e. entre um SDET sujeito e o núcleo da flexão sentencial). Assim, e no que concerne especificamente à estrutura sintagmática de um mesmo SDET, é possível prever, probabilisticamente, determinadas elaborações.

Tendo em conta, por exemplo, que a inserção de marca de género, por ocorrer primeiro (Müller, 1994a; Franseschina, 2002, 2005; Corrêa, Name & Ferrari-Netto, 2004; Montrul, 2004; White *et alii*, 2004), serve como gatilho para despoletar também a marcação de número, Lucchesi (2000a:264) constatou a aplicação do Princípio da

Coesão Estrutural (Haiman, 1983:782) no dialecto de HEL-Ba, uma vez que a marcação de género favorece categoricamente a marcação PL no constituinte analisado, enquanto a sua não-marcação restringe a inserção do número. Como nas línguas que configuram transmissão linguística irregular a concordância de número é mais afectada do que a concordância de género, é possível postular que, no PA, a regra de concordância plena de género poderá não ser inserida num SDET em que não se aplica plenamente a concordância de número (p.e. *as bananeira bem alto*), enquanto que, num SDET em que esta regra é inserida na totalidade, é improvável não acontecer concordância plena de género (p.e. *as bananeiras bem altos*). Portanto, o primeiro exemplo será comum em Almojarife, ao passo que o segundo exemplo encontra poucas possibilidades de ocorrer. Destes aspectos resulta que, se o falante inserir a marcação de número plenamente, também acontecerá, categoricamente, marcação total de género (p.e. *as bananeiras bem altas*), sendo esta igualmente possível caso não aconteça marcação plena de número (p.e. *as bananeira bem alta*).

Relativamente à questão dos nódulos que estabelecem uma relação de especificador/núcleo, abster-nos-emos de a comentar, já que cai fora do âmbito do presente trabalho, ao adequar-se mais a uma análise do tipo sintagmático, e não do género atomístico.

### **2.13. Conclusões parciais**

Face aos vários pressupostos que expusemos, verifica-se que, para se determinar se uma língua configura processo de transmissão linguística irregular, é necessário ter em consideração dois aspectos fundamentais: “*por um lado, o item da estrutura lingüística que está sendo focalizado, a sua função mais ou menos abstrata, sua transparência semântica, e o valor mais ou menos marcado do parâmetro sintático em jogo; tendo, sempre, como referência os modelos transplantados da variedade lingüística européia; e, por outro lado, a comunidade de fala em que esse item é analisado, o que implica considerar o grau de contato lingüístico por que essa comunidade passou, a sua exposição aos modelos lingüísticos difundidos institucionalmente, etc.*”<sup>116</sup> Observando-se a importância destes aspectos na determinação de uma variedade emergente em contexto de transmissão linguística irregular, estar-se-á, tal como



acontece no caso dos pidgins e crioulos, não só a perceber como a confluência de múltiplos factores, que vão desde os efeitos produzidos pelo contacto de línguas e da variação daí resultante, até ao modo como dispositivos cognitivos entram em jogo para a modelação das línguas, em geral, mas também a ter em conta que essas mesmas variantes encerram em si riquezas sociolinguísticas únicas, que necessitam de ser preservadas com recurso a planificações e políticas de línguas adequadas, em particular.

Por fim, e tendo em conta que o ponto de partida de todo o processo de transmissão linguística irregular é a não-realização, em maior ou menor grau, de morfologia flexional e das regras de concordância nominal e verbal da língua lexificadora durante a aquisição inicial da LA, verifica-se que esse é o quadro que se adequa ao PA, tal como se ajusta a variedades reestruturadas do PB. Por outro lado, como o substrato directo do PA (o santomense) configura um crioulo basilectal, marcado por irregularidade e drástica redução de complexidade morfológica, podemos assumir que, nas análises das comparações detalhadas que levaremos a cabo tendo em conta a cadeia sociohistórico-linguística estabelecida entre o PB (nomeadamente a fala do município do Rio de Janeiro, a norma culta de Salvador e o dialecto de HEL-Ba), o PA, o substrato directo deste (o santomense) e também os seus substratos ancestrais (línguas do grupo níger-congo atlântico), os conceitos de transmissão linguística irregular e crioulição/pidgnização, a par de outros, como os que giram em torno da aquisição e nativização de L1, aquisição bilingue, ASL (individual e massiva) e fossilização, ocuparão posição central para ajudarem a determinar as nossas conclusões acerca dos fenómenos que motivam a CPL-var no interior do SN do PA. Paralelamente, e tendo em conta a concepção da variação linguística dentro da estrutura gramatical, não deixaremos também de recorrer aos pressupostos inerentes ao Princípio da Coesão Estrutural (Haiman, 1983:782), observando-os no seu sentido mais fraco (o probabilístico) e mensurando-os em concomitância com os efeitos de outros factores linguísticos.

Assim sendo, no presente trabalho assumiremos, essencialmente, que a marcação PL no SN do PA depende do jogo estabelecido entre contextos linguísticos (língua-I) e extralinguísticos (língua-E), conjugando análises teóricas e quantitativas. Ainda assim, levaremos em consideração uma abordagem ecléctica, que tenha em consideração os pressupostos das diferentes metodologias propostas no presente capítulo. Face ao que se

referiu acerca da ASL e aquisição/nativização de L1's, a orientação do nosso trabalho inferirá que a GU está, de uma forma ou de outra, disponível para os adultos que aprendem uma L2, já que mantém os seus princípios activados. Paralelamente, teremos em consideração que, com a aquisição intermédia de morfologia flexional relevante e de itens lexicais, é possível restabelecer parâmetros para se atingir um estágio final de aquisição. Contudo, a variação que surge fossilizada em fase tardia pode também ser entendida como falha no restabelecimento de alguns parâmetros (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clahsen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto as categorias funcionais deixarem de estar disponibilizadas depois do período crítico do desenvolvimento. Baxter (2004, 2009), em constatações acerca da CPL-var do PT, defende que as estruturas das L1 africanas exercem, conjuntamente com as do crioulo, importante influência no ponto de entrada do item com morfema de PL que leva à configuração do SN PL do PtgL1, adquirido em situação de transmissão linguística irregular. Figueiredo (2008) avançou a hipótese de a CPL-var do PA, tal como acontece com a CPL-var do PT, estar sujeita a condicionantes que ancoram na aquisição defectiva do PtgL2 e PtgL1, em situações de contacto que envolvem estruturas de PtgL2 influenciadas pela línguas do grupo níger-congo atlântico. Como tal, para o presente trabalho, partiremos igualmente do pressuposto que a aquisição da L1 se inicia no estado FL0, enquanto que o estágio da Gramática *n* é o estado inicial para a ASL por parte de aprendentes em idade pós-adolescência. Contudo, a variação da interlíngua, até ao seu estágio final de sistematicidade, não será resultado exclusivo da refixação de parâmetros guiada pela L1, uma vez que é também consequência da interferência de outros factores, sejam eles cognitivos (maturidade, quantidade e qualidade do *input*, frequência de uso da língua ou conhecimento de outras línguas) sejam eles extralinguísticos (motivação, idade, sexo, escolaridade, aspectos culturais, contexto sociohistórico da aquisição, etc.).

No próximo capítulo daremos conta da metodologia a que recorreremos para constituição dos nossos dados, considerando quer factores já abordados em outros trabalhos, e que considerámos terem influência na marcação PL, quer ainda factores por nós constituídos, os quais nos permitirão ajudar a observar, minuciosamente, quais os elementos que motivam a pluralização no SN do PA. Com base em todos estes factores,

levantaremos hipóteses que poderão conceder não só uma visão mais abrangente sobre a CPL-var no SN do PA mas também ajudar a perceber quais as condicionantes que a determinam. As hipóteses serão posteriormente confirmadas, ou não, nos capítulos 4 e 5.

## CAPÍTULO 3

*“Não ande apenas pelo caminho já percorrido, pois ele conduz somente até onde outros estiveram antes.”*

(Alexandre Graham Bell)

### A Metodologia

No presente capítulo iremos descrever: (i) a metodologia aplicada na recolha dos dados que constituem o *corpus* do presente trabalho; (ii) a variável dependente e os grupos de factores independentes que constituímos para efectuar a análise do fenómeno da CPL-var no SN do PA; (iii) os procedimentos adoptados para se codificarem os dados; (iv) o suporte computacional utilizado para tratamento dos dados e sua análise estatística; (v) a caracterização dos informantes entrevistados para constituição do *corpus*.

#### 3.1. Aplicação dos métodos quantitativos no estudo da aquisição de línguas

A variação linguística produzida por falantes de determinadas comunidades tem sido observada com recurso a métodos matemáticos, nomeadamente aquele que permite analisar a variação segundo o paradigma de Labov (1972a), isto é, o método conhecido como “*the variable rule analysis*” (VARBRUL). Especialmente usado no estudo da ASL, seja em contexto informal (L. Dickerson, 1974; W. Dickerson, 1976; Adamson & Kovak, 1981; Adamson, 1988; Wolfram, 1989; Young, 1991; Young & Bayley, 1996; etc.) seja em contexto formal de ensino (Godinho, 2005, 2008; Martins, 2007; Silva, 2010), o pacote de programas VARBRUL, em virtude de calcular as probabilidades dos factores linguísticos e extralinguísticos que motivam a variação linguística, tem-se revelado um precioso auxiliar no estudo da aquisição. Paralelamente, a ele recorreram também pesquisadores que analisaram a variação da fala quer de comunidades linguísticas urbanas (Cedergren, 1973; Sankoff & Sankoff, 1973; Naro & Lemle, 1976; Braga, 1977; Lemle & Naro, 1977; Poplack, 1980a; Guy, 1981a; Scherre, 1988; Fernandes, 1996; L. Andrade, 2003; Lopes, 2000; Jon-And, 2008, 2009) quer de grupos rurais isolados ou semi-isolados (Emmerich, 1984, 1992; Baxter, 1992, 2004, 2009; Lucchesi & Macedo, 1997; Lucchesi, 2000a; P. Andrade, 2003; Tieppo, 2003; Figueiredo, 2008, 2009a, 2009b).

Trabalhando com uma variável dependente dotada de dois ou mais valores representativos do fenómeno em análise (no nosso caso, a ausência ou marcação de número PL no SN) e um sistema de múltiplas variáveis linguísticas e extralinguísticas (as variáveis independentes), que incluem diversos factores relacionados com contextos linguísticos e sociais da comunidade em estudo, o programa VARBRUL combina relações entre a variável dependente e as variáveis independentes, a fim de determinar o modo como os factores das segundas motivam o fenómeno em análise e representado pela primeira variável. Para o programa poder realizar as combinações, deverão ser retirados os factores ou variáveis independentes que ele não considere estatisticamente relevantes, pois só assim se obterão valores fiáveis tanto para a frequência de uso conotada a cada um dos valores dos grupos de factores independentes como para o peso que cada um dos factores exerce na motivação do fenómeno em estudo. Os resultados, fornecidos, em pesos probabilísticos, quantificam a interferência exercida por cada variável independente na aplicação da regra em estudo, revelando, deste modo, quais são os grupos de factores que contribuem ou não para a variação (Preston, 1996b:239). Como tal, o grau de fiabilidade é superior aos dos valores representados através de meras percentagens, já que estes não permitem concluir se a variação em observação é sistemática. Paralelamente, o recurso às variáveis extralinguísticas deixa observar, diacronicamente, quando e como se incorpora, aplica e desenvolve determinada regra durante o processo de aquisição de uma língua.

A multiplicidade de recursos fornecida pela ferramenta VARBRUL faz dela um dos instrumentos ideais para se observar e explicar a sistematização da variação, daí termos recorrido a tal suporte computacional para efectuarmos a análise da CPL-var detectada nos SN's da fala da comunidade de Almojarife. Para Preston (1996b:253), todos os falantes de uma determinada L1 constituem a comunidade linguística desse dialecto, uma *sibling specie* de outras línguas irmãs que compartilham a mesma estrutura, mas que se desenvolvem independentemente (cf. Croft, 2000:13-15). Com efeito, esta pode possuir subgrupos distintos, filogeneticamente ligados entre si, já que compartilham não só os mesmos traços distintivos dessa mesma L1 mas também características sociais comuns a todos esses subgrupos. Deste modo, com recurso ao VARBRUL, é possível analisar os factos linguísticos presentes na produção linguística real dos falantes do PA,

uma variedade filogenética do PE, e tentar explicar a variação linguística patente nos seus SN's, relacionando-a com factores de natureza linguística e social. Simultaneamente, e numa perspectiva da aquisição, poder-se-á observar a integração e desenvolvimento de aspectos particulares não só da LA na fala da comunidade mas também dos fenómenos conotados aos substratos que poderão estar a motivar a variação em estudo. Assim sendo, a nossa pesquisa assentou nas seguintes etapas de trabalho: observação da comunidade linguística em causa, escolha criteriosa dos informantes, captação de dados, levantamento de hipóteses de trabalho e análise e interpretação dos dados captados. Como base de partida para a nossa pesquisa acerca da CPL-var no SN tivemos em linha de conta os trabalhos pioneiros de Naro & Lemle (1976) e Lemle & Naro (1977), para o PB. Em adição, e dado que neste momento é vasta a análise sobre a CPL-var, sobretudo no PB, adoptámos também os parâmetros e métodos subjacentes a outras investigações, nomeadamente os seguidos por Braga (1977), para o português mineiro do Rio de Janeiro, Guy (1981a) e Scherre (1988), para a fala do município do Rio de Janeiro, Lopes (2000), para a NURC, P. Andrade (2003), para o dialecto de HEL-Ba, Figueiredo (2008, 2009, 2009a), para o PA, Jon-And (2008) para o PMp, Baxter (2009), para o PT e dialecto de HEL-Ba, e Jon-And (2009), para o PCV. Nas análises por nós efectuadas são estabelecidos bastantes paralelos de comparação com todos estes trabalhos,<sup>117</sup> especialmente os respeitantes às variantes africanas e brasileiras de português. Porém, outros trabalhos foram ainda levados em linha de conta, a fim de observarmos a variação no SN, nomeadamente os produzidos para o estudo de variantes de português da região sul do Brasil (Fernandes, 1996; L. Andrade, 2003) ou falas caboclas brasileiras (Tieppo, 2003). Por último, também os recentes estudos sobre a CPL-var nos SN's produzidos por aprendentes chineses do PtgL2, em contexto formal de ensino, foram por nós considerados na presente análise (Godinho, 2005, 2008, 2009; Silva, 2010).

Não obstante, a explicação dos resultados do nosso estudo não se limitará à óptica comparada, uma vez que não pode ser desconsiderada a possibilidade de a CPL-var ter como motivações os princípios universais subjacentes à aquisição das línguas. De facto, sendo o PA fruto de transmissão linguística irregular e tendo como substrato o santomense, isto é, um CP resultante da aquisição defectiva de PtgL2, os referidos

resultados terão também que ser observados numa perspectiva psico-cognitiva e generativista, nomeadamente no que concerne à aquisição das categorias funcionais por falantes adultos do PtgL2 e, mais concretamente, no que diz respeito à captação dos traços de número nos elementos do SN.

Assim sendo, o modelo teórico construído para a análise quantitativa da CPL-var nos SN's dos nossos dados segue a proposta dos estudos de Guy (1981a), Scherre (1988), Lopes (2001), P. Andrade (2003), Baxter (2004, 2009), Godinho (2005) e Jon-And (2008, 2009). Por seu lado, a análise e interpretação dos resultados assenta não só nos mesmos estudos mas também em diversos trabalhos das áreas psico-cognitiva e generativista sobre aquisição de L1 e L2, aquisição monolíngue e bilingue, fossilizações e aquisição dos traços de número nos elementos do SN do PE e do PB, referidos no capítulo 2. De acordo com o modelo teórico construído e tendo como suporte para análise dos resultados os trabalhos mencionados, procuraremos, com recurso a dados linguísticos reais, fornecer o nosso contributo para acrescentar algo mais aos conhecimentos já constatados acerca dos dispositivos inerentes aos processos de ASL por adultos em situação de contacto massivo, bem como de L1's adquiridas por transmissão linguística irregular em contexto de bilinguismo.

### **3.1.1. Recolha e transcrição dos dados do *corpus***

A abordagem meramente formalista, que determinou os estudos em linguística até às propostas labovianas, não possibilitava que as observações de idiomas levassem em consideração outros factores que não fossem os intrínsecos à própria língua. No entanto, as línguas evidenciam variação lexical, sintáctica e fonética relativamente ao que as gramáticas tradicionais, detentoras da norma-padrão, entendem como correcto, sendo alguns fenómenos da variação determinados pelas relações que a língua estabelece com fenómenos externos. Por outro lado, embora as formas orais não contempladas pela norma (e pelo registo escrito) apontem, bastantes vezes, para mudanças em curso, não significa que estas tenham que, necessariamente, ser de imediato aceites pelos organismos reguladores da língua, como as instituições de ensino. Assim, compete aos especialistas em linguística determinarem se as variações detectadas apontam, de facto, no sentido da mudança. Para se conseguir tal, é necessário comparar diferentes estádios

de uma língua, com recurso a dados linguísticos. Ora, para a maior parte do tempo de vida do português, tal acesso só era possível observando textos escritos, que na maioria das vezes não correspondem ao registo oral espontâneo. Assim sendo, estes não constituirão os *corpora* ideais para determinado tipo de investigação científica em Linguística (Histórica ou Diacrónica), mormente para identificação de fenómenos que se prendem com as mudanças decorrentes do contacto entre línguas. Todavia, hoje em dia é possível recorrer a *corpora* orais para se comparar a forma de utilização da mesma com o que a gramática tradicional prescreve ou aquilo que a norma padrão considera correcto, detectando-se e observando-se, assim, não só as mudanças em curso mas também o sentido para o qual estas se direccionam.

Ellis (1994:144) e Romaine (2003:411) defendem que a variação deve ser encarada sob dois prismas: o livre e o sistemático. O processo é desencadeado por um reduzido número de falantes relativamente a um insignificante número de situações linguísticas, pelo que a nova forma que entra na *interlíngua* apresenta uma variação não sistemática, que não é limitada por factores observáveis e ocorre em período de tempo limitado. Contudo, o fenómeno vai-se alastrando de modo progressivo, até se generalizar e atingir toda a comunidade linguística, tornando-se sistemático. Iniciando-se por razões linguísticas, psicológicas e sociológicas, a variação tem a agir por detrás dela diversas motivações, pelo que poderão actuar, no seu condicionamento, quer fenómenos como a diversidade geográfica, a tensão entre arcaísmos e neologismos, as deficiências reveladas ao nível da aprendizagem das línguas ou a movimentação em direcção à aquisição de prestígio social, quer fenómenos decorrentes do contacto entre línguas, com os falantes buscando a eficácia do seu sistema comunicativo através da simplificação do mesmo.

Relativamente ao contacto entre línguas, este decreta a emergência de uma LA que, geracionalmente, pode mesmo levar ao desaparecimento da L1 do falante, por vezes em espaço de tempo abrupto.<sup>118</sup> Trata-se, pois, de um fenómeno de mudança que tem origem em influências exógenas ao sistema linguístico, mas que o afecta na totalidade, já que ele necessita de se estabilizar para corresponder eficazmente aos anseios e necessidades comunicativas dos seus falantes. Como tal, enquanto no fenómeno de mudança que afecta os subsistemas da língua os falantes desta se mantêm monolíngues,



no caso da mudança por contacto, os falantes da “comunidade que muda” (*shifting population*) tendem a ser bilingues, mas usando uma L2 que manifesta variações decorrentes das interferências da L1 ou das mudanças inerentes à própria variedade agora falada. As interferências da L1 enraízam em várias questões, dependendo, por exemplo, de aspectos como a maior ou menor dependência política, económica ou social da “comunidade que muda” face à comunidade da LA.

Como se mencionou no capítulo anterior, a aquisição de uma nova língua é também caracterizada pela criação de “terceiras hipóteses”, transitórias ou que se cristalizam, de funcionamento de determinados subsistemas da LA, não só por influência da L1 mas também por reanálises originais da gramática da LA em fase de apropriação. Este processo de aquisição é descritível em termos de *interlínguas*, isto é, de sucessivas fases de aproximação à LA. Por outro lado, estas evidências da L2 também poderão não reflectir nenhuma das gramáticas das línguas em contacto que originaram a nova língua, mas corresponder simplesmente a idiosincrasias do falar bilingue com base nos universais linguísticos, afinal em tudo semelhantes aos das L1’s adquiridas em situação que não configura transmissão linguística irregular.

Atentando aos aspectos referidos, facilmente se infere que a Linguística não se pode reduzir apenas ao estudo das manifestações contemporâneas dos dialectos a que se convencionou classificar de normas. E se tivermos em conta que, para o linguista, os conceitos de dialecto e norma não são opostos,<sup>119</sup> então, e embora a Dialectologia tenha grandes tradições em Portugal e mantenha relações estreitas com a Linguística Histórica, se compreenderá a necessidade de direccionar também os estudos linguísticos para os fenómenos de mudança decorrentes do contacto. De facto, e embora este seja um fenómeno cujo processo é do mesmo tipo daquele que é inerente à própria língua, tem origem diferente e, conseqüentemente, desfechos diferentes.

Nesta perspectiva, impõe-se o estudo de dados linguísticos numa óptica sociolinguística, o que “*é normalmente entendido como buscar razões, no contexto social, para os diversos usos linguísticos.*” (Lopes, 2001:105). No nosso caso concreto, os dados que constituem o *corpus* da presente análise foram recolhidos com recurso a 18 entrevistas do tipo sociolinguístico-laboviano (Labov, 1972a, 1983), entre Fevereiro e Maio de 1998, no âmbito do projecto “*Semi-creolization: testing the hypothesis against*

*data from Portuguese-derived languages of São Tomé (Africa)*”, financiado pelo Australian Reserch Council e dirigido por Alan N. Baxter, com a colaboração de Dante Lucchesi. As gravações das entrevistas decorreram sob a responsabilidade do segundo e, tendo em vista a natureza das variáveis linguísticas e extralinguísticas a constituir para observação dos dados linguísticos produzidos, cada uma das conversas teve a duração de cerca de uma hora (Tabela 8.1).

Finda a recolha das entrevistas, as mesmas foram transcritas ainda em Almojarife por Dante Lucchesi. Nessa tarefa, colaborou o assistente de campo local, Francisco Paulino, que forneceu informações de valor extremo para a interpretação das produções de fala. Posteriormente, os dados auditivos e as transcrições foram entregues ao autor do presente trabalho, que recopiou os primeiros para formato digital, reviu totalmente as segundas, procedeu à sua reorganização e colectou os SN’s para análise.

Para a transcrição dos dados, foi seguida a linha metodológica da representação autêntica das realizações linguísticas proposta por Poplack (1989), dando-se conta de aspectos como os falsos inícios de discurso, as palavras interrompidas, as hesitações discursivas, as intervenções paralinguísticas, a reformulação de sentenças ou a pronúncia de palavras com características dialectais distintas da forma padronizada. Para tanto, recorreu-se à ortografia estandardizada, em detrimento da transcrição fonética (Tagliamonte, 2006:55-56), uma vez que a primeira permite uma melhor visualização do texto e, conseqüentemente, dos fenómenos que se pretendem analisar (cf. Projecto Vertentes do Português Rural do Estado da Bahia).<sup>120</sup> Contudo, a transcrição ortográfica procura reproduzir, o mais fielmente possível, as características da produção oral.

Atribuiu-se também um código identificativo a cada um dos informantes (Tabela 8.1), que surge entre parêntesis rectos após os exemplos apresentados durante o presente trabalho. Seguidamente, exemplifica-se com um pequeno excerto do diálogo entre o documentalista (DOC) e a informante (INF) [MANEVEM2], a aplicação de algumas das convenções ortográficas adoptadas na transcrição da totalidade dos registos áudio do nosso *corpus* (cf. Tagliamonte, 2006:56-64).<sup>121</sup> No referido diálogo, os SN’s por nós coligidos encontram-se grafados a negrito:

DOC: *Seora como se chama?*  
 INF: *Ma— chama Maria Pedro—**das**—Neve.*

DOC: *Hum...*  
 INF: *Quê nome co a dá... a lá é? Quê nome co a dá é? Maria—Pedro—**das**—Neve. Nã Stáva... nã Stáva. Mê nome em casa é Stáva. Mê batisado é Maria—Pedro—**das**—Neve.*

DOC: *Maria—Pedro—das—Neves?*  
 INF: *É. Maria—Pedro—**das**—Neve.*

DOC: *E nome casa é Stava?*  
 INF: *É. É Stáva.*

DOC: *Aqui as pessoas tem dois nomes normalmente, né?*  
 INF: *É. Sempre gen São Tomé tem **duas** nome. Tê nome pa chamá solteiro e casá... agora tem nome quê naturá. Naturá é quê Maria—**das**—Neve, que Stáva chamá só a toa, casa.*

DOC: *E porque derão esse nome Stava?*  
 INF: *É porque mia... mia velha é quê pôso Stáva. Que mia mãe nã pariu... pariu ho— macho. Mas três, quato pariu homé homé homé, **quato macho**. Depois é veio de trás, quê [ININT] onde quê stáva? [ININT] stáva.*

DOC: *Hum...*  
 INF: *Onde é quê stáva?*

DOC: *Hum... agora entendi. Antão era a— porque perguntaram onde é que estava?*  
 INF: *É. Onde quê stáva quê nã vi a muito tempo, que homé veio quato, pô veio agora é quinto.*

DOC: *E o marido da senhora?*  
 INF: *Mê marido nã stá.*

DOC: *Hum... ele— como ele chama?*  
 INF: *Mê marido?*

DOC: *Hum.*  
 INF: *Chama Lázaro.*

Como se pode confirmar, no presente excerto foram utilizadas as seguintes convenções ortográficas:

- (i) Falsos começos representados com um hífen:
  - (1) *ele— como ele chama?* [DOC]
- (ii) Pausas audíveis sinalizadas com reticências:
  - (2) *É porque mia... mia velha* [MANEVEM2]
- (iii) Palavras parciais marcadas com duplo hífen:
  - (3) *pariu ho— macho.* [MANEVEM2]

(iv) Nomes de pessoas, lugares, músicas, jogos, etc., que representam ocorrências lexicalizadas, grafados com hífen entre os seus elementos:

(4) *chama Maria-Pedro-das-Neve*. [MANEVEM2]

(v) Partes incompreensíveis do discurso indicadas com recurso a parêntesis:

(5) *Depois é veio de trás, quê (ININT) onde quê stáva?* [MANEVEM2]

(vi) Intervenções paralinguísticas transcritas com recurso à ortografia padronizada:

(6) *Hum... agora entendi*. [DOC]

(vii) Fenómenos linguísticos que configuram formas dialectais das palavras padronizadas ou das formas fonológicas *standard*, grafados com ortografia padrão e de modo a reproduzirem o enunciado real:

(7) *Mê nome um casa é Stáva* [MANEVEM2]  
“O meu nome em casa é Estava”

### 3.1.2. Tratamento dos dados

O dados colectados para análise foram codificados de acordo com as categorias e subcategorias morfossintácticas seleccionadas como variáveis-alvo para o nosso estudo, as quais se encontram descritas no ponto 3.3. Para possibilitar a análise dos dados áudio-transcritos e comparar os resultados obtidos com os de outros estudos levados a cabo para diferentes variedades reestruturadas de português, foi necessário submeter as unidades e categorias gramaticais a um controlo de operacionalidade descritiva, de acordo com o conjunto de noções e modelos teóricos já avançados por outros linguistas, nomeadamente em análises efectuadas para o PVB (Braga, 1977; Guy, 1981a; Scherre, 1988; Lopes, 2001; P. Andrade, 2003), para o PT (Baxter, 2004; 2009), para o PVA (Inverno, 2005), para o PMp (Jon-And, 2008),<sup>122</sup> para o PCV (Jon-And, 2009)<sup>123</sup> e para o próprio PA (Figueiredo, 2008, 2009a, 2009b). Estes passos do nosso trabalho encontram-se tratados nos capítulos 4 e 5.

### 3.1.3. Os dados dos informantes

O propósito inicial das entrevistas foi de gravar um *corpus* estratificado por sexo e idade, tendo os factores escolarização e estadia fora da comunidade sido também

contemplados, em virtude de se terem mostrado relevantes em diversos estudos de sociolinguística variacionista levados a cabo no Brasil. Assim sendo, entrevistaram-se informantes de ambos os sexos, repartidos por três faixas etárias distintas<sup>124</sup> e nascidos em Almocharife ou arredores, mas a residir nesta comunidade há longo tempo. Os informantes exercem ou exerceram também variados tipos de profissão, sendo a sua escolarização diversificada, indo desde a ausência total da mesma até distintos graus de aprendizagem. Com tal metodologia, tentámos evitar que os dados confluíssem na direcção de favorecer determinadas características linguísticas ou sociais.

No final de cada entrevista, a ficha de identificação do informante contemplava, pelo menos, os seguintes aspectos: (1) nome; (2) sexo; (3) idade; (4) línguas faladas; (5) local de nascimento; (6) profissão; (7) estado civil; (8) nível de escolaridade; (9) estadia fora da comunidade.

Refira-se ainda que todos os informantes são bilingues, uma vez que falam o santomense e o português reestruturado de aquisição recente, cuja variação no SN constitui o objecto do presente estudo. Como referimos, se aliarmos este aspecto ao facto de as características sociais dos entrevistados serem heterogéneas, entendemos como é grande o grau de variação que nos propomos analisar. A Tabela 8.1 apresenta os dados sociais dos informantes do PA.

Após efectuada a recolha e tratamento dos dados, interessa estabelecer o objecto de estudo e as etapas de trabalho que permitirão a sua análise. Para tanto, seguimos a metodologia adoptada por Godinho:

- “(i) *A determinação do fenómeno objecto de estudo, a concordância de plural no sintagma nominal (a variável dependente);*
- (ii) A determinação das variáveis independentes – os grupos de factores que estão relacionados com o fenómeno em estudo;*
- (iii) A codificação, que constitui a etapa preparatória para utilização do modelo linguístico;*
- (iv) A quantificação dos dados;*
- (v) A interpretação dos dados.” (Godinho, 2005:242).*

**Tabela 8.1.** Informantes de Almozarife: faixas etárias e seus aspectos sociais.

Nr.	Código do informante	Escalão etário	Sexo	Ensino	Estadia (Permanência fora da comunidade)	Duração da entrevista
01	[OSVALH1]	FE1 (20-40)	Masc.	4ª. Classe	Não	01:00:28
02	[ZECAH1]	FE1 (20-40)	Masc.	4ª. Classe	Não	01:10:00
03	[CASTEH1]	FE1 (20-40)	Masc.	8º. Ano	Não	00:59:14
04	[ANTOM1]	FE1 (20-40)	Fem.	8º. Ano	Não	01:04:45
05	[DULCEM1]	FE1 (20-40)	Fem.	3ª. Classe	Não	00:56:20
06	[CARMOM1]	FE1 (20-40)	Fem.	8º. Ano	Não	01:06:19
07	[ABILH2]	FE2 (41-60)	Masc.	4ª. Classe	Sim (Gabão e Libreville)	01:08:59
08	[LUIH2]	FE2 (41-60)	Masc.	Não	Sim (Príncipe)	01:25:57
09	[MAURIH2]	FE2 (41-60)	Masc.	3ª. Classe	Sim (Libreville)	01:03:28
10	[CLOTIM2]	FE2 (41-60)	Fem.	4ª. Classe	Não	01:02:36
11	[FRANCM2]	FE2 (41-60)	Fem.	Não	Não	00:56:28
12	[MANEVEM2]	FE2 (41-60)	Fem.	Não	Sim (Gabão)	01:01:10
13	[CELESH3]	FE3 (+60)	Masc.	3ª. Classe	Não	01:04:44
14	[MANOH3]	FE3 (+60)	Masc.	Não	Não	01:02:55
15	[SALVH3]	FE3 (+60)	Masc.	Não	Não	01:10:53
16	[PRINCEM3]	FE3 (+60)	Fem.	Não	Não	01:15:05
17	[MMDEUSM3]	FE3 (+60)	Fem.	Não	Sim (Angola)	01:11:51
18	[MINISM3]	FE3 (+60)	Fem.	1ª. Classe	Não	01:08:16

**Obs.:** Os nomes dos informantes não são fornecidos para salvaguarda da identidade dos mesmos.

### 3.1.4. Caracterização dos dados

Definidas as etapas do nosso estudo, começámos por colectar todos os SN's de número PL constantes nas entrevistas e procedeu-se à codificação dos dados do *corpus* que configuram SN's pluralizáveis, constituídos por pelo menos dois elementos. Refira-se que para o modelo de codificação não foram tidos em conta os SN's que apenas apresentam itens singulares (p.e. DET<sub>[-PL]</sub> N<sub>[-PL]</sub> – ponto 2.11.3.3 do presente trabalho), ainda que algum deles devesse representar o PL. Tal opção fica a dever-se ao facto de, por norma, a gramática tradicional não considerar “*como desvio a não-marcação de*

*plural nos dois últimos itens*” (P. Andrade, 2003:88). Assim sendo, SN’s como o constante nos exemplos [227c] ou [238] foram excluídos da nossa codificação:

[238] PA:  
*era forte e **branco tudo** gostava de mi* [CELESH3]  
“era forte e todos os brancos gostavam de mim”

Contrariamente, foram levados em conta, nas nossas codificações, os elementos que exibem plurais mas não são marcados na norma padrão, uma vez que tal aspecto é entendido pelos gramáticos como desvio. Portanto, itens como o numeral referido no exemplo [239] foram devidamente codificados, a fim de possibilitarem o seu estudo:

[239] PA:  
*Que morde-mo pode chegá **uns cincos abelha*** [DULCEM1]

O levantamento dos SN’s de número PL coloca ainda outras questões pertinentes, como a de eliminar SN’s que não se constituem como elementos válidos para estudos de variação linguística. Assim, foi preciso fazer valer o princípio da *Lex Parsimoniae* (Lei da Parcimónia)<sup>125</sup> em itens do tipo “*Palácio dos Congresso*” [CASTEH1], isto é, nomes próprios que funcionam como itens lexicalizados na forma PL. De facto, nestes casos a estrutura de número PL não é gerada por qualquer dos processos sintácticos que originam os plurais comuns da língua portuguesa. Igual critério tivemos em relação, por exemplo, aos SN’s que configuram locuções adverbiais envolvendo o nome “vez”, como acontece com “às vezes”, “umas vezes” ou “algumas vezes”, ainda que as mesmas registem um elevadíssimo e diversificado número de realizações no nosso *corpus*, como seguidamente se exemplifica:

[240] PA:  
*são coisa que às VEZE são coesa acontecida* [CASTEH1]

[241] PA:  
*às VEZ gen vê pexe* [OSVALH1]

[242] PA:  
*às VEZES eu vô pá bar trabalho* [CARMOM1]

[243] PA:  
*à VEZE também* [CARMOM1]

[244] PA:  
*Só é sozinho, à VEZ dois* [SALVH3]

Guy (1981a:124), ao analisar o efeito da *saliência fónica* na marcação PL, estabeleceu uma escala de saliências. Reconhecendo que ocorrências do tipo de “às vezes” podem ferir resultados finais, destinou dois factores para o elemento “vez”: um para a expressão “às vezes” e outro para as demais ocorrências em que intervém o referido item. De facto, deve ser levado em conta que a selecção de ocorrências que configurem variáveis linguísticas terá de assentar na robustez que estas possam conferir à análise do fenómeno da variação, e não apenas no critério da frequência. No caso presente, as variações à forma lexicalizada são de facto mínimas quando comparadas com a realização padrão “às vezes”, o que afecta de forma viciosa os resultados finais acerca da marcação do número.<sup>126</sup> Além do mais, é também frequente a produção de repetições sistemáticas após uma primeira realização, configurando o que se convencionou classificar de falsos inícios de discurso:

[245] PA:  
DOC: *Tem muita gente que passa mal na canoa, né?*  
INF: *Sim, várias VEZES, várias VEZES.* [DULCEM1]

Como se depreende, a codificação plena deste tipo de enunciados e sua posterior inserção nos dados para análise iria produzir um número desproporcionado de instâncias com o mesmo tipo de realização, acabando por viciar os resultados finais (Tagliamonte, 2006:94). Assim o entendeu também Scherre (1988:75), que “*desconsiderou*” igualmente a expressão “às vezes” nas suas análises.

No que diz respeito ainda à eliminação de repetições, o efeito da Lei da Parcimónia estendeu-se também às frases do seguinte tipo:

[246] PA:  
*Esses peixe piquinino piquinino piquinino.* [ZECAH1]

Depois de levantados todos os dados do *corpus relevantes para o nosso estudo*, contabilizaram-se 2.340 itens passíveis de marcação PL.



### 3.2. O suporte computacional

A análise da variação linguística inerente a determinada comunidade só é possível com recurso a metodologias que permitam quantificar as possibilidades que determinem essa variação, estando as mesmas dependentes dos contextos estruturais e das situações em que a língua é usada (Lucchesi, 2000a:146).

O uso da língua pode revelar variações em função de fenómenos como as limitações da memória, as distrações e até o registo específico de cada falante, isto é, o *idiolecto*. Contudo, este tipo de variação esporádica não se presta aos estudos em sociolinguística variacionista preconizados por Labov (1969, 1972a), os quais se debruçam sobre factores linguísticos e extralinguísticos que determinam um padrão de variação sistemático (Paolillo, 2002:2). Deste modo, foi com vista a observar este que Labov (1969) propôs a noção de regra variável, que surge condicionada pelo peso que cada factor, linguístico ou social, exerce sobre a aplicação da regra. Para tanto, Cedergren & Sankoff (1974) desenvolveram um modelo de regressão logarítmica múltipla, num pacote estatístico denominado VARBRUL, que deixa ver com precisão a influência que cada factor exerce sobre o uso da regra. Posteriormente, Pintzuk (1988) adaptou o programa ao uso em PC's.

Visto que permite não só trabalhar com dados quantificados e apreciar a distribuição geral da variável em estudo, mas também calcular o peso que cada factor exerce sobre a regra da aplicação da concordância, esta ferramenta tem sido frequentemente utilizada por linguistas em diversos estudos da CPL-var de variedades brasileiras (Braga, 1977; Guy, 1981a; Scherre, 1988; Fernandes, 1996; Lopes, 2001; L. Andrade, 2003; P. Andrade, 2003; Tieppo, 2003; Baxter, 2009) ou africanas de português, como o PT (Baxter, 2004, 2009), o PA (Figueiredo, 2008, 2009a, 2009b), o PMp (Jon-And, 2008) ou o PCV (Jon-And, 2009), dialectos que forneceram ou estão a fornecer resultados que poderão ser comparados com os do presente trabalho. A opção por este tipo de ferramenta tem a ver com o facto de ela conceder a possibilidade de se avaliar a quantidade ou a força da relação entre os elementos submetidos à análise, já que esta é representada por valores incorporados numa fórmula que, por seu lado, permite predizer o que pode acontecer num determinado conjunto de situações. No que concerne à operacionalidade e funcionalidade do VARBRUL, interessa referir que as mesmas já

foram exaustivamente descritas quer na literatura da especialidade (p.e. Guy & Zilles, 2007) quer em trabalhos anteriores que abordam a questão da concordância variável (Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Godinho, 2005), pelo que remetemos para os mesmos as explicações detalhadas sobre aquelas. Poder-se-á, apesar de tudo, avançar, de modo sintético, que o pacote de programas VARBRUL “*combina relações complexas entre variáveis múltiplas, e a variável em análise – a variável dependente – pode ter dois ou mais valores*” (Godinho, 2005:66), por forma a conceder “*valores percentuais relativos às frequências dos diversos fatores estabelecidos para cada variável, além de fornecer o peso relativo de cada um dos fatores, informando sua relevância no comportamento do fenómeno sob estudo*” (P. Andrade, 2003:89). Relativamente ao “*ponto de vista da aquisição, o método permite aferir, com uma relativa margem de precisão, a tendência para o desenvolvimento, incorporação e aplicação de uma determinada regra gramatical ao longo do processo de aquisição de uma L2*” (Godinho, 2005:68). Cabe aqui referir, ainda, que as variáveis linguísticas submetidas a análise podem ser de categoria fonológica, morfológica, sintáctica ou lexical, constituindo grupos de factores compostos por variantes (os valores possíveis das referidas variáveis) nas quais pode ocorrer, ou não, variação. Deste modo, o potencial fornecido pela ferramenta VARBRUL, associado a constatações por parte do pesquisador sobre o uso da linguagem em múltiplos contextos, permite ir para lá da fronteira do mero juízo da “gramaticalidade” ou “agramaticalidade” (Paolillo, 2002), isto é, deixa antever se o uso da variação é sistemático, no sentido de configurar mudança em curso.

Não obstante a fiabilidade do pacote de programas VARBRUL na previsão dos efeitos dos contextos sobre a variável dependente em observação e dos padrões de variação da comunidade linguística em estudo, o entendimento dos resultados está directamente relacionado com a quantidade de dados submetidos a análise. De facto, o método matemático com que a ferramenta trabalha determina que, quanto mais numerosos e sólidos forem os dados, mais rigorosa se torna a fiabilidade dos resultados finais (Young & Bayley, 1996:301). Para além destes aspectos, a ferramenta permite ainda que o seu utilizador modifique hipóteses iniciais e reanalise os dados de modo fácil (Young & Bayley, 1996; Paolillo, 2002).

A análise da variação linguística com recurso ao VARBRUL implícita, por parte do pesquisador, uma progressão metodológica que obedeça às seguintes etapas:

“(1) A colocação das hipóteses de natureza linguística e social. ... na colocação das hipóteses é preferível ser mais liberal do que conservador, uma vez que se torna mais fácil retirar factores ou grupos de factores que revelaram não afectar a variação na interlíngua do que incluir factores já depois de codificados os dados (Young & Bayley, 1996:57); (2) a codificação dos dados; (3) a verificação da codificação; (4) a recodificação de acordo com os primeiros resultados; (5) a avaliação dos factores e grupos de factores significativos para a análise e (6) a interpretação dos dados de acordo com as hipóteses colocadas” (Godinho, 2005:230).

Tecnicamente, o VARBRUL fornece a possibilidade de se efectuarem análises de uma variável dependente que tem valores nominais, ou seja, categorias que configuram as realizações possíveis dessa variável. Esta variável dependente pode ser binária (uma escolha entre duas alternativas), ou enária (uma selecção entre  $n$  alternativas). Assim sendo, o modelo pressupõe não só a existência de influências de diverso tipo (as variáveis independentes) que atingem, probabilisticamente, um ponto único de “decisão”, “opção” ou “escolha”, mas também a possibilidade de ocorrerem diversas observações, independentes do resultado da “decisão”. As variáveis independentes constituem, portanto, grupos de factores, enquanto as observações representam os dados do *corpus*, devendo estes, para tanto, ser codificados de modo a representarem uma determinada combinação de factores.

O VARBRUL trabalha na base de probabilidades, executando um modelo de regressão múltipla, fundamental para verificar o efeito que um conjunto de variáveis independentes exerce sobre a variável dependente. Deste modo, pode observar-se quais são as variáveis independentes que afectam, ou não, a regra da concordância em número. Este aspecto só é possível porque a ferramenta permite testar o grau de significância de cada grupo de factores, processo que se inicia com a rodagem de apenas um grupo de factores e ao qual o programa vai, sucessivamente, adicionando outros grupos de factores, até que todos estes tenham sido levados à análise computacional. Este processo, conhecido como análise *step-up/step-down*, fornece, no final, os valores para os pesos relativos e para o *log-likelihood* (cálculo da verosimilhança máxima),<sup>127</sup> um parâmetro estatístico que mede com elevado grau de confiança as frequências de uso de determinado registo linguístico, por forma a que este possa ser projectado como

estatisticamente representativo da comunidade. As formas como se processam quer as análises *step-up* e *step-down* quer a leitura e descodificação dos resultados fornecidos por estas encontram-se detalhadamente descritas em trabalhos específicos sobre o modo como operar com a ferramenta VARBRUL (Johnson, s.d.; Paolillo, s.d., 2002; Pintzuk, 1988; Guy & Zilles, 2007), pelo que remetemos para os mesmos a leitura sobre estas questões. Resta apenas referir que os resultados das análises, para além, de possibilitarem a confirmação, ou não, das hipóteses iniciais, também “*precisam de ser interpretados de acordo com as hipóteses colocadas: (i) um grupo de factores não é relevante e, como tal, a hipótese não se confirma; (ii) um grupo de factores é relevante, mas a influência dos factores não tem o peso inicialmente atribuído e, como tal a hipótese não se concretiza e (iii) o grupo de factores é saliente e a influência dos factores confirma a hipótese*” (Godinho, 2005:232).

Para realizar a quantificação dos dados no presente trabalho, recorreremos ao pacote de programas GOLDVARB X (Sankoff, Tagliamonte & Smith, 2005), a versão mais recente do programa VARBRUL (Pintzuk, 1988) para ser aplicada ao *software* Windows. Todavia, importa referir que lançámos também mão de outras versões do programa, quando entendemos que as mesmas forneciam melhor desempenho para obtenção de resultados a observar na análise. Foi o caso da versão GOLDVARB 2001 (Robinson, Lawrence & Tagliamonte, 2001), que produz tabelas coloridas que facilitam a detecção de interferências estatísticas entre grupos de factores. Também foi de grande utilidade o programa TSORT, do pacote VARBRUL-2, que permite isolar os dados relevantes a qualquer factor.

### **3.3. Descrição da variável dependente e das variáveis independentes**

As variáveis linguísticas do presente trabalho, independentes e dependente, adstritas a grupos de factores, foram constituídas partindo do pressuposto que, ao intervirem na marcação PL dos itens que compõem o SN actuam também na configuração da concordância variável do SN. Observando-se a influência das variáveis sobre os itens do SN, obtém-se, então, um entendimento mais sólido acerca dos fenómenos que inibem ou não a marcação PL no PA.

A exemplo da metodologia levada a cabo por Guy (1981a), Scherre (1988), Lopes (2001), P. Andrade (2003) e Baxter (2009), para estudos sobre a concordância nominal de variedades de PB, e Baxter (2004, 2009), Figueiredo (2008, 2009b) e Jon-And (2008, 2009), para as variedades africanas de português, submetemos os dados codificados (veja-se ponto 3.4) a análises, recorrendo à abordagem do tipo atomístico. Deste modo, o nosso estudo fará a observação do SN sob uma perspectiva ontogénica, se assim se pode dizer, uma vez que o analisa à escala individual, ou seja, cada elemento do SN é entendido como um dado de análise (uma ocorrência da variável dependente), definido pela presença ou ausência de marca de PL. Como tal, atribuiu-se um valor positivo aos SN's que tenham todos os seus constituintes marcados e um valor negativo aos que contenham pelo menos um elemento sem marcação de PL. Por seu lado, os dados foram codificados de acordo com quinze variáveis independentes do tipo linguístico-explanatório para serem analisadas na sua interação com duas variantes (*SN Plural com Concordância Total* e *SN Plural sem Concordância Total*) da mesma variável dependente. Pesquisas sobre a CPL-var no SN do PA, com recurso à abordagem sintagmática, serão apresentados em trabalhos posteriores, que pretendemos levar a cabo para complementar esta que agora se analisa.

Das variáveis que nos propusemos contemplar na presente análise, algumas foram já observadas em estudos sobre CPL-var (Guy, 1981a; Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Baxter, 2004, 2009; Godinho, 2005; Figueiredo, 2008, 2009b; Jon-And, 2008, 2009), concordância variável de género (Lucchesi, 2000a; Figueiredo, 2009a) ou concordância verbal variável (Martins, 2007; Silva, 2010). Porém, também constituímos factores diferentes dos observados anteriormente em análises do tipo atomístico.

### **3.3.1. Variável dependente**

As variáveis dependentes constituem categorias particulares que se distinguem qualitativamente das outras variáveis também presentes na análise. Uma das características que as diferencia particularmente é o facto de serem descontínuas, isto é, não admitirem valores intermédios. Como tal, podem configurar-se como entidades binárias, ternárias ou enárias.

Importa também referir aqui que, para o estudo da *aplicação* ou *não-aplicação da regra da variação*, deverá ser descartado o pressuposto de que *aplicação* é sinónimo de desvio à norma padrão, enquanto *não-aplicação* significa o uso desta. Efectivamente, em algumas ocorrências, a própria norma padrão aceita mais do que uma forma,<sup>128</sup> defendendo Labov (1982) que a forma mais sistemática não é a prescritiva, mas sim a mais informal, isto é, a mais utilizada pelo falante comum. Nesta perspectiva, a discussão em torno da variação não deve ser subserviente à forma padrão, muitas vezes erroneamente entendida como o alicerce incontornável das ocorrências de um determinado dialecto.

A análise atomística da presente investigação contempla uma variável dependente binária (Tabela 9.1), que permite a obtenção de resultados indicativo-percentuais ou indicativo-probabilísticos para dois tipos de ocorrências: (i) a presença de morfema de PL no item analisado; (ii) em alternativa, a ausência de morfema de PL no item analisado. Este tipo de modelo binário radica nos princípios generativo-transformacionais aplicados na determinação da *aplicação* ou *não-aplicação da regra variável*,<sup>129</sup> e implica que a existência de um factor favorecedor da *aplicação* ocasiona, necessariamente, a proporcional *não-aplicação* do outro factor.

**Tabela 9.1.** A variável dependente (análise atomística).

<b>Variável Dependente</b>
1. Presença de morfema de plural no item analisado
2. Ausência de morfema de plural no item analisado

Com base na metodologia dos trabalhos sobre CPL-var mencionados no ponto 3.3 (Descrição da variável dependente e das variáveis independentes), partimos também do pressuposto que a marcação PL não acontece de forma casual, sendo condicionada pelas variáveis independentes. Considerandos mais detalhados acerca destas serão fornecidos no ponto seguinte deste trabalho.

### **3.3.2. Variáveis independentes**

Com base nos referidos estudos, Figueiredo (2008, 2009b) procedera já à constituição de algumas variáveis independentes linguísticas e sociais para uma observação

atomística preliminar acerca do modo como as mesmas poderão exercer influência na marcação de número dos itens do SN do PA. As variáveis constituídas foram: *idade* (variável social), *posição do item analisado relativamente ao núcleo do sintagma nominal*, *posição linear* (i.e. *posição do item analisado no sintagma nominal*), *marcas precedentes*, *contexto fonológico posterior* e *saliência fônica* (variáveis linguísticas). No presente estudo, tais variáveis foram complementadas com outras nove, de forma a obter-se um quadro mais completo sobre os fenómenos que subjazem à presença/ausência de marcação PL no SN do PA. Deste modo, numa perspectiva mórfica, acabámos por avaliar a possível influência na marcação exercida por quinze variáveis independentes, onze do tipo linguístico e quatro do tipo social. Estas variáveis estão patentes na Tabela 10.1 e serão detalhadamente descritas no ponto 3.6.

Como se pode ver na tabela, as variáveis independentes do tipo linguístico constituídas para o nosso estudo atomístico acerca da CPL-var no SN do PA configuram ainda diferentes subcategorias:

- (i) 3 variáveis fonológicas: *processos morfofonológicos de formação de plural* (doravante designada *saliência fônica*); *tonicidade dos itens lexicais singulares* (doravante designada *tonicidade*); *contexto fonológico posterior*.
- (ii) 7 variáveis morfossintáticas: *marcas precedentes*; *classe gramatical*; *posição do item analisado em relação ao núcleo do SN*; *posição linear*; *ordem do item na cadeia dos constituintes flexionáveis*; *grau de concordância de número no SN*; *marcação de género*.
- (iii) 1 variável lexical: *categoria semântica do nome*.

Após a constituição das variáveis que se pretendem observar, é necessário passar à codificação dos dados para serem introduzidos no suporte computacional, a fim de serem devidamente observados. Nesta conformidade, e com base nos estudos indicados no ponto 3.3, a observação dos dados envolveu as seguintes etapas: “(1) a definição do fenómeno em estudo (a variável dependente), isto é, a concordância de plural no sintagma nominal, (2) a definição das variáveis independentes – grupos de factores que são considerados como factores condicionantes da variável dependente em estudo, (3) a codificação dos dados – etapa que

*prepara os dados para serem tratados pelo suporte computacional, (4) a quantificação dos dados e (5) a análise da quantificação.*” (Godinho, 2005:188).

**Tabela 10.1.** Variáveis independentes (análise atomística).

<b>Variáveis independentes</b>	
<b>Variáveis independentes linguísticas</b>	
1.	Saliência fónica:
1a.	Processos morfofonológicos de formação de plural
1b.	Tonicidade dos itens lexicais singulares
2.	Marcas precedentes ao item analisado
3.	Contexto fonológico posterior ao item analisado
4.	Classe gramatical do item analisado
5.	Posição do item em relação ao núcleo do sintagma nominal
6.	Posição do item no sintagma nominal (Posição linear)
7.	Marcação de género
8.	Ordem do item analisado na cadeia dos constituintes flexionáveis em número
9.	Grau de concordância de número no sintagma nominal
10.	Animacidade do nome
<b>Variáveis independentes extralinguísticas</b>	
11.	Sexo
12.	Faixa etária
13.	Estadia (Permanência fora da comunidade)
14.	Escolaridade

### **3.4. Codificação dos dados**

Finda a constituição das variáveis dependente e independentes, foi necessário elaborar os factores que ficariam envasados nestas, com vista a podermos seleccionar os dados e preparar o arquivo de ocorrências que envolvem SN's PL, antes de passarmos ao tratamento do mesmo. Cabe aqui referir que, por vezes, a constituição dos factores coloca o investigador perante questões que exigem atenção e sagacidade cuidada, sob



pena de tratar incorrectamente os dados e, com tal, viciar os resultados finais. De facto, se bem que os factores das variáveis fonológicas se revelem relativamente lineares quanto ao tratamento (a alternância pode ancorar, por exemplo, na simples diferença de um ou dois traços fonológicos ou na realização ou não de algumas consoantes em final de palavra), o mesmo não se pode dizer dos factores das variáveis morfossintácticas e semânticas. Estes podem envolver, por exemplo, variação nas flexões, alternância de itens lexicais ou diferenças sintácticas produzidas pela derivação frásica. Perante realizações como a do exemplo [247], ao codificar-se atomisticamente o item “várias”, qual o critério a seguir: considerá-lo um adjectivo, com o significado de “diferentes”, ou, pelo contrário, entendê-lo como um quantificador, com o significado de “muitas”?

[247] PA:

*Mêmo que ele pode tê várias mulher [ANTOM]*

Este tipo de questões coloca o investigador, por vezes, perante o labirinto da ambiguidade, ao ter que lidar com as problemáticas relações entre forma linguística e função linguística, e que têm gerado alguma falta de convergência entre conceitos linguísticos e sociolinguísticos, começando pela própria tentativa de tratar o termo “significado”. Como tal, o pesquisador tem que desconsiderar o dado ambíguo ou, em contrapartida, optar por um critério único de classificação. No caso presente acreditamos que, na competência do falante almojarifano, actua em primeiro lugar o conceito do termo ligado à noção do concreto, isto é, de quantificador, e não tanto o conceito que remete para as conotações abstractas. Por conseguinte, pensámos atribuir aos itens “vários/várias” a categoria morfológica de quantificador, mas levantou-se aqui novo problema, já que aos quantificadores aparece também ligada a noção de indefinido. Voltaremos a esta questão no ponto 3.6.4 do presente trabalho, dedicado à constituição e codificação dos factores atribuídos à variável independente classe gramatical.

Por outro lado, a dificuldade do empreendimento cresce consideravelmente quando se lida com dialectos vernáculos, cujo código linguístico conhece variações consideráveis, quando comparado com o das formas padronizadas. A alternância das palavras na ordem linear da frase, associada à ausência de preposições, pode, em certos casos, colocar-nos perante realizações que exigem uma análise cuidada, quando se trabalha com a codificação de dados. Note-se, por exemplo, como a ausência de preposição atribui ao

SN preposicionado a função de OD na ocorrência seguinte, se observado à luz dos conceitos sintáticos da gramática tradicional:

[248] PA:  
*Ê vende **mil dobra** cada um peixe* [CLOTIM2]  
 “Eu vendo a mil dobras cada peixe”.

Numa perspectiva geral, a atribuição de factores a todas as variáveis ou contextos linguísticos e extralinguísticos que se pressupôs condicionarem a variável dependente é, também ela, parte do processo de codificação. Contudo, como a referida atribuição é regida por questões metodológicas bastante específicas, entendemos por bem expor os passos que seguimos na selecção dos factores que incluímos em cada uma das variáveis independentes. As tabelas que acompanham a explicação do critério de selecção têm o propósito de fornecer uma visão global sobre a forma como ficou constituído cada grupo de factores, com os seus respectivos factores.

### 3.5. Variável dependente constituída

A variável dependente que se observa no presente trabalho constitui a presença ou ausência de marca PL nos elementos do SN do PA. Principalmente com base nos trabalhos sobre CPL-var levados a cabo por Guy (1981a), Scherre (1988), Lopes (2001), P. Andrade (2003) e Baxter (2004, 2009), parte-se do pressuposto que a presença da marca de número PL não acontece por mero acaso, sendo sim influenciada pelos fenómenos que contemplámos nas variáveis independentes.

**Tabela 11.1.** Factores constituídos para a variável dependente.

1) Variável dependente				
Cód.	Factor	Exemplo	Nr. de Ocorrências	%
<b>S</b>	Presença de morfema de plural no item analisado	<b>JOVENS</b> habilitoso [CASTEH1]	1.202/2.340	49
<b>0</b>	Ausência de morfema de plural no item analisado	<b>jovens</b> HABILIDOSO [CASTEH1]	1.138/2.340	51

### 3.6. Variáveis independentes constituídas

#### 3.6.1. Variável independente *saliência fónica* e suas subvariáveis

Como podemos observar na Tabela 10.1, a variável independente *saliência fónica* contempla duas subvariáveis: *processos morfofonológicos de formação de plural* e *tonicidade dos itens lexicais singulares*. Na linha dos trabalhos que orientam a nossa pesquisa, iremos tratar estas individualmente, atribuindo à primeira a designação de *saliência fónica* e à segunda a denominação de *tonicidade*.

##### 3.6.1.1. *Saliência fónica*

O grupo de factores *saliência fónica* tem marcado presença constante nas pesquisas sobre concordância variável de diferentes tipos, desde os trabalhos primordiais de Naro & Lemle (1976) e Lemle & Naro (1977) sobre concordância verbal. Pela mesma altura, Braga & Scherre (1976), tendo como base o trabalho de Naro & Lemle (1976), aplicaram a metodologia aos estudos da concordância PL, postulando que existe uma relação directa entre maior frequência de marcação em número e maior *saliência fónica*, com a segunda gerando a primeira. A escala estabelecida por Braga & Scherre (1976:465-477) para avaliar o grau de diferenciação fónica na relação singular/plural propunha os seguintes níveis: (1) inserção de *-s* final e abertura vocálica ou PL metafónico (p.e. olho > olhos/); (2) inserção de *-s* final e mudança silábica (p.e. milhão > milhões; imóvel > imóveis); (3) inserção de *-es* final em palavras terminadas em *-r* no singular (p.e. flor > flores); (4) inserção de *-s* final em palavras que formam o PL de modo regular (p.e. livro > livros); (5) inserção de *-es* final em palavras que acabam em *-s* no singular (p.e. país > países; rapaz > rapazes). Os resultados levaram as autoras a concluir que as formas mais salientes favorecem mais a inserção de PL do que as formas menos salientes. Paralelamente, observaram também que a escala de *saliência fónica* não só ganhava mais evidência nos falantes de classe média e média alta do que nos de classe baixa como também apresentava tendência para inibir a marcação PL dos itens terminados em *-s* no singular e que pluralizam com a inserção de *-es* final. O fenómeno, por nós descrito no ponto 1.1 (Português do Brasil e concordância nominal variável de número), foi justificado em termos de analogia (Braga & Scherre, 1976:474).

A pesquisa de Braga (1977), na linha das anteriores, aplica a mesma escala de diferenciação fónica e volta a referir que a *saliência fónica* detém importância na marcação PL da classe média. Por seu lado, a classe baixa apenas inibe a pluralização dos itens terminados em *-s*, favorecendo-a em todos os outros factores. Posteriormente, a análise efectuada por Scherre (1978:79-82), observando factores sociais para o MRJ, determinaria uma escala de três níveis para verificar a influência da *saliência fónica* na marcação PL.<sup>130</sup> A autora detectou então que os falantes escolarizados aplicam mais a pluralização do que os semi-escolarizados, apesar de ambos favorecerem a marcação. Simultaneamente, Scherre (1978:82) verificou que os itens em *-s* no singular apresentam menor probabilidade de concordância do que os elementos regulares.

Ponte (1979:164-166), observando o PPA, recorreu a uma escala idêntica à de Scherre (1978), tendo constatado uma escala hierárquica de marcação semelhante à dos informantes semi-escolarizados do MRJ, mas com a diferença de os seus falantes favorecerem mais a marcação PL dos itens em *-s* no singular do que a dos elementos regulares. Por seu lado, Nina (1980), com recurso a uma escala de diferenciação fónica de quatro níveis,<sup>131</sup> analisou a pluralização no SN da fala da micro-região bragantina, Pará. Apesar de os resultados apontarem para uma fraca motivação da *saliência fónica* na pluralização, os itens que pluralizam com adição de *-s* final e mudança silábica (p.e. *casal/casais*) apresentaram uma taxa de inibição total (100%). Nina (1980) indicou ainda uma escala de hierarquia fónica com influência na marcação PL, mas Scherre (1988:70) chamou a atenção para o facto de a mesma não ter sido correctamente interpretada, até porque Nina (1980), ao trabalhar com percentuais, não mediu o peso real das marcações. Assim sendo, e por estarmos de acordo com a observação de Scherre (1988:70), abster-nos-emos de efectuar quaisquer comparações entre os dialectos bragantino e de Almojarife.

Guy (1981a:181-189), aplicando a perspectiva proposta por Naro (1981) de se observar a relação entre maior frequência de marcação em número e maior *saliência fónica* tendo em conta a diferença fónica e a tonicidade, analisou a marcação PL na fala de informantes semi-escolarizados do Rio de Janeiro, mas aumentou a escala de diferenciação fónica para sete níveis.<sup>132</sup> Com tal metodologia, o autor pretendia verificar também se se estabelece uma relação directa entre o grau de diferenciação fónica e a

percepção do ouvinte no momento da aquisição, levando este a apreender mais facilmente as formas que melhor percebesse. Nesta conformidade, a variável independente *saliência fônica* prestaria uma contribuição importante para se compreenderem dois fenômenos opostos no momento da aquisição, a não-realização e a apreensão, já que será “*a partir dos itens com concordância menos saliente que a variação se inicia; e é a partir dos itens com concordância mais saliente que se dá a percepção, na fase de aquisição do mesmo fenômeno: a aquisição do fenômeno da concordância se inicia pelos itens com maior oposição singular/plural*” (Lopes, 2001:132-133). Assim sendo, a variável que aqui se aborda presta-se ao estudo quer da variação da concordância quer da aquisição de L1 e de L2, já que não existe paralelismo quanto ao efeito da *saliência fônica* quer na aquisição de diferentes variedades do mesmo dialecto (Castro & Ferrari-Neto, 2007:68) quer na aquisição registada em distintas fases etárias (Correia, 2005; Baia, 2008). Contudo, a este propósito, lembre-se que, no caso específico da ASL (pidgins e crioulos), os falantes adultos poderão dar primazia às estruturas do tipo CV em virtude da sua força lexical (Clements, 2009:22), adquirindo-as em vez de estruturas que apresentam maior *saliência fônica*. Portanto, dificilmente se confirmará o postulado de Naro & Lemle (1976:240-241) e Lemle & Naro (1977:47) de que a mudança é determinada pela deriva natural da língua, iniciando-se nos contextos menos salientes e atingindo, depois, as formas mais salientes, por “imitação”.

No caso específico do estudo de Guy (1981a), estes aspectos parecem ajustar-se também, já que os resultados finais evidenciaram que não existe uma diferença relevante na influência da marcação PL motivada pelos elementos regulares e finalizados no singular em *-ão* e *-l*, com todos eles inibindo a pluralização. Por outro lado, o item “vez”, assim como os que finalizam no singular em *-r* e *-s*, isto é, os mais salientes, exibem tendência para favorecerem a marcação PL. O pesquisador chamou ainda a atenção para a forte propensão de os elementos que pluralizam metafonicamente se apresentarem marcados, ainda que, no seu *corpus*, o seu número fosse bastante reduzido. Como tal, Guy (1981a:189) entendeu que a variável independente *saliência fônica* pode mesmo ser reduzida a uma oposição binária em termos fonéticos: itens que terminam em consoante, com tendência para favorecerem a marcação, *versus* itens que terminam em

vogal, com propensão a inibirem a pluralização. Ainda a reter da análise levada a cabo pelo autor, são os seguintes postulados: (i) não é possível comprovar conclusivamente que a *saliência fônica* exerça efeito na marcação PL; (ii) a influência da tonicidade na marcação é aparente, visto ser idêntica à exercida pela mesma variável sobre o cancelamento do –s final não-morfémico.

Buscando confirmar as premissas advogadas por Guy (1981a), Scherre (1988) reanalisa os dados do MRJ e observa a possibilidade de o eixo da *saliência fônica* se estender igualmente ao número de sílabas. Com esta reanálise, a pesquisadora procura também estudar, em pormenor, a possível sobreposição entre as três dimensões referidas (diferença fônica, tonicidade e número de sílabas) na marcação do PL e conclui que a escala da *saliência fônica* não pode ser reduzida à oposição binária proposta por Guy (1981a), já que esta variável estará correlacionada quer com as características sociais dos falantes (actua a nível do funcionamento sincrónico da gramática, estabelecendo diferenças em função dos anos de escolarização dos falantes) quer com a Condição de Distintividade da hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que postula existir uma tendência para a informação semântica relevante ser retida na estrutura de superfície (as subcategorias da *saliência fônica* estabelecem diferenças entre si na forma de motivarem a marcação PL). Nesta conformidade, Scherre (1988:78) concluiu que, por um lado, “os processos de formação do plural e a tonicidade dos itens lexicais singulares influenciam a concordância de número entre os elementos do SN”, com os itens que apresentam maior distinção na oposição singular/plural a favorecerem mais a marcação, mas, por outro lado, que “o número de sílabas dos itens lexicais singulares não exerce influência sobre a concordância de número dos elementos do SN”. Em aberto, a autora deixou ainda a possibilidade de a *saliência fônica* estabelecer ligação com a relação entre variação e mudança, que viria a ser confirmada quando ela, no mesmo trabalho, se debruçou sobre a forma como as variáveis sociais intervêm no processo de marcação de número do MRJ.

Entretanto, Lucchesi (2000a:273), observando a concordância variável de género de HEL-Ba, conclui também que a *saliência fônica* influencia a marcação, já que esta é mais evidente nos itens que formam o feminino com alteração do radical da palavra, isto é, morfologicamente mais salientes.<sup>133</sup>

Lopes (2001:137), utilizando a uma escala de diferenciação fónica de sete níveis, idêntica à de Guy (1981a), recorreu igualmente às dimensões propostas por Scherre (1988), isto é, à diferença fónica, tonicidade e número de sílabas, estudando a influência da variável independente *saliência fónica* na marcação PL de dois públicos da NURC: o grupo popular e o grupo universitário. Paralelamente, e na linha do efectuado por Scherre (1988), observa também como a mesma variável actua sobre os diversos níveis de escolaridade (Fundamental, Colegial e Universitário), bem como sobre as diferentes categorias gramaticais. Outra particularidade do estudo de Lopes (2001:155-164) é a análise do modo como o grau *saliência fónica* actua ao nível da marcação PL em falantes de descendência afrobrasileira e não-afrobrasileira.<sup>134</sup> Lopes (2001:164) conclui que, no processo de aquisição, os itens mais salientes são os que apresentam o maior grau de concordância, enquanto os elementos menos salientes são alvo de variação. Este aspecto acaba por ter reflexo na concordância revelada pelo grupo afrobrasileiro, que denota maior índice de variação pelo facto de a sua história assentar na ancestralidade escrava, isto é, em falantes cuja aquisição do português se deu a partir de contactos linguísticos diversos, marcados por grande divergência nos DLP's.

P. Andrade (2003:100-104), por seu lado, estabeleceu uma escala de diferenciação fónica para oito níveis,<sup>135</sup> tendo os resultados iniciais apontado para a possibilidade de o efeito da *saliência fónica* não se fazer sentir na marcação PL do dialecto de HEL-Ba. De facto, parecia ocorrer, neste, a tendência para uma maior incidência de marcação nos itens de alternância regular, o que contrariava a teoria de as formas mais salientes serem mais marcadas do que as menos salientes. Esta particularidade levava a ferramenta VARBRUL a rejeitar o grupo de factores *saliência fónica*, não o considerando como motivador da variação registada na pluralização do SN. Contudo, amalgamando factores que compartilham traços linguísticos,<sup>136</sup> a linguista elevou os dados de ocorrências incluídas nas novas supercategorias de factores e fez com que o grupo de factores ficasse reduzido a três factores, tornando-se relevante para a marcação e fornecendo dados sobre índices de frequência. Os achados de P. Andrade (2003) levaram a pesquisadora a confirmar outros postulados anteriores (Guy, 1981a; Lopes, 2001), que chegaram a conclusões aproximadas, apesar de terem trabalhado com factores diferentes. A

conclusão final de P. Andrade (2003:104) foi que a *saliência fónica* tem influência na aplicação da regra da marcação PL.

Baxter (2004:107-113), analisando o PT, chama a atenção para o facto de a *saliência fónica* ter vindo a ser também observada quer em estudos sobre processos de aquisição em despidginização (Emmerich, 1984) e descrioulização (Bickerton, 1975) quer em trabalhos sobre aquisição de L1 (Ingram, 1989) e L2 (Bayley, 1994; Wolfram, 1985). Nas análises efectuadas, o autor optou por estudar o comportamento da *saliência fónica* quer quanto à alternância singular/plural quer na relação que estabelece com a tonicidade. Para tanto, constituiu uma escala de diferenciação fónica com seis graus.<sup>137</sup>

Cruzando as variáveis independentes *saliência fónica* e *tonicidade*, Baxter (2004:112-113) apresenta resultados a revelaram padrões gramaticais aproximados entre a geração mais nova e a geração intermédia dos tongas, mas que se afastam substancialmente dos da faixa mais idosa da mesma comunidade. À luz dos factos, este aspecto é perfeitamente entendível, uma vez que a geração mais velha esteve sujeita a forte *input* de DLP's L2 proveniente de falantes adultos cuja L1 é africana. Paralelamente, o efeito da *saliência* que foi detectado na marcação PL do PT vai ao encontro dos achados em estudos de aquisição, os quais apontam para um desenvolvimento precoce da morfologia flexional envolvendo as oposições mais salientes (Bayley, 1994:167), por contraste às oposições que atingem as *saliências* mais fracas, adquiridas mais tardiamente. Assim sendo, o efeito da *saliência* na marcação PL é fortemente favorecido pelas distinções morfofonológicas das oposições mais marcadas, pelas combinações de sílabas plurais tónicas e pelas mudanças que afectam o radical da palavra.

Para o PMp (Jon-And, 2008) e o PCV (Jon-And, 2009), Jon-And apresentou igualmente resultados provenientes do cruzamento entre os grupos de factores *saliência fónica* e *tonicidade*. No PMp, a marcação PL revelou tendência para incidir fortemente sobre as oposições mais salientes (p.e. *casal* > *casais*; *avião* > *aviões*; *pão* > *pães*), mas com os itens terminados em *-s* ou *-r* no singular, isto é, que fazem o PL acrescentando *-es* ao radical (p.e. *mês* > *meses*; *mulher* > *mulheres*), inibindo fortemente a marcação. Quanto ao PCV, todos os factores constituídos marcam positivamente a pluralização, à excepção dos elementos em *-s* no singular, que a inibem fortemente, e dos *oxítonos ou monossílabos tónicos regulares* (p.e. *café* > *cafés*; *pai* > *pais*), com um peso próximo da



neutralidade, mas a tender ligeiramente para a inibição. Deste modo, e relativamente a esta variedade de português, a linguista considera que “*a variável morfofonológica saliência fónica tem um efeito menos forte no PCV do que em outras variedades africanas e variedades brasileiras de português*” (Jon-And, 2009:7).

Tendo como base as metodologias de trabalhos anteriores, aplicámos inicialmente aos nossos dados uma escala de diferenciação fónica para oito níveis, de acordo com a Tabela 12.1. Assim, no que diz respeito aos itens que pluralizam com inserção de *-s* e mudança silábica, seguimos o proposto por Guy (1981a), mas ampliámos para três os nossos factores: (i) um para englobar os itens finalizados em *-l*; (ii) outro para abarcar os nasais em *-ão* que pluralizam em *-ões*; (iii) e um terceiro para o mesmo tipo de nasais, mas que formam o PL em *-ães*. A separação dos itens em *-ão* que pluralizam com inserção de *-s* e alteração silábica foi primeiramente testada por P. Andrade (2003:102) no dialecto de HEL-Ba, tendo os dados apresentado percentuais de marcação distintos (*-ão/-ões* = 16%; *-ão/-ães* = 0%). Recorrendo a idêntica distinção, pretendíamos confirmar se os falantes do PA estabeleciam também escalas diferentes de marcação para os itens e se essa distinção apresentava paralelismos com os dados de P. Andrade (2003). Contudo, não foi possível observar este efeito, já que no nosso *corpus* são nulas as ocorrência de elementos em *-ão* que pluralizam em *-ães*.

Quanto aos itens nasais em *-ão* que pluralizam de forma regular, isto é, com inserção de *-s* final (p.e. irmão > irmãos), fundimos o seu factor com os dos outros elementos nasais que formam também o PL de modo regular, ou seja, com os itens que finalizam em *-em* e *-um* (p.e. homem > homens; atum > atuns).

Para a constituição dos restantes factores que compõem a nossa escala de diferenciação fónica, adoptou-se a metodologia de Scherre e Braga (1976): (i) itens com inserção de *-s* em palavras de PL regular (p.e. primeiro > primeiros); (ii) itens com inserção de *-s* e abertura vocálica, ou seja, plurais metafónicos (p.e. jogo > jogos)<sup>138</sup>; (iii) itens terminados em *-r* e que pluralizam com inserção de *-es* (p.e. jogador > jogadores); (iv) e itens terminados em *-s* e que fazem o PL com inserção de *-es* (p.e. mês > meses). Relembremos ainda mais dois aspectos relativamente à inclusão de determinados casos em alguns dos factores por nós constituídos: (1) os itens com realização fonológica parcial de PL (p.e. trabalhador > trabalhadore; mês > mese) foram

codificados como possuindo marca de PL (Scherre, 1978), sendo incluídos nos factores dos elementos terminados no singular em *-r* ou *-s*, respectivamente; (2) as expressões do tipo de “às vezes” foram excluídas das nossas codificações. Scherre (1988:75) incluiu ainda no factor dos metafónicos as ocorrências com duas marcas de PL (p.e. *papelzinho* > *papezinhos*), que no nosso *corpus* são nulas.

Na tabela que se segue, podem ser vistos os factores que compõe a nossa escala de diferenciação material fónica, bem como o seu número de ocorrências e respectivos percentuais de marcação PL:

**Tabela 12.1.** Factores constituídos para a variável independente *saliência fónica*.

2) Variável independente <i>saliência fónica</i>				
Cód.	Factor	Exemplo	Nr. de Ocorrências	%
<b>R</b>	Plurais regulares	nos PRIMEIROS jogos [CASTEH1]	995/1.973	50
<b>D</b>	Plurais metafónicos	nos primeiros JOGOS [CASTEH1]	(10/15)	67
<b>Y</b>	Plurais nasais regulares: -ão/-ãos; -em/-ens; -um/-uns	<b>HOMENS</b> fiderado [CASTEH1]	69/145	48
<b>n</b>	Singular -ão/-ães	[-]	-	-
<b>B</b>	Singular -ão/-ões	a <b>RAZÕES</b> [CASTEH1]	(21/27)	78
<b>L</b>	Singular -l	<b>bancos</b> MUNDIAIS [ABILH2]	(3/10)	30
<b>;</b>	Singular -r	<b>JOGADORES</b> mais interessado [CASTEH1]	42/79	53
<b>s</b>	Singular -s (-z)	uns MESES [CASTEH1]	61/86	71
<b>Totais:</b>			1.201/2.335	51

Com recurso à variável independente *saliência fónica* formulámos a seguinte hipótese:

**Hipótese 1 – Saliência fónica (Processos de formação de plural):** o maior peso de marcação PL incide sobre os itens que apresentam maior *saliência fónica*. Nesta conformidade, são mais marcados os plurais metafónicos, que pressupõem dupla distinção (fónica e morfológica) e os itens que pluralizam com ocorrência de mudança

silábica (palavras terminadas em *-l*, em *-r*, em *-s* e em *-ão*). Quanto aos plurais menos marcados, serão constituídos pelos itens regulares.

### **3.6.1.2. Tonicidade (Tonicidade dos itens lexicais singulares)**

A variável independente *tonicidade* estabelece ligações intrínsecas com a *saliência fónica*, uma vez que aborda questões relacionadas com a maior percepção da sílaba tónica, isto é, da sílaba que se destaca fonicamente. Como se viu, a propósito das observações acerca do grupo de factores *saliência fónica* (ponto 3.6.1.1), alguns estudos em aquisição têm apresentado evidências que esta se inicia a partir das formas silábicas mais fortes. De um modo geral, os falantes de uma L1 têm consciência das sílabas que compõem as palavras pertencentes ao seu código linguístico natural. Nesta conformidade, pode considerar-se que as sílabas representam unidades de pronúncia intuitivamente reconhecidas e que qualquer falante, mesmo que seja analfabeto, saberá não só diferenciar as sílabas constantes, por exemplo, na palavra “janela” (*ja-ne-la*), como perceber também a sílaba tónica da mesma. De facto, quando articulamos dissílabos ou polissílabos, distinguimos, normalmente, uma das sílabas (a sílaba tónica), que se destaca em relação às outras (as sílabas átonas), sendo percebida como mais “forte” devido à intensidade com que é produzida a vogal que nela ocorre.

Não obstante, no caso do português, quando ocorre um registo oral mais rápido, mesmo os falantes maternos podem experimentar alguma dificuldade em fazer a distinção silábica, especialmente em sequências fonológicas que envolvem ditongos. Tomando como exemplo a palavra “chapéu”, isto é, um termo que envolve a ocorrência de um ditongo decrescente (sequência de uma vogal e uma semivogal), constatamos que ele se revela estável, pronunciando-se sempre como um ditongo e nunca como uma sequência de duas vogais distintas. Já em situações que pressupõem o uso de termos como “seara” ou “Manuel”, ou seja, que abrangem ditongos crescentes (sequência de uma glide e uma vogal), constata-se que estes poderão ser pronunciados como ditongos, numa fala mais rápida, ou como hiatos, quando o registo oral é mais pausado.

Também a percepção das vogais (e sílabas) tónicas e átonas poderá não ser tão evidente como à primeira vista parece, porque, em algumas realizações, o falante/ouvinte depara-se com ocorrências que pressupõem quer a captação das

chamadas vogais e sílabas subtónicas quer a não-distinção da fronteira de morfema entre segmentos contíguos. Com efeito, algumas palavras derivadas, como “cafezinho” ou “facilmente”, além da acentuação fónica principal, que incide sobre a sua vogal tónica, possuem também uma acentuação fónica secundária, que recai, por norma, sobre a vogal tónica da palavra primitiva. Por outro lado, alguns segmentos átonos dependem, em termos de acentuação fónica, de um segmento contíguo, que pode ser precedente (p.e. calou-se) ou posterior (p.e. a bata).

Durante o processo de aquisição da L1, a criança vai, paulatina e inconscientemente, adquirindo o sistema fonológico do seu idioma, isto é, não só vai apreendendo as relações que se estabelecem entre os sons como também vai delimitando o campo de funcionamento de cada um deles. Todavia, os próprios falantes de uma L1 podem enfrentar dificuldades na interpretação de determinados contínuos sonoros do seu idioma natural, mesmo em fase tardia. Assim sendo, e como referimos no ponto 2.7 (Aquisição das categorias funcionais da L2: os casos do santomense e português de Almojarife), a ausência de conhecimento da LA pode conduzir à não-interpretação de articulações sonoras desta, como acontece com a não-deteção da sequência DET+N do português por falantes de L1's do grupo níger-congo atlântico, que será fonologicamente interpretada como um todo indistinto (cf. DeGraff, 1999b:482). Portanto, não surpreenderá que as formas silábicas mais salientes sejam apreendidas em primeiro lugar (J. Silva, 2005:65), contribuindo, em muitos casos, para a emersão de fenómenos de variação, que acabam cristalizados por *transmissão irregular*. Face a estes aspectos, a variável independente *tonicidade* poderá configurar-se como um elemento de importância extrema para o estudo de imensos fenómenos linguísticos, sejam eles em situação de aquisição sejam em contextos sociais de variação.

Guy (1981a:138), ao observar o efeito da tonicidade no apagamento do –s dos falantes cariocas semi-analfabetos, constatou que a variável tonicidade apresenta forte efeito sobre o fenómeno, com os itens monossilábicos e polissilábicos tónicos a resistirem acentuadamente ao referido apagamento (mais os primeiros do que os segundos), e os elementos átonos a favorecerem-no bastante. A razão para as sílabas tónicas favorecerem mais a inserção do –s do que as sílabas átonas reside, segundo o autor, no facto de as primeiras possuírem forte traço prosódico, isto é, propriedades de

intensidade, duração e altura em graus superiores aos das segundas. Como tal, são perceptualmente mais distintivas, logo mais resistentes aos apagamentos, como o demonstram ocorrências várias de manutenção e apagamentos registadas não só na história de línguas como o latim e o português mas também de CP's, como acontece com a fala caboverdiana.<sup>139</sup> Quanto à distinção entre monossílabos e polissílabos, o fenómeno não é mais do que uma extensão do anterior, já que os primeiros são acusticamente mais longos e proeminentes do que os segundos.

Guy (1981a:137) observou ainda que a inserção do *-s* final é também moderadamente favorecida pela presença da sílaba tónica na primeira posição da palavra em contexto seguinte (quando esta ocorre), um fenómeno que encontra respaldo em diferentes línguas hispânicas, casos do espanhol panamiano (Cedergren, 1973), espanhol cubano (Terrell, 1979) e EPR (Poplack, 1992 [1979]). Face às evidências sincrónicas e diacrónicas reveladas pela resistência das sílabas tónicas aos fenómenos de redução em vários dialectos, Guy (1981a:292) entende que se estará na presença de um universal linguístico, já que estes operam a nível de qualquer idioma humano. Assim sendo, não surpreende que a análise da influência do fenómeno do traço prosódico das sílabas na marcação PL venha revelando presença constante em trabalhos sobre variação.

Scherre (1988:77), para observar o efeito da variável independente *tonicidade* na inserção de marcas do MRJ, tomou como ponto de partida o item singular, estabelecendo três subdivisões prosódicas para o mesmo: (i) oxítonos e monossílabos tónicos; (ii) paroxítonos e monossílabos átonos; (iii) proparoxítonos. A constituição de um factor único que inclui paroxítonos e monossílabos átonos ficou a dever-se ao facto de a maioria dos monossílabos átonos surgir realizada sob a forma de itens proclíticos dos substantivos, o que os leva, muitas vezes, a serem interpretados como sílabas iniciais das palavras seguintes. Assim, perde-se tanto a noção de definitude como a informação de PL, embora se preserve o /s/ como partícula proclítica. Consequentemente, este elementos apresentam um elevado índice de marcação PL que, na opinião de Scherre (1988:113), e corroborada por Lopes (2001:142), é muitas vezes virtual.<sup>140</sup>

As conclusões de Scherre (1988:135) sobre o peso da tonicidade da sílaba do item singular na marcação PL são que o mesmo influencia a concordância, ocorrendo esta com mais propensão nos elementos oxítonos e menos incidências nos casos em que a

sílaba final não é tónica (paroxítonos e proparoxítonos). No entanto, convém lembrar que os resultados da autora tiveram em conta apenas as classes gramaticais em que ocorre o maior grau de diferenciação material fónica, ou seja, os substantivos, as categorias substantivadas e os adjectivos.

Lopes (2001:144-145), apesar de concordar com Scherre (1988:113) quanto à questão do mascarar da pluralização dos monossílabos átonos, separou-os dos paroxítonos em virtude de, ao contrário destes, apenas realizarem o PL de forma regular. Assim, optou por constituir factores individuais para ambos. Após trabalhar os seus dados, a autora verificou que, tal como o MRJ, a NURC revela tendência para marcar mais o PL em oxítonos e monossílabos tónicos, logo seguidos dos monossílabos átonos. A inibir a marcação surgem os proparoxítonos (os mais inibidores) e os paroxítonos (os menos inibidores). Dado que a oposição entre singular e PL mostra tendência a recair sobre a sílaba mais saliente, Lopes (2001:141) entende que tal fenómeno reflecte duas situações: “1) uma perda da concordância, inicialmente nas formas menos salientes, pois se dá mais facilmente a conservação da regra em sílabas salientes; 2) um processo de aquisição do fenómeno, ou parte dele, resultante de percepção parcial, em que formas oxítonas e monossílabos tónicos, por terem o morfema de plural mais saliente, foram as primeiras a serem percebidas e serem alvo de concordância”.

Quanto a Baxter (2004:109), cruzou as variáveis independentes *saliência fónica* e *tonicidade* para observação da influência das mesmas na marcação PL do PT, metodologia igualmente aplicada por Jon-And (2008:4, 2009:5) para estudo tanto da CPL-var do PMp como do PCV.

**Tabela 12.2.** Factores constituídos para a variável independente *tonicidade*.

<b>3) Variável independente <i>tonicidade</i></b>				
<b>Cód.</b>	<b>Factor</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>X</b>	Oxítonos/Monossílabos tónicos	minhas <b>PREOCUPAÇÕES</b> [CASTEH1]	179/282	64
<b>M</b>	Monossílabos átonos	dAS <b>ondas</b> [CASTEH1]	283/320	88
<b>P</b>	Paroxítonos	<b>JOVENS</b> habilitoso [CASTEH1]	735/1.726	43
<b>&amp;</b>	Proparoxítonos	alguns <b>MÉDICOS</b> [CASTEH1]	(5/12)	42
<b>Totais:</b>			1.202/2.340	51

A Tabela apresenta os factores da variável independente tonicidade, tendo sido aplicada, para elaboração dos mesmos, a metodologia de Lopes (2001:144-145). Assim, e observando os fenómenos de marcação PL na linha do Princípio da Saliência Fónica (Naro & Lemle, 1976:240-241; Lemle & Naro, 1977:47), o que se pretende confirmar, tal como em todos os trabalhos anteriores, é se:

**Hipótese 2 – Tonicidade:** os itens do factor *oxítonos e monossílabos de uso tónico*, por apresentarem maior *saliência fónica*, são primeiramente apreendidos e revelam maior índice de concordância. Seguidamente, procura verificar-se se são mais marcadas as palavras proparoxítonas e, por fim, as paroxítonas e os monossílabos átonos.

### 3.6.2. Variável independente *marcas precedentes ao item analisado*

Os itens em análise, exibindo ou não marcas de PL, são também anteceditos por elementos que podem estar ou não pluralizados. Deste modo, para se verificar se a presença de marcas anteriores impede ou beneficia a marcação no elemento seguinte, é necessário recorrer à variável independente *marcas precedentes ao item analisado*.

Esta variável independente é clássica em estudos sobre concordância de número,<sup>141</sup> visto permitir o vislumbramento da influência que o item ou itens situados anteriormente ao elemento analisado poderão ter na marcação deste. Os estudos sobre este grupo de factores têm optado por considerá-lo quer de modo distinto quer correlacionado com outras variáveis independentes, nomeadamente a *posição linear*. Cruzando *marcas precedentes, posição linear e classe gramatical*, Scherre (1988:183) efectuou análises aplicando a metodologia de Poplack (1980a), que observou o apagamento do –s final, marca de PL, no EPR<sup>142</sup>. Os resultados de Scherre (1988:184) indicaram que os itens flexionados em número conduzem também à flexão do elemento ulterior. Por seu lado, as ocorrências com apagamento favorecem a inibição no item que lhes é posterior. Scherre (1988:238) justificou os resultados recorrendo ao Princípio do Processamento Paralelo, baseado no Princípio da Iconicidade (Haiman, 1983:782), que determina que o falante constitui, mentalmente, grupos que compartilham traços. Por outras palavras: Scherre (1988) concluiu que são as marcas que levam a marcas, por oposição a zeros, que implicam zeros.<sup>143</sup> Nesta conformidade, contesta a teoria funcionalista das Condições de Distintividade (Kiparsky, 1972:195), defensora de que a língua exhibe

tendência para reter a informação relevante na estrutura superficial e descartar aquilo que for dispensável, isto é, inserir o –s marca de PL apenas quando ele é fundamental para indicar a ideia de pluralidade. Como tal, caso ocorra marcação no primeiro item do SN, e segundo Scherre (1988), a tendência não será para a apagar nos elementos seguintes, a fim de se evitar a redundância de informação, mas sim de inserir mais marcas, regularizando-se o padrão da CPL-var conforme o padrão icónico das línguas: quanto mais coesão, mais marcas; quanto menos coesão, menos marcas. Todavia, Lopes (2001:88-89) chama a atenção para o facto de Scherre (1988) ter incidido a atenção sobre *“estruturas do tipo ‘os meus pais’, que são vistos como casos específicos, pois a presença de marca no possessivo independe da marca no artigo ou demonstrativo que o antecede. Outro aspecto visto por Scherre é que duas ou mais marcas precedentes é fator favorecedor e não o contrário, entrando em desacordo com o previsto pela hipótese funcionalista”*. Estes aspectos levariam Lopes (2001:202) a debruçar-se sobre a variável *marcas precedentes* de modo distinto do efectuado por Scherre (1988), mas, apesar das diferenças de análises entre ambas as autoras, as conclusões seriam idênticas quanto à questão do processamento paralelo: na NURC, o *“contexto antecedente com marcas leva a mais marcas subsequentes; e, o contrário, contextos com Zero imediatamente antecedente produzem uma inibição da regra”* (Lopes, 2001:261).

Por seu lado, P. Andrade (2003:100) e Baxter (2009:282) advogariam que os seus achados, apesar de não contrariarem os de Scherre (1988) e Lopes (2001), apontam para que a incidência de marcação PL esteja conotada não a factores linguísticos, mas sim a factores de ordem social. Em HEL-Ba, e segundo P. Andrade (2003:99), os itens antecidos por elemento com marcação formal de PL revelam tendência para inibição da marcação, por oposição aos que são precedidos de realizações sem marca formal de PL, que a favorecem. Em adição, os itens que apresentam a mais elevada incidência para a incorporarem a marcação são os antecidos por ocorrências que detêm marca semântica de PL, caso dos numerais. Se bem que a hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que advoga que a informação semanticamente relevante é retida na estrutura de superfície, encontre aqui algum acolhimento no que concerne à questão dos numerais, a verdade é que a autora ancorou as suas explicações no pressuposto avançado por Guy (1981a). Com efeito, este entende que a baixa



marcação no item seguinte ao elemento portador de marca de PL tem a ver com a influência das línguas africanas, nas quais “*predomina o uso de marcadores iniciais, em forma de prefixo ou clítico*” (P. Andrade, 2003:100). Assim sendo, confirma-se a tendência para, no PVB, o PL ser indicado na primeira oportunidade (Guy, 1981a:301-302), por norma o elemento que antecede o nome núcleo do SN.

Baxter (2009:280-282), por sua vez, comparando os resultados da variável independente *marcas precedentes* do PT e do dialecto de HEL-Ba, chegou à conclusão que o princípio de que *marcas levam a marcas* é marginal à atribuição de PL nos dois dialectos, os quais exibem estruturas de SN e padrões de marcação PL bastante simples. Como tal, e de acordo com o aventado por P. Andrade (2003:100), Baxter (2009:282) acredita que a marcação PL se deve não apenas a factores eminentemente linguísticos mas também à influência quer dos substratos quer dos contextos sociolinguísticos em que ambos os sistemas linguísticos se desenvolveram.

Relativamente aos estudos de Jon-And (2008, 2009) sobre o PMp e o PCV, a autora não apresenta quaisquer valores para o grupo de factores *marcas precedentes*, em virtude de o mesmo revelar interferências com outras variáveis e, conseqüentemente, ter sido excluído pela ferramenta VARBRUL.

Quanto a nós, a fim de verificarmos como se comporta a concordância PL no SN do PA em termos de marcas precedentes, fomos constituindo factores para a variável independente *marcas precedentes* à medida que íamos coligindo os SN's do nosso *corpus*, tendo como base orientadora a metodologia aplicada em trabalhos antecedentes (Braga, 1977; Guy, 1981a; Scherre, 1988, 2001; Lopes, 2001; P. Andrade, 2003; Baxter, 2009). Para tanto, e dado que o que importa verificar é o modo como o item precedente, com ou sem marcação PL (formal ou semântica), afecta a pluralização no elemento seguinte, tivemos como ponto de partida o elemento (ou elementos) que marca(m) o início da cadeia sintagmática, pelo que os nossos diferentes factores ficaram distribuídos por três subgrupos, envasados no grupo de factores *marcas precedentes*: (i) o subgrupo das cadeias sintagmáticas iniciadas por item com marcação formal de PL; (ii) o subgrupo das cadeias sintagmáticas iniciadas por item sem marcação formal de PL; (iii) o subgrupo das cadeias sintagmáticas iniciadas por item com marcação semântica de PL (numeral). No final, levaram-se à codificação os 30 apresentados na Tabela 12.3.

**Tabela 12.3.** Factores constituídos para a variável independente *marcas precedentes ao item analisado*.

4) Variável independente <i>marcas precedentes ao item analisado</i>				
Cód.	Factor	Exemplo	Nr. de ocorrências	%
/	1ª Posição/Ausência de marca precedente:	ALGUNS médicos [CASTEH1]	-	-
<b>Cadeias sintagmáticas iniciadas por item com marcação formal de plural</b>				
P	2ª posição, marca formal na 1ª posição: S	vários <b>PONTOS</b> [CASTEH1]	222/640	35
Z	3ª posição, 2 marcas formais precedentes: SS	os meus <b>PAI</b> [OSVALH1]	(14/26)	54
S	SSM	uns <b>riscos</b> assim AMARELO [CARMOM1]	(0/1)	0
T	S0	esses <b>bicho</b> PEQUENO [OSVALH1]	0/30	0
4	S00	esses <b>peixe</b> gordo GRANDE [MANOH3]	(0/5)	0
6	S000	esses peixe gordo grande COMPRIDO [MANOH3]	(0/4)	0
W	S0M	essas <b>galinha</b> memo NATURAL [CLOTIM2]	(1/4)	25
C	SM	<b>jogadores</b> mais INTERESSADO [CASTEH1]	(2/5)	40
M	SMS	tantos e tantas <b>INFORMAÇÕES</b> [CASTEH1]	(1/1)	100
2	SN	desses dois <b>NORMAL</b> [OSVALH1]	(7/22)	32
7	SN0	umas onze <b>jovem</b> FORMADO [CASTEH1]	(0/1)	0
<b>Cadeias sintagmáticas iniciadas por item sem marcação formal de plural</b>				
B	2ª posição, zero na 1ª posição: 0	a <b>RAZÕES</b> [CASTEH1]	39/46	85
3	00	o seu <b>DONOS</b> [ABILH2]	(3/3)	100
0	0S	<b>rocha</b> todas ESSAS [OSVALH1]	(4/7)	57
i	0SS	uma otras minhas <b>PREOCUPAÇÕES</b> [CASTEH1]	(1/1)	100
b	0N	todo cinco <b>LATA</b> [ZECAH1]	(0/13)	0
9	0NN(N)0	esse vinte e cinco <b>metro</b> QUADRADO [ZECAH1]	(1/2)	50
<b>Cadeias sintagmáticas iniciadas por item com marcação semântica de plural</b>				
N	2ª posição, numeral sem -s na 1ª : N	vinte <b>ANO</b> [CASTEH1]	121/412	29
R	2ª posição, numeral em -s na 1ª : Ns	duas <b>MULHERES</b> [CASTEH1]	79/278	28
\$	NS	cinquenta <b>metros</b> QUADRADO [ABILH2]	(0/1)	0
1	NsS	[-]	(0)	-
5	N0	cem <b>metro</b> QUADRADOS [ABILH2]	(1/2)	50
@	Ns0	Dois <b>filho</b> MULHER [FRANCM2]	(0/5)	0
g	NM	noventa tal <b>CONTOS</b> [ZECAH1]	(2/5)	40
8	NsM	[-]	(0)	-
a	NMM	[-]	(0)	-
e	NsMM	duzentos e tal <b>CONTO</b> [SALVH3]	(0/1)	0
#	NN(N)	quarenta e cinco mil <b>DOBRAS</b> [OSVALH1]	14/49	29
%	NNs(N)	trinta e três <b>ANO</b> [CLOTIM2]	(2/9)	22
<b>Totais:</b>			514/1.573	33

**Obs.:**

- (i) Símbolos aplicados aos factores (coluna Factor): S = item com marcação formal de plural; 0 = item sem marcação formal de plural; N = numeral não terminado em -s; Ns = numeral terminado em -s; M = item que não possui flexão de plural;
- (ii) Símbolo / (Coluna Cód.): situação em que o critério não se aplica;
- (iii) Símbolo [-] (Coluna Exemplo): factor constituído, mas que se constatou não apresentar qualquer ocorrência no nosso *corpus*.

A maioria dos factores regista ocorrências mínimas e alguns não apresentam quaisquer realizações. No capítulo 4, dedicado à análise dos grupos de factores do tipo estrutural, voltaremos a debruçar-nos sobre os trabalhos que estudam esta variável (ponto 4.1.2.4). Iremos tratar os nossos dados de diversas formas, seguindo metodologias aplicadas por outros autores (Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Baxter, 2009), a fim compararmos os resultados destes trabalhos com os nossos. Ao constituirmos esta variável independente e seus factores buscamos confirmar a seguinte hipótese:

**Hipótese 3** – *Marcas precedentes*: a presença de marca PL em elemento anterior ao item analisado inibe a marcação neste. Portanto, os itens tenderão a exibir inibição da marcação caso sejam antecidos de quantificadores e determinantes que apresentem marca formal de PL, bem como de numeral, que já possui informação semântica de número intrínseca.

### **3.6.3. Variável independente *contexto fonológico posterior***

É já dado adquirido em trabalhos de linguística que as variáveis fonológicas podem ser influenciadas pelos contextos que se lhes seguem. Segundo Scherre (1988:241), é sustentável inferir três predições no que concerne a este tipo de influência:

- (i) a busca do padrão silábico CV provoca uma influência directa na ausência de concordância PL, uma vez que esta se realiza, na grande maioria dos casos, com recurso à inserção do morfema *-s*, isto é, com adição de um item consonantal em final de palavra;
- (ii) face ao pressuposto anterior, também é lógico predizer que uma vogal em contexto seguinte favorece a inserção da marca formal de PL, porque estarão criadas as condições para ocorrer o fenómeno de elisão. Este, facilita a ligação entre a consoante final *-s* e a vogal inicial da primeira sílaba do item posterior, de forma a originar a sílaba padrão CV;
- (iii) se a vogal favorece a inserção PL, é crível também que o oposto possa suceder, isto é, que o contexto posterior representado por consoante iniba a marcação, uma vez que o fenómeno da elisão deixa de se poder realizar (Guy, 1981a:27-29).

Contudo, diversos trabalhos têm evidenciado que os efeitos produzidos pela realização de consoante ou vogal em contexto posterior não se têm revelado significativamente diferentes para a inserção ou não do *-s* marca de PL no elemento antecedente. Scherre (1978:7-8) havia constatado que algumas consoantes em contexto seguinte revelavam tendência para favorecerem a ocorrência de marca PL em nomes terminados em *-s* ou *-r*, ou seja, substantivos que realizam flexão total ou parcial em número, no primeiro caso com a inserção final de *-es* (*meses* > *mulheres*) e no segundo caso com a inserção final de *-e* (*mese* > *mulhere*). Estes aspectos levaram Scherre (1988:243) a analisar os elementos em contexto posterior em termos de traços, principalmente as consoantes, efectuando pesquisas na linha de Braga (1977) e Guy (1981a). Braga (1977:71-72) havia concluído que os itens palatais condicionam fortemente a marcação, em virtude da sua aproximação fónica com os segmentos de PL. Guy (1981a:146), por seu lado, debruçou-se sobre o efeito da sonoridade e do ponto de articulação das consoantes na inserção do morfema *-s* em itens regulares, constatando que as velares favorecem mais esta do que as labiais e as alveolares (e/ou dentais). Ao mesmo tempo, verificou que a tendência para a marcação decresce com a ocorrência de traço sonoro e aumenta com a existência de traço surdo.

Nos seus trabalhos, Scherre (1978, 1988) agrupou velares e labiais, inviabilizando, assim, comparações com o trabalho de Braga (1977). Paralelamente, como no português é relevante o fenómeno da neutralização, Scherre (1978, 1988) absteve-se de analisar a influência das palatais na marcação PL. Contudo, e tal como Guy (1981a), observou que esta é inibida pelos segmentos consonantais sonoros e favorecida pelos surdos, tendo considerado normal que tal aconteça, uma vez que os segundos ocorrem com mais frequência nas línguas.<sup>144</sup> Simultaneamente, Scherre (1988:255) constatou ausência de efeito na oposição consoante/vogal, mas ocorrendo efeito positivo da consoante e negativo da vogal nos itens terminados em *-s* (*mês* > *meses*).

O efeito da pausa tem sido também analisado em estudos sobre inserção da marca formal de PL, com resultados diversificados, como adiante se verá. Face a esta divergência de conclusões, Scherre (1988:245) incluiu também o factor pausa nos seus estudos sobre a influência do contexto fonético/fonológico posterior na pluralização. Nas suas observações, testou várias hipóteses de agrupamentos de factores: (i) constituindo

apenas quatro factores, que implicam as oposições maiores (vogal/consoante e pausa interna/pausa final)<sup>145</sup>; (ii) separando os itens lexicais em três grupos (regulares, terminados em *-s* e finalizados em *-r*); (iii) analisando detalhadamente o conjunto dos itens regulares em termos de segmentos e traços das consoantes, de acordo com o ponto de articulação, a sonoridade e as caixas de ressonância. Os resultados levaram Scherre (1988:255) a concluir que não só não existe busca do padrão silábico universal CV como também ocorre o fenómeno da haplologia sintáctica<sup>146</sup>. Paralelamente, regista-se uma outra oposição de acção na marcação: a pausa,<sup>147</sup> peguilhando o preconizado por Poplack (1980a:61) e confirmando o defendido por Guy (1981a:166), exerce influência na inserção da marca formal PL nos itens regulares (menos acentuada) e nos elementos terminados em *-s* (mais acentuada); por seu lado, os traços dos segmentos (ponto de articulação, sonoridade e caixas de ressonância), tal como já haviam advogado Braga (1977:40) e Guy (1981a:166), exercem fraca influência na marcação, apesar de as surdas propiciarem um pouco mais a inserção das marcas PL do que as sonoras.<sup>148</sup>

Lopes (2001:225) observou igualmente o *contexto fonológico seguinte* a cada elemento do SN e a sua interferência na presença de marca PL na NURC. Inicialmente, trabalhando também com os quatro factores que envolvem as oposições maiores (vogal/consoante e pausa interna/pausa final), a autora optou por manter sempre dois factores para variável pausa, a fim de não ferir os resultados finais. Estes revelaram que a pausa final é o factor que mais favorece a marcação, que a pausa interna e a vogal exercem um peso neutro na mesma, e que a ocorrência de consoante a desfavorece ligeiramente. Contudo a proximidade de valores entre as variáveis pausa interna, vogal e consoante poderia mesmo sugerir a constituição de um grupo de factores binário: pausa final e pausa interna/vogal/consoante.

Após obtenção destes valores, Lopes (2001:228-239) avaliou também a influência na marcação em termos do traço de sonoridade consonantal, constituindo, para o efeito, duas variáveis: itens regulares e itens terminados em *-s*. A exemplo da metodologia levada a cabo para outros grupos de factores, a autora observou depois o efeito das variáveis na marcação PL das seguintes variáveis sociais: *português popular*, *português universitário*, *escolaridade (Primária, Média e Universitária)*, *sobrenomes religiosos e sobrenomes não-religiosos*. Os resultados não se revelaram uniformes para todas as

variáveis independentes, mas, ainda assim, foi possível concluir que, por exemplo, o fenómeno da haploglia sintáctica (Scherre, 1988:244) actua também nos dados de Lopes (2001:239), com o contexto posterior vocálico desfavorecendo a marcação nos itens em *-s*. Mais elementos sobre as conclusões deste estudo serão apresentados aquando da comparação dos nossos resultados com os de outros trabalhos (capítulo 5).

No PT, Baxter (2004:113-115) estabeleceu cinco factores para analisar a influência da variável independente *contexto fonológico posterior* na marcação PL do SN: consoantes vozeadas, consoantes surdas, vogal, pausa interna e pausa final. Os achados revelam que a pausa final é o elemento que mais influencia a marcação PL na fala da comunidade, observando-se esta tendência em todas as gerações. Seguem-se-lhe, com pesos quase iguais, a pausa interna e as consoantes surdas, embora estes factores revelem tendência para condicionarem mais a marcação nas gerações menos idosas. As vogais e as consoantes sonoras, por seu lado, apresentam resultados inibidores da marcação, apesar de as primeiras se estabelecerem num patamar praticamente neutro. Baxter (2004:111) atribui a responsabilidade dos vários constrangimentos fonológicos verificados no PT às tendências de aquisição e influências de transferência que ocorrem no PtgL2, via DLP's fornecidos à geração mais idosa pelos falantes adultos de L1 africanas.

Ainda relativamente ao grupo de factores *contexto fonológico posterior*, não poderemos levar a cabo comparações quer com as variedades africanas do Maputo e de Cabo Verde quer com o dialecto de HEL-Ba. De facto, esta variável independente foi excluída das análises de Jon-And (2008; 2009) por, na fala dos maputenses e dos cabo-verdianos se revelar exiguamente influente na marcação PL do SN. Por seu lado, P. Andrade (2003) não constituiu esta variável para as suas análises sobre o dialecto brasileiro em questão.

Quanto ao nosso *corpus*, se o compararmos com os de Scherre (1988) ou Lopes (2001), revela-se quantitativamente mais modesto. Assim sendo, este aspecto inviabiliza de imediato a possibilidade de se constituir um elevado número de factores em termos de segmentos e traços para a variável independente *contexto fonológico posterior*, sob pena de dispersarmos muito os nossos dados e vermos alguns dos factores rejeitados por insuficiência de ocorrências. Ainda assim, constituímos um grupo de factores inicial que

contemplasse todas as possibilidades de ocorrências no *contexto fonológico seguinte* ao do item sobre o qual incide a marcação PL.

**Tabela 12.4.** Factores constituídos para a variável independente *contexto fonológico posterior*.

<b>5) Variável independente contexto fonológico posterior</b>				
<b>Cód.</b>	<b>Factor</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
F	Pausa final	esses <b>HOMENS</b> . [CASTEH1]	134/294	46
'	Pausa interna	umas nove ou onze jovem <b>FORMADO</b> [CASTEH1]	106/349	30
V	Vogal	umas <b>VOLTA</b> e recebe [CASTEH1]	234/409	57
P	P	OS <b>pais</b> [CASTEH1]	121/197	61
X	Ch	Tem UNS <b>cheros</b> [CARMOM1]	(2/4)	50
K	K	dois <b>DIAS</b> conosco [CASTEH1]	216/312	69
R	R	A <b>razões</b> [CASTEH1]	(13/17)	76
B	B	BOAS <b>barbatana</b> [OSVALH1]	29/34	85
Z	Z	ESSAS <b>zona?</b> [OSVALH1]	(8/8)	100
D	D	um outro <b>TRABALHOS</b> de televisão [CASTEH1]	79/186	42
-	F	umas nove ou onze <b>JOVEM</b> formado [CASTEH1]	34/48	71
T	T	um <b>OUTRO trabalhos</b> [CASTEH1]	51/101	51
J	J	nos <b>PRIMEIROS jogos</b> [CASTEH1]	16/38	42
L	L	ESSAS <b>lomba</b> [OSVALH1]	19/38	50
G	G	ESSES <b>garoupa</b> assim [ZECAH1]	(6/16)	38
M	M	uma <b>OTRAS</b> minhas <b>preocupações</b> [CASTEH1]	51/86	59
N	N	UMAS nove ou onze <b>jovem</b> formado [CASTEH1]	23/66	35
S	S	<b>BRAÇOS</b> suficiente [CASTEH1]	46/102	45
=	V	UMAS <b>volta</b> [CASTEH1]	14/35	40
			<b>Totais:</b> 1.202/2.340	51

**Obs.:** De acordo com a metodologia de Lopes (2001), considerámos pausa interna a distância da realização fónica que, durante a produção oral, não admite interferências entre o som do constituinte em análise e o som do item seguinte. Por seu lado, a pausa final marca o fim de uma frase ou período.

Convém ainda referir que não optámos pela metodologia de Lopes (2001:226), já que esta entendeu não considerar os “*contextos posteriores /s/, /z/, /x/, /j/, cujos dados foram desprezados, diante da dificuldade de detectar a real presença ou ausência da*

*marca de plural no elemento anterior*”. De facto, entendemos que, nos nossos dados, é possível detectar tais distinções, como o demonstram os exemplos que coligimos e apresentamos na tabela anterior, a qual exhibe ainda os factores por nós constituídos e exemplos das suas realizações, extraídas do nosso *corpus*.

Para a variável *contexto fonológico posterior* consideramos a seguinte hipótese:

**Hipótese 4 – Contexto fonológico posterior:** sendo o PA uma língua de padrão silábico CV, serão mais marcados os itens cujos contextos subsequentes configurem segmentos vocálicos ou pausa final. Por outro lado, caso o contexto subsequente seja consonantal, então o favorecimento da inserção de marcas será mais facilitado pelas consoantes surdas do que pelas consoantes sonoras.

#### **3.6.4. Variável independente *classe gramatical do item analisado***

Os trabalhos sobre a influência da variável independente *classe gramatical* na variação de marcas de PL no SN registam já mais de duas décadas e tiveram o seu pioneirismo em estudiosos como Cedergren (1973:43), quando esta se debruçou sobre o espanhol do Panamá e concluiu que os substantivos e adjectivos retêm a marcação, enquanto os determinantes a desfavorecem. Poplack (1980a:60), estudando o PL do EPR, chegaria, no entanto, a conclusões distintas, confirmando que, nos seus dados, os substantivos revelavam um índice de marcação ligeiramente acima do dos adjectivos. Por seu lado, Lefebvre (1981:80) analisa igualmente o efeito de determinadas classes gramaticais (nomes, pronomes e interrogativos) na marcação PL do quechua de Cuzco, um dialecto oficial peruano, chegando à conclusão que o grupo de factores *classe gramatical* não exerce influência sobre a mesma. Não obstante, os estudos de Lefebvre (1981) não nos colocam face ao fenómeno da concordância em número, que não existe no quechua de Cuzco, mas simplesmente em presença de marcas de PL em nomes e pronomes.

Posteriormente, Scherre (1988:148-151) faria um estudo minucioso sobre a influência da *classe gramatical* dos itens do SN na marcação PL deste. Categorizando-os detalhadamente e efectuando, pela primeira vez, uma análise que considerava esta variável de modo separado, primeiro, e inter-relacionada com as variáveis independentes



*posição linear e marcas precedentes*, depois, Scherre (1988:151-167) comprovou que estes três grupos de factores não dão conta dos fenómenos linguísticos na sua totalidade, caso sejam observados isoladamente. Como tal, apenas a convergência entre as diferentes variáveis conseguiria fornecer explicações viáveis acerca do funcionamento das mesmas na configuração da pluralização no SN. E, ao efectuar tal, Scherre (1988) viria a não concordar com o pressuposto avançado por Guy (1981b:179), quando este postulava que os efeitos da *classe gramatical e posição linear* eram idênticos em termos de influência na marcação PL. Assim, e se Guy (1981b:179) preconizava que a equivalência na marcação se regista a nível dos determinantes e primeira posição, dos substantivos e segunda posição e dos adjectivos e terceira posição, Scherre (1988:154) buscou evidenciar que, nos seus dados, os nomes ocupam a terceira posição numa frequência superior à dos adjectivos, que estes podem igualmente aparecer na segunda posição (na qual são mais marcados) e que os possessivos e restantes elementos não-nucleares surgem, amiúde, na segunda posição. Paralelamente, a autora constatou também que a primeira posição é sempre a mais marcada, independentemente da categoria gramatical do item que a está a ocupar.

Nas análises levadas a efeito para a NURC, Lopes (2001:172-175) buscou confirmar se ocorria também uma relação directa entre *posição linear e classe gramatical* nos moldes de Guy (1981b:179), tendo concluído que este tipo de associação não é acolhida no dialecto de Salvador, apesar de os artigos serem bastante marcados, por ocorrerem, normalmente, na primeira posição. Em adição, Lopes (2001) apresenta uma escala de marcação/inibição das categorias gramaticais que abordaremos em etapa posterior do nosso trabalho, quando compararmos os resultados do PA com os da NURC (capítulo 4).

P. Andrade (2003), apesar de ver a variável independente *classe gramatical* rejeitada nas suas rodadas VARBRUL, optou atender ao princípio sugerido por Scherre (1988) de que este grupo de factores não fornece dados conclusivos, caso seja observado individualmente. Assim sendo, P. Andrade (2003:95-98) cruzou a variável *classe gramatical* com a variável independente *posição linear*, confirmando que o efeito idêntico entre os determinantes e primeira posição se ajustava também aos seus dados, com esta categoria gramatical a evidenciar um número de marcas PL superior ao dos substantivos ou adjectivos na mesma posição. Ainda assim, a alta incidência de marcas

na primeira posição confirmava, em pleno, o pressuposto de Guy (1981a:301-302) acerca da pluralização do PB ocorrer na primeira oportunidade, levantando a possibilidade da sua antecedência crioula. P. Andrade (2003:98) considerou ainda a possibilidade de este tipo de ocorrência ancorar nos princípios psicolinguísticos que advogam uma ordenação na aquisição dos morfemas que estruturam os enunciados, isto é, no Modelo dos Quatro Morfemas (Myers-Scotton & Jake, 2000a, 2000b; Myers-Scotton, 2002). Como os morfemas de PL encaixam no grupo dos sistemas morfêmicos que configuram a classe dos *early system morphemes*, são passados, nas línguas em contacto, via transmissão linguística irregular, o que não acontece com a concordância, um processo que exige a intervenção tardia de outro tipo de morfemas sistêmicos, os *bridge system morphemes*.

Baxter (2004), para o PT, e Jon-And (2008), para o PMP, não apresentaram resultados para a influência das classes gramaticais na marcação PL do SN, tendo-se limitado a analisar a relação entre *posição linear* e *posição em relação ao núcleo*. A segunda justificou o facto de não ter analisado a variável *classe gramatical* em virtude de a mesma apresentar interferências com outros grupos de factores nos seus dados. Não obstante, Jon-And (2009:4) estudou o comportamento desta variável independente no PCV, observando que os definidos apresentam um peso favorecedor da marcação superior aos dos adjectivos e nomes. Ainda assim, os valores aproximam-se todos do patamar da neutralidade, levando Jon-And (2009:7) a concluir que a *classe gramatical* exerce efeito forte na marcação PL, possivelmente devido a transferências provenientes do CCV.

Refira-se ainda que os considerando que tecemos em torno dos resultados de Baxter (2004) e Jon-And (2008, 2009) serão fornecidos no capítulo 4, quando observarmos os resultados da variável independente *posição linear*, nomeadamente na parte em que se aborda a relação que a esta estabelece com outros grupos de factores.

#### **3.6.4.1. Constituição e definição dos factores para a variável *classe gramatical***

Para a nossa análise constituímos os factores constantes na tabela do grupo de factores *classe gramatical* (Tabela 12.5), tendo como base a linha de critérios da gramática tradicional, igualmente adoptada por Scherre (1988:150-151) e por Lopes

(2001:167). Assim, optámos também por classificar os adjectivos em duas classes, uma vez que determinados itens, tipo *mesmo* ou *próprio*, podem ser vistos como adjectivos ou indefinidos, enquanto outros desempenham unicamente a função de adjectivos. Os segundos integram então o factor denominado *adjectivo 1*, enquanto os primeiros foram codificados como pertencentes ao factor *adjectivo 2*. Relativamente aos graus dos adjectivos e substantivos, no nosso *corpus* foram detectadas 8 ocorrências dos primeiros, todas no grau diminutivo, e nenhuma dos segundos.

O factor *categoria substantivada* engloba todo o tipo de classes gramaticais (adjectivos, possessivos, indefinidos ou numerais), conquanto substituam o nome na sua função de núcleo do SN. A este propósito, convém mencionar que, para a codificação dos nossos factores, também não foi estabelecido um critério diferencial entre determinantes e pronomes. Assim, pronomes substantivos ou absolutos, como o pronome indefinido do exemplo [249], foram tratados como categoria substantivada, visto desempenharem a função de núcleo do SN:

[249] PA:  
*Eu ouvi os outros a contar.* [OSVALH1]

Nos casos como o do exemplo [250], em que estamos perante pronomes adjectivos ou adjuntos que acompanham o substantivo (os apelidados de determinantes, visto determinarem a extensão do significado do nome), optámos por designá-los simplesmente como possessivos, demonstrativos ou indefinidos.<sup>149</sup>

[250] PA:  
*Como é vejo outros amigo a fazê, ê també faço.* [OSVALH1]  
“como eu vejo os outros amigos a fazerem, eu também faço”

No que diz respeito aos artigos, Scherre (1988:150) optou por considerá-los conjuntamente com os demonstrativos, mas Lopes (2001:83-84) separou as categorias, tratando os artigos definidos e indefinidos num único factor.<sup>150</sup> Nos nossos dados seguimos a metodologia de P. Andrade (2003:96), que tratou os artigos definidos e indefinidos separadamente, a fim de observar qual das categorias exercia mais peso na regra de concordância PL. Contudo, é importante salientar que, no nosso *corpus*, bastantes vezes nos vimos confrontados com dúvidas na forma de tratar os artigos indefinidos. Se por um lado adoptámos o critério de Lopes (2001:183) no que concerne à

distinção dos pronomes e artigos indefinidos “um” e “uns”, por outro lado interessava destringir também se estávamos na presença do numeral ou do artigo. O exemplo seguinte pode fornecer uma ideia sobre a dificuldade que determinados SN’s colocaram para efeitos de codificação desta categoria, uma vez que nem o próprio contexto permitia distinguir com exactidão o que o informante pretendia transmitir:

[251] PA:  
*tinha um dois home trazia pão aqui.* [MMDEUSM3]  
“havia uns dois homens que traziam o pão aqui”  
“havia um ou dois homens que traziam o pão aqui”

Como se verifica, caso o item “um” seja tratado como artigo indefinido, o SN incorpora dois elementos à codificação, ambos sem marca formal de PL (artigo e nome núcleo do SN). Contrariamente, se o mesmo elemento for tratado como numeral, o SN apenas apresenta um item sujeito à codificação, também sem marca formal de PL (o nome núcleo do SN). Esclarece-se apenas que, no caso presente (e outros semelhantes), decidimos tratar o SN de acordo com o número e ordem original de constituintes na cadeia sintagmática, o que pressupõe, neste caso, a primeira possibilidade, isto é, estarmos perante um indefinido. Este aspecto é tanto mais compreensível, se levarmos também em linha de conta as questões debatidas no ponto 1.3.2.2, acerca do determinante indefinido *ũa*, do santomense (exemplos [36] e [37]), e que transporta as especificações de número e definitude (Alexandre & Hagemeijer, 2007:40). De facto, e ao que tudo parece indicar, a realização do exemplo [251] poderá indiciar uma interferência do substrato, representando o indefinido “alguns”, oposto de “um”.

Relativamente ao numeral, o mesmo não foi tratado por Scherre (1988) nem por P. Andrade (2003). Apenas Lopes (2001:84) se debruçou sobre esta categoria, incluindo ordinais, multiplicativos e fraccionários num único factor. Por razões óbvias, que se prendem com a não flexão dos cardinais em número, a autora não considerou estes para efeitos de análise. Para o nosso estudo, seguimos exactamente a metodologia adoptada por Lopes (2001), embora em alguns casos, como aquele que é reproduzido no exemplo [239], fôssemos obrigados a considerar a existência de flexão em número em determinados cardinais, quando as realizações assim o exigiam.

Quanto aos possessivos e demonstrativos, foram tratados de acordo com os princípios advogados pela gramática normativa. Nos primeiros, foram incluídas também as formas de pronome pessoal complemento directo contraídas com a preposição *de* (*dele, dela, deles, delas*), visto serem categorizadas como determinantes possessivos pós-verbais.<sup>151</sup> Relativamente aos segundos, as formas variáveis são substituídas, por vezes, pela forma invariável *isso*, que pode surgir flexionada em número PL, como acontece no exemplo [219].

Ainda sobre este item gramatical, Baxter (2004:108-109) entende que o mesmo poderá estar a desempenhar a função do artigo definido do português. De facto, observa-se, tanto no PT como no PA, a existência de formas inovadoras que actuam no lugar do referido artigo, como acontece não só com o exemplo [219] mas também quando o artigo surge contraído com preposições, situação a que fizemos já referência no ponto 1.3.2.3.1 do presente trabalho:

[252] PA:  
*fica com dinheiro três ou quatro dia ni mão* [LUISH2]  
“fica com o dinheiro três ou quatro dias na mão”

No que respeita aos indefinidos, foi nosso critério subdividi-los em dois factores: (i) o dos indefinidos propriamente ditos, englobando os itens que se aplicam à terceira pessoa gramatical, quando esta é considerada de um modo vago e indeterminado (factor H, Tabela 12.5); (ii) o dos quantificadores, que incluem os elementos indicativos de proporções numéricas indeterminadas (factor Q, Tabela 12.5).<sup>152</sup> Perante a dificuldade levantada por alguns indefinidos, detentores de características que os podem englobar em qualquer das subdivisões, como acontece com o item *vários/várias*, ao qual fizemos referência a propósito do exemplo [247], tentámos seguir um critério uniformizado, optando por englobá-los num único factor. Assim, no que concerne ao item em questão, decidimos classificá-lo como indefinido e não como quantificador, porque, seja qual for a sua situação de uso, possui sempre o traço que estabelece a distinção, a diferença, em detrimento do traço de quantidade, que pode ser apagado.

A propósito do quantificador invariável *tudo*, o mesmo regista realizações que reflectem uma ancoragem nas formas crioulas, nas quais é usado em vez do quantificador variável *todo*, com excepção para as ocorrências fossilizadas que

concorrem conjuntamente com numerais. Assim, e a exemplo do que sucedeu com o demonstrativo invariável *isso*, o quantificador *tudo* recebeu, por vezes, o tratamento de quantificador variável, quando o contexto assim o determinava:

[253] PA:  
*tudos os dia eu assistia televisão* [OSVALH1]  
 “todos os dias eu assistia à televisão”

A Tabela 12.5 exhibe os factores constituídos para a variável independente *classe gramatical*, seus códigos, exemplos, número de realizações e respectivos percentuais de marcação PL:

**Tabela 12.5.** Factores constituídos para a variável independente *classe gramatical do item analisado*.

6) Variável independente <i>classe gramatical do item analisado</i>				
Cód.	Classe gramatical	Exemplo	Nr. de Ocorrências	%
N	Substantivo grau normal	<b>TRABALHOS</b> privado [CASTEH1]	480/1.428	34
Y	Substantivo grau diminutivo	[-]	-	-
W	Substantivo grau aumentativo	[-]	-	-
C	Categoria substantivada	os <b>OTROS</b> [OSVALH1]	17/54	32
A	Adjectivo 1	<b>trabalhos</b> PRIVADO [CASTEH1]	19/65	29
D	Adjectivo 2	esses <b>jogadores</b> MESMO [ANTOM1]	(0/8)	0
Z	Adjectivo grau diminutivo	Esses <b>peixe</b> PIQUININO [ZECAH1]	(0/8)	0
“	Adjectivo grau aumentativo		-	-
G	Artigo definido	A <b>razões</b> [CASTEH1]	260/288	90
H	Artigo indefinido	UMAS <b>volta</b> [CASTEH1]	63/80	79
L	Numeral	nos PRIMEIROS <b>jogos</b> [CASTEH1]	(3/6)	50
O	Possessivo	MINHAS <b>preocupações</b> [CASTEH1]	31/37	84
V	Demonstrativo	ESSES <b>homens</b> [CASTEH1]	174/186	94
I	Indefinido	OTRAS minhas <b>preocupações</b> [CASTEH1]	62/64	97
Q	Quantificador	TODOS <b>jogadores</b> [CASTEH1]	93/116	80
R	Relativo	[-]	-	-
T	Pronome pessoal recto	[-]	-	-
<b>Totais:</b>			1.202/2.340	51

Refira-se que foram ainda constituídos dois factores para abarcarem os seguintes casos: pronomes pessoais rectos e relativos variáveis (p.e. *qual/quais*). Contudo, estes tipo de relativos não ocorre no PA (ver Tabela 12.5), confirmando-se a tendência para o padrão das relativas confluir para uma forma comum (o uso de “que”) nas variedades crioulezantes, um fenómeno que vem também afectando o PE. De facto, neste são já tidas como confirmadas quer a supressão do relativo “cujo” quer a estratégia cortadora que vem determinando o apagamento das preposições que ocorrem em paralelo com os relativos (Mateus & Cardeira, 2007:71), fenómenos que há muito se registam no Brasil e que determinam que, neste aspecto, o padrão de realização do PE esteja cada vez mais próximo dos do PA e PB.

Com base nas conclusões dos trabalhos a que nos referimos a propósito da variável independente *classe gramatical*, pretendemos confirmar a seguinte hipótese:

**Hipótese 5 – Classe gramatical:** os artigos, os quantificadores, os indefinidos, os demonstrativos e os possessivos configuram-se como os elementos gramaticais que tendem a receber mais marcas de PL, visto desempenharem funções de determinantes, isto é, ocorrerem, por norma, antes do núcleo do SN. As outras classes gramaticais propiciam-se menos à marcação, por configurarem a posição de núcleo do SN, ou se localizarem, preferencialmente, depois deste.

### **3.6.5. Variável independente *posição do item analisado em relação ao núcleo do SN***

Algumas variáveis do tipo estrutural, casos da *posição linear*, *posição em relação ao núcleo do SN*, *marcas precedentes* ou *classe gramatical*, têm-se revelado bastante importantes para se entenderem as motivações que subjazem ao fenómeno da CPL-var no SN. Aliás, se atentarmos à hierarquia dos grupos de factores do PA seleccionados como influentes na marcação PL (Tabela 14.1), constatamos que o topo da pirâmide é ocupado por este tipo de variáveis independentes, em detrimento das variáveis do tipo semântico ou fonológico, por exemplo. Paralelamente, apercebemo-nos também que, no vértice superior da pirâmide do PA, se encontra a variável independente *posição do item analisado em relação ao núcleo do SN*, considerada pela ferramenta VARBRUL como a mais importante para a inserção/inibição do número nos elementos do SN.

Refira-se que Scherre (1988) foi a primeira a tratar os grupos de factores *posição linear*, *marcas precedentes* e *classe gramatical* quer de forma individualizada quer cruzados entre si, já que, até então, os diferentes trabalhos efectuados sobre concordância variável apenas haviam privilegiado, de modo distinto, o efeito dos dois primeiros grupos (*posição linear* e *marcas precedentes*), sem analisarem a influência exercida pelo terceiro grupo (*classe gramatical*). Posteriormente, Lopes (2001:175-198) observou também o efeito exercido na marcação PL pela variável que se estuda neste ponto do nosso trabalho. Quanto a nós, optámos por incluir também uma outra variável de características estruturais nas nossas observações, a *ordem do item na cadeia dos constituintes flexionáveis*, cujo comportamento pensamos não ter sido ainda examinado em estudos de concordância em número. Pronunciar-nos-emos detalhadamente sobre esta variável no ponto 3.6.8.

Guy (1981b:179), analisando a pressão que os constituintes morfológicos, em determinada posição na cadeia sintagmática, exercem na marcação PL das produções dos falantes semiescolarizados cariocas, concluiu existir equivalência entre classe e sua localização na cadeia sintagmática. Scherre (1988:154) viria a discordar deste pressuposto, alegando, por exemplo, que todos os itens gramaticais surgem mais marcados na primeira posição no SN. Contudo, nenhum dos autores se debruçou sobre a variável independente *posição em relação ao núcleo do SN*, apesar de Scherre (1988:219) estabelecer uma oposição entre elementos não-nucleares à esquerda do núcleo nas primeira e segunda posições. A variável, como referimos, viria então a ser tratada por Lopes (2001:171-172), que, tal como Scherre (1988:154), defende existir uma frequência diferente de concordância para os itens da mesma classe, consoante a posição que ocupem na cadeia do SN. Estranhando o alto índice de marcação revelado pelos possessivos em segundo lugar na cadeia do SN, Lopes (2001:174-175) analisou a sua posição face ao núcleo do SN, concluindo que tal lugar na cadeia do SN poderia ser anterior ou posterior ao núcleo. Como tal, a marcação incidia mais sobre a posição pré-nuclear, em detrimento da posição pós-nuclear, mostrando que a variável independente *posição em relação ao núcleo do SN* é um elemento com forte influência na concordância PL. Face a esta evidência, Lopes (2001:175-177) decidiu então observar se os outros elementos não-nucleares revelavam o mesmo tipo de



comportamento que os possessivos, acabando por verificar que os itens possuem um índice muito maior de marcação quando estão localizados à esquerda núcleo. Paralelamente, Lopes (2001:177-178) constatou também que as posições lineares podem desfavorecer a marcação, a qual aumenta a partir da segunda posição. Para a autora, a explicação assenta, em parte, no modelo psicolinguístico dos Quatro Morfemas (Lopes, 2001:183-185), cuja ordem de emanção é análoga na aquisição de L1 ou ASL.

Baxter (2004:117, 2009:276), por seu lado, observando quer o PT falado por descendentes de africanos bantóides quer o dialecto de HEL-Ba, constatou também a tendência para a marcação PL favorecer mais o elemento não-nuclear imediatamente à esquerda do núcleo, enquanto os itens pós-nucleares produzidos pelas gerações mais idosas apresentavam uma quase total falta de marcação. Reanalizando o primeiro dialecto, Baxter (2009:278) apontou para ele um perfil de variação e funcionamento dos factores condicionantes estruturais semelhantes aos do PVB. Portanto, corrobora Guy (1981a:301-302) quanto aos mecanismos responsáveis pela variação da concordância PL no SN, quando o português é adquirido por falantes de determinadas L1's africanas. Nesta perspectiva, a mudança verificada no PT seria condicionada por factores que ancoram na aquisição do PtgL2 e PtgL1 adquiridos em situação que envolve a transmissão de componentes do PtgL2 influenciados pelo umbundu, a língua ancestral dos tongas. Como tal, verificava-se que a componente sociohistórica terá também desempenhado um papel importante, via DLP's do PtgL2, na aquisição do PtgL1 nativizado, leia-se PT. E, assim sendo, a hipótese avançada por Guy (1981a) de que os substratos africanos poderão constituir-se como a principal fonte da CPL-var do PVB, escorando-a a partir da posição imediatamente pré-nuclear, ganhava sustentação com os achados de Baxter (2004).

Porém, convém também adiantar que P. Andrade (2003:98), observando igualmente o dialecto de HEL-Ba, corroborara parcialmente Guy (1981b:179) no que concerne à equivalência entre classe e posição. De facto, a autora avançou que os determinantes revelam um elevado índice de marcação quando ocorrem na primeira posição e respaldou a explicação quer no *Modelo dos Quatro Morfemas*, já antes adiantada por Lopes (2001), quer na possibilidade de as línguas transplantadas de África para o Brasil poderem estar por trás do fenómeno que leva o PB a indicar o PL na primeira

oportunidade. Quanto às segunda e terceira posições serem preferencialmente ocupadas por substantivos e adjetivos, respectivamente, como advogara Guy (1981b:179), o facto é também corroborado por P. Andrade (2003:98), que avança ainda que estes lugares apresentam índices de marcação inferiores aos da primeira posição. Paralelamente, P. Andrade (2003:97) constataria ainda que o determinante *todo* é o único elemento gramatical que ocorre indistintamente em todas as posições, com o maior número de realizações a acontecer na terceira posição, mas a maior taxa de marcação a recair igualmente na primeira posição.

Jon-And (2008:5), por seu lado, apresenta o factor *elemento imediatamente anterior ao núcleo* como o único que marca positivamente o número, com todas as outras posições a revelarem-se inibidoras da marcação. A autora verifica ainda que esta variável sintáctica tem efeitos mais fortes no PMp do que no PT e no PVB, concluindo que a causa do facto deverá ancorar na influência que as L1's bantu exercem nos dados dos seus informantes, todos eles falantes do xironga ou xichangana, em contexto doméstico, e do português, em situações de actividade diária (Jon-And, 2008:7).

Para observar o comportamento da variável independente *posição em relação ao núcleo do SN* no PCV, Jon-And (2009:6) cruzou-a com o grupo de factores *posição linear*. Os resultados revelam também que os elementos em posição imediatamente anterior ao núcleo do SN favorecem bastante a marcação, enquanto a sua localização nas outras posições inibe a pluralização. Deste modo, Jon-And (2009:7) entende que a variável conjunta *posição em relação ao núcleo do SN/posição linear* revela efeitos semelhantes no PCV e outras variedades africanas e brasileiras de português.

Trabalhando com os dados do nosso *corpus*, Figueiredo (2008:28-29) havia também cruzado as variáveis independentes *posição linear* e *posição em relação ao núcleo do SN*, confirmando a existência de um número bastante diminuto de itens em posições pós-nucleares não imediatas. Este aspecto levou-o a criar um único factor que reúne todas as posições posteriores ao núcleo do SN, a fim de o mesmo ser seleccionado pela ferramenta VARBRUL e poder fornecer valores consistentes acerca do desenvolvimento geracional da CPL-var no SN do PA:

“Refira-se que, inicialmente, esta lista estava dividida em onze factores, mas, nas duas primeiras gerações, o número de itens com marca PL nas quarta e quinta posições no SN, sejam eles Núcleo sejam item plural à direita deste, é bastante diminuto. Este aspecto levou-nos a eliminar a 5<sup>a</sup>. posição das nossas codificações, reduzindo a lista inicial para apenas sete padrões configuracionais com a criação de uma nova variável dependente: a que reúne todas as Posições Posteriores ao Núcleo” (Figueiredo, 2008:29).

Deste modo, também na presente pesquisa se optou pela constituição de um único factor para análise de todos os itens que ocorrem em posição ulterior ao núcleo do SN, integrando-o no grupo de factores constante na Tabela 12.6:

**Tabela 12.6.** Factores constituídos para a variável independente *posição do item analisado em relação ao núcleo do SN*.

7) Variável independente <i>posição do item analisado em relação ao núcleo do SN</i>				
Cód.	Posição	Exemplo	Nr. de Ocorrências	%
0	Núcleo	‘SOAS morta [DULCEM1]	497/1.482	34
A	Anterior ao núcleo, não imediatamente	UNS quatro ano [CELESH3]	57/87	66
I	Imediatamente anterior ao núcleo	ESSES lugar tudo [LUIH2]	631/681	93
U	Posterior ao núcleo	homens CAPAZES [CASTEH1]	17/90	19
<b>Totais:</b>			1.202/2.340	51

Face ao que foi exposto anteriormente, colocamos a seguinte hipótese para a variável *posição do item analisado em relação ao núcleo do SN*:

**Hipótese 6 – Posição do item analisado em relação ao núcleo do SN:** Os elementos antepostos ao núcleo do SN são mais marcados do que aqueles surgem pospostos a este. Entre os primeiros, a marcação é mais acentuada nos itens adjacentes ao núcleo.

### 3.6.6. Variável independente *posição do item na cadeia do SN (Posição linear)*

O grupo de factores *posição linear* permite analisar a frequência com que certo elemento, em determinada posição na cadeia do SN, recebe a marca PL, independentemente da sua categoria gramatical. Este grupo de factores, quando cruzado com outros grupos (caso dos grupos *classe gramatical* ou *posição do item em relação ao núcleo do SN*), deixa também perceber, claramente, a relação de equivalência e

influência na marcação de número que existe entre determinados elementos gramaticais e a posição por eles ocupada no SN. Deste modo, a constituição do grupo *posição do item na cadeia do SN*, em concomitância com outros grupos de factores, permitirá confirmar se o PA cumpre também o pressuposto, já avançado em anteriores estudos sobre a CPL-var no SN de variedades de português, de que os itens em posição pré-nuclear revelam maior pendor para favorecerem a marcação, por oposição aos itens localizados pós-nuclearmente, que tendem a inibi-la (Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Baxter, 2004, 2009; Figueiredo, 2008; Jon-And, 2008, 2009). Por outro lado, os estudos têm também indicado que a primeira posição da cadeia sintagmática tende ser a mais marcada, por oposição às outras localizações, cuja propensão é de favorecerem menos a pluralização. Importa ainda referir que, nos nossos dados, de todas as variáveis independentes que observam a questão da posição estrutural, o grupo de factores que aqui se analisa é aquele que é considerado pela ferramenta VARBRUL como o menos relevante para a inserção de marcas de PL na cadeia sintagmática do PA.

Os trabalhos pioneiros sobre o efeito desta variável independente (Braga & Scherre, 1976) concluem já sobre a oposição de marcação entre primeira posição e outras localizações, com a primeira favorecendo a inserção de marcas e as outras desfavorecendo-a. Posteriormente, Braga (1977:38) analisa o comportamento geracional do mesmo grupo de factores e conclui que os falantes das classes baixa e média aplicam mais a regra da marcação na primeira posição, e que a mesma vai perdendo gradualmente a sua força nas posições seguintes.

Seguidamente, Scherre (1978:109) observa o comportamento da variável nos estratos escolares, concluindo que o índice de favorecimento evidencia duas direcções opostas, já que, em determinadas posições, a probabilidade de aplicação da regra sobe no grupo escolarizado e desce no grupo semi-escolarizado. Ponte (1979:162), no seu trabalho sobre o PPA, constata também que a primeira posição é o factor que mais favorece a marcação PL, decrescendo esta, sucessivamente, nas posições seguintes. Igual padrão de favorecimento e desfavorecimento é também detectado por Nina (1980) na micro-região de Bragantina, e por Guy (1981b:179), nos falantes semi-escolarizados cariocas. Este, contudo, afirma que os efeitos da *posição linear* e *classe gramatical* são idênticos, já que parece existir uma equivalência entre os determinantes e a primeira posição, os

nomes e a segunda posição e os adjetivos e a terceira posição. Assim sendo, será nessas posições que os citados elementos gramaticais apresentarão mais marcas.

Até aqui, todos os trabalhos referidos apontam para uma evidência que parece inquestionável: a variável independente *posição linear* é aquela que exerce uma influência mais uniforme sobre a regra da concordância de número nos itens do SN das diferentes variedades de português no Brasil. Por outra parte, parece confirmar-se também o Princípio Funcionalista de a tendência para a informação semântica relevante ser retida na estrutura de superfície (Kiparsky, 1972:195), levando a que a marcação ocorra preferencialmente na primeira posição sintagmática, tornando a pluralização desnecessária nos elementos situados noutras localizações, visto revelar-se redundante.

Scherre (1988:183), reanalisando o fenómeno da concordância PL no MRJ, foi a primeira a tratar os grupos de factores *posição linear*, *marcas precedentes* e *classe gramatical* em conjunto, em virtude de entender que o estudo isolado das mesmas não dá conta do fenómeno da concordância PL na sua totalidade. Simultaneamente, defendeu também que o português e o espanhol são duas línguas que marcam o PL da mesma forma, isto é, sem revelarem condicionamentos ao princípio funcionalista do tipo “kiparskyano”, de acordo com o detectado por Poplack (1980a) em estudos sobre o portorriquenho. Nesta perspectiva, o que motiva a pluralização do português são fenómenos que ancoram no Princípio do Processamento Paralelo (Scherre, 1988:208), ou seja, no pressuposto de que são as marcas que levam a marcas, por oposição a zeros, que geram mais zeros. Ainda assim, Scherre (1988:213) conclui que, em termos probabilísticos, é a primeira posição que surge sempre mais marcada. E assim sendo, não se confirma o pressuposto de Guy (1981b:179), quando estabeleceu uma relação directa entre posição linear e classe gramatical. Daí que este pressuposto tenha sido minuciosamente observado por Scherre (1988), que considerou o efeito apontado por Guy (1981b:179) como virtual, já que, nos dados do MRJ, a relação entre *classe gramatical* e *posição na cadeia sintagmática* apresentou valores que não confirmam a hipótese avançada por este autor.

Lopes (2001:165-167) observou o comportamento do grupo de factores *posição linear* recorrendo a cinco factores, idênticos aos que Scherre (1988) havia utilizado nas suas análises. Os resultados confirmaram o Princípio Funcionalista de que a primeira

posição é a mais favorecedora para a inserção de marcas de PL (Kiparsky, 1972:195), visto ser esta que atribui a informação nova de número ao SN. No entanto, os dados de Lopes (2001:180-181) não acompanham os resultados de estudos anteriores que revelaram uma diminuição progressiva da marcação à medida que os itens se afastavam da primeira posição, já que a quinta posição dos SN's da NURC revelou um ligeiro aumento do favorecimento relativamente às segunda, terceira e quarta posições. Sobre este aumento nos pronunciaremos mais adiante (capítulo 4), já que discordamos da metodologia adoptada por Lopes (2001) para a codificação dos seus dados.

Por seu lado, P. Andrade (2003:94-95), não estudou esta variável de modo isolado, tendo-a cruzado com o grupo de factores *posição em relação ao núcleo do SN*. Os resultados finais indicam que a fala de HEL-Ba não constitui excepção aos achados anteriores, já que os seus falantes pluralizam mais os elementos pré-nucleares e praticamente não inserem marcas nos elementos situados depois do núcleo do SN. P. Andrade (2003:96) testou também a marcação PL em cada classe gramatical considerando a sua posição no SN, concluindo que o efeito apontado por Guy (1981b:179) só pode ser atribuído à primeira posição, a mais marcada, e que apresenta, na maioria dos casos, um determinante a ocupá-la. Este aspecto permite, contudo, que se corrobore Guy (1981a:301-302) quando este afirma que a indicação do PL, no PB, tem tendência a ocorrer na primeira oportunidade, um padrão que não tem precedentes na história do PE.

Baxter (2004:115-117, 2009:278) adoptou também a metodologia que pratica o cruzamento entre *posição linear* e *posição em relação ao núcleo do SN*. Constituindo onze factores configuracionais,<sup>153</sup> o linguista estudou a influência dos mesmos na marcação PL do dialecto dos tongas (Baxter, 2004) e na marcação de número do PT e de HEL-Ba (Baxter, 2009), numa perspectiva geracional. Considerandos sobre os resultados de Baxter (2004, 2009) e comparação entre os mesmos e os valores do PA serão levados a cabo quando nos debruçarmos pormenorizadamente sobre os nossos resultados, já que testaremos também o efeito dos nossos grupos de factores na marcação PL de acordo com a mesma metodologia utilizada para os tongas e HEL-Ba (capítulo 4).

Jon-And (2008:5), para o PMp, seguiu a mesma proposta de constituição de factores de Baxter (2004), tendo elaborado nove factores<sup>154</sup> e apresentando resultados apenas para a comunidade. O factor elemento imediatamente anterior ao núcleo é o único que favorece a marcação PL nos dados de Jon-And (2008:7), revelando efeitos mais fortes nas produções do Maputo do que nas realizações dos tongas e do Brasil. A autora atribui o facto a uma possível influência das L1's africanas dos informantes maputenses.

Relativamente ao PCV, Jon-And (2009:6) optou por cruzar apenas os grupos de factores *posição em relação ao núcleo do SN* e *posição linear*. Os resultados apresentam a primeira posição como sendo aquela que mais favorece a marcação, mesmo quando esta é núcleo do SN. Como tal, confirma-se também, no PCV, o Princípio Funcionalista (Kiparsky, 1972) de caber a esta posição a atribuição da informação de número no SN.

Para a constituição dos factores que compõem o grupo de factores *posição linear*, adoptámos a metodologia seguida por Scherre (1988) e Lopes (2001), pelo que codificámos cinco posições, conforme Tabela 12.7.

**Tabela 12.7.** Factores constituídos para a variável independente *posição do item na cadeia do SN*.

<b>8) Variável independente <i>posição do item na cadeia do SN</i></b>				
<b>Cód.</b>	<b>Posição</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>1</b>	Primeira posição na cadeia do SN	<b>TRABALHOS</b> privado [CASTEH1]	678/754	90
<b>2</b>	Segunda posição na cadeia do SN	Umas <b>DORES</b> [DULCEM1]	441/1.314	34
<b>3</b>	Terceira posição na cadeia do SN	Os meus <b>FILHOS</b> [MANOH3]	48/168	29
<b>4</b>	Quarta posição na cadeia do SN	sete ou oito <b>MESES</b> [CASTEH1]	29/83	35
<b>5</b>	Quinta ou outra posição na cadeia do SN	umas nove ou onze jovem <b>FORMADO</b> [CASTEH1]	6/21	29
<b>Totais:</b>			<b>1.202/2.340</b>	<b>51</b>

Através da influência exercida por esta variável independente na marcação PL do SN, pretendemos confirmar, ou não, a seguinte hipótese:

**Hipótese 7** – *Posição do item na cadeia do SN*: a primeira posição do SN é a que mais favorece a inserção de marcas. As outras posições tendem a marcar menos o PL, pelo facto de o traço de PL (a informação semanticamente relevante) já ter sido atribuído.

### **3.6.7. Variável independente *marcação de género***

A constituição da variável independente *marcação de género* tem como objectivo verificar se existe uma correlação entre a marcação de género e a marcação PL em cada constituinte do SN. Efectivamente, e dado que o morfema de género surge mais directamente conectado à estrutura mórfica do item analisado, em virtude de estar estreitamente ligado ao radical deste, é de esperar que a marcação PL, por ser posterior à do género, possa ser condicionada pela primeira quando ela tem que ocorrer.

Esta variável independente não foi observada nos trabalhos sobre CPL-var de Guy (1981a), Scherre (1988), Lopes (2001), P. Andrade (2003), Baxter (2004, 2009) e Jon-And (2008, 2009). Contudo, Lucchesi (2000a:264), em estudo sobre a concordância variável de género de HEL-Ba, advogou o Princípio da Coesão Estrutural (Haiman, 1983:782), constatando que, de facto, a marcação de género favorece categoricamente a marcação PL no constituinte analisado, enquanto a não-marcação de género se revela como um forte obstáculo à inserção do número. Na presente análise, pretendemos observar se, de facto, a marcação de género funciona como inibidora da marcação de número. Numa perspectiva generativista, a inserção de marca de género, por ocorrer primeiro (Müller, 1994a; Fransceschina, 2002, 2005; Corrêa, Name & Ferrari-Netto, 2004; Montrul, 2004; White *et alii*, 2004), serve como gatilho para despoletar também a marcação de número. Nesta conformidade, e conforme observado no ponto 2.12 do presente trabalho, o que se espera é que, caso ocorra marcação de número, aconteça também inserção de marca de género.

Em trabalho sobre a aquisição da concordância de PL no SN por aprendentes chineses de PtgL2 em contexto formal de ensino, Godinho (2005:345) observou o efeito desta variável na inserção da marcação PL e confirmou o Princípio de Coesão Estrutural (Haiman, 1983:782), visto que a não-expansão da cadeia da concordância nominal no SN inclui ambas as concordâncias, de género e de número, cujos traços nucleares pertencem à posição SNUM, uma categoria intermédia entre o SDET e o SN (Baxter,



Lucchesi & Guimarães, 1997:2). Deste modo, e do ponto de vista aquisicional, sendo a identificação dos traços de género e número conseguida através da concordância no SDET, a não-interpretação de um dos traços pode ter influência na não-interpretação do outro, o que implica que a categoria intermédia (SNUM) não esteja totalmente adquirida. E no caso dos aprendentes chineses, tal como nos falantes de L1's africanas do grupo níger-congo atlântico, a explicação poderá assentar no facto de a concordância nominal no português não configurar um parâmetro das suas L1's e, conforme preconizado pela FFFH, ocorrer falha no restabelecimento dos parâmetros (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), uma vez que as categorias funcionais deixam de estar disponibilizadas depois do período crítico do desenvolvimento. Por seu lado, também não deve ser descartada a possibilidade de, tal como advogam os defensores da FTFA, os falantes do PtgL2 deterem uma compreensão e um julgamento metalinguístico do género gramatical distinto daquele que acontece na LA, uma vez que determinadas línguas do grupo níger-congo atlântico, pelo facto de não possuírem género gramatical, fazem com que os seus falantes em processo de aquisição de uma L2 detenham já um conhecimento cognitivo-linguístico prévio que vai impedi-los de adquirirem completamente as propriedades relevantes da L2 (cf. Gonçalves, 2004:235; Franceschina, 2005:191).

A fim de elaborarmos os factores para o grupo de factores *marcação de género*, optámos por seguir três critérios distintos, de acordo com os seguintes parâmetros de flexão em género:

1. O item leva correcta marcação de género:

- a) itens biformes regulares, com correcta inserção de marcação (p.e. *trabalhos PRIVADO* [CASTEH1]);
- b) itens biformes irregulares, com correcta inserção de marcação (p.e. *BONS pescadores* [CELESH3]);
- c) itens uniformes com marca padrão de género e correcta inserção de marcação (p.e. *vários TRABALHOS* [CASTEH1]; *nas LOJAS* [CARMOM1]).

2. O item não leva correcta marcação de género:
- itens biformes regulares, com incorrecta inserção de marcação (p.e. *trabalhos CASEIRA* [CARMOM1]);
  - itens biformes irregulares, com incorrecta inserção de marcação (p.e. *ESSES fruta* [ANTOM1]);
  - itens uniformes com marca padrão de género e incorrecta inserção de marcação (p.e. *uns cem CONTAS*<sup>155</sup> [CELESH3]).
3. Itens uniformes sem marca padrão de género:
- itens uniformes com marcação única de género fornecida pelo determinante (p.e. *os MATERIAL* [ABILH2]; *essas POSSIBILIDADE* [CASTEH1]);
  - itens uniformes com aparente marca padrão de género, isto é, não correspondência entre género do determinante e terminação do item (p.e. *os DIAS* [CASTEH1]; *outras minhas PREOCUPAÇÕES* [CASTEH1]);
  - itens uniformes com dupla marcação de género fornecida pelo determinante (p.e. *todos JOVEM* [MMDEUSM3]).

A Tabela 12.8 apresenta os factores constituídos, suas codificações, número de ocorrências e percentuais de marcação:

**Tabela 12.8.** Factores constituídos para a variável independente *marcação de género*.

9) Variável independente <i>marcação de género</i>				
Cód.	Factor	Exemplos	Nr. de Ocorrências	%
c	O item leva correcta marcação de género	<b>trabalhos</b> PRIVADO [CASTEH1] BONS <b>pescadores</b> [CELESH3] vários <b>TRABALHOS</b> [CASTEH1] nas <b>LOJAS</b> [CARMOM1]	1.060/2.030	52
n	O item não leva correcta marcação de género	<b>trabalhos</b> CASEIRA [CARMOM1] ESSES <b>fruta</b> [ANTOM1] uns cem <b>CONTAS</b> [CELESH3]	38/59	64
/	Itens uniformes sem marca padrão de género	os <b>MATERIAL</b> [ABILH2] essas <b>POSSIBILIDADE</b> [CASTEH1] os <b>DIAS</b> [CASTEH1] outras minhas <b>PREOCUPAÇÕES</b> [CASTEH1] todos <b>JOVEM</b> [MMDEUSM3]	104/251	41
<b>Totais:</b>			1.202/2.340	51

**Obs.:** O factor marcação de género aplica-se a todos os elementos do SN, de acordo com o traço semântico do núcleo deste.

Para confirmar, ou não, os pressupostos do Princípio de Coesão Estrutural, levantámos a seguinte hipótese:

**Hipótese 8 – Marcação de género:** as probabilidades de não-marcação de número são maiores caso não ocorra marcação de género. Em contrapartida, a aquisição de marcação de género beneficia a marcação de número, já que os traços de género e número se encontram dentro do mesmo nó funcional SNUM.

### **3.6.8. Variável independente *ordem do item na cadeia dos constituintes flexionáveis em número***

Se o grupo de factores *posição linear* permite verificar a frequência com que certo item, independentemente da sua classe gramatical, recebe a marca de PL em determinada localização da cadeia sintagmática, já a variável independente *ordem do item na cadeia dos constituintes flexionáveis* surge, neste trabalho, com o intuito de observar o comportamento dos elementos flexionáveis nesse mesmo SN. Note-se que, o facto de um elemento ser o primeiro item flexionável do SN não implica, necessariamente, que esteja a ocupar a primeira posição linear da cadeia sintagmática, conforme se pode constatar no exemplo usado para o factor “P” da Tabela 12.9. Assim, e dado que na cadeia sintagmática se inserem itens passíveis e não-passíveis de flexão, poder-se-á observar se a inibição que ocorre em elementos posteriores é consequência da inserção de itens não-flexionáveis entre os elementos flexionáveis da cadeia sintagmática, não se observando o Princípio da Coesão Estrutural (Haiman, 1983:782), em virtude de se cortar o encadeamento da concordância. Relembre-se que Guy (1981b:179) advogou que o efeito da *posição linear* e *classe gramatical* são idênticos, pois parece existir uma equivalência entre ambos os grupos de factores, visto os determinantes ocorrerem preferencialmente na primeira posição, os nomes se situarem usualmente na segunda posição e os adjectivos ocorrerem, por norma, na terceira posição. Discordando do postulado de Guy (1981b:179), Scherre (1988:167) não só opinou que a marcação PL ocorre preferencialmente na primeira localização, independentemente da categoria gramatical que aí se posicione, mas também que a pluralização do SN é determinada pelo processamento paralelo (Scherre, 1988:208), isto

é, pelo princípio de que marcas conduzem a mais marcas e zeros implicam a ocorrência de mais zeros (Scherre, 1988:235).

Assim sendo, a variável independente *ordem do item na cadeia dos constituintes flexionáveis*, que não temos conhecimento de ter sido observada em outros trabalhos sobre concordância variável, analisada *per se* ou em relação com outros grupos de factores, como as variáveis linguísticas *posição linear* ou *classe gramatical* e a variável social *idade*, poderá ajudar na clarificação, entre outras, de questões como as que se seguem: (i) Marcas conduzem a marcas no item imediatamente posterior? (ii) Os itens não-passíveis de marcação afectam a cadeia de concordância quando inseridos entre elementos passíveis de marcação? (iii) Qual a posição preferencial dos diferentes itens gramaticais flexionáveis na cadeia do SN?

Para elaboração do grupo de factores *ordem do item na cadeia dos constituintes flexionáveis* constituíram-se os factores constantes na tabela seguinte:

**Tabela 12.9.** Factores constituídos para a variável independente *ordem do item na cadeia dos constituintes flexionáveis em número*.

<b>11) Variável independente <i>ordem do item na cadeia dos constituintes flexionáveis em número</i></b>				
<b>Cód.</b>	<b>Posição</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>P</b>	Primeiro elemento	duzentos e tal <b>CONTO?</b> [SALVH3]	886/1.487	60
<b>S</b>	Segundo elemento	todos OS <b>dias</b> [ANTOM1]	286/754	38
<b>T</b>	Terceiro elemento	todos os <b>DIAS</b> [ANTOM1]	28/85	33
<b>Q</b>	Quarto ou outro elemento	uma outras minhas <b>PREOCUPAÇÕES</b> [CASTEH1]	(1/10)	10
<b>Totais:</b>			1.201/2.336	51

Esta variável independente permitirá avaliar a seguinte hipótese:

**Hipótese 9 – *Ordem do item na cadeia dos constituintes flexionáveis em número*:** a inserção de itens não flexionáveis em número entre os elementos flexionáveis que vêm manifestando concordância confirma a aplicação do Princípio do Processamento Paralelo (Scherre, 1988:208), já que actua como elemento condicionador da marcação nos itens posteriores, cortando a cadeia de concordância e confirmando o pressuposto de que ausência de marcas leva à não inserção de mais marcas.

### 3.6.9. Variável independente grau de concordância de número no SN

A constituição desta variável independente visou permitir que se observasse, numa perspectiva essencialmente geracional, quais os tipos de SN's que são produzidos pelas diferentes faixas etárias de Almojarife. De facto, atendendo ao número de elementos flexionáveis inseridos nas cadeias sintagmáticas, é possível não só tirar conclusões acerca da extensão e grau de complexidade das mesmas mas também visionar quais são as gerações que as produzem. No entanto, devemos referir que este grupo de factores se adequa mais a um estudo não-atomístico, conforme realizado por P. Andrade (2003:106-107), pelo que nos limitaremos a fornecer o número de ocorrências e fazer uma breve análise sobre as mesmas.

Para a elaboração do grupo de factores grau de concordância de número no SN constituíram-se os seguintes factores:

**Tabela 12.10.** Factores constituídos para a variável independente grau de concordância de número no SN.

<b>12) Variável independente grau de concordância de número no SN</b>				
<b>Cód.</b>	<b>Factor</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>1</b>	SN's com 1 elemento flexionável no SN	duas <b>PISSOAS</b> [CASTEH1]	204/729	28
<b>2</b>	SN's com 2 elementos flexionáveis no SN	<b>TRABALHOS PRIVADO</b> [CASTEH1]	865/1.352	64
<b>3</b>	SN's com 3 elementos flexionáveis no SN	UMAS nove ou onze <b>JOVEM FORMADO</b> [CASTEH1]	125/231	54
<b>4</b>	SN's com mais de 3 elementos flexionáveis no SN	UMA OTRAS MINHAS <b>PREOCUPAÇÕES</b> [CASTEH1]	(8/28)	29
<b>Totais:</b>			1.202/2.340	51

Partindo do pressuposto que as diferentes faixas etárias de Almojarife exibem gramáticas distintas, a variável independente grau de concordância de número no SN permitirá avaliar a seguinte hipótese:

**Hipótese 10 – Grau de concordância de número no SN:** a faixa etária mais idosa de Almojarife produz SN's de estrutura mais reduzida e mais baixo grau de complexificação estrutural. Assim sendo, os SN's vão-se tornando mais extensos e

complexos com as novas gerações, confirmando-se uma aquisição geracional da regra de concordância.

### **3.6.10. Variável independente traço semântico do SN (*Animacidade*)**

No estudo dos seus dados, Scherre (1988:266) observou que a incidência de marcação PL sobre o substantivo “pessoa” (82%) era notoriamente superior à que recaía sobre o nome “coisa” (36%). Diversos trabalhos linguísticos (Ferraz, 1979; Lucchesi, 2000a; Alexandre & Hagemer, 2007; Figueiredo, 2009a) têm atribuído ao traço [+humano], por ser discursivamente mais saliente (Mollica, 1977:37), um papel relevante em vários tipos de fenómenos que ocorrem em variedades de português (p.e. a conservação da flexão de género em nomes adjectivados do tipo *macho/fêmea* ou *homem/mulher*). Assim sendo, Scherre (1988) levantou a hipótese de os itens nominais que possuem este tipo de traço apresentarem maiores índices de concordância.

Note-se que, anteriormente aos achados de Scherre (1988), já Mühlhäusler (1981), ao analisar dados do *pidgin* tok pisin, havia constatado que as propriedades dos nomes com traço semântico [+animado] poderiam ter influência na marcação de número. Ao que tudo indica, estas propriedades semânticas interagem com a sua função sintáctica, levando a que os nomes em posição de sujeito revelem maior tendência a serem marcados, uma vez que nessa posição surgem, maioritariamente, agentes do tipo animado (Young, 1991:111). Posteriormente, Sabino (1983) verificou também a relação entre nomes de traço semântico [+animado] e a função sintáctica de sujeito no crioulo de base inglesa das Ilhas Virgens, concluindo que este traço favorece a marcação de número, enquanto o traço [-animado] a inibe.

Godinho (2005) procurou também verificar se os nomes de traço [+animado] favoreciam a marcação PL nos SN's produzidos por aprendentes chineses de PtgL2 em contexto formal de ensino, mas a ferramenta VARBRUL não considerou esta variável significativa para a observação da CPL-var dos mesmos.

Por seu lado, Jon-And (2009:7) constituiu cinco factores para analisar a influência da variável independente *animacidade* na marcação PL do PCV, uma variedade que tem como substrato quer o crioulo de São Vicente (as gerações mais novas e parte dos falantes mais idosos) quer o crioulo de Santo Antão (os restantes falantes mais velhos).

Assim, além dos nomes de traço semântico [+humano], [+animado] [-humano] e [-animado], a linguista criou um factor para substantivos referentes a partes do corpo e outro para plantas. Os resultados apresentados por Jon-And (2009) permitem concluir que o factor [+humano] favorece a marcação, enquanto o factor [-animado] a inibe. Quanto aos outros três factores, exibem um reduzido número de ocorrências, pelo que os seus pesos não foram considerados pelo programa VARBRUL.

Pesquisas sobre concordância dentro do SN de línguas parcialmente reestruturadas têm também trazido à evidência que existe um paralelo entre os padrões de variação das marcações mórnicas de PL e de género. Em ambas as situações, os constituintes localizados antes do núcleo do SN revelam maior tendência a serem marcados do que os posicionados pós-nuclearmente. Apesar de os fenómenos se assemelharem, têm sido justificados distintivamente. Assim, no caso da CPL-var, tem vindo a ganhar força a possibilidade de as L1's africanas exercerem influência, conjuntamente com as dos crioulos, na configuração do SN do PtgL2 ou dos CP's. Posteriormente, caso estes dois tipos de línguas entrem em contacto, a partilha de traços comuns a ambas pode determinar a incorporação dos mesmos na nova língua que emerge (Clements, 2009:19). Assim, no caso da CPL-var registada no PA, a mesma pode estar sujeita a condicionantes que ancoram na aquisição do PtgL2 e PtgL1 em situação de contacto que envolveu estruturas de PtgL2 e CP (o santomense) influenciadas pelas línguas ancestrais africanas (cf. Guy, 1981a; P. Andrade, 2003; Baxter, 2004, 2009; Figueiredo, 2008, 2009a, 2009b; Jon-And, 2008). Idêntica situação ocorre com a CGEN-var, cujas condicionantes ancoram também nas línguas africanas, tendo estas determinado que o núcleo do SN seja o item responsável pela fixação da morfologia flexional de género e irradiação da concordância de género (Lucchesi, 2000a:189; Figueiredo, 2009a:50), indo esta marcar mais fortemente os elementos adjacentes ao núcleo, especialmente os situados imediatamente antes deste. Como tal, a inibição à marcação do género cresce à medida que os itens se encontram posicionados mais longe do núcleo na cadeia sintagmática ou à direita do mesmo (ponto 2.6.4.2 do presente trabalho).

A premissa que determina o núcleo do SN como pólo irradiador da marcação mórnic de género estabelece ainda aproximações com algumas realizações que ocorrem nos CP's da Alta Guiné e do Golfo da Guiné (Lucchesi, 2000a:189), dado que, nestes, é

também a partir do núcleo do SN que se recompõe a morfologia do género. Nesta conformidade, é visível um paralelo no plano da morfossintaxe de género entre estes CP's, o PA e o dialecto de HEL-Ba, já que, em todos eles, se constata a existência de ocorrências que revelam não-aquisição de matéria gramatical da LA, com a erosão da morfologia flexional do nome a afectar mais o mecanismo morfossintáctico da concordância do que a flexão mórfica do género dos núcleos nominais (cf. Chomsky, 1996 [1995], 2001a; Carstens, 2000) que dizem respeito a seres animados (Lucchesi, 2000a:337-338). Deste modo, com o apagamento do mecanismo morfossintáctico de função meramente gramatical para indicação do género, conservou-se a flexão deste nos nomes adjectivados do tipo *macho/fêmea* ou *homem/mulher*. No caso concreto do santomense (exemplos [53a] e [53b]) e do CCV (exemplo [54]), o valor referencial de sexo ocorre, por norma, com os referidos substantivos adjectivados sendo colocados pós-nuclearmente ao nome, que se mantém invariável na forma masculina.

Dado que este tipo de ocorrências é também recorrente no PA (exemplos [55a] e [55b]), e como em termos generativistas a causa do acontecimento pode ser atribuída ao apagamento dos traços não interpretáveis dos elementos funcionais ou não-parametrizados, adquiridos em fase pós-puberdade (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), decidimos verificar se o fenómeno da animacidade afecta também o mecanismo morfossintáctico da concordância PL.

Todavia, e independentemente da questão do paralelismo entre os padrões de variação das marcações mórficas PL e género e da possibilidade de incidência de marcação PL sobre o substantivo “pessoa” ser superior à que recai sobre o nome “coisa” (Scherre, 1988:266), o estudo da variável independente *traço semântico do SN* levantou a hipótese de o desenho da CPL-var no SN do PA ser mais condicionado por questões do tipo estrutural do que de carácter semântico, como se verá na secção dedicada à análise dos resultados desta variável (ponto 5.1.1). A fim de se ajudar a trazer mais clarificações a todas estas questões, codificaram-se os substantivos em termos da variável da animacidade do seu traço semântico, constituindo-se o grupo de factores *traço semântico do SN (animacidade)*, ao qual se incorporaram os três factores constantes na Tabela 12.11. Ainda a propósito da codificação destes itens, refira-se que elementos



gramaticais como o colectivo “equipas” foram considerados como portadores do traço semântico [+humano]. Quanto aos resultados finais, eles poderão também ajudar-nos a confirmar o pressuposto de o fenómeno da CPL-var, ao contrário da CGEN-var, não ter como pólo irradiador o núcleo do SN, mas sim a categoria funcional DET, na qual o PL é inserido através de um morfema do tipo *singleton*, o qual serve de âncora para o controlo da pluralização (Baxter, 2009:292).

Para a variável traço semântico do SN, colocamos a seguinte hipótese:

**Hipótese 11 – Traço semântico do SN:** Os nomes com traço semântico [+animado] favorecem a marcação, por oposição aos substantivos com traço semântico [–animado], que revelam tendência para desfavorecê-la.

**Tabela 12.11.** Factores constituídos para a variável independente *traço semântico do SN*.

13) Variável independente <i>traço semântico do SN</i>				
Cód.	Traço	Exemplo	Nr. de Ocorrências	%
<b>h</b>	[+humano]	os <b>HOMENS</b> [CASTEH1]	747	60
<b>a</b>	[–humano] [+animado]	esses <b>GAROUPA</b> assim [ZECAH1]	75	37
<b>z</b>	[–animado]	<b>TRABALHOS</b> privado [CASTEH1]	1.501	48
<b>Totais:</b>			1.201/2.338	51

### 3.6.11. Variável independente *sexo*

Em trabalhos de sociolinguística quantitativa sobre variação, as variáveis independentes do tipo social, apesar de incidirem sobre dados sincrónicos, surgem também intimamente ligadas à noção de prestígio, a qual, por sua vez, pode determinar comportamentos linguísticos conotados a mudanças em progresso. A discussão em torno do modo como a variável *sexo* influencia as mudanças em curso não tem sido consensual. Os estudos pioneiros com recurso a esta variável social (Chambers & Trudgill, 1980:91-93) apresentaram as mulheres como impulsionadoras das mudanças em direcção às formas de prestígio, enquanto os homens seriam os responsáveis por fenómenos de não-realização, já que apresentavam maior índice de ocorrências que se afastam dos registos padronizados. Contudo, Labov (1978:82) discordaria deste pressuposto, uma vez que atribuiu às mulheres a responsabilidade por todos os processos

de mudança. Quanto a Scherre (1988:429), relembra que o papel da variável *sexo* na questão da mudança em curso não é muito claro na literatura em linguística. Ainda assim, os falantes urbanos do MRJ, que produziram as ocorrências que Scherre (1988) observou, pareciam confirmar o pressuposto levantado por Chambers & Trudgill (1980), dado que são as mulheres quem favorece a marcação de número, enquanto os homens a inibem. A possível tendência para o favorecimento da marcação em número revelada pelo estrato masculino de comunidades urbanas seria ainda constatada por Lopes (2001:109), quando se debruçou sobre a fala culta de São Salvador.

Em contrapartida, Baxter, Lucchesi & Guimarães (1997), ao estudarem a CGEN-var das produções dos falantes rurais de HEL-Ba, apontaram uma tendência contrária para este fenómeno, com os homens a propiciarem mais a marcação do que as mulheres. Por seu lado, observando as tendências de mudança nas normas cultas e popular do PB, Lucchesi (1998a) chama a atenção para a falácia que generalizações como a de Chambers & Trudgill (1980) podem encerrar, quando afirmam que as mulheres tendem a dirigir as mudanças em direcção às formas de prestígio. Confirmando a tendência de mudança registada por Baxter, Lucchesi & Guimarães (1997) para HEL-Ba, Lucchesi (2000a:281) reafirma que não se pode generalizar qualquer tendência com vista a atribuir uma maior ou menor influência na marcação a um dos sexos, porque *“não apenas não é muito claro, como também constitui um grande equívoco considerar essas variáveis sociais como parâmetros isolados, sem se considerar um contexto mais abrangente, no qual esses elementos podem ser integrados numa interpretação totalizadora, capaz de dar conta da especificidade de cada processo histórico particular de mudança lingüística.”*. Nesta perspectiva, o estudo compartimentado das variáveis independentes, sejam elas linguísticas sejam elas sociais, pouco contributo dá para a observação dos processos de mudança registados numa determinada comunidade, enquanto todo caracterizado por determinado padrão linguístico. Portanto, só é possível concluir que certos fenómenos se encontram em situação de variação ou processo de mudança caso, a partir das evidências empíricas concedidas por cada uma das variáveis, o pesquisador consiga perceber e justificar, com base quer em fundamentações teórico-científicas já trazidas à evidência quer em inferências coerentemente argumentadas, a forma como tais evidências se articulam entre si na construção do todo coerente que é o

registro linguístico do grupo em questão. Significa isto dizer que a observação da variação e mudança da língua só é possível caso se tenha em conta que a língua é algo heterogéneo, determinado tanto por condicionantes linguísticas inerentes à sua própria estrutura como por fenómenos sociais externos, tal como o preconizam os preceitos da sociolinguística variacionista. E, neste aspecto, também os resultados de P. Andrade (2003:129-130) acabariam por evidenciar a fragilidade das generalizações avançadas por Chambers & Trudgill (1980), já que a autora confirmaria, em definitivo, que, na comunidade de HEL-BA, não são as mulheres, mas sim os homens a liderar a mudança em direcção à norma padrão, tendo, assim, a responsabilidade maior na fixação das regras de concordância. Resta ainda referir que, trabalhando igualmente com dados de informantes rurais, Baxter (2009:285) verificaria igual tendência para o PT, apresentando valores em tudo semelhantes aos de P. Andrade (2003:130).

No caso do PA, a Tabela 12.12 permite observar que as ocorrências passíveis de análise produzidas por homens (1.323) e mulheres (1.017) apresentam valores bastante aproximados. No nosso estudo, os falantes encontram-se equitativamente distribuídos por escalões etários, não se registando também grandes variações de escolarização a nível dos mesmos estratos de idade. No entanto, os falantes do sexo masculino terão tido um contacto mais estreito com o PE em determinada fase de vida da comunidade, já que eram os homens que migravam para os centros urbanos em busca de oportunidades de emprego, por um lado, e eram também integrados nos quadros do exército colonial, por outro lado.

**Tabela 12.12.** Factores constituídos para a variável independente *sexo*.

<b>14) Variável independente <i>sexo</i></b>				
<b>Cód.</b>	<b>Sexo</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>H</b>	Masculino	faço <b>TRABALHOS PRIVADO</b> [CASTEH1]	709/1.323	54
<b>M</b>	Feminino	vinho agora stá <b>DEZ CONTO</b> [MMDEUSM3]	493/1.017	49
<b>Totais:</b>			1.202/2.340	51

Tendo em conta que os aspectos mencionados, formulámos a seguinte hipótese para a variável *sexo*:

**Hipótese 12** – **Sexo**: Ambos os sexos devem apresentar um nível semelhante de favorecimento da pluralização no SN. Contudo, não é de excluir a possibilidade de os falantes do sexo masculino, por condicionantes sociolinguísticas, apresentarem ligeira tendência para realizarem mais a concordância em número do que os falantes do sexo feminino.

### **3.6.12. Variável independente *idade***

A análise da variável *idade* permite detectar possíveis mudanças linguísticas em curso na comunidade, verificáveis através do chamado *tempo aparente* (Labov, 1972a). É o que sucede, por exemplo, quando esta variável apresenta uma distribuição curvilínea, a qual revela que se está perante uma variação estável, isto é, um estágio de evidente mudança consumada. Em contrapartida, se a variável *idade* apresentar uma distribuição inclinada, com o ápice da aquisição da regra de concordância na faixa etária mais nova e o maior índice de variação na faixa etária mais velha, significa que a comunidade ainda se encontra em fase de mudança em progresso (Chambers & Trudgil, 1980; Baxter, Lucchesi & Guimarães, 1997). Contudo, importa referir que a variável *idade*, por si só, não é suficiente para indicar a mudança em progresso, pelo que é necessário recorrer a outros padrões sociais associados a ela, como o sexo dos falantes, os seus anos de escolarização e até o estrato social em que estão inseridos.

De acordo com a metodologia que tem sido aplicada em outros trabalhos sobre análise da concordância variável manifestada por diferenciados grupos linguísticos (Lucchesi, 2000a; P. Andrade, 2003; Baxter, 2004; 2009), distribuímos equitativamente os informantes do PA por faixas etárias distintas: (i) dos 20 aos 40 anos (faixa etária FE-1); (ii) dos 41 aos 60 anos (faixa etária FE-2); (iii) mais de 60 anos (faixa etária FE-3). Ao considerarmos estas diferentes gerações de informantes, é-nos possível estudar a covariação sistemática entre variáveis linguísticas e sociais, verificando em qual delas ocorre a maior frequência de formas inovadoras. Paralelamente, ao contemplarmos um período de pelo menos 20 anos para cada uma das faixas etárias, garante-se, em tempo real, a possibilidade de uma projecção para um período de tempo aparente que é, segundo os especialistas em estudos variacionistas,<sup>156</sup> suficiente para determinar a cristalização da variação das três gerações e confirmar a mudança no sentido da fixação

da regra de concordância de número. Esta, por seu lado, vai contribuir para o desenho da gramática da comunidade de Almojarife.

Dado que alguns trabalhos sobre CPL-var apresentam os informantes estratificados por escalões etários distintos dos nossos (Scherre, 1988; Lopes, 2001), remetemos para o capítulo 5 (análise dos dados das variáveis extralinguísticas) os considerandos apresentados em pesquisas anteriores acerca da variável independente *idade*, nomeadamente quando compararmos os nossos resultados com os de Scherre (1988), Lopes (2001), P. Andrade (2003) e Baxter (2004, 2009). Os factores que constituímos para a variável independente *idade* estão patentes na tabela seguinte:

**Tabela 12.13.** Factores constituídos para a variável independente *idade*.

<b>15) Variável independente <i>idade</i></b>				
<b>Cód.</b>	<b>Escalão etário</b>	<b>Exemplo</b>	<b>Nr. De Ocorrências</b>	<b>%</b>
1	FE-1 (de 20 a 40 anos)	faço TRABALHOS PRIVADO [CASTEH1]	623/984	63
2	FE-2 (de 41 a 60 anos)	são ESSES TONGA de lá. [ABILH2]	373/786	34
3	FE-3 (+60 anos)	vinho agora stá DEZ CONTO. [MMDEUSM3]	206/570	24
<b>Totais:</b>			1.202/2.340	51

Para a variável independente *idade* estabelecemos a seguinte hipótese:

**Hipótese 13 – *Idade*:** A idade e a aquisição da regra de concordância estabelecem uma relação directamente proporcional, com as gerações mais novas a favorecerem cada vez mais a marcação.

### **3.6.13. Variável independente *estadia* (*Permanência fora da comunidade*)**

A integração da variável *estadia* no nosso estudo tem a ver com o facto de muitas mudanças ocorrerem em virtude da “elasticidade” da própria língua, que aceita incorporações de outros dialectos. Estas poderão ser trazidas por falantes de dialectos distintos que visitem regularmente uma determinada comunidade, ou passem a viver integrados nela. Em complementação, as incorporações poderão ser também trazidas por membros da própria comunidade que se desloquem para fora desta por prolongados períodos de tempo e regressem com formas novas integradas no seu idiolecto, as quais são depois transmitidas aos outros elementos da sua comunidade.

O contacto com variedades urbanas, normalmente detentoras de um padrão de concordância mais rico, contribui para aumentar a inserção de marcas. A exemplo do que vem sendo adoptado em outros trabalhos sobre concordância variável (Lucchesi, 2000a; P. Andrade, 2003), dividimos os informantes do PA em duas categorias: (i) falantes que se ausentaram da comunidade durante pelo menos um período de seis meses; (ii) falantes que nunca se ausentaram da comunidade ou, caso o tenham feito, não tenham permanecido fora desta por períodos superiores a seis meses. As ausências dos membros de Almojarife são normalmente determinadas por questões laborais, já que alguns dos falantes emprestaram, sazonalmente, mão-de-obra às roças ou ter-se-ão deslocado a povoações mais expressivas ou mesmo a São Tomé, aí exercendo actividades distintas (p.e. construção civil, alfaiataria, etc.) ou servindo no exército colonial. Casos há em que os falantes estiveram em países vizinhos da região do Golfo da Guiné, como o Gabão, aonde se deslocaram para mercadejar, tendo aí prolongado a sua estadia para exercer diversas actividades. Uma das informantes regista também viagens e estadias longas em Angola, país ao qual se deslocava para visitar um filho. Como se verifica, os casos de ausência prolongada não implicaram um contacto mais próximo com a LA (cf. Tabela 8.1), pelo que as mesmas não terão sido um factor favorecedor da aquisição da regra de concordância.

A Tabela 12.14 mostra-nos os factores estabelecidos para a variável independente *estadia*:

**Tabela 12.14.** Factores constituídos para a variável independente *estadia*.

<b>16) Variável independente <i>estadia</i></b>				
<b>Cód.</b>	<b>Ausência</b>	<b>Exemplo</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
<b>S</b>	Sim	tem DOIS GÉMIA de três ano [LUISH2]	358/719	50
<b>N</b>	Não	faço TRABALHOS PRIVADO [CASTEH1]	844/1.621	52
<b>Totais:</b>			1.202/2.340	51

Com a constituição da variável independente *estadia*, procuramos confirmar a hipótese seguinte:

**Hipótese 14 – *Estadia*:** Na comunidade de Almojarife, o factor ausência não contribuiu para que falantes produzam padrões da concordância PL mais próximos dos da LA.

### 3.6.14. Variável independente *escolaridade*

A propósito da variável *idade* (ponto 3.6.12), referimos que as variáveis extralinguísticas, observadas por si só, não são suficientes para indicar a mudança em progresso, sendo necessário recorrer a outros padrões sociais associados a ela, como os anos de escolarização dos falantes ou o estrato social em que estes estão inseridos. Socialmente, a classe alta, sendo por norma escolarizada, utiliza mais as formas padronizadas, que são também sinónimo de prestígio. Por seu lado, as classes média-baixa, normalmente analfabetas ou com baixo índice de escolaridade, apresentam um registo marcado por traços linguísticos populares ou vernáculos. Assim sendo, ocorre uma tensão entre os registos das duas categorias de classes, com a primeira reflectindo uma variação estável e a segunda liderando os processos de mudança (Labov, 1982:77-78).

No ponto 3.6.11, a propósito da discussão para levantamento da hipótese a observar através da variável independente *sexo*, tecemos comentários sobre a forma como o estudo compartimentado das variáveis independentes pouco contributo dá para a observação dos processos de mudança, o que pode também ser confirmado pela variável social *escolaridade*. A constituição deste grupo de factores tem como condição primeira observar se os falantes menos escolarizados usam com menor incidência as formas-padrão do que aqueles que possuem um grau de escolarização mais elevado. De facto, é tido por aceite que a difusão dos modelos normativos concorre para o processo de mudança linguística, uma vez que não só expõe mais os falantes à norma culta da língua falada mas também os submete ao contacto com a língua padrão escrita. No entanto, convém ressaltar aqui que tanto a exposição como o contacto referido não se dão só em ambiente escolar. Com efeito, contingências como o tipo de actividade profissional exercida, a incorporação ou não de elementos do sexo masculino nos quadros do exército colonial ou a deslocação do falante para fora da comunidade podem, também, determinar que o informante, apesar de se enquadrar em determinado factor do grupo de factores *escolaridade*, tenha compartilhado experiências sociolinguísticas que determinaram a expansão do seu repertório linguístico. E se tal acontecer, o informante possuirá um idiolecto mais próximo das formas padrão da LA, contrariando, assim, a tendência generalizada do factor em que se encontra incorporado. Como tal, para

contemplar a análise dos aspectos expostos, o grupo de factores *escolaridade* tem que ser observado em concomitância com todas as outras variáveis sociais, isto é, à luz da própria sociohistória do país, em geral, e da comunidade em observação, em particular.

Durante o levantamento dos dados do nosso *corpus* não deixou também de se ter em conta a questão da “representatividade”, isto é, da quantidade de informantes a incluir em cada factor social. Assim, quanto mais informantes se possuir para cada factor (e consequentemente para cada célula VARBRUL), mais se diluem as idiosincrasias, reduzindo-se, logicamente, a possibilidade de adulteração dos resultados finais. Entre os linguistas, a questão da “representatividade” tem sido alvo de algum debate no que concerne ao número mínimo de informantes tido como representativo para indicar a direcção das tendências linguísticas de uma determinada comunidade. O consenso tem apontado para um mínimo de quatro informantes a incluir em cada célula, embora alguns pesquisadores defendam que esse mínimo deverá ser de cinco informantes. A este propósito teremos que referir que a célula *escolarização alta (frequência do ensino pós-primário)*, da variável social *escolaridade*, apenas inclui três informantes, por razões que se prendem com a própria sociohistória da comunidade de Almocharife, sendo os seus membros na quase totalidade indivíduos de estrato social humilde, logo com manifesta dificuldade em acederem ao ensino. Contudo, todos os outros grupos de factores incluídos nas diferentes variáveis sociais possuem mais de cinco informantes, estando, assim, salvaguardada a possibilidade de os nossos resultados saírem viciados por questões idiosincráticas conectadas aos falantes.

Diga-se, por fim, que as variáveis sociais, apesar de serem independentes, partilham semelhanças com as variáveis dependentes, uma vez que são também entidades nominais, revelando-se descontínuas ao não aceitarem valores intermediários. Contudo, a propósito da descontinuidade de algumas destas variáveis, é necessário um esclarecimento. Se é um facto que a variável *sexo*, por exemplo, implica apenas duas alternativas possíveis (masculino *versus* feminino), já outras variáveis, como a *idade* e a *escolaridade*, pressupõem um *continuum* que, à primeira vista, as desenquadrará do modelo nominal. De facto, tanto uma como a outra implicam uma escala de continuidade no domínio social, representada por traços distintivos (jovem/meia idade/terceira idade...; não escolarizado/menos escolarizado/mais escolarizado...), que



poderão ser seleccionados em função de critérios operacionais, com vista a enquadrá-las em factores que abranjam um, e apenas um, grupo social específico. Assim, um determinado falante que se enquadre na faixa etária FE1 (20-40 anos), não pode ser enquadrado também, por exemplo, na FE3 (+60 anos).<sup>157</sup> Da mesma forma, um falante que tenha frequentado o ciclo preparatório não pode, simultaneamente, ser analfabeto.

A inclusão da variável *escolaridade* na nossa análise ficou a dever-se ao facto de a mesma permitir observar até que ponto a difusão de modelos normativos em São Tomé terão contribuído para a aquisição da regra da concordância de PL evidenciada pelo PA. Como se sabe, a frequência do ensino leva os falantes a prestarem mais atenção às regras de concordância, pelo que passam a aplicar mais as marcas de PL. Nesta conformidade, a constituição dos factores para a variável *escolaridade* surge condicionada não só pelos contextos sociais específicos de cada comunidade mas também pela forma como o ensino se encontra estruturado no país a que a comunidade pertence. E, neste aspecto, São Tomé e Príncipe constitui caso peculiar, como se verá aquando da análise dos dados respeitantes a esta variável social (capítulo 5).

Em trabalhos sobre CPL-var no SN, apenas Scherre (1988:445), Lopes (2001:107) e Jon-And (2009:4) apresentam resultados sobre a variável escolaridade. Contudo, como as duas primeiras se debruçaram sobre variedades urbanas do PB, a tendência será para os falantes destas serem mais escolarizados que os de Almojarife, uma comunidade rural. Paralelamente, o sistema organizativo do ensino no Brasil é distinto do de São Tomé e Príncipe ou Cabo Verde, tanto na época em que estes arquipélagos eram colónias portuguesas como na actualidade. Como tal, os critérios que subjazem à constituição dos nossos factores para a variável independente *escolaridade* divergem daqueles que foram levados em conta para observação da CPL-var das variedades brasileiras de português, mas assemelham-se aos elaborados por Jon-And (2009:4) para estudo do PCV. Não obstante, estes aspectos não negam a viabilidade operacional desta variável, dado que qualquer dos ensinos em questão contribuiu para uma maior exposição dos falantes à norma padrão do português. Considerações sobre os estudos de Scherre (1988), Lopes (2001) e Jon-And (2009), bem como comparações dos resultados dos mesmos com os nossos serão apresentados no capítulo 5, dedicado à análise dos

resultados das variáveis sociais. A tabela 12.15 apresenta os factores que constituímos para a variável escolaridade, para a qual estabelecemos a hipótese seguinte:

**Hipótese 15 – Escolaridade:** Quanto mais alto é o nível de escolaridade, mais atenção é prestada pelos falantes ao mecanismo da concordância. Como tal, os falantes mais escolarizados flexionam mais os itens pluralizáveis do SN.

**Tabela 12.15.** Factores constituídos para a variável independente *escolaridade*.

17) Variável independente <i>escolaridade</i>				
Cód.	Escolaridade	Exemplo	Nr. de Ocorrências	%
a	Analfabetos	tem DOIS <b>GÉMIA</b> de três ano [LUISH2]	187/596	31
s	Escolarização baixa (Frequência do ensino primário)	leva pa OUTROS <b>LUGAR</b> também [CARMOM1]	420/695	60
e	Escolarização média (Ensino primário concluído)	pissoa mandô AS <b>MOBÍLIA</b> [OSVALH1]	347/697	50
p	Escolarização alta (Frequência do ensino pós-primário)	faço <b>TRABALHOS PRIVADO</b> [CASTEH1]	248/352	71
<b>Totais:</b>			1.202/2.340	51

### 3.7. Codificação dos itens de acordo com os códigos atribuídos aos factores

A quantificação dos dados pela ferramenta VARBRUL só é possível caso os mesmos tenham sido codificados. A codificação, feita de acordo com os factores estabelecidos, visa confirmar, ou não, as hipóteses levantadas para estes. Não nos delongaremos sobre a forma de proceder às codificações, já que este aspecto se encontra bastante contemplado na leitura da especialidade (p.e. Young & Bayley, 1996; Guy, 1988; Pintzuk, 1988; Paolillo, 2002; Scherre & Naro, 2003; Guy & Zilles, 2007). No nosso caso, será suficiente referir que os dados foram codificados de acordo com as categorias e subcategorias seleccionadas como variáveis-alvo, constituindo cadeias de códigos como as que se apresentam seguidamente (SN constituído por dois elementos):

**Tabela 13.1.** Exemplo de cadeia de códigos resultante da codificação dos elementos do SN.

	<b>Cadeia de códigos:</b>	<b>Exemplo:</b>
1.	(0DPP'AU2cS2hH1Np	<b>jovens</b> HABILIDOSO [CASTEH1]
2.	(SYP/VC01/P2hH1Np	<b>JOVENS</b> habilidoso [CASTEH1]

- Obs.:** (i) O item analisado surge grafado em caixa alta (maiúsculas), estando o núcleo do SN marcado a negrito;  
(ii) O primeiro elemento da cadeia de códigos, depois do parêntesis, diz respeito a um dos factores do grupo de factores *variável dependente*, enquanto cada um dos restantes quinze itens se refere a um factor de uma das quinze variáveis independentes constituídas.

Note-se que o processo de codificação dos dados não está concluído quando se finaliza a atribuição dos códigos a todos os itens passíveis de marcação PL, já que, como veremos seguidamente, após processamento dos dados com recurso ao suporte computacional, alguns dos factores constituídos podem apresentar-se como irrelevantes para a observação da variação. É o que sucede, por exemplo, com os factores identificados pela ferramenta VARBRUL como *knockouts* (prefiguram 0% ou 100% de marcações) ou com os grupos de factores apontados como *singletons* (variáveis independentes que apresentam um único factor), que terão que ser retirados da análise.

Refira-se ainda que a recodificação dos dados pode também ser necessária, caso determinados factores apresentem pesos relativos próximos ou certas variáveis independentes actuem conjuntamente sobre a regra. Nestes casos, a situação resolve-se com a amalgamação de grupos de factores para se constituírem grupos de factores mais robustos ou, em alternativa, com a alteração do arquivo de condições, que fornece ao VARBRUL as informações sobre o modo como ele deve conduzir a análise dos dados, isto é, quais os grupos de factores que devem ser processados em cada uma das rodadas (Paolillo, 2002:58). Para tanto, combinam-se os grupos de factores com recurso à linguagem LISP (*List Programming*) (Young & Bayley, 1996:275), especialmente programada para uso informático, e submetem-se os dados a nova análise VARBRUL.

### **3.8. Quantificação dos dados para selecção das variáveis independentes e dos factores a serem incorporados na análise final**

Depois de se constituírem as variáveis independentes e seus factores, e se codificarem também, em função destes, todos os itens em que se supõe ocorrer marcação de número PL, introduziram-se os dados no suporte computacional e procederam-se às primeiras

rodadas VARBRUL. Estas não apresentaram resultados sólidos, revelando, desde logo, haver interferências quer entre as variáveis independentes quer entre factores dos grupos de factores, como aliás se esperaria, face ao elevado número de grupos de factores e factores constituídos.<sup>158</sup> Como tal, situações houve em que alguns grupos de factores foram simultaneamente seleccionados e rejeitados, enquanto outros não apresentaram qualquer tipo de resultados. Nestes casos, a metodologia propõe que se proceda a uma análise cuidada de todas as variáveis independentes e seus factores, quer aplicando testes de significância estatística dentro dos grupos de factores quer observando a “convergência” de cada rodada, a fim de fazer valer, uma vez mais, o princípio da Lei da Parcimónia e se eliminarem os factores ou mesmo grupos de factores que não apresentem valores estatisticamente significativos. A eliminação de determinados factores poderá então passar pela retirada dos mesmos dos dados a analisar ou, em alternativa, pela sua junção a outros factores, criando-se um novo factor amalgamado.

Tido como um método bastante eficaz para testar a significância de análises alternativas dentro dos grupos de factores, o cálculo de qui-quadrado permite observar se a amalgamação de factores dentro de um mesmo grupo de factores compõe um novo factor que possa ser incorporado à análise. Reduzindo-se os factores, um a um, vão-se efectuando rodadas VARBRUL, testando-se os sucessivos efeitos das alterações no *log-likelihood*, até se obter a melhor análise, isto é, aquela que recorre ao menor número de factores para explicar os dados. No entanto, é também possível amalgamar vários factores num único, conquanto tal se justifique estatística ou teoricamente. Nos capítulos 4 e 5, ao detalharmos os passos dados no sentido de tratar cada um dos grupos de factores constituídos, referiremos também se estes foram objecto de qualquer dos tratamentos anteriores, e porquê.

Quando a análise efectuada engloba mais de dez grupos de factores – como acontece no nosso caso –, a ferramenta VARBRUL gera arquivos com muitas células, o que implica uma sobrecarga de cálculos para a obtenção do melhor modelo, isto é, para conseguir a convergência nos valores mais adequados que modelam o conjunto de dados inseridos na rodada. Para efectuar os referidos cálculos, o VARBRUL recorre a um algoritmo que, em sequências cíclicas, modifica os valores dos pesos relativos para

tentar alcançar o modelo matemático que mais fiavelmente se aproxime da distribuição constatada nos dados inseridos.

Após cada sequência cíclica, o algoritmo estabelece comparações entre o último e o penúltimo ciclo para efectuar ajustes nos valores do modelo e melhorar este. Quando a diferença entre um ciclo e outro é menor do que o valor limite estabelecido pelo algoritmo, a ferramenta VARBRUL pára as rodadas porque não consegue melhorar mais o modelo. Neste caso, o programa alcançou a convergência do algoritmo, ou seja, alcançou o modelo que mais e melhor se aproxima daquele que é verificado na fala da comunidade. Atingida a convergência, a ferramenta VARBRUL permite não só imprimir os valores dos pesos e do *input* do último ciclo mas também verificar a quantidade de ciclos (“iterações”) a que o algoritmo recorreu até alcançar a convergência.

Quanto mais elevado é o número de variáveis independentes ou de factores inseridos no mesmo grupo de factores, mais sequências cíclicas necessita o algoritmo de efectuar para obter o melhor modelo, ou seja, alcançar a convergência. Para evitar que o algoritmo produza sequências infinitas na busca dos ajustes, o programa VARBRUL permite que aquele force ciclos até à vigésima iteração, mas pára na seguinte, independentemente dos valores alcançados entre o último e o penúltimo ciclo. Nestes casos, a convergência não é alcançada e compete ao pesquisador verificar se tal foi consequência do elevado número de factores ou grupos de factores, ou, pelo contrário, resultou da falta de ortogonalidade entre duas variáveis independentes que compartilham factores, isto é, factores que definem exactamente os mesmos dados.<sup>159</sup>

As situações de falta de convergência podem ser resolvidas reduzindo-se o número de variáveis independentes a serem incluídas nas rodadas VARBRUL, isto é, omitindo os grupos de factores não significativos, ou, em alternativa, efectuando cruzamentos entre os vários grupos (*cross-tabulations*) para verificar quais os factores envasados em grupos de factores diferentes e que poderão estar a compartilhar traços idênticos. No ponto seguinte do presente trabalho (5.2.4) damos conta da metodologia aplicada nas nossas rodadas como vista a eliminar iterações e permitir a convergência dos nossos dados. Ressalve-se ainda que, apesar de termos sempre procurado identificar e resolver a causa da falta de convergência, não foi possível, em certas situações, conseguir uma

resolução definitiva. Contudo, ainda que os resultados obtidos não se revelassem os melhores possíveis nestas situações, o problema não se apresenta como fatal, já que eles se aproximam bastante dos hipotéticos melhores valores, concedendo assim uma indicação fiável dos efeitos das variáveis (cf. Guy & Zilles, 2007).

### **3.8.1. Iteração entre grupos de factores**

A redução de factores dentro de um mesmo grupo de factores pode torná-lo mais robusto para a análise que se pretende efectuar. Contudo, os próprios grupos de factores podem revelar iterações entre si, levando a ferramenta VARBRUL a produzir resultados sem significância e a eliminar da análise variáveis independentes que poderão ser significativas para o estudo em causa, ou, em contrapartida, seleccionar grupos de factores com resultados viciados. Não obstante, cruzando-se grupos que compartilhem traços ou características, podemos constituir novos grupos de factores que se revelem ajustados e significativos, permitindo descobertas importantes.

Paralelamente, poder-se-á também dar o caso de alguns grupos de factores não serem ortogonais, ocorrendo incompatibilidade entre eles. Estas situações não são resolvidas pelo cruzamento de variáveis independentes, e a metodologia aconselha que se efectuem rodadas VARBRUL retirando das mesmas, ora um, ora outro grupo, a fim de a análise alcançar resultados fiáveis e que a tornem válida. Olhando para os grupos de factores constituídos para se efectuar a análise dos nossos dados, verifica-se que as variáveis independentes *posição do item na cadeia do SN (posição linear)*, *posição em relação ao núcleo do SN* e *ordem do item na cadeia dos constituintes flexionáveis* são constituídas por factores que compartilham características. De facto, todos eles comungam traços em termos de efeito de variação, pois reflectem uma determinada posição do item analisado na cadeia linear do SN. Contudo, a amalgamação de dois ou dos três grupos de factores num único factor iria impedir que se cruzasse, independentemente, qualquer um deles com outros grupos de factores, inviabilizando-se a obtenção de alguns achados que entendemos serem bastante significativos para o estudo da CPL-var no SN do PA.

Em estudos sobre a CPL-var no SN, efectuados até à data presente, não temos conhecimento de ter sido analisado a variável independente *posição do item na cadeia dos constituintes flexionáveis*. Por outro lado, o comportamento da variável

independente *classe gramatical* tem sido observado quer conjuntamente quer em cruzamentos com os grupos de factores *posição do item na cadeia do SN* e/ou *posição em relação ao núcleo do SN*. Ora, é nosso entendimento que a variável *classe gramatical* não deixa também de representar, ela mesma, tanto a posição linear como o posicionamento relativamente ao núcleo do SN. De facto, em SN's de estrutura reduzida, como acontece maioritariamente no PA, quando se fala em determinados itens gramaticais, está a inferir-se, logicamente, os dois citados posicionamentos que o elemento gramatical ocupa preferencialmente na cadeia sintagmática. Assim, as referências aos determinantes implicam, por norma, a indicação de elementos pré-nucleares situados na primeira posição linear, as alusões ao nome ou categorias substantivadas implicam, geralmente, a indicação de itens que ocorrem em posição nuclear e em segundo lugar na cadeia linear, e a menção a adjectivos, salvo poucas excepções, está a referir-se a constituintes posicionados pós-nuclearmente e em terceira colocação no SN. Como tal, a nossa inferência leva-nos a concluir que os quatro grupos de factores referidos terão que apresentar interacções caso sejam incluídos, total ou parcialmente, na mesma rodada. E assim sendo, retira-se robustez aos mesmos, levando-os a apresentarem resultados enviesados, em vez dos não-viciados. Aliás, a idêntica conclusão terá chegado Jon-And (2008:3) nas suas análises sobre a CPL-var do PMP, uma vez que excluiu das suas rodadas a variável *classe gramatical*, “*por causa de interferência com outras variáveis*” que a autora não explicita, mas que acreditamos serem precisamente a *posição linear* e a *posição em relação ao núcleo do SN*.

### **3.8.2. Variáveis independentes a serem incorporadas na análise final**

Face à necessidade de preservarmos as quatro variáveis independentes mencionadas, que, de certa forma, reflectem uma determinada posição do item analisado na cadeia linear do SN, mas conscientes também que as mesmas apresentariam iteração caso qualquer delas fosse incluída em simultâneo na análise com uma das outras, optámos por efectuar diversas rodadas, incluindo apenas uma das variáveis independentes em cada uma destas. Assim, para efeitos de análise geral de todos os grupos mantivemos apenas um destes grupos de factores em cada rodada, a fim de conservarmos como modelo a rodada que permitisse uma visão mais abrangente sobre a totalidade dos grupos de

factores, ou seja, aquela que eliminasse o menor número de variáveis independentes. No caso das rodadas que eliminaram igual número de variáveis independentes, levou-se em consideração, para efeitos de análise geral, a que apresentou o nível de significância mais baixo.

Seguindo esta metodologia, apuraram-se resultados não enviesados, tendo as rodadas eliminado os seguintes grupos de factores:

- (i) Variável eliminada na rodada em que se manteve o grupo de factores *classe gramatical: marcação de género*.
- (ii) Variáveis eliminadas na rodada em que se conservou o grupo de factores *posição do item em relação ao núcleo do SN: animacidade do nome e marcação de género*.
- (iii) Variáveis eliminadas na rodada em que se preservou o grupo de factores *posição do item na cadeia do SN (posição linear): marcação de género*.
- (iv) Variáveis eliminadas na rodada em que se manteve o grupo de factores *ordem do item na cadeia dos constituintes flexionáveis: tonicidade e marcação de género*.

A ferramenta VARBRUL considerou a variável linguística independente *marcação de género* como não-significativa para aplicação da regra da CPL-var no SN do PA em todas as rodadas. Assim sendo, este grupo de factores, pela sua nula influência na CPL-var do SN de Almojarife, não será objecto de análise no nosso trabalho.<sup>160</sup> Relativamente às outras variáveis independentes, é possível constatar que as rodadas (ii) e (iv) eliminaram dois grupos de factores, enquanto as rodadas (i) e (iii) eliminaram apenas a variável independente *marcação de género*. Assim sendo, passamos a considerar para efeitos de análise apenas as duas últimas rodadas. De entre estas, a rodada (i) é aquela que apresenta o nível de significância mais baixo (0,011 – Tabela 11.1). Deste modo, passaremos, doravante, a tê-la como modelo para observar o efeito das diferentes variáveis independentes na marcação PL do SN do PA. As três restantes rodadas, isto é, as que mantiveram, em separado, os grupos de factores *posição do item em relação ao núcleo do SN* (nível de significância = 0,036 – Tabela 16.1), *posição do item na cadeia do SN* (nível de significância = 0,038 – Tabela 21.3) e *ordem do item na cadeia dos constituintes flexionáveis* (nível de significância = 0,017 – Tabela 17.1), serão utilizadas quer para se fazer uma leitura do modo como os grupos de factores que



foram retirados da rodada anterior intervêm na marcação PL do SN do PA quer para ajudar a definir o quadro hierárquico de importância decrescente das diferentes variáveis independentes na mesma marcação. Assim, e após observarmos não só os valores das quatro rodadas em questão mas também das rodadas gerais em que se incluíram todos os grupos de factores, estabeleceu-se a escala decrescente de importância das variáveis independentes no desenho da CPL-var do SN do PA:

**Tabela 14.1.** Escala hierárquica decrescente das variáveis independentes seleccionadas como válidas para análise da CPL-var no SN do PA.

<b>Escala hierárquica</b>	<b>Variável independente</b>
01	Posição do item em relação ao núcleo do sintagma nominal
02	Ordem do item analisado na cadeia dos constituintes flexionáveis em número
03	Classe gramatical do item analisado
04	Escolaridade
05	Marcas precedentes ao item analisado
06	Saliência fónica (Processos morfofonológicos de formação plural)
07	Grau de concordância de número no sintagma nominal
08	Posição do item na cadeia do sintagma nominal (Posição linear)
09	Idade
10	Estadia (Permanência fora da comunidade)
11	Sexo
12	Tonicidade
13	Traço semântico do sintagma nominal (Animacidade do nome)
14	Contexto fonológico posterior

No presente capítulo debruçámo-nos sobre a metodologia aplicada ao estudo da variação, feita numa perspectiva atomística, com recurso a métodos quantitativos processados pela ferramenta VARBRUL. Para tanto, não só descrevemos esta mas também a forma como foram recolhidos, tratados e codificados os dados submetidos a análise. Para levar a cabo a análise, foram constituídas uma variável dependente binária

(presença ou ausência de marca PL nos elementos do SN), onze variáveis independentes linguísticas (três fonológicas, sete morfossintáticas e uma lexical) e quatro variáveis independentes extralinguísticas. Para cada uma das variáveis independentes foram constituídos os respectivos factores e formulada uma hipótese, que pretendemos ver confirmada, ou não, nos próximos capítulos. Foi também descrita a metodologia aplicada na quantificação dos dados para, com recurso às primeiras rodadas VARBRUL, se proceder à selecção e escalonamento hierárquico das variáveis independentes que têm intervenção directa na inserção da marcação PL e configuração da regra da concordância de número PL no SN do PA. Assim, os passos seguintes do nosso estudo darão conta do modo como trabalhámos, com recurso à ferramenta VARBRUL, as diferentes variáveis independentes linguísticas e extralinguísticas tidas como confiáveis, a fim de apresentarmos e comentarmos os resultados que permitem observar o modo como estas actuam nas referidas inserção e configuração.



---

<sup>1</sup> Na época renascentista, por exemplo, Dante debruçou-se sobre 14 dialectos italianos, embora os considerasse línguas corrompidas.

<sup>2</sup> No Brasil, a “variação na concordância de número é bem significativa, chegando a ocorrer entre as camadas mais cultas, e estando amplamente generalizada entre as variedades populares da zona rural, onde o plural no SN é quase que só marcado no determinante” (Lucchesi, 2000a:19).

<sup>3</sup> Se bem que os conceitos de pidginização/crioulização e transmissão linguística irregular (*imperfect language shift*) digam respeito a processos sociohistóricos de contacto prolongado entre línguas, nos quais intervêm um idioma Y que é tomado como modelo para os falantes envolvidos no contacto, estamos perante dois processos distintos de aquisição: o primeiro envolve a aquisição/criação de uma L2, enquanto o segundo envolve a aquisição/criação de uma L1. Deste modo, o conceito de pidginização/crioulização é menos abrangente do que o conceito de transmissão linguística irregular, visto envolver processos de natureza linguística e sociohistórica que conduzem ao emergir de uma nova língua X (pidgin ou crioulo), a qual se assume como predominante na situação de contacto. Por conseguinte, a pidginização/crioulização pressupõe uma situação de reduzido acesso à língua-alvo Y, determinante para o desencadeamento de um processo de estruturação linguística, que decorre do material primário disponível adquirido e das necessidades estruturais e sociopragmáticas do que está que está a ser adquirido. Estes aspectos são determinantes para o aparecimento da nova língua X, bastante distinta das que lhe concederam os modelos iniciais para aquisição, uma vez que incorpora itens morfossintácticos, fonológicos, lexicais e semânticos de uma língua Z, e que não foram adquiridos a partir da língua Y. Quanto à transmissão linguística irregular, por configurar um *continuum* de níveis diferenciados de socialização/nativização de uma L2 (a língua X) adquirida em situação sociohistórica específica, apresenta resultados diferenciados entre si, caracterizados por não-realização de morfologia flexional que pode ir de grau mais acentuado a grau mais leve. As questões em torno da transmissão linguística irregular e tipos de variação por ela determinados encontram-se debatidos no ponto 2.11.3 do presente trabalho.

<sup>4</sup> Os falantes do PA terão ainda tido contacto quer com o PtgL2 influenciado por estruturas bantu, falado por africanos “contratados” em finais do século XIX para trabalharem nas roças de cacau e café de São Tomé, quer com o CCV, L1 de trabalhadores enviados, posteriormente, para as mesmas roças. Deste modo, não será também de desconsiderar a possibilidade de o PA apresentar influências decorrentes do contacto com estas línguas.

<sup>5</sup> Do contacto entre o português e os dialectos ameríndios resultaram, essencialmente, contribuições para o léxico do PB. Os linguistas têm sido unânimes em considerar que foi mais forte a ligação entre o PE e as línguas dos africanos, ainda que não seja de desconsiderar a intencionalidade por parte dos colonos em manter a diversidade linguística, a fim de se evitar a unidade nos grupos escravos, mantendo-os submissos (Rodrigues, 1985). Dado que a diversidade foi uma realidade, alguns estudiosos acreditam que as influências dos substratos ameríndios possam estender-se quer ao sistema vocálico, afectando a pronúncia e a entoação, quer ao sistema morfológico (Houaiss, 1985; Baxter & Lucchesi, 1997; Lucchesi, 1999, 2009a; Lucchesi & Baxter, 2006), originando “a redução de flexões, fenómeno comum aos falares crioulos” (Elia, 1979:41). Por outro lado, a CPL-var no SN estabelece-se a nível geral após a emergência de uma situação de transmissão linguística irregular no Brasil, quando os falantes dos substratos bantu e kwa, detentores de um código linguístico marcado por ausência formal de concordância a nível do SN, passaram a contactar simultaneamente com o PE e os dialectos ameríndios. Como consequência desse contacto, o sistema de marcação africano, por ser de reduzida flexão, terá sido transposto quer para o PtgL2 defectivo quer para o PBL1 (Guy, 1981a, 1989, 2005), quer ainda para dialectos ameríndios, como acontece na linguagem dos caipiras (Mendonça, 1936).

<sup>6</sup> A deriva é aqui entendida como sinónimo de mudanças fonéticas que condicionam o comportamento morfossintáctico das línguas no decorrer do tempo, originando novas formas, como dialectos, línguas ou blocos linguísticos (Sapir, 1971 [1921]). Deste modo, a *deriva românica* enquanto fenómeno inerente à própria língua, afectou a flexão gramatical das línguas europeias, em geral, e das línguas romance, em particular, levando as segundas a diferenciarem-se do latim. Trata-se, pois, de um fenómeno histórico que, no caso do português, e como esclarece Lucchesi (2000a), conduziu à simplificação morfológica do idioma luso desde as suas origens mais remotas no indo-europeu, com destaque para a perda da flexão casual dos nomes, na passagem do latim para o romance. Idêntico fenómeno ocorreu, por exemplo, nas línguas germânicas, que sofreram mudanças internas que as levaram a tornarem-se diferentes do germânico.

<sup>7</sup> Chomsky usa o “termo técnico «língua-I» (de «língua interiorizada»), opondo-se ao termo «língua-E» (de «língua exteriorizada»), que refere o conjunto de frases e expressões determinadas pela língua-I” (Raposo, 1992:27, nota 3), para se referir à linguagem, isto é, ao sistema de conhecimentos interiorizados na mente do homem, e que permite que este reconheça, interprete e produza enunciados na sua L1. Deste modo, o objecto de estudo dos generativistas é a língua-I, não a língua-E.

---

<sup>8</sup> No presente estudo, o conceito de LA é essencialmente sinónimo de idioma que é objecto de apropriação por parte de uma comunidade (aquisição massiva), que tende a substituir a sua L1 pela LA. Neste caso, a comunidade pode, diacronicamente, vir a conceder-lhe o estatuto de L1, isto é, de língua dos seus afectos.

<sup>9</sup> Disponível em: <http://www.instituto-camoes.pt/revista/descbroutro.htm>.

<sup>10</sup> Apesar de admitir que os casos de discordância são frequentes na língua oral, Bechara (1999 [1928]:544) atribui estes ao facto de o fluxo do pensamento ocorrer mais rapidamente do que a formulação e estruturação da oração. Deste modo, o gramático sugere que o falante aproveite cuidadosamente a liberdade de concordância que a língua portuguesa oferece para não prejudicar a clareza da mensagem e a harmonia do estilo.

<sup>11</sup> No estudo da concordância variável de género levado a cabo por Lucchesi (2000a) para o dialecto de HEL-Ba, o linguista deu também conta do fenómeno da analogia, avançando que o mesmo é inerente às situações de *transmissão linguística irregular* e resulta da generalização da interpretação de que as palavras terminadas em *-a* serão femininas quanto ao género.

<sup>12</sup> Guy (1981a) estabeleceu as seguintes variáveis sociais para as suas observações: *sexo, idade (adolescentes, adultos jovens e adultos velhos) e estilo de fala (informal e cuidadoso)*.

<sup>13</sup> No estudo atomístico, cada elemento do SN é visto como um dado de análise, definido pela presença ou ausência de marca de PL. Lucchesi (2000a:22) optou por apelar à análise de “mórfica”.

<sup>14</sup> O estudo sintagmático analisa o SN como um todo (uma única unidade de análise), debruçando-se sobre dois ou mais constituintes extraídos dos dados. Como tal, atribui-se um valor positivo aos SN's que tenham todos os seus constituintes marcados e um valor negativo aos que contenham pelo menos um elemento sem marcação de número, quando esta deveria figurar no mesmo.

<sup>15</sup> O Princípio do Processamento Paralelo é um dos princípios que contribuem para a elaboração do Processamento Serial (algoritmo incrementativo e determinístico), que determina a aplicação de operações mentais sobre cadeias de símbolos, em que o *output* de uma operação serve de *input* para a próxima. De acordo com o Princípio do Processamento Paralelo, o processo de compreensão engloba estágios distintos: acesso lexical, *parsing*, forma fonética, etc. Tais processamentos ocorrem um de cada vez, sem debordamento. Portanto, num modelo de processamento em paralelo, todos esses processos podem ocorrer simultaneamente, isto é, pode especificar-se foneticamente um vocábulo, ao mesmo tempo que se acessa a outro e se constrói a estrutura sintáctica. No caso concreto em que “*marcas levam a marcas e zeros levam a zeros*” (Scherre, 1988:208), é o efeito do paralelismo sintáctico, consequência talvez de um mecanismo mental, que faz o indivíduo ser levado a repetir e a imitar contextos linguísticos.

<sup>16</sup> Para os considerandos elaborados acerca do conceito de Scherre (1988) e Silva & Scherre (1998 [1996]), sobre a intervenção do Princípio do Processamento Paralelo na marcação PL dos itens do SN, P. Andrade (2003), terá partido do pressuposto que SN se encontra licenciado em SDET (Abney, 1987). Nesta conformidade, o DET é uma extensão da projecção da expressão nominal, sendo aí realizado o morfema de PL, em virtude de assumir a posição de núcleo do SDET.

<sup>17</sup> Os morfemas de conteúdo atribuem semântica aos enunciados e são os únicos cujos lemas contêm a informação necessária para a realização das estruturas de superfície. Por seu lado, os morfemas de sistema dividem-se em duas categorias: morfemas sistémicos estruturais (*early system morphemes*) e morfemas de nível funcional sistémico (*late system morphemes*) (Myers-Scotton, 2002:74). Destes, os primeiros (*early system morphemes*) são activados a nível conceptual para estabelecerem as ligações das intenções semântico-pragmáticas (*chunks*) do falante às estruturas de superfície (*content morphemes*) (Myers-Scotton, 2002:14). Como os lemas dos *early system morphemes* são activados de modo indirecto pelos lemas que suportam os morfemas de conteúdo, ambos são adquiridos de modo mais fácil. O mesmo não acontece com os *late system morphemes*, uma vez que são apenas activados a nível estrutural, ou seja, quando se procuram expandir as estruturas linguísticas.

<sup>18</sup> Baxter (2004:101-102) aponta cinco motivos que levaram o santomense a não penetrar em Monte Café: (i) o facto de os ancestrais escravos terem abandonado as roças em massa, após a abolição da escravatura, no século XIX; (ii) o facto de grande parte dos antigos capatazes das roças ser constituída por europeus; (iii) o facto de os trabalhadores transplantados continuarem a falar as suas L1's ancestrais, mantendo-as entre os seus descendentes; (iv) o facto de o sistema de controlo das roças não favorecer os contactos com outras comunidades; (v) o facto de a localização geográfica das próprias roças contribuir, também ela, para o hermetismo das mesmas.

---

<sup>19</sup> Variáveis constituídas por Baxter (2004) para análise da CPL-var do PT: (i) saliência da forma morfofonológica; (ii) contexto fonológico seguinte; (iii) marcas antecedentes; (iv) classe gramatical; (v) posição relativamente ao núcleo do SN; (vi) idade; (vii) sexo; (viii) descendência; (ix) educação.

<sup>20</sup> Em 1930, o governo português adoptou o modelo de “assimilação” francês, que procurava, através da língua colonialista, a única permitida no sistema escolar, assimilar e aculturar a população africana não-indígena aos modelos europeus (Barreto, 1977). Para além da fluência no português, outras condições impostas, para se poder ascender à condição de “assimilado”, eram: (i) ter rendimentos regulares; (ii) possuir nível de educação satisfatório (4<sup>a</sup>. classe); (iii) adoptar a monogamia (Firmino, 1995:260, nota de rodapé 8).

<sup>21</sup> A implementação de uma política de língua e educacional consistente apenas emergiu em 1930, quando o governo português adoptou o modelo de “assimilação” francês (Barreto, 1977). Deste modo, a primeira escola primária só viria a ser criada no final do século XIX, sendo o seu número, em meados do século seguinte (1945), apenas de 92 (Mateus, 1999:212).

<sup>22</sup> Na análise VARBRUL, as probabilidades devem ser observadas em função dos pesos relativos e não dos percentuais, já que estes não são indicadores fiáveis para os efeitos exercidos pelas variáveis linguísticas e extralinguísticas no fenómeno em estudo.

<sup>23</sup> Os informantes que produziram o *corpus* de Jon-And (2008) são todos falantes que têm uma língua bantu como L1 (xironga, confinado a cerca de 650000 indivíduos na região do Maputo, ou xichangana, falado por 11,4% da população moçambicana) (Instituto Nacional de Estatística de Moçambique, 2010).

<sup>24</sup> O povoamento das ilhas de Cabo Verde iniciou-se no Sul, a partir de Santiago, em 1462. Seguiu-se-lhe o do Fogo e, depois, o da Brava. Santo Antão e S. Nicolau, no lado Barlavento, foram povoadas com escravos trazidos de Santiago e do Fogo, no século XVII, ao passo que só aconteceu um verdadeiro povoamento de S. Vicente a partir dos finais do século XVIII (Carreira, 1979; Santos, 2002).

<sup>25</sup> Segundo Ferraz (1975:36), a ausência do artigo definido nos CP's da Alta Guiné e do Golfo da Guiné deve-se à influência dos substratos africanos.

<sup>26</sup> O desenquadramento da tradução em português visa conservar o espírito original da produção no crioulo.

<sup>27</sup> Definido como *Lexicon*, na perspectiva de Chomsky (1996 [1995]:6).

<sup>28</sup> Dentro destas funções encontram-se processos como a formação de palavras, armazenamento de correspondência imprevista na relação som/significado e de formas morfofonológicas irregulares ou a geração das unidades atómicas da análise sintáctica (Silva, 2010:19).

<sup>29</sup> O clássico conceito de crioulo exógeno ou de plantação (*plantation creole*), proposto por (Holm, 1988/1989), isto é, de crioulo que emerge entre falantes de distintas L1's geográfica e culturalmente deslocados por colonos, é questionado à luz da crioulistica actual (p.e. McWhorter, 2010), que só admite a emergência de crioulos endógenos (*fort creoles*). Assim, os crioulos apenas se constituem em regiões multilingues, às quais aportam grupos dominantes vindos de fora para trocas comerciais, fazendo com que os naturais dessas regiões e das áreas circundantes se associem a si cultural e economicamente. Desta forma, os indígenas, ainda que não deixem de falar as suas L1's, necessitam também de criar um código linguístico de emergência para comunicarem não só entre si mas também com os grupos estrangeiros, detentores da LA. Pelo contrário, nas situações de plantação, os escravos transplantados já são detentores de códigos linguísticos de emergência formados endogenamente (pidgins ou proto-crioulos), que se expandem e desenvolvem no seio das comunidades dos falantes de distintas L1's.

<sup>30</sup> Tendo em conta que o edo (Nigéria) é o mais antigo e mais significativo substrato dos CGG's, a etimologia de *inen* radica neste dialecto, apontando Alexandre & Hagemeyer (2007:40, nota 5) três possíveis origens para a mesma: (i) terceira pessoa PL do pronome fraco e forte *iran* “eles”, do qual *inen* é homófono; (ii) determinante definido (*n)ene* “o”; (iii) pronome demonstrativo *ene/enena* “estes/estas”, aparentemente relacionados a (ii) (cf. Melzian, 1937; Agheyisi, 1986, 1990).

<sup>31</sup> Exemplo disponível em: <http://www.abeokuta.org/yoruba.htm> [Consultado em 24 de Maio, 2010].

<sup>32</sup> Exemplos disponíveis em: <http://www.linguakimbundu.com/> [Consultado em 24 de Maio, 2010].

<sup>33</sup> Alguns nomes do santomense, a exemplo do que sucede nas línguas bantu, possuem formas masculina e feminina distintas, como é o caso do substantivo *sogru* > *sogla* (sogro > sogra) (Ferraz, 1979:60). Estes aspectos encontram-se discutidos no ponto 1.3.2.2.1 do presente trabalho.

---

<sup>34</sup> Os exemplos extraídos do *corpus* do PT e que surgem referenciados com o nome do informante foram-nos fornecidos por Alan N. Baxter, que dirigiu o projecto “*Semi-creolization: testing the hypothesis against data from Portuguese-derived languages of São Tomé (Africa)*”. Informações detalhadas sobre o projecto poderão ser consultadas no capítulo 2 do presente trabalho, dedicado à metodologia aplicada no mesmo.

<sup>35</sup> A origem etimológica de *ũa* é atribuída ao indefinido “uma” do português (Alexandre & Hagemeyer, 2007:39, nota 5), que detém o mesmo significado do santomense. Neste, o determinante indefinido e o numeral *ũa* são homónimos.

<sup>36</sup> “Ka” é um morfema aspectual largamente difundido no santomense e, ao que tudo indica, tem origem nas línguas kwa. É utilizado não só para modificar o próprio verbo mas também conjuntamente com outros verbos, já que a conjunção coordenativa copulativa “e” não existe no forro. Deste modo, as orações são ligadas umas às outras pelo método de aposição e, caso o sujeito seja o mesmo para todas elas, a sua retenção é obrigatória apenas na primeira oração.

<sup>37</sup> Exemplos disponíveis em: <http://www.linguakimbundu.com/> [Consultado em 29 de Outubro, 2009].

<sup>38</sup> Exemplo disponível em: <http://vidademacumbeiro.blogspot.com/2008/12/gramtica-yoruba.html> [Consultado em 25 de Agosto, 2010]

<sup>39</sup> Por consoante homorgânica entenda-se aquela que é idêntica a uma outra, em relação ao modo de articulação, à zona de articulação e ao papel das cavidades bucal e nasal, mas diferente quanto à sonoridade. Assim, uma é surda e a outra sonora. Tendo como exemplo a consoante /p/, vemos que ela partilha os seus fechos (traços fónicos) de *oclusividade*, *bilabialidade* e *oralidade* com a consoante /b/. Contudo, /p/ possui o traço de não-sonoridade (é surda), enquanto /b/ é marcada pelo traço de sonoridade (é sonora). O mesmo ocorre com outras consoantes, como /f/ e /v/. Ambas são constritivas-fricativas, labiodentais e orais, mas a primeira é surda e segunda sonora. Na produção de /v/ ocorre vibração das cordas vocais, o que não sucede com a articulação de /f/. Listam-se ainda outros pares homorgânicos: /t/ e /d/; /k/ e /g/; /s/ e /z/; /ʃ/ e /ʒ/.

<sup>40</sup> A vogal morfofonémica [i] é definida por Ferraz (1979:20), em ocorrências no santomense, como “*a nasal central high vowel with lip spreading. Nasality is the only linguistically significant feature of this nucleus. [i] occurs only in two instances in the language: in optional variation with /m/ ~ /n/ [n, ŋ] as a locative morpheme, and as the 1st p.s. subject pronoun. [i] is thus the representation of a morpheme whose main phonological identity is nasality*”.

<sup>41</sup> A dobra é a unidade monetária de S. Tomé, introduzida na ilha em 1977 para substituir o escudo português.

<sup>42</sup> Jornal de São Tomé e Príncipe [em linha]. Disponível em: <http://www.jornal.st/directorio.php?page=173> [Consultado 12 Dez. 2009].

<sup>43</sup> Artigo 13º da lei 53/88.

<sup>44</sup> A ausência de um conceito claramente definido para zona urbana em São Tomé e Príncipe levou o Instituto Nacional de Estatística do país a optar por considerar esta como sendo uma área que possuía uma “*aglomeração de casas com algumas infra-estruturas, tais como arruamentos, asfaltos, luz eléctrica nas casas e nas ruas, canalização de água, redes de esgotos, e em que a actividade económica predominante sejam os serviços e / ou a indústria. Também todas as capitais distritais foram consideradas como urbanas*” (Instituto Nacional de Estatística de São Tomé, 2003a:8).

<sup>45</sup> Outro dos factores que contribuiu para que o número de mulheres seja superior ao dos homens é a maior esperança de vida actual das primeiras (Instituto Nacional de Estatística de São Tomé, 2003a:14).

<sup>46</sup> Matabichar: termo usado nas variedades de português africano, com o significado de “tomar o pequeno-almoço”.

<sup>47</sup> O informante refere-se ao tempo em que o Comité de Libertação de São Tomé e Príncipe, posteriormente denominado Movimento de Libertação de São Tomé e Príncipe (MLSTP), desenvolvia acções de luta contra as autoridades coloniais portuguesas.

<sup>48</sup> Os estudos iniciais sobre ASL partiram do pressuposto que era mais fácil adquirir o que é aproximado à L1 e mais difícil aprender aquilo que se afasta dela (Lado, 1957). Corder (1967; 1971; 1974; 1981b) constataria este pressuposto, afirmando que a interferência pode provir de outra língua semelhante ou intermediária. Corder (1967) atribuiria também a denominação de *idiosincratic dialect* (dialecto idiosincrático) ao sistema linguístico da L2, termo que Selinker (1972:216) modificaria para *interlanguage* (interlíngua).

---

<sup>49</sup> A Hipótese do Período Crítico (*Critical Period Hypothesis – CPH*) para aquisição de uma língua, proposta por Lenneberg (1967), é de carácter marcadamente neurológico, advogando que as mudanças que se operam a nível cerebral impedem os adolescentes e adultos de accionarem na perfeição as capacidades inatas para aquisição de línguas, o que não sucede com as crianças. A abrangência deste conceito e as incertezas em torno da proposta do CPH não deixam que se determine com exactidão os seguintes aspectos: (i) se a habilidade para aquisição declina somente no período crítico; (ii) se a mudança no ponto crítico se dá brusca ou gradualmente; (iii) se o ponto crítico tem influência em todos os aspectos da linguagem ou se apenas em alguns; (iv) se o ponto crítico é meramente biológico ou secundado por outros aspectos que intervenham na capacidade da aprendizagem.

<sup>50</sup> As siglas *FLO* e *FLf* dizem respeito à terminologia adoptada por Raposo para a tradução de Chomsky (1999a). Chomsky (1996 [1995]) recorreu às siglas *S0* (estado inicial) e *S8* (*estado estável*), respectivamente. Por seu lado, Silva (2010) optou pelas siglas *S0* (cf. Chomsky, 1996 [1995]) e *Sp*, no sentido de Língua Produto, conforme mencionado frequentemente por João Costa durante o seminário “Tópicos em Sintaxe Generativa”, ministrado na Universidade de Macau, entre 25 e 31 de Março, 2008.

<sup>51</sup> A este propósito, o *Princípio do Subconjunto* defende que “se um parâmetro tem dois valores «+» e «-», e o valor «-» determina uma linguagem que é subconjunto da linguagem determinada pelo valor «+», então o valor «-» é o valor não-marcado, seleccionado pela GU na ausência de qualquer evidência linguística” (Chomsky, 1986a:146). É o que acontece no caso da *informação negativa* (incluindo a *informação negativa indirecta*), que não desempenha nenhum papel na aquisição da linguagem se o valor não marcado do *parâmetro* for a posição *não-pro-drop* (realização fonética obrigatória do sujeito), já que o valor inicial do *parâmetro do sujeito nulo* é a posição *não-pro-drop*. Deste modo, os parâmetros possuem um valor inicial não-marcado, que apenas é alterado durante a aquisição caso ocorra evidência positiva em contrário nos DLP’s. Nesta conformidade, o valor que desencadeia os movimentos, a partir dos núcleos funcionais, é o valor «+», uma vez que os movimentos são responsáveis por representações relativamente complexas. Por outro lado, a noção de marcação será um conceito teórico-I, uma vez que tem a ver com o modo como é aplicado o módulo das regências e ligações.

<sup>52</sup> Os termos técnicos “movimento” e “combinação” radicam nos conceitos “*move*” e “*merge*”, avançados por Radford (1997:61) na teorização sobre o minimalismo.

<sup>53</sup> Os sons que as crianças produzem nesta fase não são considerados manifestações de língua, uma vez que carecem de criatividade. De facto, representam apenas respostas involuntárias do recém-nascido a situações de desconforto (como a fome) ou conforto (por exemplo, a necessidade de ser embalado).

<sup>54</sup> As frases holofrásticas contêm uma única palavra, cujo significado se pode alargar para abranger uma classe mais ampla. Na fase holofrástica, as crianças detêm um vocabulário de cerca de 20 palavras, que utilizam para servir três funções principais: exprimir o desejo de actividade (por exemplo, “colo” pode significar o desejo de ser levado ao colo); manifestar emoção (por exemplo, “não” pode significar “*Não gosto!*”); e servir para denominar algo (por exemplo, “cão” pode significar “*Aquilo é um cão.*”). Paralelamente, a criança usa também as palavras isoladas já com fins sociais. De facto, e voltando a pegar na palavra “não”, verificamos que a mesma pode, neste caso, ser utilizada quer com sentido imperativo (“*Não me dêes sopa!*”) ou assertivo (“*Não gosto de sopa.*”).

<sup>55</sup> “... the trigger, or PLD (*primary linguistic data*) must consist only of the kinds of things that children routinely experience and includes only simple structures” (Lightfoot, 1999b:51).

<sup>56</sup> Em psicologia, o conceito de *priming* é aplicado às situações em que um estímulo precoce influencia estímulos posteriores. Os testes neste sentido têm evidenciado que, por exemplo, quando um indivíduo lê uma lista de palavras que incluiu um determinado vocábulo (p.e. comida), caso lhe seja pedido que complete uma palavra iniciada com letras pertencentes a esse vocábulo (p.e. com\_), a probabilidade de ele escrever esse mesmo vocábulo é maior do que se não o tivesse lido anteriormente. Outro exemplo de *priming* envolve pessoas a quem é mostrada uma figura inacabada, que eles não são capazes de identificar. Seguidamente, são-lhes apresentados mais elementos da figura, até que eles a consigam reconhecer. Posteriormente, se os elementos da figura lhes forem novamente exibidos, elas identificá-los-ão mais precocemente do que da primeira vez (Kolb & Whishaw, 2003:453-457).

<sup>57</sup> A alternância de códigos pressupõe o uso alternativo de duas ou mais línguas na mesma situação de conversação. Genericamente, o termo diz respeito a diferentes formas de bilinguismo, já que a alternância pode acontecer nos actos de fala de vários indivíduos durante uma conversação, nas elocuições dentro de um único acto de fala e ainda dentro de uma simples elocução (Milroy & Muysken, 1995).



---

<sup>58</sup> Os padrões morfológicos e sintáticos das alternâncias linguísticas são regidos por regras gramaticais específicas e fronteiras especiais nas quais a alternância de códigos pode ocorrer. Assim, a literatura da especialidade tem apontado alguns constrangimentos que impedem a alternância de códigos:

- (i) o “Constrangimento do Morfema Livre” (*Free-morpheme Constraint*), que determina que a alternância de códigos não pode ocorrer entre morfemas presos (Sankoff & Poplack, 1980);
- (ii) o “Constrangimento da Equivalência” (*Equivalence Constraint*), que advoga que a alternância de códigos apenas pode acontecer em posições onde a ordem de dois dos elementos da frase (um antes e outro depois da alternância) não é excluída em nenhuma das gramáticas (Sankoff & Poplack, 1980);
- (iii) o “Constrangimento das Classes Fechadas” (*Closed-class Constraint*), que entende que os itens que constituem classes fechadas não podem ser alternados (Joshi, 1985);
- (iv) o “*Matrix Language Frame Model*”, que faz a distinção entre as regras das línguas detidas pelo falante (Myers-Scotton, 1997);
- (v) “Constrangimento do Núcleo Funcional” (*Functional Head Constraint*), que indica que a alternância de códigos não pode ser aplicada entre um núcleo funcional (um determinante, uma flexão, etc.) e o seu complemento (SN, SV, etc.) (Belazi, Rubin & Toribio, 1994).

No caso do “*Constrangimento da Equivalência*”, Sankoff & Poplack (1980:18) apresentam o seguinte exemplo, para justificarem o seu pressuposto: *I like you porque eres simpatico* (“*I like you because you are likable*”). Como se verifica, ocorre alternância de código porque há uma obediência às regras de formação das orações de ambas as línguas (orações causais no exemplo fornecido para o espanhol e o inglês), que é exactamente o que sucede nas situações apontadas para o PA. Note-se ainda que o princípio do “*Constrangimento das Classes Fechadas*” (Joshi, 1985) também é contemplado no PA, já que não detectámos nenhuma situação de alternância envolvendo, por exemplo, preposições.


<sup>59</sup> O projecto DUFDE (*Deustsh und Französisch – Doppelter Erstspracherwerb*), detalhadamente descrito por Köppe (1994a), foi desenvolvido entre 1986 e 1992, sempre sob coordenação de Jürgen Meisel, e levou a cabo estudos longitudinais sobre o desenvolvimento linguístico de crianças bilingues em alemão e francês, desde o seu início (1;0 – 1;6 anos) até à idade aproximada dos 5 anos. A colecta dos dados era efectuada a cada duas semanas, tendo o projecto baseado a sua metodologia de selecção de famílias no “princípio de uma pessoa/uma língua”, enquanto pressuposto inerente ao desenvolvimento de um bilinguismo equilibrado. Desta forma, todas as crianças eram filhas de pai alemão, que com elas só comunicava em alemão, e mãe francesa, que com elas só comunicava em francês. Inicialmente, a pesquisa incidiu sobre o desenvolvimento linguístico de treze crianças, mas seis destas foram posteriormente descartadas do estudo, já que não apresentavam um bilinguismo equilibrado, como era exigido pela investigação. Todavia, os registos produzidos por estas crianças não foram abandonados, sendo também utilizados para efeitos comparativos. Quanto às restantes sete crianças, o seu desenvolvimento linguístico foi seguido até final do tempo previsto, embora apenas sejam contempladas nas análises dos estudos editados por Meisel (1994b) as produções de cinco dessas crianças.

<sup>60</sup> Alguns autores (Vainnika & Young-Scholten, 1994, 1996a, 1996b, *apud* Godinho, 2005:107) defendem mesmo que apenas a GU e parte da L1 têm papel interventivo no estado inicial da ASL.

<sup>61</sup> O Modelo do Restabelecimento dos Parâmetros proposto por Schwartz & Sprouse (1996) e desenvolvido por Sprouse (2006) assemelha-se ao da relexificação, avançado por Lefebvre (1986, 1998) para defender que a crioulição é um processo mental e sociolinguístico. Contudo, a referida semelhança dá-se unicamente a nível do conteúdo, diferindo na forma. Deste modo, na relexificação, o estágio final do modelo interlíngua (o crioulo) é semelhante ao do estado inicial da aquisição da L1. Segundo Sprouse (2006), no modelo de transferência em bloco da gramática da L1, a génese dos crioulos é uma espécie de aquisição da L2 em reverso. Assim, o crioulo é comparável a uma variedade fossilizada de L2 do superstrato. No entanto, a hipótese parece pouco razoável, porque todos os crioulos de uma determinada base teriam que reflectir, na totalidade, iguais estágios iniciais de ASL.

<sup>62</sup> A Condição de Visibilidade determina que a teoria temática se associe à teoria do caso, advogando o seguinte: *um elemento só é visível para atribuição de um papel temático se estiver numa posição com caso ou se for PRO*.

<sup>63</sup> De acordo com o processamento da sintaxe observado por Pienemann (2000) no inglês, Plag (2008a:123) elabora a seguinte tabela hierárquica de aquisição para este idioma L2:

Desenvolvimento	Estrutura	Exemplo
Estágio inicial	Realizações holofrásticas	<i>Ball.</i>
 LA	Ordem canónica (SVO)	<i>John kick ball.</i>
	Neg + V	<i>He no like coffee.</i>
	Frase iniciada com advérbio	<i>Then Bob kick ball.</i>
	Topicalização	<i>That I did'n like.</i>
	Frase iniciada com “do”	<i>Do you like it? / Do she like it?</i>
	Inversão “yes/no”	<i>Has he seen you?</i>
	Inversão da cópula	<i>Where is John?</i>
	Partículas verbais	<i>Take the hat off</i>
	“Do” / Auxiliar 2ª posição	<i>Why did she sell that car?</i>
		<i>Where has he gone?</i>
	Cancelamento da inversão	<i>I wonder why he sold that car.</i>

Os aprendentes iniciam a aquisição com realizações holofrásticas e vão, posteriormente, adquirindo estruturas mais complexas, mas seguindo sempre um padrão bastante definido. Alguns dos aprendentes conseguem chegar ao estágio em que realizam estruturas complexas e outros alcançam mesmo o patamar que lhes permite inverterem as orações subordinadas interrogativas.

<sup>64</sup> A emersão tardia da flexão inerente pode também ser constatada no desenvolvimento da flexão verbal de tempo e aspecto. Segundo Bardovi-Harlig (2000:25), a aquisição da morfologia verbal na ASL requer três estágios de desenvolvimento, adquirindo os falantes as marcas de flexão apenas na última fase. Antes, os falantes limitam-se a recorrer a estratégias pragmáticas (p.e. ordem canónica), no primeiro estágio de aquisição, passando a usar formas lexicais para marcar as distinções temporais e aspectuais na segunda etapa do desenvolvimento da interlíngua (p.e. advérbios de tempo ou conectores frásicos).

<sup>65</sup> Note-se que a noção de protótipo não implica que o mesmo seja categórico. De facto, o protótipo assume grau de variação nos membros que constituem uma certa categoria ou classe. Tal variação fica a dever-se à dinâmica dos sistemas linguísticos, da mesma forma que acontece em determinadas áreas científicas, como a biologia (Clements, 2009:4).

<sup>66</sup> Para uma consulta mais pormenorizada acerca dos tipos de desencadeamento (*bootstrapping*) envolvidos na aquisição de L1 e L2, consulte-se Bates & Goodman (1999).

<sup>67</sup> Segundo Bechara (1999 [1928]:122-123), estamos perante um caso de sincretismo e não de neutralização, pelo que “*não se pode aceitar a doutrina corrente que vê nessas formas um singular que permanece invariável no plural*”. De facto, a pluralidade surge marcada através do adjunto nominal e, em determinados casos, apesar de existir a forma PL, recorre-se normalmente àquela que é invariável (p.e. *simples* possui o PL *simpleces* e *cós* tem o PL *coses*).

<sup>68</sup> O PL deste tipo de substantivos abstractos produz uma alteração semântica, denotando uma repetição ou multiplicidade de actos, com conotação aumentativa (Brunot, 1989).

<sup>69</sup> Informações pormenorizadas sobre os substantivos massivos e seu comportamento podem ser consultadas em Oliveira (2003:151-152) e Duarte & Oliveira (2003:219-221/236-242).

<sup>70</sup> Bechara (1999 [1928]:115-116) chama a atenção para a diferença entre colectivos universais e particulares e os colectivos que representam nomes de grupos, definidos apenas como colectivos pela gramática tradicional. Assim, são colectivos os que representam nomes de grupos, os conjuntos de pessoas (p.e. *quadrilha*; *elenco*), os grupos de animais (p.e. *matilha*; *ninhada*) e os grupos de coisas (p.e. *cacho*; *mobiliário*). Nestes casos, estamos perante um conjunto de elementos contáveis de uma espécie definida (os grupos de animais), ou total ou parcialmente indefinida (os grupos de pessoas e os grupos de coisas) que, ao contrário dos colectivos universais e particulares, “*requerem a determinação explícita da espécie de objectos que compõem o conjunto*”. Deste modo, é necessário esclarecer se estamos perante *um bando de ladrões* ou *um bando de pardais*, mas não é necessário especificar que estamos perante *um casario de casas*.

<sup>71</sup> Remetemos para Bechara (1999 [1928]:127-128) informações detalhadas acerca destes plurais, nomeadamente no que respeita à regra ortográfica oficial.

---

<sup>72</sup> Os nomes próprios de pessoas, quando usados referencialmente, são obrigatoriamente precedidos de um artigo definido no PE:

- a. *O João leu um livro.*
- b. \**João leu um livro.*

São exceção à regra os nomes próprios de personagens históricos ou que fazem parte da memória colectiva portuguesa, especialmente quando referenciados em discurso literário ou histórico relativamente formal. Neste caso dispensam o uso do artigo definido:

- a. *Diogo Cão foi o primeiro navegador português a chegar à foz do Rio Zaire.*
- b. ?*O Diogo Cão foi o primeiro navegador português a chegar à foz do Rio Zaire.*

<sup>73</sup> Exemplos avançados por Godinho, 2005:115:

- Isto não passa de uma **simples/mera** dor de cabeça.*
- \**Isto não passa de uma dor de cabeça **simples/mera**.*

<sup>74</sup> Exemplos retirados de Casteleiro (2001, vol. G-Z:3417).

<sup>75</sup> Para uma visão acerca da literatura produzida sobre concordância, consulte-se:

[http://www.surrey.ac.uk/LIS/SMG/projects/agreement/agreement\\_bib\\_unicode.htm](http://www.surrey.ac.uk/LIS/SMG/projects/agreement/agreement_bib_unicode.htm).

<sup>76</sup> Conforme demonstrado por alguns autores (Rizzi, 1982; Jaeggli, 1984; Burzio, 1986; Jaeggli & Safir, 1989), as línguas de sujeito nulo evidenciam o seguinte feixe de propriedades:

a. Sujeitos sem matriz fonológica:

- (1) PTG:  
*(Eles) Telefonaram.*

- (2) INGLÊS:  
\**(They) called.*

b. A possibilidade de o sujeito ocorrer em posição pós-verbal (“inversão livre”):

- (3) PTG:
  - a. *O João telefonou.*
  - b. *Telefonou, o João.*

- (4) INGLÊS:
  - a. *John called.*
  - b. \**Called John.*

c. Ausência de *efeitos vestígio-que* (*that-trace effects*). Em caso de extracção do sujeito, esta dá-se a partir da posição pós-verbal e não da posição pré-verbal (Campos, 1997:98):

- (5) PTG:  
Quem disseste que comprou um computador?

- (6) INGLÊS:  
\**Who did you say that bought a computer?*

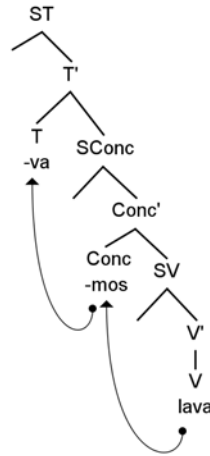
A ausência de *efeitos vestígio-que*, verificada em (5), ocorre em todas as línguas românicas de sujeito nulo (Jaeggli & Safir 1989:37). Por seu lado, Rizzi (1982:174) e Jaeggli (1984:139) entendem que a ausência de *efeitos vestígio-que* se deve ao facto de o sujeito, nas línguas de sujeito nulo, ser extraído directamente da posição pós-verbal e não da posição pré-verbal. Burzio (1986:165) comprova-o em dados do italiano, que indicam que a extracção do sujeito se dá, invariavelmente, a partir da posição pós-verbal e nunca da posição pré-verbal.

<sup>77</sup> As palavras resultantes de incorporação assemelham-se às formadas por composição, mas distinguem-se destas por manterem o seu referencial. Pelo contrário, os elementos compostos são esvaziados do seu conteúdo referencial (p.e. *o guarda-fatos > os guarda-fatos > \*as guarda-fatos*).

<sup>78</sup> Tradução nossa.

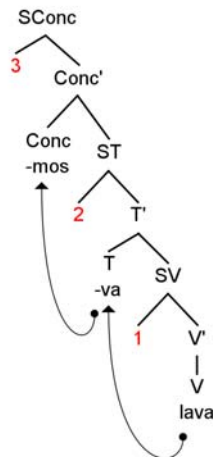
<sup>79</sup> Por *merge* entendam-se as operações de concatenação estrutural. Raposo (1992) atribui ao termo a denominação “*compor*”.

<sup>80</sup> Pollock (1989:386), verificando que o movimento do verbo estabelece correlações com a morfologia de tempo e concordância, mas também que o posicionamento necessita de dois núcleos independentes entre o SC e o SV, sugere que o verbo é formado na sintaxe. Contudo, a hierarquia proposta permite a produção de formas agramaticais, como acontece como a forma portuguesa \**lavámosva*:



- Hierarquia de Pollock (1989) -

Partindo da proposta do PEH (Baker, 1985), Belletti (1990:63) sugere a inversão da hierarquia de Pollock (1989:386), estabelecendo um movimento curto do verbo para T e estando as posições 1, 2 e 3 destinadas ao sujeito:



- Hierarquia de Belletti (1990) -

Apesar de a posição 2 estar destinada ao sujeito, este passa pela mesma sem se deter. Por outro lado, o verbo não vai até à flexão. Assim, pese embora o modelo permitir a realização da forma portuguesa *lavávamos*, não consegue justificar, por exemplo, as situações de inserção da marca de terceira pessoa do PL 's em inglês, já que esta se posiciona em FLEX, isto é, numa posição mais baixa. De qualquer forma, a questão revela aproximações às abordagens estabelecidas na MD (Halle & Marantz, 1993).

<sup>81</sup> De acordo com as assunções de Chomsky (1993), a concordância não tem que ser necessariamente visível, quando ocorre este tipo de configuração. De acordo com o quadro teórico *chomskyano*, a concordância representa uma conjunto de traços(-phi) abstractos, presentes nos núcleos CONC e nos elementos superficiais das posições SPEC, podendo a correlação que ocorre entre tais traços ser visível, ou não.

<sup>82</sup> Note-se que os movimentos não visíveis ou aparentes (*covert movements*) da FL não contam como movimentos reais.

---

<sup>83</sup> As dez categorias de mecanismos gramaticais minimalistas propostas pelo Programa Minimalista são: artigos, formas de TMA, palavras interrogativas, pluralizadores, pronomes de todas as pessoas e números, formas para marcar os casos oblíquos, preposições locativas gerais, um complementador irrealístico, partículas relativizadoras, reflexivos e recíprocos.

<sup>84</sup> Os nomes nucleares, além de próprios ou comuns, podem também ser coordenados (p.e. *Os contos e poemas que estudámos são interessantes.*) e compostos ou lexicais (p.e. *O meu guarda-fatos é de pinho.*).

<sup>85</sup> Os pronomes pessoais podem também constituir todo o SN da frase (p.e. *Tu vais ao teatro amanhã?*), já que não admitem especificadores nem complementos (à excepção de epítetos que se podem considerar um tipo especial de complementos), possuem como referência as pessoas do discurso ou aquela(s) de quem se fala (conhecida(s) dos interlocutores) e são identificadores, tal como sucede com nomes próprios (Brito, 2003a:351-352).

<sup>86</sup> Os estudos de Faria, Freitas & Miguel (1998, 2001) revelaram que a estrutura prosódica das crianças, em fase inicial da aquisição do PtgL1, se comporta de forma idêntica no SN e no SV. No caso do primeiro, “*a presença de informação morfo-sintáctica numa sílaba em final de palavra leva a criança a seleccionar a sílaba não tónica para fazer surgir a fricativa em coda de sílaba em final de palavra, que marca o plural em português, contrariando o padrão do desenvolvimento silábico*” (Godinho, 2005:139). Relativamente ao SV a fricativa morfológica em coda de sílaba é interpretada como morfema de pessoa.

<sup>87</sup> Ferrari-Neto, Corrêa & Augusto (2005) descartam a alternativa proposta por Magalhães (2004:169), que defende que o traço interpretável de número se encontra localizado apenas no DET.

<sup>88</sup> Aronoff (1976) apresenta um modelo voltado para o problema da produtividade lexical, que determina que quanto mais regulares forem os produtos de uma Regra de Formação de Palavras (RFP), maior será a sua produtividade. Como exemplo, compara as formas nominalizadas deadjectivais em inglês formadas com os sufixos *-ness* e *-ity*, apontando que, enquanto o acréscimo de *-ness* não produz qualquer alteração fonológica nas bases, o mesmo não acontece com o acréscimo de *-ity*. Além disso, as formas *X-ness* são semanticamente previsíveis, isto é, têm os significados previstos na RFP, enquanto as formas *X-ity* podem assumir significados idiossincráticos ou perder um ou outro significado previsto na Regra de Formação de Palavras. Assim, para o autor, a regra de formação de substantivos deadjectivais em *-ness* é mais produtiva do que a que forma a mesma classe de palavras em *-ity*. O autor propõe também que, diante da total regularidade das formas *X-ness*, elas não precisariam de estar listadas no léxico, uma vez que podem ser geradas por regra, ao passo que as formas *X-ity*, por não serem inteiramente regulares, têm de estar listadas no léxico. Por seu lado, Spencer (1991) refere que a produtividade também é atestada com o accionamento da regra em situações de emergência comunicativa, as chamadas formações esporádicas. De acordo com as condições morfossemânticas, o falante, perante a presença de itens lexicais envolvidos no processo mais uma situação discursiva do momento, pode criar uma palavra para designar um lugar definido pela situação. É o que acontece no português, por exemplo, nas situações em que emergem formas novas com o afixo *-dromo*, sempre proparoxítonas. Observando a formação e produtividade, Salles & Mello (s.d.) registam que, no estágio actual do PB, se está em fase de produtiva deste afixo (para a RFP *X-dromo*), o que determinou o aparecimento de palavras como sambódromo, ficódromo (lugar onde os namorados ‘ficam’, isto é, onde ficam juntos para namorar), beijódromo (lugar onde casais se beijam), fumódromo (lugar onde se pode fumar), etc. Contrariamente, formas derivadas com os afixos *-este* e *-estre*, como ‘agreste’ e ‘silvestre’, deixaram de ser atestadas diacronicamente em novas derivações. Todavia, note-se que a improdutividade que acontece na RFP *X-este* e *X-estre* constitui improdutividade da regra, e não do afixo. Estes exemplos apontam ainda para a relação entre a diacronia e a produtividade nas línguas, cuja discussão requer o exame das condições de formação dos itens no período diacrónico em que os elementos lexicais são produzidos.

<sup>89</sup> Müller (1994a), no âmbito dos estudos do projecto DUFDE (*Deutsch und Französisch – Doppelter Erstspracherwerb*), discute a emergência dos traços gramaticais de género e número na gramática de duas crianças em fase de aquisição bilingue das línguas alemã e francesa como L1’s, com vista a verificar as consequências sintácticas das opções feitas com base nas especificações lexicais dos núcleos. Segundo a autora, a gramática inicial das crianças não apresenta marcação sistemática de género e número, verificando-se a primeira por volta dos 2;0 anos e a segunda cerca dos 2;4 anos.

<sup>90</sup> As questões em torno da possibilidade de a sequência DET+N ter sido fonologicamente interpretada como um todo indistinto (Koehn, 1994:49; DeGraff, 1999b:482; Faria, Freitas & Miguel, 2001:53) voltarão a ser debatidas no ponto 4.1.2.6.2 do presente trabalho (Relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN).

<sup>91</sup> Ao contrário do que sucede no PE, que admite a estrutura VSO, no PB e no PA, a alteração de ordem padrão que determina a estrutura SV acontece apenas com verbos inacusativos.

---

<sup>92</sup> Para Selinker (1992:226), recaída é sinónimo de aparecimento repentino de um sistema da interlíngua que há muito se encontrava ausente do desempenho do falante. O fenómeno afecta estruturas que pareciam ter sido erradicadas da gramática da interlíngua, mas é perfeitamente entendido pelo falante, que pode proceder à sua correcção sem necessidade da intervenção de terceiros. Este processo distingue-se da fossilização, uma vez que esta não é uma forma desviada esporádica, antes representa uma variação inalterável, o que impossibilita a sua correcção.

<sup>93</sup> Hale (1998) levantou a questão do “problema de Orwell”, que se debruçou sobre as dificuldades da aquisição de certos tipos de conhecimentos linguísticos, ainda que a exposição à evidência seja constante, ou seja, ainda que o *input* seja bastante frequente. Chomsky (1986a) retomaria o termo, estendendo a sua denominação a “problema de Freud ou Orwell” e aplicando-o ao desenvolvimento do indivíduo no domínio da vida social e das instituições políticas. A discussão acerca das lacunas registadas entre o *input* e *performance* da interlíngua não levou a fundo as considerações teóricas do “problema de Orwell”, até que Franceschina (2005) demonstrou a sua importância enquanto ferramenta heurística para a investigação que se debruça sobre o desenvolvimento da ASL.

<sup>94</sup> Schachter (1974:208) define a esquiva como a estratégia aplicada pelo aprendente para evitar determinadas estruturas que são difíceis para ele, fazendo pressupor que a ausência de certos erros no registo do falante de L2 não reflecte, necessariamente, a sua real competência interlinguística. Kleinman (1978:162) refere que a esquiva pode ter origem em factores como a personalidade, nervosismo, falta de confiança ou inibição, entre outros.

<sup>95</sup> Ellis (1994:351) propõe cinco motivações para ocorrerem fossilizações dos itens, regras e subsistemas linguísticos: transferência linguística (*language transfer*), transferência por treinamento (*transfer of training*), estratégias de aprendizagem da L2 (*strategies of learning*), estratégias de comunicação da L2 (*strategies of second language communication*) e sobregeneralização do material linguístico da LA (*overgeneralization of TL linguistic material*) A conjugação dos cinco processos de fossilização produzirá, então, uma competência interlinguística plenamente fossilizada.

<sup>96</sup> Em trabalho sobre a aquisição do género gramatical em gramáticas fossilizadas, Franceschina (2005), tendo como ponto de sustentação os quadros teóricos dos Princípios e Parâmetros e do Programa Minimalista, debruça-se sobre as causas que originam a divergência no estágio final do desenvolvimento morfossintáctico da L2. Partindo do pressuposto que a GU e o dispositivo computacional da faculdade da linguagem se encontram totalmente disponíveis no processo de ASL (Cook, 1994; White, 1989, 2003; Bruhn de Garavito & White, 2002; Montrul, 2004; Cook & Newson, 2007 [1988]), o conjunto de suposições da autora vai girar em volta tanto do “problema de Platão” (o *input* pobre) como do “problema de Orwell ou Freud” (o *input* rico), abordados por Chomsky (1986a) no estudo sobre o conhecimento da língua. As conclusões do trabalho apontam para a possibilidade de a variação registada por falantes não-nativos não ser tão aleatória como alguma literatura de ASL sugere. Da mesma forma que o “problema de Platão” pode ter influência na questão da variação, também o “problema de Orwell” é determinante para esta, permitindo a articulação entre ambos que se observe a forma como o conhecimento interior interage (ou falha na interacção) com os estímulos externos para determinar o desenvolvimento das gramáticas mentais. Assim, enquanto a aquisição de uma L1 fornece uma perspectiva privilegiada para se observar o primeiro problema, a visão de acordo com o “problema de Orwell” ajuda a analisar, perspicazmente, as questões inerentes à ASL por falantes adultos, dado que se pode entender como uma norma, e não como um desvio ou uma fossilização, a falha em alcançar a competência da LA em determinadas áreas.

<sup>97</sup> Para Selinker (1972:218), a transferência da L1 pode ser positiva ou negativa. A primeira diz respeito às similaridades repartidas por L1 e L2 e que contribuem para facilitar a ASL. Por seu lado, a transferência negativa refere-se às diferenças entre L1 e L2, sendo responsável pelas interferências da primeira na ASL.

<sup>98</sup> Este pressuposto da Teoria da Processabilidade vai ao encontro de postulados avançados previamente em estudos de ASL, que defendem que um determinado traço só poderá ser transferido se houver lugar na estrutura da interlíngua onde o mesmo possa ser encaixado (Andersen, 1983c:182).

<sup>99</sup> Sims (1989:70) atribui às fossilizações temporárias a designação de fossilizações suaves (*soft fossilizations*) ou *jellifications*.

<sup>100</sup> Scovel (1969:248) rotulou a fossilização fonológica de “fenómeno Joseph Conrad” em homenagem ao escritor britânico de origem polaca, cujo nome de baptismo era Józef Teodor Nałęcz Korzeniowski (Berdyczew, 3 de Dezembro de 1857 – Bishopboune, 3 de Agosto de 1924). Educado na Polónia ocupada pela Rússia, Conrad viajou para Marselha, onde, com a idade de 17 anos, iniciou uma carreira de marinheiro. Em 1878 passou a servir num barco britânico, escapando assim ao serviço militar russo e conseguindo a nacionalidade inglesa, em 1884. Aos 21 anos era já fluente em inglês, língua que mais tarde dominaria com excelência, mas mantendo sempre o sotaque polaco.

---

<sup>101</sup> As pesquisas na área da voz e emoção têm demonstrado que alegria, tristeza e raiva constituem o tridente das emoções básicas, podendo ser identificadas perceptualmente em diferentes línguas e culturas, o que aponta para a possibilidade de a expressão emocional e a percepção serem regidas por determinados princípios universais. Contudo, no caso específico do português, os estudos têm evidenciado certas particularidades desta língua, que a distinguem de outras, como o inglês, o alemão ou o finlandês. De facto, no português ocorre uma grande aproximação entre os valores de emoções pertencentes a famílias distintas, no caso a expressão neutra, a tristeza e a alegria, possivelmente por razões de carácter cultural (Nunes, 2009:279). Em resultado de tal, as produções de falantes nativos do português que envolvam estas três emoções poderão não ser devidamente descodificadas por indivíduos em situação de aquisição do PtgL2, principalmente se forem detentores de códigos linguístico-culturais que se afastam bastante dos da comunidade falante da LA. Retoma-se aqui, então, a questão das fossilizações terem também como motivação o grau de aculturação do aprendente em relação ao grupo detentor da LA (*The Acculturation Model*, Schumann, 1978:34).

<sup>102</sup> Para uma discussão detalhada em torno da possibilidade de os crioulos registarem casos de flexão contextual, como a marcação tonal em certas construções com verbos seriais no saramacano (Good, 2003), as formas verbais curtas e longas no crioulo de base francesa maurício para fazer a distinção de verbos transitivos usados com ou sem objecto, a concordância de género em determinadas orações copulativas no crioulo do Haiti (DeGraff, 2001a; Lefebvre, 2003), a distinção entre pronomes sujeito e objecto no crioulo de Tobago (James, 2003) ou a marcação de caso em determinadas estruturas do crioulo caraió de base inglesa, do saramacano (Veenstra, 1994, 1996) e do crioulo do Haiti (Lefebvre, 1998; Sterlin, 1998), consulte-se Plag (2008a:126-128), já que este defende a possibilidade de se estar perante situações meramente aparentes, configurando todas elas casos efectivos de flexão inerente.

<sup>103</sup> A Hipótese da Sequência Não-marcada (*Unmarked Alignment Hypothesis*) postula que, na ASL, os aprendentes organizam inicialmente a sintaxe mapeando os papéis semânticos mais relevantes para o sujeito. Por seu lado, a expressão estrutural do sujeito ocupará a posição linear mais proeminente na estrutura-C, ou seja, a posição inicial (Pienemann, Biase & Kawaguchi, 2005:229).

<sup>104</sup> Para reforçar o postulado de que os crioulos são interlínguas convencionais de um estágio primário de aquisição, Plag (2008b) analisa ainda a questão das interrogativas e a realização de estruturas negativas, concluindo que, em ambos os casos, os padrões encontrados são sustentados em termos da HI. Quanto aos casos que têm sido entendidos como transferências pela literatura convencional, o autor entende que encontram sustento na Hipótese do Desenvolvimento Moderado da Transferência. Para uma leitura mais detalhada acerca destas questões, consulte-se Plag (2008b:320-324).

<sup>105</sup> Adoptando o quadro teórico da Função do Paradigma da Morfologia (*Paradigma Function Morphology*), que advoga que a flexão morfológica se realiza mais do que a lexical (Stump, 2001:14), Kihm (2010) estabelece comparações entre os paradigmas de flexão verbal dos crioulos da Alta Guiné e os crioulos do Golfo da Guiné, e entre os crioulos de base portuguesa da costa ocidental africana e os crioulos indo-portugueses, tendo em conta que a etapa mais evoluída de aquisição alcançada numa interlíngua relativamente estável por parte de aprendentes adultos, ou seja, a Variedade Básica (Klein & Perdue, 1997:303), é a extinção dos paradigmas flexionais e a não-realização correlativa no interior da estrutura das palavras. Kihm (2010:12) conclui que muitos dos traços morfossintácticos não são filtrados porque, simplesmente, categorias como tempo, aspecto ou número são indispensáveis para uma apropriada expressão de eventualidades a que os seres humanos necessitam de recorrer para comunicarem. Assim, alguns paradigmas limitados sobrevivem e os que se extinguem são também substituídos por paradigmas generalizados — e construções perifrásticas —, aspectos estes que contribuem especificamente para a formação dos pidgins e crioulos. Quanto à diversidade de crioulos portugueses que emergiu quer na Índia quer na costa ocidental de África, tal ficar-se-á a dever, segundo o autor, ao contacto entre o português e diferentes substratos e/ou adstratos.

<sup>106</sup> O European Science Foundation Second Language contém um arquivo computadorizado de dados coligidos por grupos de pesquisadores do European Science Foundation Project em cinco países europeus: França, Alemanha, Grã-Bretanha, Holanda e Suécia. O projecto analisa a aquisição espontânea de L2's por parte de quarenta imigrantes adultos que vivem na Europa Ocidental, bem como o código linguístico utilizado na comunicação com os falantes nativos dos seus países de acolhimento.

<sup>107</sup> Mateus & Villalva (2006:27), por exemplo, chamam a atenção para um caso de mudança que poderá estar em curso no PE, afectando a 2ª pessoa do singular do Pretérito Perfeito Simple e concorrendo para a regularização do paradigma da 2ª pessoa do singular. De facto, este é o único tempo verbal cuja 2ª pessoa do singular não termina em /s/. Daí que muitos falantes maternos utilizem, inconscientemente, as formas “*comprastes*”, em vez de “*compraste*”, “*lestes*”, em lugar de “*leste*”, ou “*dormistes*”, substituindo “*dormiste*”.

<sup>108</sup> O fanakalo, um pidgin de base zulu, inglesa e africânder, viu o seu uso extravasar para lá das zonas mineiras da África do Sul, criando pequenas bolsas de falantes na República Democrática do Congo, Namíbia, Zâmbia e Zimbábue. Apesar de usado apenas como L2, o fanakalo é utilizado como língua franca por centenas de milhares de falantes de diversas L1's. O fanakalo é também um dos poucos pidgins que tem como base uma língua não colonizadora e o único pidgin de base zulu. O nome da língua tem como étimo uma truncção de termos de origem ngungi e significa “faz isso assim”, reflectindo o uso funcional deste pidgin. Informações detalhadas sobre o fanakalo podem ser consultadas em Adendorff (2002).

<sup>109</sup> Os movimentos visíveis (*overt movements*), como a deslocação de itens do SN, podem também ocorrer nos crioulos, mesmo que não estejam presentes na sua língua lexificadora:

Exs.: FRANCÊS:

<i>Le</i>	<i>chat</i>	<i>qui</i>	<i>est</i>	<i>assis</i>	<i>sur</i>	<i>ma</i>	<i>table.</i>
DET	gato	REL	está	sentado	LOC	1SG	mesa.

“O gato que está sentado na minha mesa”

CRIOULO DO HAITI:

<i>Chat</i>	<i>ki</i>	<i>chita</i>	<i>nan</i>	<i>tab</i>	<i>mwen</i>	<i>an.</i>	(Kihm, 2008:416)
Gato	REL	senta	LOC	mesa	1SG	DET	

“O gato que está sentado na minha mesa.”

<sup>110</sup> Note-se que Chomsky (1998) conclui que, no fundo, não existe “imperfeição” no deslocamento dos itens, já que esta é a forma de a sintaxe se tornar legível às componentes fonológica e conceptual.

<sup>111</sup> No que concerne à ausência de movimento aparente do verbo para domínios funcionais mais altos nos crioulos, Roberts (1997:12-13) argumenta que, nestes, os verbos nunca se deslocam da sua posição base, enquanto núcleo do SV. Assim, a adjunção do verbo para um módulo mais alto e/ou um núcleo de concordância (“movimento de V-para-I”) encontra-se impossibilitada de acontecer, como sucede nos crioulos de base romance, ainda que o movimento exista na língua lexificadora. Contudo, Kihm (2008:414) entende que o argumento de Roberts não colhe para provar que a posição do verbo serve de pista para demonstrar a não-marcação sintáctica dos crioulos, já que não pode ser aplicado aos crioulos de base lexical francesa, nem mesmo aos CP's, conforme, na linha de Roberts (1997), o haviam postulado DeGraff (1997), para o crioulo do Haiti ou outros crioulos de base francesa e Baptista (2002), para o CCV. Assim sendo, os crioulos também não diferem das outras línguas no que respeita ao pressuposto do “movimento do verbo de V-para-I”.

<sup>112</sup> A argumentação de Roberts (1999) acerca da ausência do “movimento do verbo de V-para-I” nos crioulos de base francesa e sua comparação com idêntica ocorrência na língua inglesa assenta na Teoria dos Traços, que defende que a FLEX contém traços verbais fortes ou fracos, a fim de poder fazer a verificação das propriedades do verbo seleccionado a partir do léxico (Chomsky, 1993:28). Portanto, os traços do infinitivo e dos verbos auxiliares são interpretados na PF, mas como não estão directamente associados às matrizes fonológicas têm de atrair visivelmente o verbo principal para, com ele, produzirem o *spell out*. Contrariamente, os traços fracos, como acontece no inglês, não têm visibilidade fonológica, pelo que não necessitam de atrair visivelmente o verbo auxiliar, a fim de este ser eliminado. Estabelecendo correlação com o poder distintivo dos traços do verbo há ainda a relativa “riqueza” da flexão verbal. Este aspecto determina que os verbos *have* e *be*, em virtude de terem uma flexão mais rica do que os verbos auxiliares, se movimentem abertamente para o nóculo FLEX no inglês moderno. Esta diferença não ocorre no francês e no português, uma vez que os seus verbos possuem flexão rica.

<sup>113</sup> Note-se que o ADV funciona como especificador ou como adjunto, que é um predicado secundário. Desta forma, não é classificado com os traços das quatro categorias lexicais ou temáticas (N, V, A, F), dado que não pertence a uma categoria principal.

<sup>114</sup> A complexidade apresentada pelo posicionamento do ADV na cadeia frásica é também uma realidade no francês. De facto, Kihm (2008:420) considera como parciais as generalizações de Pollock (1989) sobre o francês, já que este entende que o “movimento do verbo de V-para-I” fixa o ADV na posição pós-verbal, daí a agramaticalidade sugerida para a frase “\**Jean souvent embrasse Marie*”. Contudo, Kihm (2008), enquanto falante nativo do francês, não considera agramatical a construção (exemplo [236b]) e outras realizações que possibilitam a colocação pré-verbal do ADV: “*Maria souvent embrassait son père avant de partir pour l'école*” (Kihm, 2008:421).

<sup>115</sup> Kihm (2008:421-422), contrariando o postulado de Roberts (1999), lista também evidências de que o posicionamento do ADV no crioulo de base francesa do Haiti é inconclusivo, principalmente quando ocorre conjuntamente com os marcadores de TMA:



---

Ex.: CRIOULO DO HAITI:

Minis la (deja) ap (deja) l diskou a (deja). (Kihm, 2008:421)  
Ministro DET (já) ASP(já) ler discurso DET (já)  
“O ministro já está a ler o discurso”

<sup>116</sup> *Projecto Vertentes do Português Rural do Estado da Bahia – Fundamentos: O Conceito de Transmissão Lingüística Irregular* [em linha]. Disponível em: <http://www.vertentes.ufba.br/conceito.htm> [Consultado 04 Dez. 2009].

<sup>117</sup> “*This facilitates substantial comparison and, in many respects, avoids ‘reinventing the wheel’*” (Baxter, 2004:105).

<sup>118</sup> O processo de mudança linguística só se completa, por norma, na terceira geração. A primeira geração é caracterizada pelo bilinguismo em estado incipiente, já que o discurso dos falantes revela bastantes traços da L1 e da criação de formas analógicas e originais. Seguidamente, a segunda geração demonstra já indícios de uma mudança parcial e, por fim, a terceira geração evidencia a mudança completa.

<sup>119</sup> Neste caso, norma será também entendida como um dos dialectos da língua, enquanto as variedades regionais e de sistemas linguísticos deterão estatuto idêntico ao da norma-padrão, isto é, não referem formas diferentes ou desprestigiadas de falar a língua, antes reflectem tão só a forma de falar a língua numa determinada região.

<sup>120</sup> *Projecto vertentes do português rural do Estado da Bahia. Chave de transcrição do projeto vertentes.* [em linha]. Disponível em: <http://www.vertentes.ufba.br/chave.htm>.

<sup>121</sup> Para uma completa identificação das convenções ortográficas aplicadas na transcrição dos nossos dados, consulte-se Tagliamonte (2006:56-64).

<sup>122</sup> As comparações com o Português Vernacular de Moçambique têm como base os achados apresentados por Jon-And (2008) no *Congresso Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE), Universidade de Lisboa, Portugal, 03-05 de Julho, 2008*, e que serão parte integrante da dissertação de doutoramento ainda em processo de elaboração.

<sup>123</sup> As comparações com o Português de Cabo Verde têm como base os achados apresentados por Jon-And (2009) no *Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE), Universidade de Colónia, Alemanha, 11-15 de Agosto, 2009*, e que serão também parte integrante da dissertação de doutoramento ainda em processo de elaboração.

<sup>124</sup> As faixas etárias registam, entre si, um intervalo temporal de vinte anos, período suficiente para possibilitar um estudo das mudanças ocorridas no tempo aparente, enquanto projecção do tempo real. Os conceitos de tempo real e tempo aparente, bem como a aplicação prática do segundo no estudo dos processos de mudança, encontram-se amplamente descritos em Lucchesi (2000a:277-280) e P. Andrade (2003:80-82).

<sup>125</sup> A *Lex Parsimoniae*, também conhecida como “Navalha de Occam” ou “Navalha de Ockham”, foi primeiramente preconizada pelo frade franciscano inglês William de Ockham, que defendeu que a explicação para qualquer fenómeno científico deve assentar na economia, parcimónia e simplicidade.

<sup>126</sup> “*Linguistic variables which are frequent but have minimal variation are less viable for investigation by this method. Although the structures themselves may be interesting, if the data at your disposal is near categorical (either 100 per cent or 0 per cent), then there is little room for quantitative investigation. (...) In this case, you may rely on the constraint ranking of factors for comparative purposes*” (Tagliamonte, 2006:86).

<sup>127</sup> O valor do *log-likelihood* representa a qualidade da aproximação entre o modelo constituído (os factores que caracterizam os contextos, os pesos que indicam o efeito dos factores sobre o uso da variação investigada, o *input* representativo do nível de uso de determinado valor da variável dependente e o modelo matemático logístico) e os dados submetidos à análise.

---

<sup>128</sup> O substantivo *aldeão*, por exemplo, possui três formas distintas de marcar o PL. Assim, se, por exemplo, tivéssemos que codificar este nome para o grupo de factores *saliência fónica*, seríamos obrigados a optar por três factores distintos, consoante a realização do informante: um para os plurais regulares (*aldeãos*); outro para os plurais –*ão/-ões* (*aldeões*); e outro ainda para os plurais –*ão/-ães* (*aldeães*). Caso ocorrências deste tipo acontecessem em número substancial no nosso *corpus*, interessaria confirmar se as mesmas se comportavam de um modo livre, isto é, se eram aplicadas pelos falantes de modo aleatório, ou, em alternativa, eram utilizadas segundo um padrão ordenado e homogêneo, constituindo, portanto, uma variação instalada na comunidade. Para estudo da segunda situação interessaria, então, tal como preconizado por Labov, 1972a, 1982 e Weinreich, Labov & Herzog (2006 [1968]:100), observar a problemática do encaixamento que estaria subjacente à variação (*embedding problem*), ou seja, determinar quais seriam os factores da estrutura linguística que estariam a motivar a referida variação.

<sup>129</sup> Para uma explicação mais detalhada acerca da determinação da *aplicação* ou *não-aplicação da regra variável* na gramática generativo-transformacional, consulte-se Guy & Zilles (2007:144-145).

<sup>130</sup> No trabalho de estudo sobre o Português do Rio de Janeiro, Scherre (1988) considerou os itens com marcação fónica parcial de PL como elementos com marca de pluralidade (p.e. *mulhere*), enquanto Braga (1977) os havia codificado como singulares.

<sup>131</sup> Nina (1980:103), alegando que todos os casos de PL em –*es* apresentam um comportamento semelhante, funde-os num único factor.

<sup>132</sup> Para o grau de diferenciação fónica na relação singular/plural, Guy (1981a:185) optou por uma escala que estabelece factores distintos para os elementos terminados no singular em –*ão* e –*l*. Relativamente ao item “vez”, como referimos a propósito do critério por nós adoptado para a caracterização dos dados (ponto 3.1.4), o linguista criou também dois factores diferentes: um para a expressão “às vezes”; e outro para as restantes realizações em que intervém o referido item.

<sup>133</sup> À semelhança dos achados de Lucchesi (2000a) no estudo do género, Emmerich (1984), ao analisar o dialecto do Alto Xingu, adquirido por contacto, concluiu que os morfemas flexionais do verbo do português xinguno tendem a ser incorporados, primeiramente, quando são fonologicamente mais salientes.

<sup>134</sup> Partindo do pressuposto que terá havido uma aquisição distinta do português pelos escravos africanos e, conseqüentemente, pelos seus descendentes, Lopes (2001) acredita que os segundos poderão, de alguma forma, fazer uso de uma linguagem que contenha traços dos seus ancestrais. Como o fenótipo (aparência racial) não se tem revelado suficiente para identificar a ancestralidade do povo brasileiro, marcado por múltiplas misturas, Lopes (2001:155-164), para caracterizar melhor os seus informantes, decidiu constituir uma variável assente nos pressupostos das pesquisas científicas em genética levadas a cabo pela Universidade Federal da Bahia. Esta, para conseguir relacionar a origem étnico-racial com algumas patologias, optou por estudar o significado dos sobrenomes, já que tanto africanos como índios abandonaram os seus sobrenomes de origem para adoptarem os da língua portuguesa. A pesquisa acabou concluindo que os africanos revelavam apetência pela escolha de sobrenomes religiosos, enquanto os índios davam preferência aos sobrenomes portugueses não-religiosos, com conexão aos reinos animal ou vegetal. O número de sobrenomes contabilizados pela UFB provou a coerência do critério e, conseqüentemente, tornou a variável do estudo de Lopes (2001) estatisticamente confiável.

<sup>135</sup> P. Andrade (2003:102) considerou três factores para os elementos que pluralizam com inserção de –*s* final e alteração silábica: itens finalizados em –*l*, itens nasais em –*ão* que pluralizam em –*ões* e itens nasais em –*ão* que pluralizam em –*ães*. Em adição, o factor dos casos regulares foi constituído com elementos que terminam em vogal oral (p.e. *mesa* > *mesas*) ou nasal (p.e. *irmã* > *irmãs*). Estes últimos foram por nós considerados como itens nasais que pluralizam de modo regular (inserção de –*s* final).

<sup>136</sup> O novo grupo de factores *saliência fónica*, após as amalgamações levadas a cabo por P. Andrade (2003:103), passou a incluir os seguintes factores: (i) –*l*, –*r*, –*s*; plurais duplos e nasais sem alteração silábica; (ii) ditongos nasais com alteração; (iii) regulares. A propósito destas amalgamações, note-se que a autora não faz referência a quaisquer testes qui-quadrado, uma vez que só com recurso a estes se poderia confirmar se as amalgamações efectuadas constituíam uma distinção linguística válida para efeitos de análise.

<sup>137</sup> Para análise dos seus dados, Baxter (2004:109) constituiu factores de acordo com as seguintes características: (i) quatro factores considerando a presença ou ausência de tonicidade na sílaba PL, e que acompanham uma mudança no radical da palavra flexionada em número; (ii) um factor referente aos monossílabos átonos; (iii) um factor englobando plurais cujos singulares terminam em –*s*.

---

<sup>138</sup> Nos nossos dados verificam-se realizações com o nome “avós”, um caso de metafonia em que ocorre marca de feminino na realização da pluralização do masculino “avô” (p.e. *mia mãe e mia avó sempre dizia que meus avós...* [CARMOM1]). Este é também o único item metafônico cujo singular é oxítono, já que todos os outros possuem a tonicidade na penúltima sílaba, configurando ocorrências paroxítonas.

<sup>139</sup> Guy (1981a:292-293:) exemplifica os fenómenos de manutenção e supressão no português com recurso a alguns fenómenos deste idioma, como a queda intervocálica da consoante [ʃ] depois de uma vogal átona (p.e. *pejõrem > pior*), ou a sua manutenção depois de uma vogal tónica (p.e. *cũjum > cujo*). Para uma observação detalhada dos fenómenos de apócope, síncope e aférese de sílabas átonas do cabo-verdiano, sugere-se a consulta de Almada (1961).

<sup>140</sup> Realizações configurando o fenómeno apontado por Scherre (1988) e Lopes (2001) são também recorrentes no PA (p.e. *eu vim praqui ê tive Z'ONZE anos* [ANTOM1]).

<sup>141</sup> Braga, (1977:30-31), no seu estudo sobre a concordância PL no SN no Triângulo Mineiro, codifica já esta variável, apresentando um grupo binário de factores: *presença de marca formal plural no segmento precedente ao elemento analisado* e *ausência de marca formal plural no segmento precedente ao elemento analisado*. Braga (1977:39) conclui que o segundo factor favorece a marcação no item seguinte.

<sup>142</sup> Nas suas observações, Poplack (1980a) levou em conta a relação entre *posição linear* e *posição dos marcadores plurais precedentes*, mas avançou que, para um melhor entendimento acerca da CPL-var do EPR, a variável independente *marcas precedentes* deveria ser também analisada em parceria com a *classe gramatical*, o *segmento fonológico seguinte* e a *saliência fónica*.

<sup>143</sup> Lucchesi (2000a), nas pesquisas efectuadas sobre a variação da concordância no género de HEL-Ba, concluiu que a marcação mórfica predomina sobre a economia funcional da língua na aquisição/fixação das regras de concordância, contrariando-se, assim, o princípio de que o processo mental de associação da língua em agrupar formas semelhantes – o *Processo Paralelo* – conduz a uma economia de língua em que “*marcas levam a marcas e zeros levam a zeros*”. Assim, não será a marca no elemento anterior da cadeia linear que leva à marcação, mas sim a marcação mais explícita de género no núcleo que favorece a marcação do género na cadeia linear. Consequentemente, quanto mais explícitas forem as marcas morfológicas presentes na posição de núcleo, mais facilmente se dará a aquisição da regra de concordância nominal no interior do SN. Paralelamente, estes resultados opõem-se às teorias funcionalistas de Martinet (1955), que determinam que os determinantes e modificadores têm maior valor funcional na definição da marcação do género, visto serem os únicos que promovem a informação referencial nos casos, por exemplo, dos nomes comuns de dois géneros. Como tal, Lucchesi (2000a) conclui que a informação se perde durante a conversação, caso não seja recuperada pelos elementos contextuais ou pragmáticos.

<sup>144</sup> Guy (1981a:146) ancora a explicação do fenómeno em estudos sobre percepção da fala, os quais avançam que as fricativas não-labiais surdas favorecem a permanência dos traços não-sonoros, uma vez que são mais distintas e mais perceptíveis do que as sonoras correspondentes.

<sup>145</sup> A constituição de dois factores para a pausa (pausa interna e pausa final), “*teve como objetivo verificar possíveis diferenças neste tipo de condicionamento de comportamento considerado arbitrário*” (Scherre, 1988:247). Assim, a pausa interna é configurada pela distância de realização fónica que trava a interferência entre sons na elocução, isto é, entre o item analisado e o elemento que se encontra em posição posterior a ele. Por seu lado, a pausa final é bastante mais longa e marca o fim de uma sentença ou período.

<sup>146</sup> De acordo com Bechara (1999 [1928]:343), a *haplogogia* ou *braquilogia* é um fenómeno comum que visa “*evitar reduplicações de sílabas que pertencem à palavra primitiva e ao sufixo*”, como em: *caridad(e) + oso > caridoso* (por *caridadoso*) ou *bondad(e) + oso > bondoso* (por *bondadoso*). Este tipo de simplificação pode ainda ocorrer em palavras formadas por composição: *trágico + cómico > tragicómico*.

<sup>147</sup> Os valores finais de Scherre (1988:249) para a pausa dizem respeito a apenas um factor, já que a linguista acabou amalgamando as pausas interna e final num único factor.

<sup>148</sup> A propósito dos seus achados, Scherre (1988:256) chama a atenção para o facto de o morfema marca de PL –s apresentar um comportamento totalmente diferente do –s não-morfémico do português, cujo cancelamento não é afectado pela pausa, ficando antes a dever-se aos itens surdos e às consoantes. Como tal, as regras que condicionam a inserção de ambos os –s terão de ser explicadas com base em pressupostos distintos.

<sup>149</sup> Igual critério foi adoptado em relação aos artigos, em que nos abstivemos de os classificar em determinantes artigos definidos ou determinantes artigos indefinidos.

---

<sup>150</sup> Lopes (2001:83) classificou como artigos indefinidos todas as formas “um” e “uns” em que não era nítida a distinção entre pronome ou artigo indefinido.

<sup>151</sup> Scherre (1988:150) tratou estes itens como pronome pessoal de 3a. pessoa, ao passo que Lopes (2001:85) os considerou como outras formas de indicar posse, incluindo-os, tal como nós, na categoria dos possessivos.

<sup>152</sup> Lopes (2001:85) e P. Andrade (2003:97) apenas consideraram como quantificador os indefinidos invariável *tudo* e variável *todo* (e suas flexões).

<sup>153</sup> Factores constituídos por Baxter (2004:115-117) para análise geracional da influência da variável *posição linear x posição em relação ao núcleo do SN* na marcação PL do PT: (i) item PL à esquerda, adjacente ao núcleo; (ii) item PL à esquerda, não-adjacente ao núcleo; (iii) núcleo em 1ª. posição; (iv) núcleo em 2ª. posição; (v) núcleo em 3ª. posição; (vi) núcleo em 4ª. posição; (vii) núcleo em 5ª. posição; (viii) item PL à direita do núcleo, 2ª. posição no SN; (ix) item PL à direita do núcleo, 3ª. posição no SN; (x) item PL à direita do núcleo, 4ª. posição no SN; (xi) item PL à direita do núcleo, 5ª. posição no SN.

<sup>154</sup> Factores elaborados por Jon-And (2008) para análise da influência da variável *posição linear x posição em relação ao núcleo do SN* na marcação PL do PMp: (i) elemento imediatamente anterior ao núcleo; (ii) elemento anterior ao núcleo, mas não imediatamente; (iii) núcleo em 1ª. posição; (iv) núcleo em 2ª. posição; (v) núcleo em 3ª. posição; (vi) núcleo nas demais posições; (vii) elemento posterior ao núcleo, 2ª. posição no SN; (viii) elemento posterior ao núcleo, 3ª. posição no SN; (ix) elemento posterior ao núcleo, demais posições.

<sup>155</sup> O falante refere-se à antiga unidade monetária portuguesa correspondente a mil escudos, isto é, ao conto de réis.

<sup>156</sup> Labov (1981) considera que o lapso de tempo ideal para se detectar uma mudança linguística em curso será o que corresponde, no mínimo, a metade de uma geração e, no máximo, a duas gerações, isto é, que contempla um intervalo temporal configurado entre 12 e 50 anos.

<sup>157</sup> Face ao exposto, resulta que a variável idade, por exemplo, pode ser tratada sob duas perspectivas: i) a da variável contínua; II) a da variável *continuum* de idades, segmentada em grupos etários distintos (a variável pela qual optámos no presente trabalho).

<sup>158</sup> Lembre-se que o grupo de factores tido como verdadeiramente fiável é o do tipo binário, que apenas permite duas alternativas. Como tal, a constituição de variáveis independentes com elevado número de factores viola este pressuposto e poderá apresentar problemas analíticos sérios devido à provável distribuição desequilibrada de dados, os quais irão ferir a ortogonalidade do grupo de factores em que estão inseridos.

<sup>159</sup> A partilha de factores idênticos por dois ou mais grupos de factores vai determinar que “*tudo o que é codificado como x também é codificado como y. Nesse caso é impossível para o programa testar como os dados do contexto x se comportam quando estão na presença do factor y, e vice-versa*” (Guy & Zilles, 2007:199).

<sup>160</sup> No ponto 3.6.7. fizemos alusão aos critérios que levaram à constituição da variável independente *marcação de género* e seus factores. Quanto ao número de ocorrências produzido para cada factor constituído, bem como os percentuais de marcação PL motivados pelos mesmos, poderão ser consultados na Tabela 12.8. Como se pode constatar, os percentuais de inserção de pluralização rondam todos o patamar da neutralidade, daí o grupo de factores *marcação de género* ter sido considerado irrelevante pela ferramenta VARBRUL, enquanto variável que intervém na configuração da CPL-var do SN do PA. Deste modo, relativamente à hipótese 8 que constituímos, verifica-se que a marcação de género não constitui condição *sine qua non* para conduzir à marcação de número, quando esta é exigida em itens flexionados em género, isto é, que a aplicação de uma primeira flexão, a do género, seja ela regular ou irregular, não implica necessariamente a aplicação de uma flexão posterior, a do número.



A CONCORDÂNCIA PLURAL VARIÁVEL NO SINTAGMA NOMINAL DO PORTUGUÊS  
REESTRUTURADO DA COMUNIDADE DE ALMOXARIFE, SÃO TOMÉ  
(Desenvolvimento das Regras de Concordância Variáveis no Processo de  
Transmissão-Aquisição Geracional)  
Vol. 2

por

**Carlos Filipe Guimarães Figueiredo**

Dissertação de Doutoramento em Linguística

2010



FACULDADE DE CIÊNCIAS SOCIAIS E HUMANAS

DEPARTAMENTO DE PORTUGUÊS

**UNIVERSIDADE DE MACAU**

A CONCORDÂNCIA PLURAL VARIÁVEL NO SINTAGMA NOMINAL DO PORTUGUÊS  
REESTRUTURADO DA COMUNIDADE DE ALMOXARIFE, SÃO TOMÉ  
(Desenvolvimento das Regras de Concordância Variáveis no Processo de  
Transmissão-Aquisição Geracional)  
Vol. 2

por

**Carlos Filipe Guimarães Figueiredo**

**Orientador:** Professor Doutor Alan Norman Baxter

Departamento de Português

Dissertação em Linguística para obtenção do grau de Doutor.

2010

Faculdade de Ciências Sociais e Humanas

UNIVERSIDADE DE MACAU

## CAPÍTULO 4

*“Todas as pressões funcional-adaptativas que dão forma à estrutura sincrónica – idealizada – da linguagem são exercidas durante o desempenho efectivo. É onde a linguagem é adquirida e onde a gramática emerge e muda. É onde a forma se ajusta – criativamente e sob a impulsão da construção oportunista momentânea do contexto – a novas funções e sentidos distendidos. É também onde a variação e indeterminação são os ingredientes necessários do mecanismo efectivo que modela e remodela a competência.”*

(In Talmy Givón, *Syntax*, p.6)

### **Análise dos resultados**

#### **(Variável dependente e variáveis independentes do tipo estrutural)**

Levando em conta as diferentes variáveis independentes linguísticas e extralinguísticas constituídas no capítulo anterior, bem como as hipóteses levantadas para todas elas, a confirmação ou não-confirmação destas passará, a partir de agora, a ser o foco da nossa atenção. Para tanto, iremos interpretar os resultados que nos foram fornecidos pelo suporte computacional VARBRUL, avaliando quais os grupos de factores e respectivos factores que se revelam significativos para o ponto de inserção da marcação PL no SN do PA e, conseqüentemente, para o desenho da sua CPL-var. A tendência para a incorporação, desenvolvimento e aplicação da CPL-var no SN do PA ao longo do processo de aquisição deste só poderá ser entendida caso se tenham também em consideração, numa perspectiva comparada, os resultados dos estudos apontados no capítulo 1 e levados a cabo anteriormente para análise da CPL-var no SN de variedades africanas e brasileiras de português, por um lado, e, por outro lado, os estudos indicados no capítulo 2 sobre a forma como se processa a aquisição das línguas, sejam eles na perspectiva generativista da intervenção da GU nesta sejam eles psico e/ou sociolinguísticos. Assim, uma correcta análise e interpretação dos dados contribuirá quer para se compreender melhor o comportamento da concordância PL em variedades reestruturadas de português que tenham um CP como substrato quer para se constatar se estas manifestam padrão de variação semelhante ao de variedades reestruturadas de português cujas línguas ancestrais não são CP's. Estes aspectos poderão, por seu lado, ajudar não só a determinar em que tipo de estruturas é que os aprendentes do PtgL2, que tenham línguas crioulas ou do grupo níger-congo atlântico como L1's, revelam tendência para colocar as marcas formais de PL, mas também a clarificar se a CPL-var de ambos os tipos de variedades de português se desenvolveu a partir de mecanismos



sintácticos dos ancestrais substratos africanos (Guy, 1981a:301-302), ou se, pelo contrário, se ficou a dever a processos de crioulição independentes de tais mecanismos, como a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a). Paralelamente, será ainda possível observar até que ponto a variação inerente a uma L1 reestruturada e compartilhada por uma determinada comunidade reflecte, ou não, aquisição de acordo com os princípios da GU.

Dado o número de variáveis independentes constituídas e a extensão de resultados fornecidos pela ferramenta VARBRUL, optámos por agrupar as primeiras de acordo com os traços principais que possam compartilhar e observar os seus resultados em capítulos distintos. Assim, no presente capítulo, iremos fazer alusão aos resultados da variável dependente do tipo atomístico constituída e incidir também a nossa atenção sobre as seis variáveis independentes linguísticas do tipo estrutural, apresentando e comentando os resultados dos seguintes grupos de factores: 1 – *Posição do item analisado em relação ao núcleo do sintagma nominal*; 2 – *Ordem do item na cadeia dos constituintes flexionáveis*; 3 – *Classe gramatical*; 4 – *Marcas precedentes*; 5 – *Grau de concordância de número no sintagma nominal*; 6 – *Posição do item na cadeia do sintagma nominal (Posição linear)*.

Quanto ao capítulo 5, será reservado à análise dos resultados tanto das restantes variáveis independentes como das variáveis extralinguísticas. Por seu lado, a observação e análise dos resultados das variáveis serão efectuadas respeitando-se a ordem decrescente de importância determinada pela ferramenta VARBRUL (Tabela 14.1). Em ambos os capítulos apresentar-se-ão, igualmente, os passos seguidos no tratamento das variáveis, quando as mesmas necessitarem de procedimentos específicos, nomeadamente para a constituição final dos seus factores e observação das relações estabelecidas com outras variáveis na influência da marcação PL. Por fim, e após comparação dos nossos resultados com os de outros estudos, verificaremos se a variável em análise confirma ou não a hipótese que procurámos testar e se esta se enquadra na teorização linguística já proposta. Para tanto, procuraremos também apresentar explicações que entendemos serem plausíveis e coadunáveis às ocorrências em questão.

## 4.1. Resultados das variáveis constituídas e respectivas análises

### 4.1.1. Variável dependente

No *corpus* que recolhemos, foram coligidas 2340 ocorrências passíveis de marcação PL, com esta revelando-se em 1201 elementos e a inibição afectando 1138 itens:

**Tabela 15.1.** *Variável dependente*: percentuais de marcas PL nos itens do SN do PA.

<i>Input desta rodada</i> : 0,504		
<i>Log-likelihood</i> : -981,009		
<i>Significância</i> : 0,011		
<b>Factores</b>	<b>Nr. de Ocorrências</b>	<b>%</b>
Presença de morfema de plural no item analisado	1.202/2.340	51
Ausência de morfema de plural no item analisado	1.138/2.340	49

A comunidade de Almojarife apresenta um grande equilíbrio entre marcação e não-marcação PL nos itens que compõem o SN, já que a diferença de percentagem entre os elementos flexionados (51%) e não-flexionados em número (49%) é mínima. O valor do *input* (0,504) mostra igualmente pouca discrepância entre a marcação e não-marcação PL.

Efectuámos igualmente uma rodada extra, tipo não-atomístico,<sup>161</sup> com todos os dados codificados para o efeito, apenas para observarmos também a concordância plena entre todos os elementos que compõem o SN. Os resultados revelam que a concordância entre todos os itens do SN cai drasticamente, sendo realizada em apenas 425 dos 1.488 SN's produzidos pela comunidade:

**Tabela 15.2.** Percentuais de SN's plenamente marcados no PA.

<i>Input desta rodada</i> : 0,223		
<i>Log-likelihood</i> : -694,131		
<i>Significância</i> : 0,001		
<b>Factores</b>	<b>Nr. de ocorrências</b>	<b>%</b>
Sintagma nominal plural com concordância plena	425/1.488	29
Sintagma nominal plural sem concordância plena	1.056/1.488	71

Os resultados evidenciam o elevado grau CPL-var no SN do PA, confirmando que a mesma, para além de consistente e sistemática, se demarca do padrão de pluralização do PE. Dos trabalhos sobre CPL-var no SN, cujos resultados comparamos com os nossos, apenas o de P. Andrade (2003:195), sobre a fala de HEL-Ba, apresenta resultados percentuais para os SN's com marcação plena de PL. Nestes, a não-inserção da regra da

concordância é categórica, enquanto no PA, a tendência é apenas para a inibição acentuada, sendo as marcas aplicadas plenamente em cerca de 1/3 dos seus SN's:

**Tabela 15.3.** Percentuais de SN's plenamente marcados: 2 variedades de português.

Factores	PA			HEL-Ba		
	Nr.	%	Input	Nr.	%	Input
Sintagma nominal plural com concordância plena	425/1.488	29	0,281	133/1.434	9	0,6
Sintagma nominal plural sem concordância plena	1.056/1.488	71		1.301/1.434	91	

Ao que os resultados indicam, a fala de HEL-Ba encontrar-se-á em estágio final de mudança (Lucchesi, 2000a; P. Andrade, 2003), tudo apontando para que a CPL-var se venha a constituir como uma das marcas que fará a distinção entre este dialecto e o PE. Quanto ao PA, parece estar ainda em situação de mudança em curso, fruto, talvez, do seu convívio mais prolongado com o PE. Contudo, sobre esta questão nos debruçaremos mais atentamente aquando da análise dos resultados das variáveis independentes extralinguísticas (capítulo 5).

Relativamente aos percentuais de marcação observados numa perspectiva atomística, pode adiantar-se que os nossos dados, à excepção do que sucede relativamente ao dialecto de HEL-Ba, indiciam um estágio menos avançado na inserção da pluralização nos itens flexionáveis em número do SN, quando confrontados com os das outras variedades de português que nos propomos comparar, nas quais a percentagens de marcação são sempre superiores às do PA (Tabela 15.4):

**Tabela 15.4.** Percentuais de marcas PL nos itens do SN: 7 variedades de português.

Factores	PA			MRJ			NURC		
	Nr.	%	Input	Nr.	%	Input	Nr.	%	Input
Itens marcados	1.202/2.340	51	0,505	7.978/11.086	72	-	11.251/13.906	81	-
Itens não marcados	1.138/2.340	49		3.108/11.086	28		2.655/13.906	19	
Factores	HEL-Ba			PT			Pmp		
	Nr.	%	Input	Nr.	%	Input	Nr.	%	Input
Itens marcados	1.310/2.893	45	0,44	1.835/3.366	55	-	2.515/2.872	88	-
Itens não marcados	1.583/2.893	55		1.531/3.366	45		357/2.872	12	
Factores	PCV								
	Nr.	%	Input						
Itens marcados	2.359/2.285	82	-						
Itens não marcados	526/2.285	18	-						

Siglas: PA: Português de Almoarifê, São Tomé; MRJ: Município do Rio de Janeiro, Brasil (Scherre, 1988); NURC: Norma urbana culta de Salvador, Brasil (Lopes, 2001); HEL-Ba: dialecto de Helvécia-Bahia, Brasil (P. Andrade, 2003); PT: Português dos tongas, Roça Monte Café, São Tomé (Baxter, 2004); Pmp: Português do Maputo, Moçambique (Jon-And, 2008); PCV: Português do Mindelo, Cabo Verde (Jon-And, 2009).

#### 4.1.2. Variáveis independentes do tipo estrutural

Conforme mencionado no ponto 3.8 do presente trabalho, a propósito da quantificação dos dados para selecção das variáveis independentes a observar, a análise dos resultados não poderá deixar de considerar também as variáveis rejeitadas pela ferramenta VARBRUL, já que estas podem conter informação relevante para o estudo do fenómeno da marcação PL no SN do PA. Deste modo, apesar de o grupo de factores *marcação de género* ter sido considerado não relevante pelo suporte computacional,<sup>162</sup> não deixaremos de ter em consideração os valores apresentados (Tabela 12.8), caso necessitemos de fazer uso dos mesmos. Posto tal, passamos não só a descrever os passos dados no tratamento de cada variável independente linguística do tipo estrutural mas também a analisar e comentar os valores resultantes deste. A ordenação deste tipo de

variáveis independentes seguirá o critério hierárquico de selecção estabelecido pelo nosso suporte computacional e apresentado na Tabela 14.1.

#### 4.1.2.1. Variável independente *posição em relação ao núcleo do SN*

A Tabela 16.1 fornece uma ideia acerca do modo como os factores envasados na variável independente *posição do item em relação ao núcleo do SN* dão o seu contributo para a CPL-var de Almojarife:

**Tabela 16.1.** Efeito da *posição em relação ao núcleo do SN* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,523		<i>Log-likelihood:</i> -962,484		<i>Significância:</i> 0,036	
<b>Posição</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Núcleo	497/1.482	34	63	0,287	
Anterior ao núcleo, não imediatamente	57/87	66	4	0,591	
Imediatamente anterior ao núcleo	631/681	93	29	0,906	
Posterior ao núcleo	17/90	19	4	0,077	
Totais:		1.202/2.340	51	100	-

A posição imediatamente anterior ao núcleo do SN favorece categoricamente a inserção de marcas (pr. 0,906), logo seguida das posições anteriores, mas não imediatas, com um efeito praticamente neutro na pluralização (pr. 0,591). As restantes posições, sejam elas nucleares (pr. 0,287) ou à direita do núcleo (pr. 0,077), desfavorecem a marcação PL, as primeiras de modo acentuado, as segundas de forma bastante categórica. Contudo, esta variável, por si só, não nos fornece um panorama total da forma como alguns elementos se comportam nas diferentes posições, caso do núcleo do SN. Esta visão só será possível de se obter observando a relação que o grupo de factores *posição em relação ao núcleo do SN* estabelece com outras variáveis, como a *posição linear*, a *categoria gramatical* e a *posição na cadeia dos constituintes flexionáveis*. Com estes procedimentos, esperamos conseguir valores que nos permitam avaliar, extensamente, a configuração do SN do PA e, conseqüentemente, quais os pesos que as diferentes categorias gramaticais exercem no desenho da CPL-var, atendendo às diferentes posições que podem ocupar na cadeia sintagmática. Estes aspectos, bem como a análise dos resultados obtidos, encontram-se contemplados quer no ponto seguinte

quer no ponto 4.1.2.6 do presente trabalho, dedicado ao estudo da variável independente *posição linear* e à relação que a mesma estabelece com todos os grupos de factores mencionados.

#### **4.1.2.1.1. Relação entre *classe gramatical* e *posição em relação núcleo do SN***

A Tabela 16.2 apresenta-nos a distribuição dos diferentes itens gramaticais em relação ao núcleo do SN em que estão inseridos, bem como os respectivos pesos na inserção de marcas de número.

Os substantivos e categorias substantivadas não só ocorrem apenas em posição nuclear como apresentam também valores que inibem consideravelmente a marcação, mais os segundos (pr. 0,215) do que os primeiros (pr. 0,306). Este aspecto não surpreende, se considerarmos que, tal como preconizado por Guy (1981a:301-302) para o PB, também o PA apresenta tendência para marcar o PL na primeira oportunidade, devido à influência do sistema de pluralização dos substratos africanos, como se procurará demonstrar. Por conseguinte, a tendência para a marcação em número ocorrerá, por norma, em qualquer dos elementos que anteceda o núcleo do SN, excepto quando acontecer inexistência dos mesmos, isto é, quando o núcleo ocupar o primeiro lugar da cadeia sintagmática. Do mesmo modo, todos os outros itens gramaticais ocorrem em posições não nucleares, sejam elas pré ou pós, com excepção dos artigos e numerais, que apenas se posicionam em lugares antes do núcleo. Paralelamente, o grande número de ocorrências em posições nucleares (1.482 realizações) e anteriores, adjacentes ao núcleo (681 realizações), por oposição a diminutos casos nos factores anterior, não adjacente ao núcleo (87 realizações) e posterior ao núcleo (90 realizações) mostra como os SN's do PA são de estrutura reduzida, compondo-se, maioritariamente, por apenas dois elementos: DET+N.

Os adjectivos do tipo 1 ocorrem, na sua grande maioria, em posição pós-nuclear, onde apresentam um peso inibidor da marcação quase categórico (pr. 0,077). Nas posições pré-nucleares as suas ocorrências são escassas (6 realizações adjacentes ao núcleo e uma ocorrência não adjacente ao núcleo), mas com um percentual categórico de marcação (100%). Quanto ao outro género de adjectivos (adjectivos do tipo 2 e adjectivos no grau

diminutivo), apresentam dezasseis realizações em posições pós-nucleares, equitativamente repartidas entre ambos e sem qualquer marcação PL.

**Tabela 16.2.** Relação entre *classe gramatical e posição em relação ao núcleo do SN*: distribuição das classes gramaticais relativamente ao núcleo do SN.

<i>Input desta rodada: 0,535</i>	<i>Log-likelihood: -950,379</i>			<i>Significância: 0,009</i>		
<b>Classe Gramatical</b>	<b>Núcleo</b>			<b>Anterior, adjacente ao núcleo</b>		
	Nr. total	% PL	Peso relativo	Nr. total	% PL	Peso Relativo
Substantivo	480/1.428	34	0,306	-	-	-
Categoria substantivada	17/54	32	0,215	-	-	-
Adjectivo 1	-	-	-	(6/6)	100	-
Adjectivo 2	-	-	-	-	-	-
Adjectivo diminutivo	-	-	-	-	-	-
Artigo definido	-	-	-	231/250	92	0,720
Artigo indefinido	-	-	-	49/57	86	0,626
Numeral	-	-	-	(3/6)	50	-
Possessivo	-	-	-	28/34	82	0,819
Demonstrativo	-	-	-	171/178	96	0,974
Indefinido	-	-	-	57/59	97	0,953
Quantificador	-	-	-	85/91	93	0,937
<b>Totais:</b>	497/1.482	34	-	630/681	93	-
<b>Classe Gramatical</b>	<b>Anterior, não-adjacente</b>			<b>Posterior ao núcleo</b>		
	Nr. total	% PL	Peso relativo	Nr. total	% PL	Peso relativo
Substantivo	-	-	-	-	-	-
Categoria substantivada	-	-	-	-	-	-
Adjectivo 1	(1/1)	100	-	12/58	21	0,077
Adjectivo 2	-	-	-	(0/8)	0	-
Adjectivo diminutivo	-	-	-	(0/8)	0	-
Artigo definido	29/38	76	0,491	-	-	-
Artigo indefinido	(14/23)	61	-	-	-	-
Numeral	-	-	-	-	-	-
Possessivo	(1/1)	100	-	(2/2)	100	-
Demonstrativo	(2/6)	33	-	(1/2)	50	-
Indefinido	(4/4)	100	-	(1/1)	100	-
Quantificador	(7/14)	50	-	(1/11)	9	-
<b>Totais:</b>	57/87	66	-	17/90	19	-

Os artigos, por seu lado, evidenciam tendência para favorecerem a marcação, na sua clássica posição pré-nuclear. Tal não surpreende se levarmos em linha de conta que os SN's do PA são maioritariamente do tipo DET+N e que o seu sistema de marcação PL faz incidir a marcação nos elementos pré-nucleares, com especial incidência no item imediatamente antes do nome núcleo, mantendo-se inalterados os restantes elementos do SN. Estes aspectos levantam, uma vez mais, não só a possibilidade de este tipo de marcação PL reflectir o sistema de marcação PL das ancestrais línguas do grupo níger-congo atlântico, substratos do PA (cf. Guy, 1981a:301-302), que fazem incidir o papel fulcral PL no classificador nominal, mas também que o desenvolvimento da concordância PL se inicia com a introdução de um PL *singleton* na categoria funcional DET, que serve de âncora para o controlo da pluralização (cf. Baxter, 2009 – Fig. 1). Em sequência destes aspectos, os definidos apresentam um peso favorecedor da inserção de marcas (pr. 0,720) em 250 ocorrências em posição imediatamente anterior ao núcleo do SN, por oposição a 38 realizações em posição não-imediatamente anterior ao núcleo do SN, que exibem efeito neutro na pluralização (pr. 0,491). Quanto aos artigos indefinidos, exibem 57 ocorrências favorecendo ligeiramente a marcação em posição imediatamente anterior ao núcleo do SN (pr. 0,626), contra 23 realizações em posição anterior ao núcleo do SN, não-imediata, e cujo percentual de pluralização se situa na casa dos 63%.

Por seu lado, os numerais registam apenas seis ocorrências na sua clássica posição imediatamente pré-nuclear, que revelam um percentual equilibrado entre marcação e não-marcação (50% de marcação vs. 50% de inibição).

Quanto aos restantes elementos gramaticais, todos eles podendo configurar determinantes, é notória também a sua propensão para serem categoricamente marcados na posição imediatamente anterior ao núcleo do SN, reforçando os pressupostos levantados de este tipo de marcação ancorar no sistema das ancestrais línguas do grupo níger-congo atlântico do PA (cf. Guy, 1981a:301-302). Assim, os possessivos mostram 34 dados imediatamente pré-nucleares, com um considerável peso na inserção de marcas (pr. 0,819), por oposição a apenas 1 ocorrência à esquerda do núcleo, mas não adjacente, e 2 realizações pós-nucleares. Por seu lado, os demonstrativos evidenciam 178 dados em localização imediatamente pré-nuclear, sendo a sua marcação bastante categórica (pr. 0,974). Em posição pré-nuclear, mas não adjacente, o número de realizações cai



drasticamente (apenas seis), acentuando-se ainda mais em localizações pós-nucleares (unicamente duas).

No que concerne aos indefinidos, seguem o mesmo padrão dos possessivos e demonstrativos, apresentando 59 ocorrências categoricamente marcadas na posição imediatamente pré-nuclear (pr. 0,953) e apenas quatro dados em posição pré-nuclear não adjacente. Pós-nuclearmente, o número de realizações é praticamente nulo, já que se regista apenas uma. Por fim, os quantificadores são, entre os quatro últimos itens gramaticais referidos, os que exibem mais realizações nas posições pré-nucleares não adjacentes (14 dados) e pós-nucleares (11 dados). Note-se que o percentual de marcação nestas posições decai bastante nos lugares pós-nucleares (posições pré-nucleares não antecedentes = 50% de marcação; posições pós-nucleares = 9% de marcação). No entanto, qualquer destes valores está bem distante daqueles que são apresentados pelo mesmo item gramatical na posição imediatamente pré-nuclear, onde o elevado número de ocorrências (91) é também acompanhado por um altíssimo favorecimento da marcação (pr. 0,937).

Confirma-se então, em pleno, a hipótese 6, que levantámos para a variável independente *posição do item analisado em relação ao núcleo do SN* (ponto 3.6.5), isto é, que os elementos antepostos ao núcleo do SN, com especial incidência para os itens adjacentes, são mais marcados do que aqueles que lhe são pospostos. De facto, como se verificou, o núcleo do SN é posição exclusiva dos substantivos e categorias substantivadas, onde revelam um peso inibidor da marcação. Por seu lado, os artigos distribuem-se pelas posições pré-nucleares, com ocorrências e peso da pluralização mais acentuados na localização imediatamente pré-nuclear. Quanto aos numerais, situam-se apenas nesta posição, com um peso neutro de pluralização. Por fim, os restantes elementos espalham-se por todas as posições pré e pós-nucleares, mas com muito maior incidência no lugar imediatamente à esquerda do núcleo, onde o seu peso de inserção de marcas PL é igualmente bastante categórico. Estes aspectos vêm ainda tornar feliz o pressuposto avançado por Guy (1981b:179) de que ocorre relação entre posição e categoria gramatical, com influência na marcação PL, já que, em SN's de estrutura reduzida (2 ou 3 elementos), se estabelece um paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjectivos e terceira posição.

Por fim, note-se que o alto índice de favorecimento à marcação evidenciado por possessivos e adjetivos em posição imediatamente pré-nuclear, bem como o desfavorecimento à marcação apresentado pelos segundos no posicionamento pós-nuclear, vêm também contribuir a favor dos pressupostos em MD debatidos no ponto 2.9 (Concordância variável ao nível do SN), que advogam não só existir uma relação de definitude entre estes itens gramaticais e o nome mas também entre o *output* da sintaxe e o *input* da morfologia, com a segunda lendo a primeira. Deste modo, como a concordância PL é funcional, vai ser determinada no PA pelos morfemas do tipo *singleton*, residindo nestes a fonte de variação. De facto, detêm o estatuto de núcleo na posição imediatamente pré-nominal, onde incide a marcação de número, mas assumem a categoria de sintagma na localização pós-nominal, na qual a marcação PL é dispensada, em virtude de se revelar redundante (cf. Kiparsky, 1972:195).

#### **4.1.2.2. Variável independente *ordem do item na cadeia dos constituintes flexionáveis***

A Tabela 12.9 evidencia um reduzido número de ocorrências para o factor *quarto ou outro elemento* (apenas dez), sugerindo que se tente a sua amalgamação com o factor *terceiro elemento*, antes de se optar pela sua rejeição. Para tanto, efectuámos rodadas simples (*one-level analysis*)<sup>163</sup> a fim de se verificar se era possível levar a cabo a junção dos dois factores. A rodada com os dois factores não-amalgamados forneceu o *log-likelihood* -1010,063, enquanto a rodada com os mesmos factores, mas amalgamados, apresentou um *log-likelihood* inferior: -1008,439. Feitos os cálculos para o teste de significância do qui-quadrado, a diferença entre os dois *log-likelihoods* facultou os seguintes valores:  $-1010,063 - [-1008,439] = -1010,063 + 1008,439 = -1,624$ . Esquecendo a marcação negativa, multiplica-se este resultado por 2 e obtém-se o valor do qui-quadrado: 3,248. No caso presente, reduzimos dois factores a um único, o que dá 1 grau de liberdade. Tendo em conta que o nível de significância aceite para as análises em questão é de 0,05, a tabela de qui-quadrado mostra um valor de 3,841. Como o qui-quadrado encontrado para o nosso grupo de factores é inferior ao do previsto na tabela, conclui-se que as posições detidas pelos terceiro, quarto ou outro item flexionáveis na cadeia sintagmática não são estatisticamente significativas para a inserção de marcas PL,<sup>164</sup> logo, representam uma distinção linguística que precisa de ser mantida na análise.

Assim sendo, como a distinção original entre os factores *terceiro elemento* e *quarto ou outro elemento* flexionáveis na cadeia sintagmática não era significativa, deve ser abandonada e substituída pela nova combinação de factores. Deste modo, a variável independente *ordem do item na cadeia dos constituintes flexionáveis* fica reduzida a apenas três factores, consoante Tabela 17.1, que inclui também os resultados da rodada VARBRUL com a constituição dos novos factores:

**Tabela 17.1.** Efeito da variável independente *ordem na cadeia dos constituintes flexionáveis* na marcação PL dos itens do SN.

<i>Input desta rodada:</i> 0,480		<i>Log-likelihood:</i> -1000,975		<i>Significância:</i> 0,017	
<b>Posição</b>	<b>Nr. total</b>	<b>% PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Primeiro elemento	886/1.487	60	64	0,727	
Segundo elemento	285/753	38	32	0,165	
Terceiro ou outro elemento	27/92	29	4	0,081	
Totais:	1.198/2.332	51	100	-	

Como se pode constatar, o primeiro elemento flexionável da cadeia sintagmática é o único que revela tendência para favorecer a marcação (pr. 0,727), independentemente do seu lugar na posição linear do SN. A marcação no segundo elemento flexionável cai depois drasticamente, passando para valores altamente inibidores da mesma (pr. 0,165) e levantando a possibilidade de marcas não levarem a marcas, contrariando o Princípio do Processamento Paralelo advogado por Scherre (1988:208). O valor verdadeiramente comprometedor da pluralização apresentado por estes itens, caso sejam segundo, terceiro ou outro elemento na cadeia dos constituintes flexionáveis do SN, faz também levantar a possibilidade de a concordância variável do PA ancorar em motivações de carácter estrutural, já que a interferência na cadeia de concordância parece não ser provocada por itens não flexionáveis que sejam eventualmente inseridos entre os elementos sujeitos a flexão.

Note-se que o peso do não-favorecimento se torna quase categórico no terceiro ou outro elemento flexionável do SN (pr. 0,081). Como em cadeias sintagmáticas reduzidas este elemento tem forte probabilidade de ocorrer pós-nuclearmente, reforça-se o já verificado anteriormente, isto é, que as posições à direita do núcleo são as que revelam tendência para apresentarem o mais elevado índice de inibição da marcação PL. Este

aspecto poderia argumentar em favor do Princípio do Processamento Paralelo de que zeros levam a zeros (Scherre, 1988:208), visto o segundo item flexionável também apresentar um baixo peso de inserção de marcas. Contudo, também é preciso ter em conta que este mesmo item pode apresentar marcação, caso seja, por exemplo, um possessivo e que, ainda assim, o terceiro elemento flexionável volta a mostrar altos valores de inibição. A discussão em torno desta questão será retomada no ponto 4.1.2.6 do presente trabalho, dedicada à análise do grupo de factores *posição linear*.

Por fim, refira-se que não foram incluídos numerais cardinais com variação de marcação (p.e. *cinco*, *duzentos*) na rodada que determinou os valores da Tabela 17.1, uma vez que os mesmos não são passíveis de flexão em número no português. Estamos, neste caso, perante casos pontuais de gramáticas em competição que não configuram realizações comunitárias. Os referidos numerais representam um total de seis ocorrências, distribuídas da seguinte forma: 3 realizações na primeira posição linear, com inibição indevida; e 3 casos indevidamente marcados, na segunda posição linear.

#### **4.1.2.2.1. Relação entre *ordem do item na cadeia dos constituintes flexionáveis e posição linear***

A fim de se tirarem dúvidas acerca da possibilidade de os itens não passíveis de marcação afectarem a cadeia de concordância, quando inseridos entre elementos passíveis de marcação, analisou-se a relação entre as variáveis independentes *ordem do item na cadeia dos constituintes flexionáveis e posição linear*. Desta forma se verificará, com mais exactidão, se os itens posteriores apresentam idêntico comportamento de marcação, sejam eles contíguos ou não ao elemento previamente marcado. Os resultados do cruzamento entre as referidas variáveis independentes estão patentes na Tabela 17.2.

O único factor que favorece categoricamente a inserção de marcas PL é o respeitante ao primeiro item flexionável em primeira posição da cadeia sintagmática (pr. 0,883). Se imaginarmos uma cadeia sintagmática sem inserção de elementos não flexionáveis entre os itens passíveis de flexão com, por exemplo, quatro elementos flexionáveis, observaríamos então que o primeiro teria uma elevadíssima probabilidade de ser marcado, decaindo a marcação do elemento flexionável em segunda posição para pesos inibidores (pr. 0,375). A inibição estender-se-ia aos restantes dois elementos flexionáveis, e de modo progressivo (terceiro elemento flexionável na terceira posição

sintagmática = pr. 0,173; quarto elemento flexionável na quarta posição do SN = 79% de probabilidade de não ser marcado). Nesta conformidade, e se atentarmos à diferença de pesos no propiciamento da marcação entre o primeiro e o segundo elemento flexionável da cadeia sintagmática, o que se verifica é que, inquestionavelmente, não se aplica o Princípio do Processamento Paralelo (Scherre, 1988:208), com marcas conduzindo a mais marcas. Por outro lado, este comportamento dos itens flexionáveis argumenta fortemente em favor do pressuposto de o PB (e, conseqüentemente, o PA) marcar o PL na primeira oportunidade onde se revele fundamental (Kiparsky, 1972:195), por influência do sistema de pluralização dos ancestrais substratos africanos (Guy, 1981a:3001-302), especialmente do grupo níger-congo atlântico (Baxter, 2009:292), acontecendo, depois, o apagamento formal da pluralidade nos outros elementos flexionáveis, em virtude de a mesma se tornar redundante, de acordo com o advogado pelo princípio funcionalista das Condições de Distintividade (Kiparsky, 1972:195).

**Tabela 17.2.** Efeito da relação entre *posição linear* e *ordem do item na cadeia dos constituintes flexionáveis* na marcação PL dos itens do SN.

<i>Input desta rodada: 0,511</i>		<i>Log-likelihood: -988,408</i>			<i>Significância: 0,037</i>		
<b>Ordem dos constituintes flexionáveis</b>	<b>1ª. posição linear</b>			<b>2ª. posição linear</b>			
	<b>Nr. Total</b>	<b>% PL</b>	<b>Peso Relativo</b>	<b>Nr. Total</b>	<b>% PL</b>	<b>Peso Relativo</b>	
Primeiro elemento	676/749	90	0,883	165/612	27	0,214	
Segundo elemento	-	-	-	274/699	39	0,375	
Terceiro ou outro elemento	-	-	-	-	-	-	
<b>Totais:</b>	676/749	90	-	439/1.311	33	-	
<b>Ordem dos constituintes flexionáveis</b>	<b>3ª. posição linear</b>			<b>4ª. ou outra posição linear</b>			
	<b>Nr. Total</b>	<b>% PL</b>	<b>Peso Relativo</b>	<b>Nr. Total</b>	<b>% PL</b>	<b>Peso Relativo</b>	
Primeiro elemento	16/51	31	0,237	29/75	39	0,226	
Segundo elemento	10/44	23	0,104	(1/10)	10	-	
Terceiro ou outro elemento	23/73	32	0,173	(4/19)	21	-	
<b>Totais:</b>	49/168	29	-	34/104	33	-	

Relativamente ao facto de o primeiro elemento flexionável poder apresentar inibição na segunda posição linear, quando deveria ser marcado por configurar a primeira oportunidade de se inserir a flexão de número PL, deve ser tido em conta que, por norma, este tipo de construções é introduzido por um numeral, portador de marca

semântica de pluralização, conforme se pode constatar nos exemplos apresentados na Tabela 12.9. Este aspecto levou-nos a fazer um levantamento destas ocorrências e, dos 612 itens que configuram casos de primeiro elemento passível de flexão em segunda posição linear, 609 são antecidos de numeral. Assim sendo, visto a marca semântica de PL estar já patente no elemento que antecede o primeiro item flexionável, não é inserida neste por se revelar redundante. Relembremos apenas que P. Andrade (2003:112) considera o numeral como sendo o item mais influente para a marcação na CPL-var, em virtude de ser um morfema semanticamente transparente, aspecto do qual discordamos parcialmente. A discussão em torno dos marcadores semânticos será abordada com mais pormenor no ponto em que se analisa a influência da variável independente *marcas precedentes* na marcação PL do SN do PA (ponto 4.1.2.4.do presente trabalho).

O mesmo tipo de panorama é aplicável ao primeiro item passível de flexão em terceira, quarta e outras posições, conforme pode ser constatado no número de realizações evidenciadas na Tabela 12.3, referente à variável independente *marcas precedentes*. De facto, com recurso aos exemplos constantes nesta tabela, é possível ver que os SN's mais extensos implicam, quase sempre, a ocorrência de numerais em posições antecedentes às dos itens flexionáveis.

Resta então referir que, relativamente à hipótese 9, levantada para a variável independente *ordem do item na cadeia dos constituintes flexionáveis* (ponto 3.6.8), não se confirma o Princípio do Processamento Paralelo de que zeros levam a zeros (Scherre, 1988:208). De facto, a ausência de marcação no elemento posterior ao item não flexionável que é inserido na cadeia sintagmática radica quer em motivações estruturais, isto é, encontrando paralelo, sim, no sistema de pluralização dos substratos africanos (Guy, 1981a:301-302) do grupo níger-congo atlântico (Baxter, 2009:292), que apenas insere a marcação no afixo pré-nominal, quer em motivações semânticas, isto é, no facto de o elemento antecedente ser um numeral, que contém a informação transparente de pluralidade (P. Andrade, 2003:112). Em correlação com estas duas motivações, vai também actuar o princípio funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que advoga que a pluralização nos elementos posteriores não é inserida, tal

como acontece nos idiomas do grupo níger-congo atlântico, pelo facto de se revelar redundante.

#### 4.1.2.3. Variável independente *classe gramatical*

A primeira rodada geral do VARBRUL (Tabela 12.5) revelou-nos que alguns factores desta variável independente apresentam ocorrências mínimas (*adjectivo 2, adjectivo grau diminutivo e numeral*) ou nulas (*substantivo grau diminutivo, substantivo grau aumentativo, adjectivo grau aumentativo, relativo e pronome pessoal recto*). Assim sendo, tais factores não foram incluídos na rodada que nos forneceu os pesos exercidos pelas diferentes classes gramaticais na inserção de marcas de número PL nos itens do SN (Tabela 18.1):

**Tabela 18.1.** Efeito da *classe gramatical* na marcação PL dos itens do SN.

<i>Input desta rodada:</i> 0,504	<i>Log-likelihood:</i> -981,009	<i>Significância:</i> 0,011		
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Substantivo grau normal	480/1.428	34	61	0,326
Categoria substantivada	17/54	32	2	0,228
Adjectivo 1	19/65	29	3	0,136
Artigo definido	260/288	90	12	0,688
Artigo indefinido	63/80	79	4	0,573
Possessivo	31/37	84	2	0,839
Demonstrativo	174/186	94	8	0,952
Indefinido	62/64	97	3	0,957
Quantificador	93/116	80	5	0,792
<b>Totais:</b>	<b>1.199/2.318</b>	<b>52</b>	<b>100</b>	<b>-</b>

O favorecimento da marcação, em maior ou menor escala, ocorre em classes gramaticais que podem desempenhar funções de determinantes, isto é, que podem ocorrer em posição pré-nuclear (indefinidos = pr. 0,957; demonstrativos = pr. 0,952; possessivos = pr. 0,839; quantificadores = pr. 0,792; artigos definidos = pr. 0,688; e artigos indefinidos = pr. 0,573). Ora, como a maioria dos nossos SN's diz respeito a estruturas reduzidas, de apenas dois ou três elementos, confirma-se, à primeira vista, que ocorre uma relação entre 1<sup>a</sup>. posição e determinantes e, conseqüentemente, uma maior

marcação nos elementos gramaticais que se encontram neste lugar. Quanto aos itens que ocupam a posição de núcleo, surgem já como desfavorecedores da marcação PL (substantivos = pr. 0,326; categorias substantivadas = pr. 0,228), desfavorecimento esse que se acentua com os adjectivos, os mais inibidores da inserção de marcas (pr. 0,136) e que se situam, usualmente, em posições pós-nucleares (Tabela 16.2). Deste modo, o que se constata nos nossos dados é que os itens gramaticais apresentam uma escala binária de marcação: elementos que, por norma, ocupam posições pré-nucleares, favorecendo a inserção de marcas; e elementos que se situam nas posições de núcleo ou pós-nucleares, inibindo a pluralização.

Na Tabela 18.2 procedemos à comparação dos padrões de marcação do PA, da NURC (Lopes, 2001:168-169) e do PCV (Jon-And, 2009:4), determinados pela influência dos seus itens gramaticais:

**Tabela 18.2.** *Categoria gramatical*: efeito da *classe gramatical* na marcação PL dos itens do SN.  
– 3 variedades de português –

Categorias gramaticais	PA		NURC		PCV	
	Pr.	%	Pr.	%	Pr.	%
Substantivo grau normal	0,326	34	0,34	67	0,45	72
Categoria substantivada	0,228	32	0,26	56	-	-
Adjectivo 1	0,136	29	0,33	74		
Adjectivo 2	(0/8)	0	0,81	97	0,50	76
Adjectivo grau diminutivo	(0/8)	0	-	-	-	-
Artigo definido	0,688	90			0,59	98
Artigo indefinido	0,573	79	0,76	99		
Indefinido	0,957	97	0,60	99	0,48	93
Possessivo	0,839	84	0,74	97	0,60	98
Demonstrativo	0,952	94	(626/626)	100	0,81	99
Quantificador	0,792	80	0,26	71	0,37	84
Numeral	(3/6)	50	0,52	87	-	-
Relativo	-	-	(2/2)	100	-	-
Pronome pessoal, caso recto	-	-	0,80	91	-	-



As três variedades de português apresentam, basicamente, o mesmo padrão de marcação motivado pelos itens gramaticais. Assim, os elementos que podem configurar casos de determinantes favorecem a inserção de marcas, enquanto os restantes casos mostram tendência para inibirem a pluralização. No caso dos numerais, apesar de Lopes (2001) apresentar um valor neutro da marcação (pr. 0,52), os dados de Salvador apenas registam quinze ocorrências. Como as nossas ocorrências também são mínimas, não incluímos estes itens gramaticais na rodada VARBRUL para apurar os pesos relativos. Ainda assim, note-se que o seu percentual de marcação é de 50%, isto é, situado igualmente no patamar da neutralidade.

O único factor que apresenta resultados completamente divergentes é o dos quantificadores, bastante favorecedores da marcação no PA (pr. 0,792) e extremamente inibidores desta na NURC (pr. 0,26) e no PVC (pr. 0,37). Acreditamos que a disparidade de resultados tem a ver, ao que tudo indica, com a forma como os dados foram codificados, já que Lopes (2001) considerou apenas o item gramatical “todo/tudo” neste factor, enquanto Jon-And (2009) constituiu um factor independente para o quantificador “muitos”. No nosso caso, lembre-se que o factor *quantificador* abarca todo o tipo de quantificadores.

O factor referente aos adjectivos apresenta um peso neutro de marcação no PCV (pr. 0,50), certamente devido ao facto de Jon-And ter incluído no mesmo factor os adjectivos dos tipos 1 e 2.

Ao cruzar as variáveis independentes *categoria gramatical* e *posição linear*, P. Andrade (2003:96-97) constatou que o quantificador “todo/tudo” aparece em todas as posições lineares, mas que apenas recebe marcas quando ocorre na 1ª. localização. A forma “tudo”, com o significado de “todo”, ancora nos crioulos, sendo utilizado sem flexão, excepto nas realizações em que ocorre com numerais. Este género de construções encontra paralelo no PA, no qual este quantificador surge flexionado em número (exemplo [253]) e género (exemplo [254]):

[254] PA:  
*põe tomato, essas coesa tuda.* [MINISM3]  
“põe tomate, essas coisas todas.”

A mesma forma é também detectada no dialecto de HEL-Ba, no qual as situações flexionadas são minoritárias e apenas em número, por ocorrerem somente com numerais: “*Em nossos dados, na grande maioria das vezes em que o quantificador é flexionado, a construção é exatamente desse tipo identificado em línguas crioulas, a exemplo de Todos dois bebo; Todos quatro; Todas três. É importante salientar que, mesmo em sintagmas que exigem flexão de gênero, a forma preferencial é “tudo”, como nos exemplos: As ôtra tudo; As coisa tudo; Ôtras moça tudo solteira*” (P. Andrade, 2003:97).

No entanto, os nossos dados registam apenas uma ocorrência com numeral, e mesmo essa não recebe marca de PL (exemplo [255]), pelo que não surpreende o baixo índice de marcação registado por Lopes (2001:169), já que as realizações da NURC não devem divergir nem do PA nem de HEL-Ba nos aspectos mencionados.

[255] PA:  
*às veze parte tudo dez (ovos) [CLOTIM2]*  
“às vezes partem-se todos os dez (ovos)”

O mesmo se poderá dizer em relação à divergência de pesos registada no favorecimento da marcação apresentada pelos indefinidos (PA = pr.0,957; NURC = pr. 0,60), já que também não foram idênticos os critérios de codificação adoptados por nós e por Lopes (2001). De facto, itens como “muito/muitos” foram por nós considerados como quantificadores, em ocorrências como a do exemplo [256], enquanto Lopes (2001:85) optou por tratá-los como indefinidos:

[256] PA:  
*Home matô muitas cobras [CARMOM1]*  
“O homem matou muitas cobras”

Um maior entendimento sobre a influência da variável independente *classe gramatical* na marcação PL do SN só é possível observando a sua relação com outras variáveis, nomeadamente a *posição linear*, a *posição em relação ao núcleo do SN* e a *posição na cadeia dos constituintes flexionáveis*. Como tal, remetemos para o ponto 4.1.2.6. do presente trabalho a observação destes aspectos, uma vez que só então se fará um estudo detalhado sobre o grupo de factores *posição linear* e sua relação com as variáveis independentes mencionadas. Ainda assim, é possível concluir que se

confirmam todos os pressupostos levantados na hipótese 5 (ponto 3.6.4.1), a propósito do comportamento da variável *classe gramatical* na marcação PL do itens do SN: os itens que configuram a categoria de determinantes tendem a receber mais marcas de PL, uma vez que ocorrem, por norma, antes do nome núcleo do SN. Hierarquicamente, propiciam a inserção de marcas de acordo com a seguinte escala decrescente de favorecimento: indefinidos, demonstrativos, possessivos, quantificadores, artigos definidos e artigos indefinidos. Quanto às outras classes gramaticais, inibem a marcação, visto configurarem categorias que ocorrem mais em posição nuclear (classes substantivadas e substantivos) ou pós-nuclear (adjectivos dos tipos 1 e 2). Por conseguinte, também se confirma, uma vez mais, a equivalência na marcação registada a nível dos determinantes e primeira posição, dos substantivos e segunda posição e dos adjectivos e terceira posição (Guy, 1981b:179).

#### **4.1.2.4. Variável independente *marcas precedentes ao item analisado***

A elevada quantidade de factores tem tendência a ferir a ortogonalidade dos grupos de factores, mas o reduzido número de muitas das ocorrências impediu que se procedesse às amalgamações entre determinados factores que, eventualmente, compartilhassem traços (Tabela 12.3). Se pegássemos, por exemplo, no factor Z (SS0\_, 3ª. posição, duas marcas formais precedentes), que regista 26 realizações, e o quiséssemos adicionar a qualquer outro factor com duas ou mais marcas formais precedentes, a fim de se constituir um factor mais robusto e com possibilidades de ser incorporado nas rodadas posteriores, estaríamos impossibilitados de o fazer, porque apenas nos restaria o factor S (SSM\_), com apenas uma ocorrência, para efectuar a amalgamação. Estes aspectos levaram a que se retirasse das rodadas seguintes a quase totalidade de cadeias sintagmáticas constituídas por três ou mais elementos, o que, em parte, se traduz na inviabilização de comparação de alguns resultados com estudos prévios sobre a CPL-var no SN. Por outro lado, tal situação deixa antever bem como o PA, ao dar preferência a SN's de cadeias estruturalmente simples, se encontra num estádio que não se afasta muito do santomense, isto é, do crioulo seu substrato.

Como se pode ver também na Tabela 12.3, o primeiro factor da lista, codificado com uma barra (/), diz respeito aos itens posicionados em primeiro lugar na cadeia

sintagmática, isto é, aos elementos que não possuem qualquer constituinte antecedente. Como o que se pretende observar é a influência que o elemento antecedente exerce na pluralização do constituinte seguinte, não seguimos o critério de Scherre (1988:176), que analisou este factor, e abstivemo-nos de o observar. Aliás, este é o critério adoptado em trabalhos sobre CPL-var no SN posteriores aos de Scherre (p.e. Lopes 2001; P. Andrade, 2003; Baxter, 2004, 2009; Figueiredo, 2008, 2009b; Jon-And, 2008), em virtude de ficar vedada a possibilidade de se analisar a motivação que determina a marcação no item posterior, porque, pura e simplesmente, este não existe.

Assim sendo, apenas puderam ser levados à análise os cinco factores seguintes (Tabela 12.3):

(i) *Cadeias sintagmáticas iniciadas por item com marcação formal de PL:*

- *Factor P:* S\_, respeitante a todos os casos em segunda posição que estejam precedidos por um elemento portador de marca formal de PL (640 ocorrências);

(ii) *Cadeias sintagmáticas iniciadas por item sem marcação formal de PL:*

- *Factor B:* 0\_, referente ao segundo elemento do SN, antecedido de um item em primeira posição não-marcado formalmente (46 ocorrências);

(iii) *Cadeias sintagmáticas iniciadas por item com marcação semântica de PL:*

- *Factor N:* N\_, relativo aos itens em segunda posição na cadeia sintagmática e que possuam, antes de si, um numeral de uma só palavra, não finalizado em –s (412 ocorrências);

- *Factor R:* Ns\_, concernente às ocorrências em segunda posição no SN e que tenham, em posição anterior, um numeral de uma única palavra, terminado em –s (278 ocorrências);

- *Factor #:* NN(N)\_, reflectindo realizações que apresentem antes de si numerais de mais de uma palavra sem –s (49 ocorrências).

Em virtude de ser reduzido o seu número de ocorrências, todos os restantes factores não foram levados às rodadas posteriores. Quanto ao *factor T* (S0\_), apesar de registar 30 ocorrências, favorece plenamente a marcação (100%), pelo que foi também excluído das rodadas seguintes.<sup>165</sup> Algumas questões pertinentes relacionadas com este factor merecerão discussão em devido tempo.

A variável independente *marcas precedentes* irá ser tratada por nós de acordo com a metodologia seguida em outros trabalhos, isto é, de modo analítico (independente), por um lado, e relacionado com outros grupos de factores, por outro lado. Contudo, o segundo método não será objecto de tratamento neste ponto do nosso trabalho, no qual nos limitaremos a analisar a variável de modo analítico. Scherre (1988:183), como se verá, cruzou o grupo de factores *marcas precedentes* com a *posição linear* e a *classe gramatical*, com vista a comprovar a premissa de que as três variáveis devem ser analisadas em conjunto, pois só assim se consegue verificar se as marcas formais ou zeros afectam de modo semelhante todas as classes gramaticais em função da sua posição no SN. Por seu lado, Guy (1981a:178) não vira necessidade de levar a cabo este tipo de cruzamentos, pois entende que *posição linear* e *marcas precedentes* espelham melhores resultados se forem tratadas individualmente. Pelo facto de concordarmos com a metodologia de Guy (1981a), não efectuaremos cruzamentos do tipo daqueles que Scherre (1988) levou a cabo. Nesta conformidade, as *marcas precedentes* apenas serão cruzadas com a variável independente *saliência* fónica (ponto 5.2.3.11 do presente trabalho), no momento em que analisarmos o grupo de factores *contexto fonológico posterior* (capítulo 5), já que iremos sustentar o pressuposto de que certas ocorrências da CPL-var do PA são motivadas não pela relação fonológica do item em análise com o seu elemento seguinte, mas sim pelo peso estrutural do elemento que o antecede.

Depois de termos limpaado a célula referente aos resultados da Tabela 12.3, efectuámos uma primeira rodada VARBRUL, a fim de observarmos os pesos relativos dos factores seleccionados (Tabela 19.1):

**Tabela 19.1.** Efeito das *marcas precedentes* na marcação PL dos itens do SN.

Input desta rodada: 0,504		Log-likelihood: -981,009		Significância: 0,011	
Factores seleccionados	Nr. total	% marcação PL	% no grupo	Peso relativo	
2 <sup>a</sup> posição, marca formal na 1 <sup>a</sup> posição: S__	222/640	35	45	0,368	
2 <sup>a</sup> posição, zero na 1 <sup>a</sup> posição: 0__	39/46	85	3	0,855	
2 <sup>a</sup> posição, numeral sem -s na 1 <sup>a</sup> posição: N__	121/412	29	29	0,607	
2 <sup>a</sup> posição, numeral em -s na 1 <sup>a</sup> : Ns__	79/278	28	20	0,556	
NN(N)	14/49	29	3	0,616	
<b>Totais:</b>	475/1.424	33	100	-	

A rodada geral que efectuámos exhibe-nos o panorama analítico do modo como os itens em posição antecedente interferem na marcação PL do elemento que lhes é posterior. Através da leitura da Tabela 19.1 é possível verificar que apenas um factor inibe a marcação PL, e de forma categórica: o que representa itens com marca formal de PL em 1ª. posição e se encontram imediatamente antes do elemento analisado (S<sub>-</sub>), isto é, adjacentes a este (pr. 0,368). Por seu lado, o item em segunda posição tem tendência a ser fortemente marcado se for antecedido de um elemento sem marca formal de PL [0<sub>-</sub> = pr. 0,855] ou de um numeral, seja ele numeral de uma só palavra, terminado em *-s*, [Ns<sub>-</sub> = pr. 0,556] seja ele numeral de apenas uma palavra, mas sem inserção de *-s* final [N<sub>-</sub> = 0,607]. Quanto aos numerais de mais de uma palavra sem *-s* final, a tendência é para favorecerem moderadamente a marcação do elemento seguinte [NN(N) = pr. 0,616]. Como se vê, caso se estabeleça um paralelo entre itens com marcação formal de PL e elementos com informação semântica de pluralização, a conclusão será de que os primeiros inibem a marcação, enquanto os segundos a favorecem.

Antes de analisarmos e compararmos os nossos resultados com os de outros trabalhos, convém ainda mencionar que, tal como o referiu Scherre (1988:169), esta variável não tem sido codificada de modo uniforme pelos diferentes pesquisadores, inviabilizando-se, assim, comparações exactas. Nina (1980:72-75), por exemplo, não apresentou números para os pesos relativos, limitando-se apenas a observar os valores percentuais do português da micro-região bragantina. Por seu lado, Braga (1977:30-31), optou pela codificação binária “presença *versus* ausência de marca formal de flexão antes do elemento analisado”, imediatamente ou não. Como tal, os marcadores semânticos (numerais), foram considerados como itens sem marcação formal de PL. A conclusão final da autora é que a ausência de flexão no item anterior favorece ligeiramente a inserção de PL no elemento posterior, e vice-versa. Nas nossas ocorrências, contudo, tanto o favorecimento como a inibição são categóricos.

Posteriormente Scherre (1978:110) codificou dados cruzando parcialmente *marcas precedentes* e *posição linear*, mas agrupou os numerais conjuntamente com os elementos portadores de marcas. A conclusão da autora é que a ausência de flexão na posição 0 (zero) favorece a marcação PL, enquanto a presença de marca formal ou semântica de PL num dos elementos precedentes desfavorece a pluralização. Os nossos

dados, ao tratarem separadamente os itens portadores de marca formal e marca semântica de PL, não confirmam este pressuposto. Por outro lado, as conclusões globais a que Scherre (1978:110) chegou, após observar as misturas de marcas que antecedem o item analisado, são de que as presenças de marcas conduzem a mais marcas, enquanto a sua ausência favorece mais zeros, ou seja, não corrobora o postulado funcionalista das Condições de Distintividade (Kiparsky, 1972:195). Ainda com base neste postulado, a autora procuraria provar, depois, não só que “*o condicionamento da variável Marcas precedentes é absolutamente regular*” (Scherre, 1988:175) mas também que “*o Espanhol e o Português têm mais semelhanças do que diferenças*” (Scherre, 1988:171).<sup>166</sup>

Entretanto, Ponte (1979:80), ao debruçar-se sobre o PPA, optou também pela metodologia de Scherre (1978), mas separou os numerais dos elementos precedentes não-flexionáveis. Ainda assim, as suas conclusões são em tudo idênticas às de Scherre (1978), isto é: elemento precedente sem marca formal ou semântica de PL favorece a inserção; item antecedente com marca formal e/ou semântica de PL, inibe-a. Nos nossos dados, os postulados de Scherre (1978, 1988) e Ponte (1978) apenas colhem parcialmente, com a ausência de marcação formal em posição imediatamente anterior (0\_) favorecendo categoricamente a pluralização (pr. 0,855) e a presença daquela (S\_) inibindo-a de forma acentuada (pr. 0,368). Por seu lado, todos os marcadores semânticos em posição imediatamente precedente, incluindo os numerais de várias palavras, mostraram-se favorecedores à inserção da marca PL, não acompanhando o padrão avançado por Scherre (1978, 1988) e Ponte (1979). Aliás, o exemplo apontado por Ponte (1979) e referido por Scherre (1988:171) para confirmar a conclusão de que o item antecedente com marca semântica de PL inibe a marcação (*duas SOBRINHA nossa*), trata-se de um SN de três itens que codificámos como Ns0\_ (item na 3ª. posição, antecedido de numeral terminado em -s na 1ª. posição e ausência de marcação formal PL na 2ª. posição). Se consultarmos a Tabela 12.3, verificamos que este factor registou apenas cinco ocorrências, revelando inibição plena (100%) da pluralização. O mesmo tipo de SN, mas iniciado com numeral sem -s (N0\_), encontra-se também presente no nosso *corpus* com duas realizações, uma com marcação no elemento seguinte e outra com inibição no item posterior. Olhando ainda para outros SN's de três elementos e

iniciados por numeral, apercebemo-nos de que ocorre ainda uma configuração NS\_ (item na 3ª. posição, antecedido de numeral sem *-s* na 1ª. posição e marca formal na 2ª. posição) sem marcação no item seguinte, bem como cinco ocorrências NM\_ (item na 3ª. posição, antecedido de numeral sem *-s* na 1ª. posição e item sem pluralização na 2ª. posição),<sup>167</sup> duas das quais com elemento posterior marcado (40% de marcação). No nosso *corpus*, o reduzido número deste tipo de estruturas configura apenas um total de sete ocorrências e a sua subdivisão apresenta menos de cinco realizações para cada factor. Como tal, não nos parece que seja o tipo de exemplo mais apropriado para sustentar a conclusão avançada por Ponte (1979). Pelo menos, em última instância, conviria apurar quantas ocorrências deste tipo estão representadas nos dados da autora, de forma a constituir-se um factor robusto e que permitisse a elaboração de conclusões sólidas.

Guy (1981a:178), por seu lado, discorda do cruzamento entre *marcas precedentes e posição linear* levado a cabo por Scherre (1978, 1988), entendendo que as duas variáveis são estatisticamente mais sólidas se forem tratadas em separado. Como tal, não aborda a questão entre ausência e presença de marca no elemento precedente, levando Scherre (1988), uma vez mais, a divergir da metodologia adoptada pelo linguista, principalmente porque Guy (1981a:180) considera os numerais como elementos de marca 0 (zero),<sup>168</sup> isto é, da mesma forma que os itens sem marca formal de PL. Por outro lado, Scherre (1988:172-73) chama também a atenção para o facto de Guy (1981a:169-170) ter tratado todas as sequências de quatro elementos da mesma forma, ou seja, independentemente de o elemento anterior ao item analisado possuir, ou não, marca de PL. Estes critérios impediam, então, que se efectuassem análises comparativas com os resultados apresentados por Poplack (1980a), vedando-se a possibilidade de se observar se um item flexionado em número conduz ou não à flexão do seu elemento ulterior.

As codificações de Scherre (1988), seguindo a metodologia de Poplack (1980a) e buscando um padrão regular para as marcas precedentes aproximam-se das nossas em alguns aspectos, permitindo comparações algo precisas no que concerne aos SN's compostos por apenas dois elementos, isto é, aqueles para os quais possuímos os valores dos pesos relativos. Quanto aos demais factores, as nossas divergências assentaram quer na forma de observar os numerais quer no modo de analisar as misturas de marcas. No



caso dos numerais, Scherre (1988:173-174), depois de concluir que não existe um padrão acentuado de variação entre os que terminam em *-s* e os que não finalizam em *-s*, acaba por tratá-los de modo homogêneo. Quanto a nós, decidimo-nos pela heterogeneidade e acabámos por incluí-los em factores distintos, a fim de percebermos quais são os que propiciam mais a marcação. Ainda sobre a questão dos numerais, adiantaremos mais considerações no ponto 4.1.2.4.6 do presente trabalho, uma vez que procuraremos, aí, efectuar observações mais pormenorizadas sobre os mesmos.

Scherre (1988:174-177) observou também o comportamento dos SPREP's dentro de um SN mais alto com marca semântica de PL, uma vez que as análises sobre este tipo de SN's postulam que os mesmos tendem a ser pouco marcados. Para tanto, a autora estabeleceu duas equações que os abarcassem e ajudassem a confirmar que os mesmos apresentam padrões regulares de marcação: (i) SPrep(x)s-, envasando realizações do tipo “*milhões de coisaS*” (SNs+Prep+S = núcleo nominal mais alto com marca formal de PL) ou “*uma série de outraS coisas*” (SN+Prep+S\_ = núcleo nominal com ou sem marca formal de PL, acompanhado de marca no primeiro elemento do SN sob análise); (ii) Sprep0(0)-, incluindo ocorrências como “*uma porção de carroØ*” (SN+Prep = núcleo nominal mais alto sem marca formal de PL) ou “*uma porção de coisaØ interessanteØ*” (SN+Prep+0\_ = núcleo nominal sem marca formal de PL, seguido de ausência de marca no primeiro elemento do SN sob análise). Estes SN's, apesar de não indicarem o PL de forma tão óbvia como os numerais, não deixam também de encerrar uma evidente carga semântica de pluralidade. Assim sendo, o pressuposto do tipo funcionalista (Kiparsky, 1972:195) defende que, pelo facto de implicarem uma duplicidade de concordância PL, estes sintagmas tendem a exhibir inibição no segundo item, evitando-se, desta forma, uma redundância de número. Scherre (1988:208), ao buscar um padrão regular de marcação para os mesmos, discorda e procura demonstrar que os mesmos são configurados segundo o Princípio do Processamento Paralelo. Estes tipos de factores representam, quanto a nós, ocorrências que têm que ser vistas como cadeias compostas por mais de um item gramatical, pelo que a sua observação deverá ser feita sob a perspectiva não-atomística, tal como efectuado por P. Andrade (2003:109). Assim, reservamos para uma oportunidade futura este tipo de análise, ou seja, quando observarmos os nossos dados sintagmaticamente.

A busca de padrões regulares de marcação por parte de Scherre (1988:180) levou-a a constituir também um factor amalgamado para os SN's com mais de dois elementos e que apresentam marcações formais plenas nos itens antecedentes ao elemento observado, representado pela equação SS(S)-. Esta, engloba as ocorrências SS\_ e SSS\_ (um percentual 70% de marcação PL para os dois factores em conjunto) e, a propósito da mesma, cabe-nos referir que o nosso *corpus* não regista qualquer ocorrência com tripla marcação formal antecedente e apresenta apenas vinte e seis realizações com dupla marcação formal antecedente, cujo percentual de pluralização é de 54% (Tabela 12.3). Como referimos no princípio da presente secção, a propósito de uma hipotética amalgamação entre este factor e outros que com ele partilhem características, registámos ainda uma realização do tipo SSM\_ — incluída por Scherre (1988:177) na equação  $y(x)x-$ , de que falaremos já de seguida —, e que não apresenta marcação no quarto elemento da cadeia sintagmática. Como se verifica, os nossos dados, apesar de não registarem, para este tipo de SN's, um mínimo de 30 realizações que permita obter valores para pesos relativos, parecem indicar um equilíbrio entre marcação e não-marcação no elemento analisado. Assim sendo, e ao que tudo indica, não se confirma aqui uma nítida tendência no sentido da inserção de marcas que permita advogar o Princípio do Processamento Paralelo (Scherre, 1988:208) para justificar tais pluralizações no PA.

Relativamente às misturas de marcas, Scherre (1988) apresentou também duas equações para tentar abarcar os padrões regulares de marcação de todas elas: (i)  $y(x)x-$ , englobando “*todos os casos de presença de pelo menos uma marca formal de plural precedendo o elemento nominal analisado, não mediada por zero em elemento que admite marca*” (Scherre, 1988:178); (ii)  $x(y)0-$ , assimilando realizações em que “*entre a última marca e o elemento analisado, tem de haver necessariamente um zero, mesmo mediado por um numeral ou por um modificador*” (Scherre, 1988:179). Quanto a nós, optámos por agrupar as configurações das marcas precedentes dos nossos dados em três subgrupos distintos (Tabela 12.3), tendo sido seleccionados para cálculos de pesos relativos (Tabela 19.1) apenas os factores que representam SN's de dois itens, com excepção da configuração NN(N)\_, que poderá dizer respeito a sintagmas de dois ou três elementos. Assim, enquanto Scherre (1988) tomou como ponto de referência a marcação

formal ou ausência desta no elemento que antecede o item analisado, nós optámos por considerar a marcação (formal ou semântica) ou ausência registada na primeira realização da cadeia sintagmática. Contudo, voltamos a referir que, ao contrário de Scherre (1988), não iniciámos as nossas observações a partir das amalgamações de factores, que apenas foram efectuadas depois de analisado o efeito das configurações individuais na pluralização do item seguinte. Contudo, a este propósito, importa salientar que Scherre (2001) viria posteriormente a analisar o comportamento individual dos factores, cujos resultados serão apresentados em devido tempo no presente estudo.

Observando agora o critério seguido por Scherre (1988:177-181) para a constituição dos factores amalgamados, convém referir alguns aspectos. Nos casos de mistura de marcas, a autora inclui na amalgamação  $x(y)0-$  (*entre a última marca e o elemento analisado, tem de haver necessariamente um zero, mesmo mediado por um numeral ou por um modificador*), por exemplo, os factores  $N0_$  (ex.: *DOIS RISCO verde*) e  $S0N_$  (ex.: *AS MINHA DUAS filha*) (Scherre, 1988:177). No primeiro caso, se entre a última marca e o elemento analisado (*verde*) tem que existir pelo menos um zero, partimos do princípio que foi tida em conta, como última marca, o numeral “dois”, isto é, uma marca semântica de PL. Contudo, no segundo exemplo, o mesmo numeral (*duas*) já não foi visto como sendo a última marca, tendo tal critério de observação recaído sobre a marca formal de pluralização do artigo definido feminino “a”. Quanto a nós, criou-se aqui um desajustamento entre os factores levados à amalgamação para constituir a equação apresentada por Scherre (1988:179). Outro aspecto importante a ter em conta é o facto de a metodologia em estudos de sociolinguística quantitativa recomendar a eliminação de todas as realizações com menos de cinco ocorrências, não devendo as mesmas ser consideradas para efeitos de amalgamações entre factores. Ora, este princípio foi violado pela autora, que levou à amalgamação dados que configuram realizações diminutas: “*Os casos de dois zeros precedentes foram todos agrupados aos de apenas um zero por serem de menor número de ocorrências. Há, por exemplo, apenas dois casos do tipo S00-*.” (Scherre, 1988:179).

Lopes (2001:204-211), por seu lado, após uma primeira rodada com quinze factores englobados individualmente no grupo de factores marcas precedentes, decidiu também proceder a amalgamações, adoptando o critério de Scherre (1988) e reduzindo a variável

independente *marcas precedentes* a sete factores. Para tanto, a autora deixou de considerar o contraste entre numerais (simples ou de mais de uma palavra) finalizados em *-s* e não finalizados em *-s* e constituiu três factores para os SN's com três ou mais itens: (i) mais de uma marca antecedente (SS\_); (ii) mistura de marcas com marca antecedente (OS\_; SM\_; NS\_; SN\_; SSM\_), configurada pela equação  $y(x)x-$  de Scherre (1988); (iii) mistura de marcas com zero antecedente (S0\_; 0M\_; N0\_; N0M\_), representada pela equação  $x(y)0-$  de Scherre (1988). Na primeira rodada, Lopes (2001:204-205) regista ainda um outro factor (S0M\_) com vinte ocorrências e para o qual não foi adiantada qualquer explicação nem em termos de amalgamação nem em termos de uma possível eliminação das rodadas posteriores. Este factor, segundo a metodologia de Scherre (1988:177), deveria ser levado à amalgamação que configurou o factor mistura de marcas com zero antecedente [ $x(y)0-$ ], mas, simplesmente não foi encaixado por Lopes (2001) em qualquer amalgamação. Por outro lado, conseguimos descortinar também que as amalgamações de Lopes (2001) enfermam do mesmo problema das de Scherre (1988), uma vez que, para estas, foram tidos em conta factores com menos de cinco realizações, como acontece com a configuração N0\_, detentora de apenas uma ocorrência (Lopes, 2001:210-211) e incluída na amalgamação mistura de marcas com zero antecedente. Outro aspecto a ter em consideração no trabalho de Lopes (2001) é o facto de, após apresentação dos resultados para os factores, vistos numa perspectiva individual, não ter sido efectuado o teste do qui-quadrado às amalgamações entre factores, com vista a confirmar a possibilidade de as mesmas poderem ou não ser realizadas. Assim a autora considerou como válidas para observação do efeito das *marcas precedentes* na marcação PL da NURC amalgamações que apenas os testes do qui-quadrado poderiam confirmar serem ou não estatisticamente significativas (cf. Guy & Zilles, 2007:194).

De seguida, e numa primeira etapa do nosso trabalho, iremos comparar os nossos resultados sobre a influência das marcas precedentes na CPL-var com os de Scherre (1988), levando-se em consideração, para os factores individuais, os valores indicados pela autora em outro trabalho (Scherre, 2001). Paralelamente, serão ainda observados quer os resultados obtidos por Lopes (2001:204-205) para os factores prévios às suas amalgamações quer os exibidos por Baxter para o PT e o dialecto de HEL-Ba (Baxter,

2009:281). Na Tabela 19.2, que apresenta estes resultados, achamos pertinente incluir também os valores de Poplack (1980a:63) para o EPR, que, segundo Scherre (1988), apresenta o mesmo padrão de variação do MRJ. Os resultados transcritos referem-se aos factores mais significativos e seus pesos relativos e/ou percentuais de marcação, visto os diferentes trabalhos não revelarem um paralelo homogêneo em termos de constituição de factores para o grupo de factores *marcas precedentes*.

Posteriormente à comparação dos resultados em termos de configurações (factores) individuais, e apesar de termos em conta que os valores exibidos por Scherre (1988) e Lopes (2001) poderão estar cientificamente feridos devido ao critério adoptado para a constituição dos factores amalgamados, efectuaremos igualmente, e conquanto os nossos dados o permitam, amalgamações de factores de acordo com as metodologias levadas a cabo pelos diferentes estudiosos. Poderemos, assim, proceder também a comparações com os números apresentados por Scherre (1988:180), Lopes (2001:211), P. Andrade (2003:99) e Baxter (2009:281). A propósito dos factores constituídos por P. Andrade (2003:99), é importante referir que a autora apenas apresentou números para três factores amalgamados, de acordo com o seguinte critério: (i) presença de marca formal antes do elemento analisado; (ii) presença de marca semântica antes do item analisado; (iii) ausência de marca no elemento anterior. A autora limitou-se a analisar, pois, a marca (formal ou semântica) ou sua ausência no elemento que antecede o item observado, não votando atenção às misturas de marcas que ocorrem na fala de HEL-Ba. Como tal, a observação do padrão individual de factores não-amalgamados foi efectuada por Baxter (2009:281) no mesmo dialecto. Este apresentou também amalgamações para as misturas de marcas em cadeias de três itens do PT, tendo, para o efeito, constituído os seguintes factores: (i) mistura de marcas com marca formal antecedente, item analisado na 3ª. posição; (ii) mistura de marcas com zero antecedente, item analisado na 3ª. posição.

Antes de efectuarmos as comparações, convém também referir que os resultados das amalgamações de Scherre (1988:180) foram, em primeira etapa, unicamente apresentados em termos percentuais, tendo a pesquisadora fornecido apenas números para os pesos relativos de rodadas VARBRUL posteriores, isto é, as respeitantes aos cruzamentos efectuados entre *marcas precedentes*, *posição linear* e *classe gramatical*,

metodologia com a qual Guy (1981a), e nós próprios, discordamos. Por outro lado, a autora apresentou números para dois tipos de informantes: crianças (7 aos 14 anos) e adultos (15 aos 71 anos). Para os nossos resultados, teremos em conta, sobretudo, o público adulto do MRJ, já que o nosso *corpus* não contempla falantes com idade inferior a 21 anos. Ainda assim, lançaremos mão dos resultados apresentados por Scherre (1988:181) para as crianças, quando estabelecermos comparações com os dados da autora, em termos geracionais.

Relativamente à questão das faixas etárias de Lopes (2001:69-71), elas estão distribuídas por informantes que prestaram informações em dois períodos distintos no tempo (década de 70 e década de 90) e se encontram distribuídos por quatro gerações distintas: F.E.I (15 aos 24 anos); F.E.II (25 aos 35 anos); F.E. III (45 aos 55 anos); e F.E.IV (+ 55 anos). A F.E.I contempla apenas informantes da década de 90.

Por seu lado, o público informante que constituiu os *corpora* de HEL-Ba, analisado por P. Andrade (2003:125), o PMp e o PCV, ambos observados por Jon-And (2008:2, 2009:2), encontra-se distribuído por três faixas etárias, absolutamente idênticas às nossas: faixa I (21 a 40 anos); faixa II (41 a 60 anos); faixa III (+60 anos). No entanto, a segunda autora não efectuou a análise da variável independente *marcas precedentes* nos seus trabalhos, embora tivesse constituído este grupo de factores para o estudo do PMp, que foi excluído das observações “*por causa de interferência com outras variáveis*” (Jon-And, 2008:3).

Quanto a Baxter (2009:283), dividiu os informantes de HEL-Ba em quatro faixas etárias distintas, critério que adoptou também para os informantes do PT: faixa 1 (21 a 40 anos); faixa 2 (41 a 60 anos); faixa 3 (61 a 80 anos); faixa 4 (+85 anos).

#### **4.1.2.4.1. *Marcas precedentes*: influência dos factores individuais na marcação PL do elemento analisado**

Na Tabela 19.2 podem ver-se os valores em pesos relativos e percentuais de marcação relativamente ao EPR e às diferentes variedades africanas e brasileiras de português indicadas no ponto anterior:

**Tabela 19.2.** *Marcas precedentes*: contribuição dos factores individuais para a marcação PL do elemento analisado: EPR e 5 variedades de português.

<i>Marcas precedentes</i>												
Item analisado na 2. <sup>a</sup> posição												
Factores	EPR		PA		MRJ		NURC		HEL-Ba		PT	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%
S_	0,56	-	0,368	35	0,50	53	0,48	56	0,499	52	0,608	52
O_	0,48	-	0,855	85	(105/105)	100	(34/34)	100	0,825	65	0,825	65
N_	-	-	0,607	29	-	58	0,57	56	-	-	-	-
Ns_	-	-	0,556	28	-	59	0,53	65	-	-	-	-
Item analisado na 3. <sup>a</sup> posição												
	EPR		PA		MRJ		NURC		HEL-Ba		PT	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%
SO_	0,32	-	(0/30)	100	0,07	6	0,06	7	-	-	-	-
SM_	-	-	(2/5)	40	-	-	0,64	61	-	-	-	-
SN_	-	-	(7/22)	32	0,52	42	0,53	54	-	-	-	-
SS_	0,60	-	(14/26)	54	0,65	70	0,62	56	-	-	-	-
OS_	-	-	(4/7)	57	0,40	53	0,49	35	-	-	-	-
OO_	0,27	-	(3/3)	100	-	-	-	-	-	-	-	-
NS_	-	-	(0/1)	0	-	-	-	-	-	-	-	-
NsS_	-	-	(0)	-	0,35	39	0,57	56	-	-	-	-
N0_	-	-	(1/2)	50	0,07	6	-	-	-	-	-	-
Ns0_	-	-	(0/5)	0	-	-	(0/1)	0	-	-	-	-
Item analisado na 4. <sup>a</sup> ou outra posição												
	EPR		PA		MRJ		NURC		HEL-Ba		PT	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%
SSM_	-	-	(0/1)	0	-	-	0,64	68	-	-	-	-
SOM_	-	-	(1/4)	25	-	-	(2/20)	10	-	-	-	-
NN(N)_	-	-	0,616	29	-	-	0,68	61	-	-	-	-
NNs(N)	-	-	(2/9)	22	-	-	0,65	70	-	-	-	-

**Obs.:** (i) Nas colunas referentes aos pesos relativos, são indicados os totais de ocorrências, caso estas não perfaçam 30 realizações ou prefigurem 0% ou 100% de marcações (*knockouts*);

(ii) Os valores para a configuração O\_ , N\_ e Ns\_ do MRJ foram colectados em Scherre (1998a:180).

Os factores que abarcam SN's de dois itens (codificados de modo bastante idêntico em todas as variedades de português) denotam algumas discrepâncias a nível de resultados da marcação formal ou ausência da mesma. Com efeito, o factor S\_ (2.<sup>a</sup> posição, marca formal na 1.<sup>a</sup> posição) regista uma acentuada tendência para inibir a

marcação no PA (pr. 0,368), acentuação esta que se dilui em direcção à neutralidade na NURC (pr. 0,48), até se tornar mesmo completamente neutra em HEL-Ba (0,499) e no MRJ (pr. 0,50). Não obstante, no PT, ao contrário do que sucede nas outras variedades de português, o mesmo factor revela moderada marcação PL do item analisado (pr. 0,608). Este aspecto poderia, aparentemente, aproximar os comportamentos do PT e do EPR, vendo-se aqui a confirmação que Scherre (1988:208) buscou, no sentido de ocorrer uma intervenção inquestionável do Princípio do Processamento Paralelo, com marcas a conduzirem a mais marcas. Contudo, Baxter (2009:282) não vê aqui a operação de um fenómeno meramente linguístico, justificando a noção de concordância apresentada pelos falantes do PT com o facto de os substratos deste manterem um contacto constante com sistemas de concordância morfológicamente elaborados da LA, sobretudo em virtude de alguns trabalhos especializados dos tongas lhes concederem um situação privilegiada e que lhes terá facilitado o contacto com falantes do PE.

Por seu lado, o factor 0\_ (2<sup>a</sup> posição, ausência de marcação formal na 1<sup>a</sup> posição) apresenta um favorecimento categórico e equilibrado da marcação PL no PT (pr. 0,825), em HEL-Ba (0,844) e no PA (pr. 0,855), a qual se torna plena (100%) na NURC e no MRJ. As constatações acerca dos dois factores referidos (S\_ e 0\_) levantam, desde já, uma questão que poderá ser pertinente em termos de comparações entre o PA e os outros dialectos, se levarmos em conta que a grande maioria dos SN's do nosso *corpus* são de estrutura reduzida: o PA aparenta ser a variedade que se encontra no estádio mais baixo da aquisição da regra de concordância PL.

Quanto aos marcadores semânticos (numerais), revelam também valores bastante equilibrados nas três variedades de português para as quais possuímos resultados. Assim, os numerais terminados em *-s* mostram um favorecimento da pluralização pouco acima do patamar da neutralidade (NURC = pr. 0,53; PA = pr. 0,556), tendência esta que ganha alguma força com os numerais que não terminam em *-s* (NURC = pr. 0,57; PA = pr. 0,607). Scherre (2001) apresentou percentuais para o MRJ que, ao que tudo indica, acompanham os padrões das outras variedades (N\_ = 58%; Ns\_ = 59%). Assim, no que aos numerais diz respeito, observa-se um paralelo entre estes e os itens portadores ou não de marcação formal, uma vez que os elementos terminados em *-s* têm também tendência para propiciarem menos a pluralização do que aqueles que não finalizam em –



s. Contudo, e tal como o observaram Scherre (2001) e Lopes (2001), as diferenças de valores reveladas quer pelos números percentuais de marcação entre ambos os tipos de marcadores semânticos quer pela sua tradução em pesos relativos são praticamente nulas, o que justificaria a eliminação do contraste entre ambos. Aliás, esta foi a linha seguida por P. Andrade (2003), para HEL-Ba e Baxter para HEL-Ba e PT (2009), mas sobre esta metodologia nos pronunciaremos no ponto seguinte deste trabalho, quando abordarmos a questão dos factores amalgamados.

A tendência para o maior ou menor propiciamento da marcação mencionada no parágrafo anterior revela-se drástica quando passamos a observar os itens que estão sujeitos a marcação formal, a ponto de a marcação ou ausência desta conduzirem a valores inibidores ou favorecedores no item seguinte. Tanto a ausência quase sistemática de marcação no elemento seguinte ao que contenha marcação na 1ª. posição (S\_) como a inserção categórica da pluralização no item posterior ao que não possua marcação na 1ª. posição (0\_) sugerem a actuação da hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que advoga a necessidade de se inserir a marcação quando ela é fundamental para indicar a ideia de pluralidade, podendo ocorrer depois o seu apagamento, em virtude de se tornar redundante. Daí que, apesar de se manifestar em favor do Princípio do Processamento Paralelo, Scherre (1988:230) acabe admitindo em configurações de dois itens, “*os ‘esses’ que conduzem a ‘esses’ são apenas os que têm a informação semântica de pluralidade*”, ou seja, os numerais. Entretanto, importa reter que o padrão de marcação formal do PA é contrário ao do EPR, uma vez que Poplack (1980a), face aos seus achados, avançou que este dialecto tem tendência para agrupar formas semelhantes, mediante processos mentais de associação, confirmando-se que o que subjaz à sua configuração sintagmática é o Princípio do Processamento Paralelo (cf. Scherre, 1988:208), responsável pelo facto de marcas levarem a marcas (S\_ = pr. 0,56) e zeros levarem a zeros (0\_ = pr. 0,48). Este aspecto permite mesmo que, segundo Poplack (1980a:64), no EPR ocorram sintagmas sem qualquer marcação formal, mas com referência PL (p.e.: *la cosa bonita* – “as coisas bonitas”). No entanto, Scherre (2001:98-99) adverte para a possibilidade de este exemplo constituir um caso de representação abstracta de pluralização, isto é, uma realização com carga semântica de pluralidade que é captada pelos falantes a partir do contexto e conduz à não-realização

da marcação formal PL. É o que sucede, por exemplo, com as ocorrências do tipo de “*a coisa agora tá feia*” (“as coisas agora estão feias” = “a situação agora está má”), possível no PB, segundo a autora, e também no PE, acrescentamos nós.<sup>169</sup> Como tal, a não constatação de tripla ausência de marcação formal ØØØ, originada pela configuração 00\_ (item analisado na 3ª. posição, com dupla ausência de marcação formal anterior) e verificada nos dados da NURC e do MRJ, vem provar que a sua existência no EPR não é senão aparente, tal como acontece no português (Scherre, 2001:99). Nos nossos dados, como se pode verificar, a estrutura ØØØ também não ocorre, já que os 3 casos registados de configuração 00\_ apresentam uma taxa plena de inserção PL no terceiro item da cadeia sintagmática. Origina-se assim a estrutura ØØS, que foi mesmo considerada inexistente por Scherre (2001:99), visto a autora apenas ter detectado tripla ausência de marcação em dois tipos de construções: (i) SPrep+0\_ (núcleo nominal mais alto sem marca formal de PL seguido de item sem marcação formal de PL), que contém três ausências formais de marcação, mas encerra informação semântica de PL (p.e. *UM MONTÃO DE NEGO velho*); (ii) estruturas cristalizadas que englobam uma relação de posse inalienável, representando mais do que um elemento (p.e. *vem lavá A MÃO, O PÉ e escová O DENTE*).

Outras das configurações em cadeias de três elementos que poderiam pressupor uma acção do referido Princípio do Processamento Paralelo (Scherre, 1988:208), com zeros a conduzirem a zeros (S0\_ e SM\_), mereceram também a nossa atenção. A segunda apresenta apenas cinco ocorrências no nosso *corpus*, manifestamente insuficientes para permitirem retirar-se quaisquer ilações com sustento científico. Contudo, os dados da NURC indiciam uma ligeira tendência para favorecer a inserção PL no terceiro elemento (pr. 0,64), o que pode também ser argumentado numa perspectiva funcionalista, caso se ponha a possibilidade de os falantes entenderem que o modificador da segunda posição (M) represente um elemento de categoria Ø, isto é, sem a marcação formal de PL. Quanto à configuração S0\_, apresenta clara tendência para inibir a pluralização no terceiro elemento (PA = 100%; NURC = 0,06; MRJ = 0,07), tornando categórica a realização da configuração SØØ e praticamente inexistente a cadeia sintagmática SØS. Esta, segundo Scherre (1988:187-188), é motivada pela influência do contexto fonológico sobre o -s final não morfémico, não reflectindo, portanto, qualquer efeito

sobre a concordância de número entre os elementos do SN. Este entendimento procura também ir ao encontro da hipótese avançada por Guy (1981a:176) de que a concordância opera copiando a informação PL da esquerda para a direita, através do nóculo do SN. Assim, as poucas realizações do tipo SØS resultarão do facto de as consoantes sonoras favorecerem mais do que as surdas o cancelamento do –s não morfémico do segundo item da cadeia, o mesmo sucedendo em relação às consoantes labiais e velares.

Relativamente à segunda configuração possível (SØØ), Scherre (1988:189-193) elabora análises no sentido de apresentar uma hierarquia de preferência em termos de estruturas sintagmáticas e aponta-a como sendo a segunda mais seleccionada pelos falantes do MRJ (25%). A ocorrência do fenómeno é justificada pela autora com a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208), sendo corroborada por Lopes (2001:212) neste aspecto. Contudo, não é este o nosso entendimento. De facto, fazendo o mapeamento das 30 ocorrências do nosso *corpus* que representam a configuração SØØ, constata-se que o terceiro item da cadeia sintagmática se encontra em posição pós-nuclear em 28 das realizações, sendo esta ocupada por adjectivos (21 casos) ou pelo quantificador “todo” (7 casos). Mesmo em exemplos apontados por Scherre [“*umas borracha grande*” (Scherre, 1988:177) e “*naqueles negócio todo*” (Scherre, 1988:185)] e Lopes [“*as sacola pronta*” (Lopes, 2001: 204)] para tal configuração, este aspecto se torna evidente. A situação remete-nos então para uma outra questão, também abordada por Scherre (1988:221-223), e que tem a ver com o facto de as classes gramaticais não-nucleares, independentemente da sua categoria, terem mais probabilidades de apresentarem marcação se estiverem antepostas ao núcleo do SN, do que se ocorrerem depois deste. Aliás, ao efectuar os cruzamentos entre *classes gramaticais* e *posição linear*, a autora havia também constatado que, em SN’s de três ou mais elementos, os adjectivos apresentam, na terceira posição, um baixo percentual de marcação (41%) (Scherre, 1988:157), enquanto os quantificadores são categoricamente marcados na primeira posição (100%) e apresentam um índice de marcas de apenas 20% na terceira posição (Scherre, 1988:160-161). Por outro lado, quer o Princípio Funcionalista de se marcar o PL na primeira oportunidade (Kiparsky, 1972:195) quer a conclusão de Guy (1981a:301-302) sobre a tendência do português popular do Brasil indicar o PL conforme os substratos africanos, em que predomina o uso de marcadores

iniciais em forma de prefixo ou clítico, pode aqui encontrar perfeita sustentação e vai ainda justificar as ocorrências SØØ, nas quais o terceiro elemento é representado por um nome, núcleo do SN. Daí que P. Andrade (2003:98-100) se tenha absterido de olhar para a totalidade da configuração das marcas precedentes e tenha centrado a sua atenção unicamente no item que antecede o elemento analisado.

O factor 0S\_ também não permite tirar conclusões nos nossos dados, uma vez que registámos apenas sete ocorrências, com um percentual de marcação de 57%. Lopes (2001:204) apresentou, para a mesma configuração, um valor que, tal como o nosso percentual, tende para a neutralização da marcação (pr. 0,49). Contudo, a linguista apenas possui vinte realizações nos seus dados e, ao considerar o factor para efeitos de cálculo de pesos relativos, não foi ao encontro daquilo que é recomendado pelos estudos em sociolinguística quantitativa, já que estes entendem que trinta é o número mínimo de ocorrências tido como confiável para a determinação do peso que um factor pode ter no grupo de factores em que está inserido (Guy & Zilles, 2007:153). Assim sendo, o único valor que poderemos ter em consideração para a marcação PL determinada pela configuração 0S\_ é a ligeira inibição apresentada para o MRJ (pr. 0,40). Embora se constate um equilíbrio evidente entre marcação e inibição nos três dialectos em questão, a configuração ØSØ parece prevalecer ligeiramente sobre a estrutura ØSS. Dado que estas configurações não deveriam ocorrer, à luz do Princípio do Processamento Paralelo (Scherre, 1988:208), a autora explica-as também em função da influência exercida pelo contexto fonológico seguinte. Como tal, a conclusão final da linguista ancora igualmente no pressuposto de Guy (1981a:169), advogando que a concordância de número entre os elementos do SN se aplica ao seu nóculo *“da esquerda para a direita sem saltar nenhuma palavra, a não ser as que não podem ser marcadas”* (Scherre, 1988:189), devido à força que o contexto fonológico seguinte exerce sobre estas em termos de oposição surdas *versus* sonoras e/ou labiais *versus* velares. Este pressuposto, na opinião da autora, vai ainda justificar o porquê do favorecimento categórico da pluralização do segundo item da cadeia sintagmática, quando a primeira posição apresenta um elemento sem PL formal (0\_), originando a configuração ØS.

Acerca do factor SS(S)\_, que revela propensão para favorecer moderadamente a marcação no terceiro elemento do SN dos dados da NURC (pr. 0,62) do MRJ (pr. 0,65),

criando a cadeia SSS, apenas regista 26 ocorrências no nosso *corpus*, com um percentual de marcação pouco acima do patamar da neutralidade (54%). O facto de a marcação no terceiro elemento se sobrepor à inibição, levou Scherre (1988:189) a advogar que os falantes revelam preferência pela construção da estrutura SSS, em detrimento da elaboração da configuração SSØ. Mais uma vez, a autora vê aqui a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208), que leva os falantes a optarem pelas estruturas em que marcas levam a mais marcas, por oposição às configurações que apresentam misturas de marcas, quer quanto à quantidade quer quanto ao tipo. Em adição, a configuração SS(S)\_ será também mais favorecedora da marcação em contexto seguinte do que a estrutura y(x)x- (existência de pelos menos uma marca formal de PL antes do elemento analisado).

O facto de os nossos dados apresentarem um percentual de marcação inferior aos da NURC e do MRJ, roçando a neutralidade, levou-nos a questionar tal comportamento. Como referimos antes, e ao que tudo indica, o PA aparenta situar-se num estágio inferior de aquisição da regra de concordância PL, quando comparado com a NURC e o MRJ. Tal não surpreende, até porque nenhum dos informantes do PA possui mais de oito anos de escolaridade, ao contrário do que acontece nestes dois últimos dialectos. Este aspecto, como é evidente, pode jogar aqui um peso determinante, visto, nas duas variantes do PB, estarmos perante SN's estruturalmente mais desenvolvidos, isto é, produções de falantes que já possuem alguma consciência acerca da aplicação da regra. Como tal, decidimos fazer um levantamento dos vinte e seis dados SS\_ do nosso *corpus*, a fim de confirmarmos o grau de escolaridade dos falantes que os produziram. As nossas suspeitas confirmaram-se, já que a grande maioria deste tipo de SN's foi realizada por informantes com um grau de escolarização médio/alto, conforme estratificação por nós adoptada para a variável independente *escolaridade*.

Na Tabela 19.3 pode ver-se a distribuição das estruturas SS\_ e seu percentual de marcação pelos falantes que as produziram. Assim, os informantes com escolaridade média/alta produziram 22 dos 26 SN's do tipo SS\_, enquanto as restantes 4 realizações se repartem em partes iguais (2+2) pelos informantes de baixa escolaridade ou analfabetos. A conclusão a que se chega é que a esmagadora maioria destes SN's é da responsabilidade de falantes que têm contacto, via ensino, quer com estruturas

sintagmáticas mais elaboradas quer com a regra da pluralização. Assim, estarão em processo mais avançado de aquisição da referida regra, visto terem consciência acerca da mesma. Portanto, não surpreende que o percentual de marcação plena (SSS) por parte dos falantes com média/alta escolarização atinja um número ligeiramente acima do patamar da neutralização (55%). Justificar então esta tendência para a marcação com recurso ao Princípio do Processamento Paralelo (Scherre, 1988:208) reflecte, quanto a nós, resultados virtuais, se observados em termos de realizações da comunidade. Os próprios resultados de Lopes (2001:213), que apresentam o efeito da marcação *marcas precedentes* nos grupos português popular (POP) e português universitário (UNI), reforçam a nossa perspectiva, visto o primeiro grupo apresentar, para elementos em 3ª. ou outra posição de cadeias com dupla marcação formal nos dois primeiros itens, um percentual de marcação (56%) que se aproxima do exibido pelo PA, enquanto o segundo grupo mostra que os informantes com escolarização superior alcançam percentuais muito próximos da pluralização plena (92%).

**Tabela 19.3.** *Marcas precedentes*: configuração SS\_ (item na terceira posição, duas marcas formais precedentes): informantes e grau de escolarização.

Informantes	Idade	Escolaridade	SS_	%	Total
[CASTEH1]	FE-1 (20-40)	8º. ano	2/5	40	4/9
[ANTOM1]	FE-1 (20-40)	8º. ano	2/4	50	
[OSVALH1]	FE-1 (20-40)	4ª. classe	0/2	0	8/13
[CARMOM1]	FE-1 (20-40)	4ª. classe	3/4	75	
[ABILH2]	FE-2 (41-60)	4ª. classe	5/5	100	
[CLOTIM2]	FE-2 (41-60)	4ª. classe	0/2	0	1/2
[MAURIH2]	FE-2 (41-60)	3ª. classe	1/1	100	
[CELESH3]	FE-3 (+60)	3ª. classe	0/1	0	
[MANEVEM2]	FE-2 (41-60)	(Analfabeta)	0/1	0	1/2
[MANOH3]	FE-3 (+60)	(Analfabeto)	1/1	100	
<b>Totais:</b>			14/26	54	14/26

Quanto aos SN's de três elementos com misturas de marcas que envolvem numerais, todos eles apresentaram menos de trinta realizações no PA. As configurações NS\_, NsS\_, N0\_ e Ns0\_ revelam, assim, um ínfimo número de ocorrências, pelo que não

mereceram atenção da nossa parte. Lopes (2001:205) apenas apresentou resultados para o factor NS\_ (NS+N<sub>s</sub>S\_), não se confirmando que a mistura precedente de marcas semântica e formal favoreça a inserção PL no elemento seguinte (pr. 0,47), de forma a configurar a estrutura SNS, que pudesse ser justificada com recurso ao Princípio do Processamento Paralelo (Scherre, 1988:208). Quanto a Scherre (1988), apresentou valores para itens antecidos pelas configurações NS\_ (pr. 0,35) e N0\_ (pr. 0,07), mas as mesmas configuram amalgamações de todos os numerais (terminados ou não em –s, fossem eles simples ou de mais de uma palavra), pelo que não podem ser tidos em linha de conta para efeitos de comparação nem com os nossos dados nem com os de Lopes (2001). Além do mais, a estrutura NS\_ apresenta um total de apenas 18 realizações, pelo que não deveria ter sido incluída na rodada VARBRUL para efeitos de pesos relativos.

Observando também os SN's constituídos por três ou mais elementos que envolvem numerais, Scherre (1988:238) refere que *“há uma preferência nítida pelas sequências n00-”*, pelo que apenas registou seis configurações do tipo NSS, dez do tipo NSØ e uma do tipo NØS. O exemplo apontado por Scherre (1988:238) para este tipo de estruturas (*“dois colega meu”*) vai, curiosamente, ao encontro de todas as ocorrências registadas nos nossos dados (apenas cinco), isto é, envolvendo numerais terminados em –s (N<sub>s</sub>0\_), as quais registam uma inibição total da marcação no terceiro elemento, originando a estrutura SØØ. Apesar de o número de realizações do nosso *corpus* não permitir conclusões sólidas, parece ocorrer uma tendência no sentido de confirmar a estrutura preferencial apontada por Scherre (1988:238). Contudo, remetemos para o ponto 4.1.2.4.6 do presente trabalho a discussão acerca de esta e outras questões envolvendo os numerais, bem como sobre a problemática do favorecimento ou inibição da marcação motivada por estas categorias gramaticais.

Por seu lado, a mistura precedente de marcas formal e semântica (SN\_) exhibe 22 realizações no nosso *corpus*, com um baixo percentual de marcação no terceiro elemento da cadeia sintagmática (32%) a fazer prever que a estrutura SNØ prevalecerá sobre a configuração SNS no PA. O MRJ e a NURC exibem pesos relativos para este tipo de ocorrências, mas tanto o valor do primeiro (pr. 0,52) como o da segunda (pr. 0,53) acabam por se situar no patamar da neutralização. Como tal, não ocorre aqui um favorecimento categórico da marcação, a fornecer evidência clara que, nestes dois

dialectos, a estrutura com mistura de marcas formal e semântica de pluralidade conduz à inserção de marcas no elemento seguinte.

Analisando, por fim, as misturas de marcas em SN's de mais de três itens, é evidente a baixa ocorrência dos mesmos no PA, tendo merecido alguma atenção apenas a configuração NN(N)\_. Em SN's de dois elementos, ficou patente que a ocorrência de um numeral sem *-s* na 1ª. posição favorece moderadamente a inserção PL no elemento da 2ª. posição (pr. 0,607). Este valor é bastante equivalente ao do mesmo tipo de ocorrências com numerais de mais de uma palavra (pr. 0,616), confirmando-se que os informantes de Almojarife tratam numerais simples e numerais de mais de uma palavra de forma idêntica. O mesmo pressuposto é também válido para a NURC, dado que o diferencial de marcação propiciado por numerais sem *-s* de uma (pr. 0,57) ou mais de uma palavra (pr. 0,68) é igualmente pouco significativa.

Quanto à estrutura NNs(N)\_, apenas apresenta nove ocorrências no nosso *corpus*, sete das quais favorecendo a marcação. Este favorecimento confirma-se nos dados de Lopes (2001:204) em valores (pr. 0,65) que não distam significativamente dos apresentados pela configuração NN(N)\_, reforçando-se a probabilidade de os falantes de ambos os dialectos tratarem numerais simples e numerais de mais de uma palavra de forma idêntica, tratamento este que elimina também a distinção entre numerais terminados em *-s* e numerais não finalizados em *-s*.

Relativamente à configuração SSM\_, que apenas regista uma ocorrência nos nossos dados, com marcação na quarta posição (SSMS), foi realizada por um informante com escolarização alta ([CARMOM1]). Na NURC, a mesma estrutura mostra-se moderadamente favorável à inserção PL no quarto elemento (pr. 0,65), fazendo prever a tendência de as estruturas do tipo SSMS se sobreporem às do tipo SSM0, em termos de realização. Todavia, a exemplo do que sugerimos para os SN's de configuração SS\_, seria de toda a conveniência verificar se as estruturas de três ou mais elementos com dupla marcação formal nos dois primeiros itens, devido ao seu grau de complexidade, foram ou não produzidas por informantes de escolaridade superior. Confirmar-se-ia, assim, se o fenómeno que subjaz ao favorecimento da marcação é o Princípio do Processamento Paralelo (Scherre, 1988:208) ou, simplesmente, uma tendência na aquisição da regra da marcação motivada pelo factor social *escolaridade*.



Scherre (1988:174-177), como se viu no ponto anterior, analisou também os SPREP'S. Acerca dos mesmos, a autora entende que apresentam idêntico comportamento ao dos outros itens, podendo ser explicados, tal como os numerais, em termos de a informação semântica dos itens com *-s* conduzirem a mais marcas. Assim, os Sprep(x)s- (pr. 0,69) conduzem a mais marcas do que os Sprep0(0)- (pr. 0,31), que, por seu lado, levam a “*mais marcas do que a existência de um zero num segmento que não tem semântica de pluralidade*” (Scherre, 1988:236). Como referimos já, abstivemo-nos de observar este tipo de configurações, por considerarmos que a sua análise não cai no âmbito do estudo atomístico. Ainda assim, gostaríamos de lembrar que o MRJ regista apenas 19 ocorrências que configuram a estrutura Sprep(x)s- (Scherre, 2001:96), um número manifestamente insuficiente para, em termos de sociolinguística quantitativa, se postular com rigor científico o que quer que seja.

#### **4.1.2.4.2. *Marcas precedentes*: contribuição dos factores amalgamados para a marcação PL do elemento analisado**

As amalgamações de factores efectuadas por Scherre (1988), Lopes (2001) e P. Andrade (2003) foram já amplamente descritas no ponto 4.1.2.4. Observações consideradas pertinentes para as mesmas foram também efectuadas em devido tempo. Não obstante, resta ainda referir que, argumentando quer contra a hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195) de que a propensão para inserir a marca formal de PL ganha força quando é necessário indicar a ideia de pluralidade (daí a sua alta probabilidade de inserção em 2<sup>a</sup> posição quando ocorre ausência de marcação na 1<sup>a</sup> posição) quer contra o pressuposto de Guy (1981b:179) de que se pode estabelecer um paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjectivos e terceira posição, Scherre (1988:183), após ter efectuado a rodada inicial com as amalgamações dos factores da variável independente *marcas precedentes*, cruzou estas com a *posição linear* e a *classe gramatical*. Com tal metodologia, a autora buscou provar que as ausências sistemáticas de marcas na primeira posição dependem de estruturas que envolvem o uso dos possessivos, já que estes revelam tendência para serem marcados quando se encontram na segunda posição do SN, mesmo que antecidos de itens com marcação formal (Scherre, 1988:163). Este comportamento dos possessivos encontra paralelo, ainda segundo a autora, em outras amalgamações: (i)

SPREP's dentro de um SN mais alto com marca semântica de PL, tipo SPrep(x)s-; (ii) realizações SS(S)\_ , ou seja, construções que incluem a presença de duas ou três marcas formais precedentes; (iii) ocorrências com pelo menos uma marca precedente não mediada por zero, isto é, que configuram a equação  $y(x)x-$ . Se a estes comportamentos aliarmos o facto de que “*a presença de um zero precedente, desde que já se tenha marca no primeiro ou no segundo elemento, provoca quase que categoricamente o aparecimento de um outro zero no segundo segmento*” (Scherre, 1988:182), então será lícito admitir que, segundo a autora, a presença de marcas leve a marcas e que a ausência desta conduza a zeros, isto é, estarmos perante motivações de processamento paralelo (Scherre, 1988:208) do tipo claramente não-funcionalista.

Ao observar a relação entre classes gramaticais e posição, Scherre (1988:154) concluíra que as primeiras apresentam comportamentos distintos em função da sua posição na cadeia sintagmática. Se a esta conclusão adicionarmos os pressupostos de que o tipo de marcas precedentes influencia a presença ou ausência de marcas seguintes e também que a ausência fora do SN ou zero na primeira posição têm efeito regular sobre todas as classes gramaticais, então será igualmente lícito aceitar que o português se assemelha ao espanhol em termos de comportamento, isto é, que nenhuma das duas línguas é regida pela hipótese funcionalista das Condições de Distintividade, nos moldes de Kiparsky (1972:195). Assim, torna-se imperativo que se efectuem amalgamações de acordo com a metodologia de Scherre (1988:177), a fim de podermos comparar o padrão de comportamento resultante da influência das marcas precedentes na marcação PL dos itens do SN no PA com o exibido pelo MRJ. Como Lopes (2001:204-205) optou por idênticas amalgamações para a NURC, as que levaremos a cabo para o PA servirão também para as comparações com a variante de São Salvador.

Em segunda etapa, e a fim de estabelecermos comparações com os resultados de P. Andrade (2003:99) e Baxter (2009:281), procederemos também à constituição de factores amalgamados conforme o critério destes autores.

#### **4.1.2.4.2.1. Factores amalgamados para estudo do efeito das *marcas precedentes* na marcação PL do elemento analisado: análise I**

As amalgamações que efectuaremos de acordo com a metodologia de Scherre (1988:177) e Lopes (2001:204-205) terão como finalidade comparar apenas os

resultados do PA com os exibidos pelas duas autoras para o MRJ e a NURC, já que tais comparações não permitirão objectivar conclusões sólidas, por motivos que já abordámos em devida altura e que se prendem, sobretudo, com o facto de ambas as autoras terem levado às amalgamações factores com menos de cinco ocorrências. Por outro lado, nunca é demais lembrar que, na análise VARBRUL, os pesos das probabilidades devem ser lidos em função dos resultados dos pesos relativos e não das percentagens, já que estas não são indicadores fiáveis para os efeitos exercidos nos grupos de factores. Um exemplo concreto desta constatação pode ser visto, por exemplo, nos valores de Lopes (2001:213) resultantes da influência das *marcas precedentes* na marcação PL dos grupos POP e UNI, em que o factor *mistura de marcas antecedentes* apresenta percentuais de marcação bastante desequilibrados entre si (POP = 43%; UNI = 91%), mas com um peso relativo igual para ambos os grupos, isto é, 0,56.

Outro aspecto a ter igualmente em conta prende-se com o facto de a análise em Sociolinguística Quantitativa apenas admitir como confiável, para a determinação do peso que um factor pode ter no grupo de factores em que está inserido, um mínimo de trinta ocorrências. Apesar de, nesta fase do nosso trabalho, não analisarmos os SPREP's dentro de um SN mais alto com marca semântica de PL, Scherre (1988:180) constituiu-os com vista a provar que os seus padrões são regulares, logo, viabilizam a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208) na marcação PL do MRJ. Contudo, o factor SPrep(x)s-, por exemplo, possui apenas dezasseis realizações, com um percentual de marcação de 69% (Scherre, 1988:180). Assim sendo, entendemos que se trata de um factor extremamente débil para ser levado à rodada VARBRUL com vista a sustentar a premissa que a autora pretende, já que a sua inclusão na mesma vai enfermar internamente a distribuição de pesos em termos da variável independente *marcas precedentes*.

Assim, e relativamente aos factores amalgamados por Scherre (1988:117) e Lopes (2001:204-205), que ajudam a sustentar a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208) na marcação PL do MRJ e, conseqüentemente, que este dialecto marca a pluralidade da mesma forma que o espanhol, é importante relembrar que os valores apresentados pelas duas autoras poderão padecer de viciação. De facto,

ambas levaram ocorrências de todo não-confiáveis quer à amalgamação de factores quer às rodadas VARBRUL, ao invés de as retirarem das mesmas.

Passando agora à primeira fase de amalgamação de factores de acordo com a metodologia de Scherre (1988:177) e Lopes (2001:204-205), fundiram-se as configurações abaixo discriminadas e constituíram-se os factores exibidos na Tabela 19.4, que inclui também os resultados da rodada VARBRUL efectuada para apurar o efeito dos mesmos na marcação PL.

- (i) Factor amalgamado *2<sup>a</sup> posição, numeral na 1<sup>a</sup> posição* [N\_]:
  - N\_ (412 ocorrências)
  - Ns\_ (278 ocorrências)
- (ii) Factor amalgamado *mais de uma marca antecedente* [SS(S)\_]:
  - SS\_ (26 ocorrências – retirado das rodadas)
- (iii) Factor amalgamado *mistura de marcas com marca formal ou semântica antecedente* [y(x)x\_]:
  - SM\_ (5 ocorrências)
  - SN\_ (22 ocorrências)
  - OS\_ (7 ocorrências)
  - ON\_ (13 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - NM\_ (5 ocorrências)
- (iv) Factor amalgamado *mistura de marcas com zero antecedente* [x(y)0\_]:
  - S0\_ (30 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - S00\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - N0\_ (2 ocorrências)
  - Ns0\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
- (v) Factor amalgamado *numeral de mais de uma palavra antecedente* [NN(N)\_]:
  - NN(N)\_ (49 ocorrências)
  - NNs(N)\_ (9 ocorrências)

Partindo do princípio que os informantes não fazem a distinção entre numerais terminados em *-s* e numerais não finalizados em *-s*, optámos por considerar como iguais as configurações N0\_ (1/1) e Ns0\_ (0/5) do factor mistura de marcas com zero

antecedente, que passam a perfazer um total de sete ocorrências. Como tal, foram levadas à amalgamação de factores e, em consequência da distribuição de marcação registada pelo novo factor (1/6), eliminou-se também o *knockout* no sentido do favorecimento da pluralização apresentado pelo factor N0\_, já que a totalidade dos dados passa a exhibir, na 3ª. posição, um dado com marcação PL e seis sem marcação. É ainda de considerar que tal distribuição possibilitou também que a totalidade do factor amalgamado não passasse a exhibir inibição plena, podendo assim ser incorporado à rodada VARBRUL.

**Tabela 19.4.** Efeito das *marcas precedentes* na marcação PL dos itens do SN do PA.  
[Factores amalgamados: Análise I]

<i>Input desta rodada:</i> 0,459		<i>Log-likelihood:</i> -966,434		<i>Significância:</i> 0,025	
<b>Factores</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
2ª posição, marca formal na 1ª posição: S__	222/640	35	42	0,378	
2ª posição, zero na 1ª posição: 0__	39/46	85	3	0,861	
2ª posição, numeral sem -s na 1ª posição: N__	200/690	29	45	0,628	
Numeral de mais de uma palavra antecedente [NN(N)__]	16/58	27	4	0,619	
Mais de uma marca antecedente [SS(S)__]	(14/26)	54	-	-	
Mistura de marcas com marca antecedente [y(x)x__]	15/52	29	3	0,339	
Mistura de marcas com zero antecedente [x(y)0__]	1/42	2	3	0,055	
<b>Totais:</b>	<b>507/1.540</b>	<b>33</b>	<b>100</b>	<b>-</b>	

Antes do mais, note-se que, na amalgamação de factores, reduziu-se o número de factores levados à amalgamação de 14 para 7, de modo a tentar formar-se um grupo de factores mais robusto para efeitos de análise. A diferença entre os *log-likelihoods* das Tabelas 19.1 e 19.4 dá-nos uma diferença de -14.575, que multiplicada por 2 fornece um valor final de -29,15. Como temos 7 graus de liberdade para uma significância de 0,025 e o valor da tabela do qui-quadrado é 19.023, confirma-se que as amalgamações segundo a metodologia de Scherre (1988:117) e Lopes (2001:204-205) se revela estatisticamente significativa, representando uma distinção linguística que não deve ser mantida para efeitos de análise. Este aspecto é relevante, no sentido de se considerar que

o estudo por nós levado a efeito e representado na Tabela 19.1 se apresenta como mais consistente e apropriado para se observar, no SN do PA, o efeito das *marcas precedentes* na marcação PL do elemento seguinte.

Como se pode confirmar também, a Tabela 19.4 revela que o reduzido número de ocorrências para a configuração SS(S)\_ impediu que constituíssemos este factor. Quanto ao factor mistura de marcas com zero antecedente [x(y)0-], referimos já que apenas foi levado à rodada devido ao facto de termos deixado de fazer a distinção entre numerais terminados em *-s* e numerais não finalizados em *-s* para efeitos de influência na aplicação da marcação. Supostamente, o alto grau de inibição da marcação revelado pelos factores que constituem esta amalgamação (pr. 0,055), constituiria forte âncora para advogar a possibilidade de que zeros levam a zeros. Todavia, as configurações S00\_ e N0\_, devido ao seu exíguo número de realizações, não dão suporte científico para a formulação de qualquer hipótese. Sobre o factor S0\_ já nos pronunciámos em devido tempo, levando em conta o facto de o mesmo abarcar, na sua maioria, realizações com adjectivo ou o quantificador “todo/tudo” em posição pós-nuclear.

Por seu lado, o factor amalgamado dos marcadores semânticos [NN(N)\_] mostra um valor (pr. 0,619) que confirma não só a tendência para estes itens gramaticais favorecerem ligeiramente a pluralização no elemento seguinte mas também o pressuposto de que os informantes do PA tratam numerais simples e numerais de mais de uma palavra de forma idêntica, não aplicando igualmente distinção no tratamento de numerais acabados em *-s* e numerais não finalizados em *-s*.

Na Tabela 19.5, apresentamos, para efeitos de comparação, os resultados da rodada com os factores amalgamados para o PA, o MRJ e a NURC. A influência dos factores marca formal ou zero na primeira posição (configurações S\_ e 0\_) na marcação PL do elemento seguinte foi já sobejamente discutida, aquando da análise dos resultados dos factores vistos numa perspectiva individual (ponto 4.1.2.4.1). Relativamente ao factor amalgamado numeral na 1ª. posição (N\_), a tendência é para confirmar o moderado favorecimento da pluralização, tanto no PA (pr. 0,628) como na NURC (pr. 0,57). Scherre (1988:229) apresenta um valor de marcação que se situa no patamar do favorecimento moderado (pr. 0,63), mas considera, para a 1ª. posição, numerais simples e numerais de mais de uma palavra.

**Tabela 19.5.** *Marcas precedentes*: contribuição dos factores amalgamados para a marcação PL do elemento analisado: 3 variedades de português.

Factores	PA		MRJ		NURC	
	Pr.	%	Pr.	%	Pr.	%
2 <sup>a</sup> posição, marca formal na 1 <sup>a</sup> posição: S_	0,378	35	0,55	55	0,48	70
2 <sup>a</sup> posição, zero na 1 <sup>a</sup> posição: 0__	0,861	85	(105/105)	100	(34/34)	100
2 <sup>a</sup> posição, numeral sem –s na 1 <sup>a</sup> posição: N_	0,628	29			0,57	62
Numeral de mais de uma palavra antecedente [NN(N)_]	0,619	27	0,63	58	0,68	67
Mais de uma marca antecedente [SS(S)_]	(14/12)	54	0,60	70	0,62	72
Mistura de marcas com marca antecedente [y(x)x_]	0,339	29	0,41	54	0,56	59
Mistura de marcas com zero antecedente [x(y)0_]	0,055	2	0,06	6	0,07	8

O factor amalgamado mais de uma marca antecedente [(SS(S)\_)] não foi por nós levado à rodada VARBRUL, devido ao seu reduzido número de ocorrências (apenas 26). O percentual de marcação situa-se no patamar da neutralidade (54%), não acompanhando plenamente a tendência percentual para favorecer a marcação do MRJ (70%; pr. 0,60), mas aproximando-se do percentual da NURC (56%; pr. 0,62). Sobre a necessidade de este factor ser analisado à luz da sua relação com a variável social *escolaridade*, já nos pronunciámos também (ponto 4.1.2.4.1).

Relativamente ao factor *mistura de marcas com marca antecedente* [y(x)x-], os nossos dados (pr. 0, 339) contrariam os da NURC, que apresenta uma tendência favorecedora da pluralização próxima da neutralização (pr. 0,56), mas revelam comportamento semelhante aos do MRJ, a inibirem também a marcação PL (pr. 0,41). Note-se ainda que, nas nossas ocorrências, a tendência para a inibição é em tudo semelhante à das cadeias sintagmáticas de apenas dois itens (S\_ = pr. 0,368), não se registando uma alteração no padrão da pluralização quando ocorre marcação na primeira oportunidade. Ao que tudo indica, o comportamento deste factor amalgamado confirma que, no PA, a actuação da hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), observada para as cadeias de dois elementos (DET+N), se estende também às estruturas de mais de dois elementos, inserindo-se a marcação apenas quando ela é fundamental para indicar a ideia de pluralidade, isto é, na primeira oportunidade.

Sobre o factor amalgamado *mistura de marcas com zero antecedente* [x(y)0-], que apresenta inibição categórica em todas as variedades, relembramos, uma vez mais, a necessidade de o mesmo ser observado sob a perspectiva da sua clássica configuração estrutural, envolvendo adjectivos ou o quantificador “todo/tudo” pós-nucleares.

Quanto ao factor amalgamado numeral de mais de uma palavra antecedente [NN(N)\_], adiantámos já algumas observações, nomeadamente em relação ao facto de exhibir tendência para favorecer ligeiramente a inserção PL (pr. 0,619) e indicar não-distinção entre numerais terminados em *-s* e numerais não finalizados em *-s*. Ainda assim, os resultados para o nosso factor amalgamado não vão totalmente ao encontro dos achados na NURC (N\_ = pr.0,57; NN(N)\_ = pr. 0,68), que sugere que “*os numerais favorecem mais a concordância do que marcas antecedentes, e mais ainda quando esses numerais são formados por mais de uma palavra*” (Lopes, 2001:212). De facto, como se pode confirmar, no PA apenas colhe a primeira parte desta premissa, já que os numerais de mais de uma palavra apresentam um peso relativo bastante idêntico ao dos numerais de uma só palavra, mas abaixo destes (N\_ = pr.0,628; NN(N)\_ = pr. 0,619). Scherre (1988:177), por seu lado, tratou numerais de uma ou mais que uma palavra de forma idêntica, considerando-os como numerais (N\_).

#### **4.1.2.4.2.2. Factores amalgamados para estudo do efeito das *marcas precedentes* na marcação PL do elemento analisado: análise II**

P. Andrade (2003:99-100) centrou a sua atenção somente no item precedente, eliminando a problemática imposta pelas configurações das misturas de marcas. A diferença entre as análises da autora e as de Baxter (2009:280-282), no que se refere ao dialecto de HEL-Ba, prende-se com o facto de a primeira não ter feito distinção entre as posições ocupadas pelos elementos analisados e o segundo ter optado por analisar, em separado, os SN’s de dois itens e os SN’s de mais de dois itens, considerando, para estes, a presença ou ausência de marca formal de PL em qualquer dos elementos que os antecedem. Contudo, nenhum dos dois refere se foram levadas às amalgamações configurações com menos de cinco realizações.

Nesta fase das nossas análises, constituímos factores amalgamados de acordo com a metodologia de P. Andrade (2003:99), já descrita no ponto 4.1.2.4, elaborando-se os seguintes factores:



- (i) Factor amalgamado *presença de marca formal antes do elemento analisado*:
- S\_ (640 ocorrências)
  - SS\_ (26 ocorrências)
  - OS\_ (7 ocorrências)
- (ii) Factor amalgamado *presença de marca semântica antes do elemento analisado*:
- N\_ (412 ocorrências)
  - Ns\_ (278 ocorrências)
  - SN\_ (22 ocorrências)
  - ON\_ (13 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - NN(N)\_ (49 ocorrências)
  - NNs(N)\_ (9 ocorrências)
- (iii) Factor amalgamado *ausência de marca no elemento anterior*:
- 0\_ (46 ocorrências)
  - S0\_ (30 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - S00\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - SM\_ (5 ocorrências)
  - N0\_ (2 ocorrências)
  - Ns0\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - NM\_ (5 ocorrências)

Novamente, voltámos a partir do princípio que os informantes não fazem a distinção entre numerais terminados em *-s* e numerais não finalizados em *-s*. Assim, aplicámos, para os factores N0\_ e Ns0\_, a metodologia de amalgamação da Tabela 19.4.

A Tabela 19.6 apresenta os resultados dos factores amalgamados, segundo metodologia de P. Andrade (2003:99). A amalgamação de factores da variável independente *marcas precedentes* apresenta agora um *log-likelihood* (-986,405) superior ao da Tabela 19.1 (-981,009), isto é, mais alto do que o do grupo de factores observados numa perspectiva individual. Nesta conformidade, a amalgamação da Tabela 19.6 é estatisticamente significativa, representando uma distinção linguística que não deve ser mantida para efeitos de análise. Este aspecto reforça o que temos vindo a referir, isto é, que a análise das *marcas precedentes* feita numa perspectiva ontogénica

permite não só observar melhor como a marcação PL é inserida, caso a caso, no elemento seguinte mas descortinar também quais são os fenómenos que poderão subjazer às inserções da pluralização. Quanto a nós, este aspecto vem também contribuir a favor do pressuposto de Guy (1981a:178), sobre a possibilidade de a variável independente *marcas precedentes* se revelar estatisticamente mais sólida se for tratada à escala individual.

**Tabela 19.6** *Efeito das marcas precedentes* na marcação PL dos itens do SN do PA.  
[Factores amalgamados – Análise II]

<i>Input desta rodada:</i> 0,450	<i>Log-likelihood:</i> -986,405	<i>Significância:</i> 0,025		
<b>Factores</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Presença de marca formal antes do item analisado	240/673	36	43	0,431
Presença de marca semântica antes do item analisado	223/783	28	51	0,541
Ausência de marca no item anterior	44/98	45	6	0,643
<b>Totais:</b>	506/1.553	33	100	-

Observemos agora os resultados comparativos entre o PA e o dialecto de HEL-Ba (P. Andrade, 2003:99):

**Tabela 19.7.** *Marcas precedentes:* contribuição dos factores amalgamados para a marcação PL do elemento analisado: 2 variedades de português.

<b>Factores</b>	<b>PA</b>		<b>HEL-Ba</b>	
	<b>Pr.</b>	<b>%</b>	<b>Pr.</b>	<b>%</b>
Presença de marca formal antes do item analisado	0,431	36	0,45	8
Presença de marca semântica antes do item analisado	0,541	28	0,59	12
Ausência de marca no item anterior	0,643	45	0,57	15

Procedendo à comparação dos resultados, é possível observar que ambos os dialectos apontam para um padrão de variação bastante idêntico no que concerne à totalidade os factores amalgamados da variável independente *marcas precedentes*. Assim, a presença de marca formal antes do elemento analisado inibe ligeiramente a pluralização nas duas variedades de português em termos bastante aproximados de pesos relativos (PA = pr. 0,431; HEL-Ba = pr. 0,45), verificando-se idêntico padrão de variação quando se analisa também o factor presença de marca semântica antes do item analisado, mas agora com um ligeiro favorecimento da marcação PL em moldes bastante próximos (PA = pr.

0,541; HEL-Ba = pr. 0,59). Relativamente ao factor ausência de marca no elemento anterior, verifica-se que favorece a inserção da pluralização nos dois dialectos, sendo o propiciamento ligeiramente mais acentuado no PA (PA = pr. 0,643; HEL-Ba = pr. 0,57), mas sem comprometer o diferencial geral de pesos registado entre ambos. No cômputo geral, é então possível corroborar P. Andrade (2003:100), que entende que, mais do que uma determinação interna à própria língua, como o defende Scherre (1988), o maior índice de marcas de PL parece estar ligado a factores de ordem social, como por exemplo a *escolaridade*, não se constatando que o Princípio do Processamento Paralelo (Scherre, 1988:208) detenha influência de monta na CPL-var de ambos os dialectos.

O estudo comparado do modo como o grupo de factores *marcas precedentes* intervém na marcação PL do item posterior de diferentes variedades parcialmente reestruturadas de português permitem considerar a forte probabilidade de o português adquirido em situação de contacto por falantes de um CP manifestar os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Esta conclusão vem reforçar, igualmente, a probabilidade de, no PA e no PVB, a CPL-var ser fortemente condicionada pela indicação do PL na primeira oportunidade (cf. Kiparsky, 1972:195), por norma o item colocado na posição imediatamente anterior ao núcleo do SN, devido à influência que ancora nas línguas de substrato africano (Guy, 1981a:301-302) do grupo níger-congo atlântico (Baxter, 2009:292). Nestas, como se viu (ponto 1.4.2 do presente trabalho), predomina o uso de classificadores nominais em posição inicial, em forma de prefixo ou clítico. Deste modo, interessa também comparar factores amalgamados não só do PA com os do PT mas também do PA com os de HEL-Ba, seguindo-se a metodologia de Baxter (2009:280-282), a fim de podermos verificar se o padrão de variação entre o PA e HEL-Ba se mantém idêntico, estendendo-se tal similaridade de variação ao padrão do PT. Confirmar-se-á assim, ou não, a predição de que o português adquirido em situações de transmissão linguística irregular manifesta o mesmo padrão de variação havendo, ou não, interferência de um CP no processo de aquisição.

#### 4.1.2.4.2.3. Factores amalgamados para estudo do efeito das *marcas precedentes* na marcação PL do elemento analisado: análise III.

Baxter (2009:280-282) incluiu SN's de dois itens e SN's de mais de dois itens em factores distintos. De acordo com esta metodologia, elaborámos os seguintes factores amalgamados:

- (i) Factor amalgamado *numeral na 1<sup>a</sup>. posição, item analisado na segunda posição*:
  - N\_ (412 ocorrências)
  - Ns\_ (278 ocorrências)
  - NN(N)\_ (49 ocorrências)
  - NNs(N)\_ (9 ocorrências)
- (ii) Factor amalgamado *mistura de marcas precedentes com marca formal, item analisado na 3<sup>a</sup>. posição*:
  - SS\_ (26 ocorrências)
  - SN\_ (22 ocorrências)
  - SM\_ (5 ocorrências)
- (iii) Factor amalgamado *mistura de marcas precedentes com zero, item analisado na 3<sup>a</sup>. posição*:
  - S0\_ (30 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - S00\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - 0S\_ (7 ocorrências)
  - 0N\_ (13 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - N0\_ (2 ocorrências)
  - Ns0\_ (5 ocorrências = *knockout* – 100% de ausências de marcação PL)
  - NM\_ (5 ocorrências)

Note-se que, uma vez mais, não fizemos distinção entre numerais terminados em *-s* e numerais não finalizados em *-s*. Nesta perspectiva, os factores N0\_ e Ns0\_ foram tratados como um só, seguindo-se a metodologia utilizada nas amalgamações descritas para as Tabelas 19.4 e 19.6. Os numerais de mais de uma palavra foram também entendidos como numerais de uma só palavra, visto termos concluído que os falantes não fazem distinção entre ambos. Por conseguinte, os primeiros (N0\_) foram encarados

como configurando um todo lexical que, na totalidade dos nossos dados, representam apenas duas realizações (retiradas das nossas rodadas) e não se antepõem imediatamente ao item analisado (Tabela 12.3). Assim sendo, os numerais de mais de uma palavra (NN(N)\_ e NNs(N)\_) foram incluídos no factor amalgamado *numeral na 1ª. posição, item analisado na segunda posição*.

Apresentam-se, na Tabela 19.8, os valores da rodada VARBRUL com os factores amalgamados de acordo com a metodologia de Baxter (2009:280-282):

**Tabela 19.8.** Efeito das *marcas precedentes* na marcação PL dos itens do SN do PA.  
[Factores amalgamados – Análise III]

<i>Input desta rodada:</i> 0,467		<i>Log-likelihood:</i> -972,255		<i>Significância:</i> 0,028	
<b>Factores</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
2ª. posição, marca formal na 1ª. posição: S__	222/640	34	41	0,367	
2ª. posição, zero na 1ª. posição: 0__	39/46	85	3	0,854	
2ª. posição, numeral sem -s na 1ª. posição: N__	216/748	29	48	0,622	
3ª. posição, mistura de marcas precedentes com marca formal	23/53	43	4	0,467	
3ª. posição, mistura de marcas precedentes com zero	7/67	10	4	0,190	
<b>Totais:</b>		507/1.554	33	100	-

O *log-likelihood* da Tabela 19.8 (-972,255) é inferior ao da Tabela 10.2 (-981,009), fornecendo uma diferença de -8,754, que, multiplicada por dois, concede um valor de -17,508 para o qui-quadrado. Como o número de factores foi reduzido de 14 para 5, temos 9 graus de liberdade para uma significância de 0,028. A tabela do qui-quadrado exhibe o valor 16,919 para esta situação, confirmando-se que as amalgamações da Tabela 19.9 são estatisticamente significativas. Nesta conformidade, representam uma distinção linguística que não pode ser considerada para efeitos de análise. Reforça-se então, mais uma vez, a vantagem de a análise das *marcas precedentes* ser feita numa perspectiva ontogénica.

A presente amalgamação de factores confirma o padrão de influência na marcação que recai sobre os itens em 2ª. posição do PA. Assim, as rodadas VARBRUL efectuadas apresentam pesos relativos aproximados para os mesmos factores, quer em rodadas com factores individuais quer em análises com factores amalgamados, asseverando que o

item com marca formal na 1ª. posição inibe consideravelmente a pluralização no elemento seguinte (Tabela 19.1 = pr. 0,368; Tabela 19.4 = pr. 0,378; Tabela 19.8 = pr. 0,367), enquanto a presença na mesma posição de um marcador semântico a favorece moderadamente (Tabela 19.1 = pr. 0,607; Tabela 19.4 = pr. 0,628; Tabela 19.8 = pr. 0,622). Por seu lado, a ausência de marcação formal na 1ª. posição conduz a um favorecimento categórico da marcação PL do item em 2ª. posição no SN (Tabela 19.1 = pr. 0,855; Tabela 19.4 = pr. 0,861; Tabela 19.8 = pr. 0,854).

Relativamente aos factores amalgamados, o padrão de marcação sofre alterações consideráveis, principalmente no que concerne ao factor amalgamado com zero antecedente. Observando o factor amalgamado que engloba a mistura de marcas com marca antecedente (se assim se pode dizer), verifica-se que o mesmo se apresenta como inibidor da marcação no PA, indo o grau de inibição desde a ausência categórica de marcação, segundo a metodologia de Scherre (1988:177) e Lopes (2001:205-205) (Tabela 19.4 = pr. 0,339), até à inibição suave, de acordo com a metodologia de P. Andrade (2003:99) (Tabela 19.6 = pr. 0,431) e Baxter (2009:280-282) (Tabela 12.8 = pr. 0,467).

No que concerne à mistura de marcas com zero antecedente, ancora neste factor uma discrepância de resultados que analisaremos em detalhe. As metodologias de Scherre (1988:177), Lopes (2001:204-205) e P. Andrade (2003:99) assemelham-se bastante, já que tiverem em conta a presença de zero antes do item analisado. Contudo, os valores obtidos revelam-se completamente divergentes, já que a amalgamação segundo Scherre (1988:177) e Lopes (2001:204-205) se mostra categoricamente inibidora da pluralização (Tabela 19.4 = pr. 0,055), enquanto a junção de factores de acordo com o entendimento de P. Andrade (2003:99) produz um resultado que aponta para a inserção moderada da marcação PL (Tabela 19.6 = pr. 0,643). Contudo, esta discrepância não surpreende, já que P. Andrade (2003:99), ao incidir a sua atenção unicamente no item que antecede o elemento analisado, acabou por constituir um factor amalgamado em tudo idêntico ao da 2ª. posição, zero na 1ª. posição (0\_), que, aliás, foi incorporado à amalgamação em causa. Este factor, como foi sobejamente observado, é categoricamente propiciador da marcação, acabando os resultados finais das misturas de marcas, também favorecedores, por espelharem, redundantemente, os valores do factor 2ª. posição, zero na 1ª. posição.

Baxter (2009:281), por seu lado, levou em consideração, para as misturas de marcas, a ocorrência em qualquer posição antecedente. Ainda assim, os resultados a que chegou indiciam uma inibição categórica tanto para HEL-Ba (Tabela 19.9 = pr. 0,099) como para o PT (Tabela 19.9 = pr. 0,190).

Na Tabela 19.9 exibem-se os resultados comparativos entre o PA, o dialecto de HEL-Ba e o PT (Baxter, 2009:281). Comparando os dados do PA com os dos outros dois dialectos, note-se que, para a mistura de marcas com marca antecedente, HEL-Ba, à semelhança do PA, apresenta sempre um padrão de inibição, embora menos oscilante [Tabela 19.7, metodologia de P. Andrade (2003:99) = pr. 0,45; Tabela 19.9, metodologia de Baxter (2009:181) = pr. 0,412]. Olhando atentamente para a flutuação apontada para o PA nas três metodologias de amalgamação, constata-se que as que seguiram os pontos de vista de P. Andrade (2003:99) e Baxter (2009:181) apresentam valores aproximados (Tabela 19.6 = pr. 0,432; Tabela 19.8 = pr. 0,467), sendo a discrepância maior decretada pela amalgamação segundo o entendimento de Scherre (1988:177) e Lopes (2001:204-205) (Tabela 19.4 = pr. 0,339).

**Tabela 19.9.** *Marcas precedentes*: contribuição dos factores amalgamados para a marcação PL do elemento analisado: 3 variedades de português.

Factores	PA		HEL-Ba		PT	
	Pr.	%	Pr.	%	Pr.	%
2 <sup>a</sup> . posição, marca formal na 1 <sup>a</sup> . posição: S__	0,367	41	0,499	8	0,583	48
2 <sup>a</sup> . posição, zero na 1 <sup>a</sup> . posição: 0__	0,854	3	0,844	56	0,651	36
2 <sup>a</sup> . posição, numeral sem -s na 1 <sup>a</sup> . posição: N__	0,622	48	0,641	14	0,413	21
3 <sup>a</sup> . posição, mistura de marcas precedentes Com marca formal	0,467	4	0,412	12	0,557	52
3 <sup>a</sup> . posição, mistura de marcas precedentes com zero	0,190	4	0,099	3	0,190	7

Continuando a analisar o mesmo factor (*mistura de marcas com marca antecedente*), detecta-se também um paralelo de marcação, agora favorecendo muito ligeiramente a marcação PL em pesos idênticos, entre a NURC (Tabela 19.5 = pr. 0,56) e o PT (Tabela 19.9 = pr. 0,557). Este aspecto vem, confirmar, quanto a nós, que a inserção de marcas determinada pelo Princípio do Processamento Paralelo, conforme o advogam Scherre (1988:208) e Lopes (2001:212), não é senão aparente. De facto, não restam dúvidas que

a NURC e o PT estarão num estágio mais avançado de aquisição da regra de concordância PL, quando comparados com HEL-Ba e o PA. Estágio esse que se fica a dever, não a factores linguísticos, mas antes a condicionantes estritamente sociais, como o grau de *escolaridade* (no caso da NURC) ou o prolongado contacto com o PE (na situação do PT).

No que diz respeito ao factor *mistura de marcas com zero antecedente*, nota-se que os resultados de Baxter (2009:181) (HEL-Ba = pr. 0.099; PT = pr. 0,190) se demarcam claramente dos de P. Andrade (2003:99) (Tabela 19.7 = pr. 0,57), revelando-se bastante mais consentâneos com os das outras variedades de português: PA (Tabela 19.8 = pr. 0,190); NURC (Tabela 19.5 = pr. 0,07). A discrepância do resultado de P. Andrade (2003:99) exibida neste factor tem a ver com o facto de a amalgamação levada a cabo por esta não ser mais do que uma redundância do factor *2ª. posição, zero na 1ª. posição (0\_)*, como apontámos já.

As incongruências no que se refere às amalgamações do grupo de factores *marcas precedentes* chegam mesmo a ser surpreendentes, se levarmos em conta também que os vários testes do qui-quadrado por nós efectuados consideram como não estatisticamente significativas as amalgamações segundo a metodologia de P. Andrade (2003:99), logo, uma distinção linguística válida para efeitos de análise. Estes aspectos vêm, quanto a nós, trazer à evidência que as amalgamações de factores não podem ser levadas a cabo de “ânimo leve”, exigindo-se, para a sua constituição, reflexão cuidada e metodologia científica apropriada, sob pena de se obterem resultados que não traduzem a realidade dos factos linguísticos inerentes à comunidade em observação. Assim, e em última análise, reforça-se, indubitavelmente, a nossa perspectiva de a variável independente *marcas precedentes* ser analisada de forma ontogénica (cf. Guy, 1981a:178), sob pena de se elaborarem, para resultados viciados, pressupostos linguísticos falaciosos.

De qualquer forma, e independentemente das incongruências registadas, é nosso entender que os resultados fornecidos para o modo como o grupo de factores *marcas precedentes* intervém na marcação PL do item posterior das diferentes variedades parcialmente reestruturadas de português apontam para um padrão de variação bastante idêntico em todas elas. Nesta conformidade, reforçamos também o nosso entendimento de que o português adquirido em situação de contacto por falantes de um CP manifesta



os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Pressuposto este que nos leva, de novo, à questão da CPL-var ser fortemente condicionada, nos diferentes dialectos observados, pela indicação do PL na primeira oportunidade (Kiparsky, 1972:195), por influência dos sistemas estruturais africanos (cf. Guy, 1981a:301-302), sobretudo do grupo níger-congo atlântico, e não pelo Processamento Paralelo, como advoga Scherre (1988:208). Assim sendo, estará também descartada a possibilidade de as variedades de português em análise manifestarem o mesmo padrão de variação do EPR, em que marcas levam a marcas e zeros conduzem a zeros.

#### **4.1.2.4.3. Relação entre *marcas precedentes*, *posição linear* e *classe gramatical* na marcação PL dos itens do SN**

Scherre (1988:183) efectuou ainda, como se adiantou já, extensos testes em que cruza *marcas precedentes*, *posição linear* e *classe gramatical*. Com estes, a autora buscou confirmar não só se as marcas formais ou zeros afectam de modo idêntico todas as classes gramaticais em função da sua posição na cadeia sintagmática, inclusive as situações de SPREP'S, mas também se existe realmente um paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjectivos e terceira posição (cf. Guy, 1981b:179).

Ao cruzar apenas *classe gramatical* e *posição linear*, Scherre (1988:156) concluiu que as classes gramaticais apresentam comportamentos distintos em função das diferentes posições que ocupam na cadeia do SN. Os substantivos, por exemplo, são menos marcados na segunda do que na terceira posição, enquanto os adjectivos e o possessivos sofrem marcação inversa, isto é, surgem mais marcados na segunda do que na terceira posição. Por seu lado, o quantificador é pouco marcado em qualquer destas posições. Ao observar, depois, o comportamento das *marcas precedentes*, Scherre (1988:168) referiu mais dois aspectos a ter em conta nos seus dados: (i) o tipo e quantidade de marcas precedentes têm influência directa na presença ou ausência de marcas seguintes; (ii) a ausência de marca precedente (/\_) e a não-marcação na primeira posição (0\_) têm efeito regular e quase categórico sobre qualquer classe gramatical.

Analisando a relação entre os três grupos de factores referidos, Scherre (1988:183) procura respostas não só para o comportamento dos itens morfológicos em função das restantes posições ocupadas na cadeia sintagmática mas também para o das próprias configurações sintagmáticas. Assim, no que concerne às terceiras, quarta e quinta posições, todas as classes gramaticais apresentam, segundo a autora, um comportamento regular em termos de efeito exercido pelas *marcas precedentes*, confirmando-se o Princípio do Processamento Paralelo (Scherre, 1988:208). Como tal, e segundo Scherre (1988:190-191), os substantivos demonstram preferência por realizações do tipo SS\_ (duas marcas formais antecedentes) ou yx- (marcação ou não-marcação formal na 1ª. posição e item representado por marcação formal ou semântica na 2ª. posição), enquanto os adjectivos são uniformemente realizados em sequências do género SS\_, yx- e x0- (item antecedente não-adjacente representado por marcação formal ou semântica na 1ª. posição e não-marcação na 2ª. posição). Quanto aos quantificadores e possessivos, mostram apetência por sequências x0- e praticamente não ocorrem em configurações do tipo yx-.

As observações de Scherre (1988:192) dirigem-se depois para o comportamento das classes gramaticais em SN's mais altos, a fim de constatar se as mesmas são afectadas pelas marcas precedentes. Como não analisamos este tipo de configurações, limitamo-nos a referir que, segundo a autora, a configuração SPrepS- (núcleo nominal mais alto com marca formal de PL) favorece mais a marcação no substantivo seguinte do MRJ, enquanto a ocorrência SPrep+S\_ (núcleo nominal mais alto com marca formal de PL seguido e item com marcação formal de PL) propicia a marcação em substantivos e adjectivos. Em oposição, o SPrep0\_ (núcleo nominal mais alto sem marca formal de PL) desfavorece a inserção da pluralização em substantivos, tal como a estrutura SPrep+0\_ (núcleo nominal mais alto sem marca formal de PL seguido de item sem marcação formal de PL) inibe a marcação em adjectivos. Contudo, a configuração SPrep0\_ revelou que, caso a preposição seja seguida de indefinido (segunda posição, portanto), este apresenta uma marcação plena. Perante este facto, Scherre (1988:194) decidiu constatar se as outras classes gramaticais em segunda posição sofrem igual afectação por parte dos diferentes elementos que as precedem, o que acabou por não se confirmar. Assim, caso ocorra marcação na primeira posição, o efeito da mesma sobre o elemento

em segunda posição varia de acordo com a classe gramatical. Por outro lado, a ausência de marcação na primeira posição conduz à inserção de marca no elemento seguinte, independentemente da classe gramatical a que este pertence. Neste aspecto, e como referimos no final do ponto 4.1.2.4.2.3, as variedades de português por nós observadas diferenciam-se do espanhol, já que, neste, a não-marcação na primeira posição leva a zeros tanto no item seguinte como no terceiro elemento da cadeia sintagmática (Poplack, 1980a, 1981), sendo a informação semântica da pluralização, nestes casos, captada em função do contexto.

Por fim, Scherre (1988:197) procurou ainda confirmar qual a influência dos numerais em primeira posição sobre os elementos gramaticais em segundo lugar. Não obstante, como aqueles ocorrem apenas com substantivos ou categorias substantivadas, a análise inviabilizou-se no que concerne às restantes categorias gramaticais. Quando efectuarmos a análise da influência da variável independente *contexto fonológico posterior* (capítulo 5) na marcação do elemento antecedente, iremos abordar a questão da influência dos numerais sobre os substantivos, uma vez que a mesma interessa para questões que motivam tal marcação.

#### **4.1.2.4.4. Análises alternativas considerando *marcas precedentes*, *posição linear* e *classe gramatical***

Scherre (1988:198) levou ainda a cabo mais quatro tipos de análises alternativas, considerando diferentes factores para as variáveis independentes *marcas precedentes*, *posição linear* e *classe gramatical*, a fim de tentar provar que a posição isolada de *marcas precedentes* e *classe gramatical* não dá conta da totalidade dos fenómenos linguísticos que se pretendem estudar. Contudo, é de referir que, devido não só ao facto de discordarmos da metodologia das amalgamações de Scherre (1988:177) e Lopes (2001:204-205), como temos vindo a dar conta, mas também por concordarmos com o pressuposto de Guy (1981a:178) no que concerne ao facto de a variável independente *marcas precedentes* se apresentar mais sólida quando observada ontogenicamente, como confirmámos nos pontos 4.1.2.4.1, 4.1.2.4.2 e 4.1.2.4.3, não vislumbramos vantagem em buscar um paralelo entre os nossos factores e os agora constituídos por Scherre, a fim de se compararem resultados. Aliás, voltamos a referir que, pelo facto de as amalgamações de factores levadas a cabo por Scherre (1988:177) incluírem ocorrências que

configurarão meros idiolectos, poderão apresentar resultados viciados. Como tal, importará então referir apenas os tipos de análises efectuados pela autora e as conclusões que a mesma delas retirou. Em etapas posteriores do nosso trabalho, poderemos lançar mão de algumas destas conclusões, caso elas se revelem pertinentes para sustentarmos pontos de vista das nossas análises.

A fim de efectuar comparações com os resultados apresentados por Guy (1981a:168-180), Scherre (1988:198) levou em conta, para a sua primeira análise, quatro factores para a variável independente *posição linear* e cinco factores para o grupo de factores *marcas precedentes*. Ainda assim, estes factores apresentam ligeiras diferenças, se comparados com os de Guy (1981a:170-177) – que não efectuou cruzamentos com a variável independente *classe gramatical* –, justificadas por Scherre (1988:199) com a necessidade de permitirem também comparações com os dados da sua segunda análise. Nesta, mantiveram-se inalteráveis as variáveis independentes *posição linear* e *classe gramatical*, mas foram incluídos nove factores no grupo de factores *marcas precedentes*, constituídos com base na linha observada por Poplack (1980a), e que torna este grupo uma quase subdivisão da variável *posição linear* (Scherre, 1988:199).

Na primeira análise, os resultados de Scherre (1988:204) e Guy (1981a:179) apresentam-se bastante semelhantes e confirmam números anteriores, com a ausência de marcação formal na primeira posição favorecendo a concordância e a presença de pelo menos uma marca antes do item analisado desfavorecendo-a. Este desfavorecimento acentua-se caso a marca formal não se encontre em posição adjacente. Em adição a estas observações, foi também possível constatar que tanto o numeral como o SPREP, ambos detentores de marcas semânticas de PL, apresentam comportamentos distintos, com o primeiro favorecendo ligeiramente a marcação e o segundo inibindo-a suavemente. Com vista a conseguir uma resposta acerca das motivações que conduzem a tais comportamentos, Scherre (1988:205) levou então a cabo a segunda análise.

Os resultados obtidos não só ajudaram a autora a tentar sustentar os pressupostos do Princípio do Processamento Paralelo (Scherre, 1988:208) como lhe permitiram lançar um olhar sobre estruturas que estariam parcialmente encobertas na primeira análise. Assim, quer os SPREP's que contenham um *-s* (SPreps) quer os que são seguidos de *-s* (SPrep+s) conduzem à inserção de marcas nos elementos analisados (pr. 0,70 e pr. 0,77,

respectivamente), enquanto os SPREP's que contêm um zero (SPrep) ou são seguidos de zero (SPrep+0) favorecem a inibição no item que lhes é posterior (pr. 0,40 e 0%, respectivamente). Este comportamento regular de marcas levarem a marcas e zeros originarem zeros apenas é contrariado pela configuração zero na primeira posição (0\_), que favorece a marcação (100%), tal como havia já sido verificado na primeira análise alternativa.

Com base nas subdivisões efectuadas aos SPREP's, Scherre (1988:207) postula que não é a configuração SPrep que desfavorece a marcação, como até então havia sido aventado por todos os estudos em sociolinguística variacionista, mas sim o SPrep0(0)\_. Como estes incluem a grande maioria de ocorrências (92%), condicionariam a observação do grupo SPrep, em que estariam inseridos conjuntamente com os SPrep(x)s-, que favorecem a inserção PL. Contudo, no que concerne a estes, achamos pertinente voltar a referir que, nos dados de Scherre (1988:205), os mesmos apenas apresentam um total de 16 realizações. Tendo em conta que um valor abaixo das trinta ocorrências deixa de ser confiável para a determinação do peso que um factor pode ter no grupo de factores em que está inserido (Guy & Zilles, 2007:153), e face à impossibilidade de este factor ser amalgamado com outro(s), deveria ter sido excluído das rodadas. Como tal não sucedeu, o peso relativo apresentado por Scherre (1988:205) poderá ser questionado quanto à sua transparências em termos de realização comunitária. Por outro lado, a metodologia seguida pela autora, no que respeita à questão da subdivisão dos SPREP's, acaba por reforçar ainda mais a nossa perspectiva sobre a necessidade de se analisar o grupo de factores *marcas precedentes* à escala ontogénica, a fim de se descortinarem fenómenos que poderão estar a ser encobertos pelas amalgamações. Aliás, Tagliamonte (2006:157) alerta para o encobrimento de generalizações linguísticas válidas originadas pelo fenómeno de “*kitchen sink effect*”, caso, guiado apenas pelo esquisso de que as amalgamações fortalecem mais a variável independente, o pesquisador opte por colocar “tudo” o que compartilhe traços ou características no mesmo factor do grupo de factores.<sup>170</sup>

Regressando aos achados de Scherre (1988:2007), os mesmos permitiram que a autora continuasse a buscar explicações que justificassem, em termos do Princípio do Processamento Paralelo (Scherre, 1988:208), as motivações que levam o numeral e o

SPREP a apresentarem comportamentos distintos: a carga semântica anexa ao SPrep(x)s- favorece mais a aplicação da regra do que a dos numerais, porque duas marcas levam a mais marcas. Por seu lado, o numeral na primeira posição, devido à carga semântica de pluralidade que carrega, favorece mais a inserção de PL do que a configuração SPrep(0)\_. Esta, por sua vez, vê o seu peso semântico de pluralidade neutralizado pelo zero formal, registando então valores inibidores. Por fim, o factor x(y)0\_, que não contém intrinsecamente qualquer marca semântica de PL, apresenta um índice superior de inibição, se contrastado com o factor SPrep(0)\_.

Em termos de processamento paralelo (Scherre, 1988:208) foi também avançada a explicação para as divergências exibidas entre as diferentes configurações que encerram marcas formais de PL: mais marcas de uma só natureza [SS(S)\_] levam a mais marcas do que uma só marca [S\_] ou marcas de natureza distinta [y(x)x\_ (pelo menos um s- precedente não mediado por zero) e x(y)0\_ (pelo menos um zero precedente não mediado por s-, com marca na primeira posição)] (Scherre, 1988:208).

Concluídas as análises e observados os seus resultados, Scherre (1988:209) parte para a análise do comportamento da *classe gramatical* face às duas variáveis independentes estudadas. Os resultados evidenciaram probabilidades semelhantes entre as duas análises, mas a autora percebeu que não ocorria uma linha geracional natural, em termos das realizações de crianças e adultos, envolvendo classes gramaticais. Adicionalmente, apesar de verificar que a primeira posição é aquela que mais favorece a regra de concordância, Scherre entende que a variável independente *posição linear* não só não é, isolada, a mais forte das três em discussão como provoca também interferências com outras variáveis, que conduzem a resultados indesejáveis. Como tal, volta a contrariar a suposição de Guy (1981a:179) de que *marcas precedentes* não devem ser cruzadas com *posição linear* e reforça o pressuposto de que todas elas devem ser estudadas em conjunto. Assim, decide transformar os grupos de factores *classe gramatical* e *posição linear* numa só variável independente, que vai observar em paralelo com o grupo de factores *marcas precedentes*. A partir deste momento, Scherre (1988:212) opta também por abandonar a primeira análise alternativa, não só por considerar que esta provoca uma desvinculação entre as variáveis *marcas precedentes* e *posição linear* mas também por entender que a segunda análise é a reveladora de melhores resultados linguísticos.

Ainda assim, e antes de proceder à fusão das duas variáveis, Scherre (1988:214) efectua um cruzamento entre ambas e torna a concluir que não se pode estabelecer identidade ou paralelo entre elas, conforme advogado por Guy (1981a:179). A conclusão da autora tem como sustentação o facto de o comportamento das categorias não-nucleares do SN dever ser observado em função do modo como elas se distribuem em torno do núcleo e não da sua classe gramatical. Este aspecto reforça, uma vez mais, a necessidade de se fundirem classe e posição, o que leva à constituição da variável independente *relação entre os elementos do SN*, observada em sintonia com o grupo de factores *marcas precedentes* da segunda análise (Scherre, 1988:219). Para a criação dos nove factores da nova variável independente, a autora amalgamou os elementos não-nucleares em função da localização pré ou pós-nuclear, por um lado, e considerou os itens nucleares em função da sua posição na cadeia do SN, por outro lado. No entanto, mais uma vez, não foi efectuado o teste do qui-quadrado para se confirmar se as amalgamações são ou não estatisticamente significativas.

Os resultados mostram que os substantivos e categorias substantivadas são sempre menos marcados na segunda posição do SN, enquanto as classes gramaticais não-nucleares, independentemente da sua categoria, têm mais probabilidades de apresentarem marcação se estiverem antepostas ao núcleo do SN, do que se ocorrerem depois deste. Como tal, o SN apresentará uma coesão inquestionável, já que “*os elementos não-nucleares se agrupam em função da sua relação com o núcleo do SN e os núcleos se agrupam em termos da posição que ocupam dentro do SN*” (Scherre, 1988:221). Paralelamente, os únicos itens que admitem elementos entre eles serão os substantivos e os elementos que se lhes pospõem, já que entre as realizações antepostas e os seus núcleos não são admitidos, ou raramente são admitidos, outros intervenientes. Scherre (1988:222-223) justifica este comportamento apelando ao Princípio da Iconicidade ou da Motivação Icónica (Haiman, 1983:782), que defende que, quanto mais coesão existe entre os itens, mais marcas co-ocorrem (os elementos pré-nucleares estabelecem uma relação mais formal com o seu núcleo, daí deterem mais inserções para marcarem a inseparabilidade). Como tal, e correspondendo ao oposto do princípio (quanto menos marcas, menos coesão), os itens pós-nucleares marcam a sua menor coesão sintagmática com o núcleo, admitindo maior inibição formal do número.

Scherre (1988:223) aborda também o comportamento dos pronomes pessoais, explicando a sua forte tendência para favorecerem a marcação em virtude de pertencerem a uma classe com paradigma bem definido. Os nossos dados não revelam ocorrências deste tipo, o que impossibilita quaisquer comparações entre o PA e MRJ, neste aspecto. De acordo com os seus resultados, Scherre (1988:227), postula que não se pode estabelecer um paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjetivos e terceira posição, como fora preconizado por Guy (1981b:179), sem que se percam factos linguísticos importantes.

Na quarta análise, Scherre (1988:227-228) constata ainda mais algumas ocorrências: (i) o possessivo tem mais tendência a aceitar o artigo antes de si quando se encontra em posição pré-nuclear; (ii) o quantificador posposto surge, por vezes, realizado com recurso à forma invariável “tudo”, em substituição de “todos”; (iii) o adjetivo anteposto é, por norma, do tipo “avaliativo”, o que pode fazer pressupor que a sua maior tendência de marcação esteja relacionada com a questão do traço semântico e não da sua distribuição em torno do núcleo;<sup>171</sup> (iv) a presença de SS(S)\_ continua a favorecer mais a aplicação da marcação no elemento seguinte do que as marcas de natureza distinta [configurações y(x)x\_ (pelo menos um s- precedente não mediado por zero) e x(y)0\_ (pelo menos um zero precedente não mediado por s-, com marca na primeira posição)]. Em termos gerais, Scherre (1988: 233-235) conclui que a influência da variável independente *marcas precedentes* não encontra qualquer suporte funcionalista do tipo “kiparskiano”. Como tal, o que acontece é que ocorre um processamento paralelo de marcação, em que “*marcas conduzem a marcas, inclusive quando esta marca é semântica*” (Scherre, 1988:235). Este aspecto leva-nos, assim, à questão dos marcadores semânticos e à confrontação que Scherre estabeleceu entre os numerais e os SPREP’s que encerram em si informação do tipo semântico [SPrep(0)\_ e SPrep0(0)\_], aspectos que abordaremos no ponto 4.1.2.4.6 do nosso trabalho.

Depois de efectuadas as quatro análises alternativas e observados os seus resultados, Scherre (1988:240) conclui que, do ponto de vista estatístico, é de facto a análise que transforma as três variáveis independentes *marcas precedentes*, *posição linear* e *classe gramatical* em apenas duas – *marcas precedentes em função da posição* e *relação entre elementos não nucleares e posição dos elementos nucleares dentro do SN* – aquela que



se revela mais adequada para o estudo dos fenómenos linguísticos referidos durante as etapas do estudo por si realizadas. Esta análise, ainda segundo a autora, será a que melhor conta dá da descrição de aspectos regulares que envolvem, por exemplo, casos de poucas ocorrências. Aspectos estes que acabam também por trazer à evidência que o português e o espanhol, “*duas línguas que marcam a pluralidade da mesma forma, apresentam mais similaridades do que diferenças, no que diz respeito a um condicionamento forte não funcionalista no sentido de Kiparsky (1972)*” (Scherre, 1988:240-241).

#### **4.1.2.4.5. Faixas etárias e perfil de marcação motivado pelas marcas precedentes**

Ao analisar as probabilidades de concordância originadas pelas distintas classes gramaticais nas diferentes posições da cadeia linear do SN, Scherre (1988:208-210) constatou que, em crianças e adultos, as referidas classes não seguiam uma linha natural em termos de agrupamentos. Impossibilitada de estabelecer agrupamentos gramaticais homogêneos que lhe permitissem observar a influência dos mesmos na marcação PL exibida pelos SN's dos dois grupos etários, a autora optou mesmo por transformar *classe gramatical e posição linear* numa só variável, para a cruzar com as *marcas precedentes*, como referimos no ponto anterior. Esta metodologia vem ao encontro das próprias convicções de Scherre (1988:428), que entende que a simples distribuição etária não é suficiente para indicar estar-se perante uma mudança em curso.

Lopes (2001:210) analisa também o efeito da variável independente *marcas precedentes* na concordância PL da faixa etária mais idosa da NURC (+55 anos), mas comparando o grupo popular com o grupo universitário. Os resultados comprovam que ambos possuem diferentes gramáticas, com o primeiro apresentando um percentual de marcação próximo da neutralidade (58%) e o segundo um percentual roçando a concordância plena (97%). Em termos gerais, Lopes (2001:129-130) detectou também que os falantes menos idosos aplicam mais a regra de concordância e chega a aventar a possibilidade de existir conexão entre inibição e falantes de ancestralidade negra.

P. Andrade (2003:124-125), apenas apresenta valores gerais acerca da realização da concordância em número no SN em função das faixas etárias de HEL-Ba, concluindo

que a regra de concordância se encontra em processo de mudança aquisicional e que os falantes mais velhos são os que menos marcam o PL.

Baxter (2009:283), ao comparar os dados do PT e de HEL-Ba, chegou às mesmas conclusões de P. Andrade (2003:124-125): não só o PT e HEL-Ba estão em processo de mudança aquisicional como também são os falantes mais idosos que menos aplicam a regra de concordância. Porém, Baxter (2009:283) constata que os falantes das faixas etárias intermédias (faixas 2 e 3) constituem duas classes distintas: filhos de casais africanos e filhos de casais mistos (africano+tonga, isto é, nascido na Roça Monte Café). Assim, Baxter (2009:204) procede a uma análise para observar como ambos os grupos aplicam a concordância em número no SN e conclui que a origem dos pais tem influência nesta, uma vez que os filhos de pais africanos inibem a marcação (pr. 0,36) enquanto os filhos dos casais mistos a favorecem (pr. 0,63).

Jon-And (2008:4), tal como P. Andrade (2003:124-125), exhibe unicamente valores totais para a aplicação da regra de concordância. Os resultados indicam que, em Maputo, a idade menor favorece a aplicação da pluralização, muito provavelmente porque o português é, cada vez mais, a língua dominante na vida quotidiana da cidade, especialmente entre as gerações mais jovens. Para o PCV, Jon-And (2009) não observou o efeito do factor *idade* na marcação PL do SN.

Relembramos que o que se pretende observar nesta etapa do nosso trabalho é a influência das *marcas precedentes* na regra de concordância PL, ao longo das gerações. Como tal, os nossos dados apenas poderão ser comparados com os de Scherre (1988) e Lopes (2001), ainda assim parcialmente, dado que foram estas autoras as únicas a apresentar valores para este tipo de estudo, embora de forma distinta da nossa: Scherre (1988:209), para crianças e adultos; Lopes (2001:212-213), para a geração mais idosa das classes popular e universitária.

Exibem-se, na Tabela 19.10, os resultados que permitem aquilatar a forma como as *marcas precedentes* influenciam, geracionalmente, a regra da concordância em número no PA. Como se poderá verificar, optámos por incluir na tabela de resultados todos os factores com pelo menos cinco realizações a nível comunitário, de forma a conceder um panorama geral sobre o desenvolvimento da concordância ao longo das três gerações.

Observando-se a evolução geracional da regra de concordância PL motivada pelas *marcas precedentes*, detecta-se um perfil no PA que permite retirar algumas conclusões pertinentes. Antes de mais, note-se como é constante na comunidade, ao longos das gerações, a tendência para a não realização de SN's que envolvam três ou mais itens, isto é, sintagmas cuja concordância em número pressupõe um grau mais elevado de aquisição da respectiva regra. Relativamente aos SN's de estrutura simplificada (dois itens), a realização com marca formal na 1<sup>a</sup>. posição (S\_) espelha uma linha geracional em direcção à aquisição da regra de concordância, com os mais idosos (FE-3) não favorecendo categoricamente a marcação (pr. 0,193), a geração intermédia (FE-2) inibindo-a moderadamente (pr. 0,370) e os mais jovens (FE-1) beneficiando-a mais, mas ainda assim sem conseguirem atingir o patamar da marcação positiva (pr. 0,462). Esta linha natural de aquisição é justificada, quanto a nós, pela entrada em jogo de factores sociais como a *escolaridade* ou a inclusão dos elementos masculinos nos quadros do exército colonial, que proporcionaram aos falantes almozarifanos um contacto mais acentuado com o PE.

No que concerne ainda a este factor, Scherre (1988:180) apresenta, para o MRJ, percentuais que indiciam uma linha oposta à do PA, uma vez que as crianças (41%) aparentam marcar menos o PL do que os adultos (55%). Não obstante, a leitura em termos de percentuais não oferece a mesma fiabilidade do que a dos pesos relativos, quando se pretende observar o comportamento de determinados fenómenos em análises quantitativas. Relativamente aos dados de Lopes (2001:215), constata-se que os falantes mais idosos do grupo popular (pr. 0,44) estarão já num estágio de aquisição da regra de concordância bastante próximo daquele que é exibido pela geração mais jovem do PA (pr. 0,462). Olhando também para os informantes universitários da mesma faixa etária desta variedade brasileira de português, é possível constatar que a variável social *escolaridade* leva a que a regra da concordância se manifeste já, embora de modo bastante ténue (pr. 0,53). Contudo, a generalidade dos dados vêm confirmar, uma vez mais, o que temos vindo a afirmar: o PA estará num estágio mais baixo do que estas duas variedades, no que concerne à aquisição da regra PL.

**Tabela 19.10.** Efeito das marcas precedentes na marcação PL dos itens do SN do PA: *faixas etárias*.

	<i>Input desta rodada: 0,504</i> <i>Log-likelihood: -981,009</i> <i>Significância: 0,011</i>			<i>Input desta rodada: 0,196</i> <i>Log-likelihood: -187,390</i> <i>Significância: 0,034</i>		
	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>		
<b>Marcas Precedentes</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
2 <sup>a</sup> posição, marca formal na 1 <sup>a</sup> posição: S	222/640	35	0,368	22/130	17	0,193
3 <sup>a</sup> posição, 2 marcas formais precedentes: SS	(14/26)	54	-	(1/2)	50	-
S0	(0/30)	0	-	(0/8)	0	-
S00	(0/5)	0	-	(0/1)	0	-
SM	(2/5)	40	-	-	-	-
SN	(7/22)	32	-	(1/5)	20	-
2 <sup>a</sup> posição, zero na 1 <sup>a</sup> posição: 0	39/46	85	0,855	(1/2)	50	-
0S	(4/7)	57	-	-	-	-
0N	(0/13)	0	-	(0/8)	0	-
2 <sup>a</sup> posição, numeral sem –s na 1 <sup>a</sup> : N	121/412	29	0,607	25/145	17	0,748
2 <sup>a</sup> posição, numeral em –s na 1 <sup>a</sup> : Ns	79/278	28	0,556	12/95	13	0,574
Ns0	(0/5)	0	-	(0/2)	0	-
NM	(2/5)	40	-	(0/2)	0	-
NN(N)	14/49	29	0,616	(0/6)	0	-
NNs(N)	(2/9)	22	-	(0/1)	0	-
	<i>Input desta rodada: 0,371</i> <i>Log-likelihood: -357,976</i> <i>Significância: 0,018</i>			<i>Input desta rodada: 0,738</i> <i>Log-likelihood: -411,403</i> <i>Significância: 0,039</i>		
	<b>FE-2 (41 – 60 Anos)</b>			<b>FE-1 (20 – 40 Anos)</b>		
<b>Marcas Precedentes</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
2 <sup>a</sup> posição, marca formal na 1 <sup>a</sup> posição: S	67/197	34	0,370	133/313	43	0,462
3 <sup>a</sup> posição, 2 marcas formais precedentes: SS	(6/9)	67	-	(7/15)	47	-
S0	(0/15)	0	-	(0/7)	0	-
S00	(0/2)	0	-	(0/2)	0	-
SM	-	-	-	(2/5)	40	-
SN	(3/4)	75	-	(3/13)	23	-
2 <sup>a</sup> posição, zero na 1 <sup>a</sup> posição: 0	29/33	88	0,817	(9/11)	82	-
0S	(1/4)	25	-	(3/3)	100	-
0N	(0/2)	0	-	(0/3)	0	-
2 <sup>a</sup> posição, numeral sem –s na 1 <sup>a</sup> : N	27/139	19	0,533	69/128	54	0,571
2 <sup>a</sup> posição, numeral em –s na 1 <sup>a</sup> : Ns	20/88	23	0,602	47/95	50	0,527
Ns0	(0/1)	0	-	(0/2)	0	-
NM	(1/2)	50	-	(1/1)	100	-
NN(N)	(8/27)	30	-	(6/16)	38	-
NNs(N)	(1/3)	33	-	(1/5)	20	-

Para a configuração *ausência de marca formal na 1<sup>a</sup> posição (0\_)*, possuímos apenas o peso relativo da segunda geração (pr. 0,817), que favorece acentuadamente a marcação, num valor muito próximo daquele que é exibido pela comunidade (pr. 0,855). O percentual de marcação que ocorre na geração mais nova (82%) deixa antever que este será um padrão que não revela flutuações geracionais. Todavia, o baixo número deste tipo de configuração na geração mais idosa (apenas duas ocorrências) faz pressupor uma forte tendência de esta faixa etária marcar funcionalmente o PL logo na primeira oportunidade, a categoria DET, originando, de preferência, a estrutura que apresenta marca formal na 1<sup>a</sup> posição (S\_). O reduzido número de ocorrências da configuração 0\_ é também uma realidade no MRJ e na NURC, variedades em que, como vimos, todas as gerações beneficiam a pluralização. Estes aspectos permitem que se estabeleça um paralelo com os pressupostos avançados por Baxter (2009:293) para o PT e HEL-Ba, advogando-se idêntico perfil de inserção e desenvolvimento de marcas PL entre estes dialectos e o PA (Fig. 1): o desenvolvimento da concordância PL inicia-se com a introdução de um PL *singleton* na categoria funcional DET, que serve de âncora para o controlo da pluralização. Por conseguinte, a concordância inicia-se de modo parcial, sendo atribuída a partir da posição DET, ou seja, a partir do elemento pré-nuclear adjacente. Posteriormente, o PL desenvolve-se, ainda com características de concordância parcial, com recurso a morfemas dissociados, que são atribuídos aos elementos mais próximos do *fulcro*. A sequência de atribuição incide sobre os elementos do SDET, primeiramente, e sobre os itens à direita do núcleo, depois. Todavia, o perfil da âncora pré-nuclear é conservado geracionalmente.

Observando também o padrão de evolução dos numerais simples e de mais de uma palavra, com ou sem *-s* final, constata-se que o mesmo mantém um favorecimento moderado da marcação, constante ao longo das gerações. Contudo, é de ter em conta que o factor numeral sem *-s* na 1<sup>a</sup> posição (N\_) tende a propiciar mais a marcação na faixa etária mais velha de Almojarife (pr. 0,748).

No que ao MRJ diz respeito, os percentuais apontados por Scherre (1988:180) parecem indicar que existe uma menor sensibilidade à marca semântica de PL por parte das crianças (37%), já que os adultos apresentam um maior índice de realizações com concordância (58%). Quanto à NURC (Lopes, 2001:216), a menor sensibilidade ocorre

com mais frequência em relação aos numerais simples (POP 4 = pr. 0,53; UNI 4 = pr. 0,38) do que relativamente aos numerais de mais de uma palavra (POP 4 = pr. 0,90; UNI 4 = 100%). Estes aspectos determinam que se discutam também as questões relacionadas com a marcação motivada pelos numerais, o que efectuaremos já de seguida.

#### **4.1.2.4.6. *Marcas precedentes representadas por marcador semântico***

No que concerne à marcação motivada pelos numerais, as análises por nós efectuadas vão ao encontro das de diversos pesquisadores, que têm revelado que estes itens gramaticais em posição antecedente, sejam eles simples ou formados por mais de uma palavra, tendem a favorecer mais a pluralização do que os itens portadores de marca formal de PL. Contudo, as explicações para esta premissa não têm merecido consenso por parte dos teorizadores, na medida em que apontam para duas hipóteses: a funcionalista das Condições de Distintividade, defendida por Kiparsky (1972:195); e a não-funcionalista, advogada por Guy (1981a:180). Ancorando parcialmente na primeira, Scherre (1988:175) entende que os numerais favorecem mais a marcação, sobretudo nas classes mais escolarizadas do MRJ. Assim, nas faixas analfabetas ou semianalfabetas, a marca semântica de PL, por não ser percebida, leva a que se pluralizem morfemicamente os numerais em determinadas realizações. Nesta conformidade, ocorrências deste tipo só tendem a diminuir à medida que aumenta a escolarização, isto é, a partir do momento em que os falantes vão tomando consciência da pluralização semântica inerente aos numerais. A este propósito, refira-se que detectámos nos nossos dados, precisamente nos falantes de baixa escolarização, resquícios de ocorrências que podem indiciar a não percepção da noção semântica de PL. Assim, no exemplo [239] é possível ver que a falante [DULCEM1], que apenas possui a 3ª. Classe (Tabela 8.1), flexiona indevidamente o numeral cardinal, enquanto a informante analfabeta [MMDEUSM3] (Tabela 8.1), deixa de inserir a marcação no mesmo tipo de categoria gramatical, quando tal é exigida:

[257] PA:  
*começô tustão a subi... duzento, duzento, duzento cin-- cinco stens*  
[MMDEUSM3]

No MRJ, o percentual de marcações condicionado pelo numeral na 1ª. posição parece indicar que existe uma menor sensibilidade à marca semântica de PL por parte das crianças (Scherre, 1988:180). Se considerarmos que estas ainda estão a frequentar a escolaridade e que muitos dos adultos informantes do MRJ já a concluíram, então é lícito depreender que o pressuposto da autora justifica a diferença de percentual de marcação entre estas duas faixas etárias, visto os adultos já terem adquirido a noção de pluralização semântica intrínseca ao numeral. Não obstante, os dados de Lopes (2001:215-216) contrariam a justificação avançada por Scherre (1988:180), já que na faixa etária mais idosa da NURC, são precisamente os informantes do grupo universitário aqueles que marcam menos a pluralização em SN's de dois elementos começados por numeral, inibindo-a mesmo (pr. 0,38), enquanto os falantes do grupo menos escolarizado revelam tendência para inserirem muito ligeiramente a marcação (pr. 0,53). Fazendo-se ainda uma análise interdialectal acerca da influência do numeral sem *-s* em 1ª. posição na marcação PL do elemento seguinte nas duas variedades de português de São Tomé, objecto da nossa análise (PA e PT), detecta-se um fenómeno de certa forma idêntico (Tabela 19.9): na primeira, em estágio menos desenvolvido de aquisição da regra da concordância PL, ocorre um favorecimento da marcação (pr. 0,622); na segunda, que revela um estágio mais avançado da aquisição da regra de concordância PL, acontece o oposto, já que a tendência é para inibir ligeiramente a marcação (pr. 0,413). Paralelamente, uma análise intradialectal do PA sobre o mesmo factor (Tabela 19.10) mostra-nos que, na geração mais idosa, em estágio mais primitivo de aquisição da regra de concordância PL, o propiciamento da marcação (pr. 0,748) é superior ao das gerações menos idosas (FE-2 = 0,533; FE-1 = pr. 0,571), situadas num patamar mais avançado da aquisição da mesma regra. A análise intradialectal do PA poderá mesmo encontrar paralelo no MRJ, se levarmos agora em conta a possibilidade de os informantes adultos deste serem menos escolarizados do que as crianças, já que o analfabetismo decai na proporção directa do índice de frequência escolar, que revela tendência a acentuar-se, por norma, nas gerações mais jovens. Em adição, também não podem ser descartadas outras questões de carácter social que pressionam o falante adulto no sentido de realizar a concordância, como acontece, por exemplo, com o mercado de emprego.

Observando agora a hipótese não-funcionalista advogada por Guy (1981a:180), a mesma aponta para a possibilidade de a influência do numeral na marcação espelhar apenas automonitorização e autocorreção, ou seja, a probabilidade de o informante estar a reproduzir ou adaptar as suas realizações ao discurso padrão do entrevistador. Neste aspecto, o autor não é corroborado por P. Andrade (2003:112), que considera o numeral o item mais influente para a marcação na CPL-var, em virtude de ser um morfema semanticamente transparente, isto é, um morfema de conteúdo, logo mais facilmente apreensível em situações de contacto linguístico. Quanto a nós, o pressuposto desta autora apresenta-se mesmo como contra-funcionalista, não justificando o porquê de a geração mais idosa de Almojarife favorecer consideravelmente a marcação, quando ocorre um numeral na 1ª. posição do SN. De facto, existindo captação semântica do PL inerente ao primeiro item, isto é, marcação deste, o que deveria suceder era uma inibição da marcação no segundo elemento, de forma a evitar a redundância (Kiparsky, 1972:195). Só assim se justificaria também o facto de P. Andrade (2003:112) ancorar as suas conclusões no pressuposto de Guy (1981a:3001-302), que atribui a responsabilidade maior pela CPL-var do dialecto de HEL-Ba ao sistema de marcação de número dos ancestrais substratos africanos deste.

Alguns trabalhos sobre a variável independente *marcas precedentes* concluíram também que um SPREP dentro de um SN mais alto com marca semântica de PL tende a ser pouco marcado. Este tipo de SN's, apesar de não indicar o PL de forma tão óbvia como os numerais, não deixa também de encerrar uma carga semântica de pluralidade evidente. Assim, pelo facto de implicarem uma duplicidade de concordância PL, tenderão a exibir inibição no segundo item, evitando-se, desta forma, uma redundância de número (Kiparsky, 1972:195). Este pressuposto, do tipo claramente funcionalista, não foi por nós testado, devido ao facto de não termos analisado este género de dados, pelos motivos referidos em tempo devido (ponto 4.1.2.4.4. do presente trabalho). Contudo, Scherre (1988:227-228) fê-lo detalhadamente na sua quarta análise alternativa, avançando explicações para os resultados em termos de processamento paralelo (Scherre, 1988:208). Assim, acaba por contrariar, na generalidade, as Condições de Distintividade, preconizadas por Kiparsky (1972:195), isto é, que sendo marcada a primeira posição do SN, o número de marcas seguintes é função da forma de processar a



informação. Assim, segundo Scherre (1988:222-223), o facto de o numeral em 1ª. posição (N<sub>1</sub>) favorecer mais a marcação seguinte do que a marcação formal na 1ª. posição (S<sub>1</sub>) não é justificado com a probabilidade de este item gramatical ser um morfema semanticamente transparente, mas sim com a possibilidade de as formas gramaticais semelhantes se agruparem, mostrando tendência para ocorrerem juntas (Haiman, 1983:782). E, caso tal não suceda, a restrição poderá ser originada pela articulação mecânica, salvo se houver algum motivo que justifique a mudança (cf. Schiffrin, 1981).<sup>172</sup>

Na mesma análise, Scherre (1988:173) questiona-se igualmente sobre a possibilidade de ocorrer uma diferença no comportamento dos numerais, caso terminem ou não em *-s*, já que os falantes poderão entender que os primeiros possuem marcação formal de PL e, como tal, aplicarem o Princípio do Processamento Paralelo (Scherre, 1988:208), favorecendo mais a marcação nos itens que lhes são posteriores. As análises por nós efectuadas, tal como as de Scherre (1988:173) e de Lopes (2001:207), debruçando-se sobre os numerais terminados em *-s* e não finalizados em *-s*, ocorrendo na primeira posição dos SN's de dois elementos, comprovaram que a diferença no favorecimento da marcação provocada por ambos não é considerável. Estes resultados fizeram Scherre (1988:236) concluir também que, no que concerne aos itens que carregam informação semântica de pluralidade, o Princípio do Processamento Paralelo (Scherre, 1988:208) apenas actua sobre aqueles que terminam em *-s*. Por outro lado, a restrição mecânica advogada por Schiffrin (1981) também não sustentará a questão da inserção da pluralização, uma vez que as marcas formais de número em 1ª. posição não conduzem a mais marcas. Assim, segundo Scherre (2001:97), o que acontece é um paralelismo linguístico a nível frásico, evidenciando-se que a variação morfológica e sintáctica é controlada pela tendência de preservação de estruturas paralelas (Labov, 1994:550).

A questão dos numerais terminados ou não em *-s* foi por nós também levada em consideração na perspectiva da análise geracional. Em termos gerais (Tabela 19.10), isto é, reflectindo o seu peso na marcação PL a nível comunitário, a diferença no favorecimento à marcação provocada por ambos não se revela considerável (N<sub>1</sub> = pr. 0,607; Ns<sub>1</sub> = pr. 0,556). Contudo, olhando especificamente para a faixa mais idosa de Almojarife, salta à evidência que a diferença de propiciamento se acentua, já que os

numerais não terminados em *-s* apresentam um peso que favorece consideravelmente a marcação (pr. 0,748), ao passo que o favorecimento provocado pelos numerais terminados em *-s* se situa pouco acima do patamar da neutralização (pr. 0,574). De certa forma, o que aqui se verifica é um padrão que segue as linhas que regem o comportamento dos elementos em 1ª. posição com ou sem *-s*, marca formal de PL, e os leva a favorecerem ou não a marcação de forma homogênea em todas as variedades de português observadas neste estudo.

Se associarmos estas verificações às constatadas quer nas análises interdialectal e intradialectal do PA quer na observação do efeito *escolaridade* levantado por Scherre (1988:180) e contrariado tanto por Lopes (2001:216) como pelos nossos dados (veja-se como a geração mais idosa, logo menos escolarizada, apresenta pesos de favorecimento da marcação superiores aos da geração menos idosa, isto é, mais escolarizada), ganha força a probabilidade de, nas faixas analfabetas ou semianalfabetas, a marca semântica de PL não ser percebida (Scherre, 1988:173), considerando-se ainda que o informante reproduza ou adapte realizações ao discurso padrão do entrevistador (Guy, 1981a:180). Desta forma, o falante, ao não perceber a informação semântica PL presente no primeiro elemento da cadeia sintagmática, vai realizar ocorrências com incidência de marcação no segundo item (pr. 0,748 – Tabela 19.10), tal como o faz quando se regista ausência de marcação formal da pluralização no primeiro elemento do SN. Note-se, a propósito deste factor (0\_), como na FE-3 de Almoxarife quase não se realiza este tipo de ocorrências – apenas duas – (Tabela 19.10), confirmando-se que, em situações de contacto linguístico que envolvem línguas de substrato africano, os falantes adultos revelam tendência para inserirem a marcação na primeira oportunidade (Guy, 1981a:301-302). No caso dos numerais finalizados em *-s* e em primeira posição, é de considerar que os falantes os interpretem como itens que carregam marca formal de PL e não apliquem a pluralização no segundo elemento, por influência do princípio funcionalista das Condições de Distintividade (Kiparsky, 1972:195), que advoga que o falante revela tendência para evitar a redundância. Nesta conformidade, não estarão senão a repetir o processo de menos favorecimento à marcação que acontece igualmente quando o item em primeira posição possui marca formal de PL (S\_).

Posteriormente, as gerações mais jovens, devido ao peso de factores extralinguísticos, como a *escolaridade*, o mercado de emprego ou a busca de ascensão social, vão estar mais sujeitas à pressão de cima (superstrato) para baixo, que afecta o PAL1. Desencadeia-se assim a variação em direcção à mudança (Tarallo & Alkmin, 1987), e os falantes, ao contactarem com formas sintagmáticas gramaticalmente mais elaboradas, ampliam as noções sintáctico-semânticas de pluralização, passando a realizar SN's mais extensos e complexos, mas também marcados por incertezas e hesitações na aplicação da regra da concordância. E se é uma realidade que estes falantes exibem alguma expansão da regra de concordância, passando a propiciar mais a marcação em itens na segunda posição quando acontece marcação na primeira ocorrência, também é verdade que a inserção da pluralização não chega nunca ao patamar do favorecimento, quedando-se, em termos comunitários, pela inibição moderada (pr. 0,368). Paralelamente, ao realizar ocorrências com ausência de marcação no primeiro elemento, o falante tem noção de que precisa de aplicar a pluralização, caso ela não seja recuperada pelos elementos contextuais ou pragmáticos, sob pena de se perder a informação de pluralidade. Esta é então inserida no segundo elemento da cadeia sintagmática, em virtude de, na *transmissão irregular*, haver grande probabilidade de se terem perdido, na geração que transmite a L1, os traços não interpretáveis dos elementos funcionais ou não parametrizados que decretariam a marcação no primeiro elemento. Este aspecto foi também constatado por Figueiredo (2009a:50) em análise que abordou a CGEN-var e a CPL-var no SN PL do PA. Buscando verificar quais os constituintes que favorecem (ou não) a aplicação da concordância em Almojarife, o autor deparou-se com achados que contrariam não só o processamento paralelo advogado por Scherre (1988:208) de que a presença de marca no determinante conduz à marcação nos itens posteriores mas sustentam também quer a correlação entre Princípio da Simplicidade e Princípio da Integração (Lucchesi, 2000a) (ponto 1.3.1.1 do presente trabalho) quer o princípio advogado pela FFFH (ponto 2.4.1.1 do presente trabalho), o qual preconiza o apagamento dos traços não interpretáveis dos elementos funcionais ou não parametrizados, adquiridos em fase pós-puberdade (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003). Aspecto igualmente relevante na mesma pesquisa foi o facto de o princípio continuísta

defendido pela FTFA (ponto 2.4.1.1 do presente trabalho) não ter sido confirmado, já que alguns itens morfológicos do PA sugerem expansão da regra de concordância, enquanto outros indicam o oposto. Face a estes aspectos, Figueiredo (2009a:51) concluiu que as gerações mais novas de Almojarife não levam a cabo a expansão da regra de concordância, acabando, isso sim, por confirmar a mudança. No entanto, convém aqui referir que, a análise mais pormenorizada das variáveis sociais levadas a cabo no presente estudo apontam não para uma mudança em curso, mas sim para uma variação estável, como tentaremos sustentar no capítulo 5, em que se analisarão os resultados dos grupos de factores extralinguísticos.

Deste modo, no que respeita aos numerais, ao adquirirem a noção de que eles carregam intrinsecamente a semântica de PL, os falantes passam a inserir menos a marca de número no segundo elemento, num processo normal em que intervêm os aspectos funcionais que determinam que, após marcada a primeira posição do SN, o número de marcas seguintes se torna redundante (Kiparsky, 1972:195). Em certa medida, ocorre aqui um processo na inversa àquele que afecta os elementos que carregam marca formal de PL, mas que é perfeitamente entendível se levarmos em linha de conta que ambos os procedimentos se direccionam no sentido de cristalizar um padrão de concordância, ao que tudo indica, afectado pela interferência das L1's ancestrais. Por conseguinte, estabelece-se um nivelamento entre os dois padrões de marcação (semântico e estrutural) na geração mais jovem, já que esta se revela incapaz de repor os traços não interpretáveis dos elementos funcionais ou não parametrizados perdidos pelas gerações adultas que estabeleceram os primitivos contactos linguísticos.

No fundo, o que aqui se verifica não é senão o espelho do fenómeno também apontado por Lucchesi (2000a) e Figueiredo (2009a) em pesquisas efectuadas sobre a CGEN-var em HEL-Ba e Almojarife, respectivamente, e que levou ambos a concluir que, na aquisição/fixação das regras de concordância de género, também não actua o princípio mental da língua em agrupar formas semelhantes – o processamento paralelo (Scherre, 1988:208) – que conduz a uma economia de marcas levarem a marcas e zeros conduzirem a zeros. Como tal, não será a marca de género ou número no elemento anterior da cadeia linear que conduz a mais marcas posteriores. Assim, no caso do género, conforme ambos os autores constataram, será a marcação mais explícita de

gênero no núcleo que favorece a marcação na cadeia linear. No caso do número, e tendo em conta que a gramática da L1 é o ponto de partida para a ASL (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), ocorre pluralização de acordo o sistema das ancestrais L1's africanas (Guy, 1981a:301-302), em que a marcação é inserida em apenas um dos elementos da cadeia sintagmática, de preferência no primeiro (Kiparsky, 1972:195), ou seja, na categoria DET (Baxter, 2009), mantendo-se os outros itens do SN inalterados, visto que a informação de PL se torna redundante, caso seja inserida nestes (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75). Por outro lado, como a ancestral ASL teve como intervenientes os falantes que haviam ultrapassado o período crítico para a aquisição, os traços de PL das categorias funcionais também não são adquiridos (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003). Posteriormente, os DLP's deficitários vão originar a instanciação em cadeia dos novos/errados parâmetros, que originam a variação e impossibilitam o reajustamento paramétrico, fazendo com que os mesmos sejam tomados como modelos para uma nova reinstanciação paramétrica na L1 nativizada (Gonçalves, 2004:235). Em consequência de tal, encontram-se satisfeitos os requisitos para se consumir uma variação estável ou a mudança (Schiffrin, 1981), que ancora a marca de PL no primeiro item da cadeia sintagmática e restringe a marcação nos outros elementos desta.

A questão da influência exercida pelos numerais na marcação do elemento seguinte será por nós retomada posteriormente quer no próximo ponto quer quando nos debruçarmos sobre a variável independente *contexto fonológico posterior* (capítulo 5), nomeadamente no que concerne ao cruzamento desta com o grupo de factores *saliência fónica* (ponto 5.2.3.10).

Resta então referir que, relativamente à hipótese 3, levantada para a variável independente *marcas precedentes* (ponto 3.6.2. do presente trabalho), confirma-se que apenas a marca de PL formal inibe a marcação nos itens seguintes, já que os numerais revelam tendência para favorecerem a inserção de marcas, tendência essa que decresce geracionalmente a fim de estabilizar o padrão da regra ao encontro do padrão estrutural de marcação. Nesta conformidade, também salta à evidência que o padrão de marcação não é condicionado pelo Princípio do Processamento Paralelo (Scherre, 1988:208).

Paralelamente, é fundamental também referir que a amalgamação de factores praticada por Scherre (1988:177), com vista a buscar uma confirmação para o pressuposto de que marcas levam a marcas e zeros conduzem a zero, não foi sustentada, no nosso trabalho, pela metodologia científica dos testes do qui-quadrado, que acabaram revelando a vantagem de a análise das *marcas precedentes* ser efectuada numa perspectiva ontogénica (cf. Guy, 1981a:178), e não filogénica. Por fim, é ainda fundamental chamar a atenção para o facto de a análise dos resultados da variável *marcas precedentes* ter levantado a hipótese de o português adquirido em situação de contacto por falantes de um CP manifestar os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a).

#### **4.1.2.5. Variável independente grau de concordância de número no SN**

A primeira rodada VARBRUL mostrou que, neste grupo de factores, o factor 4 apresenta apenas 28 ocorrências (Tabela 12.10). Face a tal, poderíamos testar a amalgamação dos factores 3 e 4 num único factor e conseguir um grupo de factores que se aproximasse mais do sistema binário, mas optámos por levar à análise as 28 ocorrências do factor *mais de três elementos flexionáveis no SN*, a fim de elaborarmos uma melhor visão sobre as configurações sintagmáticas das diferentes faixas etárias. Embora sendo certo que a aproximação ao sistema binário fortalece as hipóteses linguísticas, não deve ser descartada a possibilidade de este modelo, em determinadas circunstâncias, não encaixar na perfeição nos dados linguísticos que se coligiram. Assim sendo, a Lei da Parcimónia não deve pura e simplesmente fazer valer de modo irracional o seu efeito de “navalha”, sob pena de se inviabilizar a análise de determinadas generalizações preciosas (Guy & Zilles, 2007:153), que acreditamos ocorrerem na constituição dos factores do grupo de factores que aqui se analisa. Portanto, conservaram-se as distinções originais para efeitos de análises VARBRUL, que apresentaram os valores seguintes:

**Tabela 20.1.** Efeito do grau de concordância de número no SN na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504	<i>Log-likelihood:</i> -981,009	<i>Significância:</i> 0,011		
<b>Factores seleccionados</b>	<b>Nr. de itens</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
SN's com 1 elemento flexionável no SN	204/729	28	31	0,268
SN's com 2 elementos flexionáveis no SN	865/1.352	64	58	0,628
SN's com 3 elementos flexionáveis no SN	125/231	54	10	0,563
SN's com mais de 3 elementos flexionáveis no SN	(8/28)	29	1	0,274
<b>Totais:</b>	<b>1.202/2.340</b>	<b>51</b>	<b>100</b>	<b>-</b>

O factor *SN's com um elemento flexionável no SN* apresenta-se como o mais inibidor à inserção de marcas (pr. 0,268), o que não surpreende, se tomarmos em conta que este tipo de configuração reflecte, na sua quase totalidade, SN's em que o primeiro elemento é composto por um numeral não flexionável em número, na primeira posição, e um nome, depois. Como vimos no ponto anterior, o que aqui se espelha é o padrão de comportamento que nivela o sistema de marcação PL no SN do PA, de acordo com as linhas que regem o comportamento dos elementos em 1<sup>a</sup>. posição com ou sem *-s*, marca formal de PL. De facto, à medida que os falantes almoxarifanos vão adquirindo a noção de que o numeral encerra em si a marca semântica de PL, em termos gerais, e de acordo com o princípio funcionalista das Condições de Distintividade (Kiparsky, 1972:195), vão também inibindo a marcação no elemento seguinte a esta categoria gramatical. Assim, caso ocorram marcas (semânticas ou estruturais) no primeiro elemento do SN, a tendência será de as não inserir nos elementos posteriores, ao passo que, não estando o primeiro elemento marcado, a tendência será para propiciar a marcação nos itens posteriores (Tabela 19.10). Estes aspectos são confirmados pelos outros factores, já que os *SN's com dois elementos flexionáveis no SN*, configurando na sua grande maioria SN's de estrutura reduzida do tipo DET+N, vão apresentar um incremento na marcação (pr. 0,628), que ocorrerá sistematicamente no primeiro elemento, podendo o segundo item ser também marcado, mas de modo esporádico. Tal foi constatado através do levantamento e observação pormenorizada de todos estes sintagmas, que revelaram as seguintes realizações:

1.  $DET_{[+PL]} N_{[+PL]} = 80$  ocorrências;
2.  $DET_{[+PL]} N_{[-PL]} = 367$  ocorrências;
3.  $DET_{[-PL]} N_{[+PL]} = 28$  ocorrências.

As configurações permitem também vislumbrar como é altamente variável a concordância de número neste tipo de SN's, situando-se o valor da concordância plena bastante abaixo do da concordância parcial.

Quanto ao factor *SN's com três elementos flexionáveis no SN*, por norma configurando a estrutura do tipo DET+N+ADJ, reforça os pressupostos que temos vindo a apresentar e confirma que a distensão e complexificação do SN retrai a inserção de marcas (pr. 0,563), a qual afecta, essencialmente, os itens em posição pós-nuclear. Ainda assim, o padrão geral da marcação continua a manter-se no ponto da neutralidade, comprovando alguma aquisição da regra da concordância PL, certamente por influência de questões do tipo socioeconómico, já que a referida aquisição se dá com mais pendor entre as gerações mais novas da comunidade, portanto também mais escolarizadas. Estes aspectos serão debatidos já de seguida, a propósito dos valores exibidos na Tabela 20.2, que apresenta o total geral dos itens que integram os SN's de um elemento, dois elementos, três elementos e mais de três elementos flexionáveis, tanto para a comunidade como para as diferentes faixas etárias de Almojarife. Porém, note-se ainda como o peso altamente inibidor revelado pelo factor *SN's com mais de três elementos flexionáveis no SN* (pr. 0,274) comprova a dificuldade na aquisição plena da regra da concordância PL, à medida que os SN's passam a requerer mais elementos flexionáveis na sua cadeia.

A variável independente *grau de concordância de número no SN* permite que se observe também o total de SN's que contêm, incorporados em si, um, dois, três ou mais de três itens flexionados. Para tanto, basta que se divida o total de itens de cada factor pelo número de itens que determina o tipo de factor. Exemplificando: como o factor *SN's com dois elementos flexionáveis no SN* regista um total de 1352 itens flexionáveis para a comunidade, se dividirmos este total por dois (o número de itens flexionáveis existentes em cada um deste tipo de SN's) obtemos o resultado final de 676, que representa o número de SN's com dois itens flexionáveis produzidos pelos falantes das



três faixas etárias de Almojarife. Procedendo-se de igual modo para todos os factores, obtém-se o total de SN's registados para cada um deles, não só a nível comunitário mas também no que concerne a cada uma das faixas etárias, isoladamente.

As rodadas gerais para o grupo de factores *idade*, relativas às faixas etárias FE-1 e FE-2, rejeitaram o grupo de factores *grau de concordância de número no SN*, revelando que o mesmo não exerce peso significativo na inserção de marcas PL nestas duas faixas etárias (Tabela 20.2). Quanto à faixa etária mais idosa de Almojarife, esta variável independente apresenta um padrão de marcação em tudo semelhante ao da comunidade. Confirmam-se, então, dois pressupostos: (i) a grande maioria dos SN's produzidos pela FE-3 é de estrutura mínima, composta apenas por um ou dois itens flexionáveis; (ii) o padrão de marcação PL no SN, não sofre, em termos gerais, um desenvolvimento geracional acentuado no sentido da aquisição da regra de concordância PL.

Observando-se pormenorizadamente as produções da FE-3, verifica-se que regista 243 itens flexionáveis em estruturas sintagmáticas que apenas incluem um elemento flexionável, traduzindo-se igualmente em 243 SN's, na sua grande maioria antecedidos de numeral. No ponto 4.1.2.4.6 deste trabalho (*marcas precedentes* representadas por marcador semântico) avançou-se a possibilidade de os falantes mais idosos de Almojarife não percepcionarem a marca semântica de PL presente no numeral que antecede o elemento flexionável, o que os leva a realizarem ocorrências com incidência de marcação no segundo item (pr. 0,748 – Tabela 19.10). Contudo, a Tabela 20.2 indicamos que o percentual de marcação neste tipo de SN's é de apenas 14% (34 itens marcados e 209 não marcados), com um peso bastante inibidor de inserção de marcas (pr. 0,220), o que parece contrariar os resultados e conclusões que retirámos aquando da análise das marcas precedentes representadas por marcador semântico. A este propósito, convém então clarificar que o reduzido número de configurações desenhadas pela FE-3 determinou a ocorrência de um grupo de factores ternário para observação do efeito das marcas precedentes na marcação PL (um com marca formal de PL e dois com marca semântica de PL) composto pelos seguintes factores: (i) 2<sup>a</sup>. posição, marca formal na 1<sup>a</sup>. posição (S\_\_); (ii) 2<sup>a</sup>. posição, numeral sem –s na 1<sup>a</sup>. (N\_\_); (iii) 2<sup>a</sup>. posição, numeral em –s na 1<sup>a</sup>. (Ns\_\_). Deste modo, é em relação às outras ocorrências em 2<sup>a</sup>. posição da FE-3, essencialmente a 2<sup>a</sup>. posição com marca formal na 1<sup>a</sup>. posição (S\_\_), que os numerais

vão determinar um peso mais favorecedor da marcação no elemento imediatamente seguinte, apesar de esse percentual de marcação ser relativamente baixo.

Com dois elementos flexionáveis surgem os segundos SN's mais produtivos desta geração, compostos por um total de 278 itens, dos quais 55% receberam marcação de número, com um peso moderadamente favorecedor de inserção de marcas (pr. 0,701). Temos então 139 SN's, na sua maioria reflectindo a estrutura DET+N, confirmando-se que a inserção de marcas no primeiro elemento ocorrerá desde o início da aquisição do PtgL2. A partir daqui, e em consequência do aumento de itens passíveis de flexão a inserir na cadeia sintagmática, detecta-se uma queda na realização de SN's extensos e pluralização dos seus elementos. Assim, a faixa etária mais idosa de Almojarife produziu apenas 42 itens em SN's que incluem três elementos passíveis de flexão. O percentual de inserção de marcas em 14 SN's quedou-se pelos 45%, reduzindo o peso de marcação PL para o plano da neutralidade (pr. 0,530) e comprovando a baixa aquisição da regra de concordância PL. Por fim, registam-se apenas 5 itens flexionáveis, em cadeias sintagmáticas com mais de 3 itens flexionáveis. Neste caso, estamos perante apenas um SN, composto por 5 itens passíveis de flexão, mas com marca de pluralização somente no primeiro item, como pudemos confirmar.

O que se observa nas gerações seguintes, com recurso aos percentuais de marcação, é que o número de pluralizações em SN's com apenas um elemento flexionável se mantém estabilizado, na FE-2, reflectindo o padrão já evidenciado pela FE-3 (22% de itens pluralizados, em 253 SN's). Contudo, a FE-1 leva já a cabo uma expansão da regra de concordância PL para o plano da neutralidade, inserindo marcas em 49% dos elementos dos 233 SN's que a requerem. Para além de equilibrar os totais de itens marcados (115 itens) e não marcados (118 itens), esta geração revela que percepção a marca semântica de PL. Por seu lado, os SN's nos quais estão inseridos dois itens passíveis de marcação PL, e que reflectem, maioritariamente, a estrutura DET+N, acentuam o seu percentual de marcação, com a FE-2 inserindo marcas em 63% dos elementos que compõem os 209 SN's produzidos, enquanto a FE-1 leva a cabo 68% de pluralizações nos itens integrados em 328 SN's. Estes percentuais, apesar de serem superiores aos do factor *SN's com um elemento flexionável no SN*, não podem ser vistos como um indicativo da aquisição da regra de concordância PL, evidenciando apenas que a

inserção da marca de PL ocorrerá no elemento imediatamente pré-nuclear, isto é na categoria DET, reflectindo o sistema de concordância PL dos ancestrais substratos do PA, ou seja, idiomas do grupo níger-congo atlântico (cf. Baxter, 2009:292).

**Tabela 20.2.** Efeito do grau de concordância de número no SN no percentual de marcação PL dos itens do SN: comunidade e faixas etárias.

		<i>Input desta rodada: 0,504 Log-likelihood: -981,009; Significância: 0,011</i>				<i>Input desta rodada: 0,429 Log-likelihood: -192,116 Significância: 0,000</i>			
		<b>Comunidade</b>				<b>FE-3 (+60 anos)</b>			
<b>Factores seleccionados</b>	<b>Nr. de itens</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. de SN's</b>	<b>Nr. de itens</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. de SN's</b>	
SN's com 1 elemento flexionável no SN	204/729	28	0,268	279	34/243	14	0,220	243	
SN's com 2 elementos flexionáveis no SN	865/1.352	64	0,628	676	152/278	55	0,701	139	
SN's com 3 elementos flexionáveis no SN	125/231	54	0,563	77	19/42	45	0,530	14	
SN's com mais de 3 elementos flexionáveis no SN	(8/28)	29	0,274	6	(1/5)	20	-	1	
<b>Totais:</b>	<b>1.202/2.340</b>	<b>51</b>	<b>-</b>	<b>1.488</b>	<b>206/568</b>	<b>36</b>	<b>-</b>	<b>397</b>	
		<i>Input desta rodada: 0,562 Log-likelihood: -366,835 Significância: 0,002</i>				<i>Input desta rodada: 0,723 Log-likelihood: -458,671 Significância: 0,042</i>			
		<b>FE-2 (41 – 60 anos)</b>				<b>FE-1 (20 – 40 anos)</b>			
<b>Factores seleccionados</b>	<b>Nr. de itens</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. de SN's</b>	<b>Nr. de itens</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. de SN's</b>	
SN's com 1 elemento flexionável no SN	55/253	22	-	253	115/233	49	-	233	
SN's com 2 elementos flexionáveis no SN	264/418	63	-	209	449/656	68	-	328	
SN's com 3 elementos flexionáveis no SN	53/108	49	-	36	53/81	65	-	27	
SN's com mais de 3 elementos flexionáveis no SN	(2/9)	22	-	2	(5/14)	36	-	3	
<b>Totais:</b>	<b>374/788</b>	<b>47</b>	<b>-</b>	<b>500</b>	<b>622/984</b>	<b>63</b>	<b>-</b>	<b>591</b>	

Obs.: FE-2 e FE-1 rejeitadas pela ferramenta VARBRUL.

No que concerne aos SN's com três itens flexionáveis, verifica-se que, na FE-2, foram produzidos 36 SN's que apresentam 49% de itens marcados com PL. O percentual de marcação decai ligeiramente neste factor, daqui se inferindo que a responsabilidade da estagnação na aquisição da regra de concordância será dos elementos pós-nucleares, que marcam presença considerável em SN's de três elementos flexionáveis. Quanto à faixa etária mais nova, produziu 81 itens inseridos em 27 SN's com três elementos passíveis de flexão PL, tendo o percentual de marcação atingido 65%. Este aspecto denota que a aquisição da regra de concordância PL não acontece de forma acentuada entre os elementos da faixa mais nova de Almojarife, que continuam a produzir considerável CPL-var.

Por fim, note-se que a realização de itens inseridos em cadeias sintagmáticas com mais de três elementos flexionáveis é praticamente insignificante. De facto, a FE-2 apenas produziu 2 SN's, para um total de 9 elementos deste tipo. Relativamente a estes SN's, confirmámos que um possui quatro elementos flexionáveis e o outro cinco. Os dois itens com inserção de pluralização apresentados na Tabela 20.2 repartem-se equitativamente por ambos os SN's e correspondem ao primeiro elemento flexionado, evidenciando a tendência funcionalista de não se repetir a marcação, em virtude de ela se revelar redundante (Kiparsky, 1972:195). Relativamente à FE-1, realizou 3 SN's com mais de três itens flexionáveis: um SN com quatro elementos passíveis de flexão, três destes com marcação PL, mas o primeiro apresentando inibição (*tive uma otras minhas preocupações* [CASTEH1]); dois SN's com cinco elementos flexionáveis cada, sendo pluralizado apenas o primeiro elemento em ambas as cadeias sintagmáticas.

Os resultados apresentados pelas diferentes faixas etárias fornecem-nos, então, o desenho da aquisição da marca PL e seu desenvolvimento na comunidade, permitindo concluir que, em relação à hipótese 10 levantada por nós para a variável independente *grau de concordância de número no SN*, a geração mais idosa de Almojarife produz, de facto, SN's maioritariamente simplificados (um ou dois elementos flexionáveis na cadeia sintagmática) e com baixa percentagem de marcação PL. Lembre-se que esta geração teve também contacto com o PtgL2 falado por africanos “contratados”, o qual poderá, de certa forma, ter contribuído para ajudar a estabilizar o actual padrão de marcação do PA. Já a marcação PL em itens passíveis de flexão inseridos em SN's com

apenas um elemento flexionável mantém-se estável na FE-2, mas aumenta na FE-1, denotando que esta já percebe a marca semântica de PL. Por seu lado, a realização de SN's com dois, três ou mais de três itens flexionáveis é exponenciada depois, revelando que as faixas etárias mais novas de Almojarife não só produzem cadeias sintagmáticas de maior complexificação estrutural como também aplicam mais a regra da concordância em número. Confirma-se uma ligeira aquisição geracional desta, com o incremento da percentagem de marcação PL no PA podendo ser atribuída à influência de diversos factores sociolinguísticos e socioeconómicos, que exerceram pressão sobre as gerações mais novas de falantes do PA, sobretudo na segunda metade do século XX, levando a um maior nivelamento linguístico a nível da fala da comunidade. Contudo, é também de ter em consideração que as gerações mais novas de Almojarife continuam a produzir CPL-var no SN, comprovando dois aspectos: (i) a marcação PL no SN continua a seguir o padrão já revelado pelos antecessores dos actuais membros da comunidade, isto é, inserindo marcas PL preferencialmente apenas no primeiro elemento da cadeia sintagmática, com especial incidência no item imediatamente pré-nuclear, ou seja, na categoria DET; (ii) em termos comunitários, a fala de Almojarife fornece evidências para se concluir que, de acordo com hierarquia da “processabilidade” (Plag, 2008a, 2008b), não foram adquiridas plenamente as regras da flexão contextual, instanciadas a partir do nível 4 do desenvolvimento da interlíngua (Tabela 3.1). Desta forma, o PA foi-se estabilizando geracionalmente entre os patamares 3 e 4 do desenvolvimento sintáctico proposto pela referida hierarquia, fase em que se deverá manter em virtude de não se encontrarem reunidas condições socioeconómicas em São Tomé que façam prever um continuado desenvolvimento da aquisição da regra da concordância PL, em direcção ao PE. Estes aspectos serão mais pormenorizadamente debatidos no capítulo 5, quando fizermos a análise dos resultados fornecidos pela ferramenta VARBRUL para as variáveis extralinguísticas.

Em adição, e de um modo geral, os resultados referentes à variável independente *grau de concordância de número no SN* vêm também reforçar o que temos vindo a afirmar ao longo das nossas análises: (i) a nível dos itens que compõem a cadeia dos SN's do PA e contribuem para a sua CPL-var, não se entrevê qualquer intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208) a motivar esta, com marcas conduzindo a

mais marcas e, conseqüentemente, zeros levando a zeros; (ii) constata-se que ocorre tendência para a marcação PL ser inserida na primeira oportunidade, por norma no DET antes do nome núcleo do SN, por influência do padrão de pluralização dos substratos africanos (cf. Guy, 1981a:301-302); (iii) os falantes evitam a marcação redundante, inibindo a redundância da pluralização nos elementos posteriores, de acordo com o advogado pelo princípio funcionalista das Condições de Distintividade (Kiparsky, 1972:195); (iv) por influência de variáveis extralinguísticas, ocorre ligeira aquisição geracional das regras de concordância PL, o que leva as faixas etárias mais novas de Almojarife a inserirem mais marcas de PL nos itens do SN. Esta última constatação vem contrariar fortemente o pressuposto de o apagamento da marca formal de PL poder ser justificado em termos de fenómenos inerentes à própria língua portuguesa, como a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a), já que esta determina uma eliminação diacrónica do *-s* morfémico PL, ao passo que, o que acontece realmente no PA (e variedades africanas e brasileiras de português) é um fenómeno completamente oposto, ou seja, uma inserção geracional do referido morfema.

#### **4.1.2.6. Variável independente *posição do item na cadeia do SN (Posição linear)***

Conforme se pode verificar na Tabela 12.7, após gerarmos a primeira célula VARBRUL, constatou-se que o factor *quinta ou outra posição na cadeia do SN* apresenta um total de ocorrências que não permite a sua incorporação nos dados a serem analisados (21 realizações). Contudo, como os factores 4 e 5 compartilham traços,<sup>173</sup> optou-se por testar a recodificação do grupo de factores, efectuando a amalgamação dos referidos factores num só (*quarta ou outra posição na cadeia do SN*) e estabelecendo um novo grupo de factores, de acordo com a Tabela 21.1.

Para se confirmar se o novo grupo de factores representa uma distinção linguística que pode ser incorporada à análise, levámos a cabo o teste do qui-quadrado, medindo-se o nível de significância do novo grupo de factores. Efectuou-se então uma rodada simples (*one-level analysis*), a fim de se obter o valor para o novo *log-likelihood*, compará-lo com o da rodada feita com os factores da Tabela 12.7, efectuar os cálculos para achar o qui-quadrado e constatar se os novos factores poderiam ser incorporados à análise.

**Tabela 21.1.** Grupo de factores *posição do item na cadeia do SN* e respectivos factores, para estudo da configuração do SN PL do PA: grupo de factores a testar para incorporação na análise de dados.

<b>Factores a testar</b>	<b>Exemplo</b>	<b>Nr. de ocorrências</b>	<b>%</b>
1 Primeira posição na cadeia do SN	<b>TRABALHOS</b> privado [CASTEH1]	754	90
2 Segunda posição na cadeia do SN	Umas <b>DORES</b> [DULCEM1]	1.314	34
3 Terceira posição na cadeia do SN	O <sub>s</sub> meus <b>FILHOS</b> [MANOH3]	168	29
4 Quarta ou outra posição na cadeia do SN	sete ou oito <b>MESES</b> [CASTEH1]	34/104	33
<b>Totais:</b>		1.202/2.340	51

Feito o teste do qui-quadrado a partir da diferença registada entre os dois *log-likelihoods* (Tabela 21.2), e tendo em conta o valor fornecido pela tabela de qui-quadrado, concluiu-se que o resultado achado para o novo grupo de factores é bem inferior ao inicial. Como tal, as diferenças entre as posições lineares não são estatisticamente significativas, logo, representam uma distinção linguística que deve ser mantida na análise. Assim sendo, como a distinção inicial não era significativa, foi abandonada e substituída pela nova combinação de factores, ou seja, a que engloba 4 factores (Tabela 21.1).

**Tabela 21.2.** Teste de significância estatística dentro do grupo de factores *posição do item na cadeia do SN*.

<b>Factores a testar [5 Factores]</b>	<b>Nr. total</b>	<b>% no grupo</b>
Primeira posição na cadeia do SN	678/754	33
Segunda posição na cadeia do SN	441/1.314	56
Terceira posição na cadeia do SN	49/168	7
Quarta posição na cadeia do SN	29/83	3
Quinta ou outra posição na cadeia do SN	5/21	1
<i>Log-likelihood: -936,081</i>	<b>Totais:</b> 1.202/2.340	100

<b>Factores a testar [4 Factores]</b>	<b>Nr. total</b>	<b>% no grupo</b>
Primeira posição na cadeia do SN	678/754	32
Segunda posição na cadeia do SN	441/1.314	56
Terceira posição na cadeia do SN	49/168	7
Quarta ou outra posição na cadeia do SN	34/104	5
<i>Log-likelihood: -936,181</i>	<b>Totais:</b> 1.202/2.340	100

Após constituição definitiva dos factores a incorporar na variável independente *posição linear*, procedeu-se a uma rodada VARBRUL para apuramentos dos pesos dos referidos factores na marcação PL no SN do PA (Tabela 21.3).

Os resultados mostram que o PA não foge à norma dos outros dialectos, já que a 1ª. posição é a única que marca positivamente o PL, e de modo categórico (pr. 0,878). Todas as outras posições apresentam-se como desfavorecedoras da marcação, sendo drástica a redução registada da primeira para a segunda posição. A escala de desfavorecimento acentua-se progressivamente (2ª. posição = pr. 0,312; 3ª. posição = pr. 0,161), mas, ao contrário do que acontece no MRJ e a exemplo do que sucede na NURC, o factor que representa a posição mais distante do SN, isto é, a quarta posição (pr. 0,169), não revela maior inibição dos que os dois que o antecedem, mantendo-se estável em relação à marcação na 3ª. posição. Estes aspectos não surpreendem, se levarmos em consideração que a grande maioria dos SN's produzidos pela comunidade de Almojarife é de estrutura reduzida (DET+N) e que a marcação tende a incidir apenas no elemento pré-nuclear, consoante sistema de pluralização das línguas africanas (cf. Guy, 1981: 301-302) do grupo níger-congo atlântico (Baxter, 2009:292), como temos vindo a salientar, mantendo-se o nome inalterado, em virtude de a marcação neste se revelar redundante (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75).

**Tabela 21.3.** Efeito da *posição linear* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,513	<i>Log-likelihood:</i> -991,577		<i>Significância:</i> 0,038	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Primeira posição na cadeia do SN	678/754	90	32	0,878
Segunda posição na cadeia do SN	441/1.314	34	56	0,312
Terceira posição na cadeia do SN	49/168	29	7	0,161
Quarta ou outra posição na cadeia do SN	34/104	33	5	0,169
<b>Totais:</b>	1.202/2.340	51	100	-

Seguidamente, apresentam-se os resultados do PA e da NURC (Lopes, 2001:166), a fim de se estabelecerem comparações entre os padrões de marcação PL determinados pela variável independente *posição linear* nestas duas variedades de português:



**Tabela 21.4.** *Posição linear*: contribuição dos factores individuais para a marcação PL do elemento analisado: 2 variedades de português.

Factores	PA		NURC	
	Pr.	%	Pr.	%
Primeira posição na cadeia do SN	0,878	90	0,84	99
Segunda posição na cadeia do SN	0,312	34	0,23	69
Terceira posição na cadeia do SN	0,161	29	0,25	62
Quarta posição na cadeia do SN	0,169	33	0,23	59
Quinta ou outra posição na cadeia do SN			0,44	75

O padrão de marcação é idêntico nas duas variedades de português, confirmando-se totalmente os achados de outros trabalhos sobre o favorecimento da marcação determinado pela primeira posição da cadeia sintagmática. Mais uma vez, verifica-se também que o português adquirido em situação de contacto por falantes de um CP manifesta os mesmos padrões de variação do português adquirido em situação na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a).

Para continuarmos a perceber como é que os pesos das diferentes posições sintagmáticas actuam na marcação PL, procederemos, no próximo ponto, ao cruzamento entre a variável independente *posição linear* e outros grupos de factores, nomeadamente as variáveis *classe gramatical* e *posição em relação ao núcleo do SN*.

#### 4.1.2.6.1. Relação entre *posição linear* e *classe gramatical* na marcação PL dos itens do SN

O estudo da relação entre as variáveis independentes *posição linear* e *classe gramatical* na marcação PL dos itens do SN foi feito por Scherre (1988:151) e Lopes (2001:170), não só com o intuito de observar quais os itens gramaticais que se situam nas diversas posições sintagmáticas mas também com o propósito de avaliar se se estabelece uma relação directa entre os dois grupos de factores nos moldes defendidos por Guy (1981b:179), isto é, entre determinantes e a primeira posição, nomes e a segunda posição e adjectivos e a terceira posição.

A Tabela 21.5 permite visualizar o total de ocorrências para cada item gramatical nas distintas posições da cadeia sintagmática do PA, bem como a sua influência na marcação PL:

**Tabela 21.5.** Relação entre *posição linear* e *classe gramatical*: distribuição das classes gramaticais pelas diversas posições do SN.

<i>Input desta rodada: 0,552</i>			<i>Log-likelihood: -985,404</i>		<i>Significância: 0,007</i>	
<i>Classe gramatical</i>	<b>1ª. posição</b>			<b>2ª. posição</b>		
	<b>Nr. total</b>	<b>% PL</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>% PL</b>	<b>Peso relativo</b>
Substantivo	(24/28)	86	-	384/1.199	32	0,311
Categoria substantivada	(1/1)	100	-	15/49	31	0,218
Adjectivo 1	(7/7)	100	-	(7/23)	30	-
Adjectivo 2	-	-	-	-	-	-
Adjectivo diminutivo	-	-	-	-	-	-
Artigo definido	250/278	90	0,629	(10/10)	100	-
Artigo indefinido	63/80	79	0,578	-	-	-
Numeral	(0/3)	0	-	(3/3)	100	-
Possessivo	(14/16)	88	-	(14/18)	78	-
Demonstrativo	170/178	96	0,969	(3/6)	50	-
Indefinido	59/60	98	0,982	(3/4)	75	-
Quantificador	90/103	87	0,883	(2/2)	100	-
<b>Totais:</b>	<b>678/753</b>	<b>90</b>	<b>-</b>	<b>441/1.314</b>	<b>34</b>	<b>-</b>
<i>Classe gramatical</i>	<b>3ª. posição</b>			<b>4ª. ou outra posição</b>		
	<b>Nr. total</b>	<b>% PL</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>% PL</b>	<b>Peso relativo</b>
Substantivo	40/114	35	0,263	32/87	37	0,241
Categoria substantivada	(1/3)	33	-	(0/1)	0	-
Adjectivo 1	(4/24)	17	-	(1/11)	9	-
Adjectivo 2	(0/8)	0	-	-	-	-
Adjectivo diminutivo	(0/4)	0	-	(0/4)	0	-
Artigo definido	-	-	-	-	-	-
Artigo indefinido	-	-	-	-	-	-
Numeral	-	-	-	-	-	-
Possessivo	(2/2)	100	-	(1/1)	100	-
Demonstrativo	(1/2)	50	-	-	-	-
Indefinido	-	-	-	-	-	-
Quantificador	(1/11)	9	-	-	-	-
<b>Totais:</b>	<b>49/168</b>	<b>29</b>	<b>-</b>	<b>34/104</b>	<b>33</b>	<b>-</b>

Os resultados mostram o porquê de Scherre (1988:152) ter discordado de Guy (1981b:179) no que concerne à relação directa estabelecida entre item gramatical e posição linear. De facto, se olharmos para a distribuição dos substantivos e adjetivos, os primeiros, a exemplo do que sucede no MRJ, apresentam percentuais de ocorrências nas terceira (8% do total de ocorrências, correspondendo a 114 realizações) e quarta posições (6%, do total das realizações, respeitantes a 87 casos) da cadeia sintagmática, que não podem ser desconsiderados. No caso específico dos adjetivos, as suas realizações são sempre inferiores às dos substantivos, seja em que posição for. Contudo, esta inferior quantidade de realizações é algo uniforme em todas as posições, com especial incidência entre as segunda (23 ocorrências) e terceira posições (24 realizações), o que não permitirá dizer que estes elementos gramaticais ocorrem preferencialmente na terceira posição.

Scherre (1988:155) chama ainda a atenção para o facto de Guy (1981a:168) apenas ter trabalhado com a variável independente *posição linear* e não ter estudado a influência da classe gramatical no desenho da concordância PL do SN. Como tal, a autora decidiu observar esta influência nos dados dos adultos e crianças, levando a efeito análises quer em SN's de três elementos (supostamente compostos por DET+N+ADJ) quer em SN's de mais de três elementos. Scherre (1988:157) verificou então que os efeitos dos grupos de factores *classe gramatical* e *posição linear* não estabelecem paralelismos, aspecto que também parece ser confirmado nos nossos dados, já que, na terceira posição, por exemplo, não há sequer sintonia aproximada entre o total dos adjetivos (24 ocorrências) e o total de dados gramaticais (168 realizações).

Face a estas constatações, Scherre (1988:156) avançou três suposições, que não se confirmam totalmente nos nossos dados: “1) *determinantes na segunda posição são até mais marcados do que na primeira; 2) os substantivos têm mais chances de serem marcados na terceira posição do que na segunda; 3) os adjetivos, inversamente aos substantivos, se apresentam mais marcados na segunda do que na terceira posição*”. Ora, como se pode verificar, as duas primeiras suposições não encontram eco no PA, já que, nos nossos dados, por norma, os determinantes são mais marcados na primeira posição do que na segunda. Há, de facto, a excepção dos determinantes artigos definidos (90% de marcação na primeira posição, por oposição a 100% de marcação no segundo

lugar), mas, como se pode constatar, o número de realizações na segunda posição, apenas dez, é bastante diminuto para se poder avançar com qualquer conclusão. Quanto ao facto de os substantivos serem mais marcados na terceira posição do que na segunda, é verdade que os percentuais apontam para um ligeiro propiciamento da terceira (35%) sobre a segunda (32%), mas os pesos relativos indicam exactamente o contrário (3<sup>a</sup>. = pr. 0,263; 2<sup>a</sup>. = pr. 0,311). Relativamente aos adjectivos, é realmente um dado adquirido que o seu percentual de marcação é mais elevado na segunda (30%) do que na terceira posição (17%). Contudo, quando se cruzarem as variáveis independentes *posição linear* e *posição relativamente ao núcleo do SN* (ponto 4.1.2.6.2 do presente trabalho), perceber-se-á que intervém aqui a hipótese funcionalista das Condições de Distintividade (Kiparsky, 1972:195), já que os adjectivos em segunda posição são pré-nucleares, assumindo a função de núcleo, portanto com maior peso de marcação, enquanto os adjectivos em terceira posição se localizam pós-nuclearmente, detendo a função de sintagma, logo com probabilidade praticamente nula de serem pluralizados (Castro & Costa, 2003:108).

Outro aspecto em que os nossos dados não acompanham as suposições de Scherre (1988:157) tem a ver com o facto de que “*os substantivos são menos marcados do que os adjectivos na segunda posição e mais marcados do que estes na terceira*”. De facto, apenas a segunda parte da premissa é válida para o PA (substantivos = 35% de marcação; adjectivos = 17% de marcação), uma vez que, na segunda posição, os substantivos apresentam um percentual de marcação ligeiramente mais elevado do que os adjectivos (substantivos = 32% de marcas; adjectivos = 30% de marcas).

Debruçando-se seguidamente sobre o facto de qualquer classe gramatical que esteja em primeira posição apresentar propensão para ser mais marcada, Scherre (1988:158) reforça, uma vez mais, que o paralelismo entre *classe* e *posição*, de acordo com o avançado por Guy (1981b:179), não pode ser estabelecido. Efectivamente, a maioria dos casos em primeira localização constitui ocorrências apenas com alguns determinantes, caso dos artigos, demonstrativos e indefinidos. Quanto aos outros determinantes, como o possessivo e o quantificador, podem ocorrer em diferentes lugares do SN, dependendo a sua frequência de marcação das posições em que surgem realizados. Nesta perspectiva não se poderá estabelecer uma relação entre primeira posição e determinantes, sendo

mais correcto que, e de acordo com Scherre (1988:163), se postule apenas que a primeira posição do SN é aquela que é mais marcada, independentemente da classe gramatical. Tal é também confirmado por Lopes (2001:174) e P. Andrade (2003:96-97) para a quase totalidade dos itens gramaticais, exceptuando-se os artigos, cuja frequência de marcação se mantém equilibrada em todas as posições, tanto na NURC (1ª. posição = 99% de marcação; 2ª. posição = 100% de marcação; 3ª. posição = 100% de marcação) como no dialecto de HEL-Ba (1ª. posição = 98% de marcação; 2ª. posição = 100% de marcação). Nos nossos dados, o artigo definido revela tendência semelhante à destas duas variedades (1ª. posição = 90% de marcação; 2ª. posição = 100% de marcação).

Concretamente no que diz respeito ao possessivo, Scherre (1988:161) chama a atenção não só para a sua baixa frequência na terceira posição, tal como acontece nos nossos dados (apenas duas realizações), mas também para a grande percentagem de ocorrências na segunda posição em SN's do tipo [ART ou DEM+POS+N]. Nos dados do MRJ (Scherre, 1988:160), esta frequência de realizações na segunda posição (135/331 dados = 41% do total das ocorrências) é inferior à do possessivo em primeira posição (184/331 dados = 56% do total das ocorrências). No entanto, este aspecto é compreensível, dada a expletividade do definido no PB, isto é, ocorrer frequentemente a sua omissão quando antecede os possessivos que estão em posição pré-nuclear.<sup>174</sup> Porém, no PE, em SN's do género daqueles que são referidos por Scherre (1988:160), a construção sem artigo é agramatical, pelo que a posição do possessivo em cadeias sintagmáticas reduzidas que envolvam o seu uso pré-nuclear é, por norma, a segunda. O maior contacto dos falantes de Almojarife com a estrutura gramatical do PE determina que as construções envolvendo possessivos reflectam, também, respaldo na norma europeia, embora as influências provenientes dos substratos, nomeadamente do santomense, estabeleçam uma correlação de forças com a norma do PE na motivação para a inserção de marcas. Deste modo, são realizadas quer construções do tipo das que se encontram no PB, isto é, sem artigo (exemplo [258]), quer frases com artigo, como acontece no PE (exemplo [259]):

- [258] PA:  
... *depois veio companhá meus pai.* [DULCEM1]  
“... depois veio acompanhar os meus pais”

[259] PA:  
*É os meus pai que me deu essa idea.* [OSVALH1]  
“Foram os meus pais que me deram essa ideia.”

Assim, nota-se que, nos nossos dados, o número de realizações que envolvem o uso do possessivo é bastante equilibrado nas duas localizações referidas (1ª. posição = 16 ocorrências; 2ª. posição = 18 ocorrências), com o percentual de pluralizações da primeira (88%) superando o da segunda (78%), dado ocorrer a tendência para se marcar o número na primeira oportunidade, por influência dos substratos africanos (Guy, 1981a:301-302), como se vem acentuando neste trabalho. No entanto, note-se que Lopes (201:174) estranha o facto de os possessivos serem alvo de tão alto percentual de marcação na segunda posição (95% nos dados da NURC), “*já que essa é uma posição que, como se viu, desfavorece a marca de plural em itens de outra classe*”. Debruçando-se então sobre estes elementos gramaticais, a autora conclui que a concordância é feita, principalmente, quando o possessivo se encontra em posição pré-nuclear. Estudos sobre a CPL-var, à luz dos pressupostos da MD (Castro & Costa, 2003; Costa & Silva, 2006b), atribuem a marcação pré-nominal ao facto de o possessivo ser forte portador de definitude e ter aqui a categoria de elemento nuclear, enquanto na posição pós-nominal detém a categoria de sintagma, portanto, com tendência para não ser marcado (ponto 2.9 do presente trabalho). Seja como for, estas particularidades alertam para a necessidade de se observar a influência da variável independente *posição em relação ao núcleo do SN* sobre a marcação de todos os itens gramaticais. Ainda assim, gostaríamos de chamar também a atenção para o facto de, à excepção dos substantivos e categorias substantivadas (elementos nucleares, portanto), todos os outros elementos gramaticais dos nossos dados em segunda posição, ao contrário do pretendido por Lopes (2001:174), apresentarem elevado percentual de marcação, como se pode verificar na Tabela 21.5. Este aspecto é, aliás, observado também por Scherre (1988:165) nas suas análises: “*o percentual de marcas de todas as outras categorias que não o substantivo, à excepção do quantificador, é bastante elevado, inclusive o dos adjetivos*”. Tal não surpreende, visto a expansão dos SN’s ser levada a cabo pelas gerações mais jovens, consequentemente mais escolarizadas, e que aplicam mais a regra da concordância PL.

Quanto aos quantificadores, os dados de Scherre (1988:160) apresentam um total bastante elevado de realizações na terceira posição (108/248 dados = 43% das

realizações totais), e que não se distancia muito do registado na primeira posição (133/248 dados = 53% das realizações totais). Por outro lado, os referidos itens revelam-se categoricamente marcados na primeira posição (100% de marcas), caindo essa tendência de modo drástico na segunda posição (20% marcas). Scherre (1988:161-162) postula então que o “*comportamento do possessivo e do quantificador na terceira posição assemelha-se mais ao dos adjectivos do que ao dos determinantes. A característica distribucional que têm em comum é exactamente a de estarem na terceira posição ou, em outras palavras, estarem pospostos ao núcleo do SN*”. Lembre-se também que as observações da CPL-var, de acordo com pressupostos da MD (Castro & Costa, 2003; Costa & Silva, 2006b), chamam a atenção para o paralelo estabelecido entre possessivos e adjectivos, no que concerne à correlação entre posição e definitude determinada pelas posições pré e pós-nominais, que constroem as categorias do tipo nuclear ou sintagmático (ponto 2.9 do presente trabalho). P. Andrade (2003:96), por seu lado, confirma que, nos seus dados, metade dos quantificadores se concentra na terceira posição, mas que nenhum deles recebe marcação PL. Os nossos dados, contudo, não confirmam Scherre (1988:161-162) e P. Andrade (2003:96) no que respeita às ocorrências na terceira localização, já que os possessivos, como se viu, se distribuem mais pelas primeira e segunda posições. Quanto aos quantificadores, também não ajudam a corroborar a opinião de ambas, já que registam apenas onze dados (9% das realizações) na terceira posição, embora o seu percentual de marcação seja também baixo (apenas um dado marcado). Assim, no PA a tendência é para estes itens gramaticais serem maioritariamente realizados na primeira posição (103 dados = 89% das realizações), revelando-se o peso da marcação nesta localização bastante categórico (pr. 0,883). Scherre (1988:161) chama ainda a atenção para o facto de os quantificadores ocorrerem esporadicamente na segunda posição, o que se confirma tanto nos nossos dados (apenas duas realizações) como nos de HEL-Ba (uma única ocorrência).

Como se vê, Lopes (2001:171) e P. Andrade (2003:98) corroboram Scherre (1988:152) na generalidade das observações desta, concluindo a primeira que os seus resultados “*derrubam a hipótese*” apontada por Guy (1981b:179), que defende a ocorrência de uma relação directa entre classe e posição. Tal tem a ver com o facto de, na NURC, apesar de os artigos ocorrerem preferencialmente na primeira posição (95%)

e os substantivos com mais incidência na segunda (85%), a maioria dos adjectivos do tipo 1 cair fora da terceira posição (3ª. posição = 41% das realizações). Por outro lado, a quarta posição é também maioritariamente ocupada por substantivos (56,5% dos itens nesse lugar). P. Andrade (2003:98), por seu lado, entende que o efeito virtualmente idêntico entre posição e classe se aplica à primeira posição e os determinantes, mas que os substantivos e adjectivos, apesar de aparecerem preferencialmente nas posições mencionadas por Guy (1981b:179) — segunda e terceira posições, respectivamente —, não apresentam a mesma proporção de marcas evidenciada pelos determinantes na primeira posição. Não obstante, “*a prevalência de marcas nos elementos que ocupam a primeira posição permite-nos corroborar a teoria de Guy sobre a indicação de plural, no português do Brasil, ocorrer na primeira oportunidade*” (P. Andrade, 2003:98), em virtude de o morfema que expressa o número se tratar de um *early system morpheme* (morfema sistémico estrutural), presente nos itens que são adquiridos por transmissão linguística irregular. Já no que diz respeito à concordância, a mesma não acontece porque, segundo P. Andrade (2003:98), a sua concretização se faz com recursos aos morfemas adquiridos mais tardiamente, isto é, os *late system morphemes* (morfemas do nível funcional sistémico) (cf. ponto 2.4.2.3 do presente trabalho).

Relativamente ao PA, o que se pode dizer é que os nossos dados acompanham os da NURC e de HEL-Ba na perspectiva do raciocínio de Lopes (2001:171) e P. Andrade (2003:98). De facto os artigos do nosso *corpus* ocorrem também com mais incidência na 1ª. posição, sejam eles definidos (1ª. posição = 96%; total das outras posições = 4%) ou indefinidos (1ª. posição = 100%), enquanto os substantivos são maioritariamente mais realizados na segunda localização (2ª. posição = 84%; total das outras posições = 16%). Quantos aos adjectivos do tipo 1, registam uma maior ocorrência nas outras posições do que no terceiro lugar (3ª. posição = 37%; total das outras posições = 63%).

Ao analisar os percentuais de marcação dos vários itens gramaticais nas distintas posições, Scherre (1988) e Lopes (2001) procuram rebater o pressuposto avançado por Guy (1981b:179) de que não se pode estabelecer uma relação directa no efeito entre *posição linear* e *classe gramatical*. Contudo, não só os nossos dados apresentam resultados que não vão totalmente ao encontro das suposições avançadas por Scherre (1988) como também cremos que Guy (1981b:179), ao referir-se a esta relação, estará



apenas a chamar a atenção para as posições preferenciais em que os referidos elementos gramaticais ocorrem. E neste aspecto, os resultados da Tabela 21.5 dão-lhe total razão, uma vez que os determinantes registam um total de 715 realizações na primeira posição, contra apenas 43 na segunda, 15 na terceira e uma na quarta ou outra posição. Por seu lado, os substantivos e categorias substantivadas exibem 1248 dados na segunda posição, contra somente 29 na primeira, 117 na terceira e 88 na quarta posição. Quanto aos adjectivos, de facto parece existir um equilíbrio de realizações entre a segunda e a terceira posição. Contudo, se contabilizarmos também os adjectivos do tipo 2 e as flexões em grau, verificamos que o total deste tipo de itens gramaticais é de 36 ocorrências na terceira posição, contra 7 na primeira, 23 na segunda e 11 na quarta ou outra posição. Por outro lado, é preciso não esquecer que a importância na marcação deve ser avaliada não em termos de percentagens, mas em função de pesos relativos. Estes, como se sabe, são estabelecidos entre elementos do mesmo grupo de factores que ocorrem na mesma posição e não avaliando um factor, *per se*, em função das suas ocorrências nas diferentes posições da cadeia sintagmática. Daí que os substantivos, por exemplo, registem um peso relativo mais alto na segunda (pr. 0,311) do que na terceira posição (pr. 0,263), apesar de apresentarem um percentual de marcação mais baixo naquela (32%) do que nesta (35%). Todavia, é um facto que os resultados da Tabela 21.5 pouco poderão adiantar sobre a influência na marcação motivada pela relação entre *posição linear* e *classe gramatical*, se não tivermos em linha de conta o modo como o favorecimento e a inibição actuam no desenho da configuração da CPL-var. Como tal, este entendimento só poderá acontecer caso cruzemos estas duas variáveis independentes, por si só e relacionadas entre si, com o grupo de factores *posição em relação ao núcleo do SN*, já que, ao que tudo indica, parece ser determinante para a marcação o facto de os itens gramaticais ocorrerem em posição pré ou pós-nuclear. Aliás, Scherre (1988:162) refere, igualmente, que a grande diferença entre o possessivo e o quantificador em segunda posição “*se dá em função também da sua relação com o núcleo do SN. Dos 235 casos de possessivo desta posição, 130 são antepostos. Dos sete casos de quantificador desta mesma posição, todos são pospostos*”. Daí que a autora tenha decidido observar também a variável *relação entre os elementos do SN*, de forma a achar uma solução para o controverso postulado de Guy (1981b:179) acerca da

equivalência entre classe gramatical e posição. Deste modo, a observação da forma como a variável independente *posição em relação ao núcleo do SN* se relaciona com os grupos de factores *posição linear* e *classe gramatical* constitui o próximo passo do nosso estudo.

#### **4.1.2.6.2. Relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN**

Tendo em conta o facto de os itens pré-nucleares favorecerem mais a marcação do que os elementos em posição pós-nuclear, como se viu no ponto 4.1.2.4. do presente trabalho, dedicado à análise dos resultados da variável independente *marcas precedentes*, interessa agora saber não só qual é o peso na marcação PL desempenhado pelos diferentes itens gramaticais nos distintos lugares do SN mas também como se comporta o núcleo nas várias posições, igualmente em termos de pluralização. Para a constituição da variável que cruza os grupos de factores *posição linear* e *posição em relação ao núcleo do SN*, Lopes (2001:172-174) constituiu onze factores, que foram adoptados também por Figueiredo (2008:28-29) nas suas pesquisas iniciais sobre a CPL-var no SN do PA. Contudo, em virtude de os nossos dados apresentarem apenas noventa ocorrências nas diferentes posições pós-nucleares (Tabela 12.6), a distribuição das mesmas pelas diferentes faixas etárias levou a que algumas das realizações tivessem de ser eliminadas por apresentarem menos de trinta dados. Figueiredo (2008:30) procedeu então à amalgamação de todos os factores pós-nucleares num único factor (*itens plurais à direita do núcleo do SN*), tendo o grupo de factores que faz o cruzamento das variáveis independentes em questão ficado reduzido aos sete factores apresentados na Tabela 21.6 e que serão também utilizados por nós para análises mais detalhadas do que as apresentadas no trabalho inicial de Figueiredo (2008) sobre o PA.

**Tabela 21.6.** Efeito da relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN do PA: total de ocorrências e percentuais de marcação.

<b>Factores</b>	<b>Exemplo</b>	<b>Nr. total</b>	<b>%</b>
Anterior ao núcleo, não imediatamente	UN <u>S</u> quatro <b>ano</b> [CELESH3]]	57/87	66
Imediatamente anterior ao núcleo	UMAS <b>dores</b> [DULCEM1]	631/681	93
Núcleo na 1ª. posição do SN	‘ <b>SOAS</b> morta [DULCEM1]	(25/29)	86
Núcleo na 2ª. posição do SN	umas <b>DORES</b> só [DULCEM1]	399/1.248	32
Núcleo na 3ª. posição do SN	o <u>s</u> meus <b>FILHOS</b> [MANOH3]	41/117	35
Núcleo na 4ª. ou outra posição do SN	sete ou oito <b>MESES</b> [CASTEH1]	32/88	36
Posterior ao núcleo (2ª., 3ª., 4ª. ou outra posição)	HOMENS <b>capazes</b> [CASTEH1] PRÉDIOS <b>bem grandes</b> [ANTOM1]	17/90	19
<b>Totais:</b>		1.202/2.340	51

À excepção do factor *núcleo na 1ª. posição do SN*, com apenas 29 ocorrências, todos os outros factores possuem mais de trinta dados. A fim de podermos observar também o comportamento do factor que tem menos de trinta realizações, decidimos incluí-lo na rodada para obtenção de pesos relativos. A tabela 21.7 apresenta os resultados do efeito da relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN do PA:

**Tabela 21.7.** Efeito da relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN do PA.

<i>Input desta rodada: 0,527</i>		<i>Log-likelihood: -952,926</i>		<i>Significância: 0,028</i>	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Anterior ao núcleo, não imediatamente	57/87	66	4	0,565	
Imediatamente anterior ao núcleo	631/681	93	29	0,903	
Núcleo na 1ª. posição do SN	(25/29)	86	1	0,678	
Núcleo na 2ª. posição do SN	399/1.248	32	53	0,300	
Núcleo na 3ª. posição do SN	41/117	35	5	0,210	
Núcleo na 4ª. ou outra posição do SN	32/88	36	4	0,213	
Posterior ao núcleo (2ª., 3ª., 4ª. ou outra posição)	17/90	19	4	0,069	
<b>Totais:</b>		1.202/2.340	51	100	-

Sobre o comportamento dos itens pré e pós-nucleares já nos debruçamos no ponto 4.1.2.6.1 do presente trabalho. Relativamente ao efeito do núcleo do SN na marcação PL do SN do PA, constata-se a sua tendência para favorecer moderadamente a inserção de marcas quando se encontra na primeira posição da cadeia sintagmática (pr. 0,678). Comparando os presentes resultados com os da Tabela 16.2 (relação entre *classe gramatical* e *posição em relação ao núcleo do SN*) e 21.5 (relação entre *posição linear* e *classe gramatical*) é possível confirmar que o peso inibidor revelado pelos substantivos e categorias substantivadas não actua na marcação quando estes itens se encontram na primeira posição sintagmática, reforçando-se o pressuposto funcionalista acerca de a marcação ocorrer preferencialmente na primeira oportunidade (Kiparsky, 1972:195), conforme sistema de pluralização dos ancestrais substratos africanos (Guy, 1981a:301-302). Isto mesmo pode ser confirmado através da inibição apresentada pelo núcleo na segunda posição do SN (pr. 0,300), passando a pluralização a incidir sobre o elemento que o antecede, independentemente da sua classe gramatical (Tabela 16.2). Como se pode ver ainda, a inibição aumenta à medida que o núcleo se distancia da primeira posição, apresentando valores bastante inibidores e que apenas não se distanciam entre si nas terceira e quarta ou outra posição (núcleo na 3ª. posição = pr. 0,210; núcleo na 4ª. ou outra posição = pr. 0,213).

Analisando a CGEN-var do dialecto de HEL-Ba, Lucchesi (2000a:249-250) refere que, em situação de transmissão linguística irregular, os falantes exibem uma variação significativa, condizente com a fase inicial e mais precária da aquisição do PtgL2. Deste modo, a expansão da regra de concordância, levada a cabo pelas gerações afectadas por pressões sociolinguísticas, é norteadas por três princípios gerais: (i) Princípio da Simplicidade; (ii) Princípio da Integração (ambos referenciados no ponto 1.3.1.1 do presente trabalho); (iii) Princípio da Saliência (Naro & Lemle, 1976:240-241; Lemle & Naro, 1977:47), que aponta para a ocorrência de uma maior concordância de género quer para as estruturas de género com marcas mais explicitadas quer para os SN's com núcleo preenchido (por oposição aos núcleos de categoria vazia). Estes aspectos levam a concordância de género a expandir-se a partir do nome núcleo, para se fixar na estrutura do SN, orientando-se para os elementos à esquerda do nome e, marginalmente para os itens à direita deste. Contudo, o que se observa na concordância de número é que os

itens em posição pré-nuclear são os que têm tendência para receberem a marcação mórfica de concordância, com especial incidência para o elemento colocado na posição imediatamente anterior ao núcleo do SN, ou seja, a categoria DET. Esta tendência sugere, por um lado, a hipótese de a inserção da marca de número se iniciar neste elemento gramatical (Longobardi, 1994; Baxter, 2009), sob a influência do sistema de marcação PL das ancestrais línguas do substrato africano (cf. Guy, 1981a:301-302), nas quais predomina o uso de classificadores nominais em forma de prefixo ou clítico, que vão também a motivar a CPL-var de HEL-Ba e do PA. Por outro lado, o facto de os elementos pré-nucleares terem um papel activo na construção estrutural da concordância relega os itens pós-nucleares para funções sintácticas de mera adjunção (Lucchesi, 2000a; Figueiredo, 2009a), não se inserindo marcas de concordância de PL nestes, em virtude de as mesmas se revelarem redundantes (Kiparsky, 1972:195). Numa perspectiva generativista (ponto 2.6.6 do presente trabalho), a não-realização fonética da marca PL licenciada por {s} obedece a uma operação de identificação de conteúdo semântico-funcional (Cerqueira, 1994:125), que garante ao SDET, enquanto projecção máxima de N e SN, a possibilidade de projectar os traços categoriais e flexionais destes dois núcleos alargados, permitindo que os mesmos estabeleçam o pareamento de traços entre si. Por seu lado, a marca PL em DET vai permitir que a mesma não seja visível em N, ou seja, vai tornar possível o licenciamento de {s} nulo neste. Por outro lado, como determinadas línguas não seleccionam traços [-interpretáveis] de género e número (Franceschina, 2002:76), como acontece com os dialectos do grupo níger-congo atlântico, os falantes adultos destes experimentam dificuldades na aquisição dos mesmos traços do português (cf. Franceschina, 2003:104), o que os leva a estabelecerem novos/errados parâmetros (cf. Gonçalves, 2004:235), responsáveis pela reinstanciação paramétrica no PAL1 e consequente aquisição das regras de concordância variáveis de género e de número a nível do SN.

As análises da relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL levadas a cabo para o PT (Baxter, 2004:117) e HEL-Ba (Baxter, 2009:278) revelaram a existência de gramáticas distintas para as diferentes gerações de ambas as comunidades. Nesta conformidade, para se perceber se existe uma regra quantificadora que determine com exactidão o ponto de entrada da marcação de número no PA e qual a

variação que a mesma sofre em situação de transmissão linguística irregular ao longo das gerações, decidimos verificar também o efeito destas duas variáveis independentes na marcação PL das diferentes faixas etárias de Almojarife. A Tabela 21.8 apresenta os resultados obtidos quer para a comunidade quer para as três gerações em estudo.

**Tabela 21.8.** Relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação PL dos itens do SN do PA: comunidade e faixas etárias.

		<i>Input desta rodada: 0,527</i> <i>Log-likelihood: -952,926</i> <i>Significância: 0,028</i>			<i>Input desta rodada: 0,348</i> <i>Log-likelihood: -148,154</i> <i>Significância: 0,027</i>		
<b>Factores</b>	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>			
	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	
Anterior ao núcleo, não imediatamente	57/87	66	0,565	(11/19)	58	-	
Imediatamente anterior ao núcleo	631/681	93	0,903	(133/140)	95	0,973	
Núcleo na 1ª. posição do SN	(25/29)	86	0,678	-	-	-	
Núcleo na 2ª. posição do SN	399/1.248	32	0,300	57/357	16	0,230	
Núcleo na 3ª. posição do SN	41/117	35	0,210	(3/28)	11	0,078	
Núcleo na 4ª. ou outra posição do SN	32/88	36	0,213	(0/10)	0	-	
Posterior ao núcleo	17/90	19	0,069	(2/16)	13	-	
<b>Totais:</b>	<b>1.202/2.340</b>	<b>51</b>	<b>-</b>	<b>206/570</b>	<b>36</b>	<b>-</b>	
		<i>Input desta rodada: 0,403</i> <i>Log-likelihood: -347,080</i> <i>Significância: 0,005</i>			<i>Input desta rodada: 0,811</i> <i>Log-likelihood: 398,848</i> <i>Significância: 0,030</i>		
<b>Factores</b>	<b>FE-2 (41 – 60 anos)</b>			<b>FE-1 (20 – 40 anos)</b>			
	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	
Anterior ao núcleo, não imediatamente	(14/29)	48	0,365	32/39	82	0,685	
Imediatamente anterior ao núcleo	196/231	85	0,856	302/310	97	0,947	
Núcleo na 1ª. posição do SN	(4/5)	80	-	(21/24)	88	-	
Núcleo na 2ª. posição do SN	127/417	31	0,367	215/474	45	0,183	
Núcleo na 3ª. posição do SN	16/35	46	0,260	22/54	41	0,222	
Núcleo na 4ª. ou outra posição do SN	12/40	30	0,208	20/38	53	0,280	
Posterior ao núcleo	(4/29)	14	0,064	11/45	24	0,075	
<b>Totais:</b>	<b>373/786</b>	<b>48</b>	<b>-</b>	<b>623/984</b>	<b>63</b>	<b>-</b>	

Os resultados evidenciam dois aspectos a ter em conta: (i) a existência de gramáticas distintas nas diferentes gerações de Almojarife; (ii) perfis de marcação geracional muito semelhantes entre o PA, o PT e HEL-Ba, como se verá mais detalhadamente. Assim, as distintas gramáticas geracionais poderão ser observadas, por exemplo, pelo comportamento do factor *posição anterior ao núcleo do SN, mas não imediatamente*, que caminha no sentido da aquisição da concordância, pois começa por ser um elemento com escassas ocorrências na FE-3 (apenas 19), para passar, depois, a ser mais usado na faixa etária intermédia, mas com um peso inibidor da marcação (pr. 0,365), e acabar por se fixar, na geração mais nova, como item que favorece moderadamente a marcação (pr. 0,685). Este aspecto, para além de revelar um padrão de desenvolvimento geracional oposto ao da deriva românica, evidencia também a preferência da geração mais idosa do PA pela realização de SN's que envolvem apenas uma posição pré-nuclear, isto é, de estrutura reduzida, um aspecto que encontra paralelo nas outras duas comunidades observadas. Como tal, são escassos os SN's que apresentam três ou mais itens gramaticais, tanto em Almojarife como em Monte Café ou HEL-Ba. Paralelamente, a tendência para fixar a marcação em número na posição imediatamente à esquerda do núcleo do SN é logo estabelecida na faixa etária mais idosa (PA = pr. 0,973; PT = pr. 0,958; HEL-Ba = 0,960), com o núcleo desfavorecendo sempre a inserção de marcas, independentemente da sua localização, mas com maior incidência quando se encontra no terceiro lugar do SN, excepto em HEL-Ba (núcleo na 2<sup>a</sup>. posição: PA = pr. 0,230; PT = 0,260; HEL-Ba = 0,111 / núcleo na 3<sup>a</sup>. posição: PA = pr. 0,078; PT = 0% marcação em 13 ocorrências; HEL-Ba = 0,268). Estes aspectos dão consistência, uma vez mais, ao pressuposto de que é a categoria funcional DET o ponto de inserção da marca de número PL (Longobardi, 1994; Baxter, 2009), sob a forma de morfema *singleton*, nestas três variedades de português. Note-se também a baixíssima marcação PL apresentada pelos itens pós-nucleares neste estrato etário de todas as variedades em análise (PA = 13% de marcação; PT = pr. 0,024; HEL-Ba = 0% de marcação).

O perfil de marcação evidenciado pela FE-3 de Almojarife transita, em grande parte, para a FE-2, sendo acompanhado, na maioria das suas especificidades, pelos contornos sintagmáticos do PT e de HEL-Ba. Assim, no PA e em HEL-Ba, os pontos de entrada da marcação PL continuam a manter-se unicamente no elemento imediatamente pré-nuclear

(PA = pr. 0,856; HEL-Ba = pr. 0,939), mas no PT, o *elemento anterior ao núcleo, mas não imediatamente*, apresenta já um elevado peso de inserção de marcas (imediatamente anterior ao núcleo = pr. 0,913; anterior ao núcleo, mas não imediatamente = pr. 0,884). Em Almojarife e em HEL-Ba, contudo, o factor *anterior ao núcleo, mas não imediatamente* revela-se moderadamente inibidor da pluralização, mais no primeiro dialecto do que no segundo (PA= pr. 0,365; HEL-Ba = pr. 0,421).

Nesta faixa etária, os SN's apresentam já uma estrutura relativamente mais desenvolvida, com os itens pré-nucleares não adjacentes mostrando mais realizações. Por seu lado, o núcleo aumenta também as suas ocorrências em todas as posições da cadeia sintagmática, exibindo, nas variedades africanas, produções na primeira posição do SN, algo que não sucedia nas faixas etárias mais idosas destas línguas. Não obstante, o desfavorecimento do núcleo mantém-se nesta geração, revelando, na quase exclusividade dos dados, um peso cada vez mais acentuado à medida que se afasta da primeira posição, isto é, à medida que a estrutura do SN se distende e se torna mais complexa (1ª. posição: PA = 5 dados com 80% marcação; PT = 6 dados com 100% marcação; HEL-Ba = ausência de realizações / 2ª. posição: PA = pr. 0,367; PT = pr. 0,243; HEL-Ba = pr. 0,179 / 3ª. posição: PA = pr. 0,260; PT = pr. 0,320; HEL-Ba = pr.0,001 / 4ª. posição: PA = pr. 0,208). Quanto aos itens pós-nucleares, confirmam também a tendência para inibirem fortemente a pluralização já verificada na geração anterior (PA = pr. 0,064; PT = pr. 0,084; HEL-Ba = 0% de marcação em 26 dados).

Esta faixa etária, apesar de apresentar estruturas sintagmáticas mais desenvolvidas, começa a evidenciar que a marcação PL revela tendência a fixar-se na primeira posição do SN, como o demonstra o facto de os nomes nesta localização surgirem marcados, o que não acontece quando estão em segundo lugar, visto a inserção de marcas já ter sido levada a cabo no elemento que se encontra imediatamente à sua esquerda. Estes aspectos evidenciam que, quando essa posição não é ocupada por um elemento DET, o SN se torna o alvo de PL, em virtude das suas características lexicais (Baxter, 2009:292). Por seu lado, os itens pré-nucleares adquiridos mais tardiamente, e que ajudam a construir estruturas sintagmáticas mais complexas (*anterior ao núcleo, mas não imediatamente*), revelam-se mais receptivos à inserção de marcas do que os elementos pós-nucleares. Antevê-se aqui o início da aquisição da regra de concordância, sobretudo por influência



de factores sociais, como a escolarização pública, a migração periódica de homens em busca de trabalho ou sua inserção nos quadros do exército colonial, o melhoramento do sistema rodoviário e de transportes ou o acesso aos meios de comunicação de massa. Deste modo, o PL inicia o seu desenvolvimento com características de concordância parcial, através de morfemas dissociados, que são atribuídos aos elementos mais próximos do *fulcro* inicial. A ordem de atribuição contempla os elementos do SDET, primeiro, e os itens à direita do núcleo, depois. Não obstante, o perfil da âncora pré-nuclear é mantido geracionalmente, denotando que o sistema de marcação adquirido por transmissão irregular se começa a cristalizar (elemento imediatamente pré-nuclear fortemente marcado, núcleo e elemento pós-nuclear altamente desfavorecedores da pluralização), confirmando uma reinstanciação paramétrica no PAL1, com base no *input* marcado pelos novos/errados parâmetros instanciados na L2 (cf. Gonçalves, 2004:235).

Finalmente, a FE-1 estabiliza a variação, na qual é patente a herança da marcação PL adquirida por transmissão linguística irregular. Deste modo, apesar de os SN's se apresentarem ainda ligeiramente mais complexos do que os da geração antecedente (veja-se, por exemplo, como continua a crescer o número de realizações com núcleo em primeira posição), o padrão de marcação mantém-se parcial, uma vez que apenas os elementos pré-nucleares não adjacentes revelam um aumento de concordância, passando a favorecer moderadamente a marcação (pr. 0,685). Este aumento de concordância em direcção ao favorecimento de inserção de marcas ocorre também em HEL-Ba (pr. 0,673), mas mantém-se constante no PT (p. 0,770) que, como vimos, já marca fortemente o número nesta posição na faixa etária intermédia (pr. 0,884). Quanto aos restantes factores, confirmam que o padrão geral da pluralização iniciado na FE-3 não sofre grandes alterações relativamente ao grau de inserção/inibição já revelado pelas gerações antepassadas. Desta forma, a primeira posição continua a revelar-se como favorecedora da marcação, independentemente do item que nela se encontre, mesmo que seja o núcleo do SN (88% de realizações marcadas). Este favorecimento é idêntico no PT (pr. 0,770), mas HEL-Ba revela um estágio mais avançado na aplicação da regra, já que apenas regista uma ocorrência com nome na primeira posição. O favorecimento maior da inserção de marcas continua, contudo, a recair no elemento imediatamente

anterior ao núcleo, que é categoricamente marcado em todas estas variedades de português (PA = pr. 0,947; PT = 0,835; HEL-Ba = pr. 0,919).

Por seu lado, a inibição da marcação também não revela alterações significativas relativamente à realizada nas gerações primordiais, pois continua a actuar categórica e uniformemente quer sobre o núcleo em posições que não a primeira (2<sup>a</sup>. posição: PA = pr. 0,183; PT = pr. 0,243; HEL-Ba = pr.0,153 / 3<sup>a</sup>. posição: PA = pr. 0,222; PT = pr. 0,273; HEL-Ba = pr. 0,174 / 4<sup>a</sup>. posição: PA = pr. 0,280) quer sobre os itens pós-nucleares (PA = pr. 0,075; PT = pr. 0,144; HEL-Ba = 0% de marcação em 23 dados). Desta forma, a CPL-var no SN das três variedades de português aqui observadas acaba, em última instância, por confirmar os princípios advogados pela FFFH, que preconizam a dificuldade na ASL dos traços não interpretáveis de número dos elementos funcionais ou não parametrizados (Franceschina, 2002:76), adquiridos em fase pós-puberdade, e seu conseqüente apagamento geracional, em virtude de ocorrer reinstanciação paramétrica de novos errados/errados parâmetros na L1 nativizada (Gonçalves, 2004:235).

Lopes (2001:187) observou também o efeito da escolaridade no padrão de variação que cruza as variáveis independentes *posição linear* com a *posição em relação ao núcleo do SN*. O perfil em todos os estratos escolares é idêntico entre si e não se afasta daquele que é revelado pela comunidade: (i) itens pré-nucleares favorecendo a pluralização, com destaque para a posição imediatamente anterior ao núcleo do SN; (ii) núcleo favorecendo a marcação ou mostrando-se neutro na primeira posição, mas desfavorecendo-a nas restantes localizações sintagmáticas; (iii) elementos pós-nucleares inibindo categoricamente a inserção de marcas. Se levarmos em linha de conta que o postulado (i) se aplica maioritariamente aos SN's de dois elementos (DET+N), acreditamos que, no PA, o padrão de variação não se afastará deste, até porque, como se viu, existe uma forte aproximação entre os resultados dos estratos etários e da escolaridade nas duas variedades. Assim sendo, abstivemo-nos de observar a relação entres as referidas variáveis linguísticas e a variável social *escolaridade*, uma vez que os valores de tal relação não se afastarão consideravelmente daqueles que estão patentes na Tabela 21.8, com os resultados da FE-3 correspondendo aos das faixas analfabetas e de escolarização baixa, os da FE-2 evidenciando valores respeitantes aos dos falantes de

escolarização média e a FE-1 fornecendo números para os membros da comunidade com escolarização alta.

Porém, importa ainda realçar que os padrões de variação das faixas etárias mais novas das três variedades de português observadas (PA, PT e HEL-Ba) acabam por reflectir também os das próprias comunidades, já que todas elas exibem bastantes semelhanças entre si. Por outro lado, estas semelhanças são ainda extensíveis aos padrões da CPL-var de Salvador (Lopes, 2001:178-180) e do Maputo (Jon-And, 2008:5), tal como se pode verificar na Tabela 21.9.

No caso específico do PVA, não temos conhecimento de trabalhos com recurso a métodos quantitativos. No entanto, Inverno (2005) recolheu um *corpus* na região do Dundo (Província da Lunda Norte), com recurso a entrevistas levadas a cabo a informantes de diferentes escalões etários, sexo e níveis de escolaridade, a fim de proceder a uma observação das ocorrências variáveis no mesmo. No caso concreto dos itens que compõem o SN, Inverno (2005:130) concluiu que o PVA não difere do PE no que concerne quer ao número de elementos quer às suas categorias gramaticais. Porém, no que diz respeito à distribuição destes elementos pela cadeia sintagmática, são evidentes as diferenças, principalmente quando se trata de possessivos e marcação da pessoa. Mas não só, já que as concordâncias de género e o número evidenciam também variações relativamente ao PE (ponto 1.3.1.2 do presente trabalho).

No caso concreto da CPL-var, o fenómeno é generalizado e vai ao encontro de outros registados por Marques (1985 [1983]:219), Gärtner (1989:159) e Vilela (1999:56), que atribuem o fenómeno da ausência de marcação do núcleo do SN ao sistema das línguas bantu, cuja pluralização se faz com recurso a afixos aglutinados pré-nominalmente aos núcleos dos SN's. Porém, Inverno (2005) vai mais longe nas suas observações, detectando também marcações parciais de número em palavras que terminam em *-es* no PE (exemplo [260]). A autora atribui esta variação à conservação do /e/ da sílaba final e sua alternância vocálica em [i], um fenómeno que acontecerá em virtude de as sílabas das línguas bantu serem sempre abertas:

[260] PANG:  
*outro bebé só tem dez mesi* (Inverno, 2005:133)  
“o outro bebé só tem dez meses”

Paralelamente, Inverno (2005:135) avança que as suas observações permitem lançar a hipótese de que a marcação de número é feita nos elementos mais à esquerda do SN, sejam eles núcleos ou itens não nucleares, de acordo com as seguintes possibilidades: “(a) quando o SN é composto por dois elementos, somente o elemento mais à esquerda recebe a marca de número; e (b) quando o SN é composto por três elementos, somente os dois primeiros elementos recebem a marcação de Plural”<sup>175</sup> (Inverno, 2005:135). A autora indica ainda que tanto a categoria gramatical dos elementos como a função sintáctica do próprio SN parecem não condicionar directamente a marcação de número.

As hipóteses avançadas por Inverno (2005) para o PVA, salvaguardada a possibilidade de se ter que observar a questão mencionada em (b) com recurso a métodos quantitativos, encaixam nos achados quer do PA quer ainda do PT, PMp, PCV e PB, como se tem vindo a constatar.

Procedendo agora à análise da Tabela 21.9, comparem-se então os padrões de variação da marcação PL do PA, NURC, HEL-Ba, PT, PMp e PCV. Como se pode constatar em todas estas variedades de português, os elementos pré-nucleares propiciam mais a pluralização, com especial incidência para o item imediatamente à esquerda do núcleo do SN, cuja marcação é sempre categórica. Contudo, veja-se que a inserção de marcas na posição pré-nuclear não adjacente, que ocorre mais tardiamente (*late system morphemes*), revela os diferentes estádios gramaticais dos dialectos em questão. Salvador e HEL-Ba, que cessaram o intenso contacto directo com o PE mais precocemente, apresentam valores próximos do patamar da neutralidade, com a variedade mais culta, isto é, a NURC, mostrando um peso mais favorecedor da inserção de marcas (pr. 0,64) do que a variedade vernácula (pr. 0,479). Por seu lado, o PA e o PCV, apesar de terem passado por um processo de criouliização, mantiveram um contacto mais prolongado com o PE do que as variedades brasileiras, daí o peso da sua marcação se aproximar do de HEL-Ba (PA = pr. 0,565; PCV = pr. 0,44). Já o PT, apreendido mais tardiamente e sem sofrer um processo de criouliização, apresenta o peso mais favorecedor de todas as variedades, a denotar já um elevado índice de aquisição da regra de concordância (pr. 0,795). Quanto ao PMp, exhibe um valor moderadamente inibidor da marcação (pr. 0,38), já que, como Jon-And (2008:2) refere, os seus informantes “têm uma língua bantu como L1 (*xironga ou xixangana*)” e “falam uma

*língua bantu em casa e português numa actividade diária*”. Condição esta que poderá igualmente estar na base da ligeira inibição apresentada pelo PCV, já que os falantes desta variedade têm todos o CCV de São Vicente/Santo Antão como L1 (Jon-And, 2009:2).

**Tabela 21.9.** Relação entre *posição linear* e *posição em relação ao núcleo do SN* na marcação plural dos itens do SN: 6 variedades de português.

Factores	PA		NURC		HEL-Ba		PT		PMp		PCV	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%
Anterior ao núcleo, não imediatamente	0,565	66	0,64	97	0,479	85	0,795	95	0,38	95	0,44	90
Imediatamente anterior ao núcleo	0,903	93	0,90	99	0,914	98	0,915	95	0,78	98	0,78	97
Núcleo na 1ª. posição do SN	0,678	86	0,54	96	(0/1)	0	0,798	93	48/48	100	51/51	100
Núcleo na 2ª. posição do SN	0,300	32	0,17	66	0,153	9	0,223	32	0,30	78		
Núcleo na 3ª. posição do SN	0,210	35	0,11	62					0,29	81	0,32	71
Núcleo na 4ª. posição do SN			0,16	64	0,397	17	0,267	50				
Núcleo na 5ª. posição do SN	0,213	36	(16/16)	100					0,34	88		
Posterior ao núcleo, 2ª. posição do SN			0,17	80					0,26	76		
Posterior ao núcleo, 3ª. posição do SN			0,10	60	0/90	0	0,112	24	0,17	67	0,18	61
Posterior ao núcleo, 4ª. posição do SN	0,069	19	0,05	52					0,21	75		
Posterior ao núcleo, 5ª. posição do SN			0,07	67								

Ainda a propósito da inibição apresentada por alguns elementos pré-nucleares não adjacentes, Scherre (1988:162) e Lopes (2001:175) chamam a atenção para o facto de os SN's em questão serem semelhantes, uma vez que exibem elementos em segunda

posição sempre marcados, antecedidos por elementos não-marcados na primeira posição. Scherre (1988:163) observou ainda que, nos seus dados, os possessivos preenchem sempre a segunda posição. Este fenómeno levou a autora a estabelecer uma oposição entre elementos não nucleares à esquerda do núcleo, localizados nas primeira e segunda posições. Posteriormente, Lopes (2001:182-183) confirmou que a esmagadora maioria dos seus casos tinha perfil idêntico aos do MRJ e estabeleceu um paralelo com os achados de Lucchesi (2000a) sobre a CGEN-var de HEL-Ba, que advoga uma relação entre a maior proximidade ao núcleo, à esquerda deste, e mais concordância. Por outras palavras, Lopes (2001) considera o aspecto da contiguidade como um factor altamente influente na concordância e que leva os elementos adjacentes ao núcleo a serem bastante marcados. Neste aspecto o nosso entendimento diverge do de Lopes (2001), já que achamos que as motivações que levam o elemento pré-nuclear a ser categoricamente marcado assentam no pressuposto de a CPL-var ser influenciada pelo sistema de marcação em número dos substratos africanos (cf. Guy, 1981a:301-302). Estamos então perante dois pontos distintos de entrada da marcação: um para o género, com o sistema de concordância a irradiar a partir do nome núcleo do SN (Lucchesi, 2000a; Figueiredo, 2009a; Godinho, 2009), e afectando primeiramente os elementos à sua esquerda; e outro para o número, com a marca de PL a ser inserida no elemento imediatamente pré-nuclear (Guy, 1981a; P. Andrade, 2003; Baxter, 2004, 2009; Figueiredo, 2008, 2009a), isto é, na categoria DET (Baxter, 2009:292), deslocando-se depois para os elementos à sua esquerda, embora de forma ténue. Aliás, os dados não marcados do nosso *corpus* não confirmam as constatações de Scherre (1988) e Lopes (2001) quer quanto ao tipo de SN's (segunda posição sempre marcada, antecedida por não-marcação na primeira [0S(0)]), visto registarmos ocorrências não marcadas nas duas primeiras posições e que centram a marcação apenas no terceiro elemento, o núcleo do SN [00(S)] (cf. Tabela 12.3 e exemplos [261], [262] e [263]), quer quanto ao item gramatical em posição pré-nuclear, mais concretamente ao possessivo em segunda localização (exemplos [262] e [263]), que apenas regista quatro realizações no nossos *corpus*. Aliás, Lopes (2001:185) revela que Baxter lhe avançara verbalmente que esta também não era a situação do PT, uma vez que neste dialecto apenas se detectara uma ocorrência com possessivo em segunda posição. Transcrevem-se, seguidamente, alguns SN's compostos por três

elementos gramaticais, o primeiro dos quais sem marcação, e que não configuram as constatações das duas autoras:

[261] PA:  
*pra apreciar um... um outro trabalhos de televisão* [CASTEH1]  
“para apreciar uns... uns outros trabalhos de televisão”

[262] PA:  
*uma empresa tem o seu patrões* [ABILH2]  
“uma empresa tem os seus patrões”

[263] PA:  
*cada uma dessa dependência tem o seu donos* [ABILH2]  
“cada uma dessas dependências tem os seus donos”

Quanto à marcação que incide sobre o núcleo nas diferentes localizações sintagmáticas, a propensão para a primeira posição favorecer a inserção da marca de número em todas as variedades de português é também uma constatação geral. No que respeita às restantes posições nucleares, bem como às localizações pós-nucleares, verifica-se que o sistema de marcação tem respaldo no das gerações antepassadas, inibindo categoricamente a pluralização em todos os dialectos que aqui se observam. Note-se, contudo, o valor da NURC para o factor *núcleo na quinta posição do SN*, com dezasseis realizações a favorecerem plenamente a marcação (100%). Este aspecto mereceu a nossa atenção, por constituir uma excepção relevante no paradigma de marcação da posição nuclear de todas as variedades de português apresentadas. Porém, depois de observarmos o exemplo avançado por Lopes (2001:180) para este factor (“*as mesmas quer dizer assim PESSOAS que são*”), somos obrigados a discordar da codificação adoptada pela autora, já que, em nossa opinião, ocorrerão nesta cadeia discursiva pausas que visam estabelecer um encadeamento de ideias e consequente reformulação do discurso, não se consubstanciando o referido exemplo como um SN uno e indivisível.<sup>176</sup> Assim sendo, o valor apontado para a NURC terá que ser descartado do padrão geral de marcação revelado pelo núcleo do SN nas diferentes posições que ocupa, sugerindo-se que as codificações para esta variedade sejam revistas no que concerne a este factor, a fim de se obterem resultados conclusivos para o mesmo.

Face à evidência da importância desempenhada pela posição imediatamente pré-nuclear na marcação geracional do PL no SN de Almojarife (e também de todas as outras variedades de português), levanta-se uma questão: porque é que tal acontece logo

nas gerações mais antepassadas, mantendo-se o padrão estabilizado até às faixas etárias actuais? No que concerne especificamente ao PT, Baxter (2004:117) avançou três sugestões para justificar a emergência da marcação irregular no PT, e que se ajustam perfeitamente à situação do PA: (i) a evidência de a marca PL ser tão difusa nos DLP's que se pode considerar ter sido adoptada uma solução de não-marcação, evidenciando-se a quantificação apenas na posição imediatamente à esquerda do núcleo (que coincide com a primeira posição no SN de estrutura mínima da primeira geração, isto é, na categoria DET); (ii) os DLP's contendo dados do PtgL2 com evidências desta estrutura irregular provirem de estruturas simplificadas do PtgL1 (DET+N); (iii) os DLP's contendo dados do PtgL2 se revelarem fortemente influenciados por transferências provenientes da L1.

A quase inexistência em São Tomé de informantes tanto do PTL2<sup>177</sup> como do PAL2 impediram a recolha, para o segundo, de dados que permitissem confirmar *in loco* as hipóteses avançadas. A explicação fornecida em (ii) pode ser uma opção que tivesse funcionado como gatilho para impelir a marcação variável, mas, na ilha, o rácio entre europeus e africanos era desproporcionado, com evidente desvantagem numérica para os primeiros. Entrevistados por Baxter (2004:125, nota de fim de texto 38), três portugueses europeus, antigos feitores em roças de São Tomé, garantiram que os europeus, para efeitos de facilidade comunicativa, moldavam o seu discurso ao dos falantes de origem africana, levando a cabo simplificações nas estruturas do PE, traduzidas no uso de algumas formas analíticas (p.e. pronomes sujeito no lugar de pronomes objecto). Não obstante, os entrevistados fizeram questão de salientar que o grosso das divergências emergia no discurso dos africanos, e não dos europeus. Nesta conformidade, e apesar de não se descartar a possibilidade de a variação poder ancorar esporadicamente no PE, parece ser dado adquirido que a responsabilidade maior da sua emergência caberá a outras fontes, isto é, às línguas do substrato. Daí que as estruturas do tipo DET+N apresentem elevado grau de ausência de concordância no PA ou no PT, o que não sucede no dialecto HEL-Ba, que se encontra em situação de mudança no sentido de adquirir as regras padrão, e para o qual as análises, numa perspectiva sintagmática, revelam que as estruturas simplificadas (2 elementos) apresentam tendência para favorecerem a marcação PL (P. Andrade, 2003:142; Tieppo, 2003:6).



O argumento do desenvolvimento apontado em (i), ou seja, aquele que é activado pela não captação da evidência de marca de PL, aparenta ser bastante plausível. Face a estas ocorrências, o falante em situação de ASL constrói SN's de estrutura mínima com um item quantificador na posição à esquerda, ou seja, na primeira oportunidade com que se depara para inserir a marcação (cf. Kiparsky, 1972:195). Não obstante, aspectos relacionados com a demografia ou questões fonológicas, que se abordarão no capítulo 5, sugerem que se considere seriamente a possibilidade de ocorrerem transferências que emanam das L1. As gerações mais idosas de Almojarife são as que evidenciam ter recebido DLP's mais marcados pelo *continuum* da L2 antecessora, não constituindo excepção aos resultados amplamente documentados sobre aquisição de L2's que evidenciam a influência proveniente das L1's, sobretudo nos estádios rudimentares de desenvolvimento da *interlândia*.

Como se tem referido amplamente neste trabalho, Guy (1981a:3001-302) avançou a possibilidade de a CPL-var do PVB, que marca preferencialmente a posição pré-nuclear, ter origem nos escravos falantes de L1's africanas, nas quais os classificadores semânticos, e não os elementos nucleares, detêm as marcas de PL. Esta hipótese parece colher perfeitamente no caso do PA e foi também contemplada como sendo bastante válida nos estudos efectuados por P. Andrade (2003), para HEL-Ba, por Baxter (2004; 2009) para o PT e HEL-Ba, por Figueiredo (2008, 2009a, 2009b, 2010), para o PA, e por Jon-And (2008, 2009), tanto para o PMp como para o PCV. Como se viu no ponto 1.4.1 do presente trabalho, o kimbundu e o umbundu constroem a pluralização recorrendo a classificadores situados em posição pré-nominal e que remetem para classes semânticas, podendo os mesmos carregar ainda funções de carácter deictico. Por outras palavras: a maioria das línguas do grupo níger-congo atlântico é caracterizada por um sistema de concordância em número determinado pelo classificador do nome, que actua como prefixo deste e estende a concordância aos adjectivos, determinantes e mesmo alguns conectores, que mantêm o seu radical inalterado (exemplos [76a]) e [76b]). Deste modo, a indicação de PL no classificador nominativo funciona como a chave mórfica que dá o mote para a pluralização dos classificadores dos outros itens inseridos na cadeia sintagmática.

Face a estes aspectos, é muito natural que os falantes adultos de uma L1 do grupo níger-congo atlântico, como o balanta, o wolof, o kimbundu ou o umbundu, submetidos à ASL, se revelem sensíveis à marca de PL em posição imediatamente pré-nuclear, dando também continuidade, de certa forma, à aplicação do Princípio da Simplicidade, que leva a marcação a incidir inicialmente sobre as estruturas mais simples dos SN's. Como será também improvável que os mesmos falantes percepcionem que a marca de PL actua numa posição distinta nessa mesma L2, isto é, em forma de sufixo. As pesquisas levadas a cabo por Koehn (1994), no âmbito do projecto DUFDE (*Deustsh und Französisch – Doppelter Erstspracherwerb*) sobre aquisição bilingue das línguas alemã e francesa, dão conta que o fenómeno da aquisição de morfologia de género e número no SN sofre um processo de maturação (Borer & Wexler, 1987; Genesee, 1989; Radford, 1990, 1997a, 2000; Vainikka, 1993/1994; Meisel, 1994a, 1997) de acordo com o Modelo de Esquemas (Koehn, 1994:35, ponto 2.6.4.2 do presente trabalho), pelo que não existem apenas relações entre formas básicas e derivadas em termos de regras de afixação, ocorrendo igualmente interligações a nível fonológico e que levam os afixos a não serem armazenados independentemente dos radicais com os quais podem ocorrer. Como o estado FL0 da aquisição é composto apenas por conceitos semânticos, ou seja, caracteriza-se pela ausência sistemática de marcação de género e de número, os falantes em processo de ASL, para além de lidarem inicialmente com as palavras como um todo, reconhecem a estrutura interna das formas morfologicamente complexas por comparação com outras formas, recorrendo a conexões lexicais (cf. Koehn, 1994:41), ou seja, estabelecem ligações morfofonológicas entre a LA e a sua L1 africana, para elaborarem novas formações na interlíngua.

Baxter (2004:118), ao estudar o fenómeno da CPL-var no SN do PT, recorre a exemplos do seu *corpus* para confirmar a possibilidade de certos determinantes poderem ser fonologicamente reinterpretados como classificadores semânticos, como acontece com o artigo definido “o”, cuja realização fonológica é igual à do classificador semântico “u” da classe II do umbundu, referente a pessoas. Este é prefixado nominalmente com os afixos *u-*, no singular, e *a-*, no PL, mantendo-se inalterado o radical do nome: **u**~lume > **a**~lume (homem > homens) (Valente, 1964:68). Evidências do mesmo tipo de fenómeno estão presentes quer nos nossos dados quer em ocorrências

do PANG, em que o artigo definido foi fonologicamente aglutinado ao nome, como se de um marcador semântico se tratasse:

[264] PA:  
*U'omé do... de Benfica* [CASTEH1]  
“O homem do... do Benfica”

Constatações de que os falantes dos dialectos bantu interpretam o substantivo do português como invariável e os elementos não-nucleares do SN como equivalentes dos afixos bantu são-nos fornecidas pelos empréstimos lexicais incorporados nestas línguas. No umbundu, por exemplo, os termos provenientes do português caem dentro da classe dos afixos OMU–, que pluralizam em OMA– (pessoas) e AMU– (coisas): **omu**~kanda > **amu**~kanda (carta > cartas) (Valente, 1964:81). Chatelain (2001), por seu lado, regista, no kimbundu, uma série de vocábulos portugueses aos quais são aglutinados os afixos bantu, que podem não surgir na forma singular da palavra, mas marcam presença constante nas pluralizações para fornecerem a distinção na marcação de número. Listamos, seguidamente, alguns desses termos: **ki**~ngeleji > **i**~ngeleji (inglês > ingleses); pultukeji > **ji**~pultukeji (português > portugueses); mosa > **ji**~mosa (moça > moças); **ku**~mezai > **ji**~mesa (mesa > mesas); **di**~letá > **ma**~letá (leitão > leitões); **di**~kalu > **ma**~kalu (carro > carros).

Mas a aglutinação de afixos do kimbundu não se limita aos termos importados por este dialecto, registando-se também o mesmo tipo de fenómeno no PANG. Um exemplo concreto deste tipo de aglutinação acontece com alguns diminutivos do português, que registam a incorporação do afixo “ka-”, pertencente à classe X do kimbundu e que tem a função de construir os diminutivos neste idioma:

[265] PANG:  
*peixe pequeno come também peixe capiquinito.* (Cardoso, 1980:86)  
“o peixe pequeno come também o peixe pequenito”

Como se viu também no ponto 1.4.1 do presente trabalho, a concordância entre os elementos que constituem o SN é controlada nas línguas do ramo bantu pelo classificador pré-nominal e concordantes aglutinados em posição prefixal nos restantes

itens, daí resultando que o radical das palavras se mantenha inalterado em todos os elementos que compõem a cadeia sintagmática.

O tipo de estrutura frásica que conserva a marca de PL no afixo pré-nominal e concordantes aglutinados em posição prefixal nos restantes itens, mantendo os radicais inalterados, foi transposta para a *interlíngua* dos falantes das L1's bantu em situação de aquisição do PtgL2, determinando, numa fase inicial da aquisição, o uso de construções em que a pluralização era inserida pré-nominalmente no nome, mantendo-se o seu radical inalterado. Quanto aos restante itens do SN, apresentavam também o mesmo tipo de concordante do substantivo, aglutinado em posição prefixal, construindo-se estruturas em tudo idênticas às das línguas bantu. De facto, são comuns, em finais do século XIX, os registos que atestam o emprego generalizado no Brasil, entre os falantes de línguas bantu, do morfema /Zi-/ , como prefixo, quer a nível isolado quer a nível dos itens que constituíam a cadeia frásica. Segundo Ribeiro (1897:271), estes falantes das L1's bantu, ao exprimirem-se em português, repetiam “*por alliteração a particula prefixa inicial em todo o corpo da phrase*”, construindo cadeias frásicas do seguinte tipo:

[266] PTG:  
Z'ere z'mandou z'dizé (Ribeiro, 1897:271)  
“Ele mandou dizer”

A presença deste tipo de prefixo, na sua forma isolada, é ainda uma constante na fala de Almojarife, mesmo no registo das gerações menos idosas:

[267] PA:  
Ah! Z'omé disse ê recebi terra [ABILH2]  
“Ah! Os homens disseram que eu recebi terras”

Observando as ocorrências registadas nos exemplos [264], [266] e [267], parece resultar evidente um claro cruzamento entre funções sintácticas e semânticas desempenhadas pelo artigo definido do português **o** > **os** e o referido marcador semântico (com funções sintácticas de marcação PL) do kimbundu **ji-**, e que se traduz na amalgamação **zi-** e sua posterior realização fonológica em **z'-**. Contudo, Bonvini (2008:50) chama a atenção para se notar que, “*malgrado sua semelhança, o morfema Zi- [~zi] dos “negros do Brasil” não coincide com o classificador ji- ‘plural’ da classe 10 do quimbundo, que era sua função inicial. Em seu emprego recente, esse morfema*

*não é mais uma marca de classe, mas um morfema plurifuncional, suscetível de intervir em diversos níveis de estrutura gramatical de uma maneira generalizada: nomes, sintagmas de determinação, enunciados*". De facto, o pressuposto de Bonvini faz todo o sentido se levarmos em linha de conta que, segundo o princípio do Constrangimento do Núcleo Funcional (*Functional Head Constraint*), a alternância de códigos não pode ser aplicada entre um núcleo funcional, no caso a categoria DET, e o seu complemento (Belazi, Rubin & Toribio, 1994).

Independentemente destes aspectos, dúvidas parecem não subsistir que, relativamente quer ao tipo de estrutura frásica representada no exemplo [265] quer às que revelam ausência de redundância na marcação e que levam à fixação da pluralização unicamente no primeiro elemento pré-nominal (cf. Kiparsky, 1972:195), mantendo-se os radicais de todos os itens gramaticais da cadeia frásica inalterados, a sua realização ficar-se-á a dever à influência das L1's africanas (cf. Guy, 1981a:301-302), sobretudo do grupo níger-congo atlântico (Baxter, 2009:292).

Lopes (2001:183), por seu lado, levanta a hipótese de a concordância nominal de número entre os itens do SN poder ancorar na Teoria dos 4 Morfemas, avançada em trabalhos sobre estruturas mistas de afasias bilingues, aquisição e *code-switching* bilingue, levados a cabo por Myers-Scotton (1997; 2001) e Myers-Scotton & Jake (2000a; 2000b). Estas autoras demonstram que os morfemas de conteúdo (*content morphemes*) de uma língua podem ser encaixados em estruturas morfossintáticas de outro idioma e que, quando esse encaixamento acontece, a estrutura mista que actua como matriz pode também conter morfemas sistémicos do tipo *early*, isto é, morfemas intrínsecos ao significado do morfema de conteúdo encaixado e que são gerados conjuntamente com esse encaixamento. Convém aqui esclarecer que as diferenças do modo como e momento em que os morfemas são adquiridos surgem reflectidas nas realizações das estruturas de superfície (Myers-Scotton, 2001:243). Desta forma o Modelo dos 4 Morfemas contempla morfemas de quatro classes, isto é, morfemas de conteúdo e três tipos de morfemas sistémicos (*early systeme morphemes*, *bridge system morphemes* e *outsider system morphemes*). As diferentes classes de morfemas são activadas em níveis distintos da produção linguística, com os *early system morphemes* actuando a nível do léxico mental (o nível *lema*) como um tipo de satélite dos morfemas

de conteúdo e fornecendo a estes significados específicos, como acontece com a marca de PL –s ou a referência de definitude. Por seu lado, os *late system morphemes* surgem relacionados com os requisitos de carácter estrutural (Myers-Scotton & Jake, 2000a). Desta forma, iniciando-se a produção linguística a um nível pré-linguístico, com a activação de conjuntos de traços semântico-pragmáticos linguisticamente específicos, são primeiramente seleccionados os morfemas lexicais (*content morphemes*). O sentido destes, por seu lado, é complementado pelos morfemas gramaticais (*system morphemes*), responsáveis pela “*definição das estruturas gramaticais superficiais*” (Lopes, 2001:93). Os primeiros morfemas gramaticais a serem activados são os *early system morphemes*, em simultâneo com os morfemas de conteúdo (os lemas), para darem resposta a necessidades funcionais imediatas, ou seja, contemplarem a intenção comunicativa do falante. Assim, no caso concreto da CPL-var dos diferentes dialectos que temos vindo a abordar, a oposição *early system morphemes versus late system morphemes* ganha pertinência em virtude de a mesma, ao que tudo indica, surgir conectada à possibilidade de alguns morfemas de PL serem percebidos de forma mais básica, ou seja, configurarem *early system morphemes*, enquanto outros se opõem a eles em virtude do seu carácter ser meramente estrutural, isto é, consubstanciarem-se como *late system morphemes*. Como tal, estes últimos, por serem tardiamente apreendidos, são alvo de maior variação.

O tipo de transferências detectado por Myers-Scotton (1997; 2001) e Myers-Scotton & Jake (2000a; 2000b) é bastante comum em estádios iniciais de aquisição, tanto de L1’s como de L2’s. Os nossos dados revelam isso mesmo no caso das segundas, já que se detecta uma forte presença do padrão predominante da marcação PL das L1’s africanas nos DLP’s da fala das gerações mais idosas, quer das variedades africanas quer das variedades brasileiras de português. Nesta perspectiva, a “*transferência na aquisição da L2 rudimentar dos adultos compreenderia, em parte, um grau de fabricação de uma estrutura morfosintáctica envolvendo encaixes da L1 e morfemas de conteúdo da L2, especialmente em situação na qual o estágio inicial de aquisição de uma L2 se dá com forte presença da L1*” (Baxter, 2004:119).<sup>178</sup> E nos casos em questão, os falantes das L1’s do grupo níger-congo atlântico, ao lidarem fonologicamente com a estrutura DET+N como um todo fonoprosódico na fase inicial da aquisição da LA (cf.

DeGraff, 1999b:482), não só não percebem a segmentação entre as diferentes sequências morfológicas desta como também não estabelecem a correspondência entre as características formais e as formas morfofonológicas dos itens em questão (cf. Prévost & White, 1999 – MSIH). Desta forma, os classificadores nominais das L1's do grupo níger-congo atlântico poderão ter sido associados aos determinantes do português, concentrando as marcas de PL neste, ou seja no primeiro elemento do SN, mais concretamente no item imediatamente pré-nuclear, em detrimento das demais pluralizações na cadeia sintagmática. Este aspecto revela não apenas que este elemento configura um *early system morpheme* gerado conjuntamente com os morfemas de conteúdo mas igualmente que ele é acompanhado pela não-aquisição de material morfosintático nos outros itens do SN, e que se detecta quer nos CP's quer nas variedades reestruturadas de português.

Dando sequência a este entendimento, a inserção de pluralização na posição imediatamente pré-nuclear da geração mais nova pode então reflectir o intrometimento de morfemas de conteúdo do português na configuração estrutural da L2 rudimentar dos falantes adultos em situação de aquisição do português, especialmente em estruturas nominais que transitaram geracionalmente via DLP's até à FE-1. Tais configurações nominais requeriam marcação PL imediatamente à esquerda, posição em que apareceria o classificador nominal do grupo níger-congo atlântico, tendo pertencido às gerações seguintes a responsabilidade de reforçar este traço, assimilando-o como parte integrante no desenvolvimento da estrutura sintagmática quer do PA quer do PT, quer ainda do dialecto de HEL-Ba, isto é, do PVB. Assim, na gramática dos falantes bilingues do português e de outras variedades de origem africana, a posição adjacente à esquerda do núcleo, isto é, a categoria DET, foi assumida como o ponto da marcação PL *singleton* (Baxter, 2009:292), mesmo em sintagmas de configuração desenvolvida. E no caso da pluralização dos outros elementos que compõem o SN, o que se nota é que a mesma tem apenas a função de orientar gramaticalmente a concordância, sendo esta possível graças à intervenção dos *bridge late system morphemes*. Nesta conformidade, a explicação de Baxter (2004:119) é complementada pela descrição avançada por P. Andrade (2003:79) sobre a forma como se estruturam as línguas crioulas: “o suporte gramatical é tomado das línguas de substrato; os morfemas de conteúdo são predominantemente da língua

*alvo e estes podem sofrer processos de gramaticalização de modo a suprir as necessidades da estrutura abstracta; os early system morphemes do superstrato podem ser absorvidos por essas línguas; e os late system morphemes não passam para as línguas crioulas*". E não sendo todos os morfemas de PL activados ao mesmo nível, a tendência é para o núcleo e elementos pós-nucleares inibirem fortemente a marcação, já que a posterior activação dos *late system morphemes* é dispensada em virtude de estes configurarem unicamente uma "*questão estilística consciente*" (P. Andrade, 2003:79-80). No entanto, note-se que, os *late system morphemes* actuam mais nos itens flexionáveis à esquerda do núcleo do SN e menos nos elementos à direita deste, configurando-se, assim, a aplicação do Princípio da Integração (Lucchesi, 2000a:250).

O debate sobre a intervenção da GU na ASL tem dirigido a sua atenção para a questão de os falantes em processo de aquisição restabelecerem ou não os parâmetros instanciados na aquisição da L1 (ponto 2.4 do presente trabalho). Neste aspecto, as teorias vão desde as mais radicais até às mais moderadas. Assim, enquanto alguns especialistas entendem que a GU não está disponível para actuar na ASL, visto ter desempenhado já o seu papel na aquisição da L1 ou aquisição bilingue (Clahsen & Muysken, 1986; Muysken, 2001), outros defendem uma posição totalmente oposta, assente na FTFA e advogando que a GU actua de forma igual na aquisição da L1 e ASL (Cook, 1994; White, 1989, 2003; Bruhn de Garavito & White, 2002; Montrul, 2004; Cook & Newson, 2007 [1988]). Mais moderados são os defensores de que a GU actua na L2 através da L1 (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), mas de forma parcial. Assim, os erros tardios produzidos pelos falantes de L2's revelam fossilizações provenientes da falha no restabelecimento de parâmetros (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clahsen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto as categorias funcionais deixarem de estar disponibilizadas depois do período crítico do desenvolvimento. De facto, e ao que tudo indica, a capacidade inata que permite à criança monolingue ou bilingue discriminar os sons do *input* a nível dos padrões trocaicos e jâmbicos, filtrando e reconhecendo aqueles que pertencem ao(s) sistema(s) linguístico(s) em processo de aquisição, parece desaparecer em fase pós-adolescência. Exames sintácticos feitos com recurso a orações relativas (sujeito *versus* objecto directo), em crianças que eram surdas e recuperaram a



audição tardiamente, parecem confirmar este aspecto, visto terem revelado que a aquisição tardia da L1 afecta a capacidade fisiológica de fazer uso da variação rítmica, a qual permanece activa apenas até aos seis meses de idade (Kuno, 1974). Assim, a partir desta fase, o falante em processo de aquisição deixa de prestar atenção às subtilezas da língua, perdendo, conseqüentemente, a sensibilidade às alterações dos padrões de som. Deste modo, aprendentes de L2's em idade pós-puberdade, ao entenderem diferentes segmentos fráscicos (p.e. DET+N) como sequências fonológicas indistintas (cf. DeGraff, 1999b:482), realizam construções cruzando sintaxe da L1 e itens lexicais da L2, isto é, formas morfofonológicas da L2 em especificações de traços da L1. Prévost & White (1999) vão mais longe nas suas apreciações, defendendo que, devido à actuação do fenómeno da MSIH, os aprendentes de uma L2, ainda que continuamente expostos à mesma, não conseguem adquirir as mesmas representações sintácticas dos falantes nativos, isto é, não conseguem alcançar o grau máximo (*ultimate attainment*) do desenvolvimento da interlíngua e que coloca a L2 adquirida no mesmo patamar de aquisição da LA pelos falantes nativos desta. Desta forma, os falantes da L2, ainda que adquiram competência no que concerne às propriedades sintácticas distribucionais dos itens funcionais, poderão não conseguir reproduzi-las sistematicamente, em virtude de não conseguirem fazer corresponder as características formais às formas morfofonológicas, uma vez que não adquirem os traços gramaticais abstractos da L2, que são diferentes dos da L1. E se quiséssemos buscar realizações para apoiar, com exemplos reais, as sustentações da FFFH e MSIH, verificaríamos que os exemplos [264] e [267] assentam aqui perfeitamente.

Para Koehn (1994), o estágio FL0 da aquisição é caracterizado pela ausência de marcação de género e de número,<sup>179</sup> muito embora os conceitos semânticos já estejam presentes no mesmo, isto é, os processos que instanciam o processamento da flexão inerente já estejam disponíveis no estágio inicial do desenvolvimento da interlíngua, permitindo que a marcação do género e número dos nomes se inicie no estágio 2 do desenvolvimento da interlíngua (Plag, 2008a:124 – Tabela 3.1). Porém, as noções gramaticais de número e de género (a flexão funcional) só são desenvolvidas posteriormente, numa fase em que a gramática do falante já está mais amadurecida (cf. Koehn, 1994; Müller, 1994a),<sup>180</sup> levando a que a concordância no interior do SN se

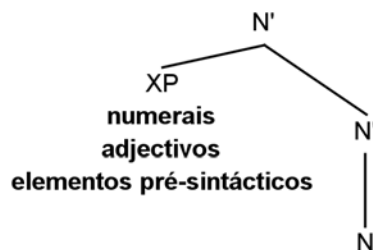
inicie apenas no estágio 3 do desenvolvimento da interlíngua (Pienemann & Johnson, 1987; Plag, 2008a:124 – Tabela 3.1), estágio esse que não é ultrapassado pelos crioulos (nem pelo PA), por razões que se prendem, essencialmente, com aspectos psico e sociolinguísticos (Plag, 2008b). Na óptica minimalista de Chomsky (1996 [1995], 2001a), o género e o número são tidos como traços de categoria  $\emptyset$  (traços de concordância que incluem informação necessária para a interpretação semântica) presentes no núcleo do SN e que estabelecem uma relação com outros traços correspondentes localizados algures na cadeia sintagmática. Em português, como sabemos, o número é indicado normalmente pela terminação /s/, concordando o adjectivo e o determinante com o núcleo do SN. Todavia, como o traço da pluralização não é interpretável (Franceschina, 2002:76), pode sofrer apagamento (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashes, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), como de facto acontece no núcleo sintagmático do PtgL2 adquirido por falantes adultos de L1's do grupo níger-congo atlântico, originando então a concordância variável na *interlíngua*, isto é, a falta de correlação de traços correspondentes do SN, como referimos na análise dos valores da Tabela 21.7. Assim sendo, verifica-se que o número pode ser adquirido isoladamente, como acontece no quechua de Cuzco, o que já não pode suceder com a concordância de número.

Tendo confirmado que, nos nossos dados, a esmagadora maioria dos SN's produzidos pela FE-3 é de apenas dois ou três elementos, optámos por observá-los sob a perspectiva generativista que inclui a categoria funcional NUM localizada entre o DET e o SN, aí se situando também os traços de número (Picallo, 1991; Carstens, 1991; Ritter, 1991; Valois, 1991; Bernstein, 1993; Cinque, 1994; Montrul, 2004). Müller (1994a), ao defender que a aquisição do alemão e do francês como L1's em bilingues sofre um processo de maturação, concorda com Koehn (1994) no que concerne ao facto de as gramáticas das crianças não apresentarem marcação de género ou número no estado FL0, sendo estas noções gramaticais disponibilizadas pela GU num estágio posterior de desenvolvimento gramatical. Nesta conformidade, a autora, a exemplo dos que defendem a localização do traço de número na categoria funcional NUM, parte do pressuposto que as gramáticas adultas possuem pelo menos uma projecção funcional

acima do SN, no qual se encontram alojados os traços flexionais dos nominais que não estão disponibilizados na fase FL0 da aquisição.

Debruçando a sua atenção sobre os artigos definidos e indefinidos produzidos pelas referidas crianças bilingues, Müller (1994a) procura delinear como ocorre o processo de integração dos traços flexionais na gramática destas. Como os falantes em início de aquisição não apresentam marcas de género e número, a autora entende que se está perante a evidência de que estes não possuem S<sub>DET</sub>, ou que, em contrapartida, a sua gramática contém uma categoria funcional subespecificada durante algumas etapas do seu funcionamento. Verifica-se então que a hipótese da maturação é levada ao extremo por Müller (1994a), a ponto de o próprio DET estar sujeito a desenvolvimento geracional. E assim sendo, como os traços gramaticais de género e número só são disponibilizados em fase posterior à FL0, os falantes usam o artigo indefinido como numeral, sendo a ausência de DET nas suas estruturas sintagmáticas atribuída à indisponibilidade dos referidos traços. Esta, por sua vez, tem como consequência a não-projecção da categoria funcional.

Face aos seus postulados, Müller (1994a:62) propõe uma representação do SN na fase FL0 da aquisição próxima da gramática-alvo e em que os itens em XP são mutuamente excludentes:



**Fig. 34.** Estrutura do SN na FL0 da aquisição.

Quanto aos “elementos pré-sintácticos”, dizem respeito a propriedades do objecto em referência, e não a propriedades da categoria sintáctica do N, o que origina a falta de concordância. Desta forma, a integração de DET na gramática ocorre posterior e faseadamente, despoletando-se no momento em que os falantes passam a entender a categoria de artigos indefinidos como distinta da dos numerais e evidenciando-se o seu aparecimento não só através do uso gramatical de género e número nos definidos,

indefinidos e possessivos mas também pela distinção de número em nomes e SN's complexos (DET+N+ADJ).<sup>181</sup>

Porém, a perda da capacidade de discriminar determinados sons em fase pós- puberdade leva a que se entendam palavras distintas (DET+N) como sequências fonológicas que constituem um todo indistinto (Koehn, 1994:49; DeGraff, 1999b:482; Faria Freitas & Miguel, 2001:53), ou seja, como uma unidade plurilexemática que, sendo caracterizada pela cristalização estrutural e semântica, é armazenada como um todo na mente dos falantes (Ribeiro, 2009:84). A falta de noção de que o nome já se encontra marcado em número faz reunir as condições para que os falantes em processo de aquisição na idade pós-puberdade transfiram o sistema de marcação da L1 (Andersen, 1983c), feito com recurso a afixos semânticos, visto este tipo de marcação ser inexistente na LA e configurar-se o traço *a*. das Tabelas 5.1 e 6.1 (Plag, 2008b:311). Desta forma, os falantes realizam formas morfofonológicas da L2 em especificações de traços da L1 (DeGraff, 1999b:482), isto é, construções relexificadas que cruzam sintaxe das L1's do grupo níger-congo atlântico e itens lexicais do PtgL2, como as retratadas nos exemplos [264], [265], [266] e [267]. Por outras palavras, os falantes adultos em situação de aquisição do PtgL2 fixam fonologicamente o determinante da LA no nome da *interlíngua*, de acordo com o verificado, por exemplo, no romeno (Fig. 35).

- [268] ROMENO:  
 a. *un lup*  
     *um lobo*  
 b. *lup~ul*  
     *lobo~o* (Brito, 2003b:501)
- [269] ROMENO:  
 a. *un frumuas lup*  
     *um formoso lobo*  
 b. *lup~ul frumuas*  
     *lobo~o formoso* (Brito, 2003b:501)

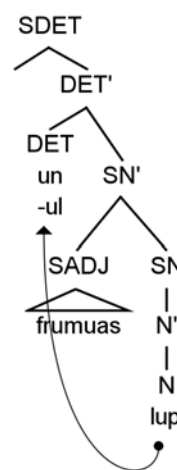


Fig. 35. Romeno: movimento de N e conseqüente fixação de DET (a partir de Brito, 2003b:501).

De facto, neste idioma ocorre contraste entre o uso dos determinantes definidos e dos determinantes indefinidos (exemplos [268a] e [268b]), já que os primeiros evidenciam

que pode ocorrer um movimento do nome para se fixar neles, levando-os a funcionar como seus enclíticos (exemplos [269a] e [269b]) (Brito, 2003b:501). Este movimento (Fig. 35) estabelece, então, semelhanças com o movimento do verbo para a flexão (ponto 2.11.3.3 do presente trabalho).

Para além dos aspectos referidos, os falantes adultos de L1's do grupo níger-congo atlântico em aquisição de PtgL2 levam igualmente a cabo o apagamento dos traços não-interpretáveis do substantivo, de acordo com o preconizado pela FFFH. Como as sequências não-marcadas não implicam o mapeamento directo nem requerem nenhum armazenamento para iniciarem o processamento sintáctico (Pienemann, Biase & Kawaguchi, 2005), não só a categoria DET deixa de ser percebida, e consequentemente de sofrer desenvolvimento maturacional normal, como também a noção gramatical de número, que deveria ocorrer posteriormente, deixa de ser disponibilizada (Koehn, 1994; Müller, 1994a). Assim, e apesar de as motivações psicolinguísticas e sociolinguísticas sofrerem um incremento considerável (a LA é vista como idioma de prestígio e fornece perspectivas de ascensão socioeconómica), o desenvolvimento moderado da transferência (Pienemann *et alii*, 2005a, 2005b; Pienemann & Håkansson, 2007) continua comprometido, pelo que os afixos flexionáveis de valor semântico expresso, ainda que deixem de sobreviver com esta função devido à pressão de cima para baixo (Plag, 2008b), continuam a operar de forma plurifuncional, intervindo de maneira generalizada ao nível das diversas estruturas gramaticais (Ribeiro, 1897:271). Daí resulta que tanto a noção de distinção fonológica entre determinante e nome como o apagamento do morfema de PL do nome não sejam totalmente recuperados para a sintaxe e esta, ao estabelecer as relações de concordância, depois de se juntarem os elementos que compõem a frase, face à amalgamação pré-nuclear do determinante no nome que determina a ocorrência de duas categorias na posição de especificador/núcleo, deixa definitivamente de perceber a concordância por não ser morfologicamente visível, originando-se o licenciamento de {s} nulo e que determina que a variação seja caracterizada pela presença de número apenas no morfema *singleton* do determinante proclítico no nome. Para as estruturas que envolvem o uso de possessivos e adjectivos pré-nucleares, tendentes a receberem a marca de PL, vale o que foi já referido anteriormente a propósito da correlação entre posição e definitude (ponto

2.9 do presente trabalho), de acordo com a estrutura patente na Fig. 26. Assim sendo, adquirem a categorização de núcleo e são marcados com morfema de número, caso estejam antes do nome. Se, pelo contrário, ocorrerem pós-nuclearmente, detêm a categoria de sintagma, sendo a sua inibição originada pela ocorrência do morfema *singleton* que fixa a pluralização unicamente no elemento DET.

No entanto, convém lembrar aqui lembrar que alguns generativistas, defensores da FTFA, entendem que estas estruturas são passíveis de regularização em estádios mais avançados de aquisição da L2 (Franceschina, 2001; Bruhn de Garavito & White, 2002; White *et alii*, 2004). Contudo, é também de considerar que os iniciadores da variedade reestruturada de Almojarife continuavam incapacitados de analisar os artigos definidos (cf. Kihm, 2008:433), já que estes não marcavam presença no santomense e revelavam ausência de *saliência fónica* na língua lexificadora. Estes aspectos, associados à falta de motivação inicial para aprender a LA numa situação de transmissão linguística irregular em que o contexto socioeconómico era extremamente precário e desfavorecedor para os falantes do PtgL2, fez com que as gerações mais novas de Almojarife apenas levassem a cabo uma ligeira expansão da regra de concordância de número nos itens pré-nucleares em posição não imediata, ou seja, nos *late system morphems*, confirmando que o apagamento dos traços não-interpretáveis ocorrido na *interlíngua* que lhes foi irregularmente transmitida como L1 pelos seus antepassados se cristalizara, visto estes terem sido incapazes de restabelecer os parâmetros que lhes disponibilizariam as categorias funcionais, tal como preconizam os defensores da FFFH. Efectivamente, e de acordo com os princípios generativistas, a fossilização resulta do facto de os parâmetros da L1 não se ajustarem a determinados parâmetros da LA (Flynn, 1989:107), que não são irradicados na ASL, instanciando-se uma nova parametrização na L2. Como esta instanciação não é posteriormente alterada, “fossiliza” os novos/errados parâmetros na L1 nativizada (Gonçalves, 2004:235), pelo que a fossilização é uma consequência da dificuldade que os aprendentes mais idosos de Almojarife revelaram na refixação dos parâmetros na L2 (cf. Hale, 1998:33).

Embora as fossilizações possam também configurar meras estabilizações, isto é, serem temporárias (Selinker, 1979; Sims, 1989), a verdade é que a sociohistória que determinou a aquisição e reestruturação do PtgL2 em Almojarife, bem como a sua

consequente nativização, propiciou também a conjugação de factores sociais, psicológicos e interactivos no sentido de impedirem que determinadas formas transferidas (cf. MEP, Selinker & Lakshamanan, 1993) se desfossilizassem (ponto 2.10.1 do presente trabalho), tornando-as permanentes (Hale, 1998:33). Assim, a observação geracional do perfil de marcação PL no SN do PA mostra que, em virtude de as gerações primordiais do PA terem perdido os traços virtuais não especificados das categorias funcionais, as faixas etárias mais novas, sobretudo as menos escolarizadas, não conseguem suplantar as variações fossilizadas que herdaram e apenas se limitam a expandir a concordância referente aos morfemas de carácter estrutural adquiridos mais tardiamente (os elementos pré-nucleares não imediatos), quando começam a elaborar SN's mais complexos, devido à pressão sociolinguística sofrida. E partindo do pressuposto que determinadas características das variedades criouliizantes configuram instanciações inatas de parâmetros por defeito dessas mesmas propriedades (ponto 2.11.3.2 do presente trabalho), a questão da variação no SN tem que ser igualmente entendida como resultante da alta dependência em relação aos traços idiossincráticos das línguas lexificadoras, ou seja, que não há instanciação de parâmetros por defeito neste domínio (Kihm, 2008:434). Refira-se ainda que, em trabalho comparativo sobre a concordância variável do PA e de HEL-Ba, Figueiredo (2009a:50) chegara a idênticas conclusões no que concerne à questão do restabelecimento dos parâmetros para aplicação da regra do género, pelo que se confirma que o que as faixas etárias mais novas de Almojarife conseguem fazer é apenas perpetuar determinadas características da L2 na L1 nativizada, reproduzindo perfis sintagmáticos que configuram criouliização leve e cujas CGEN-var e CPL-var são praticamente idênticas às já previamente desenhadas pelas suas gerações antepassadas.

Ainda assim, e face ao anteriormente exposto, confirma-se o pressuposto que levantámos na hipótese 7, constituída para a variável independente *posição linear*, isto é, que o primeiro lugar da cadeia sintagmática, com especial relevância para a posição imediatamente pré-nuclear, é o que mais favorece a inserção de marcas, revelando-se as outras posições como tendentes a marcar menos a pluralização, em virtude de a informação semântica já ter sido atribuída (Kiparsky, 1972:195) e de as gerações mais novas do PA, sujeitas a DLP's deficitários, não recuperarem os traços virtuais não

especificados das categorias funcionais que não foram detectados pelas gerações mais antigas (cf. Hawkins & Chan, 1997; Prévost & White, 1999). Nesta conformidade, também se descarta completamente a possibilidade de o sistema de inserção de marcas no SN ser guiado pelo Princípio do Processamento Paralelo (Scherre, 1988:208), isto é, que marcas levam a marcas e zeros conduzem a zeros. Paralelamente, a grande semelhança registada no padrão de variação da marcação dos itens flexionáveis em número do SN de sete variedades de português (PA, MRJ, NURC, HEL-BA, PT, PMp e PCV) vem sustentar fortemente a evidência de que, em situação de transmissão linguística irregular, o português adquirido por falantes de um CP manifesta os mesmos padrões de variação do português adquirido quando a língua ancestral não é um CP (Figueiredo, 2008; 2009a).

No presente capítulo analisaram-se os resultados da variável dependente do tipo binário por nós estabelecida e de 6 variáveis independentes do tipo estrutural, sendo estas observadas de acordo com a escala hierárquica decrescente das variáveis independentes consideradas como válidas, pela ferramenta VARBRUL, para estudo da CPL-var no SN do PA: *posição em relação ao núcleo do SN, ordem do item na cadeia dos constituintes flexionáveis, classe gramatical, marcas precedentes ao item analisado, grau de concordância de número no SN e posição linear* (cf. Tabela 14.1). Algumas destas variáveis foram cruzadas entre si, bem como com a variável extralinguística *idade*, a fim de se determinar, com mais exactidão, quer o ponto de inserção das marcas de PL no SN do PA quer o desenvolvimento geracional da regra de concordância PL neste. As análises por nós efectuadas confirmaram na totalidade as hipóteses 1, 3, 10 e 7, colocadas no capítulo 3, e parcialmente a hipótese 4, levantada no mesmo capítulo. Estas observações dão suporte aos aspectos que levaram à não confirmação da hipótese 2, isto é, a hipótese que advogava como válido o Princípio do Processamento Paralelo (Scherre, 1988:208) na inserção ou inibição das marcas de PL no SN do PA. Pelo contrário, a CPL-var registada ao nível do seu SN e de mais quatro variedades africanas de português (PMp, PANG, PT e PCV), bem como de uma variedade brasileira de português (HEL-Ba), todas elas adquiridas por transmissão linguística irregular será, sim, motivada pelo sistema de marcação PL das línguas africanas que subjazem, de forma directa ou ancestral, às mesmas (cf. Guy, 1981a:301-302). Assim, e embora Scherre (1988:208),



para o MRJ, e Lopes (2001:212), para a NURC, advoguem a intervenção do processamento paralelo na questão da inserção/inibição das marcas de PL nestas duas variantes de PB, que terão um perfil idêntico ao do espanhol (Poplack, 1980a), a conclusão das análises que levamos a cabo neste capítulo dificilmente ajudarão a sustentar cientificamente estes pressupostos, até porque, como foi referenciado, alguns dos dados destas autoras terão sido incorrectamente codificados ou indevidamente inseridos em grupos de factores levados às rodadas VARBRUL. Paralelamente, a análise do desenvolvimento geracional da regra de concordância PL no SN do PA contraria o pressuposto de a não-realização da marca formal de PL poder ser justificado com recurso ao fenómeno da deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a), já que esta determina um apagamento diacrónico do *-s* morfémico PL ao passo que o que sucede no PA é um fenómeno completamente oposto, ou seja, a inserção geracional do referido morfema.

De todas as variedades de português observadas, resultou também evidente que o PA é aquela que se encontra em estágio menos avançado no que concerne à aquisição da regra de concordância PL no SN. Todavia foi confirmado que o padrão de variação neste é idêntico em todas as variedades, independentemente de os seus ancestrais falantes em situação de aquisição do PtgL2 terem tido ou não um CP como língua ancestral. Deste modo, o desenvolvimento da concordância PL inicia-se, em todas elas, com a introdução de um PL *singleton* na categoria funcional DET, que actua como a âncora que irá controlar, depois, a pluralização. Se essa posição não for ocupada por um elemento DET, então será o SN a tornar-se alvo de PL, devido às suas características lexicais. Em virtude de a inserção ocorrer com um morfema do tipo *singleton*, não só ocorre licenciamento de {s} nulo como a concordância se torna também parcial, sendo atribuída a partir da posição DET, isto é, a partir do elemento pré-nuclear adjacente. Assim, o desenvolvimento do PL conserva as características de concordância parcial, através de morfemas dissociados, que não são uniformemente atribuídos aos elementos mais próximos do *fulcro*. A ordem de atribuição, quando ocorre, contempla os elementos do SDET, primeiro, e os itens à direita do núcleo, depois. Não obstante, o perfil da âncora pré-nuclear conserva-se geracionalmente.

No próximo capítulo iremos analisar os resultados quer das variáveis independentes do tipo semântico e fonológico quer das variáveis extralinguísticas do PA.



## CAPÍTULO 5

*“Os outros passam a escrita a limpo. Eu passo a escrita a sujo.  
Como os rios que se lavam em encardidas águas. Os outros  
têm caligrafia, eu tenho sotaque. O sotaque da terra.”*

(O Barbeiro de Vila Longe)

(In, Mia Couto, *O Outro Pé da Sereia*, p. 271)

### **Análise dos resultados**

#### **(Variáveis independentes do tipo semântico e fonológico; variáveis extralinguísticas)**

No presente capítulo iremos observar, analisar e comentar os resultados quer das restantes variáveis linguísticas quer das variáveis extralinguísticas por nós constituídas e discriminadas no capítulo 3. Subdividiremos o capítulo em três partes, cada uma delas dedicada a um grupo distinto de variáveis que compartilhem traços. Deste modo, foram constituídos três agrupamentos de variáveis, de acordo com os seguintes critérios:

- (i) Uma variável independente do tipo semântico:
  - 1 – *Traço semântico do sintagma nominal (Animacidade do nome);*
- (ii) Duas variáveis independentes do tipo fonológico:
  - 1 – *Saliência fónica (subdividida em Processos morfofonológicos de formação de plural e Tonicidade dos itens lexicais singulares);*
  - 2 – *Contexto fonológico posterior;*
- (iii) Quatro variáveis extralinguísticas:
  - 1 – *Escolaridade;*
  - 2 – *Idade;*
  - 3 – *Estadia (Permanência fora da comunidade);*
  - 4 – *Sexo.*

Os resultados serão analisados de acordo com a metodologia aplicada no capítulo anterior e descrita no início deste, respeitando-se, em cada agrupamento, a ordem decrescente de importância das variáveis na marcação PL dos itens do SN do PA determinada pela ferramenta VARBRUL e apresentada na Tabela 14.1.

## 5.1. Variável independente do tipo semântico

### 5.1.1. Variável independente *traço semântico do SN (Animacidade)*

Os resultados da variável independente *traço semântico do SN* estão patentes na Tabela 22.1:

**Tabela 22.1.** Efeito do *traço semântico do SN* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood</i> :-981,009		<i>Significância:</i> 0,011
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
[+humano]	451/748	60	32	0,562
[-humano] [+animado]	290/79	37	3	0,277
[-animado]	721/1.511	48	65	0,482
<b>Totais:</b>	1.201/2.338	51	100	-

Os dados de Scherre (1988:272) revelaram a tendência de a marcação PL sobre os nomes referentes a pessoas (pr. 0,55) ser superior à dos nomes comuns referentes a coisas (pr. 0,45), um aspecto que a Tabela 22.1 parece confirmar. De facto, no PA ocorre tendência para se marcarem mais os itens de traço [+humano] (pr. 0,562), que se revelam ligeiramente favorecedores à inserção da pluralização, em detrimento dos itens de traço [-humano]. Contudo, entre os elementos de traço [-humano], ambos desfavorecedores da marcação, verifica-se que os itens de traço [-animado] propiciam mais a marcação (pr. 0,482) do que os elementos respeitantes a seres de traço [+animado] (pr. 0,277). Não obstante, este aspecto é entendível, já que os estudos sobre CPL-var têm confirmado a tendência para a inibição incidir fortemente sobre os itens colocados pós-nuclearmente. Ora, esta é a posição em que surgem os adjectivos substantivados do tipo *macho/fêmea* ou *homem/mulher*. Assim sendo, poderá estar encontrada uma das justificações para a utilização destes itens na forma invariável, isto é, sem flexão de número. Portanto, não surpreende que sejam os itens de traço [-humano] [+animado] a revelarem também o peso mais baixo na inserção de marcas, situados num patamar de inibição acentuada. Nesta conformidade, é possível concluir que a inibição não ocorre pelo facto de estes elementos representarem uma certa categoria semântica, mas sim por estarem dependentes de determinadas posições lineares na cadeia sintagmática, a qual vai originar a não inserção de marcas nos

mesmos. Portanto, é possível prever que o desenho da CPL-var do PA será determinado mais por condicionalismos estruturais do que por questões de carácter semântico.

Seguidamente apresentam-se os valores referentes à influência na marcação PL motivada pela categoria semântica, tanto para o PA como para o MRJ (Scherre, 1988:272) e para o PCV (Jon-And, 2009:7):

**Tabela 22.2.** *Traço semântico do SN: contribuição dos factores individuais para a marcação PL do elemento analisado: 3 variedades de português.*

Factores	PA		MRJ		PCV	
	Pr.	%	Pr.	%	Pr.	%
[+humano]	0,562	60	0,55	62	0,60	89
[-humano] [+animado]	0,277	37	0,33	36	(9/11)	82
[-animado]	0,482	48	0,44	52	0,41	77

Como se pode ver, existe uma analogia total entre os resultados do PA, do MRJ e do PCV, apesar de não se possuir o peso relativo do factor [-humano] [+animado] referente à última destas variedades de português. Deste modo, confirma-se um padrão de marcação bastante semelhante no que concerne às motivações da categoria semântica que desenharam a CPL-var dos três dialectos. Assim, nos dados do MRJ e do PCV, os itens com traço [+humano] apresentaram também índices que favorecem levemente a marcação (pr. 0,55; para o primeiro, e 0,60, para o segundo), enquanto os elementos de traço [-humano] da variedade brasileira se quedam pela inibição. Esta, tal como no PA, é bastante suave nos itens de traço [-animado] tanto do MRJ (pr. 0,44) como do PCV (pr. 0,41), mas mais acentuada nos elementos que categorizam o traço [+animado] da variedade carioca (pr. 0,33).

Scherre (1988:272), devido à distinção ocorrida essencialmente entre as categorias de traço [+humano] e [-humano], testou uma rodada em que apenas incluiu estes dois factores, tendo concluído que as diferenças probabilísticas entre ambas as categorias não é muito significativa ([+humano] = pr. 0,55; [-humano] = pr. 0,45). No nosso caso, dado que este grupo de factores foi o penúltimo a ser seleccionado pela ferramenta VARBRUL em termos de significância, e também pelo facto de os nossos valores serem todos idênticos aos do MRJ, abstivemo-nos de fazer um teste semelhante ao da autora,

pois acreditamos que este levaria a conclusões idênticas às avançadas por Scherre (1988:273): a diferença entre ambas as categorias semânticas não é significativa para o desenho da configuração da CPL-var. Deste modo, uma vez mais se reforça o pressuposto de que a CPL-var do PA será essencialmente condicionada por factores estruturais.

Scherre (1988:334) observou ainda a influência na marcação PL provocada pela variável *animacidade* numa perspectiva sintagmática, constatando que, de entre os SN's com núcleo [+humano], 59% apresentam marcação total no sintagma. Por seu lado, nas cadeias sintagmáticas com núcleo [-humano], este valor decai muito ligeiramente, já que apenas 55% dos SN's apresentam marcação PL em todos os elementos flexionáveis. Contudo, como se pode verificar, uma vez mais, a diferença percentual entre ambas as categorias semânticas revela que o seu efeito não é estatisticamente relevante para a CPL-var. Face à semelhança dos resultados atomísticos obtidos entre esta variedade e o dialecto de Almojarife, somos levados a crer que o PA não apresentaria também diferenças em relação ao resultados dos falantes cariocas, caso procedêssemos a uma análise sintagmática para os factores do grupo de factores *traço semântico do SN*.

Relativamente ao pressuposto que avançámos na hipótese 11, constituída no ponto 3.6.10 para a variável independente *traço semântico do SN*, confirma-se que os nomes com traço [+humano] têm tendência a favorecerem a marcação, enquanto os substantivos com traço [-humano] e [-animado] a desfavorecem. Entre os nomes [-humanos], os que contêm traço [-animado] propiciam mais a inserção de marcas do que os que possuem traço [+animado], não por condicionalismos semânticos, mas sim por motivações de carácter estrutural. Por seu lado, o paralelismo de marcação estabelecido entre o PA, o MRJ e o PCV vem reforçar, novamente, a possibilidade de o português adquirido por falantes de um CP manifestar os mesmos padrões de variação do português adquirido quando a língua ancestral de substrato não é um CP (Figueiredo, 2008; 2009a).

Seguidamente, passaremos a analisar os resultados das variáveis independentes do tipo fonológico.

## 5.2. Variáveis independentes do tipo fonológico

### 5.2.1. Variável independente *saliência fónica* (*Processos morfofonológicos de formação de PL*)

A Tabela 12.1 evidencia que, devido a um reduzido número de ocorrências, quatro dos nossos factores tiveram que ser afastados da rodada VARBRUL efectuada para efeitos de captação dos pesos relativos: os *singulares em -l*, com dez ocorrências, sendo três delas portadoras de marca de PL (30%); os *plurais metafónicos*, com 15 ocorrências, dez das quais marcadas com PL (67%); os *singulares em -ão que pluralizam com inserção de -s e alteração silábica em -ões*, com 27 realizações, vinte e uma das quais com marca de PL (78%); e os *singulares em -ão que pluralizam com inserção de -s e alteração silábica em -ães*, que não registam qualquer ocorrência. Desde já fica inviabilizada a possibilidade de confirmarmos se os metafónicos, em virtude de pressuporem dupla distinção fónica e morfológica, estão sujeitos a um maior peso de marcação PL.

Na tabela 23.1 podemos observar os resultados da variável independente *saliência fónica*, após ter ficado reduzida a apenas quatro factores representativos da escala de diferenciação material fónica do PA:

**Tabela 23.1.** Efeito da *saliência fónica* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Plurais regulares	995/1.973	50	86	0,491
Plurais nasais regulares: -ão/-ãos; -em/-ens; -um/-uns	69/145	48	6	0,369
Singular -r	42/79	53	4	0,463
Singular -s (-z)	61/86	71	4	0,862
<b>Totais:</b>	1.166/2.283	51	100	-

O perfil de marcação PL do PA, motivado pela *saliência fónica*, revela um padrão que não parece ir totalmente ao encontro dos achados de Guy (1981a:181-199), isto é, dos falantes cariocas semi-escolarizados. De facto, os itens apontados pelo autor como inibidores da marcação mostram-se também desfavorecedores da mesma em Almojarife (regulares = pr. 0,491; nasais regulares = pr. 0,369; elementos em *-l* = 30% de marcação



positiva). Contudo, o mesmo não sucede com os casos em *-r*, que favorecem a pluralização nos dados do Rio de Janeiro, mas exibem tendência para inibirem ligeiramente a inserção de marcas no PA (pr. 0,463). Assim, nos nossos dados, a tendência para o favorecimento da marcação, bastante acentuado, diga-se, é um exclusivo dos itens que finalizam em *-s* no singular (pr. 0,862), que apresentam apenas um grau ligeiro de distinção fónica singular/plural. Sobre a motivação que poderá subjazer à tendência para a inserção de marcas nestes elementos nos pronunciaremos em etapa posterior do presente trabalho. Na Tabela 12.1 é possível ainda ver que os plurais metafónicos se apresentam como elementos que poderão possibilitar a inserção de marcas, mas moderadamente (67% de realizações marcadas positivamente), e não de forma acentuada, como referiu Guy (1981a:187).

Face aos nossos resultados, parece ser possível avançar já com duas constatações: (i) no PA, a inibição evidenciada pelos itens em *-r* impossibilita a redução da variável independente *saliência fónica* a uma oposição binária nos moldes de Guy (1981a:166), isto é, consoantes favorecedoras da marcação, por um lado, e vogais inibidoras da pluralização, por outro lado; (ii) não se comprova conclusivamente, tal como avançou Guy (1981a:189), que a *saliência fónica* exerça efeito na marcação, uma vez que o maior grau de concordância não parece incidir sobre os itens cuja *saliência fónica* é mais perceptível na oposição singular/plural. Ainda sobre os itens que finalizam em *-r*, iremos adiantar mais considerando quando testarmos cruzamentos entre as variáveis independentes *saliência fónica* e *tonicidade* (pontos 5.2.2.2 a 5.2.2.7 do nosso trabalho).

Mas, antes de procedermos a comparações e conclusões, atentemos aos valores que nos permitem estabelecer contrastes acerca do modo como a variável independente *saliência fónica* intervém na marcação PL do SN dos diferentes dialectos que temos vindo a analisar. A influência da *saliência fónica* na marcação PL dos elementos do SN do MRJ, segundo Scherre (1988:78), e conforme o demonstram os resultados da tabela 23.2, exhibe o seguinte padrão: mais marcas nos plurais metafónicos e menos marcas nos plurais regulares. Quantos aos restantes factores, apresentam-se colocados nas categorias intermédias. Apesar de não possuímos o peso relativo dos plurais metafónicos, é possível constatar que o seu percentual de marcação (67%) é inferior quer ao dos itens terminados em *-ão* e que pluralizam com inserção de *-s* e alteração silábica (78%)

(Tabela 12.1) quer ao dos elementos em *-s*, cujo PL se faz através da inserção de *-es* ou *-e* (71%) e se apresentam na Tabela 23.1 como aqueles que mais beneficiam a inserção da pluralização (pr. 0,862). Assim sendo, o padrão de marcação motivado pela *saliência fónica* apresentado por Scherre (1988:79) evidencia algumas diferenças relativamente ao padrão do PA, uma vez que, neste, não parecem ser os plurais metafónicos aqueles que mais propiciam a inserção de marcas de PL, como também não são os itens regulares que apresentam o maior peso na inibição, cabendo este, como se viu, aos elementos nasais que pluralizam regularmente (pr. 0,369). Contudo, a autora amalgamou os itens que pluralizam de modo regular conjuntamente com os elementos nasais que também fazem a pluralização de forma regular (Scherre, 1988:75), pelo que acreditamos que, no MRJ, o peso inibidor dos itens nasais tenha contribuído para reforçar, ainda mais, a tendência para a inibição demonstrada pelo factor plurais regulares. Assim, e mesmo que se exceptue o caso dos plurais regulares em *-ão*, bem como as ocorrências em *-l*, cuja tendência para o não-favorecimento da marcação se ficará a dever à incerteza na forma de os pluralizar,<sup>182</sup> constatamos que o Princípio da Saliência Fónica (Naro & Lemle, 1976:240-241; Lemle & Naro, 1977:47) encontra eco débil no PA, com as formas com maior distinção fónica singular/plural a não apresentarem a maior probabilidade de marcação PL, não se confirmando, na totalidade, o observado por Scherre (1988:79) para o MRJ: “*os falantes adultos, com referência a Processos, evidenciam mais marcas de concordância nos itens lexicais que apresentam mais diferença material fónica entre as suas respectivas formas singulares e plurais*”.

Scherre (1988:77) analisou ainda o efeito do número de sílabas na marcação PL, tendo concluído que o mesmo não exerce influência sobre a concordância PL entre os elementos do SN do MRJ, visto os monossílabos, os dissílabos e os itens de mais de duas sílabas apresentarem, todos, valores idênticos e que se situam no patamar da neutralidade da inserção da pluralização. A autora adianta duas explicações para justificar o fenómeno: “1) 90% dos casos de monossílabos ocorrem antepostos ao núcleo, sendo 81% artigos e 09% possessivos, os quais são marcados por esta razão e não por serem monossílabos; 2) os 10% restantes são todos tónicos, e são marcados por causa da Tonicidade e não, também, pelo número de sílabas” (Scherre, 1988:82). Tendo

como bastante consistentes as justificações da linguista, absteremo-nos de analisar este tipo de efeito no presente trabalho.

**Tabela 23.2.** *Saliência fónica*: contribuição dos factores individuais para a marcação plural do elemento analisado: 7 variedades de português.

Factores	Saliência fónica							
	PA		MRJ		NURC		HEL-Ba	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%
Plurais regulares	0,491	50			0,50	81	0,47	49
Nasais regulares	0,369	48	0,24	76	0,21	49		
Plurais metafónicos	(10/15)	67	0,86	93	0,84	88		
Singular –r	0,463	53	0,48	88	0,58	90	0,66	32
Singular –s (-z)	0,862	71	0,38	83	0,39	75		
Singular –l	(3/10)	30	0,56	86	0,63	88		
Singular –ão/–ões	(21/27)	78	0,42	86	0,54	87	0,52	7

Factores	PT		PMp		PCV	
	Pr.	%	Pr.	%	Pr.	%
Plurais regulares			-	-	-	-
Nasais regulares	0,52	51	-	-	-	-
Plurais metafónicos	0,59	59	-	-	-	-
Singular –r	0,46	45	0,33	77	0,36	72
Singular –s (-z)	0,35	25	0,22	64	0,08	24
Singular –l			-	-	-	-
Singular –ão/–ões	0,88	82	-	-	-	-

Contudo, estes aspectos levaram Scherre (1988:82) a excluir das suas rodadas as categorias gramaticais cuja diferenciação material fónica não é considerada relevante para a inserção de marcas PL, visto serem itens que pluralizam de modo regular. Como tal, a linguista passou a incluir nas suas análises apenas os substantivos, adjectivos e categorias substantivadas, elementos nos quais ocorre a maior diferenciação material fónica entre singular e PL. Comparando os resultados de todos os dados com os que incluem apenas estas três categorias gramaticais, (Scherre, 1988:85) constatou que, a segunda análise condizia melhor com “a realidade dos factos, porque a variável

*Processo é a que se aplica especificamente ao fenómeno da concordância*”. Esta conclusão ficou a dever-se ao facto de ocorrer uma inversão na selecção das variáveis independentes, tendo o grupo de factores *saliência fónica* passado a ser seleccionado em segundo lugar entre todas as variáveis, ficando a *tonicidade* relegada para a sexta posição. Deste modo, a autora passou a estudar os efeitos da *saliência fónica* sobre a marcação, considerando apenas as três categorias gramaticais em questão.

A conclusão de que a variável independente *saliência fónica* é aquela que se aplica especificamente ao fenómeno da concordância encontra paralelo nos nossos dados. Não obstante, os nossos resultados revelaram este aspecto desde o início, visto que este grupo de factores foi seleccionado imediatamente como mais importante do que a variável independente *tonicidade*, mesmo com todos os itens gramaticais incluídos. Nesta conformidade, não vemos necessidade de se excluirmos das nossas rodadas VARBRUL os elementos gramaticais que apresentam menos diferenciação material fónica entre singular e PL.

Comparativamente aos resultados de Lopes (2001:137-138), verifica-se a ocorrência de um paralelo de marcação entre a NURC e o PA no que concerne aos itens que pluralizam regularmente, já que ambos se encontram no patamar da neutralização da marcação PL (PA = pr. 0,491; NURC = pr. 0,50). Este paralelo de marcação dos itens regulares estende-se ainda aos que finalizam em *-ão* no singular, já que revelam uma tendência inibidora da marcação em ambos os dialectos, embora esta seja mais acentuada no português de Salvador (PA = pr. 0,369; NURC = pr. 0,21).

Relativamente a outros itens com *saliência fónica* na distinção singular/plural, observa-se que os elementos finalizados em *-r* no singular apresentam tendência para propiciar ligeiramente a inserção de marcas na NURC (pr. 0,58), conquanto no PA a tendência seja para a inibição suave (pr. 0,463). A diferença de comportamento que motiva a variação da marcação PL vai acentuar-se ainda mais com os casos terminados em *-s* no singular, que exibem comportamentos diametralmente opostos nos dois dialectos, ao apresentarem-se quer como os mais inibidores da pluralização na NURC (pr. 0,39) quer como os mais favorecedores da inserção de marcas no PA (pr. 0,862). Lopes (2001:138), a exemplo de Braga & Scherre (1976:474), justifica a inibição com base na interpretação analógica de que o /s/ final pode ser entendido em Salvador como

marca de PL, mas, ao que tudo indica, este é um fenómeno que não ocorre no PA. Contudo, convém aqui referir que a discrepância de marcação entre a NURC e o PA poderá muito bem assentar no facto de Lopes (2001:137), contrariando o sugerido pela metodologia de análise em Sociolinguística Quantitativa, ter codificado e incluído o item “vez/vezes” no factor dos elementos terminados em *-s*, enquanto nós o excluimos das nossas análises. Por fim, resta referir que, mais uma vez, os dados do PA não permitem corroborar, pelo menos nesta fase das nossas análises, o princípio de que os itens mais salientes são os que demonstram o maior grau de concordância, enquanto os itens menos salientes são afectados pela variação.

Nos dados de P. Andrade (2003:103), a categoria de factores que agrupa os itens terminados em consoante (*-l*, *-r* e *-s*), os plurais metafónicos e os elementos terminados em ditongos nasais sem alteração fónica na flexão PL, é aquela que apresenta maior índice de marcação (pr. 0,66), apesar de ocorrer em menor frequência do que a categoria dos plurais regulares, os mais inibidores da inserção PL (0,47). Em plano intermédio encontram-se os elementos constituídos por ditongo nasal com alteração fónica acentuada (*-ão/-ões* e *-ão/-ães*), a exibirem ligeira tendência para a marcação (0,52). Relativamente ao primeiro factor amalgamado, isto é, o que mais beneficia a inserção de marcas, verificamos que nele estão também incluídos os itens que mais marcam o PL no PA (*-s* = pr. 0,862). No que diz respeito ao factor intermédio de HEL-Ba (ditongos nasais com alteração fónica), apenas possuímos percentuais de marcação (78%), pelo que não é possível estabelecer comparações entre os dois dialectos. No entanto, atente-se ao paralelo de inibição determinado pelos itens regulares, bastante próximo nos dois dialectos (PA = pr. 0,491; HEL-Ba = pr. 0,47). Estamos convictos que se efectuássemos uma amalgamação dos nossos factores conforme metodologia adoptada para HEL-Ba, baixáramos não só o peso favorecedor dos itens em *-s* como elevaríamos também o peso do favorecimento dos itens em *-r*, colocando-os a um nível aproximado do valor apresentado por P. Andrade (2003:103). Como tal, entendemos que os padrões de marcação PL do PA e HEL-Ba se assemelham bastante, o que nos permite concordar com a autora quando refere que a *saliência fónica* tem influência na aplicação da regra da marcação PL (P. Andrade: 2003:120).

Os achados de Baxter (2004:109) apontam para o efeito da saliência na marcação PL ser fortemente favorecida pelas distinções morfofonológicas das oposições mais marcadas, pelas combinações de sílabas plurais tónicas e pelas mudanças que produzem alteração no radical da palavra (pr. 0,88). Com afectação mais neutra na marcação surgem os monossílabos (pr. 0,61), as formas metafónicas (pr. 0,59) e as realizações que envolvem uma pluralização tónica, mas sem alteração no radical da palavra (pr. 0,52). Desfavorecendo a marcação surgem quer as palavras sem mudança no radical e com sílaba PL átona (pr. 0,46) quer as palavras com singular em *-s* (pr. 0,35). Contudo, as comparações com o PA teriam que ser observadas aqui sob alguma reserva, visto que os resultados do PT reflectem valores do cruzamento entre as variáveis independentes *saliência fónica* e *tonicidade*, enquanto os dados do PA se referem apenas ao grupo de factores *saliência fónica*. Assim sendo, remetemos as comparações entre Almojarife e Monte Café para o momento em que levarmos a cabo idêntico tratamento de dados (pontos 5.2.2.2 a 5.2.2.7 do presente trabalho).

Igual procedimento será adoptado em relação aos resultados apresentados quer para o PMp (Jon-And, 2008:4) quer para o PCV (Jon-And, 2009:5), já que os mesmos dizem também respeito ao cruzamento entre as mesmas variáveis independentes. No entanto, chamamos aqui a atenção para o facto de os itens que pluralizam com inserção de *-es* final se revelarem altamente inibidores da marcação tanto no PMp (*-s* = pr. 0,22; *-r* = pr. 0,33) como no PCV (*-s* = pr. 0,08; *-r* = pr. 0,36), enquanto no PA apresentam comportamentos distintos, um favorecendo a pluralização e o outro inibindo-a ligeiramente (*-s* = pr. 0,862; *-r* = pr. 0,468). Contudo, é de ter em conta que a autora não optou por uma codificação do nosso tipo, já que considerou como pluralizados os itens com marcação fonológica parcial (mês/mese),<sup>183</sup> demarcando-se também da metodologia de Braga (1977:83, nota 20), que os codificou como elementos sem marcação de número.

Das comparações entre o PA e as diferentes variedades de português aqui abordadas, salta à evidência um padrão absolutamente uniforme dos itens regulares em todos os dialectos, com valores sempre situados no patamar da neutralização da marcação. Por outro lado, é também claro que ocorre uma discrepância quanto à forma de os itens em *-s* marcarem o PL, com os valores do PA favorecedores da marcação encontrando

respaldo nos resultados de Guy (1981a:187) e também de P. Andrade (Andrade, 2003:103), mas demarcando-se completamente dos pesos inibidores apresentados por Scherre (1988:78), Baxter (2004:109) e Jon-And (2008:4, 2009:5). Quanto aos dois outros itens para os quais possuímos pesos relativos, registam-se também algumas flutuações, mas em menor grau do que a apresentada pelos elementos em *-s*. Assim, os itens nasais regulares inibem a marcação no PA, no MRJ e na NURC, mas favorecem-na no PT. Quanto ao itens em *-r*, apresentam-se todos próximos do ponto da neutralidade, umas vezes desfavorecendo ligeiramente a inserção de marcas (PA, MRJ, PT e PMp), outra vezes favorecendo-a suavemente (NURC e HEL-Ba). Contudo, estas ligeiras flutuações poderão estar dependentes, quanto a nós, das diferentes formas de constituir os factores amalgamados nos quais estes elementos foram inseridos.

#### **5.2.1.1. Relação entre *saliência fónica* e *escolaridade* na marcação PL do PA**

Scherre (1988:85) abordou a correlação entre a *saliência fónica* e as características sociais dos falantes do MRJ, referindo que os anos de escolarização actuam a nível do funcionamento sincrónico da gramática, o que faz com que a influência da *saliência fónica* sobre a marcação PL dependa do grau de ensino dos falantes. A motivação que levou a autora a optar por este tipo de análise assentou, essencialmente, no facto de os itens em *-ão*, ao inibirem a pluralização, não obedecerem à regra que determina a hierarquia de marcação dos itens com maior distinção fónica singular/plural. Em termos gerais, as análises de Scherre (1988:86) apresentaram algum distanciamento na forma de marcar o PL em apenas alguns itens, especialmente os elementos finalizados em *-ão* e *-l* produzidos pelos níveis Primário e Ginásial,<sup>184</sup> aspecto que levou a autora a observar também a correlação entre a *tonicidade* e o *grau de escolarização*. Nesta, o distanciamento é menos acentuado, tendo-se notado uma uniformidade na forma de marcar os oxítonos (os que mais beneficiam a marcação) e paroxítonos (os que mais inibem a inserção de marcas), mas alguma irregularidade na forma de pluralizar os proparoxítonos (o nível Primário inibe a marcação, enquanto o Colegial a favorece ligeiramente). A conclusão final de Scherre (1988:90-91) é que “*Processos e Tonicidade influenciam a presença/ausência da concordância nominal embora apresentem algumas diferenças na forma de atuar em função do grau de escolarização dos falantes*”.

Lopes (2001:146) analisou também o mesmo tipo de correlação, mas fê-lo cruzando os grupos de factores *saliência fónica e tonicidade*. A conclusão da autora é que, apesar de as probabilidades de marcação não serem muito semelhantes nos três grupos analisados, verifica-se em todos eles a oposição entre mais salientes e mais concordância, *versus* menos salientes e menos concordância. Outro aspecto a ter em conta nos resultados de Lopes (2001:154), é o facto de o aumento da escolarização conter, e por vezes reduzir, a influência da *saliência fónica* na marcação PL.

Baxter (2004) e Jon-And (2008, 2009) não observaram a interacção entre as variáveis independentes *saliência fónica e escolaridade* na marcação PL, mas o primeiro analisou o efeito do cruzamento entre *saliência fónica e tonicidade* no desenvolvimento geracional do PL no SN dos tongas. Como tal, remetemos as comparações com os números de Baxter (2004:109) para o momento em que procederemos também a idêntico tratamento de dados (ponto 5.2.2.6 do presente trabalho).

Em termos de anos de escolarização, a estratificação dos falantes de Almojarife difere dos grupos de informantes de Scherre (1988) e Lopes (2001), pelo que é difícil estabelecer comparações precisas a nível dos factores que compõem a variável social independente *escolaridade* dos dialectos em questão. Contudo, é possível observar o padrão geral de variação da pluralização, bem como detectar os contrastes de marcação entre falantes com maior ou menor grau de escolarização, seja na perspectiva da intraescolaridade no PA seja na óptica da interescolaridade no PA, MRJ e NURC. Os resultados da correlação entre a *saliência fónica* e o *grau de escolarização* dos falantes do PA, na inserção de marcas de PL, estão patentes na Tabela 23.3:

A exiguidade dos nossos dados não permite obter resultados para a grande maioria dos factores, quando se pretende observar a sua distribuição pelos diferentes níveis de escolarização. Note-se que, para conseguirmos resultados mínimos que permitissem uma observação da relação entre *saliência fónica e escolaridade*, tivemos mesmo que lançar mão de dados que registam menos de trinta ocorrências, como aconteceu com o factor *singular –s* da faixa com escolarização média, com apenas 26 ocorrências, e o factor *nasais regulares*, dos falantes com escolarização alta, que regista somente 28 realizações. Só assim conseguimos evitar que estes dois grupos escolares fossem rejeitados pela ferramenta VARBRUL por configurarem *singletons*, isto é, por



apresentarem valores apenas para um factor e ficar inviabilizada a hipótese de uma observação binária dos grupos de factores. Ainda assim, apenas conseguimos escassos valores, e somente para dois grupos (analfabetos e escolarização média), já que os outros dois (escolarização baixa e escolarização alta) foram considerados não significativos para a marcação PL no SN do PA.

**Tabela 23.3.** Relação entre *saliência fónica* e *escolaridade* na marcação PL dos itens do SN do PA.

		<i>Input desta rodada: 0,153</i> <i>Log-likelihood: -169,044</i> <i>Significância: 0,045</i>			<i>Input desta rodada: 0,616</i> <i>Log-likelihood: -296,569</i> <i>Significância: 0,009</i>		
<i>Saliência fónica</i> x <i>Escolaridade</i>	<b>Analfabetos</b>			<b>Escolarização baixa</b>			
	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	
Plurais regulares	152/486	31	0,484	353/589	60	-	
Nasais regulares	3/33	9	0,122	37/60	62	-	
Plurais metafónicos	(4/5)	80	-	(2/3)	67	-	
Singular –r	(3/19)	16	-	(15/26)	58	-	
Singular –s (-z)	20/41	49	0,912	(7/7)	100	-	
Singular –l	-	-	-	(0/2)	0	-	
Singular –ão/–ões	(5/9)	56	-	(5/7)	71	-	
Total	187/593	32	-	419/694	60	-	
		<i>Input desta rodada: 0,464</i> <i>Log-likelihood: -337,006</i> <i>Significância: 0,007</i>			<i>Input desta rodada: 0,617</i> <i>Log-likelihood: -168,068</i> <i>Significância: 0,023</i>		
<i>Saliência fónica</i> x <i>Escolaridade</i>	<b>Escolarização média</b>			<b>Escolarização alta</b>			
	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	
Plurais regulares	298/616	48	0,468	92/282	68	-	
Nasais regulares	(8/24)	33	-	(21/28)	75	-	
Plurais metafónicos	(0/1)	0	-	(4/6)	67	-	
Singular –r	(12/19)	63	-	(12/15)	80	-	
Singular –s (-z)	(22/26)	85	0,954	(12/12)	100	-	
Singular –l	(2/5)	40	-	(1/3)	33	-	
Singular –ão/–ões	(5/5)	100	-	(6/6)	100	-	
Total	347/696	50	-	248/352	71	-	

**Obs.:** Escolarização baixa e escolarização alta rejeitadas pela ferramenta VARBRUL.

Os informantes analfabetos apresentam uma distinção acentuada na forma de marcar o PL entre os itens que finalizam em *-s* no singular, categoricamente favorecedores da marcação (pr. 0,912), e os itens nasais que pluralizam regularmente, os quais inibem acentuadamente a pluralização (pr. 0,122). Em plano intermédio encontram-se os elementos que fazem o PL de modo regular, com uma tendência suavemente inibidora da pluralização (pr. 0,484).

Note-se que o padrão do elemento mais favorecedor da marcação, isto é, dos itens que terminam em *-s* no singular, se mantém constante ao longo de todas as gerações, apresentando uma taxa de marcação plena (100%) nas escolaridades baixa e alta (apesar não atingir trinta ocorrências em qualquer delas) e um peso na escolaridade média (pr. 0,954) que se situa no mesmo patamar do favorecimento categórico exibido pela classe analfabeta. O mesmo factor apresenta-se como inibidor da marcação em todos os níveis de escolaridade do MRJ, embora o grau de inibição varie entre os diferentes estratos.<sup>185</sup> Relativamente aos falantes da NURC, Lopes (2001:153) apresenta valores favorecedores da marcação PL, mas que vão decrescendo do Fundamental até ao Universitário (Fundamental = pr. 0,72; Colegial = pr. 0,63; Universitário = 0,58). Como tal, é possível constatar que, no que respeita à pluralização dos itens em *-s*, o MRJ é o único dialecto a apresentar um padrão inibidor da marcação, sendo o do PA mais categórico ao favorecimento do que o da NURC. Gostaríamos ainda de abrir aqui um parêntesis para chamar a atenção para as análises e observações que efectuaremos a propósito deste fenómeno durante o estudo da variável independente *contexto fonológico posterior* (ponto 5.2.3 do presente trabalho), nomeadamente quando nos debruçarmos sobre questões fonológicas relacionadas com a assimilação regressiva e a haplologia sintáctica, as quais, segundo Braga & Scherre (1976), Braga (1977) e Scherre (1988) estarão na origem da inibição da marcação PL evidenciada pelos itens em *-s*.

O padrão de regularidade na marcação PL pode ser também observado para o factor intermédio nos falantes analfabetos e de escolarização média do PA, ou seja, os itens que pluralizam de forma regular. De facto, o peso ligeiramente inibidor da inserção de marcas evidenciado pelos analfabetos (pr. 0,484) é bastante semelhante ao do evidenciado pelos informantes com escolarização média (pr. 0,468), não chegando a ultrapassar o patamar da neutralização. O mesmo factor exhibe também um padrão

bastante regular no MRJ (pr. 0,24 para todos os níveis de escolaridade), sendo o mais inibidor da marcação neste, e em valores mais acentuados do que os do PA. Contudo, não podemos esquecer que Scherre (1988:75) juntou neste factor os elementos regulares e os itens que finalizam em *-ão* e que pluralizam também de forma regular, o que terá inflacionado o peso de inibição da marcação. Quanto aos valores dos itens regulares oxítonos, paroxítonos e proparoxítonos da NURC, comentá-los-emos quando os compararmos com os números exibidos pela relação entre as variáveis independentes *tonicidade* e *escolaridade* das nossas ocorrências (ponto 5.2.2.1 do presente trabalho).

Relativamente aos itens nasais regulares, apenas obtivemos valores para o grupo analfabeto, no qual o mesmo se revela como altamente inibidor da marcação (pr. 0,122). Scherre (1988) e Lopes (2001) não apresentam valores para este tipo de factor, já que a primeira os amalgamou com os elementos que pluralizam de modo regular, enquanto a segunda nos traz apenas valores para os itens regulares oxítonos, paroxítonos e proparoxítonos.

Dada a escassez dos nossos resultados, iremos proceder a mais considerandos quando cruzarmos a variável independente *tonicidade* com o grupo de factores *escolaridade* (ponto 5.2.2.1 do presente estudo), pelo que remetemos para esta fase do trabalho as conclusões acerca da possível influência na presença/ausência da concordância nominal motivada pelo eixo *saliência fónica/tonicidade*.

Ainda assim, e a propósito da hipótese 1, que levantámos para a variável independente *saliência fónica* (ponto 3.6.1.1), é possível adiantar que, embora este grupo de factores revele influência na aplicação da regra de concordância do PA, esta não acompanha totalmente a predição avançada de que os itens que apresentam maior distinção fónica tendem a favorecer a inserção da pluralização. De facto, além de se constatar que os elementos terminados em *-s* no singular favorecem categoricamente a marcação no PA, verificamos também que, entre os itens que pluralizam regularmente, a inibição é mais acentuada nos que apresentam maior distinção fónica, isto é, nas palavras que terminam em *-ão* no singular. O maior propiciamento destes à não-marcação ficará a dever-se às diferentes formas de inserir as marcas de número nos itens em *-ão* (Scherre, 1988:124), que levará os falantes a hesitarem na forma de os pluralizar. As baixas realizações registadas para os elementos metafónicos impedem que

confirmemos se são estes os que favorecem mais a inserção da regra, mas o seu percentual de marcação, situado no patamar neutro, indicia que, no PA, o seu propiciamento à pluralização não alcance a *performance* apresentada pelos itens terminados em *-s* no singular.

### 5.2.2. Variável independente *tonicidade* (*Tonicidade dos itens lexicais singulares*)

Como se constata na Tabela 12.2, o nosso *corpus* possui apenas doze itens proparoxítonos. Pode também verificar-se na Tabela 24.4, na qual se apresentam resultados para o cruzamento entre as variáveis independentes *saliência fónica* e *tonicidade*, que todos estes elementos pluralizam de forma regular. Assim sendo, decidimos manter esta distinção e não amalgamá-los com os itens paroxítonos, que podem pluralizar de modo irregular (p.e. *difícil/difíceis*). Com esta metodologia, anulou-se a possibilidade de se constituir um factor único para ambos os tipos de itens, tendo a insuficiência de realizações proparoxítonas determinado que as mesmas fossem retiradas da rodada VARBRUL para apuramento de pesos relativos dos factores que constituem o grupo de factores *tonicidade*. Os resultados desta rodada podem ser observados na Tabela 24.1:

**Tabela 24.1.** Efeito da *tonicidade* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Oxítonos e monossílabos tónicos	179/282	64	12	0,730
Monossílabos átonos	283/320	88	14	0,690
Paroxítonos	735/1.726	43	74	0,423
<b>Totais:</b>	<b>1.197/2.328</b>	<b>51</b>	<b>100</b>	<b>-</b>

O padrão de marcação do PA é em tudo semelhante aos do MRJ (Scherre, 1988:80) e da NURC (Lopes, 2001:140), mostrando que, em termos de comunidade, os oxítonos e os monossílabos tónicos detêm a primazia na marcação PL em todos estes dialectos, enquanto os paroxítonos e os proparoxítonos se apresentam como os elementos que mais inibem a pluralização (Tabela 24.2). Assim sendo, ganha forte sustentação o pressuposto de que as formas mais salientes favorecem mais a concordância PL do que os itens em que essa percepção não é tão evidente.

**Tabela 24.2.** *Tonicidade*: contribuição dos factores individuais para a marcação PL do elemento analisado: 3 variedades de português.

Factores	Tonicidade					
	PA		MRJ		NURC	
	Pr.	%	Pr.	%	Pr.	%
Oxítonos/Monossílabos tónicos	0,730	64	0,66	86	0,72	85
Monossílabos átonos	0,690	88			0,65	99
Paroxítonos	0,423	43	0,39	70	0,37	71
Proparoxítonos	(5/12)	42	0,44	52	0,41	75

Guy (1981a:166) sugeriu o uso de uma escala binária para estudo da influência da *saliência fónica* na marcação, ou seja, palavras terminadas em vogal *versus* palavras terminadas em consoante. Com esta oposição binária, torna-se lógico que a substituição de uma forma por outra se dê a partir daquela que é mais saliente, mas, ao contrário do que o autor defende (Guy, 1981a:298), a redução que configura processos de descrioulização não é motivada unicamente por factores internos à língua, isto é, não se inicia somente pela morfologia, descartando, conseqüentemente, as motivações fonológicas.

P. Andrade (2003:114), no seu estudo não-atomístico do dialecto de HEL-Ba, abordou a concordância PL em função da *saliência fónica* na oposição singular/plural, tendo constituído, para o efeito, um grupo de factores binário com SN's que contêm apenas itens regulares, por um lado, e SN's contendo itens que apresentam diferenciação fónica quando pluralizados, por outro lado. Os resultados mostram que o segundo factor favorece moderadamente a marcação PL (pr. 0,66), enquanto os SN's que apenas possuem itens regulares a inibem muito ligeiramente (pr. 0,46). Deste modo, a autora conclui que os seus informantes adquirem formas da língua a partir de contextos linguísticos mais salientes, mas que há toda uma rede de elementos internos e externos envolvidos nos processos de mudança que determinam não só a aquisição de marcas mas também a perda de antigas características provenientes de processos de transmissão linguística irregular (P. Andrade, 2003:115). Tendo em conta estes aspectos, decidimos observar o modo como as variáveis sociais *escolaridade* e *idade* intervêm, em conjunto com a *tonicidade*, na concordância PL do PA. Para efeitos de comparações com os diferentes dialectos em observação, foi necessário recorrer também a cruzamentos entre

as variáveis independentes *saliência fônica* e *tonicidade*, segundo diferentes metodologias adotadas pelos pesquisadores que estudaram tais variedades de português.

#### **5.2.2.1. Relação entre *tonicidade* e *escolaridade* na marcação PL do PA e do MRJ**

Scherre (1988:89) foi a única a apresentar valores para a tonicidade da sílaba do item lexical singular por escalões escolarizados, uma vez que os resultados de Lopes (2001), Baxter (2004) e Jon-And (2008, 2009) levaram em conta apenas o grupo de factores *saliência*, isto é, o cruzamento entre as variáveis independentes *saliência fônica* e *tonicidade*. Antes de compararmos os nossos números com os do MRJ, observemos, com recurso à Tabela 24.3, os valores relativos à relação entre *tonicidade* e *escolaridade* no PA.

Entre os indivíduos analfabetos, os oxítonos e monossílabos tónicos são aqueles que são primeiramente apreendidos e mostram maior índice de concordância em número (pr. 0,654). Por seu lado, os elementos com menor *tonicidade* surgem menos marcados, com os monossílabos átonos a revelarem-se os mais inibidores da concordância (pr. 0,439), enquanto os paroxítonos, que possuem uma sílaba tónica, ou seja, com *saliência tónica* superior à dos anteriores, inibem um pouco menos a marcação PL (pr. 0,479). Como tal, e ao que tudo indica, as informações que pressupõem maior tonicidade são as que são percebidas primeiramente, reflectindo o PA mais marcas de aprendizagem nos elementos mais salientes e ausência destas nos itens menos salientes.

Não obstante, os informantes com baixa escolarização evidenciam um padrão totalmente distinto, já que, nestes, os monossílabos átonos ganham protagonismo no favorecimento à concordância (pr. 0,841), em valores que ultrapassam mesmo os dos oxítonos e monossílabos tónicos (pr. 0,726). Em sequência de tal, os paroxítonos acabam por cair para a cauda da marcação, mas mantendo a sua característica de elementos inibidores da concordância (pr. 0,361). Neste estrato de ensino, salta à evidência que actua uma gramática distinta, com o padrão da pluralização perdendo uniformidade, já que são os elementos de menor *tonicidade* que passam a deter o maior índice de marcação PL.

**Tabela 24.3.** Relação entre *tonicidade* e *escolaridade* na marcação PL dos itens do SN do PA.

	<i>Input desta rodada:</i> 0,153 <i>Log-likelihood:</i> -169,044 <i>Significância:</i> 0,045			<i>Input desta rodada:</i> 0,616 <i>Log-likelihood:</i> -294,148 <i>Significância:</i> 0,005		
	<b>Analfabeto</b>			<b>Escolaridade baixa</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Oxítonos/Monossílabos tónicos	37/81	46	0,654	49/69	71	0,726
Monossílabos átonos	38/49	78	0,439	126/128	98	0,841
Paroxítonos	112/465	24	0,479	244/493	50	0,361
Proparoxítonos	(0/1)	0	-	(1/5)	20	-
<b>Totais:</b>	<b>187/596</b>	<b>31</b>	<b>-</b>	<b>420/695</b>	<b>60</b>	<b>-</b>
	<i>Input desta rodada:</i> 0,464 <i>Log-likelihood:</i> -337,006 <i>Significância:</i> 0,007			<i>Input desta rodada:</i> 0,587 <i>Log-likelihood:</i> -164,734 <i>Significância:</i> 0,046		
	<b>Escolaridade média</b>			<b>Escolaridade alta</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Oxítonos/Monossílabos tónicos	52/86	61	-	(41/46)	89	0,858
Monossílabos átonos	78/98	80	-	41/45	91	0,201
Paroxítonos	217/511	43	-	162/257	63	0,480
Proparoxítonos	(0/2)	0	-	(4/4)	100	-
<b>Totais:</b>	<b>347/697</b>	<b>50</b>	<b>-</b>	<b>248/352</b>	<b>71</b>	<b>-</b>

Obs.: Escolaridade média rejeitada pela ferramenta VARBRUL.

Quanto ao grupo com escolarização média, as suas realizações não foram consideradas relevantes para a marcação da concordância. Por seu lado, a faixa com escolarização alta revela um padrão de concordância de acordo com o da comunidade, isto é, itens com maior *tonicidade* apresentando também maior índice de concordância PL (oxítonos e monossílabos tónicos = pr. 0,858) e elementos com menos *tonicidade* revelando inibição da concordância, mais os elementos plenamente átonos (monossílabos átonos = pr. 0,201) do que os casos em que ocorre uma sílaba tónica (paroxítonos = pr. 0,480).

Em termos de relação entre *tonicidade* e *escolaridade*, o padrão de marcação do PA diverge bastante do padrão do MRJ (Scherre, 1988:89), no qual o factor dos oxítonos e monossílabos tónicos é sempre os mais marcado e o dos paroxítonos o menos marcado, numa oposição uniforme que se estende ao longo dos diferentes estratos de escolarização. No entanto, Scherre (1988:90) constatou que os proparoxítonos

apresentam alguma diferença de comportamento na marcação, que a autora entende advir do facto de estes elementos serem instáveis na língua portuguesa, havendo mesmo casos de passagem de proparoxítonos a oxítonos, com maior frequência nas classes não escolarizadas. No PA, o comportamento deste tipo de itens não pode ser aquilatado, devido ao seu reduzido número de realizações. No entanto, face ao padrão registado pelos analfabetos e indivíduos de escolarização alta, é possível observar que *saliência fónica* e *tonicidade* parecem ter influência na presença/ausência da marcação PL, apesar de se registarem flutuações acentuadas de alguns itens em função do grau de escolarização dos falantes que os realizam.

Como os níveis de escolarização surgem associados aos diferentes escalões etários do PA (menos escolarizados sinónimo de mais idosos e mais escolarizados igual a menos idosos), decidimos observar também o comportamento das variáveis independentes *saliência fónica* e *tonicidade* em função da idade dos informantes, a fim de tirarmos conclusões mais sólidas sobre a influência destas variáveis na pluralização dos itens do SN da fala de Almocharife. Por outro lado, este tratamento de dados permitirá que se estabeleçam também comparações com o PT, cuja influência da *saliência* foi observada por Baxter (2004:109) numa óptica geracional. Para tanto, teremos que ajustar os nossos factores aos deste autor, procedendo a cruzamentos entre factores dos grupos de factores *saliência fónica* e *tonicidade*. Este tipo de metodologia vai ainda permitir que estabeleçamos comparações não só com os níveis de escolarização do MRJ e da NURC mas também com os resultados do PMp.

#### **5.2.2.2. Efeito da *saliência* (*saliência fónica* x *tonicidade*) na marcação PL do dos itens do SN do PA**

A sobreposição entre as variáveis independentes *saliência fónica* e *tonicidade* é evidente, já que a maioria dos itens regulares e metafónicos são proparoxítonos, enquanto os elementos que terminam em *-l*, *-r*, *-s* e *-ão* são quase todos oxítonos. Deste modo, e embora Guy (1981a:163) considere que o efeito da *tonicidade* seja aparente, especialmente nos itens não regulares, Scherre (1988:91) e Lopes (2001:143) procederam a cruzamentos entre estes dois grupos de factores, por considerarem que esta variável é “*um traço integrante do eixo da saliência fónica, com repercussões importantes sobre o índice de presença de marca formal de plural nos itens analisados*”



(Scherre, 1988:101). Lopes (2001), ao dar seguimento às sugestões de Scherre (1988) no sentido de considerar só uma variável, advinda da junção de *processos* e *tonicidade*, constatou também que apenas há possibilidade de se recorrer à segunda variável como elemento de comparação entre os itens regulares, uma vez que só estes têm palavras com os três tipos de tonicidade. Desta forma, a autora seguiu a metodologia de Scherre (1988), transformando as referidas variáveis numa só, apelidada de *saliência*, e que cruza a *tonicidade* unicamente com os itens que fazem o PL de modo regular. Dado ainda que os outros factores do grupo de factores *saliência fónica* ou não apresentam mais do que uma forma quanto à *tonicidade* ou, se a apresentam, a distinção não é tão forte quanto a dos itens regulares, Lopes (2001) não efectuou cruzamentos para os mesmos e optou por manter as suas distinções iniciais, incluindo as dos itens finalizados em *-l* e dos elementos terminados em *-ão*. Contudo, a autora apenas se limitou a dar seguimento à metodologia avançada por Guy (1981a) e corroborada por Scherre (1988), já que ambos os elementos, apesar de pluralizarem com inserção de *-s* e alterações silábicas na maioria dos casos, revelaram comportamentos distintos na marcação PL das realizações dos falantes cariocas.

Na Tabela 24.4 apresentamos o número de ocorrências e respectivos percentuais de marcação, resultantes do cruzamento que efectuámos entre todos os itens das variáveis independentes *saliência fónica* e *tonicidade*. Os resultados permitem verificar que o PA regista perfil idêntico ao da NURC, já que apenas os itens regulares apresentam distintividade entre os três tipos de tonicidade. Quanto aos metafónicos, registam uma única ocorrência oxítone, por oposição a catorze paroxítonas. Como o total de ocorrências não perfaz 30 realizações, é um factor que não será tido em conta nos cruzamentos que se pretendem efectuar, pelo que se manterá a distintividade inicial. Relativamente aos nasais regulares, possuem três tipos de itens: monossílabos átonos; oxítonos ou monossílabos tónicos; e paroxítonos. Inicialmente, pensámos constituir dois factores distintos para estes elementos: (i) factor englobando os monossílabos átonos e os oxítonos, visto a *saliência* recair sobre a última sílaba em ambos; (ii) factor envasando os paroxítonos, cuja *saliência* incide sobre a penúltima sílaba. Contudo, depois de observarmos melhor os últimos, apercebemo-nos que estamos perante casos típicos de sílabas subtónicas finais, isto é, sílabas finais com acentuação fónica secundária.

**Tabela 24.4.** *Saliência* resultante do cruzamento entre *saliência fónica e tonicidade* (todos os dados) [Grupo de factores constituído consoante metodologia de Lopes (2001:143)].

<b>Factores</b>	<b>Exemplo</b>	<b>Nr. total</b>	<b>%</b>
Regulares monossilábicos átonos	o/os	249/274	91
Regulares oxítonos ou monossilábicos tónicos	cacau/cacaús; pai/pais	39/61	64
Regulares paroxítonos	onda/ondas	702/1.626	43
Regulares proparoxítonos	médico/médicos	(5/12)	42
Metafónicos oxítonos	avô/avós	(1/1)	100
Metafónicos paroxítonos	jogo/jogos	(9/14)	64
Nasais regulares monossilábicos átonos	um/uns	34/46	74
Nasais regulares oxítonos	irmão/irmãos; atum/atuns	(12/17)	71
Nasais regulares paroxítonos	homem/homens	23/82	28
-ão/-ões (oxítonos)	razão/razões	(21/27)	78
-l (oxítonos)	mundial/mundiais	(3/10)	30
-r (oxítonos)	jogador/jogadores	42/79	53
-s (oxítonos)	mês/meses; rapaz/rapazes	61/86	71
<b>Totais:</b>		1.201/2.335	51

Scherre (1988:90), ao debruçar-se sobre a inconstância da pluralização apresentada pelos itens proparoxítonos, constatou que a maior distância entre a sílaba tónica e a que vai receber a marca de PL não é condição *sine qua non* para conduzir a menos marcação. Nesta conformidade, não se pode ter como adquirido o princípio de que os itens nasais regulares paroxítonos, pelo facto de não terem a sílaba tónica na última posição, terão que ser menos marcados do que aqueles que a possuem no final da palavra. Se levarmos em linha de conta que a última sílaba detém também uma tonicidade secundária nos nasais regulares paroxítonos, é muito natural que os falantes os entendam como portadores de *saliência* final. Logo, não se torna lógico constituir diferentes factores para os nasais regulares oxítonos e paroxítonos, visto ambos serem marcados, em maior ou menor intensidade, com *tonicidade* na sílaba final. Assim sendo, optámos por adoptar na íntegra a metodologia de Lopes (2001:143) com vista não só à elaboração dos factores que passarão a constituir o grupo de factores *saliência* mas também à possibilidade de

levarmos a cabo as análises e comparações que pretendemos. Ainda assim, convém chamar a atenção para o facto de, no factor dos itens nasais regulares, a autora apenas ter incluído os elementos finalizados em *-ão*, tendo os casos em *-em* ou *-um* sido classificados como oxítonos ou paroxítonos.

A Tabela 24.5 exhibe os resultados da *saliência*, após constituição de factores segundo metodologia de Lopes (2001:143):

**Tabela 24.5.** Efeito da *saliência* (*saliência fónica x tonicidade*) na marcação PL dos itens do SN do PA. [Grupo de factores constituído consoante metodologia de Lopes (2001:143)].

<i>Input desta rodada:</i> 0,523	<i>Log-likelihood:</i> -984,967		<i>Significância:</i> 0,029	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Regulares monossilábicos átonos	249/274	91	12	0,738
Regulares oxítonos ou monossilábicos tónicos	39/61	64	2	0,599
Regulares paroxítonos	702/1.626	43	72	0,415
Regulares proparoxítonos	(5/12)	42	1	-
Nasais regulares	69/145	48	6	0,398
- r	42/79	53	3	0,700
- s	61/86	71	4	0,943
<b>Totais:</b>	<b>1.167/2.283</b>	<b>51</b>	<b>100</b>	<b>-</b>

Os resultados do cruzamento entre os grupos de factores *saliência fónica* e *tonicidade*, de acordo com a metodologia de Lopes (2001:143), evidencia que os oxítonos finalizados em *-s* são os que continuam a propiciar mais a marcação em número, favorecendo-a categoricamente (pr. 0,943). No plano oposto, inibindo moderadamente a pluralização, situam-se dois factores: os nasais regulares (pr. 0,398) e os paroxítonos que pluralizam também regularmente (pr. 0,415). Em plano intermédio, e favorecendo acentuadamente a marcação, encontram-se monossilábicos átonos (pr. 0,738) e os elementos que terminam em *-r* no singular (pr. 0,700). Por fim, e num patamar de neutralidade, encontramos os oxítonos ou monossilábicos tónicos, com tendência mínima para propiciarem a marcação (pr. 0,599).

Nos nossos dados saltam a evidência dois aspectos: (i) o elevado peso que os elementos em *-s* detêm na inserção de marcas, a denotar que o fenómeno da analogia

(Braga & Scherre, 1976:474) não é acolhido no PA; (ii) o também elevado índice de marcação dos monossílabos átonos, comprovando-se que os proclíticos dos substantivos têm tendência a serem interpretados como sílabas iniciais destes, perdendo-se a sua noção de definitude. O fenómeno da pluralização dos itens em *-s* é abordado mais detalhadamente nos pontos que observam a influência na marcação PL motivada pelas variáveis independentes *contexto fonológico posterior* (ponto 5.2.3 do presente trabalho) e *marcas precedentes* (ponto 4.1.2.4 do presente trabalho), uma vez que estes elementos são, na maioria, antecidos por numerais. Quanto à questão da pluralização dos monossílabos átonos, não concordamos com Lopes (2001:141), quando a mesma afirma que este fenómeno revela também perda da informação de PL. Em nossa opinião, ocorre apenas não-realização a nível da estrutura mórfica na marcação do substantivo núcleo do SN, reflectindo, tão-só, o licenciamento fonético de {s} nulo, e não no plano semântico, no qual a mesma permanece activa. De facto, o elevado peso da marcação constatado nos monossílabos átonos, isto é, nos itens proclíticos dos substantivos, é tão somente o reflexo do fenómeno apontado por Guy (1981a:301-302) para o PBV, cuja CPL-var, a exemplo do que sucede com o PA, é fortemente condicionada pelo item colocado na posição imediatamente anterior ao núcleo do SN. O fenómeno, como temos vindo a mencionar, encontra respaldo no sistema de pluralização das línguas do grupo níger-congo atlântico, nas quais predomina o uso de marcadores iniciais, em forma de prefixo ou clítico. Assim, e apesar de não possuímos os pesos relativos para a totalidade dos itens que apresentam elevada distinção fónica na oposição singular/plural, é possível observar que o sistema de marcação PL do PA assenta fortemente na motivação estrutural, que compete, neste aspecto, com motivações do âmbito da *saliência fónica*, sobrepondo-se mesmo à grande maioria destas.

### **5.2.2.3. Efeito da *saliência* (*saliência fónica x tonicidade*) na marcação PL do PA e da NURC**

A Tabela 24.6 permite comparar os resultados do efeito da *saliência* na marcação PL dos itens do SN do PA com os do MRJ (Scherre, 1988:139) da NURC (Lopes, 2001:144-145):

**Tabela 24.6.** *Saliência (saliência fónica x tonicidade)*: contribuição dos factores individuais para a marcação PL no elemento analisado: 3 variedades de português.

Factores	<i>Saliência (saliência fónica x tonicidade)</i>					
	PA		MRJ		NURC	
	Pr.	%	Pr.	%	Pr.	%
Regulares monossilábicos átonos	0,738	91			0,65	99
Regulares oxítonos ou monossilábicos tónicos	0,599	64	0,38	72	0,72	92
Nasais regulares	0,398	48			0,40	49
Regulares paroxítonos	0,415	43	0,17	50	0,37	71
Regulares proparoxítonos	(5/12)	42	0,21	52	0,41	75
Metafónicos	(10/15)	67	0,80	93	0,78	88
–ão/–ões	(21/27)	78	0,59	86	0,71	87
–l	(3/10)	30	0,69	86	0,80	88
–r	0,700	53	0,65	88	0,78	90
–s	0,943	71	0,56	83	0,63	76

Pode dizer-se, de um modo geral, que os sistemas de marcação PL do PA, do MRJ e da NURC, motivados pela *saliência*, evidenciam paralelismo quanto aos itens que beneficiam ou inibem a pluralização, apenas se denotando ligeira dissemelhança nos pesos apresentados. Salvaguardando a diferença exibida pela amalgamação efectuada por Scherre (1988:91), é possível constatar que, em ambos os dialectos, o favorecimento da marcação é determinado, em maior ou menor escala, pelos seguintes factores: monossílabos átonos, oxítonos regulares ou monossílabos tónicos e itens finalizados no singular em *–r* ou *–s*. Opostamente, inibindo a marcação, surgem dois factores: regulares paroxítonos e nasais regulares.

Relativamente aos elementos para os quais não possuímos pesos relativos, é possível observar que os percentuais da marcação do PA deixam antever também um paralelismo entre estas três variedades de português. Assim, os itens que apresentam percentuais de marcação positivos no PA (metafónicos = 67%; *–ão/–ões* = 78%), também inserem positivamente as marcas no MRJ (metafónicos = pr. 0,80; *–ão/–ões* = pr. 0,59) e na NURC (metafónicos = pr. 0,78; *–ão/–ões* = pr. 0,71), enquanto os proparoxítonos, que evidenciam percentual negativo em Almojarife (42%), demonstram ser também inibidores no Rio de Janeiro (pr. 0,21) e em Salvador (pr. 0,41). A única diferença de

monta é evidenciada pelos elementos que finalizam em *-l* no singular, com percentual de marcação negativo no PA (30%) e peso positivo no MRJ (pr. 0,69) e na NURC (pr. 0,80). Já nos pronunciámos sobre as diferentes formas de se pluralizarem os itens finalizados em *-l*, caso sejam oxítonos ou paroxítonos, e o modo como tal poderá levar os falantes a hesitarem na forma de inserirem a marcação de número (nota de fim de texto 182). Assim, é provável que tanto a norma urbana carioca como a norma culta de Salvador, por se encontrarem em estádios mais avançado de aquisição da noção da regra de concordância do que o PA, evidenciem uma maior tendência para a inserção de marcas neste factor.

Se exceptuarmos o caso dos monossílabos átonos, cujo elevado índice de marcação reflecte o respaldo da fala almoxarifana nas línguas do substrato, não deixa de ser evidente que o maior peso de marcação incide sobre os factores relativos a maiores graus de saliência, embora com ligeiras diferenças nos resultados apresentados pelas variedades de Almojarife e Salvador. Analisando a influência da variável independente *saliência* na marcação PL dos grupos popular e universitário, Lopes (2001:146) concluiu que, no primeiro grupo, parece ocorrer uma oposição mais nítida entre os elementos mais e menos salientes, a apontar para uma história de aquisição mais radical de dados menos precisos, insuficientes e variáveis, por transmissão geracional. Como tal, a variação ocorrerá, neste grupo, em maior escala nas formas menos salientes, por oposição à maior incidência de concordância nos elementos em que a oposição singular/plural é mais perceptível.

A fim de confirmar melhor os aspectos relacionados com a relação entre *saliência* e concordância, Lopes (2001:148) procedeu também a comparações entre os grupos popular e universitário da faixa mais idosa de Salvador. Os resultados evidenciaram a existência de gramáticas diferentes nos dois conjuntos de falantes, com a *saliência* proporcionada pela *tonicidade* dos itens a não interferir na concordância do estrato universitário. Este aspecto evidencia que a análise de regras variáveis deverá também ter em conta a relação entre a competência gramatical e a natureza das comunidades discursivas (Bayley, 1994:167), já que estas podem apresentar diferentes frequências de *performance*. Contudo, diferentes efeitos dos factores denunciam a existência de gramáticas distintas e, nesta conformidade, o falante detentor de uma gramática distinta

já não pertence à comunidade discursiva em questão. Assim sendo, diferentes variações motivadas pela mesma variável, no caso a *saliência*, apontarão para situações distintas de transmissão linguística em que os grupos se encontram envolvidos, daí resultando que estes se mostrem mais/menos sensíveis à referida variável.

Tendo em conta que o factor *escolaridade* poderia justificar a diferença de concordância entre grupos, Lopes (2001:152), a exemplo do já efectuado por Scherre (1988:96) para os grupos Primário, Ginásial e Colegial, estudou e comparou o efeito da *saliência* em falantes detentores dos níveis Fundamental, Colegial e Universitário. As nossas comparações com estes níveis escolares terão de ser apenas parciais, visto não possuímos qualquer informante do nível Universitário. Contudo, as mesmas permitirão complementar os resultados das análises já colhidos nas Tabelas 23.3 (relação entre *saliência fónica* e *escolaridade*) e 24.3 (relação entre *tonicidade* e *escolaridade*) e concluir com mais precisão acerca da possibilidade de os informantes do PA possuírem diferentes concordâncias motivadas pela mesma variável: a *saliência fónica*.

#### **5.2.2.4. Relação entre *saliência* (*saliência fónica* x *tonicidade*) e *escolaridade*: PA e NURC**

Na Tabela 24.7 exibimos os resultados do cruzamento entre as variáveis independentes *saliência* e *escolaridade*, tendo em conta a metodologia de constituição de factores adoptada por Lopes (2001:143). Como se vê, apenas se possuem valores em todas as faixas escolares para os monossílabos átonos e paroxítonos regulares, o que não nos permite comparar os nossos resultados como os do MRJ, visto estes factores terem sido amalgamados num único. No estrato analfabeto, outros dois factores evidenciam resultados, mais concretamente, os nasais regulares e os elementos que finalizam em *-s* no singular. Estes últimos apresentam elevados valores de favorecimento à pluralização em todos os escalões de escolarização, apesar de só possuímos o peso da faixa analfabeta (pr. 0,954). Contudo, o alto percentual de marcação nos outros estratos mantém-se constante, fazendo prever que a *escolaridade* não exerce influência na *saliência* destes elementos, uma vez que surgem já bastante marcados desde as gerações antepassadas (escolarização baixa = 100%; escolarização média = 85%; escolarização alta = 100%).

**Tabela 24.7.** Relação entre *saliência* e *escolaridade* na marcação PL dos itens do SN do PA.  
[Grupo de factores constituído consoante metodologia de Lopes (2001:143)]

	<i>Input desta rodada: 0,156</i> <i>Log-likelihood: -166,863</i> <i>Significância: 0,037</i>			<i>Input desta rodada: 0,609</i> <i>Log-likelihood: -295,586</i> <i>Significância: 0,008</i>		
	<b>Analfabetos</b>			<b>Escolaridade baixa</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Regulares monossilábicos átonos	36/40	90	0,517	105/106	99	0,948
Regulares oxítonos ou monossilábicos tónicos	(9/11)	82	-	(15/19)	80	-
Regulares paroxítonos	107/434	25	0,465	232/459	51	0,335
Regulares proparoxítonos	(0/1)	0	-	(1/5)	20	-
Nasais regulares	3/33	9	0,120	37/60	62	0,530
Metafónicos	(4/5)	80	-	(2/3)	68	-
- ão/- ões	(5/9)	56	-	(5/7)	71	-
- l	-	-	-	(0/2)	0	-
- r	(3/19)	16	-	(15/26)	58	-
- s	20/41	49	0,954	(7/7)	100	-
<b>Totais:</b>	<b>187/593</b>	<b>32</b>	<b>-</b>	<b>419/694</b>	<b>60</b>	<b>-</b>
	<i>Input desta rodada: 0,465</i> <i>Log-likelihood: -350,116</i> <i>Significância: 0,010</i>			<i>Input desta rodada: 0,577</i> <i>Log-likelihood: -165,979</i> <i>Significância: 0,015</i>		
	<b>Escolaridade média</b>			<b>Escolaridade alta</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Regulares monossilábicos átonos	71/89	80	0,806	37/39	95	0,041
Regulares oxítonos ou monossilábicos tónicos	(11/27)	41	-	(4/4)	100	-
Regulares paroxítonos	216/498	43	0,437	147/235	63	0,628
Regulares proparoxítonos	(0/2)	0	-	(4/4)	100	-
Nasais regulares	(8/24)	33	-	(21/28)	75	-
Metafónicos	(0/1)	0	-	(4/6)	67	-
- ão/- ões	(5/5)	100	-	(6/6)	100	-
- l	(2/5)	40	-	(1/3)	33	-
- r	(12/19)	63	-	(12/15)	80	-
- s	(22/26)	85	-	(12/12)	100	-
<b>Totais:</b>	<b>347/696</b>	<b>50</b>	<b>-</b>	<b>248/352</b>	<b>71</b>	<b>-</b>

As questões apontadas confirmam que o fenómeno da analogia fenómeno da analogia (Braga & Scherre, 1976:474) não actua sobre os falantes de Almojarife, encontrando-se



a matéria em torno deste fenómeno discutida tanto no estudo das variáveis independentes *contexto fonológico posterior* (ponto 5.2.3 do presente estudo) e *marcas precedentes* (ponto 4.1.2.4 do presente trabalho) como na análise que adiante se fará, quando levarmos a cabo a comparação entre os resultados do PA e do PT.

Quanto aos itens nasais regulares, fornecem também valores para a faixa com escolarização baixa. Assim, estes elementos, que se apresentam nas camadas analfabetas como aqueles que mais inibem a pluralização (pr. 0,120), passam a beneficiar ligeiramente a marcação PL na produção dos indivíduos com escolarização baixa (pr. 0,530), fazendo pressupor que ganham consciência acerca do modo de pluralizar os itens em *-ão*, o que ajudará a subir o peso da concordância dos mesmos.

Relativamente aos paroxítonos regulares, constata-se que os mesmos mantêm um padrão regular ao longo das faixas analfabeta e menos escolarizadas, inibindo ligeiramente a marcação (analfabetos = pr. 0,465; escolarização baixa = pr. 0,335; escolarização média = pr. 0,437), mas que esse padrão se altera no estrato com maior escolarização, que passa a inserir mais as marcas de PL (pr. 0,628). Esta alteração do padrão faz pressupor, tal como acontece na NURC, a existência de uma gramática distinta entre os indivíduos portadores de mais elevado grau de escolarização e os de menor índice de escolaridade, aspecto que se confirma quando olhamos para o peso da marcação dos monossílabos átonos. Efectivamente, estes começam por apresentar um peso neutro no favorecimento da marcação PL em falantes analfabetos (pr. 0,517), mas que sobe drasticamente quando os indivíduos iniciam a sua escolarização (escolaridade baixa = pr. 0,948; escolaridade média = pr. 0,806). Este aspecto é bem elucidativo de como o grau de *saliência fónica* na distinção singular/plural não detém a primazia na influência da pluralização do PA, uma vez que, como referimos antes, os falantes deste ancoram o seu sistema de marcação PL no dos substratos do grupo níger-congo atlântico, que concede ao item colocado na posição imediatamente anterior ao núcleo do SN a primazia na inserção de marcas. Nesta conformidade, o elevado índice de marcação PL revelado por estes elementos não será tão virtual, como o defendem Scherre (1988:113) e Lopes (2001:142). Contudo, o alto grau de escolarização coloca os falantes do PA em contacto com um sistema de concordância mais elaborado, pelo que estendem a concordância aos outros elementos do SN. Deste modo, o protagonismo da

marcação transita dos monossílabos átonos para os elementos que anteriormente se apresentavam como inibidores desta, isto é, os paroxítonos regulares (pr. 0,628), fazendo decair o peso positivo apresentado pelos monossílabos átonos para valores bastante baixos (pr. 0,041), não tanto porque estes elementos inibam a marcação, mas sim pelo facto de passarem a ter outros itens a competirem consigo na inserção de marcas dentro do mesmo grupo de factores. Ainda assim, esta questão vem confirmar na íntegra o pressuposto avançado por Lopes (2001:154) de que, apesar de existir o efeito da saliência na concordância do grupo mais escolarizado, “*a escolarização consegue conter e reduzir a influência dessa variável no fenómeno estudado*”

Para perceber como o fenómeno da *saliência* estabelece paralelos entre o PA e outras variedades de português, procederemos a comparações, em etapas distintas, com os resultados de Baxter (2004:108), para o PT, e Jon-And (2008:4, 2009:5), para o PMp e PCV.

#### **5.2.2.5. Efeito da *saliência* (*saliência fónica x tonicidade*) na marcação PL do PA e do PT**

Para estudo do efeito da *saliência* (*saliência fónica x tonicidade*) na marcação PL do PT, Baxter (2004:108) optou por uma metodologia distinta da dos outros estudos, tendo em linha de conta, para a constituição dos seus factores, não uma leitura apenas sobre a *tonicidade* dos itens singulares mas também o resultado final da pluralização, isto é, se os plurais são tónicos ou átonos. Assim, se atentarmos, por exemplo, ao factor *plural tónico sem mudança no radical* (Tabela 24.8), apercebemo-nos que a forma singular dos itens englobados no mesmo possui, em todos, a tonicidade na última sílaba (p.e. cacau; irmão; atum). Contudo, se observarmos o factor *plural átono sem mudança no radical*, constatamos que o mesmo engloba itens que no singular poderão terminar tanto em sílabas tónicas (p.e. jogador) como em sílabas átonas (p.e. casaa; jovem; médico), apesar de os plurais de todos eles configurarem apenas sílabas finais átonas. Para tanto, Baxter (2004:109) constituiu 6 factores na totalidade: (i) quatro factores que consideram não só a presença ou ausência de sílaba tónica no PL mas também se esta é ou não acompanhada por mudança no radical do item; (ii) um factor para os monossílabos átonos; (iii) e um factor para os plurais cujos singulares terminam em –s. A Tabela 24.8 apresenta os resultados que obtivemos para factores elaborados de acordo com a

metodologia de Baxter (2004:109) e que levou em linha de conta quer o grau de distância fonética entre singular e PL quer a tonicidade das formas plurais:

**Tabela 24.8.** *Saliência* resultante do cruzamento entre *saliência fónica* e *tonicidade* (todos os dados). [Grupo de factores constituído consoante metodologia de Baxter (2004:109)]

<b>Factores</b>	<b>Exemplo</b>	<b>Nr. total</b>	<b>%</b>
Plural tónico com mudança no radical			
1. singular – l	casal/casais	24/37	65
2. – ão/– ões	razão/razões		
Plural tónico sem mudança no radical			
1. oxítonos	cacau/cacaus	51/78	65
2. nasais regulares oxítonos	irmão/irmãos; atum/atuns		
Monossilábicos átonos	o/os	283/320	88
Metafónicos	jogo/jogos	(10/15)	67
Plural átono sem mudança no radical			
1. paroxítonos	casa/casas	772/1.799	43
2. nasais regulares paroxítonos	jovem/jovens		
3. singular – r	jogador/jogadores		
4. proparoxítonos	médico/médicos		
Singular – s (– z)	mês/meses; rapaz/rapazes	61/86	71
<b>Totais:</b>		<b>1.201/2.335</b>	<b>51</b>

Os itens metafónicos registam apenas quinze ocorrências, pelo que não foram incluídos na rodada que visa fornecer os pesos relativos dos factores envasados na variável independente *saliência* (Tabela 24.9).

**Tabela 24.9.** Efeito da *saliência* (*saliência fónica* x *tonicidade*) na marcação PL dos itens do SN do PA. [Grupo de factores constituído consoante metodologia de Baxter (2004:109)].

<i>Input desta rodada: 0,502</i>		<i>Log-likelihood: -988,170</i>		<i>Significância: 0,010</i>	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Plural tónico com mudança no radical	24/37	65	2	0,871	
Plural tónico sem mudança no radical	51/78	65	3	0,620	
Monossilábicos átonos	283/320	88	14	0,613	
Plural átono sem mudança no radical	772/1.799	43	78	0,431	
Singular – s (– z)	61/86	71	3	0,945	
<b>Totais:</b>	<b>1.191/2.320</b>	<b>51</b>	<b>100</b>	<b>-</b>	

Os resultados do grupo de factores constituído de acordo com a metodologia de Baxter (2004:109) apresentam diferenças notáveis em relação aos do grupo de factores elaborado de acordo com a proposta de Lopes (2001:143). De facto, se excluirmos os itens finalizados em *-s* no singular, que continuam a deter o maior peso na marcação PL (pr. 0,945), constata-se agora que os plurais tónicos propiciam mais a pluralização (plurais tónicos com mudança no radical = pr. 0,871; plurais tónicos sem mudança no radical = pr. 0,620) do que os plurais átonos (monossílabos átonos = pr. 0,613; plurais átonos sem mudança no radical = pr. 0,431). Deste modo, os monossílabos átonos deixam de ser o segundo factor mais propiciador da marcação, para cederem o seu lugar aos plurais tónicos, com ou sem mudança no radical do item. Por conseguinte, constata-se que a tonicidade exerce uma maior influência na inserção de marcas nos factores constituídos para estudo do PT, sendo esta mais acentuada quando ocorre, em simultâneo, uma mudança no radical da palavra.

Seguidamente apresentam-se, na Tabela 24.10, os resultados comparados entre o PA e o PT (Baxter, 2004:109):

**Tabela 24.10.** *Saliência (saliência fónica x tonicidade)*: contribuição dos factores individuais para a marcação PL do elemento analisado [2 variedades de português].

Factores	<i>Saliência (saliência fónica x tonicidade)</i>			
	PA		PT	
	Pr.	%	Pr.	%
Plural tónico com mudança no radical	0,871	65	0,88	82
Plural tónico sem mudança no radical	0,620	65	0,52	51
Monossilábicos átonos	0,613	88	0,61	96
Metafónicos	(10/15)	67	0,59	59
Plural átono sem mudança no radical	0,431	43	0,46	45
Singular <i>-s</i> ( <i>-z</i> )	0,945	71	0,35	25

Se exceptuarmos o caso dos itens finalizados em *-s* no singular, simultaneamente os mais favorecedores da marcação no PA (pr. 0,945) e os mais inibidores da mesma no PT (pr. 0,35), todos os outros factores apresentam um paralelismo de marcação, não só na questão do favorecimento ou inibição mas também em termos de valores de pesos relativos. Assim, é notória a semelhança para os itens plurais com terminação tónica, em ambos os dialectos, propiciarem mais a marcação do que os elementos plurais com

sílaba átona no final. O propiciamento da marcação é também mais evidente nas duas variedades quando a pluralização envolve mudança no radical da palavra.

Face à oposição registada entre itens mais e menos salientes em alguns estratos escolarizados, e que aponta para uma história de aquisição mais radical de dados menos precisos, insuficientes e variáveis, por transmissão geracional, Lopes (2001:148) comparou dados dos falantes populares e universitários da geração mais idosa de Salvador. Na mesma linha de pensamento, Baxter (2004:109) decidiu analisar a influência da saliência na marcação PL de todas as gerações de Monte Café, tendo observado que a faixa mais idosa apresenta um padrão de variação distinto do das faixas menos idosas, em virtude de a primeira ter recebido um *input* mais acentuado de DLP's da L2 transmitida por falantes adultos das L1 africanas.

#### **5.2.2.6. Relação entre *saliência* (*saliência fónica* x *tonicidade*) e *idade* na marcação PL dos itens do SN do PA e do PT**

As análises efectuadas para apurar o modo como a saliência actua na marcação produzida pelas diferentes classes escolarizadas mostram que, no PA, existe também um padrão distinto de marcação entre falantes mais escolarizados e menos escolarizados ou analfabetos. Ora, estando os diversos níveis de escolarização conotados, por norma, às diferentes faixas etárias, optámos por analisar o comportamento geracional da *saliência* na marcação PL do PA. Deste modo, seguimos a metodologia de Baxter (2004:109) para a constituição dos factores, já que a mesma permite uma comparação mais precisa com os resultados obtidos para os tongas.

Na tabela 24.11 os valores para a comunidade e diferentes escalões etários de Almojarife. Como se tem visto, os itens que finalizam em *-s* no singular revelam tendência para serem os que mais beneficiam a marcação PL no PA. A fim de confirmarmos esta tendência, optámos por levar à análise VARBRUL as realizações registadas por este factor nas FE-1 e FE-3, ainda que as mesmas não perfaçam um total de trinta ocorrências. Os resultados confirmam, uma vez mais, a forte influência destes itens na inserção de marcas, mantendo-se esta constante ao longo das gerações (FE-3 = pr. 0,978; FE-2 = 0,907; FE-1 = 0,982). Este padrão de propiciamento fora já evidenciado por Braga (1977), para os falantes do triângulo mineiro, Ponte (1979), para

a comunidade de Porto Alegre e Guy (1981a), para os informantes semianalfabetos cariocas, mas não encontra paralelo no PT (Baxter, 2004:109), no qual surge sempre como elemento bastante inibidor da marcação PL (faixa etária idosa = pr.0,26; faixa etária intermédia = pr. 0,35; faixa etária jovem = pr. 0,28). Baxter (2004:112), na linha de Scherre (1988:244), justifica a inibição em termos aquisicionais sob as perspectivas da analogia e da haplologia sintáctica. A este propósito, gostaríamos de acrescentar que a questão em torno da marcação dos itens em *-s* no singular mereceu análises exaustivas por parte de Scherre (1988:115), tendo a mesma chegado a testar a influência na marcação registada pelos elementos que apresentam ausência de concordância (p.e. *mês/mês*), concordância não padrão (p.e. *mês/mese*) e concordância padrão (*mês/meses*). A autora conjecturou ainda em torno não só da falta de ligação entre a inserção do *-s* final e a fronteira de morfema mas também de certas particularidades da formação do PL destes itens, algumas delas de ocorrência rara no português e, diremos mesmo, nulas nos nossos dados. Após as suas reflexões, a autora concluiu que se estará perante a actuação da *saliência fónica* (Naro & Lemle, 1976:240-241; Lemle & Naro, 1977:47), a qual relega os elementos terminados em *-s* para a base da hierarquia da inserção de marcas, em virtude de “*estarem envolvidos num processo haplológico que consiste em eliminar sílabas ou palavras iguais ou parecidas... e por apresentarem comportamento bastante especial*” (Scherre, 1988:120) que os leva a não admitirem pluralização em determinadas situações (p.e. *cós, giz*). Contudo, nunca em nenhuma fase das suas análises a autora aventou a possibilidade de a presença/não-presença de marcas poder ancorar em interferências da responsabilidade dos substratos. Quanto ao elevado peso da marcação exibido pelos falantes do PA, podemos referir que qualquer dos fenómenos apontados por Baxter (2004:112) não é acolhido por estes, estando a discussão em volta da haplologia também debatida nas análises que efectuámos no presente trabalho sobre as variáveis independentes *contexto fonológico posterior* (ponto 5.2.3) e *marcas precedentes* (ponto 4.1.2.4).

Por seu lado, os monossílabos átonos surgem como o factor que mais inibe a marcação na faixa etária mais velha (pr. 0,236) do PA, registando esta, tal como acontece no PT, a tendência para a omissão dos definidos (exemplo [206]). Em adição, outro aspecto que concorrerá para a inibição provocada pelos monossílabos átonos na

FE-3 terá a ver com o facto de quer determinados idiomas do grupo níger-congo atlântico quer o santomense possuírem poucos monossilabos e, funcionalmente, os que não são aglutinados aos substantivos representarem preposições ou advérbios.

**Tabela 24.11.** Relação entre *saliência* e *idade* na marcação PL dos itens do SN do PA.  
[Grupo de factores constituído consoante metodologia de Baxter (2004:109)].

	<i>Input desta rodada: 0,502</i> <i>Log-likelihood: -988,170</i> <i>Significância: 0,010</i>			<i>Input desta rodada: 0,609</i> <i>Log-likelihood: -295,586</i> <i>Significância: 0,008</i>		
	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Plural tónico com mudança no radical	24/37	65	0,871	(10/16)	63	-
Plural tónico sem mudança no radical	51/78	65	0,620	(9/13)	69	-
Monossilábicos átonos	283/320	88	0,613	68/78	87	0,236
Metafónicos	(10/15)	67	-	-	-	-
Plural átono sem mudança no radical	772/1.799	43	0,431	104/433	24	0,494
Singular – s (– z)	61/86	71	0,945	15/27	56	0,978
<b>Totais:</b>	<b>1.201/2.335</b>	<b>51</b>	<b>-</b>	<b>206/567</b>	<b>36</b>	<b>-</b>

	<i>Input desta rodada: 0,375</i> <i>Log-likelihood: -359,215</i> <i>Significância: 0,009</i>			<i>Input desta rodada: 0,779</i> <i>Log-likelihood: -412,838</i> <i>Significância: 0,004</i>		
	<b>FE-2 (41 – 60 anos)</b>			<b>FE-1 (20 – 40 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Plural tónico com mudança no radical	(5/9)	56	-	(9/12)	75	-
Plural tónico sem mudança no radical	16/35	46	0,522	26/30	87	0,802
Monossilábicos átonos	77/100	77	0,609	138/142	97	0,418
Metafónicos	(4/6)	67	-	(6/9)	67	-
Plural átono sem mudança no radical	252/605	42	0,452	416/761	55	0,463
Singular – s (– z)	19/30	63	0,907	27/29	93	0,982
<b>Totais:</b>	<b>373/785</b>	<b>48</b>	<b>-</b>	<b>622/983</b>	<b>63</b>	<b>-</b>

Nas gerações seguintes, os monossilabos átonos revelam alguma flutuação, passando a favorecer ligeiramente a marcação na FE-2 (pr. 0,609), para voltarem, depois, a inibi-la, mas mais suavemente do que na geração longeva (pr. 0,418). O aumento do contacto com o sistema mais elaborado de concordância do PE transmitirá à geração intermédia uma maior noção do uso dos determinantes, bem como do seu sistema de pluralização. Posteriormente, com o distender da regra de concordância aos restantes elementos do

SN, a noção de marcação da pluralização poderá oscilar entre os vários itens que compõem a cadeia sintagmática, prevalecendo o princípio de que a mesma se torna redundante caso seja aplicada em mais do que um elemento (Kiparsky, 1972:195).

Refira-se que, nos tongas (Baxter, 2004:109), o padrão revelado pelos monossílabos átonos em direcção à aquisição da regra de concordância é mais consistente, começando por se mostrar ligeiramente desfavorecedor à inserção do PL (faixa etária idosa = pr. 0,44), para passar depois a favorecedor da mesma nas gerações seguintes (faixa etária intermédia = pr. 0,82; faixa etária jovem = pr. 0,79).

No nossos dados, apenas mais um factor concedeu valores para a FE-3: os plurais átonos sem mudança no radical, que exercem um efeito neutro na marcação PL (pr. 0,494). Este efeito mantém-se consistente ao longo de todas as gerações, mas sem atingir valores que favoreçam a inserção de marcas (FE-2 = pr. 0,452; FE-1 = 0,463). Neste factor estão incluídos os paroxítonos, que constituem a maioria das ocorrências da língua portuguesa e representam, neste grupo, 77% do total das realizações. Note-se também que esta classe de palavras não só prefigura os graus de saliência singular/plural menos distintos do português como vai também ao encontro do grosso das ocorrências de determinadas línguas do grupo níger-congo atlântico, sobretudo do ramo bantu, e do santomense, isto é, de paroxítonas que não possuem marca de PL, visto esta incidir sobre os marcadores semânticos que as antecedem ou ser meramente referencial. Assim, não surpreende que a pluralização de idênticos tipos de ocorrências surja comprometida na *interlíngua* que emerge em situação de contacto por parte de falantes de L1's do grupo níger-congo atlântico ou santomense L1.

Note-se ainda que a maior parte dos informantes desta geração é analfabeta ou semianalfabeta e que não revela tendência a produzir palavras esdrúxulas (apenas 1 ocorrência não marcada, cf. Tabela 24.7) nem a pluralizar os oxítonos em *-r*, incluídos neste factor em virtude de apresentarem PL átono (três elementos marcados em dezanove ocorrências, cf. Tabela 24.7). Aliás, esta faixa etária chega mesmo a tratar estes itens da mesma forma que os terminados em vogal (cf. exemplo [270]), mostrando que, por vezes, não faz a distinção entre o som [r] (inexistente nos substratos do grupo níger-congo atlântico do santomense e sujeito a fenómeno de lambdacismo neste, cf.



exemplo [210]) e os elementos vocálicos em final de proposição, que, ocupam sempre a última posição das palavras nos primeiros:

[270] PA:  
*compra duas vadô, são três conto e tal* [MANOH3]  
“compra dois voadores, são três contos e tal”

Assim, será só nas FE-2 e FE-1, em virtude de estas sofrerem maior pressão social de cima para baixo, que os falantes passam a realizar mais proparoxítonos – sempre em reduzido número – e a aplicar a marcação nos elementos finalizados em *-r* no singular.

O factor *plural átono sem mudança no radical* apresenta no PT (Baxter, 2004:109) um comportamento bastante idêntico ao do PA, com valores ligeiramente abaixo do patamar da neutralidade, isto é, sem nunca atingirem um grau de favorecimento da marcação em qualquer das gerações (faixa etária idosa = pr.0,40; faixa etária intermédia = pr. 0,39; faixa etária jovem = pr. 0,46).

No que concerne aos plurais metafónicos, dado o seu número reduzido de ocorrências, não nos é permitido tirar conclusões sustentáveis. No PT (Baxter, 2009:109), a quantidade de realizações também é mínima ao longo das gerações, mas o seu total perfaz vinte e nove ocorrências para a comunidade. O resultado final aponta para um ligeiro favorecimento da marcação PL (pr. 0,59), a que não será alheio o facto de apresentarem um grau acentuado de saliência na realização da pluralização. No PA, o percentual da comunidade aponta também para a possibilidade de um moderado propiciamento à inserção de marcas (67%).

O último factor a ser analisado diz respeito aos plurais tónicos. Os que implicam mudança no radical da palavra apenas apresentam valores para a comunidade, na qual o resultado final denuncia um favorecimento acentuado da pluralização (pr. 0,871). Este valor é bastante semelhante ao do PT (pr. 0,88), que apresenta ainda resultados para as duas faixas etárias mais recentes, sempre categoricamente favorecedores da inserção de marcas e a determinarem um padrão que se mantém geracionalmente constante (faixa etária intermédia = pr. 0,93; faixa etária jovem = pr. 0,84).

Quanto aos plurais tónicos sem mudança no radical da palavra, são estes os elementos que apresentarem o valor que mais propicia a marcação na faixa etária mais idosa de Monte Café (pr. 0,93). Em Almoxarife, contudo, o mesmo factor detém apenas treze

realizações na FE-3, pelo que possuímos somente valores percentuais de marcação (70%), mas que sugerem que este tipo elementos são marcados desde as gerações mais longevas. Lembremos aqui, uma vez mais, que o maior acervo vocabular dos substratos do santomense é constituído por palavras graves, pelo que vocábulos oxítonos do português serão imediatamente percebidos pelos falantes em situação de contacto devido à sua “estranha” tonicidade. Deste modo, numa perspectiva aquisicional, faz todo o sentido que esta saliência marque presença evidente nos DLP’s da L2 dos falantes adultos, através quer dos plurais tónicos sem mudança no radical da palavra quer dos singulares terminados em *-s*. Aliás, face à ausência de pesos para os factores dos plurais tónicos na FE-3, o protagonismo da inserção de marcas é assumido, neste estrato etário, pelos itens terminados em *-s* no singular (pr.0,978).

Relativamente ao padrão geracional dos plurais tónicos sem mudança no radical, veja-se que o mesmo revela uma tendência para aquisição progressiva da marcação PL (FE-2 = pr. 0,522; FE-1 = pr. 0,802). No PT, a transição geracional é feita com perda de marcação (faixa etária intermédia = pr. 0,57; faixa etária jovem = pr. 0,45), justificada com o incremento do favorecimento da pluralização que implica mudança no radical da palavra (Baxter, 2004:112).

Antes de nos pronunciarmos sobre o panorama geral do efeito da saliência na marcação PL, torna-se pertinente tecer alguns considerandos em torno da questão que se prende com a relação estabelecida entre pesos relativos e percentuais de marcação, sobretudo na FE-3. Como se pode constatar, os valores apresentados pelos factores monossilábicos átonos e itens em *-s* no singular permitem verificar as seguintes correlações: maior percentagem de marcação igual a menor peso relativo e menor percentagem de marcação sinónimo de maior peso relativo. Tendo também em atenção os resultados patentes nas Tabelas 24.5 e 24.11, verifica-se que esta particularidade acontece porque, como temos vindo a referir, a maioria dos monossílabos átonos serem artigos em primeira posição sintagmática (relembre-se que os SN’s produzidos pela faixa etária mais idosa são de estrutura bastante simplificada), na qual são marcados numa frequência que varia dos 87%, na geração mais velha, até aos 97%, no estrato etário mais jovem. Assim, é em relação aos outros itens que ocorrem em primeira posição que os monossílabos átonos vão apresentar um peso relativo ligeiramente

desfavorável. Contrariamente, os singulares em *-s* configuram elementos nucleares na quase totalidade das realizações, que apenas muito esporadicamente ocorrem na primeira posição da cadeia sintagmática. Assim sendo, em relação aos outros elementos nucleares que ocorrem também em posição não-inicial do SN, a percentagem de elementos finalizados em *-s* acaba por deter um peso verdadeiramente categórico na marcação PL.

Mesmo não possuindo alguns valores para determinados factores, é possível concluir que o efeito da *saliência* na marcação PL do PA, ao longo do seu percurso geracional, ancora nos achados dos estudos de aquisição, que têm defendido que o desenvolvimento da morfologia flexional tem o seu início nas oposições salientes (Bayley, 1994:167). Deste modo, e conforme nos revelam os valores gerais da comunidade, as oposições mais salientes são adquiridas precocemente, sendo a marcação favorecida pelas distinções morfofonológicas plural/singular mais acentuadas, como acontece com os oxítonos terminados em *-s* no singular (pr. 0,945), os plurais tónicos com mudança no radical (pr. 0,871) e os plurais tónicos sem mudança no radical (pr. 0,620).

Contrariamente, a aquisição do PL das formas menos salientes ocorre em período mais tardio, como nos revelam os valores dos monossílabos átonos, que devem o seu moderado favorecimento da pluralização (pr. 0,613) à aquisição do artigo nas FE-2 e F-1, e dos plurais átonos sem mudança no radical (pr. 0,431), que, geracionalmente, nunca atingem números que apontem para um favorecimento da inserção de marcas de PL.

#### **5.2.2.7. Efeito da *saliência* (*saliência fónica x tonicidade*) na marcação PL dos itens do SN: 3 variedades de português**

Os factores constituídos por Jon-And (2008:4, 2009:5), tanto para o PMp como para o PCV, são idênticos aos de Baxter (2004:109), embora a primeira tenha atribuído outras designações a alguns deles. De facto, a única alteração aqui registada prende-se como o factor denominado pelo segundo autor como PL átono sem mudança no radical, que inclui quer os paroxítonos e proparoxítonos regulares quer os itens finalizados em *-r* no singular. Jon-And (2008:4, 2009:5) optou por constituir dois factores para estes elementos: um incluindo as palavras graves e esdrúxulas regulares; outro destinado aos elementos finalizados em *-r* no singular, visto serem oxítonos.

Os totais de ocorrências e respectivos percentuais de marcação para cada um dos factores, segundo a metodologia de Jon-And (2008:4, 2009:5), podem ser observados na Tabela 24.12:

**Tabela 24.12.** *Saliência* resultante do cruzamento entre *saliência fónica* e *tonicidade* (todos os dados). [Grupo de factores constituído consoante metodologia de Jon-And (2008:4, 2009:5)]

<b>Factores</b>	<b>Exemplo</b>	<b>Nr. total</b>	<b>%</b>
Oxítonos ou monossilábicos tónicos regulares	cacau/cacaus; irmão/irmãos; pai/pais	51/78	65
Oxítonos ou monossilábicos tónicos irregulares	razão/razões; mundial/mundiais	24/37	65
Paroxítonos/proparoxítonos regulares	onda/ondas; médico/médicos	730/1.720	42
Paroxítonos irregulares	fácil/fáceis; jogo/jogos	(10/15)	67
Monossílabos átonos	o/os	283//320	88
– r oxítonos	jogador/jogadores	42/79	53
– s oxítonos	mês/meses; rapaz/rapazes	61/86	71
<b>Totais:</b>		<b>1.201/2.335</b>	<b>51</b>

Os paroxítonos irregulares resumem-se aos plurais metafónicos, visto não possuímos itens terminados em *-l* no singular que registem este tipo de *saliência fónica*. Assim, uma vez mais tivemos que retirar estas ocorrências da rodada seguinte. Os pesos relativos referentes aos factores incluídos na variável independente *saliência*, segundo metodologia de Jon-And (2008:4, 2009:5), encontram-se disponíveis na Tabela 24.13:

**Tabela 24.13.** *Efeito da saliência (saliência fónica x tonicidade)* na marcação PL dos itens do SN do PA. [Grupo de factores constituído consoante metodologia de Jon-And (2008:4, 2009:5)].

<i>Input desta rodada: 0,507</i>	<i>Log-likelihood: -980,222</i>		<i>Significância: 0,011</i>	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Oxítonos ou monossilábicos tónicos regulares	51/78	65	3	0,613
Oxítonos ou monossilábicos tónicos irregulares	24/37	65	2	0,877
Paroxítonos/proparoxítonos regulares	730/1.720	42	74	0,417
Monossílabos átonos	283//320	88	14	0,615
– r oxítonos	42/79	53	3	0,712
– s oxítonos	61/86	71	4	0,946
<b>Totais:</b>	<b>1.167/2.283</b>	<b>51</b>	<b>100</b>	<b>-</b>

Devido à semelhança dos factores constituídos por Baxter (2004:109) e Jon-And (2008:4, 2009:5), os valores da Tabela 24.13 são em tudo idênticos aos da Tabela 24.9, como se compreende. Contudo, é possível observar que os itens finalizados em *-r* no singular se apresentam, em termos comunitários, como moderadamente favorecedores da inserção de marcas de número (pr. 0,712). Avançámos a possibilidade de estes itens serem tratados da mesma forma que os terminados em vogal na geração mais idosa (exemplo [270]), concluindo-se que, por vezes, os falantes deste escalão etário não fazem a distinção entre o som *-r* e as vogais em final de palavra. Os resultados do PA vêm agora mostrar que, de facto, nas FE-2 e FE-1, em virtude maior pressão social sofrida de cima para baixo, se passa a aplicar mais a pluralização nestes elementos. Deste modo, confirma-se, uma vez mais, a influência da saliência na marcação, já que os itens paroxítonos e proparoxítonos regulares nunca chegam a atingir o patamar do favorecimento à inserção de marcas, enquanto os elementos em *-r*, por serem oxítonos, isto é, serem mais percebidos pelos falantes, acabam por se projectar para valores que beneficiam a pluralização. Contudo, interessa aqui fazer uma ressalva relativamente ao PMp (Jon-And, 2008:4) e ao PCV (Jon-And, 2009:5), nos quais o comportamento dos mesmos elementos evidencia um desenvolvimento geracional distinto, já que apresenta valores moderadamente inibidores da marcação em termos comunitários (PMp = pr. 0,33; PCV = 0,36). Nesta conformidade, pressupõe-se que o tratamento análogo dado à vogal final e ao item /r/ se cristaliza desde as gerações antepassadas e se mantém constante ao longo do tempo em Maputo e Cabo Verde.

Na Tabela 24.14 apresentamos os valores para as quatro variedades africanas de português que temos vindo a tratar, a fim de mais facilmente se estabelecerem comparações entre os diferentes padrões de marcação determinados pela variável *saliência*. Constata-se que ocorre uma grande semelhança entre os padrões de marcação determinados pela variável *saliência* do PT, do PMp e do PCV, quer no que concerne aos factores que beneficiam a marcação PL quer no que diz respeito aos factores inibidores da mesma. Relativamente ao PA, o padrão é também extensível a este para a maioria dos factores. Contudo os falantes de Almoxarife mostram-se mais sensíveis a captarem a saliência na totalidade dos oxítonos, o que não acontece com o PT e o PMp, no que respeita aos elementos que pluralizam com a inserção de *-es* final, isto é, os itens

cujos singulares apresentam *-r* ou *-s* no seu final. Ainda assim, é possível constatar tudo o que foi referido aquando da comparação que estabelecemos entre o PA e o PT (ponto 5.2.2.5): (i) as oposições mais salientes são adquiridas primeiro e as formas menos salientes em tempo mais tardio; (ii) a marcação PL é guiada pelas distinções morfofonológicas plural/singular mais acentuadas.

**Tabela 24.14.** *Saliência (saliência fónica x tonicidade):* contribuição dos factores individuais para a marcação PL do elemento analisado [4 variedades de português].

Factores	PA		PT		PMp		PCV	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%
Oxítonos ou monossilábicos tónicos irregulares	0,877	65	0,88	82	0,82	97	0,45	65
Oxítonos ou monossilábicos tónicos regulares	0,613	65	0,52	51	0,52	92	0,56	92
Monossílabos átonos	0,615	88	0,61	96	0,78	99	0,69	97
Paroxítonos irregulares	(10/15)	67	0,59	59	(5/5)	100	-	-
Paroxítonos/proparoxítonos regulares	0,417	42			0,35	82	0,46	79
Singular em <i>-r</i>	0,712	53	0,46	45	0,33	77	0,36	72
Singular em <i>-s</i>	0,946	71	0,35	25	0,22	64	0,08	24

Acerca da hipótese 2, que levantámos para a variável independente *tonicidade*, é possível confirmar que, tal como havíamos predito, em termos de comunidade, os falantes de Almocharife inserem a marcação PL preferencialmente nos oxítonos e nos monossílabos de uso tónico. A hierarquia da inserção da pluralização nos itens que menos beneficiam a marcação não seguiu, contudo, a linha que propuséramos, já que os monossílabos átonos propiciam mais a inserção de marcas do que os paroxítonos. Para os proparoxítonos não foi possível obter a confirmação do peso exercido na pluralização, devido ao seu exíguo número de realizações. Ainda assim, é possível observar que o seu percentual de marcação (42%) se situa ao mesmo nível dos paroxítonos (43% – Tabela 24.1).

O estudo da relação entre *tonicidade* e grau de escolarização dos falantes veio confirmar que, no PA, as formas que apresentam maior diferenciação material fónica na oposição singular/plural, por serem mais salientes, favorecem mais a concordância PL do que os itens em que essa percepção não é tão óbvia. Logo, e ao que tudo indica, os

informantes de Almojarife adquirem as formas da língua a partir de contextos linguísticos mais salientes.

A observação do comportamento da *saliência* (*saliência fónica x tonicidade*) na inserção de marcas revelou que os monossílabos átonos, ao contrário do que se pressupunha, apresentam elevado índice de marcação em termos comunitários. A causa, quanto a nós, tem respaldo no sistema de marcação PL dos ancestrais substratos do grupo níger-congo atlântico, e que leva os proclíticos dos substantivos a serem interpretados como sílabas iniciais destes e a carregarem a marca de número PL, perdendo-se assim a sua noção de definitude. Nesta conformidade, reforça-se o pressuposto de que o sistema de marcação PL do PA assenta fortemente na motivação estrutural, que se sobrepõe às do tipo fónico ou semântico.

A análise da influência da *saliência* na marcação PL do itens do SN do PA levantou ainda mais duas questões: (i) o elevado índice de inserção de marcas nos itens finalizados em *-s* no singular não faz pressupor a intervenção do fenómeno da haplologia sintáctica na fala de Almojarife, tendo o mesmo sido apontado por Scherre (1988:244) como uma das causas principais da inibição da pluralização nos referidos elementos; (ii) o peso da marcação nos itens finalizados em *-r* no singular faz pressupor que os informantes mais idosos do PA os tratam da mesma forma que os terminados em vogal. A causa terá a ver não só com o facto de o som [r] não existir em determinadas línguas do grupo níger-congo atlântico mas também com a relação estabelecida com as palavras destes idiomas, que terminam sempre em elemento vocálico.

Gostaríamos ainda de referir que, tal como mencionou Scherre (1988:141), esta é uma variável polémica e que continuam a permanecer “*em aberto o número de subcategorias e a ‘sistematicidade’ da influência de algumas delas*”. Nesta conformidade, gostaríamos ainda de ter analisado a influência da *saliência* subtónica na marcação PL, pelos menos para confirmar não só se esta tem alguma palavra a dizer na organização hierárquica das categorias intermédias, cujos valores desenham padrões inconstantes de factor para factor, mas também para confirmar melhor até que ponto a concordância entre os elementos do SN do PA se encontra em processo de mudança linguística. Na falta de oportunidade de levar a cabo tais estudos no presente trabalho, propomo-nos efectua-los em pesquisas futuras sobre o PA ou outras variedades de português.

### 5.2.3. Variável independente *contexto fonológico posterior*

A fim de podermos observar o modo como a variável independente *contexto fonológico posterior* actua na configuração da CPL-var do SN do PA, começámos por efectuar uma rodada geral em que incluimos todos os factores da Tabela 12.4, à excepção dos que apresentavam ocorrências reduzidas (*factor X* = 4 realizações, *factor R* = 17 realizações; *factor Z* = 8 realizações; *factor G* = 16 realizações). Os resultados finais tinham como finalidade fazer actuar a Lei da Parcimónia, reduzindo a extensão do grupo de factores e tornando-o mais robusto à análise. Como se verifica, todos estes factores dizem respeito a consoantes, pelo que voltaremos a lançar mão de alguns deles quando necessitarmos de amalgamar factores. A este propósito, lembre-se que os factores com menos de cinco ocorrências não são de todo confiáveis, pelo que o *factor X* será o único que nunca fará parte das amalgamações que efectuaremos.

Os resultados da primeira rodada VARBRUL podem ser confirmados na Tabela 25.1:

**Tabela 25.1.** Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN do PA.

<i>Input desta rodada: 0,504</i>		<i>Log-likelihood: -981,009</i>		<i>Significância: 0,011</i>	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Pausa final	134/294	46	13	0,656	
Pausa interna	106/349	30	15	0,533	
Vogais	234/409	57	18	0,492	
P	121/197	61	9	0,476	
K	216/312	69	14	0,569	
B	29/34	85	1	0,562	
D	79/186	42	8	0,374	
F	34/48	71	2	0,552	
T	51/101	51	4	0,386	
J	15/35	42	2	0,444	
L	19/38	50	2	0,279	
M	51/86	59	4	0,450	
N	23/66	35	3	0,380	
S	46/102	45	4	0,334	
V	14/35	40	1	0,486	
<b>Totais:</b>	<b>1.172/2.295</b>	<b>51</b>	<b>100</b>	<b>-</b>	



Partindo do princípio que os factores referentes às pausas interna e final (cf. metodologia de Lopes, 2001:227) e às vogais estão impossibilitados de sofrerem qualquer tipo de amalgamação, para as efectuarmos, teremos que dirigir a nossa atenção para o comportamento das consoantes, seguindo a metodologia dos trabalhos científicos anteriormente levados a cabo para estudo da influência da variável independente *contexto fonológico posterior* na marcação PL dos itens do SN. Assim, à medida que necessitarmos de fazer amalgamações entre factores respeitantes a consoantes, apontaremos os trabalhos nos quais nos basearemos. Como se pode constatar também na Tabela 25.1, os próprios pesos relativos referentes aos factores que dizem respeito a algumas consoantes apresentam valores bastante aproximados, o que, desde logo, aconselha que se teste a amalgamação de factores.

Mas antes de nos debruçarmos sobre as consoantes, convém referir que, acerca das pausas, Cedergren (1973:43) entende que elas não exercem qualquer influência na marcação PL. Por seu lado, Poplack (1980a:61; 1981:61) defende que a inserção é desfavorecida pela pausa e favorecida pela ocorrência de consoante e vogal. Quanto a Guy (1981a:27-28), considera que o efeito da pausa é arbitrário e imprevisível, mas que, no caso do PB, favorece a inserção, por oposição à consoante e à vogal. Aliás, esta preconização vai ao encontro não só daquilo que foi defendido por Braga (1977:40), que havia já apresentado resultados condizentes não só com os de Guy (1981a) no que concerne à inserção mas também com os de Scherre (1988:249), que defende um favorecimento ligeiro das pausas na marcação PL.

Nos nossos dados, como se pode ver, as pausas parecem influenciar a marcação PL do SN, confirmando aquilo que havia sido sugerido por Guy (1981a:27-28) para o PB e Scherre (1988:249) para o MRJ. Assim sendo, os princípios teorizados por Cedergren (1973) e Poplack (1980a, 1981) não encontram eco no PA. De facto, ao que tudo indica, a pausa final é o elemento que mais favorece a marcação nos nossos dados (pr. 0656), estabelecendo-se, neste aspecto, um paralelo de marcação entre o PA, a NURC e o PT. Por seu lado, a pausa interna, de acordo com os nossos resultados, revela paralelos com o PB, o MRJ, a NURC e o PT, já que mostra um valor mais neutro, mas com ligeira tendência para favorecer a inserção PL (pr. 0,533).

Como se referiu no ponto 3.6.3, no qual descrevemos os critérios que subjazeram à constituição dos factores para a variável independente *contexto fonológico posterior*, desde os estudos de Braga (1977:31) que vem sendo demonstrado que as vogais em *contexto fonológico seguinte* têm tendência para inibir a marcação. Guy (1981a:166), Scherre (1988:248) e Baxter (2004:114) chegariam a idênticas conclusões, mas Lopes (2001:226) apresentaria resultados ligeiramente distintos, uma vez que, nos seus achados, este factor se mostrou neutro para o favorecimento da pluralização, sustentando, de certa forma, os resultados de Poplack (1980a:61). No nosso caso, os valores aproximam-se bastante dos que foram apresentados por Baxter (2004:114), uma vez que, tanto no PA (pr. 0,492) como no PT (pr. 0,48), as vogais surgem como itens ligeiramente inibidores da marcação, mas a roçarem o patamar da neutralização.

No que concerne às consoantes, apenas três surgem nos nossos dados a favorecerem levemente a marcação: a oclusiva oral dorso-palatal [k] (pr. 0,569), a oclusiva oral bilabial [b] (pr. 0,562) e a constrictiva fricativa lábio-dental [f] (pr. 0,552). Como se vê, estamos perante dois pares distintos de consoantes não só quanto ao modo de articulação (oclusivas orais *versus* constrictivas fricativas) mas também quanto aos pontos onde são articulados (labiais *versus* palatais). Contudo, identificam-se todas umas com as outras no que respeita ao traço de classe principal não-soante. Por outro lado, não parece confirmar-se o pressuposto de Braga (1977:71-72) de que as palatais desfavorecem a marcação, já que os nossos resultados apresentam a dorso-palatal [k] como favorecendo a mesma. Relembre-se que Scherre (1988:255), por sua vez, apesar de reconhecer que as consoantes surdas favorecem um pouco mais a marcação do que as sonoras, considerou débil a influência do ponto de articulação, da sonoridade e da caixa de ressonância na inserção da marca PL. Os resultados de Lopes (2001:228-229) e Baxter (2004:114) confirmariam igual padrão na NURC e no PT e, ao que tudo indica, poderão ser corroborados também na nossa pesquisa. Efectivamente, se exceptuarmos o comportamento da líquida lateral ápico-alveolar [l], bastante inibidora da marcação (pr. 0,279), as restantes consoantes apresentam valores moderadamente inibidores ou que se aproximam da neutralidade, sejam elas surdas ou sonoras. Ainda assim, duas das consoantes do nosso *corpus* que favorecem ligeiramente a inserção de marca PL são

surdas: a oclusiva oral dorso-palatal [k] (pr. 0,569) e a constrictiva fricativa lábio-dental [f] (pr. 0,552).

### **5.2.3.1. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo os traços de fonte**

Os resultados apresentados na Tabela 25.1 tinham como finalidade ajudarem-nos a definir parâmetros acerca da compartilha de traços das consoantes, com vista a amalgamá-las em factores mais robustos. Face à diversidade e incongruência dos resultados obtidos, optámos por ter em linha de conta os traços distintivos compartilhados pelas três consoantes que marcam positivamente a pluralização e agrupar todas segundo os mesmos.<sup>186</sup> Contudo, numa primeira etapa, seguindo a linha metodológica de outras pesquisas anteriores, decidimos observar igualmente o comportamento sincrónico e diacrónico das consoantes dos nossos dados segundo os seus traços de fonte, a fim de termos uma panorama que não só nos permita comparações mais precisas com trabalhos anteriores sobre CPL-var no SN mas também nos faculte uma visão do desenvolvimento geracional desta. Nesta conformidade, para o grupo de factores que inclui os traços de fonte das consoantes, distribuímo-las por dois factores, isto é, surdas ([p], [t], [k], [f], e [s]) e sonoras ([b], [d], [g], [m], [n], [v], [z], [ʒ], [R] e [l]). O traço que as divide em estridentes ou não-estridentes não foi por nós levado em linha de conta, uma vez que não é uniforme nas três consoantes que tomámos como ponto de referência. As rodadas que forneceram valores para o grupo de factores que englobam factores de acordo com os traços de fonte das consoantes apresentaram os resultados constantes na Tabela 25.2.

Braga (1977:31) e Guy (1981a:166) consideram que a consoante e a vogal desfavorecem a marcação. O segundo, contudo, indica que, no que se refere às consoantes, o traço surdo tem tendência a marcar positivamente o PL, enquanto o traço sonoro o inibe. Scherre (1988:254) confirmaria estes achados e Baxter (2004:115) apontaria a neutralidade das surdas, embora com valores ligeiramente positivos no que concerne à marcação, por oposição à ténue inibição provocada por vogais e consoantes sonoras. Lopes (2001:228), por seu lado, seria a única a analisar dados em que as vogais mostram ligeira tendência para favorecerem a marcação (pr. 0,54), opondo-se à

moderada inibição das consoantes (pr. 0,48). Ainda assim, ambos os elementos possuem valores que nos permitem observar a neutralidade da sua influência na concordância PL.

Baxter (2004:111), por seu lado, defende que grande parte das inibições fonológicas à marcação PL que ocorre na geração mais idosa do PT pode ser explicada em termos de tendências aquisicionais ou influências na transferência do PtgL2 dos africanos adultos, que providenciaram os DLP's aos falantes desta faixa etária. Geracionalmente, porém, não pudemos confirmar este pressuposto, visto não termos obtido resultados para a faixa etária mais idosa, já que a ferramenta VARBRUL considerou o *contexto fonológico seguinte* irrelevante para a marcação PL na mesma.

**Tabela 25.2.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços de fonte), na marcação PL dos itens do SN.

	<i>Input desta rodada:</i> 0,468 <i>Log-likelihood:</i> -961,643 <i>Significância:</i> 0,019			<i>Input desta rodada:</i> 0,202 <i>Log-likelihood:</i> -182,718 <i>Significância:</i> 0,047		
	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	134/294	46	0,672	17/61	28	-
Pausa interna	106/349	31	0,537	19/114	17	-
Vogais	234/409	57	0,490	52/85	61	-
Consoantes surdas	468/760	62	0,478	74/172	43	-
Consoantes sonoras	258/521	50	0,415	45/136	33	-
	<i>Input desta rodada:</i> 0,351 <i>Log-likelihood:</i> -354,271 <i>Significância:</i> 0,008			<i>Input desta rodada:</i> 0,704 <i>Log-likelihood:</i> -412,044 <i>Significância:</i> 0,001		
	<b>FE-2 (41 – 60 Anos)</b>			<b>FE-1 (20 – 40 Anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	49/112	44	0,698	68/121	56	0,557
Pausa interna	40/123	33	0,546	47/112	42	0,481
Vogais	59/120	49	0,520	123/204	60	0,430
Consoantes surdas	161/275	59	0,440	233/313	74	0,579
Consoantes sonoras	64/154	42	0,404	149/231	65	0,433

**Obs.:** FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

Contudo, as FE's-2 (41-60 anos) e FE's-1 (20-40 anos) do PA e do PT exibem valores bastante díspares. Assim, a FE-2 apresenta um paralelo com a regra geral de marcação PL não só com comunidade em si mas também com idêntica faixa etária do

PT. De facto, estabelece-se uma hierarquia similar de marcação em todas estas variáveis, a qual reflecte, segundo Baxter (2004:111), dois aspectos a ter em conta: (i) nivelamento morfológico dos DLP's da L2 dos adultos que funcionaram como *input* para a actual geração; (ii) L1 actual revelando uma tipologia fonotáctica traduzida em sílabas marcadamente abertas, a qual impõe, simultaneamente, restrições às sequências consonantais. Face a estas condições, o ponto de entrada mais forte para a marcação PL acaba por cair em itens na posição final da cadeia sintagmática, isto é, antecedendo a pausa final (comunidade PA = pr. 0,672; FE-2 PA = 0,698; FE-2 PT = pr. 0,64) e a pausa interna (comunidade PA = pr. 0,537; FE-2 PA = 0,546; FE-2 PT = pr. 0,53).

Como terceiro elemento mais propiciador da marcação em número surgem as vogais (comunidade PA = pr. 0,490; FE-2 PA = 0,520; FE-2 PT = pr. 0,50), com valores muito próximos da neutralidade, sejam eles ligeiramente propiciadores ou inibidores da inserção PL. Por fim, surgem as consoantes, com as sonoras a desfavorecerem mais a marcação (comunidade PA = pr. 0,415; FE-2 PA = 0,404; FE-2 PT = pr. 0,39) do que as surdas (comunidade PA = pr. 0,478; FE-2 PA = 0,440; FE-2 PT = pr. 0,49). O decréscimo no propiciamento à marcação apresentado pelas vogais fica a dever-se, no caso do PA, à FE-1, já que este item passa de terceiro elemento mais favorável à marcação para elemento mais inibidor da mesma (pr. 0,430). Como tal, o valor estabiliza-se na comunidade perto do ponto da neutralidade, mas com tendência a inibir suavemente a inserção de marca PL. Refira-se que, relativamente à tendência crescente para a inibição apresentada pelas vogais, o PA acaba por se assemelhar ao MRJ, cujo público infantil (a geração menos idosa) propicia menos o PL (pr. 0,39) do que o público adulto (pr. 0,48), quando o *contexto fonológico seguinte* é representado por um item vocálico (Scherre, 1988:248).

Continuando a observar o padrão de comportamento da geração mais jovem do PA, no que diz respeito à marcação PL no SN, vê-se que as consoantes surdas passam a factor mais favorecedor desta (pr. 0,579), relegando a pausa final para a segunda posição no que concerne ao favorecimento (pr. 0,557), logo seguida da pausa interna, que apresenta agora um peso a tender para a ligeira inibição (pr. 0,481). Por fim, com um valor bastante próximo da inibição daquele que é apresentado pelas vogais, surgem as consoantes sonoras (pr. 0,433), embora propiciando um pouco mais a marcação. Nos

dados de Baxter (2004:114), estes dois elementos também se apresentam como os mais inibidores na geração mais jovem (pr. vogais 0,46; pr. consoantes sonoras 0,21). O autor entende que a explicação para o facto poderá assentar na fonotáctica do umbundu, o substrato do PT (Baxter, 2004:113). De facto, como neste idioma não há sequências consonantais, estas são substituídas por obstruções pré-nasalizadas à passagem do ar no tracto bucal. Aliando este aspecto à ausência do som [z] no mesmo dialecto, os falantes adultos de umbundu poderão experienciar dificuldades na inserção do /s/ antes de uma vogal ou consoante sonora no momento da aquisição do PtgL2, acabando este aspecto por se traduzir numa restrição à fixação da sequência PL+CONSOANTE SONORA ou VOGAL. A este propósito, gostaríamos de chamar a atenção para alguns paralelos entre os aspectos referidos por Baxter (2004:113) e a estrutura do santomense (o substrato do PA): (i) tendência para a estrutura dissilábica padrão CVCV, isto é, ausência de encontros consonantais; (ii) não distinção singular/plural para a maioria dos substantivos e seus qualificadores, ou seja, pluralização sem recurso ao /s/ final, sendo a mesma efectuada com utilização de estratégias lógico-semânticas (reduplicação ou semântica do determinante pós-nominal) ou inserção pré-nuclear do item *inen*, que funciona como marcador PL para nomes de traço [+humano], expressando um significado definido que diz respeito a um grupo não-individualizado familiar ao ouvinte e ao falante (ponto 1.3.2.2.1 do presente trabalho).

Abordando agora o padrão da comunidade de Almojarife, constata-se que apenas as pausas favorecem a marcação, com a pausa final a exhibir uma tendência moderada para ajudar à inserção de flexão de número no SN (pr. 0,672) e a pausa interna a apresentar um valor quase neutro (pr. 0,537). Ao constituirmos os factores de acordo com os traços de fonte [+vozeado] e [-vozeado], ficámos a perceber que as consoantes sonoras se apresentam como o item que mais inibe a marcação (pr. 0,415), logo seguidas das consoantes surdas (pr. 0,478). Assim sendo, a ligeira tendência revelada pelas consoantes [k], [b] e [f] para favorecerem a marcação foi diluída pela força inibidora das outras consoantes surdas ou sonoras, levando mesmo a que, tanto vozeadas (pr. 0,478) como não-vozeadas (0,415) inibam ligeiramente a inserção PL no SN, em tendência que se superioriza à das vogais, uma vez que estas se apresentam a roçar a neutralidade (pr. 0,490). Assim sendo, o PA estabelece paralelos com o PB e o PMn no que concerne ao

facto de vogais e consoantes desfavorecerem a marcação PL. Também possui similaridades com estes, com o MRJ e com o PT em virtude de o traço surdo das consoantes propiciar mais a marcação do que o traço sonoro. Aliás, este último dialecto (Baxter, 2004:114) apresenta um comportamento muitíssimo similar ao do PA, no que diz respeito ao peso exercido na marcação tanto por vogais (PA = pr. 0,490; PT = pr. 0,48) como por consoantes sonoras (PA = pr. 0,415; PT = pr. 0,40). Contudo, como se disse já, as consoantes surdas do PA não atingem o limiar na marcação positiva (pr. 0,478), como acontece na NURC e no dialecto dos tongas (pr. 0,53 para ambas as variedades).

Mas, de tudo o que se referiu, o que se torna evidente é que, à excepção do peso relativo da pausa final (pr. 0,672), todos os outros valores se mostram muito próximos do ponto de neutralização. Como tal, confirma-se na íntegra o postulado de Scherre (1988:255) de a sonoridade exercer fraca influência na marcação PL.

#### **5.2.3.2. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando as consoantes num único factor**

Nos dados de Scherre (1988:248) as vogais e consoantes possuem pesos equilibrados e a roçarem a neutralidade (pr. vogais 0,48; pr. consoantes 0,43). Contudo, no estudo de Lopes (2001:228) as vogais alcançam valores tenuemente favorecedores da inserção PL (pr. 0,53), por oposição à tendência ligeiramente inibidora revelada pelas consoantes (pr. 0,48). Contudo, esta diferença não é significativa, pelo que se pode dizer que a NURC, tal como o MRJ, apresenta pesos equilibrados e próximos do patamar da neutralidade para vogais e consoantes. Assim sendo, interessa também observar como se comporta o PA neste aspecto, pelo que, a exemplo da metodologia levada a cabo nos trabalhos das duas autoras, e a fim de percebermos o padrão geral das consoantes face ao das vogais, decidimos amalgamar o factores consoantes surdas e consoantes sonoras num único factor e comparar os valores deste com os dos elementos vocálicos. Os resultados obtidos podem ser observados na Tabela 25.3:

**Tabela 25.3.** Efeito do *contexto fonológico posterior* (englobando consoantes num único factor), na marcação PL dos itens do SN.

	<i>Input desta rodada:</i> 0,468 <i>Log-likelihood:</i> -962,812 <i>Significância:</i> 0,017			<i>Input desta rodada:</i> 0,202 <i>Log-likelihood:</i> -182,718 <i>Significância:</i> 0,047		
	Comunidade			FE-3 (+60 anos)		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	134/294	46	0,674	17/61	28	-
Pausa interna	106/349	31	0,539	19/114	17	-
Vogais	234/409	57	0,492	52/85	61	-
Consoantes	726/1.281	57	0,451	119/308	39	-

	<i>Input desta rodada:</i> 0,3541 <i>Log-likelihood:</i> -355,301 <i>Significância:</i> 0,007			<i>Input desta rodada:</i> 0,698 <i>Log-likelihood:</i> -412,257 <i>Significância:</i> 0,049		
	FE-2 (41 – 60 Anos)			FE-1 (20 – 40 Anos)		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	49/112	44	0,695	68/121	56	0,558
Pausa interna	40/123	33	0,547	47/112	42	0,488
Vogais	59/120	49	0,519	123/204	60	0,428
Consoantes	225/429	52	0,428	382/544	70	0,517

**Obs.:** FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

De um modo geral, as consoantes (pr. 0,451) continuam a propiciar menos a marcação do que as vogais (pr. 0,492) na comunidade, apesar de ambos os factores se revelarem quase neutros, mas com incipiente tendência para a inibição. No entanto, “*diversos trabalhos sobre o cancelamento do –s morféxico, e até do não morféxico, têm mostrado que a diferença probabilística entre o efeito da consoante e o da vogal seguintes não é muito forte e, às vezes, inexistente*” (Scherre, 1988:242). Por conseguinte, pode dizer-se que o padrão do PA, no que concerne ao comportamento de vogais e consoantes em *contexto fonológico seguinte*, revela as mesmas tendências dos dados estudados por Cedergen (1973:43), Poplack (1981a:61, 1981:376), Braga & Scherre (1976:476), Braga (1977:31), Guy (1981a:166), Scherre (1988:248), Lopes (2001:226) e Baxter (2004:114) na marcação PL dos itens do SN.

Contudo, os itens vocálicos dos nossos dados continuam a marcar mais o PL do que os consonantais, pelo que, neste aspecto o PA apresenta o mesmo tipo de comportamento do EPR (Poplack, 1980a), do PB (Guy, 1981a), do MRJ (Scherre, 1988)



e da NURC (Lopes, 2001). Ainda assim, os nossos resultados mostram que as vogais não chegam a atingir o patamar positivo de favorecimento à marcação observado no EPR e na NURC (pr. 0,53), quedando-se pela inibição ténue (0,492), tal como acontece no PB e no MRJ (pr. 0,48). Aliás, para o EPR, Poplack (1980a) apresentou também valores favorecedores da marcação para as consoantes, o que não acontece em qualquer destas variedades de português.

Geracionalmente, os valores respeitantes à pausa final, pausa interna e vogais são muito semelhantes aos exibidos na Tabela 25.3, como se compreenderá. Dado que não se obtiveram valores para a FE-3, considerada irrelevante para a marcação, apenas nos podemos concentrar nos valores fornecidos pela duas faixas etárias mais jovens de Almojarife. Observando ambas, é possível constatar que o padrão entre elas é totalmente oposto, uma vez que o único factor a exhibir alguma constância é o da pausa final, que marca diacronicamente de forma positiva a inserção da pluralidade (FE-2 = pr. 0,695; FE-1 = pr. 0,558). Quanto aos outros factores, invertem completamente a tendência revelada inicialmente na FE-2. Assim, os factores que começam por favorecer a marcação nesta faixa etária (pausa interna = pr. 0,547; vogais = pr. 0,519) passam depois a inibidores na faixa etária seguinte (pausa interna = pr. 0,488; vogais = pr. 0,428). Inversamente, o item que desfavorece a inserção PL na FE-2 (consoantes = pr. 0,428) torna-se elemento favorecedor da mesma na FE-1 (consoantes = pr. 0,517).

Quando comparados com os dados do MRJ (Scherre, 1988:248), as nossas ocorrências revelam alguma aproximação aos mesmos apenas no que concerne ao comportamento diacrónico das vogais, que desfavorece a inserção PL (pr. adultos 0,48; pr. crianças 0,39). Quanto aos outros factores, ou se mantêm estáveis ou apresentam uma tendência inversa à do PA. Assim, no MRJ, as consoantes (pr. adultos 0,43; pr. crianças 0,43) e a pausa interna (pr. adultos 0,56; pr. crianças 0,56) conservam a tendência geracional anterior, mas a pausa interna aumenta o seu peso favorecedor da inserção PL (pr. adultos 0,53; pr. crianças 0,62), o que não acontece no PA.

Nesta rodada, a nossa atenção incidiu essencialmente sobre as consoantes, tendo-se procedido à amalgamação dos factores consoantes surdas e consoantes sonoras num único factor, o das consoantes. Como se constata, apesar de a geração menos idosa mostrar que estas têm influência ligeiramente positiva na marcação PL do item

anterior (pr. 0,517), tal marcação não é suficientemente forte para projectar as consoantes para o patamar do favorecimento à marcação na comunidade (pr. 0,450). Assim sendo, as consoantes mantêm a tendência ligeira para inibir a marcação já revelada na FE-2 (pr. 0,428), confirmando-se como o factor que menos favorece a inserção PL na comunidade. Neste aspecto, convém ainda apontar a proximidade de valores revelada entre o PA (pr. 0,451), o MRJ (pr. 0,43) e a NURC (pr. 0,48), todos desfavorecendo ligeiramente a marcação PL.

Ainda relativamente aos nossos números, foi possível verificar, com o teste do qui-quadrado, que os valores gerais dos *log-likelihoods* registados para a comunidade nas Tabelas 25.2 e 25.3, apesar de serem inferiores ao do *log-likelihood* da Tabela 25.1, não permitem a incorporação dos grupos de factores com 5 ou 4 factores à análise geral, já que os novos grupos de factores se revelaram estatisticamente significativos. Deste modo, a variável independente *saliência fónica* que incorpora os factores da Tabela 25.1 deve ser mantida para efeito do cálculo dos pesos que actuam na marcação PL dos itens do SN.

### **5.2.3.3. Efeito do contexto fonológico posterior na marcação plural dos itens do SN, englobando as pausas num único factor**

Em Scherre (1988:248) é possível constatar também que os valores das pausas interna e final são bastante aproximados (pr. pausa interna 0,56; pr. pausa final 0,53), o que a levou a amalgamá-las num único factor. Os nossos resultados desaconselham esta metodologia, mas, ainda assim, e a exemplo do efectuado por Lopes (2001:227/228), optámos por testar as consequências de idêntica amalgamação, reduzindo o grupo de factores *contexto fonológico posterior* a apenas três factores (Tabela 25.4).

Se nos dados de Lopes (2001:228) a amalgamação dos factores pausas num único desfez a diferença entre os factores iniciais, tornando o *factor pausas* neutro (pr. 0,53) e nivelado com o das vogais (pr. 0,54), o mesmo não sucedeu nas nossas ocorrências, as quais confirmam que, na comunidade de Almocharife, o *factor pausas* se revela como o único que tem algum peso para a inserção PL no SN do PA (pr. 0,605). Quanto às vogais e consoantes, apresentam valores idênticos e próximos da neutralização (pr. 0,491 e pr. 0,450, respectivamente), mostrando que estes itens pouco interferem na pluralização, apesar de as vogais inibirem menos a marcação do que as consoantes. Neste aspecto, a

configuração da marcação do PA assemelha-se mais à do MRJ do que à da NURC (vogais = pr. 0,54; consoantes = pr. 0,48), uma vez que, nos dados de Scherre (1988:249) e Guy (1981a:166), as vogais apresentam um peso neutro (pr. 0,50 para ambos os casos) e as consoantes revelam um resultado a inibir muito pouco a inserção PL (pr. 0,44 e pr. 0,46, respectivamente). Já no que concerne à pausa, esta favorece levemente a marcação (pr. 0,56 e pr. 0,55, respectivamente). Contudo, um aspecto é de ter em conta, quando se olham para os resultados em termos de pausa interna e pausa final: nos dados de Scherre (1988:248), a segunda (pr. 0,53) tem uma influência ligeiramente inferior à primeira (pr. 0,56) na marcação PL, enquanto nas ocorrências do PA (pausa interna = 0,539; pausa final = 0,674), da NURC (pausa interna = 0,52; pausa final = 0,63) e do PT (pausa interna = 0,54; pausa final = 0,58) sucede o inverso.

**Tabela 25.4.** Efeito do *contexto fonológico posterior* (englobando pausas num único factor), na marcação PL dos itens do SN.

	<i>Input desta rodada: 0,466 Log-likelihood: -966,857 Significância: 0,012</i>			<i>Input desta rodada: 0,200 Log-likelihood: -187,756 Significância: 0,005</i>		
	<b>COMUNIDADE</b>			<b>FE-3 (+60 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausas	240/643	37	0,605	36/175	21	-
Vogais	234/409	57	0,491	52/85	61	-
Consoantes	726/1.281	57	0,450	119/308	39	-
	<i>Input desta rodada: 0,354 Log-likelihood: -357,155 Significância: 0,007</i>			<i>Input desta rodada: 0,684 Log-likelihood: -412,153 Significância: 0,040</i>		
	<b>FE-2 (41 – 60 Anos)</b>			<b>FE-1 (20 – 40 Anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausas	89/235	38	0,621	115/233	49	0,527
Vogais	59/120	49	0,516	123/204	60	0,430
Consoantes	225/429	52	0,428	382/544	70	0,515

**Obs.:** FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

Relembre-se que Cedergren (1973:43) entendeu que as pausas não exercem qualquer influência na marcação PL. Poplack (1981a:61), por seu lado, apresentou um padrão totalmente diferente no que concerne à influência da pausa na marcação PL do EPR, já que a mesma a condiciona notoriamente (pr. 0,37). Em oposição a esta, são as vogais

que vão favorecer ligeiramente a inserção PL (pr. 0,60), enquanto as consoantes se quedam pela quase neutralidade da marcação (pr. 0,53). No que concerne às pausas, as nossas realizações não acompanham os resultados de nenhum destes trabalhos.

Geracionalmente, não obtivemos valores para o peso das pausas na marcação PL da FE-3. Nos nossos dados, há uma tendência para as pausas favorecerem mais a marcação na FE-2 (pr. 0,621) do que na FE-1 (pr. 0,527), estabilizando no patamar do favorecimento moderado, em termos de comunidade (pr. 0,605). Este aspecto revela um percurso inverso ao do MRJ (Scherre, 1988:249), cuja tendência é para um aumento do favorecimento à marcação (adultos = pr. 0,57; crianças = pr. 0,62).

Quanto a uma possível amalgamação de factores de acordo com a Tabela 25.4, ela também está fora de hipótese, uma vez que o cálculo do qui-quadrado com recurso aos valores dos *log-likelihoods* para a comunidade (Tabelas 25.1 e 25.4) permitiu constatar que o grupo de factores amalgamado se revela estatisticamente significativo, pelo que não deve ser considerado para efeitos de análise.

#### **5.2.3.4. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo os traços de classe principal**

As tabelas 25.2, 25.3 e 25.4 não só evidenciaram que a diferença probabilística entre os efeitos da consoante e da vogal em contexto posterior não é relevante para a inserção da marca PL mas revelaram também que o peso de ambos os itens para essa mesma inserção é quase neutro. Assim sendo, Scherre (1988:254) optou por observar o comportamento das consoantes em contexto seguinte em termos de traços distintivos, a fim de tentar perceber o tipo de influência que estes têm na inserção do *-s* morfémico PL. Refira-se, a este propósito, que outros autores, nomeadamente Braga (1977:40) e Guy (1981a:166) haviam já procedido a análises detalhadas do mesmo tipo.

Dando seguimento a este tipo de metodologia, decidimos constituir um grupo de factores que reunisse as consoantes de acordo com os traços de classe principal soante e não-soante, visto o segundo ser compartilhado por todas as consoantes que favorecem positivamente a marcação PL nos SN's dos nossos dados (Tabela 25.1). O traço soante caracteriza os sons que não registam uma obstrução significativa à passagem do fluxo de ar, mas implicam uma vibração espontânea das cordas vocais. Assim sendo, os sons estão sujeitos a um vozeamento que pode ser mais pronunciado [+sonoro] ou mais

inibido [-sonoro]. Contudo, estes traços distintivos não devem ser confundidos como os analisados anteriormente (ponto 5.3.2.1 do presente trabalho), isto é, [+vozeado] e [-vozeado], uma vez que estes se referem à zona onde os sons são produzidos, apesar de implicarem também vibração das cordas vocais.

Como consoantes soantes temos então a líquida lateral [l], a líquida vibrante [R] e as nasais [m] e [n]. Como não-soantes surgem-nos quer as oclusivas orais [p], [b], [t], [d], [k] e [g] quer as constrictivas fricativas [f], [v], [s], [z] e [ʒ].

O grupo de factores que inclui as consoantes de acordo com os seus traços de classe principal forneceu os resultados da Tabela 25.5:

**Tabela 25.5.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços de classe principal), na marcação PL dos itens do SN.

Factores	Comunidade			FE-3 (+60 anos)		
	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
	Pausa final	134/294	46	0,674	17/61	28
Pausa interna	106/349	31	0,539	19/114	17	-
Vogais	234/409	57	0,491	52/85	61	-
Consoantes soantes	106/206	51	0,415	17/48	35	-
Consoantes não-soantes	620/1.075	58	0,468	102/258	40	-

Factores	FE-2 (41 – 60 Anos)			FE-1 (20 – 40 Anos)		
	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
	Pausa final	49/112	44	0,696	68/121	56
Pausa interna	40/123	33	0,547	47/112	42	0,486
Vogais	59/120	49	0,519	123/204	60	0,432
Consoantes soantes	31/68	46	0,469	58/90	64	0,385
Consoantes não-soantes	195/360	54	0,420	323/457	70	0,541

Obs.: FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

Os valores da Tabela 25.5 pouco diferem dos da tabela 25.2, tornando evidente que o traço de sonoridade, comum aos traços de cavidade e traços de classe principal,

condiciona o comportamento das consoantes em ambos os grupos de factores, aproximando-as em termos de influência na marcação PL da comunidade. Assim sendo, as consoantes soantes (traço distintivo [+sonoro]) revelam tendência a inibirem mais a marcação (pr. 0,415), tal como sucede com as consoantes sonoras (traço distintivo [+vozeado][+sonoro]), cujo pr. é 0,415. Por seu lado, as consoantes não-soantes (traço distintivo [-sonoro]) propiciam um pouco mais a marcação PL (pr. 0,468), acontecendo o mesmo com as consoantes surdas (traço distintivo [-vozeado][-sonoro]), que apresentam o pr. de 0,478. Contudo, como se pode ver, todos estes números continuam muito próximos do valor da neutralidade em termos de influência na marcação PL.

Os restantes factores deste grupo continuam a desenhar o padrão já antes verificado: a pausa final é o elemento que mais favorece a marcação (pr. 0,674), logo seguida da pausa interna (pr. 0,539), cujo peso de marcação, apesar de positivo, já rasa a neutralidade em termos de influência na inserção de marcas PL. Por fim, as vogais situam-se no patamar da inibição, mas em valor insignificante e igualmente bastante próximo do ponto neutro (pr. 0,491).

Observando agora os valores das faixas etárias de Almojarife, pelos valores percentuais confirma-se também que os traços de classe principal das consoantes não exercem peso na marcação PL do SN, tendo a faixa etária FE-3 sido rejeitada pela ferramenta VARBRUL. Na geração intermédia (41-60 anos), é possível verificar que apenas as consoantes inibem a inserção da pluralidade, com as soantes a propiciarem-na um pouco mais (pr. 0,469) do que as não-soantes (pr. 0,420). Por seu lado, a pausa final apresenta o valor mais favorecedor da marcação (pr. 0,696), seguida da pausa interna (pr. 0,547) e das vogais (0,519), estando estas muito próximas do ponto da neutralidade. Por seu lado, a geração mais jovem leva a cabo uma inversão na forma de marcar a pluralidade, fazendo baixar a influência da pausa final (pr. 0,562) e projectando as consoantes não-soantes para o segundo lugar na hierarquia do favorecimento à marcação (pr. 0,541). Em consequência de tal, a pausa interna e as vogais pendem para valores inibidores da marcação (pausa interna = pr. 0,486; vogais = pr. 0,432), acompanhando as consoantes soantes neste aspecto, que se apresentam como os elementos que menos favorecem a inserção PL (pr. 0,385).

A variável independente *contexto fonológico posterior*, englobando consoantes segundo os traços de classe principal, não foi analisada em trabalhos precedentes, pelo que não é possível estabelecer comparações com os mesmos neste aspecto.

Atentando ainda ao valor do *log-likelihood* para a comunidade (Tabela 25.5), após efectuado o teste do qui-quadrado, verifica-se também que o grupo de factores *contexto fonológico posterior* englobando as consoantes segundo os traços de classe principal não pode ser incorporado à análise para obtenção de resultados gerais, já que a amalgamação em questão se revela estatisticamente significativa.

### **5.2.3.5. Efeito do *contexto fonológico posterior* na marcação plural dos itens do SN, englobando consoantes segundo os traços coronal e não-coronal**

Seguidamente, observámos também o comportamento das consoantes segundo os traços de cavidade, uma vez que três das que tomámos como azimutes orientativos para a constituição dos novos factores, isto é, as consoantes [k], [f] e [b], compartilham o traço não-coronal. Portanto, são produzidas sem elevação da língua acima da posição neutra e na direcção dos dentes incisivos superiores, à região alveolar ou ao palato duro. Por outro lado, se atentarmos a alguns dos nossos resultados, torna-se visível que existe uma evidente proximidade de pesos relativos de acordo com os pontos de articulação quer das coronais quer das não-coronais:

(i) coronais:

(a) ápico-dentais: [d] = pr. 0,380; [n] = pr. 0,373; [t] = pr. 0,327;

(b) ápico-alveolares: [s] = pr. 0,326; [l] = pr. 0,294;

(ii) não-coronais:

(c) bilabiais: [m] = pr. 0,422; [p] = pr. 0,414; [b] = pr. 0,411;

(d) lábio-dentais: [f] = pr. 0,555; [v] = pr. 0,472.

Assim sendo, para esta rodada constituíram-se dois grupos de factores que englobam as consoantes, ou seja, o factor das coronais, incluindo os seus pontos de articulação (ápico-dentais [t], [d] e [n]; ápico-alveolares [s], [z] e [l]; e ápico-palatal [ʃ]), e o factor das não-coronais, abarcando igualmente os seus pontos de articulação (bilabiais [p], [b] e [m]; lábio-dentais [f] e [v]; dorso-velares [k] e [g]; e dorso-uvular [R]). O que se

espera obter com esta rodada é a confirmação de que as não-coronais são mais favoráveis à inserção do PL do que as coronais, visto que as três consoantes que favorecem positivamente a marcação ([k], [b] e [f] – cf. Tabela 25.1) compartilham o segundo traço distintivo.

O grupo de factores formado com as consoantes de acordo com os seus traços de cavidade permitiu obter os resultados que se transcrevem na Tabela 25.6, e que confirmam o que se esperava na comunidade: as consoantes não-coronais (pr. 0,522) propiciam mais a marcação PL no SN do PA do que as coronais, ultrapassando o peso das vogais (pr. 0,489) e atingindo mesmo um valor positivo bastante próximo do da pausa interna (pr. 0,533). Este grupo de factores apresenta já alguma distintividade, uma vez que a pausa final continua a exercer boa influência na marcação (pr. 0,670), enquanto as consoantes coronais caem para valores consideravelmente inibidores da inserção PL (0,364). Assim sendo, ao que tudo indica, parecem asilar nos traços distintivos de cavidade as causas para a inserção da marca PL no SN do PA, quando se analisa esta em termos de *contexto fonológico seguinte*, e não nos traços de fonte nem nos traços de classe principal. Por outras palavras, a marcação PL no SN poderá ser condicionada pelo modo, zonas e pontos de articulação das consoantes em *contexto fonológico posterior*, parecendo, assim, não se confirmar, no PA, o postulado de Scherre (1988:255) de que os pontos de articulação das consoantes em *contexto fonológico posterior* exercem pouca influência na inserção da pluralidade. Deste modo, importa agora confirmar, ou não, esta premissa.

Contudo, antes de se observar a forma como o modo, as zonas e os pontos de articulação das consoantes em *contexto fonológico posterior* actuam sobre a marcação PL, convém também perceber como os traços de cavidade se comportam geracionalmente. Olhando para o padrão da comunidade, é possível constatar a sua semelhança com o da FE-2. De facto, ambos os padrões apresentam a pausa interna como o elemento mais favorecedor da inserção PL (comunidade = pr. 0,670; FE-2 = pr. 0,698), logo seguida da pausa interna (comunidade = pr. 0,533; FE-2 = pr. 0,545). Também o factor menos favorecedor da marcação, isto é, as consoantes coronais, é comum a ambas (comunidade = pr. 0,364; FE-2 = pr. 0,324). Quanto ao não-paralelismo, é exibido pelas vogais (comunidade = pr. 0,489; FE-2 = pr. 0,516) e consoantes não-coronais (comunidade = pr. 0,522; FE-2 = pr. 0,509), com as primeiras passando de elementos inibidores



a favorecedores da marcação e as segundas a manterem-se no patamar do favorecimento, mas perdendo alguma força.

**Tabela 25.6.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços coronal e não-coronal), na marcação PL dos itens do SN.

	<i>Input desta rodada: 0,470</i> <i>Log-likelihood: -955,557</i> <i>Significância: 0,013</i>			<i>Input desta rodada: 0,169</i> <i>Log-likelihood: -145,642</i> <i>Significância: 0,004</i>		
	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	134/294	46	0,670	17/61	28	-
Pausa interna	106/349	31	0,533	19/114	17	-
Vogais	234/409	57	0,489	52/85	61	-
Consoantes não-coronais	484/742	65	0,522	82/166	49	-
Consoantes coronais	242/539	46	0,364	36/129	28	-

	<i>Input desta rodada: 0,362</i> <i>Log-likelihood: -349,688</i> <i>Significância: 0,007</i>			<i>Input desta rodada: 0,702</i> <i>Log-likelihood: 409,879</i> <i>Significância: 0,004</i>		
	<b>FE-2 (41 – 60 Anos)</b>			<b>FE-1 (20 – 40 Anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	49/112	44	0,698	68/121	56	0,551
Pausa interna	40/123	33	0,545	47/112	42	0,477
Vogais	59/120	49	0,516	123/204	60	0,425
Consoantes não-coronais	155/252	62	0,509	247/324	76	0,600
Consoantes coronais	72/182	40	0,324	134/228	58	0,410

**Obs.:** FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

O factor das não-coronais, pelo facto de incluir no seu seio as consoantes [k], [b] e [f], apresenta um valor favorecedor da marcação superior ao das vogais e consoantes coronais, que a inibem. Como na geração FE-2 o peso das não-coronais é inferior ao das vogais, a responsabilidade da variação cabe à FE-1, que projecta estas para o topo do favorecimento à marcação (pr. 0,600), deixando para a pausa final um lugar secundário na pluralização (pr. 0,551). Assim, se na FE-2 surgem 4 factores a favorecerem a marcação e apenas um a desfavorecê-la, na faixa etária mais jovem já é possível constatar a existência de apenas três factores inibindo a inserção PL, de acordo com a seguinte ordem decrescente de propiciamento: pausa interna (pr. 0,477), vogais (pr. 0,425) e consoantes coronais (pr. 0,410).

### **5.2.3.6. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo o modo de articulação: 6 factores**

Face aos achados de que a marcação PL poderá ser condicionada pelo modo, zonas e pontos de articulação das consoantes em *contexto fonológico posterior*, passamos então a analisar a influência que o primeiro poderá ter na inserção da pluralidade. Desta feita, constituímos três grupos de factores no que concerne às consoantes: o factor das obstruintes oclusivas [p], [b], [t], [d], [k], [g], [m] e [n]; o factor das obstruintes constrictivas [f], [v], [s], [z] e [ʒ]; e o factor das soantes líquidas [l] e [R]. A rodada com estas variáveis independentes concedeu os resultados que transcrevemos na Tabela 25.7, os quais evidenciam que na comunidade de Almocharife, se exceptuarmos a tendência para favorecer a marcação revelada pela pausa final (pr. 0,674), os restantes pesos relativos apresentam pouca diferença probabilística entre si, já que todos eles se aproximam do ponto neutro da inserção de marcas de PL. Assim sendo, apenas a pausa interna apresenta um valor favorecedor da marcação, embora tenuemente (pr. 0,539). Vogais e consoantes descem já para o degrau da inibição, com as primeiras (pr. 0,491) propiciando mais a marcação do que as segundas. As consoantes, por seu lado, ajudam mais à inserção PL quando a obstrução à passagem do ar através da cavidade bucal é também maior. Assim sendo, quando ocorre um fechamento total da boca para interromper momentaneamente a passagem do ar, isto é, quando se produzem as obstruintes oclusivas em *contexto fonológico posterior*, o propiciamento à marcação PL, apesar de ligeiramente inibidor (pr. 0,457), não deixa de ser superior ao das consoantes produzidas apenas com compressão da corrente de ar no tracto bucal, ou seja, ao das obstruintes constrictivas (pr. 0,436). Por outro lado, quando a corrente de ar na região da boca não é interceptada nem comprimida, produzindo-se consoantes soantes líquidas, a tendência é para aumentar a inibição, mas sem cair ainda em valores comprometedores (pr. 0,403).

Geracionalmente, o peso do modo de articulação das consoantes em *contexto fonológico seguinte* não se mostrou preponderante para a inserção da pluralização no item antecedente. Como tal, as faixas etárias mais idosa e mais nova foram rejeitadas pela ferramenta VARBRUL, apenas nos restando os valores da FE-2 para serem comparados com os da comunidade. O aspecto mais saliente centra-se nas vogais, que

perdem a sua ligeira tendência de favorecerem a marcação na FE-2 (pr. 0,507), para se tornarem levemente inibidoras da inserção PL na comunidade (pr. 0,491). Quanto aos restantes valores, a hierarquia do favorecimento mantém-se: (i) pausa final: FE-2 = pr. 0,703; comunidade = pr. 0,674; (ii) pausa interna: FE-2 = pr. 0,551; comunidade = pr. 0,539. Por outro lado, ocorre uma insignificante inversão hierárquica nos valores da inibição: (i) obstruintes constritivas: FE-2 = pr. 0,427; comunidade = 0,436; (ii) obstruintes oclusivas: FE-2 = pr. 0,421; comunidade = 0,457). As soantes líquidas apresentam um reduzido número de ocorrências na FE-2, inviabilizando comparações com o valor da comunidade.

**Tabela 25.7.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo o modo de articulação), na marcação PL dos itens do SN: 6 factores.

Factores	Comunidade			FE-3 (+60 anos)		
	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
	Pausa final	134/294	46	0,674	17/61	28
Pausa interna	106/349	31	0,539	19/114	17	-
Vogais	234/409	57	0,491	52/85	61	-
Obstruintes oclusivas	576/996	58	0,457	97/242	40	-
Obstruintes constritivas	118/231	51	0,436	20/56	36	-
Soantes líquidas	32/54	59	0,403	(2/7)	29	-

Factores	FE-2 (41 – 60 Anos)			FE-1 (20 – 40 Anos)		
	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
	Pausa final	49/112	44	0,703	68/121	56
Pausa interna	40/123	33	0,551	47/112	42	-
Vogais	59/120	49	0,507	123/204	60	-
Obstruintes oclusivas	181/326	56	0,421	298/428	69	-
Obstruintes constritivas	31/78	40	0,427	67/97	69	-
Soantes líquidas	(14/25)	56	-	(16/22)	73	-

Obs.: FE-1 (20-40 anos) e FE-3 (+60 anos) rejeitadas pela ferramenta VARBRUL.

### **5.2.3.7. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo o modo de *articulação*: 7 factores**

Observando os pesos da comunidade apresentados na Tabela 25.7, se compararmos o efeito das vogais com o das consoantes soantes líquidas na marcação PL, o comportamento das últimas não deixa de constituir, de certa forma, um paradoxo. De facto, do ponto de vista acústico, estas consoantes têm a particularidade de se aproximarem dos sons vocálicos, de tal modo que, em certas línguas, se tornam o centro de sílaba. Assim sendo, a tendência de marcação deveria ser de aproximação à das vogais e não de afastamento em relação às mesmas. Face a este comportamento, partimos para a análise da influência que as subcategorias das consoantes, em termos de modo de articulação, poderão ter sobre a inserção ou inibição das marcas de PL no SN do PA. Para tanto, foram constituídos os seguintes factores: (i) factor das obstruintes oclusivas orais [p], [b], [t], [d], [k] e [g]; (ii) factor das obstruintes oclusivas nasais [m] e [n]; (iii) factor das constrictivas fricativas [f], [v], [s], [z] e [ʒ]; (iv) factor da líquida lateral [l]; (v) e factor da líquida vibrante [r]. Como este último factor apenas regista 17 ocorrências (Tabela 12.4), foi por nós retirado das rodadas, cujos resultados podem ser lidos na Tabela 25.8. Para as rodadas que visaram obter valores por estrato etário, foram também retirados os itens com menos de 30 realizações.

Basicamente, os resultados desta rodada permitem perceber que, na comunidade de Almorixe, é a líquida lateral [l] a condicionar fortemente a marcação PL (0,292), apesar de não podermos observar o comportamento geracional da mesma em termos de pesos relativos, devido ao reduzido número de realizações nas diferentes faixas etárias. Por outro lado, apesar de também não possuímos valores para a líquida vibrante [r], é possível inferir que será esta a diluir a inibição da marcação, como o comprova o peso relativo fornecido pelo factor em que se amalgamaram estas duas vogais (soantes líquidas = pr. 0,403 – Tabela 25.7). Pormenores sobre a influência destas duas consoantes na marcação PL do SN do PA serão observados no ponto 5.2.3.9 do presente trabalho.

Quanto à pausa final, continua a revelar-se como o elemento que mais favorece a marcação (pr. 0,676) e, próximas do patamar da neutralidade, estão a pausa interna (pr. 0,542), as vogais (pr. 0,492) e todas as outras consoantes, com as oclusivas orais a

propiciarem um pouco mais a marcação (pr. 0467) e as oclusivas nasais a favorecerem menos esta (pr. 0407). Entre ambas, situam-se as constritivas fricativas (pr. 0,437), de certa forma contrariando a tendência apontada anteriormente, uma vez que as nasais são caracterizadas por uma interrupção total e momentânea da corrente respiratória, obrigando o som a ressoar na cavidade nasal, enquanto as constritivas apenas comprimem a passagem de ar.

**Tabela 25.8.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo o modo de articulação), na marcação PL dos itens do SN: 7 factores.

Factores	<i>Input desta rodada: 0,470</i> <i>Log-likelihood: -962,283</i> <i>Significância: 0,019</i>			<i>Input desta rodada: 0,204</i> <i>Log-likelihood: -185,006</i> <i>Significância: 0,007</i>		
	Comunidade			FE-3 (+60 anos)		
	Nr. Total	%	Peso Relativo	Nr. Total	%	Peso Relativo
Pausa final	134/294	46	0,676	17/61	28	-
Pausa interna	106/349	31	0,542	19/114	17	-
Vogais	234/409	57	0,492	52/85	61	-
Obstruintes oclusivas orais	502/844	59	0,467	82/204	40	-
Obstruintes oclusivas nasais	74/152	47	0,407	15/41	37	-
Obstruintes constritivas fricativas	118/231	51	0,437	20/56	36	-
Soante líquida lateral	19/38	50	0,292	(0/5)	0	-

Factores	<i>Input desta rodada: 0,355</i> <i>Log-likelihood: -354,549</i> <i>Significância: 0,007</i>			<i>Input desta rodada: 0,680</i> <i>Log-likelihood: -405,747</i> <i>Significância: 0,044</i>		
	FE-2 (41 – 60 Anos)			FE-1 (20 – 40 Anos)		
	Nr. Total	%	Peso Relativo	Nr. Total	%	Peso Relativo
Pausa final	49/112	44	0,704	68/121	56	0,573
Pausa interna	40/123	33	0,552	47/112	42	0,494
Vogais	59/120	49	0,507	123/204	60	0,440
Obstruintes oclusivas orais	164/281	58	0,416	257/359	71	0,554
Obstruintes oclusivas nasais	17/43	40	0,451	42/68	62	0,317
Obstruintes constritivas fricativas	31/78	40	0,427	67/97	69	0,478
Soante líquida lateral	(8/17)	47	-	(11/16)	69	-

Obs.: FE-3 (+60 anos) rejeitada pela ferramenta VARBRUL.

A propósito das nasais, refira-se ainda que Braga (1977:40) estudou também a sua influência na marcação PL, apresentando um peso relativo (pr. 0,41) para o PMn que se aproxima bastante daquele que é revelado pelo PA. Scherre (1988:254) observou também a dicotomia orais/nasais, apresentando um valor ligeiramente favorecedor da marcação para as primeiras (pr. 0,54) e, para as segundas, uma tendência suavemente inibidora (pr. 0,46). A autora justifica esta inibição referindo que a mesma se fica a dever à “*incidência bastante perceptível da aspiração e queda do –s diante da nasal bilabial principalmente*” (Scherre, 1988:255). Sobre as bilabiais nos pronunciaremos mais adiante. Contudo, comparando os valores apresentados por Braga (1977) e Scherre (1988) com os nossos, no que concerne ao modo de articulação das consoantes em *contexto fonológico posterior*, é lícito concordar com a segunda, quando esta afirma que a caixa de ressonância exerce fraca influência na inserção PL (Scherre, 1988:255).

Geracionalmente, as subcategorias das consoantes, em termos de modo de articulação, não se revelam importantes para a marcação PL na FE-3. Não obstante, já foi possível obter valores para a faixa etária mais jovem, que se analisarão posteriormente.

No que concerne à FE-2, o padrão hierárquico dos três elementos mais propiciadores da marcação é idêntico ao da comunidade, embora o peso das vogais seja ainda levemente favorecedor da marcação nesta faixa etária: (i) pausa final: FE-2 = pr. 0,704; comunidade = pr. 0,676; (ii). pausa interna: FE-2 = pr. 0,552; comunidade = pr. 0,542; (iii) vogais: FE-2 = pr. 0,507; comunidade = pr. 0,492. Relativamente às consoantes obstruintes, ligeiramente inibidoras da marcação em ambos os factores, apresentam alguma flutuação entre oclusivas orais e oclusivas nasais. As oclusivas orais, o terceiro factor mais propiciador da inserção PL entre as consoantes na FE-2 (pr. 0,416), ganham peso e passam a consoantes mais propiciadoras da pluralização na comunidade (pr. 0,467). Paralelamente, as nasais perdem importância, passando de obstruintes mais propiciadoras da marcação na FE-2 (pr. 0,451) para obstruintes menos propiciadoras da inserção PL na comunidade (pr. 0,407). As constrictivas fricativas, por seu lado, revelam-se estáveis, surgindo sempre como as segundas obstruintes a propiciarem mais a marcação, seja na FE-2 (pr. 0,427) seja na comunidade (pr. 0,437). Note-se ainda que a líquida lateral [l], devido à dispersão das suas ocorrências pelas três faixas etárias, não

apresenta realizações que deixem aferir o seu comportamento ao longo das diferentes gerações, como referimos no início da análise dos valores da Tabela 25.8.

Na FE-1, apenas dois factores se mostram favoráveis à marcação PL: a pausa final (pr. 0,573) e as consoantes oclusivas orais (pr. 0,554), que se tornam o segundo elemento mais favorecedor da inserção e levam a que, na comunidade (pr. 0,467), o seu propiciamento seja superior ao da FE-2 (pr. 0,416). Nos dados de Scherre (1988:254), esta tendência é muito menos pronunciada, uma vez que os adultos já marcam suavemente a pluralização, com as orais em *contexto fonológico seguinte* (pr. 0,54), pouco se acentuando nesta geração e também na das crianças (pr. 0,57). A pausa interna, por seu lado, apesar de revelar nos nossos dados uma ligeira queda no favorecimento, que começa por se apresentar na FE-2 (pr. 0,552), mantém-se a roçar a neutralidade na FE-1 (pr. 0,494), impedindo, assim, que o seu valor caia no patamar do desfavorecimento na comunidade (pr. 0,542). Simultaneamente, o resvalar acentuado das vogais de valores ligeiramente favorecedores na FE-2 (pr. 0,507) para resultados algo desfavorecedores na FE-1 (pr. 0,440), vai fazer com que este factor se torne um pouco inibidor da marcação na comunidade (pr. 0,492). Note-se também que o desfavorecimento apresentado pelas nasais na geração mais jovem (FE-2 = pr. 0,451; FE-1 = pr. 0,317) acaba por colocar estas consoantes num degrau a revelar inibição na inserção de marcas PL no SN do PA (pr. 0,407). Neste aspecto, a tendência para acentuar o não-favorecimento da marcação PL é quase imperceptível nos dados de Scherre (1988:254), já que os adultos e crianças revelam uma ligeira inibição, que se aproxima em termos de pesos relativos (adultos = pr. 0,46; crianças = pr. 0,43). Uma última palavra para as constrictivas fricativas, sem dúvida o item que menos flutuação apresenta diacronicamente (FE-2 = pr. 0,427; FE-1 = pr. 0,478), já que o mesmo mantém uma posição constante de desfavorecimento ligeiro à marcação PL ao longo das gerações, acabando este por se reflectir em termos de comunidade (pr. 0,437).

Os *log-likelihoods* achados para a comunidade e que constam nas Tabelas 25.6, 25.7 e 25.8, após efectuados os cálculos do qui-quadrado, mostram que os factores constantes nas referidas tabelas são estatisticamente significativos. Como tal, não podem ser incorporados no suporte computacional para obtenção de valores gerais, mantendo-se como válido, para este efeito, o grupo de factores apresentado na Tabela 25.1.

### **5.2.3.8. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo as zonas de articulação**

Braga (1977:71-72), Guy (1981a:166) e Scherre (1988:254) levaram também a cabo análises detalhadas em termos de traços seguintes no que concerne à zona e ponto de articulação das consoantes, e que interessam também para o nosso estudo. A fim de continuarmos a perceber como se processa a influência da consoante seguinte na marcação PL e compararmos os nossos achados com os resultados das análises destes três autores, constituímos um grupo de factores com cinco factores para as consoantes, de acordo com as suas zonas de articulação: (i) factor das labiais [p], [b] e [m]; (ii) factor das dentais [t], [d], [f], [v] e [n]; (iii) factor das alveolares [s], [z] e [l]; (iv) factor das palatais [k], [g] e [ʃ]; (v) e factor da uvular [R]. Como o último grupo de factores (a uvular [R]) apenas regista 17 ocorrências, foi retirado da próxima rodada. Os resultados podem ser observados na Tabela 25.9, mostrando que todas as faixas etárias passaram a ser consideradas relevantes pela ferramenta VARBRUL. Deste modo, reforça-se a possibilidade de ancorarem nos traços distintivos de cavidade, mais propriamente nas zonas e pontos de articulação, as causas para a inserção da marca PL no item que antecede as consoantes.

Relativamente à pausa final, verifica-se que é o elemento em contexto seguinte que mais favorece a marcação PL nas duas gerações mais idosas (FE-3 = pr. 0,709; FE-2 = pr. 0,707), apenas perdendo alguma preponderância na faixa etária mais nova (FE-1 = pr. 0,553), na qual ocupa o segundo lugar na hierarquia dos elementos que favorecem positivamente a inserção de número. Ainda assim, este aspecto não impede que a pausa final se revele como o factor que mais marca a pluralidade na comunidade de Almojarife (pr. 0,675). Contudo, as pausas deixam de ser agora os dois únicos elementos a inserirem positivamente a marca de número, uma vez que o factor consoantes palatais vai aumentando, diacronicamente, a sua influência sobre a marcação PL, passando de item ligeiramente inibidor na FE-3 (pr. 0,464) a elemento neutro, mas com alicive para a marcação, na FE-2 (pr. 0,501), até se tornar aquele que mais favorece a inserção da pluralização na geração mais nova (FE-1 = pr. 0,702). Em resultado de tal, o factor das consoantes palatais surge na comunidade como o segundo elemento que mais favorece a marcação PL (pr. 0,560), logo depois da pausa final (pr. 0,707).



**Tabela 25.9.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo as zonas de articulação), na marcação PL dos itens do SN.

	<i>Input desta rodada: 0,471</i> <i>Log-likelihood: -956,627</i> <i>Significância: 0,009</i>			<i>Input desta rodada: 0,204</i> <i>Log-likelihood: -176,733</i> <i>Significância: 0,007</i>		
	<b>Comunidade</b>			<b>FE-3 (+60 anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	134/294	46	0,675	17/61	28	0,709
Pausa interna	106/349	31	0,539	19/114	17	0,449
Vogais	234/409	57	0,490	52/85	61	0,668
Labiais	201/317	63	0,454	35/70	50	0,556
Dentais	201/436	46	0,401	38/119	32	0,439
Alveolares	73/148	49	0,327	6/29	21	0,099
Palatais	238/364	65	0,560	38/89	43	0,464
	<i>Input desta rodada: 0,354</i> <i>Log-likelihood: -352,334</i> <i>Significância: 0,018</i>			<i>Input desta rodada: 0,681</i> <i>Log-likelihood: -397,087</i> <i>Significância: 0,048</i>		
	<b>FE-2 (41 – 60 Anos)</b>			<b>FE-1 (20 – 40 Anos)</b>		
<b>Factores</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
Pausa final	49/112	44	0,707	68/121	56	0,553
Pausa interna	40/123	33	0,556	47/112	42	0,477
Vogais	59/120	49	0,510	123/204	60	0,428
Labiais	66/111	59	0,430	100/136	74	0,542
Dentais	65/138	47	0,408	98/179	54	0,336
Alveolares	21/56	38	0,290	46/63	73	0,529
Palatais	68/114	60	0,501	132/161	82	0,702

Os nossos dados mostram ainda que o desenvolvimento geracional da influência da pausa interna na marcação PL é bastante irregular. Assim, este elemento começa por inibir ligeiramente a pluralização na FE-3 (pr. 0,449), para a favorecer depois (FE-2 = pr. 0,556). Não obstante, a FE-1 vem mostrar uma regularização do item para a tendência inibidora (pr. 0,477), embora sem cair nos níveis da geração mais idosa. Em consequência de tal, a pausa interna acaba por se estabilizar como item suavemente favorecedor da marcação PL na comunidade (pr. 0,539). Comportamento identicamente irregular é revelado pelas consoantes labiais, mas na inversa. De facto, este factor inicia-se na FE-3 como levemente favorecedor da marcação PL (pr. 0,556), passando a

desfavorecedor na geração seguinte (FE-2 = pr. 0,430). Seguidamente, a FE-1 recupera a sua tendência para o suave favorecimento da marcação PL (pr. 0,542), traduzindo-se o resultado final, em termos de comunidade, numa ligeira inibição à inserção de PL (pr. 0,454).

Acompanhando a inibição das labiais na comunidade, temos ainda quer as vogais quer as restantes consoantes, isto é, as dentais e as alveolares. As dentais revelam-se inibidoras da marcação, acentuando-se esta geracionalmente. Como tal, apresentam-se já a marcar negativamente a inserção PL desde a geração mais idosa (FE-3 = pr. 0,439), numa tendência que se acentua nas duas faixas etárias seguintes (FE-2 = pr. 0,408; FE-1 = pr. 0,336). O resultado final, em termos de comunidade, só se poderia traduzir, pois, em desfavorecimento da pluralização (pr. 0,401). Por seu lado, as alveolares apresentam um comportamento oposto, isto é, tendente a tornar-se cada vez mais propiciador da marcação. Como se pode observar, este factor é fortemente inibidor da pluralização na FE-3 (pr. 0,099), mas essa inibição perde preponderância na FE-2 (pr. 0,290), desaparecendo mesmo na FE-1, que apresenta valores a favorecerem ligeiramente a marcação PL (pr. 0,529). Contudo, a tendência para os falantes mais jovens da comunidade marcarem a pluralidade, quando ocorre uma alveolar em *contexto fonológico seguinte*, é diluída pelas outras faixas etárias que compõem a comunidade, traduzindo-se o resultado final numa inibição considerável da pluralização (pr. 0,327), sobretudo provocada, ao que tudo indica, pela consoante líquida [l] (pr. 0,292 – Tabela 25.8).

Por fim, falta fazer alusão à conduta das vogais. Como se pode confirmar, a sua tendência é para acentuar, diacronicamente, a inibição à pluralização. De facto, este factor revela-se algo favorecedor da marcação na geração mais idosa (FE-3 = pr. 0,668), mas o favorecimento decai na FE-2 para valores já próximos da neutralidade (pr. 0,510), consumando-se a tendência para o desfavorecimento à inserção PL na FE-1 (pr. 0,428). Assim sendo, as vogais em *contexto fonológico seguinte* passam a inibir ligeiramente a pluralização em termos de comunidade (pr. 0,490), mas em valores muito próximos do patamar da neutralidade.

Quando comparados com os resultados de Braga (1977:71-72), os nossos números revelam algum paralelo com o PMn, no que às consoantes dentais diz respeito (pr. PA =

0,401; pr. PMn = 0,42), mas que vai sumindo quando se olha para os valores das labiais, mais propiciadoras da marcação nos nossos dados (pr. PA = 0,454; pr. PMn = 0,36). Por fim, o paralelo referido desaparece mesmo, face à discrepância constatada nas palatais, que favorecem levemente a marcação nas nossas ocorrências (pr. 0,560), mas condicionam fortemente esta no PMn (pr. 0,11). A própria autora adiantaria uma explicação para o não-favorecimento da concordância sintagmática PL originada pelas palatais nos seus dados: a semelhança destas com os segmentos que marcam o PL conduzem a processos de assimilação regressiva, isto é, à modificação de sons por influência dos sons vizinhos posteriores que com eles passam a partilhar traços articulatorios (as palatais em contexto seguinte). Contudo, convém aqui referir um aspecto pertinente, a merecer alguma reflexão: os contextos fonéticos (i.e. as palavras concretas onde ocorrem as mudanças fonéticas) mais propícios à assimilação são os nasais, os anteriores e os intervocálicos. Assim sendo, importaria apurar se eram exactamente estes o tipo de elementos que antecediam as palatais do *corpus* do PMn.

Guy (1981a:166) e Scherre (1988:254) apenas analisaram a influência das zonas e pontos de articulação das consoantes em itens que formam o PL de forma regular. Como tal, remetemos as nossas comparações com os resultados destas variáveis para a análise das Tabelas 25.17 e 25.18.

#### **5.2.3.9. Efeito do *contexto fonológico posterior* na marcação PL dos itens do SN, englobando consoantes segundo os pontos de articulação**

A última etapa das análises sobre a influência das consoantes em *contexto fonológico seguinte* na marcação PL no SN do PA passou pela observação do comportamento das mesmas, segundo os seus pontos de articulação (cf. Tabela 25.10). Assim, depois de efectuada esta rodada pôde entender-se em pormenor como actuam na marcação PL os subfactores que estão inseridos nos factores consonantais da Tabela 25.9. A constituição do grupo de factores *contexto fonológico posterior* segundo os pontos de articulação consonantais permitiu elaborar os seguintes factores para as consoantes: factor das ápico-dentais [t], [d] e [n]; factor da ápico-alveolares [s], [z] e [l]; factor da ápico-palatal [ʃ]; factor das bilabiais [p], [b] e [m]; factor das lábio-dentais [f] e [v]; e factor das dorso-velares (ou dorso-palatais) [k] e [g]. O factor da dorso-uvular [R] não foi incluído nesta rodada, visto possuir apenas 17 realizações. Igual procedimento foi adoptado em

relação à ápico-palatal [j], detentora de apenas 4 ocorrências no nosso *corpus*. Para as rodadas que visaram estabelecer os valores por estrato etário, foram também retirados os itens com menos de 30 realizações.

Se nos abstrairmos dos pesos apresentados pela pausa final, pausa interna e vogais, que continuam a manter, diacronicamente, valores idênticos aos registados anteriormente (comparem-se as tabelas 25.9 e 25.10), e nos concentrarmos nos números apresentados pelos factores das consoantes, é possível aquilatar que, na comunidade de Almozarife, a maior responsabilidade pelo valor minimamente favorecedor da marcação registado pelas palatais (Tabela 25.9 = pr. 0,560) cabe às dorso-velares [k] e [g] (pr. 0,576), especialmente à primeira (Tabela 20.1), e a menor responsabilidade à ápico-palatal [ʃ] (pr. 0,444). O ligeiro favorecimento à marcação levado a cabo pelas primeiras na comunidade fica a dever-se, como se pode confirmar, à tendência diacrónica em aumentar o peso de favorecimento à inserção da pluralização. Assim as dorso-velares surgem como factor inibidor da marcação na FE-3 (pr. 0,437), mas essa inibição vai-se diluindo com o tempo, passando pelo degrau da neutralidade na geração intermédia (FE-2 = pr. 0,501) e quedando-se no patamar do favorecimento evidente na geração mais nova (FE-1 = pr. 0,723). Dado que a região palatal favorece a inserção PL, a ligeira inibição registada pela ápico-palatal [ʃ] fica, então, a dever-se ao ponto apical, que, como se pode verificar, vai condicionar bastante a marcação PL, conforme o demonstram as ápico-dentais (pr. 0,369) e as ápico-alveolares (pr. 0,329). A dispersão de realizações pelas faixas etárias não nos permite estabelecer uma radiografia geracional para a ápico-palatal [ʃ], mas, no que concerne às ápico-dentais, é possível observar que as mesmas se apresentam sempre como elementos que inibem a marcação PL, e numa escala de desfavorecimento cada vez mais acentuado (FE-3 = pr. 0,437; FE-2 = pr. 0,351; FE-1 = pr. 0,342). Já as ápico-alveolares revelam um tendência inversa, uma vez que iniciam o seu percurso como itens bastante desfavorecedores da marcação (FE-3 = pr. 0,103), desfavorecimento este que se vai diluindo (FE-2 = pr. 0,293), até se tornar ligeiramente favorecedor da inserção PL na geração menos idosa (FE-1 = pr. 0,528). Contudo, esta ligeira tendência para o favorecimento não é suficiente para contrariar a tendência registada das gerações mais idosas, pelo que, em termos de comunidade, o peso acaba por resultar em desfavorecimento considerável da inserção PL (pr. 0,329).

O facto de o ponto apical se apresentar como inibidor da inserção PL ajuda a perceber, agora, o porquê dos condicionamentos à marcação levados a cabo quer pela consoante /s/, o segundo elemento mais inibidor (Tabela 25.1 = pr. 0,334), quer pela consoante /l/, o item que menos propicia a marcação (Tabela 25.1 = pr. 0,279), apesar de este último se aproximar das vogais em termos de modo de articulação. Braga (1977:72) explicou o condicionamento produzido pela consoante /s/ em termos de assimilação regressiva, aspecto que discutiremos adiante, mais propriamente aquando da análise dos resultados da Tabela 25.11.

Quanto às ápico-dentais apresentam-se ligeiramente menos inibidoras do que as ápico-alveolares porque a zona dental surge como favorecedora à inserção PL, como se pode notar pelo peso das lábio-dentais (pr. 0,534), o segundo factor mais propiciador da pluralização no item antecedente. Não obstante, não é possível estudar a evolução das lábio-dentais, uma vez que estas consoantes apenas apresentam valores para a FE-1, por sinal a registarem um peso moderadamente inibidor da inserção PL (pr. 0,384). Ainda assim, o peso apresentado pela comunidade ajuda a perceber o motivo pelo qual a consoante lábio-dental /f/ se assume como o terceiro elemento consonantal que mais ajuda à pluralização (Tabela 25.1 = pr. 0,562). Paralelamente, pode ainda constatar-se que o peso das lábio-dentais nos ajudam a inferir, igualmente, que os pontos labiais propiciam mais a marcação do que os apicais, contribuindo para que a bilabial /b/ apresente um valor ligeiramente favorecedor da marcação PL (Tabela 25.1 = pr. 0,562). Convém referir, apesar de tudo, que as bilabiais exibem uma conduta geracional um tanto difusa, já que emergem como elementos em *contexto fonológico seguinte* levemente favorecedores da inserção PL (FE-3 = pr. 0,553), para passarem depois a itens desfavorecedores desta (FE-2 = pr. 0,436). Todavia, na geração mais jovem retomam a sua posição de factor que ajuda à inserção da pluralidade (FE-1 = pr. 0,558), fazendo, então, com que o valor da comunidade se aproxime do ponto da neutralidade, ainda que ligeiramente inibidor (pr. 0,455).

**Tabela 25.10.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os pontos de articulação), na marcação PL dos itens do SN.

<i>Input desta rodada: 0,472</i> <i>Log-likelihood: -954,225</i> <i>Significância: 0,009</i>							<i>Input desta rodada: 0,186</i> <i>Log-likelihood: -177,148</i> <i>Significância: 0,044</i>						
Comunidade				FE-3 (+60 anos)									
Factores	Nr. total	%	Peso relativo	Nr./total	%	Peso relativo							
Pausa final	134/294	46	0,674	17/61	28	0,698							
Pausa interna	106/349	31	0,538	19/114	17	0,468							
Vogais	234/409	57	0,489	52/85	61	0,658							
Consoantes ápico-dentais	153/353	43	0,369	27/96	28	0,437							
Consoantes ápico-alveolares	73/148	49	0,329	6/29	21	0,103							
Ápico-palatais	16/38	42	0,461	(3/9)	33	-							
Consoantes bilabiais	201/317	63	0,455	35/70	50	0,553							
Consoantes lábio-dentais	48/83	58	0,534	(11/22)	50	-							
Consoantes dorso-velares	222/326	68	0,576	35/76	46	0,437							
<i>Input desta rodada: 0,355</i> <i>Log-likelihood: -351,261</i> <i>Significância: 0,012</i>							<i>Input desta rodada: 0,724</i> <i>Log-likelihood: -403,496</i> <i>Significância: 0,032</i>						
FE-2 (41 – 60 Anos)				FE-1 (20 – 40 Anos)									
Factores	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo							
Pausa final	49/112	44	0,705	68/121	56	0,541							
Pausa interna	40/123	33	0,556	47/112	42	0,463							
Vogais	59/120	49	0,524	123/204	60	0,417							
Consoantes ápico-dentais	48/111	43	0,351	78/146	53	0,342							
Consoantes ápico-alveolares	21/56	38	0,293	46/63	73	0,528							
Ápico-palatais	(2/13)	15	-	(11/16)	69	-							
Consoantes bilabiais	66/111	59	0,436	100/135	74	0,558							
Consoantes lábio-dentais	(16/27)	59	-	21/34	62	0,384							
Consoantes dorso-velares	67/110	61	0,501	120/140	86	0,723							

Após o estudo das Tabelas 25.9 e 25.10, percebeu-se exactamente quais são as zonas de articulação que mais propiciam a marcação, como se verificou também que os valores presentes na primeira destas tabelas estarão condicionados pela correlação de pesos entre zona e pontos de articulação. Assim sendo, e em termos de resultado final, poderemos estabelecer a seguinte ordem hierárquica decrescente de favorecimento à marcação PL,

em termos de *contexto fonológico seguinte*, no que às consoantes diz respeito: (i) palatais, com especial acento para as velares e declive para as apicais; (ii) labiais; (iii) dentais, com acento para as lábio-dentais e declive para as ápico-dentais; (iv) alveolares. Ainda no que concerne às velares, também Braga (1977:72) as indicou como sendo as mais favorecedoras da marcação nas realizações do PMn (pr. 0,65). Contudo, uma incongruência fica por esclarecer nos dados da autora: estando o ponto de articulação velar inserido na zona de articulação palatal, como podem as velares surgir como os elementos que mais favorecem a marcação, por oposição às palatais, os elementos que ela indicou como sendo os mais condicionadores da inserção PL?

Refira-se, por fim, que os grupos de factores das Tabelas 25.9 e 25.10 também apresentam *log-likelihoods* para a comunidade que não permitem que as referidas amalgamações sejam levadas à análise geral, visto que as mesmas se revelam estatisticamente significativas. Como tal, é o grupo de factores constituído inicialmente (Tabela 25.1) que deve ser incorporado no suporte computacional para efeitos de observação do peso que a variável independente *contexto fonológico seguinte* exerce na marcação PL dos itens do SN do PA.

#### **5.2.3.10. Efeito do *contexto fonológico posterior* na marcação PL da *saliência fónica***

Scherre (1988:252) entendeu que seria pertinente observar o efeito do *contexto fonológico seguinte* em nomes morfologicamente distintos, uma vez que os regulares fazem o PL com inserção do *-s* morfémico, ao passo que os irregulares o constroem com acréscimo de *-es* no final. Os resultados da autora apontaram para a possibilidade de as consoantes em *contexto fonológico seguinte* favorecerem a inserção PL nos itens terminados em *-s* (pr. 0,74) e *-r* (pr. 0,69). Contrariamente, as vogais desfavorecem bastante a marcação (pr. *-s* = 0,26; pr. *-r* = 0,30). A fim de compararmos este tipo de comportamento entre o MRJ e o PA, efectuámos diversos cruzamentos, levando em consideração as variáveis independentes *contexto fonológico posterior* e *saliência fónica*.

### 5.2.3.10.1. Efeito do *contexto fonológico posterior* na *saliência fónica*, englobando consoantes segundo os traços de fonte, para a marcação PL dos itens do SN

Para os cruzamentos entre os grupos de factores *contexto fonológico posterior* e *saliência fónica*, começámos por constituir um factor para as consoantes terminadas em *-s* e outro para as que finalizam em *-r*. Porém, todas as rodadas que realizámos revelaram exiguidade de ocorrências quer para um quer para o outro factor, inviabilizando, no nosso trabalho, a possibilidade de se comparar, com outros dialectos, o perfil da marcação PL nos itens terminados em *-s* ou *-r* no singular. Face a este panorama, e dado que ambos os itens formam o PL com inserção de *-es* final, decidimos amalgamá-los em um único factor, tendo então sido possível obter resultados para alguns factores dos grupos de factores constituídos para efeitos de cruzamentos em que intervém a variável *itens singulares em -r e -s*. Refira-se ainda que, em virtude da dispersão de dados pelas diferentes faixas etárias tornar também inviável a análise para a grande maioria dos factores, apenas se efectuaram rodadas gerais, a fim de se obterem resultados mais sólidos para efeitos de leitura.

**Tabela 25.11.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços de fonte), na marcação PL da *saliência fónica*.

<i>Input desta rodada:</i> 0,459		<i>Log-likelihood:</i> -991,159			<i>Significância:</i> 0,006		
<i>Contexto fonológico posterior</i>		PL's regulares			PL's nasais (regulares)		
X		Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
<i>Saliência fónica</i>							
	Pausa final	98/240	41	0,629	(2/7)	29	-
	Pausa interna	75/286	26	0,468	(5/20)	25	-
	Vogais	191/399	56	0,459	(9/20)	45	-
	Consoantes surdas	412/657	63	0,449	32/52	62	0,592
	Consoantes sonoras	219/438	50	0,392	19/45	42	0,415

<i>Contexto fonológico posterior</i>		Singulares <i>-r e -s</i>		
X		Nr. total	%	Peso relativo
<i>Saliência fónica</i>				
	Pausa final	27/35	77	0,958
	Pausa interna	(18/27)	67	-
	Vogais	24/36	67	0,888
	Consoantes surdas	17/35	49	0,827
	Consoantes sonoras	18/33	55	0,832



A rodada em que cruzámos os grupos de factores *saliência fónica* e *contexto fonológico posterior*, levando em conta as consoantes segundo os seus traços de fonte, mostra que, no caso dos plurais regulares, a pausa final mantém a sua homogeneidade na inserção da pluralização, favorecendo-a moderadamente (0,629). Contudo, a pausa interna, embora se situe como o segundo elemento que mais propicia a marcação, perde alguma força e desce para o patamar da ligeira inibição, apesar de os seus valores se revelarem praticamente neutros (pr. 0,468). Abaixo da pausa interna, mas ainda no patamar da neutralidade, surgem as vogais (pr. 0,459) e consoantes surdas (pr. 0, 449). Quanto às consoantes sonoras, têm já algum peso na inibição PL dos itens regulares (pr. 0,392). Ainda no que diz respeito aos itens regulares, os valores do MRJ (Scherre, 1988:249) apresentam-se neutros quer para os elementos consonantais (pr. 0,44) quer para os vocálicos (pr. 0,50). Relativamente à pausa, o valor registado revela que a mesma ajuda a inserir tenuemente a pluralização (pr. 0,56).

Lopes (2001:231-232), por seu lado, apresentou resultados um pouco distintos dos nossos para os itens regulares. De facto, nos nossos dados apenas um factor ajuda a marcar positivamente o PL (a pausa final), mas, nas ocorrências de Salvador, somente um item marca negativamente a pluralização: a consoante sonora (pr. 0,44). Como tal, na NURC, todos os outros factores apresentam pesos positivos para a inserção do PL, de acordo com a seguinte escala hierárquica de favorecimento: pausa final = pr. 0,60; consoante surda = pr. 0,53; vogal = pr. 0,52; e pausa interna = pr. 0,50. Como se pode ver, a panorâmica geral dos itens regulares da NURC apresenta valores mais próximos do ponto da neutralização do que os do PA, cujos factores exibem pesos mais distintivos entre favorecimento e inibição à inserção do PL.

Voltando aos nossos dados, mais concretamente aos plurais nasais regulares, vemos que estes mostraram apenas resultados para as consoantes, uma vez que os outros factores apresentam realizações reduzidas. Assim, é apenas possível confirmar que os elementos consonantais surdos favorecem ligeiramente a marcação PL (pr. 0,592), enquanto os sonoros a desfavorecem exiguamente (pr. 0,415).

Relativamente aos itens terminados em *-s*, o MRJ (Scherre, 1988:253) exhibe valores que confirmam que o *contexto fonológico seguinte* tem tendência a propiciar mais a marcação quando é consonantal (pr. 0,58), do que quando é vocálico (pr. 0,14). Nos

elementos que finalizam em *-r*, dá-se o oposto, uma vez que são as vogais a propiciarem mais a inserção PL (pr. 0,42) do que as consoantes (pr. 0,36), se bem que ambas se situem no degrau do não favorecimento. Por seu lado, a pausa é o elemento que favorece mais acentuadamente a marcação (pr. terminados em *-s* = 0,81; pr. terminados em *-r* = 0,71).

Lopes (2001:231-232), por sua vez, apenas nos traz resultados para os elementos terminados em *-s*. Nestes, as consoantes apresentam uma alternância de hierarquia no propiciamento da inserção PL, quando comparadas com as dos itens regulares. De facto, agora são as surdas que emergem como elementos que mais desfavorecem a marcação (pr. 0,44), enquanto as sonoras se situam no ponto da neutralidade absoluta (pr. 0,50). Como segundo elemento que menos favorece a inserção PL aparecem as vogais (pr. 0,46), ao contrário das pausas, as quais ajudam a inserir a pluralização (pausa final = pr. 0,80; pausa interna = pr. 0,58).

Convém também referir que Lopes (2001:232-240) efectuou o estudo do modo como o *contexto fonológico seguinte* actua na pluralização do elemento anterior registado nas realizações produzidas por quatro grupos de falantes distintos, reunidos segundo duas categorias: informantes que utilizam o português popular *versus* informantes que usam o português universitário; e informantes com sobrenome religioso *versus* informantes sem sobrenome religioso. Como os informantes que produziram o nosso *corpus* não se enquadram nestas categorias, pelo menos no que respeita às três últimas, não nos é possível exibir resultados que possibilitem comparações com as mesmas.

A exiguidade dos nossos dados não permitiu que se estabelecesse a oposição entre itens singulares terminados em *-s* e itens singulares terminados em *-r*. Ainda assim, é possível constatar que a amalgamação dos dois factores evidencia que estes elementos propiciam fortemente a marcação, a qual é sempre bastante acentuada. A pausa final apresenta valores praticamente absolutos de marcação (pr. 0,958), logo seguida das vogais (pr. 0,888). As consoantes surdas e sonoras apresentam resultados de favorecimento bastante nivelados, pelo que a diferença entre ambas não deve ser considerada significativa (pr. consoantes surdas = 0,827; pr. consoantes sonoras = 0,832). O reduzido número de ocorrências em que intervém a pausa interna, apenas 27,

não permitiu que incorporássemos estes dados à análise. Contudo, é possível adiantar, para estes, um percentual de marcação positiva, situado nos 67%.

Nos nossos dados, uma evidência é inquestionável: o *contexto fonológico posterior*, segundo os traços de fonte, propicia muito mais a marcação PL nos itens cujos singulares terminam em *-s* e *-r*, do que nos regulares. O facto de não acontecer a busca do padrão silábico CV é explicado por Scherre (1988:251-252) e confirmado por Lopes (2001:239) em termos de haplologia sintáctica, isto é, eliminação de segmentos idênticos ou iguais dentro da palavra ou em fronteira de palavras, traduzida em inibição produzida pelas vogais e tendência para a marcação levada a cabo pelas consoantes. O fenómeno acontece especialmente em realizações em que intervêm nomes do tipo *mês/meses*, em que a inserção da marca PL em itens aos quais se segue uma vogal iria provocar uma ocorrência de duas sílabas contíguas muito semelhantes (Scherre, 1988:244). Lopes (2001:231), que detectou iguais ocorrências nos seus dados, refere-se ao fenómeno ilustrando-o com o exemplo “*Umas vezes ele vem*”, em que a repetição do som /ze/ provoca a ocorrência *vezezele*. Assim sendo, e com vista a evitar a reduplicação de sons, os falantes deixam de efectuar a concordância, como acontece em “*Algumas vez ele vem*”. Nos nossos dados, as ocorrências com os substantivos iguais ao de Lopes não foram consideradas, visto entendermos que estes configuram itens lexicalizados, isto é, elementos cuja estrutura PL não é gerada por qualquer dos processos sintácticos que originam os plurais comuns da língua portuguesa (ponto 3.1.4 do presente trabalho, dedicado à caracterização dos nossos dados). Contudo, o nome apontado por Scherre (*mês/meses*) é frequentemente utilizado pelos informantes do PA, que levam também a cabo a não concordância ou apenas a concordância parcial, supostamente com o fim de evitar repetições fonológicas do tipo das que foram apontadas para o MRJ (exemplos [271] e [272]):

[271] PA:  
*três MESE ESSA porta tá fechado* [MANEVEM2]

[272] PA:  
*ficô aqui em casa nove MESE E tal, moreu.* [MMDEUSM3]

Contudo, se observarmos outros exemplos com o mesmo nome, constata-se que ocorrem também inibições da marcação PL, mas que não são determinadas por *contexto fonológico posterior* vocálico:

[273] PA:  
*três MESE SÓ, já dá milho.* [OSVALH1]

[274] PA:  
*mais de quato... cinco MES SEM com... sem tocá* [LUISH2]

Nestes exemplos, temos ocorrência consonantal seguinte a operar “metaplasmente” na terminação PL do elemento antecedente. Por outro lado, é preciso não esquecer que a haplologia actua ao nível de sílabas da mesma estrutura e contíguas, em que a menos saliente é suprimida. Portanto, estaremos aqui em presença das assimilações regressivas preconizadas por Braga (1977:72) e originadas pela presença de dois fonemas ápicos e contíguos, que compartilham traços de semelhança: o [s] áptico-alveolar em contexto seguinte e o [j] áptico-palatal, que marca o PL no item anterior.

Contudo, idêntico tipo de metaplasmos vai actuar sobre os mesmos itens, quando o *contexto fonológico posterior* é distinto dos anteriormente apontados. Os exemplos que listamos seguidamente apresentam algumas dessas ocorrências:

[275] PA:  
*fiquei cerca de quatro MESE DOENTE.* [ZECAH1]

[276] PA:  
*quatro MÊS VAI picá dedo, tem paludismo* [LUISH2]

[277] PA:  
*Creche rápido, só três MESE.* [CLOTIM2]

[278] PA:  
*Ficou seis MÊS.* [FRANCM2]

O metaplasmo que ocorre na realização do exemplo [275] é idêntico ao do exemplo [273], embora as consoantes em *contexto fonológico seguinte* sejam distintas, inviabilizando-se assim a possibilidade de uma assimilação no exemplo [275]. Por seu lado, a queda da pluralização que acontece no exemplo [276] é estruturalmente semelhante à do exemplo [274], apesar de os contextos consonantais posteriores serem

desiguais. Mais uma vez, o fenómeno verificado no exemplo [276], não pode ser explicitado em termos de assimilação. Aliás, análogos metaplasmos voltam a ser detectados nos exemplos [277] e [278], mas agora com contextos fonológicos posteriores traduzidos por pausa final. Como tal, uma vez mais, nem a assimilação postulada por Braga (1977:72) nem a haplologia apontada por Scherre (1988:251-252) e Lopes (2001:239) se coadunam às justificações que pretendam explicar os metaplasmos dos dois últimos exemplos.

Por outro lado, os tipos de metaplasmos a que nos vimos referindo ocorrem também com os itens singulares em *-r*, e que se pluralizam de modo semelhante aos singulares terminados em *-s*, isto é, com a inserção final de *-es*.

- [279] PA:  
*tem quatro MULHER EM numa casa* [ANTOM1]
- [280] PA:  
*Dois filho MULHER QU'Ê tem.* [FRANCM2]
- [281] PA:  
*é por isso as MULHER FICA tando dá assi muita coesa* [CELESH3]
- [282] PA:  
*nós stamos três MULHE, um rapaz, um home* [MMDEUSM3]
- [283] PA:  
*É! Um homé, dois MULHER.* [FRANCM2]

Como se pode observar, o exemplo [279] fornece-nos uma ocorrência com inibição total de pluralização antes de um contexto vocálico, mas que não pode ser explicada em termos de haplologia sintáctica, visto não se estar em presença de reduplicação de sons. Como também não podem ser justificados em termos de assimilação regressiva os metaplasmos que acontecem antes dos diferentes contextos fonológicos consonantais apresentados nos exemplos [280] e [281]. Aspectos estes que acabam reforçados pelos contextos seguintes de pausas representados em [282] e [283], para os quais não se encontra qualquer explicação na perspectiva quer da assimilação regressiva quer da haplologia sintáctica. Contudo, dos resultados exibidos na Tabela 25.11 e dos vários exemplos apresentados retiramos duas conclusões:

- (i) Os itens singulares terminados em *-s* e *-r* tendem fortemente a ser marcados, qualquer que seja a ocorrência *fonológica posterior*;
- (ii) Os vários metaplasmos em questão têm um padrão definido e semelhante, apesar de os *contextos fonológicos posteriores* que, supostamente, os motivaram serem bastante heterogêneos.

Resta então uma pergunta: se os itens referidos tendem fortemente a ser pluralizados, o que motiva realmente os escassos metaplasmos que se traduzem em marcação parcial ou inibição total da marcação?

Estudando a influência da *saliência fônica* na inserção da marca PL no SN do português da área urbana do Rio de Janeiro, Braga & Scherre (1976:465-477) estabeleceram uma escala de diferenciação material fônica em cinco níveis, que viria a ser ampliada por Guy (1981a:181-199) para sete níveis. A constituição de tais escalas foi justificada com o facto de a concordância de número entre os elementos do SN ser condicionada pelo grau de *saliência fônica* que ocorre na oposição singular/plural. Assim, duas dessas escalas incluíam, precisamente, os itens singulares terminados em *-s* e em *-r*, que fazem o PL com inserção final de *-es*. A conclusão do primeiro estudo foi que as formas salientes favorecem mais as marcas de PL do que as formas menos salientes. Os baixos valores probabilísticos de inserção PL nos itens terminados em *-s* foram justificados, no estudo das autoras, com o facto de estes itens apresentarem, no final, uma consoante semelhante à do PL. Contudo, outros trabalhos posteriores, como sucedeu com o de Ponte (1979), que estudou dados de falantes semi-analfabetos de Porto Alegre, Rio Grande do Sul, não apresentaram idênticos resultados. De facto, Ponte (1979:164-166) concluiu que o factor que menos favorecia a marcação PL era o dos nomes regulares, aspecto que viria a ser corroborado por Guy (1981a:189). Este último viria ainda a indicar que os nomes terminados em *-s* e *-r* favoreciam a presença de PL (pr. 0,64). Scherre (1988:252) voltaria a referir que os itens em *-s* desfavorecem a presença PL (pr. 0,13) no MRJ, apresentando, por outro lado, valores favorecedores à inserção para os nomes em *-r* (pr. 0,70). Contudo, os nossos resultados praticamente confirmam os pressupostos de Ponte (1979) e Guy (1981a), já que os itens em *-s* favorecem acentuadamente a marcação (pr. 0,862), enquanto os elementos em *-r* (pr.

0,463) apresentam um peso de inserção de marcas praticamente igual ao dos itens regulares (pr. 0,491), com os valores de ambos situando-se próximos do patamar da neutralização PL (Tabela 23.1). Paralelamente, se compararmos os elementos em *-s* e *-r* entre si, constatamos que os primeiros são os que mais propiciam a marcação PL, contrariando a tendência observada no MRJ.

Convém ainda referir que, tal como referimos a propósito do estudo da variável independente *saliência fónica* (ponto 5.2.1 do presente trabalho), Naro & Lemle (1976:240-241) e Lemle & Naro (1977:47) haviam apresentado o Princípio da Saliência para justificar que as formas mais salientes, logo mais perceptíveis, detêm maior probabilidade de serem marcadas do que as formas menos salientes. Assim sendo, não surpreende que, nos nossos dados, os itens em *-s* e *-r* apresentem maior peso na inserção PL do que os itens regulares. Contudo, interessa aqui perceber porque é que, em algumas das realizações que envolvem o uso destes itens, vai ocorrer a inibição da pluralização, se bem que de forma escassa. E neste aspecto, tanto as Condições de Distintividade (Kiparsky, 1972:195), em termos de hipótese funcionalista, como o pressuposto de Guy (1981a:3001-302) de que a causa para a ocorrência da CPL-var no PB deverá ancorar nos substratos africanos, podem ajudar a trazer alguma luz à interrogação que subsiste. Assim, e nesta perspectiva, ter-se-á em conta quer o facto de a informação semanticamente relevante ter tendência para ser retida na estrutura de superfície (Kiparsky, 1972:195) quer a possibilidade de a marcação PL ocorrer na primeira oportunidade, por norma antes do nome núcleo do SN, evitando-se depois a sua repetição, por se tornar redundante.

As constatações avançadas por Braga (1977), Scherre (1988) e Lopes (2001) apontam todas para a influência de fenómenos do tipo superstratal no PB, ainda que Scherre (1988:257) admita que, no caso de a variação de número de marcas formais de PL no SN do MRJ “*estiver refletindo um processo de mudança linguística, certamente não será um tipo de mudança preconizada pelos neogramáticos, a qual se supõe altamente regular e não condicionada por fatores outros que não os fonéticos*”. Contudo, as explicações das autoras para justificarem a inibição das marcas de PL em termos de *contexto fonológico posterior* revelaram-se bastante incongruentes, motivando a nossa reflexão acerca das mesmas. No ponto 1.1 aludimos à evidência das influências

fonológicas dos substratos na L2, pelo que interessa verificar, então, até que ponto ocorrem interferências específicas do sistema fonológico dos substratos directos e ancestrais do PA neste, a ponto de afectarem os morfemas em final de palavra e, conseqüentemente, o seu sistema de marcação PL.

Ferraz (1979), para o santomense, Mingas (2000), para o PLd, Figueiredo (2003) e Inverno (2005), para o PVA, Gonçalves (1997) para o PMp e Azevedo (2003) para o PM, elencam uma série de interferências fonológicas em que ocorrem, na L2, transferências ou substituições de fonemas da LA por outros das L1's. No caso específico do santomense, Ferraz (1979:21-24) chama a atenção não só para o facto de o seu sistema consonantal ser marcadamente diferente do português mas também para a realidade de diversas ocorrências da fonologia do primeiro serem mais o espelho de realizações encontradas nos substratos africanos do que no PE. As consoantes líquidas vibrantes, por exemplo, não existem no santomense, estando sujeitas a fenómenos de síncope (exemplo [284]) e apócope (exemplo [285]), ou, em alternativa à sua troca pela líquida lateral /l/, isto é, ao lambdacismo (exemplo [286]).

[284]	FORRO:	PTG:		
	[dlete ]	[dƏRƏter]	>	derreter (Ferraz, 1979:22)
	↓			
	Ø			

[285]	FORRO:	PTG:		
	[vεdε]	[vƏrdadƏ]	>	verdade (Ferraz, 1979:22)
	↓			
	Ø			

[286]	FORRO:	PTG:		
	[fa <u>l</u> uza]	[fƏRuʒɛ]	>	ferrugem (Ferraz, 1979:23)

Estamos, evidentemente, perante o reflexo de ocorrências que acontecem na generalidade das línguas africanas atlântico-ocidentais, também desprovidas de vibrantes. Bentley (1967 [1887]), em trabalho levado a cabo em finais do séc. XIX, listou alguns empréstimos portugueses incorporados pelo kishikongo, no qual o fenómeno do lambdacismo é também evidente:





Apesar de tudo, tais exceções não encobrem, de forma alguma, a realidade de determinadas ocorrências fonológicas representarem herança das realizações que acontecem nas L1's, nada tendo a ver com fenómenos registados na LA. Refira-se ainda que, no PA, são igualmente bastante comuns os fenómenos de lambdacismo, sobretudo nas gerações mais idosas (exemplos [211] e [291]), configurando-se aqui casos de fossilização temporária (ponto 2.10.4.1 do presente trabalho):

[291] *Ê nō vō lepará mais não* [MINISM3]  
 “Eu não vou reparar mais, não”

Outro dos fenómenos característicos do santomense que assenta nos substratos, mais propriamente nos idiomas do grupo níger-congo atlântico, como o kwa, é o da harmonização vocálica em duas sílabas consecutivas,<sup>188</sup> observado por Ferraz (1979:46) sob dois prismas: o fonémico e o morfofonémico. Segundo o autor, imensos exemplos do forro evidenciam que esta realização não só harmoniza as vogais átonas de acordo com o núcleo da sílaba tónica do português mas detém também preferência sobre os outros processos de assimilação em que há não-realização de fonemas, a ponto de, morfofoneticamente, acontecerem casos de inserção vocálica, como a epêntese (exemplo [292]) ou a paragoge (exemplo [293]), para proporcionar a ocorrência da harmonização:

[292]	FORRO:	PTG:			
	[kele]	[kɾer]	>	crer	(Ferraz, 1979:45)
	⊕				

[293]	FORRO:	PTG:			
	[liʔi]	[nari]	>	nariz	(Ferraz, 1979:45)
	⊕				

Dado que os falantes em ASL não recorrem à simplificação de formas, aplicam estratégias que visam retomar a estrutura da L1 (Tarone, 1978, 1995), recorrendo a epênteses ou paragoges para desmontarem encontros consonantais, um fenómeno que, ao que tudo indica, será regido por princípios universais no sentido de aplicar a estrutura silábica CV (Tarone, 1978:328). Desta forma, o recurso à harmonização vocálica possuirá também primazia sobre os fenómenos de síncope, como a haplologia, em que ocorre a queda, isto é, a não-realização da primeira de duas sílabas seguidas iniciadas por um mesmo fonema consonantal. E mais curiosamente: das três vogais portuguesas,

/a/, /e/ e /u/, aquela que mais frequentemente surge afectada pelo fenómeno da harmonização é precisamente a segunda, isto é, a vogal /e/, apontada nos exemplos de Scherre (1988:244) e Lopes (2001:231) para justificarem a haplologia. Como se pode ver ainda, o fenómeno de harmonização, enquanto inabilidade dos falantes em ASL para pronunciarem correctamente os sons da LA (Ellis & Beaton, 1995:164), conduziu a uma fossilização fonológica (Tarone, 1978, 1995; Levelt, 1989) que se mantém em diacronia e activa no PA (exemplos [294] e [295]), em virtude de os parâmetros desactivados das matrizes fonéticas não terem sido reactivados posteriormente (cf. Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003):

[294] PA:  
*dois **senhoro** e uma responsável de Centro.* [CLOTIM2]  
“dois senhores e uma responsável do Centro”.

[295] *vá pa mar pegá **vadoro*** [MANOH3]  
“vai para o mar pegar (apanhar) voadores”

Olhando para o exemplo “vezezele”, apontado por Lopes (2001:231) para justificar a haplologia, se o transportássemos para o santomense, enquanto substrato do PA, veríamos que estariam reunidas todas as condições para se proceder à harmonização vocálica do /e/ pertencente à sílaba que antecede a sílaba tónica, e nunca à síncope da sílaba na qual o mesmo está inserido. Mas a verdade é que os nossos dados revelam que, por vezes, há queda de fonemas nos casos apresentados. Continua por esclarecer, então, o que motiva tais fenómenos de supressão.

Configurando a epêntese uma estratégia para ajudar os falantes em ASL a pronunciarem os sons difíceis (Tarone, 1978, 1995), resulta num fenómeno comum na aquisição fonológica das L2's. Assim, a inserção epentética faz com que o fenómeno se configure como potenciador da fossilização fonológica (Tarone, 1978, 1995). Mas, a nível morfofonémico, as epênteses ou paragoges têm ainda outra finalidade: possibilitar a construção da estrutura silábica padrão da L1 (Tarone, 1978:328), ou seja, a estrutura dissilábica CVCV, no caso do santomense. Como se verifica, este tipo de estrutura silábica representa uma simplificação em comparação com as do português, embora reflecta igualmente traços da estrutura deste. Não obstante, é também óbvio, segundo



irreversível: (i) síncope do [r], enquanto som inexistente nas línguas africanas atlântico-ocidentais e conseqüente inabilidade psicológica de produzir uma nova fonologia (Tarone, 1978, 1995; Ellis & Beaton, 1995), que leva à interferência do sistema fonológico da L1 na L2; (ii) apócope do *-s* final, marca de PL, por transferência do sistema de pluralização das ancestrais L1's africanas (Guy, 1981a:301-302), que apenas marcam o número no classificador pré-nominal, permanecendo os radicais das palavras inalterados; (iii) estratégia fonética para articulação de encontros fonológicos mais complexos, guiada pela tendência/preferência universal pela sílaba estruturada na forma CV (Tarone, 1972:328); iv) e transferência da estrutura dissilábica padrão da L1, tendo em conta que a L2 é construída a partir do sistema da L1 previamente adquirida (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006). Por fim, a fossilização mantém-se em diacronia, em virtude de ocorrer transmissão linguística irregular e não se terem refixado, na L2, os parâmetros desactivados das matrizes fonéticas (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) nem se terem erradicado determinados parâmetros presentes nas ancestrais L1's e no PtgL1 adquirido defectivamente (Hale, 1988:32; Gonçalves, 2004:235).

Em alguns dos exemplos apresentados em que acontece o apagamento parcial da forma *-es* PL, ou seja, os números [271], [272], [273], [275] e [282], verifica-se que ocorre a construção silábica ao encontro da estrutura dissilábica padrão CVCV não só do santomense mas também das línguas binji e kwa. Assim, poderá residir aqui uma das explicações para o fenómeno que apaga apenas o /s/ final, principalmente em realizações em que o *contexto fonológico seguinte* é representado por uma pausa, seja ela interna seja final. Porém, continuámos sem uma resposta acerca das motivações que levam aos outros metaplasmos, fazendo com que os monossílabos do tipo de “mês” e os dissílabos de estrutura silábica CVCVC, do tipo de “mulher”, permaneçam inalterados.

Em resumo: para se perceber quais os contextos fonológicos seguintes que motivariam os apagamentos parciais ou totais do *-es* final em itens singulares terminados em *-s* e *-r*, verificámos que quer a haplologia sintáctica, postulada por Scherre (1988:244) e corroborada por Lopes (2001:239) ou Baxter (2004:112), quer a assimilação regressiva, advogada por Braga (1977:72) e secundada por Scherre

(1988:255) ou Baxter (2004:112), não concedem resposta plenamente justificativa para as síncopes em questão. Face ao facto de tais itens enquadrarem realizações cuja *saliência fónica* é bastante acentuada, logo com tendência a serem mais marcadas, tentámos buscar nos substratos as motivações que subjazem às inibições em questão, uma vez que têm sido comprovadas as suas interferências fonémicas e morfofonémicas em L2's que emergem do contacto entre línguas tipologicamente distintas, mesmo em palavras incorporadas da LA. O fenómeno do lambdacismo e a estrutura dissilábica padrão CVCV do santomense e línguas de substrato africanas forneceram-nos uma possibilidade de resposta para se justificarem algumas das ocorrências analisadas, principalmente quando acontecem marcações parciais que originam itens do tipo de “mese”. Contudo, como prevalece o impasse sobre o que poderá estar a motivar os restantes casos de inibição, decidimos dirigir a nossa atenção para outras possibilidades, isto é, outras variáveis que poderão estar a correlacionar forças com os fenómenos analisados, gerando as síncopes apontadas. Como tal, abstraímos-nos do *contexto fonológico seguinte* e passámos a observar as marcas que precedem os itens afectados pelo fenómeno do apagamento PL, total ou parcial, a fim de verificarmos a possibilidade de a motivação para a inibição PL ancorar nestas. Em mente, mantivemos também a Condição de Distintividade advogada por Kiparsky (1972:195), que defende que a informação semanticamente relevante tem tendência a ser retida na estrutura de superfície. Estes e outros aspectos serão discutidos em etapa posterior do nosso trabalho, mais propriamente no ponto 5.2.3.11, que observará a influência das *marcas precedentes* na marcação PL dos itens terminados em *-s* e *-r* no singular.

#### **5.2.3.10.2. Efeito do *contexto fonológico posterior* na *saliência fónica*, englobando consoantes segundo os traços de classe principal, para a marcação PL dos itens do SN**

Continuando a análise das questões relacionadas com o grupo de factores *contexto fonológico posterior*, efectuámos mais rodadas VARBRUL, a fim de se poder observar os possíveis efeitos desta variável independente na marcação PL dos itens morfológicamente distintos. O peso que os traços de classe principal das consoantes exercem sobre os elementos que pluralizam regularmente ou com inserção de *-es* final constitui, assim, o passo seguinte do nosso estudo. Para obtenção dos resultados das

rodadas que visaram obter os valores por estrato etário, foram retirados os itens com menos de 30 realizações. Os resultados destas rodadas VARBRUL podem ser vistos na Tabela 25.12.

**Tabela 25.12.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços de classe principal), na marcação PL da *saliência fónica*.

<i>Input desta rodada:</i> 0,454		<i>Log-likelihood:</i> -994,532			<i>Significância:</i> 0,006		
<i>Contexto fonológico posterior</i>		PL's regulares			PL's nasais (regulares)		
X		Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
<i>Saliência fónica</i>							
Pausa final		98/240	41	0,632	(2/7)	29	-
Pausa interna		75/286	26	0,492	(5/20)	25	-
Vogais		191/399	56	0,471	(9/20)	45	-
Consoantes soantes		93/178	52	0,400	(5/11)	46	-
Consoantes não-soantes		538/917	59	0,432	46/86	54	0,527

<i>Contexto fonológico posterior</i>		Singulares <i>-r e -s</i>		
X		Nr. total	%	Peso relativo
<i>Saliência fónica</i>				
Pausa final		27/35	77	0,958
Pausa interna		(18/27)	68	-
Vogais		24/36	67	0,888
Consoantes soantes		(8/16)	50	-
Consoantes não-soantes		27/52	52	0,837

Os itens regulares continuam a ter na pausa final o seu mais forte factor de favorecimento à inserção de PL (pr. 0,632). Por seu lado, todos os outros factores apresentam tendências inibidoras, apesar de roçarem o plano da neutralidade. Ao analisar-se a dicotomia itens com inserção final de PL *-s versus* itens com inserção final de PL *-es*, dois aspectos são de reter em linha de conta: (i) a hierarquia do favorecimento e/ou inibição de marca PL não se altera; (ii) a pausa interna, apesar de ser o segundo elemento mais propiciador da inserção PL, deixa de possuir um peso favorecedor, para se tornar mais neutra e apresentar um valor que inibe muito ligeiramente a marcação PL (pr. 0,492). Como tal, todos os outros factores são também inibidores da marcação, embora em valores próximos da neutralidade. As vogais (pr. 0,471) continuam a propiciar mais a marcação do que as consoantes e, entre estas, as soantes apresentam-se como as que mais inibem a marcação (pr. 0,400). Por seu lado, as

consoantes não-soantes são o segundo elemento mais inibidor da pluralização nos itens regulares (pr. 0,432).

À primeira vista, poderia pressupor-se que a consoante /r/, por ser uma soante, favorece mais a inibição PL do que o /s/, uma consoante não-soante. Todavia, verificámos já que uma outra soante, o /l/, é a que mais inibe a inserção de número. Assim sendo, o valor do factor consoantes soantes pode estar a ser condicionado pelas realizações dos itens que terminam nesta consoante. Infelizmente, não detemos também um número de ocorrências consistente de consoantes soantes em *contexto fonológico seguinte* para verificar como as mesmas actuam na inserção do PL *-es*. Não obstante, é possível constatar que, também nos itens singulares terminados em *-s* e *-r*, continuam a ser as consoantes que menos propiciam a marcação PL, neste caso as não-soantes (pr. 0,837). Seguidamente, o factor que mais favorece a inserção da pluralização é o das vogais (pr. 0,888), antecedendo o peso da pausa final, que se mostra como o elemento que mais categoricamente favorece à inserção (pr. 0,958). Por seu lado, a pausa interna não apresentou realizações que nos permitissem observar a sua hierarquia no que respeita ao proporcionamento da marcação PL. Pode observar-se, isso sim, que os contextos fonológicos posteriores continuam a aparentar um forte favorecimento da inserção PL em itens terminados em *-s* e *-r* no singular.

Relativamente aos itens que admitem a pluralização nasal regular, apenas se obtiveram resultados para as consoantes não-soantes. Estas actuam de modo praticamente neutro sobre a marcação PL, mas favorecendo ligeiramente a inserção da marca (pr. 0,527).

#### **5.2.3.10.3. Efeito do *contexto fonológico posterior na saliência fónica, englobando consoantes segundo os traços coronal e não-coronal, para a marcação PL dos itens do SN***

A observação do peso dos traços de cavidade das consoantes na pluralização dos itens regulares ou terminados em *-s* e *-r* foi seguidamente analisada por nós, conforme Tabela 25.13. Para obtenção dos resultados desta tabela foram retirados os itens que perfazem menos de 30 ocorrências.

A análise contrastiva entre itens que constroem o PL de modo regular e aqueles que o fazem com inserção de *-es* em final de palavra forneceu-nos valores que revelam uma



ligeira alteração, em termos de hierarquia dos factores que favorecem a pluralização. Assim, enquanto a pausa final (elemento mais favorecedor da marcação = pr. 0,630), as vogais (segundo elemento menos favorecedor da inserção PL = pr. 0,464) e as consoantes coronais (elemento que menos favorece a marcação = pr. 0,329) mantêm os seus lugares hierárquicos no que concerne ao propiciamento da marcação, ocorre uma alternância de posições entre a pausa interna e as consoantes não-coronais. Estas mantêm praticamente o valor ligeiramente favorável à pluralização já anteriormente verificado (rodada geral pr. 0,522 – Tabela 25.6; rodada contrastiva pr. 0,509 – Tabela 25.13), mas a pausa interna vê decrescer o peso propiciador da marcação (pr. 0,533 – Tabela 25.6) que a colocava no segundo lugar do favorecimento à inserção PL, transitando agora para um patamar ligeiramente inibidor e situado no terceiro posto na hierarquia da inserção PL (pr. 0,469 – Tabela 25.13).

**Tabela 25.13.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os traços coronal e não-coronal), na marcação PL da *saliência fónica*.

<i>Input desta rodada:</i> 0,458		<i>Log-likelihood:</i> -982,980			<i>Significância:</i> 0,004		
<i>Contexto fonológico posterior</i>		PL's regulares			PL's nasais (regulares)		
X		Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
<i>Saliência fónica</i>							
	Pausa final	98/240	41	0,630	(2/7)	29	-
	Pausa interna	75/286	26	0,469	(5/20)	25	-
	Vogais	191/399	56	0,464	(9/20)	45	-
	Consoantes não-coronais	438/655	67	0,509	26/44	59	0,564
	Consoantes coronais	193/440	44	0,329	25/53	47	0,455
<i>Contexto fonológico posterior</i>		Singulares <i>-r</i> e <i>-s</i>					
X		Nr. total	%	Peso relativo			
<i>Saliência fónica</i>							
	Pausa final	27/35	77	0,957			
	Pausa interna	(18/27)	67	-			
	Vogais	24/36	67	0,886			
	Consoantes não-coronais	(14/27)	52	-			
	Consoantes coronais	21/41	51	0,818			

Relativamente aos singulares terminados em *-s* e *-r*, continua a verificar-se que todos os factores são bastante favoráveis à inserção de pluralização. A pausa final conserva a sua liderança, traduzida num favorecimento categórico da marcação (pr. 0,957), logo

seguida das vogais (pr. 0,886) e das consoantes coronais (pr. 0,818). A pausa interna e as consoantes não-coronais registam ocorrências que não permitem fornecer resultados em termos de pesos (27 realizações), mas o valor percentual de marcação é superior para a primeira (18 itens marcados = 67%) e inferior para as segundas (14 itens marcados = 52%).

Os itens nasais regulares apenas forneceram resultados para as consoantes. Apesar de tudo, a tendência para estas propiciarem a marcação segue o padrão dos plurais regulares, com as consoantes não-coronais a favorecerem muito ligeiramente a inserção do PL (pr. 0,564) e as coronais a inibirem-na suavemente (pr. 0,455).

#### **5.2.3.10.4. Efeito do *contexto fonológico posterior na saliência fónica, englobando consoantes segundo o modo de articulação, para a marcação PL dos itens do SN: 6 factores***

A forma como o modo de articulação das consoantes influencia a marcação PL nos itens regulares e singulares acabados em *-s* e *-r* concedeu o mote para o próximo passo das nossas constatações. Uma primeira rodada VARBRUL foi efectuada, agrupando as consoantes em três factores, de acordo com a Tabela 25.14. Nesta rodada não foram processados os itens que perfazem menos de 30 realizações.

Os resultados mantiveram a hierarquia de propiciamento à marcação PL verificada na rodada geral (Tabela 25.7). Contudo, a pausa final revela-se agora como o único factor que favorece a marcação (pr. 0,634), acabando os restantes factores por inibi-la. Como tal, a pausa interna deixa de propiciar positivamente a marcação (pr. 0,539 – Tabela 25.7) para a desfavorecer, embora de modo quase neutro (pr. 0,474). O terceiro elemento mais propiciador da marcação continua a ser o factor das vogais (pr. 0,465), sobrepondo-se a todo o tipo de consoantes, segundo o seu modo de articulação. Entre estas confirma-se o seguinte paralelo: quanto mais forte é a obstrução à passagem da corrente de ar na cavidade bucal, menor é a inibição verificada. Assim, o peso consonantal de não favorecimento à pluralização começa por ser menor com as obstruintes oclusivas (pr. 0,440), para aumentar com as obstruintes constrictivas (pr. 0,403) e tornar-se já consideravelmente acentuado com as soantes líquidas (pr. 0,368).

**Tabela 25.14.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo o modo de articulação), na marcação PL da *saliência fónica*: 6 factores.

<i>Input desta rodada:</i> 0,454		<i>Log-likelihood:</i> -996,453			<i>Significância:</i> 0,006		
<i>Contexto fonológico posterior</i>		<b>PL's regulares</b>			<b>PL's nasais (regulares)</b>		
<b>X</b>		<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Saliência fónica</i>							
Pausa final		98/240	41	0,634	(2/7)	29	-
Pausa interna		75/286	26	0,474	(5/20)	25	-
Vogais		191/339	56	0,465	(9/20)	45	-
Obstruintes oclusivas		506/859	59	0,440	39/72	54	0,526
Obstruintes constrictivas		96/190	51	0,403	(10/20)	50	-
Soantes líquidas		29/46	63	0,368	(2/5)	40	-

<i>Contexto fonológico posterior</i>		<b>Singulares –r e –s</b>		
<b>X</b>		<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Saliência fónica</i>				
Pausa final		27/35	77	0,956
Pausa interna		(18/27)	40	-
Vogais		24/36	67	0,886
Obstruintes oclusivas		24/47	51	0,797
Obstruintes constrictivas		(10/18)	56	-
Soantes líquidas		(1/3)	33	-

O peso da inibição consonantal parece diminuir à medida que a *saliência fónica* se acentua. Assim, as obstruintes oclusivas tornam-se já um factor que favorece ligeiramente a inserção PL em itens nasais regulares (pr. 0,526), favorecimento esse que se acentua bastante nos elementos terminados no singular em *–s* e *–r* (pr. 0,797). Apesar de tudo, este aspecto não pode ser totalmente verificado, já que não se obtiveram valores para as restantes consoantes, quer sejam dos primeiros itens quer sejam dos segundos. Ainda relativamente aos elementos terminados no singular *–s* e *–r*, é possível apenas confirmar que todos os factores que concederam valores para a análise continuam a favorecer fortemente a inserção PL, com a pausa final, uma vez mais, a liderar tal favorecimento (pr. 0,956), logo secundada pelas vogais (0,886). Para a pausa interna também não se obtiveram resultados, visto possuímos apenas 27 realizações.

#### **6.2.3.10.5. Efeito do *contexto fonológico posterior* na *saliência fónica*, englobando consoantes segundo o modo de articulação, para a marcação PL dos itens do SN: 7 factores**

Após análise sobre a maneira como o modo de articulação das consoantes actua, de modo geral, na marcação PL dos itens regulares e singulares acabados em *-s* e *-r*, passámos a observar, de forma mais minuciosa, como esta se processa. Assim, constituíram-se quatro factores de acordo com as subcategorias consonantais e procedeu-se a mais uma rodada VARBRUL, cujos resultados se encontram expressos na Tabela 25.15. De acordo com a metodologia que temos vindo a aplicar, mais uma vez não foram processados os elementos que perfazem menos de 30 ocorrências.

Relativamente aos itens que pluralizam de modo regular, continua a verificar-se que a pausa interna se assume como o único elemento favorecedor da marcação PL (pr. 0,634). Em termos de propiciamento da marcação, mas já no patamar da ligeira inibição, seguem-se-lhe a pausa interna (pr. 0,474) e vogais (pr. 0,467), ambas com pesos superiores aos de todas as consoantes. Entre estas, mais uma vez se confirma que a obstrução da passagem do ar no tracto bucal é um elemento que leva a propiciar mais a marcação. Desta forma, oclusivas orais (pr. 0,447) e oclusivas nasais (pr. 0,416) vão apresentar um valor de inibição inferior ao das constrictivas fricativas (pr. 0,405). Lembre-se que a rodada geral, com 7 factores segundo o modo de articulação das consoantes (Tabela 25.8), nos havia fornecido valores que contrariavam ligeiramente esta tendência (oclusivas nasais = pr. 0,407; constrictivas fricativas = pr. 0,437). Agora, o padrão é novamente repostado para os itens cujos plurais são regulares, pelo que não residirá nestes a motivação que levou a tal flutuação. Infelizmente, como veremos já de seguida, não obtivemos valores para confirmar se esta se fica a dever aos elementos que pluralizam com a inserção final de *-es* ou aos itens nasais que formam o PL com acréscimo de *-s* final. Resta, apenas referir que, ainda no concernente aos itens de pluralização regular, se confirma que o elemento mais condicionador da inserção PL é, uma vez mais, a líquida lateral [l] (pr. 0,276).

Em termos de itens que pluralizam com inserção de *-es* final, mais uma vez parece confirmar-se que, qualquer que seja o seu *contexto fonológico posterior*, a tendência é para serem bastante marcados. Assim, a pausa final favorece de novo, e categoricamente, a inserção PL (0,956), sendo secundada pelas vogais (pr. 0,887) e pelas

consoantes oclusivas orais (0,801). Para as restantes consoantes e pausa interna não se obtiveram resultados, devido ao seu reduzido número de realizações.

**Tabela 25.15.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo o modo de articulação), na marcação PL da *saliência fónica*: 7 factores.

<i>Input desta rodada</i> : 0,451		<i>Log-likelihood</i> : -996,725			<i>Significância</i> : 0,006		
<i>Contexto fonológico posterior</i>		PL's regulares			PL's nasais (regulares)		
X		Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
<i>Saliência fónica</i>							
Pausa final		98/240	41	0,634	(2/7)	29	-
Pausa interna		75/286	26	0,474	(5/20)	25	-
Vogais		191/339	56	0,467	(9/20)	45	-
Obstruintes oclusivas orais		442/727	61	0,447	36/66	55	0,558
Obstruintes oclusivas nasais		64/132	49	0,416	(3/6)	50	-
Obstruintes constrictivas fricativas		96/190	51	0,405	(10/20)	50	-
Soante líquida lateral		18/32	56	0,276	(0/3)	0	-
<i>Contexto fonológico posterior</i>		Singulares <i>-r e -s</i>					
X		Nr. total	%	Peso relativo			
<i>Saliência fónica</i>							
Pausa final		27/35	77	0,956			
Pausa interna		(18/27)	67	-			
Vogais		24/36	67	0,887			
Obstruintes oclusivas orais		17/34	50	0,801			
Obstruintes oclusivas nasais		(7/13)	54	-			
Obstruintes constrictivas fricativas		(10/18)	56	-			
Soante líquida lateral		(1/3)	33	-			

Os itens nasais que pluralizam de modo regular apenas apresentaram valores para as oclusivas orais, sendo os mesmos levemente favorecedores da marcação (pr. 0,558). Uma vez mais, este aspecto levanta a possibilidade de haver tendência para se marcar mais a pluralização em itens de maior *saliência fónica*.

**5.2.3.10.6. Efeito do *contexto fonológico posterior* na *saliência fónica*, englobando consoantes segundo as zonas de articulação, para a marcação PL dos itens do SN**

A pormenorização das nossas análises levou-nos a observar, seguidamente, o modo como as zonas de articulação podem influenciar a marcação PL em itens morfológicamente distintos. A Tabela 25.16 fornece-nos uma visão dos factores constituídos, bem como a leitura dos resultados desta rodada VARBRUL, na qual não foram processados os itens com menos de 30 realizações:

**Tabela 25.16.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo as zonas de articulação), na marcação PL da *saliência fónica*.

<i>Input desta rodada: 0,445</i>		<i>Log-likelihood: -989,959</i>			<i>Significância: 0,006</i>		
<i>Contexto fonológico posterior</i>		<b>PL's regulares</b>			<b>PL's nasais (regulares)</b>		
<b>X</b>		<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>	<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>
<i>Saliência fónica</i>							
Pausa final		98/240	41	0,640	(2/7)	29	-
Pausa interna		75/286	26	0,480	(5/20)	25	-
Vogais		191/339	56	0,472	(9/20)	45	-
Labiais		183/283	65	0,438	(6/12)	50	-
Dentais		163/358	46	0,384	23/46	50	0,524
Alveolares		62/123	50	0,287	(5/13)	39	-
Palatais		212/317	70	0,569	(15/24)	63	-
<i>Contexto fonológico posterior</i>		<b>Singulares -r e -s</b>					
<b>X</b>		<b>Nr. total</b>	<b>%</b>	<b>Peso relativo</b>			
<i>Saliência fónica</i>							
Pausa final		27/35	77	0,957			
Pausa interna		(18/27)	67	-			
Vogais		24/36	67	0,886			
Labiais		(10/14)	71	-			
Dentais		(12/28)	43	-			
Alveolares		(5/10)	50	-			
Palatais		(8/16)	50	-			

Os valores obtidos mostram-nos que o padrão geral observado na Tabela 25.9 se repete com os itens de pluralização regular. Assim, apenas dois factores favorecem positivamente a marcação PL: a pausa final (pr. 0,640) e as consoantes palatais (0,569).

Ainda assim, e apesar de manter a sua posição hierárquica, isto é, surgir como o terceiro elemento mais propiciador da pluralização, a pausa interna deixa de favorecer esta de maneira positiva (pr. 0,539 – Tabela 25.9) e cai para valores ligeiramente inibidores (pr. 0,480). Quanto às vogais, apresentam resultados mais propiciadores à marcação do que todas as restantes consoantes (pr. 0,472). Estas, por sua vez, mantêm a escala hierárquica de inibição já verificada anteriormente, que se traduz, em termos crescentes, na seguinte ordem: consoantes labiais, levemente inibidoras (pr. 0,438); consoantes dentais, moderadamente inibidoras (pr. 0,384); e consoantes alveolares, acentuadamente inibidoras (pr. 0,287).

Guy (1981a:166) e Scherre (1988:254) também analisaram a influência das zonas de articulação das consoantes em *contexto fonológico seguinte* na marcação PL dos itens regulares. Para tanto, os autores constituíram um único factor para as dentais e alveolares. Nenhum deles forneceu valores para as palatais, tendo a segunda decidido não as analisar, devido ao fenómeno da neutralização que ocorre no português (Câmara Jr., 1975:83-84). Por conseguinte, não nos é possível estabelecer comparações com estas consoantes. Já o mesmo não acontece em relação às labiais, que nos dados dos autores apresentam um ligeira tendência para inibirem a marcação (Guy, 1981a = pr. 0,49; Scherre, 1988 = pr. 0,44), tal como acontece nas realizações do nosso *corpus* (pr. 0,438). Relativamente às dentais e alveolares, a tendência dos resultados de Guy (1981a) e Scherre (1988) é para a quase neutralização da marcação PL (0,48 e pr. 0,49, respectivamente), enquanto nas nossas ocorrências, quer umas quer outras, apresentam propensão para a inibição da pluralização (dentais = pr. 0,384; alveolares = pr. 0,287).

Infelizmente não se obtiveram quaisquer resultados para observar como é que as zonas de articulação das consoantes agem sobre os itens terminados em *-s* e *-r*, já que as realizações do PA se dispersaram bastante entre as diferentes zonas. Assim, apenas se conseguiram resultados para a pausa final (pr. 0,957) e vogais (pr. 0,886), com ambos os factores a favorecerem fortemente a marcação PL, uma vez mais.

Idêntico problema se pôs para os itens nasais que realizam plurais de modo regular, uma vez que se obtiveram valores unicamente para as consoantes dentais (pr. 0,524). Estas, ao contrário dos itens regulares, favorecem ligeiramente a marcação PL, apesar de se situarem no nível da neutralidade.

### 5.2.3.10.7. Efeito do *contexto fonológico posterior* na *saliência fónica*, englobando consoantes segundo os pontos de articulação, para a marcação PL dos itens do SN

O passo final das nossas análises teve como objectivo constatar a forma como os pontos de articulação interagem com itens morfológicamente distintos, no que diz respeito à influência da marcação PL destes, seguindo o mesmo critério de constituição de factores estabelecido no ponto 5.2.3.9. A Tabela 25.17 dá-nos um retrato quer dos factores constituídos quer dos valores obtidos com o tratamento dos dados.

**Tabela 25.17.** Efeito do *contexto fonológico posterior* (englobando consoantes segundo os pontos de articulação), na marcação plural da *saliência fónica*.

<i>Input desta rodada:</i> 0,444		<i>Log-likelihood:</i> -986,563			<i>Significância:</i> 0,008		
<i>Contexto fonológico posterior</i>		PL's Regulares			PL's Nasais (Regulares)		
X		Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
<i>Saliência fónica</i>							
Pausa final		98/240	41	0,638	(2/7)	29	-
Pausa interna		75/286	26	0,479	(5/20)	25	-
Vogais		191/339	56	0,472	(9/20)	64	-
Consoantes ápico-dentais		121/288	42	0,345	19/38	50	0,519
Consoantes ápico-alveolares		62/123	50	0,289	(5/13)	39	-
Ápico-palatais		10/29	35	0,435	(1/2)	50	-
Consoantes bilabiais		183/283	65	0,439	(6/12)	50	-
Consoantes lábio-dentais		42/70	60	0,538	(4/8)	50	-
Consoantes dorso-velares		202/288	70	0,589	(14/22)	64	-
<i>Contexto fonológico posterior</i>		Singulares –r e –s					
X		Nr. total	%	Peso relativo			
<i>Saliência fónica</i>							
Pausa final		27/35	77	0,957			
Pausa interna		(18/27)	67	-			
Vogais		24/36	67	0,884			
Consoantes ápico-dentais		(11/24)	46	-			
Consoantes ápico-alveolares		(5/10)	50	-			
Ápico-palatais		(5/7)	71	-			
Consoantes bilabiais		(10/14)	71	-			
Consoantes lábio-dentais		(1/4)	25	-			
Consoantes dorso-velares		(3/9)	33	-			



Para efeitos de obtenção de resultados que permitam comparações com os da Tabela 25.10, optámos por incluir nos plurais regulares o factor das ápico-palatais, ainda que estas registem apenas 29 realizações. Na tabela anterior é possível verificar que três factores favorecem a marcação PL dos itens regulares, de acordo com a seguinte ordem decrescente de influência: pausa final (pr. 0,638), consoantes dorso-velares (pr. 0,589) e consoantes lábio-dentais (0,538). Note-se que a pausa interna, que surgia como o terceiro elemento que mais favorecia a inserção nos dados gerais (pr. 0,538 – Tabela 25.10), desce agora para o patamar da ligeira inibição (pr. 0,479). Exceptuando as consoantes já mencionadas, todas as outras apresentam valores inibidores da marcação mais fortes do que as vogais, as quais exibem um peso moderadamente desfavorável à pluralização (pr. 0,472). Os pressupostos que avançámos acerca do modo como os pontos de articulação poderão influenciar a marcação PL (ponto 5.2.3.9) encontram aqui eco, já que apicais, uma vez mais, se assumem como os elementos mais inibidores da inserção (pr. ápico-dentais = 0,345; pr. ápico-alveolares = 0,289). Note-se que, de novo, o peso inibidor das apicais é atenuado pela zona palatal, levando a que as ápico-palatais (pr. 0,435) e as bilabiais (pr. 0,439) apresentem um teor de inibição muito aproximado, caminhando para o ponto da neutralidade.

Guy (1981a:166) e Scherre (1988:254) analisaram também o comportamento das velares em itens regulares, concluindo que as mesmas apresentam, tal como no PA (pr. 0,589), tendência para favorecerem a marcação em valores que se aproximam bastante dos obtidos por nós (pr. 0,54 e pr. 0,56, respectivamente).

Os resultados apresentados tanto para os itens nasais que formam o PL de forma regular como para os elementos que o produzem com inserção de *-es* final revelaram-se pouco esclarecedores, dada a escassez de ocorrências. Assim, para os primeiros apenas se obteve o peso relativo das consoantes ápico-dentais, que deixa de ser acentuadamente inibidor para passar a ligeiramente favorecedor da marcação PL (pr. 0,519). Este aspecto vem realçar, de novo, a tendência para os falantes marcarem os itens de maior *saliência fónica*, o que se confirma com os elementos que realizam a pluralização com inserção de *-es* final. Ainda assim, para estes somente se apuraram resultados relativamente à pausa final (pr. 0,957) e vogais (pr. 0,884), ambas favorecendo fortemente a marcação PL.

Das várias rodadas que efectuámos para perceber como o *contexto fonológico seguinte* pode influenciar a marcação PL em itens morfológicamente distintos, alguns aspectos se tornaram evidentes, se bem que não tivéssemos obtido dados para todos os factores envasados quer nos itens nasais que pluralizam regularmente quer nos elementos que constituem o PL com inserção de *-es* final. De qualquer forma, é possível inferir que a marcação se acentua nos itens de *saliência fónica* mais perceptível, tornando-se altamente elevada quando os elementos finalizam no singular em *-s* ou *-r*, independentemente do tipo de *contexto fonológico posterior*. Se este se traduzir em vogal, o peso de marcação é bastante significativo para os elementos que pluralizam com inserção de *-es* final, variando entre 0,884 e 0,888 (cf. Tabelas 25.12 a 25.17). Quanto a nós, estes valores reforçam a possibilidade de as motivações que conduzem aos apagamentos parciais ou totais do *-es* final em itens singulares terminados em *-s* e *-r* não ancorarem no *contexto fonológico seguinte* traduzido por vogal. Como tal, os pesos significativos de marcação retiram também força ao pressuposto de tais síncope assentarem no fenómeno da haplogogia sintáctica postulada por Scherre (1988:251-252).

Analisando agora o comportamento das consoantes em *contexto fonológico seguinte* no que respeita ao favorecimento ou não da marcação PL, e face ao pressuposto da possível assimilação regressiva defendida por Braga (1977:72), é possível verificar que os valores obtidos nas referidas tabelas, quando viáveis, apontam também para um favorecimento significativo, embora inferior ao das vogais. Incidindo a nossa atenção nas realizações que possam conduzir à assimilação, constata-se que o /s/ em final de palavra apenas se pode realizar como coronal não-soante, seja ele o segmento [s] (fricativa, alveolar, surda), o segmento [ʃ] (fricativa, palatal, surda), o segmento [z] (fricativa, alveolar, sonora), ou o segmento [ʒ] (fricativa, palatal, sonora). Se bem que para os modos, zonas e pontos de articulação não se tenham obtido valores relativamente a este tipo de consoantes, é possível verificar que as coronais em contexto posterior favorecem acentuadamente a marcação PL nos itens antecedentes terminados em *-s* ou *-r* (pr. 0,818 – Tabela 25.13), como também o favorecem categoricamente as consoantes não-soantes (pr. 0,837 – Tabela 25.12). Relativamente aos traços de fonte (Tabela 25.11), quer as surdas (pr. 0,827) quer as sonoras (pr. 0,832) surgem também como elementos bastante favorecedores da inserção PL nos elementos que pluralizam com

acrescento de *-es* no seu final. Desta forma, parece reforçada a nossa intuição de que também não estaremos perante casos de assimilação quando acontecem as sínopes totais ou parciais que afectam a terminação PL *-es* dos itens singulares em *-s* e *-r*.

Com a análise dos resultados das variáveis independentes do tipo fonológico procurámos entender pormenorizadamente as motivações que poderão levar à inibição PL registada por itens terminados em *-s* e *-r* no singular, tendo essa procura incidido sobre as ligações que os mesmos estabelecem com o *contexto fonológico seguinte*. Como se verificou também, este, por si só, não concede resposta satisfatória para a totalidade dos apagamentos detectados. Assim sendo, optámos por confirmar se tais motivações terão a ver com outros factores, que não os observados. Nesta conformidade, dirigimos a nossa atenção para a variável independente *marcas precedentes*, encarando como princípio orientativo a possibilidade de tais inibições reflectirem apenas a funcionalidade da língua apontada por Kiparsky (1972:195), pressupondo-se que o falante retém unicamente a informação relevante, descartando aquilo que é dispensável. Portanto, o que poderá estar a acontecer é, tão-somente, o fenómeno que ancora nos substratos africanos de se marcar o PL na primeira oportunidade (cf. Guy, 1981a:131-132), ocorrendo depois o apagamento formal da pluralidade nos elementos seguintes, em virtude de esta se tornar redundante (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75).

#### **5.2.3.11. Efeito das *marcas precedentes* na *saliência fónica***

A presente rodada não surge englobada no ponto 5.2.1, no qual se analisa o efeito da variável *saliência fónica* na marcação PL dos itens do SN do PA, visto, naquela etapa do nosso trabalho, a mesma não encadear de modo lógico nas considerações que temos vindo a tecer acerca das motivações que poderão subjazer à inibição PL revelada pelos itens terminados em *-s* e *-r* no singular. A este propósito, relembremos que Poplack (1980a:62-67), ao debruçar-se sobre o EPR, observou os resultados da relação entre *posição linear* e *posição dos marcadores plurais precedentes*, concluindo que uma correcta análise da CPL-var do referido dialecto teria que ter também em conta estudos que relacionassem a segunda variável independente (*marcas precedentes*) com outros grupos de factores, como o *contexto fonológico seguinte* ou a *saliência fónica*. Nesta

conformidade, e dado que o que queremos observar é precisamente a relação de influência na marcação estabelecida entre os itens que pluralizam em *-es* e o seu contexto antecedente, optámos por proceder ao cruzamento entre as variáveis independentes *saliência fónica* e *marcas precedentes*. Os resultados deste cruzamento estão patentes na Tabela 25.18:

**Tabela 25.18.** Efeito das *marcas precedentes* na marcação PL da *saliência fónica*.

<i>Input desta rodada:</i> 0,494	<i>Log-likelihood:</i> -993,213			<i>Significância:</i> 0,014		
<i>Marcas precedentes</i> X <i>Saliência fónica</i>	PL's regulares			PL's nasais (regulares)		
	Nr. total	%	Peso relativo	Nr. total	%	Peso relativo
2 <sup>a</sup> . posição, marca formal na 1 <sup>a</sup> . posição: S__	154/492	31	0,332	15/57	27	0,250
2 <sup>a</sup> . posição, zero na 1 <sup>a</sup> . posição: 0__	33/39	85	0,835	-	-	-
2 <sup>a</sup> . posição, numeral sem <i>-s</i> na 1 <sup>a</sup> : N__	84/362	23	0,529	(2/5)	40	-
2 <sup>a</sup> . posição, numeral em <i>-s</i> na 1 <sup>a</sup> : Ns__	41/202	20	0,491	(6/20)	30	-
NN(N)	14/47	30	0,604	(0/1)	0	-
<i>Marcas precedentes</i> X <i>Saliência fónica</i>	Singulares <i>-r e -s</i>					
	Nr. total	%	Peso relativo			
2 <sup>a</sup> . posição, marca formal na 1 <sup>a</sup> . posição: S__	37/70	53	0,619			
2 <sup>a</sup> . posição, zero na 1 <sup>a</sup> . posição: 0__	(2/2)	100	-			
2 <sup>a</sup> . posição, numeral sem <i>-s</i> na 1 <sup>a</sup> : N__	28/37	76	0,963			
2 <sup>a</sup> . posição, numeral em <i>-s</i> na 1 <sup>a</sup> : Ns__	30/47	64	0,947			
NN(N)	(0/1)	0	-			

Os resultados da tabela anterior indicam uma realidade inquestionável nos nossos dados, para a qual já havíamos chamado a atenção: no cômputo geral, os itens que apresentam alguma distintividade na forma de construir o PL, isto é, os que terminam em *-s* e *-r* no singular e pluralizam com a inserção de *-es* no seu final, apresentam uma taxa de marcação consideravelmente superior à dos itens regulares. Note-se que, nestes últimos, o único factor que favorece categoricamente a marcação PL (pr. 0,835) é o da configuração 0\_ (item analisado na 2<sup>a</sup>. posição, ausência de marcação formal na 1<sup>a</sup>. posição), que, como sabemos, tem tendência em favorecer a inserção da pluralização. A justificação para tal comportamento foi avançada em termos de funcionalidade da língua (Kiparsky, 1972:195). Os outros factores dos elementos regulares seguem também a

linha padrão detectada na comunidade: (i) ausência de distinção não só entre numerais terminados em *-s* e não finalizados em *-s* mas também entre numerais de uma palavra e numerais de mais de uma palavra, com todos eles apresentando valores próximos do patamar da neutralização, mas com ligeiro aclave para favorecerem a marcação; (ii) tendência para a estrutura S\_ (item analisado na 2ª. posição, presença de marcação formal na 1ª. posição) inibir a marcação (pr. 0,332), também explicada nos moldes funcionalistas de Kiparsky (1972:195) e respaldo nas línguas ancestrais africanas (Guy, 1981a:301-302).

Por seu lado, os itens nasais regulares apresentam valores apenas para o factor item analisado na 2ª. posição, presença de marcação formal na 1ª. posição (S\_), a exhibir, tal como nos elementos regulares, uma tendência acentuada para inibir a marcação (pr. 0,250). A fraca tendência para favorecer a pluralização revelada por estes elementos dever-se-á ao facto de os falantes terem dúvidas quanto às diferentes maneiras de pluralizar as palavras em *-ã*, hesitando na selecção da forma a aplicar (Scherre, 1988:124).

Quanto aos itens que fazem o PL com inserção de *-es*, continuam a revelar que os falantes não fazem a distinção entre os diferentes tipos de numerais, que nesta variável favorecem categoricamente a marcação PL (N\_ = pr. 0,963; Ns\_ = pr. 0,947). Por sua vez, a estrutura item analisado na 2ª. posição, presença de marcação formal na 1ª. posição (S\_) revela também tendência para ajudar a inserir a pluralização, embora mais moderadamente (pr. 0,619). Esta tendência, que contraria o rumo apresentado pela mesma configuração em outras ocorrências (a predilecção pela inibição) e, aliada ao facto de os numerais dispararem em termos de favorecimento para patamares muito próximos da marcação plena, parece encontrar explicação na Condição de Distintividade (Kiparsky, 1972:195), já que os falantes do PA, ao que tudo indica, captam e retêm na superfície as formas de pluralização mais salientes. Por outro lado, o acentuado pendor para a marcação PL exibido pelos itens em *-s* e *-r* faz também prever que serão escassas as situações em que ocorre a inibição dos mesmos. Este aspecto, só por si, poderá também debilitar pressupostos avançados na tentativa de explicar as inibições em causa, se sobre as mesmas não for lançado um olhar à lupa. Isto mesmo não o terão feito Braga (1977), Scherre (1988) e Lopes (2001), visto terem tentado generalizar, com recurso a

explicações de cariz meramente fonológico, uma manifestação que não só se nos afigura minoritária como também motivada por fenómenos de índole variada. E quando se trata de efectuar análises em sociolinguística variacionista, aconselha a prudência que se desconfie, logo à partida, das realizações minoritárias. Nesta perspectiva, não poderá também ser descartada a possibilidade de algumas das ocorrências marcadas pela inibição reflectirem apenas variações diastráticas de uma determinada camada etária ou escolar, ou não passem mesmo de uma mera variação idiolectal. E, se num universo de escassas ocorrências houver um único falante, ou restrito grupo de falantes, que reincida em determinado tipo de realização, os resultados finais poderão reflectir uma panorâmica virtual em termos de variação quantitativa, caso sejam acolhidos como representativos da comunidade.

Face ao exposto, decidimos analisar pormenorizadamente os itens em *-s* e *-r* que apresentam inibição da inserção PL, a fim de se procurar escortinar as motivações que levam a esta. O nosso *corpus* possui um total de 167 realizações de itens terminados em *-s* (87 ocorrências, 26 não-marcadas) e *-r* (80 ocorrências, 37 não-marcadas). Temos então um total de 63 ocorrências sem marca de PL, para um universo de 2.340 realizações, o que representa escassos 1,5% do nosso *corpus*. A tabela 25.19 apresenta uma visão geral do modo como estes itens estão distribuídos pelo nossos *corpus*. A este propósito, refira-se que, apesar de termos considerado os itens com marcação parcial de PL como elementos marcados na nossa codificação, para uma melhor análise da distribuição que agora pretendemos observar, optámos por apresentá-los geracionalmente de acordo com os três tipos de realização fonológica produzidos pelos nossos informantes: (i) inibição total da marcação PL (p.e. *quatro MÊS vai picá dedo, tem paludismo* [LUISH2]); (ii) realização parcial da marcação PL (p.e. *fiquei cerca de quatro MESE doente* [ZECAH1]); (iii) e realização da marcação PL (p.e. *durante três MESES já tem milho* [ANTOM1]). Como se pode ver na mesma tabela, devido ao baixo número de realizações totais, limitar-nos-emos a apresentar quer a quantidade de ocorrências quer os percentuais que tanto as faixas etárias como a comunidade levaram a cabo para cada tipo de marcação.

**Tabela 25.19.** Itens com singular em *-s* e *-r* (Marcação fonológico-geracional).

Idade	Tipo de realização					
	Marcação PL		Marcação parcial PL		Ausência marcação PL	
	Nr. total	%	Nr. total	%	Nr. total	%
FE-3 (+60 Anos)	6/46	13	15/46	33	25/46	54
FE-2 (41 – 60 Anos)	17/64	26	17/64	27	30/64	47
FE-1 (20 – 40 Anos)	44/57	77	6/57	11	7/57	12
Comunidade (Total)	67/167	40	38/167	23	62/167	37

Abstraindo agora das marcações de PL, que interessam menos para o fenómeno que pretendemos observar, centremos a nossa atenção nas realizações que apresentam apócope da forma do PL *-es*, ou simplesmente do *-s* final. No que diz respeito à marcação parcial de PL (apócope do *-s* final), verifica-se que, à faixa etária mais idosa, cabe a produção de 15 ocorrências (39% das realizações comunitárias), em números praticamente idênticos aos da FE-2, com 17 ocorrências (45% das realizações da comunidade). Quanto à geração mais jovem de Almoхарife, é responsável por apenas 6 marcações parciais de PL (16% das realizações da comunidade). À primeira vista, o equilíbrio de ocorrências entre as duas gerações mais idosas pode indiciar a emersão do fenómeno na faixa etária mais idosa, que se mantém constante na geração seguinte. Por fim, com a aquisição das regras de concordância em período mais tardio, a geração mais jovem regulariza o padrão da comunidade, acentuando o índice de marcação, que passa a aproximar-se mais do padrão da LA em termos de favorecimento da pluralização.

Contudo, uma observação pormenorizada das ocorrências vai mostrar que esta dedução é falaciosa. Atentando aos falantes que produziram a concordância parcial na FE-3, apercebemo-nos que, do total de 15 realizações registadas, 14 são da responsabilidade da informante [MMDEUSM3]. Confirmaram-se então as nossas suspeitas sobre a viciação que dados minoritários podem encerrar, caso sejam encarados como representativos de determinado fenómeno. Estamos aqui face a um aspecto da problemática da implementação (*actuation problem*) (Labov, 1972a, 1982; Weinreich, Labov & Herzog, 2006 [1968]), acerca das potenciais causas que determinam o fenómeno da mudança linguística, porque, de facto, o que ocorre na FE-3 é uma mera variação idiolectal, que acaba inflacionando, de modo virtual, os percentuais totais de marcação parcial desta geração, em primeira instância, e de todas as ocorrências da

comunidade, em etapa segunda. Ainda a ter em conta nos dados de marcação parcial da FE-3 é o facto de apenas um deles poder configurar um caso de haplogia sintáctica (*ficô aqui em casa nove MESE E tal, moreu [MMDEUSM3]*), se observado na perspectiva do grupo de factores *contexto fonológico posterior*. Por outro lado, não se detectou qualquer realização que pudesse perspectivar a possibilidade de uma assimilação regressiva. Com efeito, as restantes ocorrências possuem os seguintes contextos fonológicos posteriores: 6 casos de pausa final; 6 de pausa interna; e 2 com a consoante [p]. Em adição, se olharmos para o item que antecede as marcações parciais de PL, salta à evidência um aspecto que deverá ser levado em consideração, já que todas elas prefiguram a configuração N\_ (item analisado na 2ª. posição, numeral na 1ª. posição).

Olhando agora para a marcação parcial da FE-1, verificamos que apenas 6 ocorrências (16% do total da comunidade) foram levadas a cabo por este estrato etário. Mais uma vez, apenas uma das realizações poderá configurar um caso de haplogia sintáctica (*cerca de duas ou três VEZE ESSA outra minha também tinha problema [OSVALH1]*). As restantes ocorrências são todas de contexto consonantal posterior, mas nenhuma representa qualquer situação que possa determinar a ocorrência de assimilação regressiva. Dando atenção às *marcas antecedentes*, mais uma vez se regista uma ocorrência maioritária da configuração N\_ (item analisado na 2ª. posição, numeral na 1ª. posição), com 5 casos, sendo do tipo S\_ (item analisado na 2ª. posição, marca formal na 1ª. posição), a realização que resta. O baixíssimo número de marcações parciais registado nesta faixa etária não permite conclusões sólidas, ainda que três delas (50%) sejam da responsabilidade do informante [OSVALH1]. Como tal, parece evidente que os falantes da geração menos idosa de Almojarife tendem para outro tipo de realização: a que implica marcação PL do item analisado, como se pode confirmar na Tabela anterior (44 casos, em 57 possíveis = 77%).

No que concerne à FE-2, a mesma é responsável por 17 realizações parciais da pluralização dos itens em *-s* e *-r* (45% do total da comunidade), dissimiladas pelos seis informantes deste estrato geracional. Como tal, será esta faixa etária a mais representativa, diremos mesmo, a única que espelha o fenómeno numa perspectiva comunitária. Como tal, e em termos de comunidade, a pluralização parcial dos itens em



–s e –r acaba por se configurar como um fenómeno marcadamente diastrático e idiolectal, não reflectindo propriamente uma variação em termos comunitários. Logo, no PA este tipo de variação não pode ser considerado em termos de haplologia sintáctica (cf. Scherre, 1988:251-252) nem de assimilação regressiva (cf. Braga, 1977:72) para efeitos da generalização de pressupostos que poderão motivar a queda do –s final, ainda que se registem alguns casos minoritários de metaplasmos que possam ser atribuídos a estes fenómenos. De facto, das 17 realizações parciais de pluralização na FE-2, apenas duas poderiam configurar haplologia sintáctica (p.e. *quatro MESE ANTÃO eu fui* [ABILH2]) e outras duas assimilações regressivas (p.e. *um ano e três MESE SEÔR já tem dinheiro* [MAURIH2]). As restantes ocorrências de *contexto fonológico posterior* estão distribuídas da seguinte forma: 5 casos de pausa final, 1 de pausa interna, 5 de consoantes diversas, que não conduzem à assimilação, e 2 de vogais, que não possibilitam a haplologia sintáctica. Quanto às marcas antecedentes, temos, mais uma vez, uma situação maioritária de 13 casos configurando a estrutura N\_ (item analisado na 2ª. posição, numeral na 1ª. posição) e 4 situações representando a forma S\_ (item analisado na 2ª. posição, marcação formal na 1ª. posição). Finalmente, cabe ainda referir que, em nossa opinião, a ocorrência de marcação parcial se ficará a dever, nesta faixa etária, a fenómenos socioeconómicos que possibilitam um contacto mais estreito com o sistema de regras de concordância do PE (escolaridade, migração para os centros urbanos, etc.). Este contacto poderá, então, conduzir a um período de variação marcado por incertezas e hesitações na forma de realizar a marcação PL, que apenas estabiliza em fase posterior, isto é, na FE-1.

Relativamente aos contextos fonológicos posteriores que ocorrem conjuntamente com a marcação parcial PL dos itens anteriores em –s e –r de todas as faixas etárias, nota-se que são bastante heterogéneos, não permitindo uma conclusão sólida sobre o peso da sua interferência na marcação fonológica parcial do elemento anterior. Ainda assim, não é de descartar totalmente a possibilidade de, em situações pontuais, condicionarem tal marcação. Contudo, se olharmos agora para as marcas que precedem o item parcialmente pluralizado, nota-se, pelo contrário, uma grande homogeneidade das mesmas, com apenas duas configurações possíveis: N\_ (item analisado na 2ª. posição, numeral na 1ª. posição, com 33 ocorrências) e S\_ (item analisado na 2ª. posição,

marcação formal na 1ª. posição, com 5 realizações). Assim, é possível que sejam as *marcas precedentes* a constituir-se como a motivação maior que leva à marcação parcial PL dos itens em *-s* e *-r*, e não o oposto, isto é, os contextos fonológicos posteriores. De facto, estes poderão não reflectir a realidade das ocorrências em termos comunitários, visto estarem a ser condicionados pelas variações meramente diastráticas ou idiolectais de algumas faixas etárias. Por seu lado, a homogeneidade do comportamento das *marcas precedentes* permite, essa sim, explicações sólidas e que vão ao encontro de outras já avançadas no presente trabalho. Nos casos em que a marca formal de PL precedente S\_ (item analisado na 2ª. posição, marcação formal na 1ª. posição) conduz à queda do *-s* final no elemento seguinte, se bem que a explicação possa ser discutida em termos funcionais da língua, isto é, de os informantes do PA mostrarem tendência para marcar o PL na primeira oportunidade e evitarem a informação redundante subsequente (Kiparsky, 1972:195), a metodologia em Sociolinguística quantitativa propõe que não se considerem ocorrências diminutas para análise de aspectos variacionistas. Como tal, resta-nos a estrutura N\_ (item analisado na 2ª. posição, numeral na 1ª. posição), a única que parece operar realmente de modo sistemático na queda do *-s* do item subsequente.

Como foi observado aquando da análise da influência na marcação motivada pelas *marcas precedentes* representadas por um marcador semântico (ponto 4.1.2.4.6 do presente trabalho), concluímos que, ao que tudo indica, as gerações mais idosas não percebem a marca semântica de PL (Guy, 1981a:180; Scherre, 1988), levando a que marquem o segundo elemento da cadeia sintagmática, quando se confrontam com a estrutura N\_ (item analisado na 2ª. posição, numeral na 1ª. posição). O reduzido número de marcações fonológicas totais (6 casos) ou parciais (1 caso, se excluirmos as realizações da informante [MMDEUSM3]), detectado nos itens em *-s* e *-r* na FE-3, não permite conclusões sólidas. Contudo, é facto adquirido que, no mesmo estrato geracional, a inibição total parece ser condicionada tanto pelo item com marcação formal na primeira posição (13 casos), que conduz o falante no sentido de, funcionalmente, evitar a redundância PL (Kiparsky, 1972:195), como pelo numeral em 1ª. posição (12 casos), cuja noção semântica de pluralização começa a ser percebida por alguns dos informantes, caso dos mais escolarizados como [CELESH3], responsável por mais de 1/3 das inibições registadas. Em sequência de tal, as situações de ausência

de marcação (total ou parcial) ganham peso, sobrepondo-se à das marcações, uma vez que os falantes passam a evitar a redundância da pluralização (Kiparsky, 1972:195).

Quanto à informante detentora da forma idiolectal que inflaciona os valores de marcação parcial na FE-3 ([MMDEUSM3]), é de notar que se trata de uma falante analfabeta (Tabela 8.1), mas também a única que se ausentou da comunidade, tendo permanecido em Angola por prolongados períodos de tempo, onde terá contactado assiduamente com falantes do PtgL2 influenciado por estruturas das línguas bantu.

Debruçando-nos agora sobre a ausência de marcação PL dos itens em *-s* e *-r*, detecta-se a ocorrência de 62 realizações na comunidade. Destas, 25 couberam à FE-3 (40% do total da comunidade), 30 são pertença da FE-2 (49% do total da comunidade) e apenas 7 são da responsabilidade da FE-1 (11% do total da comunidade). Um olhar geral sobre a panorâmica das marcas que antecedem o item com inibição total da pluralização em todas as faixas etárias, vai revelar, uma vez mais, o predomínio absoluto das configurações S\_ (item analisado na 2ª. posição, marca formal na 1ª. posição) e N\_ (item analisado na 2ª. posição, numeral na 1ª. posição). Por outro lado, observando-se os itens do *contexto fonológico posterior* ao elemento que não apresenta marcação PL, detecta-se, novamente, que estes são caracterizados por um extenso leque de realizações distintas, com maior incidência em consoantes, mas que muito raramente poderão originar situações de assimilação regressiva.

No que diz respeito à FE-1, o padrão é em tudo semelhante ao da marcação fonológica parcial, quer no que concerne ao predomínio da marca antecedente N\_ (item analisado na 2ª. posição, numeral na 1ª. posição), com 7 ocorrências (contra 3 da estrutura S\_, item analisado na 2ª. posição, marca formal na 1ª. posição), quer no que respeita ao número total de realizações, bastante exíguo em termos de comunidade. Como tal, abstivemo-nos de analisar este escalão etário em pormenor, devido à baixa solidez que revela para se retirarem conclusões pertinentes em termos de comunidade. Ainda assim, deverá ser levado em conta que esta geração é responsável por 66% dos casos de marcação PL da comunidade (23 casos de estrutura S\_, item analisado na 2ª. posição, marca formal na 1ª. posição, e 21 situações de configuração N\_, item analisado na 2ª. posição, numeral na 1ª. posição), justificando-se os baixos percentuais de não-marcação e marcação parcial do PL com a expansão das noções sintáctico-semânticas de

pluralização dos falantes deste escalão etário. Em adição, gostaríamos igualmente de chamar a atenção para o facto de os contextos fonológicos posteriores ao item com ausência de marcação de PL não configurarem qualquer caso de haplologia sintáctica e, de três casos de consoante, apenas um poder possibilitar a realização de assimilação regressiva (*menos dois PAR SAPATO* [ZECAH1]). Aliás, nesta faixa etária são mesmo detectados casos que poderiam conduzir a situações de metaplasmos, como a haplologia sintáctica, mas que não revelam quedas de segmentos fonológicos (p.e. *nem coesas de dois MESES EU já tive tudo* [CASTEH1]). Este aspecto poderá então fazer prever que a tendência para não eliminar segmentos, detectada em substratos do grupo níger-congo atlântico e no santomense, continua ainda a operar nos falantes do PA. Para finalizar a questão concernente às realizações de *contextos fonológicos posteriores*, registaram-se também as seguintes ocorrências: 2 casos de vogal (que não favorecem a concretização de haplologia sintáctica), 1 de pausa final e 1 de pausa interna.

Acerca da geração mais idosa (FE-3), as realizações com inibição de marcação PL apresentam-se agora em números já aceitáveis para uma análise em termos de Sociolinguística quantitativa, estando distribuídos pelos diferentes falantes da faixa etária, embora em ocorrências desequilibradas. Assim, cabe a cinco dos seis entrevistados a responsabilidade das realizações registadas, reflectindo estas, mais uma vez, a influência homogénea das duas estruturas de marcas antecedentes já apontadas para a inibição da marcação PL: 13 ocorrências com a configuração S\_ (item analisado na 2ª. posição, marca formal na 1ª. posição) e 12 com a forma N\_ (item analisado na 2ª. posição, numeral na 1ª. posição). Quanto à possibilidade de a ausência de marcação ser determinada pelos *contextos fonológicos posteriores*, constata-se, novamente, que estes se revelam bastante heterogéneos, não possibilitando que se lhes atribua um sólido paradigma motivador da inibição da pluralização: 7 casos de pausa interna, 5 de vogais (com três podendo originar situações de haplologia sintáctica, p.e. *havia uns TRABALHADOR TAVA ubabudo aqui* [CELESH3]), 3 de consoantes (com um podendo reflectir uma assimilação regressiva: *compra duas VADÔ SÃO três conto e tal* [MANOH3]) e 1 de pausa final.

Apesar de as formas mais salientes revelarem uma maior tendência para ajudarem à inserção das marcas de PL, este aspecto não se confirma na FE-3, certamente por se

encontrar em fase primitiva de aquisição das noções sintáctico-semânticas da pluralização. Este aspecto reflecte-se na hesitação em realizar a pluralização de itens irregulares, como bem o demonstra o reduzido número de marcações: apenas 6 casos, distribuídos por dois informantes.

Quanto à FE-2, apresenta igualmente um considerável número de inibições da marcação PL, partilhado por cinco dos seus informantes. Observando-se o comportamento das *marcas precedentes* ao item que apresenta inibição da marcação PL, constata-se, de novo, a prevalência das estruturas já anteriormente mencionadas: 17 casos de configuração S\_ (item analisado na 2ª. posição, marca formal na 1ª. posição) e 12 ocorrências com a estrutura N\_ (item analisado na 2ª. posição, numeral na 1ª. posição). Nesta faixa etária regista-se ainda outra estrutura, mas englobando também um numeral na primeira posição: N0\_ (item analisada na 3ª. posição, antecedido de numeral na 1ª. posição e ausência de marcação na 2ª. posição). Relativamente às ocorrências em *contexto fonológico posterior*, deparámo-nos novamente com um extenso e variado número de possibilidades, tendo sido listadas as seguintes realizações: 15 casos de consoante (apenas 1 podendo configurar uma situação de assimilação regressiva: *mais de quato... cinco MÊS SEM... com... sem tocá* [LUISH2]), 7 de vogais (nenhum possibilitando a concretização de haplogogia sintáctica), 6 de pausa final e 2 de pausa interna.

Lançando novamente mão da análise efectuada sobre a influência na marcação motivada pelas *marcas precedentes* representadas por um marcador semântico (ponto 4.1.2.4.6 do nosso trabalho), verificou-se que as gerações mais velhas não percebem a marca semântica de PL, acontecendo esta percepção em fase mais tardia, sobretudo via escolaridade (cf. Scherre, 1988:175). Como tal, o índice de marcação nos elementos posteriores aos itens com semântica de pluralização decai na FE-2 (apenas 5 casos marcados, contra 12 não marcados), reforçando a consistência da não-marcação que visa evitar a redundância da pluralização (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75), também presente nas estruturas com marca formal na 1ª. posição (S\_). Este aspecto é, pois, revelador de que o falante, com a aquisição das noções sintáctico-semânticas de pluralização, começa a evitar a redundância da marcação PL (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75), originando um padrão que corresponde ao

nivelamento entre os padrões de inibição motivados por elemento com marca semântica de pluralização e elemento com marca formal de PL, com o primeiro acentuando a inibição da marcação para ir ao encontro da constância do segundo e passar a estabelecer-se num patamar próximo do ocupado por este. Posteriormente, a FE-1 estabiliza o padrão, que se mostra, em termos gerais, favorável à inserção do PL, visto conjugar três aspectos: (i) o propiciamento da marcação PL que emerge na FE-3, em estruturas que envolvem a marca antecedente realizada por marcador semântico, que nunca chega a atingir o patamar da inibição, apesar de os falantes ganharem a noção da pluralização semântica e passarem a evitar mais a redundância da marcação; (ii) a aquisição da regra da concordância, que leva a subir ligeiramente a tendência para se marcarem, redundantemente, os elementos posteriores ao item com marca formal de PL; (iii) a aquisição da noção de como realizar correctamente o PL dos itens que apresentam distintividade fonológica mais acentuada.

A análise geral do modo como a variável independente *contexto fonológico posterior* influencia a marcação PL do item que a antecede indicou-nos que a predição levantada na hipótese 4 (ponto 3.6.3) apenas se confirma parcialmente. De facto, a pausa final é o elemento que mais favorece a inserção de marcas, aspecto que apenas não seria corroborado aquando da observação da influência na marcação PL exercida pelas zonas e pontos de articulação das consoantes pela FE-1. A pausa interna, por seu lado, apresenta-se como o segundo elemento que mais favorece a inserção PL, apesar de exibir quer um valor próximo da neutralidade quer algumas oscilações geracionais. Ainda em termos da predição avançada, as vogais não evidenciam um favorecimento à marcação diferenciado das consoantes, já que ambas se situam no patamar da inibição ligeira. Quanto à influência das consoantes na marcação PL, confirma-se a predição avançada na hipótese 4, já que a inserção de marcas será mais facilitada pelas surdas do que pelas sonoras. Ainda assim, a diferença probabilística entre vogais e consoantes, no que concerne ao efeito para a inserção da pluralização no item antecedente, não é relevante. Paralelamente, as nossas análises permitiram constatar que a não-obstrução à passagem de ar pelo tracto bucal favorece a inibição, o que nos permitiu estabelecer a seguinte escala hierárquica das consoantes no favorecimento à marcação, de acordo com as suas zonas e pontos de articulação: (i) palatais, com especial aclave para as velares e

declive para as apicais; (ii) labiais; (iii) dentais, com afixação para as lábio-dentais e declive para as ápico-dentais; (iv) alveolares. Por seu lado, a observação do comportamento das ocorrências em que intervêm os itens singulares que finalizam em *-s* e *-r* e pluralizam com a inserção de *-es* final revelou que, em alguns estudos anteriores, possivelmente se terão generalizado pressupostos com base em ocorrências minoritárias para justificar as inibições motivadas pelos mesmos, assentando tais pressupostos em fenómenos como a haploglia sintáctica (Scherre, 1988:244) ou a assimilação regressiva (Braga, 1977:72). Como tal, a motivação maior para a marcação/inibição PL dos itens do SN não é de carácter foneticofonológico, resultando essencialmente de aspectos estruturais conotados às marcas precedentes ao item analisado. Note-se que a fraca influência das variáveis do tipo fonológico na marcação PL foi também detectada por Jon-And (2008:7, 2009:7) nos estudos sobre o PMp e o PCV, nos quais a motivação maior para a inserção de marcas assenta, igualmente, nas variáveis do tipo estrutural (pontos 1.3.1.3 e 1.3.2.1 do presente trabalho). Portanto, a questão da fraca influência fonológica na marcação, um fenómeno também identificado na aquisição do PtgL1 (Faria, Freitas & Miguel, 2001:53), vem argumentar contra a possibilidade de as alterações nos padrões de concordância assentarem em processos de alteração fónica, históricos e internos ao próprio PE, como a deriva românica, que consubstancia a perda das consoantes finais *-s* e *-m* (Naro, 1981; Scherre, 1988; Naro & Scherre, 2000). Em contrapartida, a mesma influência fonológica na marcação argúi a favor da hipótese da influência das línguas de substrato, levantada por Guy (1981a:3001-302), e que sugere que a variação na concordância PL ancora em um ou mais substratos.

Após termos analisado os resultados das variáveis independentes do tipo fonológico, passaremos a observar a influência das variáveis extralinguísticas na marcação PL do SN da fala de Almojarife.

### **5.3. Variáveis extralinguísticas**

À excepção dos trabalhos de Figueiredo (2008, 2009a, 2009b), não temos conhecimento de terem sido efectuados quaisquer outros estudos linguísticos sobre a fala da comunidade de Almojarife, nem que existam descrições de estágios passados deste dialecto. Nesta conformidade a observação da evolução diacrónica do mesmo terá que

ser coerente e consistentemente inferida a partir dos dados sincrónicos que constituem o nosso *corpus*, por forma a definirmos, a partir dos resultados exibidos pelas nossas variáveis, qual será o estágio actual da sua variação.

O estudo da influência das variáveis extralinguísticas na configuração da CPL-var do SN do PA irá ser levado a cabo com base em duas perspectivas: (i) analisando o efeito que cada uma das variáveis exerce, isoladamente, sobre a concordância de número; (ii) buscando inferir, sustentadamente, o modo como as diferentes evidências empíricas concedidas pelas distintas variáveis sociais se articulam entre si.

Seguidamente, passamos a descrever as etapas do tratamento dado às variáveis independentes extralinguísticas, a que se seguirá a análise dos respectivos resultados. Todas as variáveis foram consideradas relevantes para o estudo da CPL-var do SN do PA e serão analisadas de acordo com a ordem hierárquica de importância estabelecida pelo suporte VARBRUL, patente na Tabela 14.1.

### **5.3.1. Variável independente *escolaridade***

A variável independente *escolaridade* tem de ser observada em função de condicionalismos específicos que afectaram o sistema de ensino de São Tomé, em geral, e a comunidade de Almocharife, em particular. De acordo com os dados apresentados na Tabela 8.1, a classe mais idosa revela elevado nível de analfabetismo, com apenas dois informantes tendo frequentado o ensino primário – um do sexo masculino (3ª. classe) e outro do sexo feminino (1ª. classe) –, mas sem o concluir. Esta é a geração que, à época em que teria idade para frequentar a escolaridade primária, foi proscrita pelo Estado Novo, sentindo os efeitos das medidas exploratórias e discriminatórias impostas pelo *Acto Colonial* (ponto 1.3.2.3 do presente trabalho), as quais reduziram também a relativa autonomia de que dispunham as administrações dos territórios colonizados. Atirada para a condição de indigência e isolamento, esta geração poucas possibilidades teve para aceder ao ensino formal da LA, ou seja, do PE.

Posteriormente, com o desenvolvimentos das estruturas socioeconómicas da colónia, e esbatidas que estavam também determinadas tensões entre nativos e autoridades administrativas, massificaram-se os programas de alfabetização e o acesso ao ensino tornou-se mais acessível para os nativos. Motivada também pelo prestígio da LA, a faixa etária intermédia dos nossos informantes revela um grau de escolaridade mais elevado,



com o índice de analfabetismo a cair ligeiramente (apenas 3 informantes não frequentaram a escola) e alguns a concluírem já o ensino primário. Contudo, é de ter em conta que os informantes do sexo masculino detêm, em ambas as faixas etárias referidas, um índice de escolarização superior ao das mulheres (Tabela 8.1).

Quanto à geração mais nova de Almojarife, é toda ela escolarizada. Contudo, os níveis de escolarização variam do baixo ao alto. Neste aspecto, poderia causar perplexidade o facto de os informantes mais idosos desta faixa etária serem os mais escolarizados. Contudo, se atendermos às realidades sociais colonial e pós-colonial, constata-se que estes se enquadram exactamente na percentagem dos naturais que frequentavam *o Liceu e Escola Técnica* no período imediatamente anterior à independência, conforme nos dão conta os relatórios do exército português à época (ponto 1.3.2.3 do nosso trabalho). Por outro lado, os falantes menos idosos deste estrato etário possuem um grau de escolarização inferior (ensino primário), confirmando a desarticulação que afectou as estruturas de ensino após a retirada definitiva dos portugueses. Estamos em crer que, se este aspecto tem influência directa na aquisição da regra de concordância por parte dos indivíduos desta faixa etária, não menos o terá também o facto de o padrão de registo (o português de São Tomé) do novo corpo docente, constituído por professores nativos que substituíram os do quadro colonial, se aproximar mais do dos seus próprios alunos. De facto, se tivermos em conta que a existência de uma prática padronizada transmitida pelo ensino é reconhecida pelos falantes de uma determinada língua, conclui-se também que a ausência de *feedback* negativo relativamente à variação produzida pelos alunos (Vigil & Oller, 1976:287; Corder, 1981b:72), contribui com transferência por treinamento (*transfer of training*) (Higgs & Clifford, 1982:62; Sims, 1989:66) para a interiorização e estabilização/fossilização das formas não-padronizadas (Gaies, 1977, 1979; Ellis, 1985; Valette, 1991).<sup>189</sup>

Por outro lado, os aspectos apontados levam ainda a dar atenção ao pressuposto observado por Lucchesi (2000a:293) de que a variável *escolaridade* não pode ser observada à margem de variáveis sociais como a *idade* ou o *sexo*, nas quais se encontra intrinsecamente embrionada. Portanto, ao constituirmos o grupo de factores *escolaridade*, optámos, inicialmente, por envasar quatro factores no mesmo grupo de

factores, de acordo com o representado na Tabela 12.15. Contudo, após uma observação atenta das características sociais de cada informante, apercebemo-nos que um deles, [MINISM3], possui um baixo grau de escolarização (frequência da 1ª. classe), bastante próximo do analfabetismo, portanto. Como se viu também no ponto 2.3.11, estatisticamente dever-se-á ter em conta a possibilidade de os informantes possuírem idiolectos marcados, o que leva a pequenas flutuações nos dados, que acabam por se desviar da tendência generalizada do grupo. Como os desvios só podem ser reduzidos ou anulados pela junção, no mesmo factor, de vários indivíduos que compartilhem características linguísticas ou sociais, optámos também por testar a inclusão da informante de baixa escolarização no factor dos analfabetos, efectuando o teste do qui-quadrado a este novo factor. Comparando-se os resultados de ambos os grupos de factores, seria então possível verificar em qual deles se deveria incluir a informante, de forma a incorporar o factor seleccionado na análise.

Da inclusão da informante [MINISM3] no grupo dos analfabetos resultou o grupo de factores descrito na Tabela 26.1:

**Tabela 26.1.** Grupo de factores *escolaridade* e respectivos factores, para estudo da configuração do SN PL do PA: grupo de factores para teste de factores a incorporar na análise de dados.

<b>Factores</b>	<b>Exemplo</b>	<b>Nr. total</b>	<b>%</b>
Analfabetos/semianalfabetos (1ª. Classe)	tem DOIS <b>GÉMIA</b> de três ano [LUISH2]	232/672	35
Escolarização média (3ª. e 4ª. classes)	pissoa mandô AS <b>MOBÍLIA</b> [OSVALH1]	722/1.316	55
Escolarização alta (Frequência do ensino pós-primário)	faço <b>TRABALHOS PRIVADO</b> [CASTEH1]	248/352	70
<b>Totais:</b>		1.202/2.340	51

Efectuámos então as rodadas simples (*one-level analysis*) para cada um dos grupos de factores, que forneceu os seguintes *log-likelihoods*: Tabela 12.15 = -912,246; Tabela 26.1 = -918,096. Comparando ambos os valores, verifica-se que o resultado referente à rodada efectuada com o grupo de quatro factores se revelou inferior ao da rodada efectuada com o grupo de três factores. Como tal, está inviabilizada a possibilidade de se efectuarem os cálculos para achar o qui-quadrado, uma vez que o grupo amalgamado

possui significância estatística, não devendo ser incorporada à análise. Para efeitos de análises posteriores, preservou-se, então, o grupo de factores inicial (Tabela 12.15).

Após efectuada a rodada em que se inseriu, no suporte computacional, a variável social *escolaridade* composta pelos factores iniciais, obtiveram-se os resultados indicados na Tabela 26.2. Lembremos que, de entre as variáveis sociais, a variável *escolaridade* foi indicada pela ferramenta VARBRUL como sendo aquela que mais influência tem no desenho da CPL-var do SN do PA (Tabela 14.1), confirmando que o número de anos de estudo constitui um dos factores que mais influência tem na aquisição da regra de concordância redundante. Contudo, os resultados mostram que são os informantes com frequência do ensino primário os que mais realizam a concordância PL (pr. 0,690), sobrepondo-se mesmo aos falantes com frequência do ensino pós-primário (pr. 0,663). Por seu lado, os analfabetos surgem como aqueles que mais inibem a concordância, e de um modo bastante significativo (pr. 0,296). Como se vê também, é a partir do momento em que os falantes passam a frequentar o ensino que se dá a maior expansão da aplicação da regra, já que o ápice desta acontece na transição do estado de analfabetismo para o de escolarizado. Este aspecto, por si só, vem confirmar, mais uma vez, como é forte a influência da variável *escolaridade* no sentido de os falantes começarem a direccionar o seu padrão linguístico com vista à aplicação da concordância.

**Tabela 26.2.** Efeito da *escolaridade* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Analfabetos	187/593	32	25	0,296	
Escolarização baixa (Frequência do ensino primário)	180/338	53	15	0,690	
Escolarização média (Ensino primário concluído)	433/862	50	37	0,401	
Escolarização alta (Frequência do ensino pós-primário)	402/540	74	23	0,663	
<b>Totais:</b>	<b>1.202/2.333</b>	<b>51</b>	<b>100</b>	<b>-</b>	

A ausência de uma linha curvilínea ou de uma inclinação progressiva na aquisição da regra da concordância de número vem levantar a possibilidade de o PA não estar em

estádio de mudança em progresso, situação que se procurará confirmar com a observação das outras variáveis sociais. Comparando ainda os segundo e terceiro grupos, isto é, os falantes com frequência do ensino primário (pr. 0,690) e informantes que o concluíram (pr. 0,401), verifica-se que ocorre uma convivência de forças contrárias, ou seja, de aquisição e posterior perda da marcação PL. Os dados revelam a mistura de influências em direcções distintas, característica do quadro de sociedades cuja sociohistória é marcada por períodos de apogeu e declínio, como sucedeu, pelo menos por duas vezes, em São Tomé: ascensão e queda dos engenhos do açúcar e ascensão e queda das roças do cacau e café. Deste modo, os factores que constituem a variável *escolaridade* não podem ser linearmente entendidos como escolaridade baixa igual a gerações mais idosas e escolaridade alta sinónimo de faixa etária mais nova. De facto, basta atentar ao período que se sucedeu à independência de São Tomé, com a consequente desarticulação do seu sistema de ensino, para perceber que muitos dos jovens de Almojarife, enquanto membros de uma comunidade rural que foi votada ao abandono e isolamento durante esta fase da vida do país, estarão incorporados nos factores dois e três, devido à impossibilidade de prosseguirem os seus estudos, (cf. Tabela 8.1). Paralelamente, os elementos da comunidade que tiveram oportunidade de levar a cabo estudos pós-primários, apesar de estarem enquadrados na faixa etária mais nova do nosso estudo, constituem a classe mais idosa deste factor, e que teve a oportunidade de beneficiar das condições de escolaridade fornecidas pela administração colonial nos anos imediatamente anteriores à independência do país. Por outro lado, se levarmos em linha de conta que o ensino possível em São Tomé, após a independência, se terá demarcado mais do padrão do ministrado anteriormente (o PE), sendo assegurado por docentes que apresentam um registo linguístico bastante próximo do dos aprendentes, compreende-se melhor o porquê da flutuação registada na aquisição da regra, com os falantes do antigo sistema de ensino a revelarem maiores índices na aplicação da regra da pluralização do que os informantes que concluíram o ensino primário durante o sistema de ensino pós-independência. Ainda a este propósito, relembremos aqui que, quanto menos escolarizada for uma comunidade, menos resistente se mostram os seus falantes à aceitação da variação, já que lhes falta o vector que mais apego revela em relação às formas padronizadas.

Paralelamente, os jovens falantes que constituem os factores três e quatro, com a partida dos colonos e a extinção da administração portuguesa, deixaram de estar sujeitos à pressão de cima para baixo que fora experimentada pelos adultos que compõem o factor dois. Estes últimos, como se sabe, são os representantes de um grupo que lutava pela melhoria das condições de vida, migrando para os centros urbanos, onde as oportunidades para servir na administração pública, e até no exército, eram mais facilmente concedidas àqueles que dominavam razoavelmente a língua portuguesa. Como tal, estes falantes tiveram uma história de aprendizagem do português através de DLP's mais próximos do PE, bastante diferente da dos elementos que permaneciam na comunidade.

Comparando os valores do PA com os de outras três variedades de português, verificamos que o primeiro apresenta um quadro bastante distinto, com as segundas denotando um estágio de mudança em curso no sentido da aquisição da regra de concordância, enquanto em São Tomé estará a acontecer um estágio de variação estável, que pode levar à de retracção da regra da pluralização.

**Tabela 26.3.** *Escolaridade*: contribuição dos factores individuais para a marcação PL do elemento analisado [4 variedades de português].

Factores	Escolaridade							
	PA		MRJ		NURC		PCV	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%
Analfabetos	0,296	32	-	-	-	-	-	-
Escolarização baixa	0,690	53					0,36	69
Escolarização média	0,401	50	0,32	62	0,18	64	0,39	78
Escolarização média-alta	0,663	74	0,53	72			0,59	87
Escolarização alta	-	-	0,65	82	0,45	82	-	-
Ensino superior	-	-	-	-	0,83	96	-	-

Scherre (1988:444) aplicou três factores à variável social *escolaridade*: 1 a 4 anos de ensino (Primário), 5 a 8 anos de ensino (Ginasial) e 9 a 11 anos de ensino (Colegial). Para além desta distribuição, a autora analisou estes grupos em função do seu estrato social, metodologia que vai ao encontro da que fora aplicada por Naro & Lemle (1976) no estudo para o estabelecimento de uma hipótese geral sobre a mudança sintáctica

operada no PB. Nas nossas observações, tal metodologia não foi usada, uma vez que os informantes do PA configuram um único tipo de factor social: o de ambiente humilde.

Ao estabelecer a correlação entre *escolaridade* e ambiente social, Scherre (1988:448) indica-nos que no grupo de ambiente humilde e concordância baixa, a linha é bastante inclinada, com ápice de formas não-padrão no segmento mais jovem. Quanto ao grupo de ambiente não-humilde e concordância alta, apresenta uma linha curvilínea, com as faixas etárias mais nova e mais idosa produzindo formas não-padrão bastante idênticas e a faixa etária intermédia apresentando o ápice das formas de prestígio. Assim sendo, o primeiro grupo denuncia uma mudança linguística na direcção de um sistema sem marcas de concordância, enquanto o segundo evidencia que a variação se vai estabilizando, à medida que os falantes vão também aumentando a sua idade. Paralelamente, os falantes de ambiente humilde e concordância alta, que a autora classifica como possíveis “descriolizados” (Scherre, 1988:456), revelam também que ocorre uma proporção directa entre aumento de concordância e incremento da escolarização, já que os seus índices de captação da regra são determinados pela influência elitizante da escola e não pelo avançar da idade. Por outro lado, os falantes de ambiente não-humilde e concordância baixa são influenciados tanto pela escolarização como pela idade, o que os leva a apresentarem, geracionalmente, um padrão de perda.

Através dos resultados do MRJ, é possível perceber como é complexo o jogo entre variáveis sociais, confirmando-se que estas não podem ser observadas de forma compartimentada, conforme entendimento de Chambers & Trudgill (1980:97-98), a propósito da variável social *sexo*. De facto, se olharmos para os valores do MRJ (Tabela 26.3), verificamos que o incremento do uso da regra de concordância se dá na transição do Primário para o Ginásial, visto situar-se aí o ápice da inovação. Por outro lado, seríamos também levados a concluir que o processo de mudança se encontra ainda em curso, já que os alunos do nível Colegial continuam a apresentar uma tendência para aquisição da regra de concordância PL. Contudo, com recurso à observação da relação que a variável *condição social* estabelece com as variáveis *escolaridade* e *idade*, foi possível a Scherre (1988:454) concluir que o MRJ exhibe dois tipos de padrão: um apontando para uma tendência natural de mudança linguística, nos moldes preconizados

por Naro & Lemle (1976); outro exibindo uma variação sociolinguística estável, de acordo com o defendido por Guy (1981a, 1986).

Os resultados de Lopes (2001:107), por seu lado, apontam para uma relação directa e proporcional entre o aumento de tempo de exposição à pressão escolar e a maior probabilidade de ocorrer concordância no SN. Note-se que o ápice da inovação se situa na transição do Colegial para o Universitário, conotado à norma culta, a qual, por sua vez, se aproxima bastante da do PE em termos de concordância. Este aspecto vem ainda complementar os resultados de Scherre (1988:454), que apontavam para o facto de os elementos integrados no factor *escolaridade alta* (Ginasial) se encontrarem ainda em estágio de mudança em progresso.

P. Andrade (2003:126-129) não apresenta valores para a variável *escolaridade*, tendo observado esta de uma forma intrínseca com a variável *idade*, isto é, estabelecendo uma relação entre esta variável e a sociohistória da comunidade de HEL-BA. Baxter (2004, 2009) também não observou o efeito que esta variável social exerce na concordância PL do PT, ao passo que Jon-And (2009:4) o fez apenas relativamente ao PCV.<sup>190</sup> Neste, a inibição da marcação mantém-se estável até aos 4 anos de escolarização, isto é, desde a fase de analfabetismo até ao final do período primário. Posteriormente, os falantes detentores de escolarização média-alta passam a beneficiar a marcação PL, ocorrendo o ápice da inovação na fase de transição do ensino primário para o ensino liceal. A linha de ascensão no sentido de aquisição da regra faz também pressupor que, em Cabo Verde, esta se encontra em estágio de mudança em progresso.

No caso concreto do PA, as nossas observações apontam para a confirmação do pressuposto que levantámos para a hipótese 15, constituída para a variável *escolaridade* (ponto 3.6.14). Assim, a atenção dos falantes para com os mecanismos de concordância aumenta com o incremento da frequência escolar. Após ingressarem na escola, os falantes de Almoxarife começam a flexionar mais os itens pluralizáveis do SN. Contudo, com a frequência do ensino, os membros da comunidade passam também a produzir SN's de estrutura mais complexa, o que se traduz na hesitação da aplicação da concordância numa etapa inicial da segunda fase de escolarização. Numa terceira fase, em que se poderiam consolidar as regras de concordância das estruturas sintagmáticas mais extensas e complexas, ocorreu uma desarticulação do sistema socioeconómico do

país, que impediu que a regra de concordância se continuasse a desenvolver. Este processo de desaceleração e posterior estabilização da variação no SN encerra todas as características que fazem ainda pressupor a possibilidade de vir a ocorrer fossilização (Long, 2003:490). De facto, alguns estudos sobre ASL defendem que a principal causa de fossilização de formas incorrectas na interlíngua tem a ver com a falta de instrução formal na LA e, conseqüentemente, de *feedback* negativo sobre a variação produzida (Vigil & Oller, 1976:287; Corder, 1981b:72). Assim sendo, Valette (1991:329) entende que a aquisição por mero contacto leva o aprendente a fabricar uma interlíngua ou línguas idiossincráticas com regras que diferem das do padrão da LA. Segundo o autor, os “falantes de rua” de uma L2 apresentam mais formas fossilizadas do que “os aprendentes de sala-de-aulas”, uma vez que contactam mais facilmente com padrões lexicais e sintácticos que são inapropriados na LA. Assim, a fossilização é agravada pelo facto de os “falantes de rua” não serem corrigidos nem se autocorrigirem. Embora Ellis (1988:307) entenda que uma das características do ensino é evitar a fossilização, quer Higgs & Clifford (1982:62) quer Sims (1989:66) compartilham o entendimento de Valette (1991:329), afirmando que, por vezes, a própria escola se encontra impossibilitada de desfossilizar ocorrências de determinados aprendentes, uma vez que a aprendizagem formal da L2 compete com a “aprendizagem de rua”, sendo esta mais intensa e inibindo o normal progresso da primeira. Por outro lado, apesar de o ambiente em sala-de-aulas se propiciar à correcção (ou desfossilização) de determinadas estruturas, Richards (1971:216) e Stentson (1974:183) referem a possibilidade de um ensino ministrado por um instrutor com um padrão de proficiência também variável contribuir para a fossilização de determinadas formas. Ora, este é exactamente o caso de São Tomé, desde que os colonos portugueses se retiraram em definitivo da ilha.

Apesar de tudo, afirmar que a aprendizagem da L2 em ambiente informal motiva mais a fossilização do que aquela que acontece em situação formal, ou vice-versa, é, no mínimo, deter uma visão superficial acerca deste fenómeno. Em ambos os tipos de aquisição devem ser consideradas não só as situações particulares em que a aprendizagem se dá mas também as participações dos intervenientes envolvidos nesta. Estes aspectos conduzem, por exemplo, a questões como a monotorização. Ellis (1985:85) salientou que a linguagem a que os docentes recorrem no processo de



interacção com os alunos em ASL é tratada como um registo nas suas especificidades. O facto de determinados professores utilizarem modelos interaccionais que se afastam das formas coloquiais ou a ajustes discursivos, sejam eles excessivamente padronizados sejam eles simplificados, leva os alunos a interiorizarem (e mesmo fossilizarem) registos que se distanciam das formas usuais da LA (Gaies, 1977, 1979; Ellis, 1985).

A este propósito, Krashen & Terrell (1983) e Krashen (1985) advogam que dois sistemas de conhecimento distintos concorrem para o maior ou menor grau de desempenho na L2: o sistema de aquisição e o sistema de aprendizagem. O primeiro consiste de um conhecimento subconsciente da gramática da L2, assemelhando-se àquele que os falantes detêm em relação à L1. Quanto ao sistema de aprendizagem, detêm importância inferior e resulta da instrução formal, implicando o conhecimento consciente das regras. Seliger (1988:39), por seu lado, defende que o falante só tem acesso à gramática aprendida caso respeite três condições: i) tenha tempo para reflectir sobre a regra da gramática que armazenou; ii) durante a comunicação, concentre a sua atenção no conteúdo da mensagem e na forma a aplicar; iii) conheça a regra em uso. Este último pressuposto não se vem verificando em São Tomé, pelas razões que já apontámos e se prendem com o registo produzido pelos actuais docentes, falantes do português santomense. Portanto, quando o aprendente memoriza determinada produção linguística ou realiza uma produção que requeira atenção consciente para a sua execução, está a accionar o monitor, ou seja, a parte do sistema interno responsável pela realização do processo linguístico consciente. Para determinados autores, que concordam com os postulados de Krashen & Terrell (1983) e Krashen (1985), o grau de utilização do monitor varia em função de diversos aspectos, como a idade, as características do aprendente e a quantidade e qualidade do *input* (Dulay, Burt & Krashen, 1982). Ainda assim, alguns autores demarcam-se das teorias de Krashen & Terrell (1983) e Krashen (1985), alegando que os autores não estabelecem uma distinção exacta entre os domínios da aquisição e aprendizagem, pelo que é impossível determinar, empiricamente, qual delas está a actuar em determinado momento da aquisição (McLaughlin, 1989). Por outro lado, também não pode deixar de ser considerado que a estabilização é permeável e reversível, podendo ser corrigida (Long, 2003:490). Nesta conformidade, a concordância variável registada no PA pode também

caminhar no sentido da aquisição da regra, caso o panorama do ensino de São Tomé se altere no futuro, isto é, sejam incluídos docentes falantes do PE no quadro de professores do país.

### 5.3.2. Variável independente *idade*

Como se pode constatar na Tabela 27.1, a FE-3 é aquela que produz menos realizações para análise do desenho da CPL-var no SN do PA. Depois, o número de ocorrências vai aumentando gradualmente, confirmando que as novas gerações incrementam também a aquisição de estruturas da LA, a ponto de a FE-1 produzir quase o dobro de realizações relativamente à FE-3. O maior número de ocorrências não significará necessariamente a produção de mais SN's, já que estes valores apenas podem ser constatados com recurso à análise não-atomística. Contudo, ao contrário do que sucede na FE-3, as FE-1 e FE-2 deixam antever a produção de SN's de estruturas cada vez mais complexas, envolvendo mais do que dois ou três itens, e que, atomisticamente, englobam mais elementos passíveis de serem analisados.

A variável independente *idade* não necessitou de qualquer tratamento adicional, uma vez que os factores inicialmente constituídos (Tabela 12.13) se apresentam todos como relevantes para a marcação PL dos itens do SN. Depois de levados à análise VARBRUL, os dados do grupo de factores *idade* apresentaram os seguintes valores:

**Tabela 27.1.** Efeito da variável *idade* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011	
<b>Escalão etário</b>	<b>Nr. Total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
FE-1 (de 20 a 40 anos)	623/984	63	42	0,658	
FE-2 (de 41 a 60 anos)	373/786	34	34	0,404	
FE-3 (+60 anos)	206/570	24	24	0,357	
<b>Totais:</b>	1.202/2.340	51	100	-	

A diferença de pesos relativos entre as três gerações revela que é a FE-1 aquela que mais aplica a regra da concordância (pr. 0,658), sendo mesmo a única que apresenta um valor moderadamente favorecedor do uso da pluralização. Note-se ainda como esta faixa etária se distancia das outras duas em termos de favorecimento à marcação, revelando as FE-2 (pr. 0,404) e FE -3 (pr. 0,357) uma maior semelhança em termos de padrão

linguístico, com tendência para inibir a pluralização. Pelos valores exibidos é também possível inferir que, com o aumento do grau de escolarização, os falantes mais novos de Almojarife vão adquirindo formas novas e configurando uma variação livre, até atingirem nova sistematização, agora de modo mais próximo da gramática do PE. Este aspecto indicia também a existência de gramáticas distintas nas diferentes faixas etárias da comunidade, com os valores de marcação a revelarem uma tendência progressiva no sentido de as novas gerações incorporarem mais a regra da concordância PL. Este aspecto pode ser confirmado através dos valores dos SN's plenamente marcados, por escalão etário:

**Tabela 27.2.** SN's plenamente marcados no PA: valores por *idade*.

<i>Input desta rodada:</i> 0,223		<i>Log-likelihood:</i> -694.131		<i>Significância:</i> 0.001	
<b>Escalão etário</b>	<b>Nr. Total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
FE-1 (de 20 a 40 anos)	247/588	42	40	0,606	
FE-2 (de 41 a 60 anos)	118/497	24	34	0,529	
FE-3 (+60 anos)	60/404	15	26	0,411	
<b>Totais:</b>	<b>425/1.488</b>	<b>29</b>	<b>100</b>	<b>-</b>	

Em Almojarife, o peso da marcação plena para a FE-3 situa-se em valores levemente inibidores (pr. 0,411), mas o percentual da mesma revela que os membros mais idosos da comunidade raramente realizam a concordância entre todos os elementos do SN, já que apenas 15% dos SN's produzidos por esta faixa etária se apresentam completamente marcados. Como esta faixa etária realiza, essencialmente, SN's de estrutura reduzida (DET+N ou numeral+N), levanta-se, de imediato, a questão de a marcação ser inserida em apenas um dos elementos do SN, um fenómeno que Guy (1981a:301-302) advoga ter respaldo no sistema de marcação PL das línguas dos ancestrais substratos africanos.

Seguidamente, a geração intermédia demonstra que possui uma maior consciência acerca da regra da concordância, que se situa já em plano neutro, mas a propiciar de forma leve a marcação plena (pr. 0,529). Esta faixa etária passou a beneficiar de um maior contacto com o PE, muito por força da frequência escolar de nível médio, como se viu no ponto 5.3.1. Com o grau de escolarização a aumentar na geração mais jovem, esta passa já a apresentar um valor a beneficiar moderadamente a concordância plena (pr. 0,606). O facto de ser a FE-1 a realizar mais a regra de concordância PL faz pressupor

que, à partida, o PA se encontra ainda em processo de mudança em curso. Com efeito, se espelharmos no passado os valores apresentados para as diferentes faixas etárias, apercebemo-nos que o maior incremento em direcção à aquisição da regra de concordância de número terá ocorrido no período de tempo que separa a FE-2 da FE-1. Seguindo o raciocínio lógico de que o padrão de comportamento linguístico se fixa até à adolescência, então os falantes com 20 anos tê-lo-ão cristalizado até cinco anos antes da recolha do *corpus*, isto é, até 1993. Deste modo, como a FE-1 engloba os informantes com idades compreendidas entre os 20 e os 40 anos, é lícito concluir que a variação da concordância PL teve o seu maior incremento no último terço do século XX, mais propriamente entre a fase de transição da década de 60 para década de 70 e a passagem dos anos 80 para os anos 90. Como vimos, os relatórios periódicos dos efectivos militares de São Tomé permitem-nos confirmar que o período que marca o início da maior aplicação da regra de concordância de número do PA condiz, exactamente, com aquele em as autoridades coloniais faziam referência à boa cobertura do ensino, com consequente elevação do nível da educação da população até patamares considerados excelentes e que fizeram com que praticamente toda ela falasse o português (ponto 1.3.2.3 do presente trabalho).

Paralelamente, os mesmos relatórios, ao apresentarem os dados do recenseamento da população, permitem não só constatar que é igualmente nesta fase que se dá o aumento do movimento migratório das zonas rurais para a capital do arquipélago mas também que os elementos que integram esses fluxos são, na sua grande maioria, do sexo masculino. O caso específico de Santana, capital do Distrito de Cantagalo, no qual se insere Almojarife, é paradigmático desta situação. Os jovens rapazes da região, em número bastante mais elevado do que as raparigas, enfrentavam então dois cenários: por um lado a falta de oferta de trabalho, já que não mostravam predisposição para aceitar as duras condições de trabalho impostas nas roças a troco de míseros salários; por outro lado, a dificuldade em constituir família, dado o reduzido leque de elementos femininos livres para o efeito. A busca de melhores oportunidades no meio urbano afigura-se então como a única alternativa possível para se alcançarem índices mínimos de sobrevivência, no que terão sido acompanhados por chefes de família mais idosos, isto é, indivíduos da FE-2, que tentam também obter um maior desafogo para sustentar os elementos dos seus

agregados. As mulheres, por seu lado, ficam para trás, entregues às lides caseiras, prestando apoio aos filhos e lavrando pequenas parcelas de terra para colherem o parco sustento diário de que necessitam.

A grande maioria dos migrantes do sexo masculino vai trabalhar para a construção civil, como bem o demonstra o ritmo de crescimento de todo o tipo de construções na capital a que aludem os relatórios do exército colonial, mas os mais jovens espreitam também a oportunidade de servir neste, evidenciando a *“manifestação de um desejo tradicional do nativo de São Tomé, e sobretudo porque vêm na Tropa o único processo de angariarem meios de subsistência dignos...”* (CTI de São Tomé, 1968a:1). Os migrantes que vão encontrando trabalho na cidade acabam por entrar em contacto com a variante mais padronizada do português, situação que se aplica igualmente aos mancebos que são incorporados no exército, o qual *“tem continuado a subsidiar os estudos dos militares no Curso Nocturno do Liceu, dentro das suas possibilidades monetárias.”* (CTI de São Tomé, 1973:4). Como tal, estamos em crer que, a exemplo do que sucedeu em HEL-Ba, a aceitação do padrão mais normatizado por parte dos deslocados terá sucedido de forma livre, já que ocorreu também uma *“visão positiva acerca dos benefícios trazidos com o progresso, a melhoria nas condições de transporte, no atendimento médico e a possibilidade de integração cultural proporcionada pelos meios de comunicação de massa. Os mais jovens são os mais atingidos por esse sentimento, buscando de uma maneira mais ativa a sua integração no novo modus vivendi”*(Lucchesi, 2000a:286). Isto mesmo nos confirmam também os relatórios da CTI, quando especificam que as relações entre população e autoridades eram bastante sociáveis, muito por força das acções de captação da simpatia da população, levada a cabo por actividades do exército, entre as quais se destacavam as emissões radiofónicas para o efeito.

Servidos também por uma boa rede rodoviária e de transportes, os migrantes deslocam-se regularmente a Almojarife, a fim de reverem os familiares e trazer-lhes o provento angariado. Ocorre então a típica situação de que a variação, enquanto processo de *interlíngua*, se dá primeiramente a nível individual, ocorrendo, posteriormente, a formação do dialecto, que acontece ao nível da fala da comunidade e ao nível do indivíduo (Plag, 2008a:115). De facto, os elementos da comunidade, encarando o padrão

linguístico dos migrantes como sinónimo de progresso, ascensão social e bem-estar, recebem-no de forma cordial. Como consequência, o próprio santomense passa a ser estigmatizado, sendo substituído pelo PtgL1 adquirido defectivamente. Assim, e uma vez mais a exemplo do que sucedeu com HEL-Ba, a variação detectada na camada mais idosa da FE-1 acaba por reflectir um movimento em direcção ao padrão do PE, tanto do ponto de vista social como a nível da consciência de alguns falantes. Desta maneira, a mudança teve também como elementos motivadores a pressão linguística que a comunidade sofreu quer de cima para baixo quer de fora para dentro.

Contudo, a expansão da regra de concordância viria a ser precocemente cerceada, já que, com a independência do país, o sistema socioeconómico deste se desarticulou, atingindo, inevitavelmente, todo o tipo de infra-estruturas erigido anteriormente pelo governo colonial. Os jovens aprendentes, que adquiriram o PtgL1 defectivamente, viram dificultadas as condições de acesso ao ensino, que é também ministrado por docentes falantes de um padrão mais próximo do seu (o português da São Tomé) e, com tal, não estão reunidas condições para continuar a levar a cabo a expansão da regra de concordância. Simultaneamente, a deterioração da rede rodoviária e de transportes aumentou o isolamento das comunidades rurais e disso se ressentiu Almozarife, com as crianças a não poderem atingir o nível de *escolaridade* das gerações anteriores e a verem dificultado, cada vez mais, o contacto com formas cujo padrão se encontra mais próximo do PE. Como tal, acreditamos que não estará em curso um processo de mudança linguística na comunidade, encontrando-se esta, pelo contrário, num estágio de variação estável, já que não vemos serem criadas condições para, num futuro próximo, se alterar radicalmente o actual estado em que se encontra a estrutura socioeconómica do país. Lembremos ainda, a propósito da variação estável, que esta surge intimamente ligada à relação directamente proporcional entre escolarização e uso de formas de prestígio (Labov, 1972a, 1981; Guy, 1981a, 1986), estando a expansão das regras de concordância do PA condicionada quer pelas condições em que se encontra o actual sistema de ensino de São Tomé quer pelo estado de isolamento a que a comunidade de Almozarife se encontra votada. A análise das variáveis sociais *estadia* e *idade* virão dar mais força ao pressuposto aqui defendido, como se verá.

Comparando agora os nossos resultados com os de Scherre (1988:454), Lopes (2001:113), P. Andrade (2003:125), Jon-And (2008:4) e Baxter (2009:283), constata-se que apenas os dois últimos optaram por distribuir os seus informantes por faixas etárias idênticas às nossas. Observando os estratos etários constituídos por Scherre (1988), verifica-se que variam ligeiramente dos nossos. Ainda assim, é possível observar que o MRJ, ao contrário do PA, apresenta um padrão bastante curvilíneo, mas com a geração mais jovem a aplicar menos a concordância do que as gerações intermédias. Deste modo, a variável social *idade* constituiu um dos suportes que levou a autora a concordar parcialmente com Guy (1981a, 1986), quando este postulou que o PB se encontra em estágio de variação estável.

**Tabela 27.3.** *Idade*: contribuição dos factores individuais para a marcação PL do elemento analisado [6 variedades de português].

Faixas etárias	Idade							
	PA		MRJ		NURC		HEL-Ba (P. Andrade, 2003)	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%
7 aos 14 anos	-	-	Pr.	%	-	-	-	-
15 aos 25 anos	-	-	0,45	-	0,49	75	-	-
20/25 aos 35/40	0,658	63	0,51	-	0,44	81	0,66	47
40/45 aos 55/60	0,404	34	0,56	-	0,46	80	0,46	47
+60 anos	0,357	24	0,48	-	0,61	85	0,37	42
+ 80/85 anos	-	-	-	-	-	-	-	-

Faixas etárias	HEL-Ba (Baxter, 2009)		PT		PMp	
	Pr.	%	Pr.	%	Pr.	%
	7 aos 14 anos	Pr.	%	Pr.	%	Pr.
15 aos 25 anos	-	-	-	-	-	-
20/25 aos 35/40	-	-	-	-	-	-
40/45 aos 55/60	0,638	49	0,792	76	0,63	93
+60 anos	0,483	45	0,425	47	0,55	90
+ 80/85 anos	0,393	40	0,112	26	0,29	78

Por seu lado, os dados da NURC revelam flutuação na realização da concordância PL, que se assemelha à que detectámos na variável *escolaridade* do PA (Tabela 26.2).

Em Salvador, é a faixa etária mais idosa a única que favorece a concordância, acontecendo, depois, a ausência de uma sequência curvilínea ou de uma inclinação progressiva. Estes fenómenos revelam que não está a ocorrer um crescimento da aplicação da regra de concordância, pelo que também não se pode postular uma tendência de mudança na aplicação da concordância de número para a NURC. De facto, os resultados desta apontam para um estágio de variação estável, a exemplo do postulado por Guy (1981a, 1986) e Scherre (1988) para o MRJ.

Observando os resultados da variável social *idade* em função da variável *escolaridade*, Lopes (2001:115) conclui pela existência de histórias sociais distintas que incidiram sobre os diferentes grupos etários e que conduziram a uma situação de gramáticas em competição: (i) o grupo mais velho, detentor de escolaridade superior, a realizar quase sempre a concordância; (ii) o grupo com escolaridade mais baixa, com uma gramática que realiza bastante menos a concordância; (iii) o grupo das pessoas que detêm as duas gramáticas, e que realizam a concordância de forma intermediária, isto é, posicionando-se entre os dois extremos.

Seguidamente, Lopes (2001:119-120) compara dados sincrónicos da NURC recolhidos em duas épocas distintas: década de 70 e fala actual. Os resultados revelam a manutenção do mesmo índice de realização da concordância no grupo mais idoso (década de 70 = 0,67; fala actual = pr. 0,68), mas o grupo universitário apresenta uma perda nítida na aplicação da regra, descendo para valores que chegam a ser suavemente inibidores (década de 70 = 0,55; fala actual = pr. 0,47). Ainda assim, a autora mostra-se cuidadosa na sua conclusão, chamando a atenção para o facto de não estar a acontecer uma mudança, mas apenas uma variação etária ou geracional.

Os dados de HEL-Ba, analisados por P. Andrade (2003:125) e Baxter (2009:283) revelam que a fala da comunidade se encontra “*em processo de mudança em curso no sentido de adoptar por completo a concordância*” (P. Andrade, 2003:126), já que a frequência das formas inovadoras da fala é evidenciada pela faixa etária mais nova. A idêntica conclusão chegaria ainda Baxter (2009:283) para o PT. Estes pressupostos seriam confirmados por ambos os autores, que levaram em conta a relação estabelecida entre a variável *idade*, a sociohistória das duas comunidades e os fenómenos gramaticais evidenciados por elas. Note-se ainda que os padrões diacrónicos de marcação PL das



duas comunidades se assemelham não só entre si mas também ao do PA, evidenciando uma distribuição que se inclina em direcção à aquisição da regra. Contudo, as actuais realidades socioeconómicas do Brasil e de São Tomé são bem distintas, com o primeiro país conhecendo uma grande pujança económica que se reflecte, evidentemente, na melhoria das suas infra-estruturas, rodoviárias, comunicativas e de ensino. Contrariamente, São Tomé mergulhou num grave recessão socioeconómica depois de 1975, que desarticulou as estruturas de suporte de desenvolvimento do país. Nesta conformidade, enquanto no Brasil estão reunidas as condições para que os falantes de HEL-Ba continuem a expandir a regra da concordância de número, em Almojarife a situação é de estagnação, e mesmo algum retrocesso, como nos deixou antever a variável *escolaridade* e se confirmará com os dados dos restantes grupos de factores sociais, isto é, a *estadia* e o *sexo*. Ainda assim, é preciso ter também em conta que os falantes do PA e do PT apresentam sociohistórias distintas, já que os segundos viveram sempre uma situação de maior isolamento em relação aos primeiros, como se pode inferir da descrição apresentada sobre tongas e forros em relatórios da CTI: “*Torna-se interessante frisar que, enquanto o forro abandonou por completo o trabalho da roça, os descendentes dos tongas se mantiveram mais ou menos ligados à propriedade onde os seus avós serviram.*” (CTI, 1973:6).

Observando também a variável social *parentesco* para o PT, Baxter (2009:283-284) conclui que a faixa etária com mais de sessenta anos e parte da faixa com idades compreendidas entre os 41 e 60 anos são falantes nascidos de casais africanos ou casais mistos (africano + tonga). Como durante o “*seu processo de aquisição de L1, essas pessoas teriam acesso a um forte componente de português L2 falado por africanos, um português aprendido em situação de contacto, além do modelo de língua africana, neste caso o umbundo*” (Baxter, 2009:283), acabam por revelar um registo linguístico no qual a variação é mais acentuada. Por outro lado, os falantes com pais nascidos na roça Monte Café favorecem mais a concordância PL (pr. 0,63) do que os informantes cujos pais nasceram em África (pr. 0,36) (Baxter, 2009:284), confirmando-se que a presença africana exerce influência na não-marcação PL, efeito esse que o autor atribui ao PtgL2 influenciado por estruturas bantu.

Por seu lado, os mais jovens levam a cabo uma maior concordância PL no seus registos em virtude de diversos factores socioeconómicos ocorridos durante a segunda metade do século XX, e que contribuíram para um nivelamento linguístico entre o PT e o PA. De facto, a ligeira liberalização do sistema administrativo das roças permitiu que os serviçais e seus descendentes passassem a ter mais liberdade de movimento, com os segundos a poderem frequentar a escola. Como tal, as características linguísticas das gerações antepassadas foram desaparecendo do registo dos mais novos, que se aproximou mais do padrão do PE, em virtude das influências externas à comunidade. Posteriormente, com a desarticulação que afectou o sistema socioeconómico de São Tomé após a sua independência, as estruturas das roças entraram em total degradação e o seu sistema produtivo extinguiu-se. Os mais velhos ainda permanecem ligados à vida rural, dedicando-se a agricultura de subsistência, mas os mais novos migraram para os centros urbanos em busca de melhores condições de vida, passando a contactar com a variedade do português santomense, cujo sistema de concordância PL é mais próximo do PE. Nesta conformidade, enquanto o PA se encontra em estágio de variação estável, os jovens tongas, que só tardiamente abandonaram o sistema de semi-isolamento das roças, acabam por encontrar ainda uma margem de progressão na realização da concordância de número em direcção ao padrão de outra variedade de português presente na ilha, o português de São Tomé.

A linha inclinada do padrão diacrónico do PMp, com a faixa etária mais jovem a ser a única que favorece a concordância PL, ainda que moderadamente, revela que este dialecto se encontra em estágio de mudança em curso. Jon-And (2008:4) entende que o favorecimento à aplicação da regra de concordância por parte dos menos idosos tem a ver com o facto de o português dominar, cada vez mais, as situações de contacto quotidianas no Maputo, principalmente entre as gerações mais jovens. Relativamente ao PCV, Jon-And (2009) não analisou a marcação PL realizada pelas diferentes faixas etárias. No entanto, a autora constituiu a variável extralinguística *idade de início de aquisição de português (onset age)*, composta por três factores (Jon-And, 2009:4): (i) aquisição aos 0 anos; (ii) aquisição entre os 6 e 7 anos; (iii) aquisição entre os 10 e 14 anos. O estudo permite perceber que o factor (ii) constitui a idade ideal para aquisição da L2, já que beneficia a inserção de marcas (pr. 0,56), não sendo alheio a este

favorecimento o facto de, nesta fase, se iniciar a escolarização. Em plano ligeiramente inferior (pr. 0,43), surge o factor 0 anos, comprovando a facilidade da ASL na idade crítica da aquisição, mas que regista variação que se pode fixar, caso não receba *input* correctivo. Por fim, o factor início de aquisição de português entre os 10 e os 14 anos confirma que, muito dificilmente o aprendente suplanta os erros de concordância (pr. 0,35), dada a incapacidade de aquisição dos traços virtuais não especificados das categorias funcionais (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashes, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), visto estas deixarem de estar disponibilizadas depois do período crítico do desenvolvimento.

Relativamente à premissa que levantámos para a hipótese 13, constituída para a variável extralinguística idade (ponto 3.6.12), confirma-se que as diferentes gerações detêm gramáticas distintas, já que as gerações mais novas aplicam mais a regra de concordância. No caso específico do PA, e por condicionalismos conotados à desestruturação do sistema socioeconómico do país que ocorreu no período pós-independência, poder-se-á estar agora a assistir a uma situação de variação estável, com tendência para manter o actual padrão de CPL-var no SN.

### **5.3.3. Variável independente *estadia* (*Permanência fora da comunidade*)**

Como se pode verificar na Tabela 8.1, dos dezoito informantes de Almojarife, apenas cinco estiveram ausentes da comunidade por períodos superiores a seis meses. Este número salvaguarda a questão da “representatividade”, mas reflecte-se na quantidade de ocorrências submetidas a análise em cada um dos grupos (Tabela 12.14), com os informantes que se ausentaram a produzirem menos de metade de ocorrências (719 itens) do que o grupo de informantes cujas ausências foram nulas ou inferiores a seis meses (1.621 itens).

Tal como aconteceu para a variável independente *idade*, também a variável social *estadia* não necessitou de qualquer tratamento extra para os seus dados, até por que se trata de um grupo de factores binário. Na Tabela 28.1 é possível observar o peso que as produções dos dois grupos têm na configuração da CPL-var do SN do PA:

**Tabela 28.1.** Efeito da *estadia* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011	
<b>Factores seleccionados</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>	
Sim	358/719	50	31	0,620	
Não	844/1621	52	69	0,446	
<b>Totais:</b>	1.202/2.340	51	100	-	

No ponto 5.3.2, dedicado ao estudo da variável social *idade*, abordámos já as consequências para o padrão de concordância do PA advindas dos movimentos migratórios dos membros da comunidade, principalmente por parte dos informantes do sexo masculino. Deste modo, as variações na fala a nível comunitário terão mais a ver com situações de migração do que com ausências para fora do país, já que as últimas não implicaram um contacto directo com falantes do PE (Tabela 8.1). Ainda assim, os informantes de Almojarife que passaram períodos prolongados fora da comunidade tendem a favorecer moderadamente a concordância PL (pr. 0,620), enquanto os que não registaram ausências significativas evidenciam tendência para inibir a pluralização, mas em valores que se aproximam do patamar da neutralização (pr. 0,466). Note-se que, ao contrário de HEL-Ba, cujo efeito da concordância parece ser mais significativo entre os jovens, no PA essa incidência cairá sobre os informantes mais idosos, especialmente da FE-2 (41-60 anos). No contexto específico das comunidades rurais de São Tomé, tal não surpreende, por três razões óbvias: (i) a faixa etária mais idosa (FE-3, informantes com mais de 60 anos) viveu um período conturbado da vida da colónia, marcado por medidas segregativas e discriminatórias do Estado Novo, como o *Acto Colonial*, e que a empurraram para uma situação de precariedade, indigência e afastamento linguístico relativamente ao PE; (ii) por seu lado, a geração entre os 20 e os 40 anos (FE-1) foi vítima do desmembramento do sistema socioeconómico da ilha, que ocorreu após a independência desta e lançou as comunidades rurais para o isolamento e maior empobrecimento, em virtude da desarticulação de todo o tipo de infra-estruturas erigidas pela administração colonial; (iii) não detectámos no nosso *corpus* informantes desta faixa etária que se tenham ausentado para São Tomé durante o período em que ocorreu o fluxo de migração para a capital e que tenham, posteriormente, regressado para a comunidade com a finalidade de voltar a fixar residência nesta. Ainda assim, em termos

gerais, a ligeira tendência para favorecer a concordância evidenciada pelos elementos que passaram mais tempo fora da comunidade pouco ou nada contribuiu para aproximar o padrão de concordância PL do PA daquele que se regista no PE.

Resta referir que os estudos sobre as outras variedades de português não apresentaram valores para a variável *estadia*. P. Andrade (2003:123) refere que a mesma não foi seleccionada pelo programa VARBRUL na abordagem atomística, mas mesmo assim, e a exemplo dos outros estudiosos do fenómeno da CPL-var no SN, observa o modo como o grupo de factores *estadia* estabelece relações intrínsecas com a variável *idade*, para o desenho da concordância dentro do SN.

No que concerne à hipótese 14, constituída para a variável *estadia*, confirma-se que o factor ausência não contribuiu para a realização de padrões de concordância PL mais próximos dos do PE.

#### 5.3.4. Variável independente *sexo*

A exemplo do que sucedeu com as duas variáveis anteriores, também os dados da variável *sexo* não necessitaram de mais nenhum tratamento adicional, visto ser um grupo de factores binário. Na Tabela 29.1 é possível verificar qual o peso exercido por ambos os sexos na CPL-var do SN do PA:

**Tabela 29.1.** Efeito da variável social *sexo* na marcação PL dos itens do SN do PA.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009		<i>Significância:</i> 0,011
<b>Sexo</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Masculino	709/1.323	54	56	0,550
Feminino	493/1.017	49	44	0,436
<b>Totais:</b>	<b>1.202/2.340</b>	<b>51</b>	<b>100</b>	<b>-</b>

Como referimos no ponto 5.3.1, a discussão em torno da variável social *sexo* não tem sido consensual, já que o pressuposto de Chambers & Trudgill (1980:97-98), que atribuiu às mulheres a responsabilidade pelas mudanças em direcção ao padrão da LA, não é corroborado por outros estudiosos (Scherre, 1988:491; Lucchesi, 2000a:289; P. Andrade, 2003:129). No que diz respeito aos valores de Almojarife, são os homens que usam a forma mais próxima do padrão europeu (pr. 0,550), cabendo às mulheres o maior índice de variação (0,436). Este padrão de favorecimento estabelece semelhanças com o

de HEL-Ba e o do PT (Tabela 29.4), sendo rurais todas estas comunidades. Por outro lado, não deixa de ser curioso que as comunidades estudadas por Labov (1978), Chambers & Trudgill (1980), Scherre (1988) e Lopes (2001), nas quais as mulheres lideram o favorecimento da marcação, são todas do tipo urbano. Note-se ainda que os números apontados por Lucchesi (2000a:289), para a CGEN-var da comunidade de HEL-Ba, mostram que são também os homens os responsáveis pela produção das formas com maior índice de concordância (homens = pr. 0,55; mulheres = pr. 0,45).

Observando as marcações percentuais por sexo nas diferentes faixas etárias, a Tabela 29.2 apresenta-nos os resultados respeitantes ao estrato sexo que faz parte de cada um dos escalões etários:

**Tabela 29.2.** Marcação PL dos itens do SN do PA: valores por *idade e sexo*.

<i>Input desta rodada:</i> 0,504		<i>Log-likelihood:</i> -981,009	<i>Significância:</i> 0,011	
<b>Faixa etária</b>	<b>Sexo</b>	<b>Nr. total</b>	<b>% marcação PL</b>	<b>% no grupo</b>
FE-1 (20 – 40 Anos)	M	289/499	58	51
	F	334/485	69	49
FE-2 (41 – 60 Anos)	M	281/485	58	62
	F	92/301	31	38
FE-3 (+60 Anos)	M	139/339	41	59
	F	67/231	29	41
<b>Totais:</b>		1.202/2.340	51	100

Os homens lideram a tendência para inserir a pluralização nas duas faixas etárias mais idosas (FE-3: H = 41%, M = 29%; FE-2: H = 58%, M = 31%), mas depois acontece uma inversão no padrão da marcação por sexo no escalão mais jovem, no qual os membros do sexo feminino passam a aplicar mais a marcação do que os do sexo masculino (FE-1: H = 58%, M = 69%). Note-se ainda que, na FE-3, ambos os sexos marcam negativamente a pluralização, mas que na FE-2 os homens passam a revelar uma tendência para beneficiar a marcação, enquanto as mulheres, apesar de apresentarem uma maior aquisição da regra, ainda a inibem. Por fim, na FE-1, ambos os sexos passam a marcar positivamente a pluralização, sendo as mulheres quem mais propicia esta.

Quanto ao total de SN's com marcação plena produzidos pela comunidade de Almojarife, encontra-se repartido por ambos os sexos de acordo com a tabela seguinte 29.3:

**Tabela 29.3.** SN's plenamente marcados no PA: valores por *sexo*.

<i>Input desta rodada:</i> 0,223		<i>Log-likelihood:</i> -694.131		<i>Significância:</i> 0.001
<b>Sexo</b>	<b>Nr. Total</b>	<b>% marcação PL</b>	<b>% no grupo</b>	<b>Peso relativo</b>
Masculino	250/819	30	55	-
Feminino	175/669	26	45	-
<b>Totais:</b>	<b>425/1.488</b>	<b>29</b>	<b>100</b>	<b>-</b>

**Obs.:** Variável rejeitada pela ferramenta VARBRUL na rodada extra realizada (rodada geral do tipo não-atômico).

Os homens revelam maior tendência para aplicar a regra de concordância plena do que as mulheres, uma vez que produziram 819 SN's, com um percentual total de marcação de 30% (250 SN's plenamente marcados). O maior intercâmbio dos homens com o mundo exterior, preconizado através dos fluxos de migração para os centros urbanos, a integração de mancebos nas forças do exército ou o maior índice de escolarização dos elementos deste sexo são, por si só, condicionantes suficientes para entendermos o porquê de ser do sexo masculino a responsabilidade maior na aplicação da regra de concordância PL, já que cabe a este estrato o contacto mais próximo com as formas normatizadas. Por seu lado, as mulheres, mais fixadas à comunidade, foram responsáveis pela produção de 669 SN's, os quais apresentam um percentual de marcação que se queda pelos 26% (175 SN's com marcação em todos os itens da cadeia sintagmática, quando ela é exigida). Contudo, note-se que a pequena diferença nas frequências e percentagens de marcação PL podem indiciar estar a ocorrer uma tendência para a hegemonia dos padrões de comportamento destes dois segmentos sociais, daí a variável não ter sido considerada relevante pela ferramenta VARBRUL. Embora alguns autores entendam o fenómeno como indiciador de a comunidade se encontrar em estágio final de mudança linguística (cf. Baxter, Lucchesi & Guimarães, 1997:46), a verdade é que, no caso de Almojarife, tal sinal terá que ser entendido com alguma reserva. De facto, o actual panorama social de Almojarife, face ao desarticular das estruturas socioeconómicas do país após a sua independência, mostra-nos que os habitantes de São Tomé tiveram que recorrer a alternativas de sobrevivência, devido à

falta de oferta de emprego generalizado. Neste contexto, os elementos da comunidade de Almojarife passaram a dedicar-se a actividades novas, com os homens tirando partido da benesse píscea oferecida pela baía onde se encontram radicados e as mulheres ajudando-os no escoamento do produto capturado:

DOC: *Mas vai sempre a cidade?*

INF: *Ê vô sim, ê vô sempre sim.*

DOC: *Vai ao mercado?*

INF: *Ê vô pa mercado sim.*

DOC: *É que aqui tem uma tradição que quando o marido é pescador a mulher que vende peixe.*

INF: *Sim as veze ê vô vendê sim, as veze quando el trás peixe, ê vô ven... Quando el traz ê vô vendê, quando tem muito ê levo. [CLOTIM2]*

Deste modo, a actividade piscatória deixou de ser um exclusivo dos angolares, como as mulheres de Almojarife deixaram também de se dedicar unicamente à lavoura de pequenas parcelas de terra, actividade que as fixava na comunidade e as levava a conservarem características linguísticas mais primitivas determinadas pela transmissão linguística irregular. Assiste-se então ao inverter da situação que outrora levava os homens a protagonizarem o contacto com o mundo exterior, já que agora permanecem na comunidade após regressarem da faina do mar, repousando e preparando os aprestos marítimos para nova saída. Quanto às mulheres, deslocam-se para fora da comunidade para venderem o pescado, animais de criação e excedentes colhidos nas lavras e hortas, contactando com registos linguísticos que poderão, porventura, estar mais próximos do padrão do PE. Razão pela qual, na fase actual da vida da comunidade, se registará uma hegemonia nos padrões linguísticos de ambos os sexos, não a prefigurar um estágio de mudança consumada, mas uma aproximação ao registo masculino, que se encontrará em estágio de variação estável. Contudo, estes aspectos só poderão ser devidamente confirmados com a recolha e análise de dados linguísticos produzidos pelas actuais gerações mais novas de Almojarife, já que *“muitas mudanças discutidas sob este rótulo não estão provavelmente em progresso, mas em variação que pode permanecer estável durante séculos”* (Labov, 1981:177).<sup>191</sup> Este raciocínio é complementado por outros autores, quando procuram estabelecer uma hipótese geral sobre a mudança sintáctica, admitindo-se que qualquer nova estrutura *“pode ser parcialmente bloqueada por um certo período de tempo. É somente a longo prazo, não no seu início ou progresso, que*



*mudança lingüística produz resultados relativamente ‘naturais’ e uniformes”* (Naro, 1981:97).

Para podermos comparar o actual estágio de variação lingüística do PA com o de outras variedades de português, atente-se aos resultados apresentados na Tabela 29.4:

**Tabela 29.4.** Variável social *sexo*: contribuição dos factores individuais para a marcação plural do elemento analisado [5 variedades de português].

Sexo	Variável social <i>sexo</i>									
	PA		MRJ		NURC		HEL-Ba		PT	
	Pr.	%	Pr.	%	Pr.	%	Pr.	%	Pr.	%
Masculino	0,550	54	0,41	65	0,47	80	0,57	46	0,573	53
Feminino	0,436	49	0,59	77	0,53	82	0,43	44	0,401	58

O padrão da CPL-var do SN do PA apresenta similaridades com o de outras comunidades rurais, casos de HEL-Ba (P. Andrade, 2003:129) e do PT (Baxter, 2009:285), confirmando os seus valores o pressuposto que levantámos para a variável social *sexo* (hipótese 12). De facto, em todas estas são os homens que revelam um uso maior da regra de concordância em número, enquanto as mulheres apresentam um maior índice de variação. Estes valores encontram ainda paralelo naqueles que foram apresentados por Lucchesi (Lucchesi 2000a:289) para a CGEN-var de HEL-Ba. Contrariamente, nos dialectos urbanos do MRJ (Scherre, 1988:491) e de Salvador (Lopes, 2001:109), são as mulheres a liderar o uso das formas mais próximas do português padrão, cabendo aos homens apresentar o maior índice de variação. Estes aspectos fazem pressupor que os padrões de comportamento lingüístico em direcção à realização da regra de concordância do português são orientados por diferentes agentes, quando se trata de dialectos do tipo rural ou urbano. Para termos mais confirmações sobre este pressuposto, gostaríamos de possuir dados acerca do PMp, uma variedade urbana de português africano. Contudo, a variável social *sexo* não foi considerada significativa pelo programa VARBRUL nos estudos de Jon-And (2008), pelo que não podemos estabelecer comparações entre este dialecto e os constantes na Tabela 29.4.

No presente capítulo analisaram-se os resultados de uma variável do tipo semântico, três variáveis do tipo fonológico e quatro variáveis extralingüísticas. A variável do tipo semântico confirmou a hipótese levantada para a mesma. Por seu lado, as hipóteses constituídas para as variáveis fonológicas foram todas confirmadas apenas parcialmente.

Ainda assim, foi possível constatar que fenómenos como a assimilação regressiva (Braga, 1977:72) ou a haplologia sintáctica (Scherre, 1988:251-252) não evidenciam consistência para justificar a inibição da pluralização em termos de *contexto fonológico posterior*, sendo esta mais do tipo estrutural e condicionada pelas marcas precedentes. Este aspecto não só abona em favor da hipótese da CPL-var no SN ser influenciado pelo sistema de pluralização das ancestrais línguas africanas (cf. Guy, 1981a: 301-302) como retira também força à possibilidade de tal fenómeno ser motivado por derivas históricas e internas à própria língua (cf. Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a). Por fim, os resultados das variáveis sociais confirmam todas as hipóteses levantadas e revelam que a fala de Almojarife se encontra em situação de variação estável, com tendência para manter o actual padrão de CPL-var no SN.

Seguidamente, apresentamos as conclusões finais do presente estudo.



## CONCLUSÕES FINAIS

*“O desafio é interpretar a língua-E em relação à língua-I dos seus falantes.”*

(In Susan Pintzuck, George Soulas & Anthony Warner, *Diachronic Syntax: Models and Mechanisms*, p. 63).

Durante o desenrolar do presente trabalho procurámos determinar o quadro da CPL-var na estrutura do SN do PtgL1 adquirido por falantes bilingues, que têm um CP (o santomense) como outra L1. Paralelamente, procurou-se perceber como ocorre a aquisição massiva do PtgL2 por falantes monolingues de L1's africanas do grupo níger-congo atlântico, que têm o PE como LA. O objectivo global é perceber as tendências do processo de aquisição em contextos de transmissão linguística irregular, quer de L2's por falantes monolingues quer de L1's por falantes bilingues. Recorrendo à comparação com os perfis de variação registados no SN de diversas variedades africanas e brasileiras de português, procurámos confirmar também se o português adquirido em situação de contacto por falantes de um CP manifesta padrões de variação semelhantes aos do português resultante de uma situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Observados estes aspectos, seria possível concluir se as variedades parcialmente reestruturadas em situação de contacto terão sido desenvolvidas por influência de uma língua crioula (ou por línguas crioulas), ancorando a CPL-var do SN em mecanismos sintácticos dos substratos (Guy, 1981a:3001-302), ou, em contrapartida, se ficaram a dever a processos internos à própria língua portuguesa, independentes de tal influência, como a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a).

Para levarmos a cabo estas observações, foram analisados dados produzidos por 18 informantes bilingues em santomense e português, pertencentes à comunidade crioula de Almojarife, São Tomé. Os informantes foram equitativamente repartidos por três faixas etárias, de ambos os sexos. A nossa análise partiu do pressuposto que a variação observável na fala almoxarifana é condicionada por factores linguísticos e extralinguísticos e que, para além disso, reflecte características tipológicas, actuais e ancestrais, da L1 dos aprendentes, bem como a existência de características universais, independentes da L1 e da LA, que actuam sobre a estrutura da L2 e conseqüente L1 nativizada. O estudo apoiou-se no modelo variacionista laboviano (Bayley 1994; Bayley & Preston, 1996; Labov 1972a, 1982) e recorreu, essencialmente, ao pacote estatístico

GOLDVARB X (Sankoff, Tagliamonte & Smith, 2005), a versão recente do programa VARBRUL (Pintzuk, 1988), com apoio de outras versões do programa, como o GOLDVARB 2001 (Robinson, Lawrence & Tagliamonte, 2001) ou o TSORT, do pacote VARBRUL-2. Para tanto, construímos um modelo teórico seguindo a proposta dos estudos de Guy (1981a), Scherre (1988), Lopes (2001), P. Andrade (2003), Baxter (2004, 2009), Godinho (2005) e Jon-And (2008, 2009). O recurso aos pacotes e programas referidos permitiu não só que quantificássemos os dados mas estimássemos também o peso probabilístico das variáveis linguísticas e sociais na aplicação das regras de concordância de plural no SN do PA. Por seu lado, a metodologia facultou-nos a possibilidade de se observar o modo como os mecanismos morfossintáticos da concordância de número se vão incorporando, geracionalmente, no sistema linguístico dos falantes de Almojarife. Assim sendo, perspectivou-se que a aquisição dos mecanismos morfossintáticos da concordância de PL estivesse relacionada com onze variáveis linguísticas e quatro variáveis sociais (Tabela 10.1).

A análise das variáveis tidas como relevantes na aplicação da concordância de PL nos elementos do SN assentou não só nos estudos que permitiram elaborar o modelo teórico mas também em diversos trabalhos psicolinguísticos e generativistas sobre aquisição de L1 e L2 em contextos monolíngue e/ou bilingue, fossilizações e aquisição dos traços de número nos elementos do SN do PE e de variedades africanas e brasileiras de português. De entre os estudos que serviram de suporte para as referidas análise e interpretação dos resultados poder-se-ão apontar, entre outros, os de White (1989, 2003), Cerqueira (1993, 1994), Hawkins (1993, 1998, 2001), Bruhn de Garavito (1994, 2003, 2005), Koehn (1994), Meisel (1994a), Müller (1994a), Hawkins & Chan (1997), Bruhn de Garavito & White (2002), Franceschina (2002, 2003, 2005), Costa & Santos (2004 [2003]), Montrul (2004, 2006, 2008), White *et alii* (2004), Côrrea & Augusto (2005), Godinho (2005), Bruhn de Garavito & Atoche (2006), Castro & Pratas (2006), Costa & Silva (2006b), Castro (2007), Castro & Ferrari-Neto (2007) e Silva (2010), trabalhos esses que poderão ter em conta quer os princípios preconizados na Gramática Generativa (Chomsky & Lasnik, 1999 [1993], Chomsky, 1996 [1995]; Radford, 1997; Carnie, 2007 [2002]), quer modelos psicolinguísticos como o dos Quatro Morfemas (Myers-Scotton & Jake, 2000a, 2000b; Myers-Scotton, 2002) quer os pressupostos da MD (Halle & Marantz, 1993;

Noyer, 2006). Para além destes estudos, não deixámos também de considerar aspectos que se relacionam com a mudança evolucionista da língua (Givón, 1998; Croft, 2000; Mufwene, 2001; Clements, 2009) ou a incrementação dos processamentos linguísticos em interlínguas (Pienemann, 1988, 2005b). A observação dos resultados demonstrou que nem todas as variáveis constituídas exercem influência sobre a concordância. Com efeito, a variável independente *marcação de género* não se apresenta como elemento preponderante para a inibição ou favorecimento da inserção de marcas de PL, em virtude de, em algumas situações, ocorrer não-realização do mecanismo morfossintáctico de função meramente gramatical para indicar o género no PA, que passa a ser do tipo referencial. Por seu lado, a variável *ordem do item na cadeia dos constituintes flexionáveis em número* exerce um efeito oposto ao que prevíamos de início. Paralelamente, foi ainda possível constatar que, das variáveis linguísticas, o peso mais determinante na configuração da CPL-var no SN do PA cabe às variáveis do tipo estrutural. Em relação ao efeito da variável semântica *animacidade*, embora não seja de desconsiderar o seu peso na marcação PL de SN's com numerais, tem menos relevância do que as anteriores. Quanto às variáveis do tipo fonológico, exceptuando a variável independente *saliência fónica*, o seu efeito é nitidamente neutro para o desenho da CPL-var no SN. Portanto, a consistência de pressupostos que pretendem atribuir a responsabilidade da CPL-var no SN a fenómenos internos à própria língua portuguesa, como a assimilação regressiva (Braga, 1977), a haplologia sintáctica (Scherre, 1978, 1988) ou a própria deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a), é, desde logo, questionada pela escala hierárquica de favorecimento à inserção de marcas PL dos nossos achados.

Das hipóteses estipuladas para as variáveis linguísticas, confirmam-se totalmente as seguintes:

**Hipótese 5** – *Classe gramatical*: os itens que configuram a categoria de determinantes tendem a receber mais marcas de PL, uma vez que ocorrem, por norma, antes do nome núcleo do SN. As outras classes gramaticais inibem a marcação, visto configurarem categorias que ocorrem mais em posição nuclear (classes substantivadas e substantivos) ou pós-nuclear (adjectivos dos tipos 1 e 2).

**Hipótese 6** – *Posição do item analisado em relação ao núcleo do SN*: os elementos antepostos ao núcleo do SN, com especial incidência para os itens adjacentes, são mais marcados do que aqueles que lhe são pospostos.

**Hipótese 7** – *Posição do item na cadeia do SN*: o primeiro lugar da cadeia sintagmática, com especial relevância para a posição imediatamente pré-nuclear, é o que mais favorece a inserção de marcas. As outras posições tendem a marcar menos a pluralização, em virtude de a informação de número PL já ter sido atribuída (Kiparsky, 1972:195).

**Hipótese 10** – *Grau de concordância de número no SN*: a grande maioria dos SN's produzidos pela faixa etária mais idosa de Almojarife é de estrutura mínima (um ou dois itens flexionáveis) e baixo grau de complexificação estrutural. Quanto às faixas etárias mais novas da comunidade, não só produzem cadeias sintagmáticas mais extensas e de maior complexificação estrutural como também aplicam mais a regra da concordância em número.

**Hipótese 11** – *Traço semântico do SN*: os nomes com traço [+humano] têm tendência a favorecerem a marcação, enquanto os substantivos com traço [-humano] e [-animado] a desfavorecem. Entre os nomes [-humanos], os que contêm traço [-animado] propiciam mais a inserção de marcas do que os que possuem traço [+animado].

Confirmaram-se parcialmente as seguintes hipóteses:

**Hipótese 1** – *Saliência fônica (Processos de formação de plural)*: os itens que apresentam maior distinção fônica não tendem, necessariamente, a favorecer mais a inserção da pluralização, já que, entre os elementos que pluralizam regularmente, a inibição é mais acentuada nos que apresentam maior distinção fônica, isto é, nas palavras que terminam em *-ão* no singular. Por seu lado, os elementos terminados em *-s* no singular favorecerem categoricamente a marcação no PA.

**Hipótese 2** – *Tonicidade*: a inserção da marcação PL incide preferencialmente sobre os oxítonos e os monossílabos de uso tônico. Contudo, a hierarquia da inserção da

pluralização nos itens que menos beneficiam a marcação não seguiu a linha inicialmente prevista, já que os monossílabos átonos propiciam mais a inserção de marcas do que os paroxítonos. Para os proparoxítonos não foi possível obter a confirmação do peso exercido na pluralização, devido ao seu exíguo número de realizações nos nossos dados.

**Hipótese 3** – *Marcas precedentes*: apenas se confirma que a marca de PL formal em quantificadores e determinantes inibe a marcação nos itens seguintes, já que os numerais revelam tendência para favorecerem a inserção de marcas. Esta tendência decresce geracionalmente, a fim de estabilizar o padrão da regra ao encontro do padrão estrutural de marcação.

**Hipótese 4** – *Contexto fonológico posterior*: a pausa final é o elemento que mais favorece a inserção de marcas, aspecto que apenas não foi corroborado aquando da observação da influência na marcação PL exercida pelas zonas e pontos de articulação das consoantes pela FE-1. A pausa interna, por seu lado, apresenta-se como o segundo elemento que mais favorece a inserção PL, apesar de exibir quer um valor próximo da neutralidade quer algumas oscilações geracionais. Ainda em termos da predição avançada, confirma-se a influência das consoantes na marcação PL, sendo a inserção de marcas mais facilitada pelas surdas do que pelas sonoras. Por seu lado, a não obstrução à passagem de ar pelo tracto bucal favorece a inibição, o que nos permitiu estabelecer a seguinte escala hierárquica das consoantes no favorecimento à marcação, de acordo com as suas zonas e pontos de articulação: (i) palatais, com especial aclave para as velares e declive para as apicais; (ii) labiais; (iii) dentais, com aclave para as lábio-dentais e declive para as ápico-dentais; (iv) alveolares. Relativamente às vogais, não evidenciam um favorecimento à marcação mais acentuado do que a totalidade das consoantes, contrariando o que havíamos predito inicialmente.

Não se confirmou a seguinte hipótese:

**Hipótese 9** – *Ordem do item na cadeia dos constituintes flexionáveis em número*: não se detecta a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208), com zeros levando a zeros no SN do PA, já que a ausência de marcação no elemento



posterior ao item não flexionável que é inserido na cadeia sintagmática radica quer em motivações estruturais quer em motivações semânticas. Estas, por seu lado, determinam que a pluralização nos elementos posteriores, por questões funcionalistas, não venha a ser inserida, em virtude de se revelar redundante (Kiparsky, 1972:195).

No que diz respeito às variáveis sociais, confirmaram-se totalmente as hipóteses levantadas:

**Hipótese 12** – *Sexo*: por condicionantes sociolinguísticas, os homens usam mais a regra de concordância em número, enquanto as mulheres apresentam um maior índice de variação. Contudo, o nível de propiciamento da pluralização no SN tem vindo a nivelar-se entre ambos.

**Hipótese 13** – *Idade*: as diferentes gerações de Almojarife detêm gramáticas distintas, com as gerações mais novas a aplicarem mais a regra de concordância PL do que as gerações mais idosas.

**Hipótese 14** – *Estadia*: os membros da comunidade que se ausentaram desta durante prolongados períodos de tempo não contribuíram para que o padrão de concordância PL do PA se aproximasse mais do padrão do PE.

**Hipótese 15** – *Escolaridade*: a atenção dos falantes para com os mecanismos de concordância aumenta com o incremento da frequência escolar, levando-os a flexionar mais os itens pluralizáveis e a produzir SN's de estrutura mais complexa.

No que concerne às variáveis do tipo estrutural, para percebermos o modo como os elementos da cadeia sintagmática se comportam nas diferentes posições desta, observámos os efeitos exercidos na marcação PL pelos grupo de factores *posição em relação ao núcleo do SN, posição linear, posição na cadeia dos constituintes flexionáveis e categoria gramatical*, quer de forma individual quer estabelecendo cruzamentos diversos entre eles. No geral, constatámos a sua forte influência na configuração da CPL-var no SN do PA, com as posições anteriores ao núcleo do SN tendendo a ser as mais marcadas, sobretudo a posição adjacente a este. A mais alta

inibição da marcação, por seu lado, afecta categoricamente os elementos pospostos ao núcleo.

Os cruzamentos entre variáveis permitiram também confirmar que os substantivos e categorias substantivadas revelam um peso inibidor da marcação em posição nuclear. Contudo, esta inibição perde força quando o núcleo ocupa o primeiro lugar da cadeia sintagmática, já que passa a favorecer moderadamente a inserção de marcas por questões funcionalistas, que levam a pluralização a ser inserida na primeira oportunidade (Kiparsky, 1972:195). Relativamente aos artigos, distribuem-se pelas suas clássicas posições pré-nucleares, com ocorrências e peso da pluralização mais acentuados na localização imediatamente antes do núcleo. Quanto aos numerais, situam-se apenas nas posições pré-nucleares, com um peso neutro na pluralização, mas com os numerais que terminam em *-s* a propiciarem menos a pluralização do que aqueles que não finalizam em *-s*. Desta forma, estabelece-se um paralelo entre estes elementos e os itens portadores ou não de marcação formal de PL, uma vez que a análise do grupo de factores *marcas precedentes* permitiu observar que o factor S\_ (2<sup>a</sup> posição, marca formal na 1<sup>a</sup> posição) regista uma acentuada tendência para inibir a pluralização nos itens seguintes, enquanto o factor 0\_ (2<sup>a</sup> posição, ausência de marcação formal na 1<sup>a</sup> posição) conduz a um favorecimento categórico da marcação PL nos elementos que se seguem. Estes aspectos contrariam a hipótese do Princípio do Processamento Paralelo, confirmando que este não tem intervenção na marcação PL dos elementos do SN do PA, conforme preconizado por Scherre (1988:208) para a fala do MRJ. Observando a questão da marcação determinada pelos numerais nas diferentes faixas etárias, foi possível constatar que, inicialmente, os falantes almorixanos não percebem que estes itens carregam intrinsecamente a noção semântica de PL. Contudo, as faixas etárias mais novas de Almorixe, e conseqüentemente mais escolarizadas, adquirem essa noção e estabelecem um nivelamento entre os dois padrões de marcação, semântico e estrutural: (i) se ocorrem marcas semânticas ou estruturais de PL no primeiro elemento do SN, a tendência será de não inseri-las nos elementos posteriores (ii) se tais marcas não ocorrerem, a tendência será para propiciar a marcação nos itens posteriores. Ainda assim, importa referir que as diferenças de valores reveladas quer pelos pesos relativos quer pelos percentuais de marcação entre ambos os tipos de numerais são praticamente

nulas, validando também a eliminação do contraste entre ambos. Por outro lado, foi também possível constatar que os informantes de Almojarife tratam de forma idêntica os numerais simples e numerais de mais de uma palavra.

No que diz respeito aos restantes elementos gramaticais, espalham-se por todas as posições pré e pós-nucleares, mas com muito maior incidência no lugar imediatamente à esquerda do núcleo, onde o seu peso de inserção de marcas PL é bastante categórico. De entre estes, os possessivos e adjectivos em posição imediatamente pré-nuclear evidenciam alto índice de favorecimento à marcação, apresentando os segundos um desfavorecimento quando em posição pós-nuclear, onde ocorrem maioritariamente. Estes aspectos confirmam os pressupostos da MD que advogam existir uma relação não só de definitude entre ambos os itens gramaticais e o nome mas também entre o *output* da sintaxe e o *input* da morfologia, com a segunda lendo a primeira. Como tal, a concordância PL é funcional e determinada no PA por morfemas do tipo *singleton*, responsáveis pela variação, em virtude de deterem o estatuto de núcleo na posição imediatamente pré-nominal, onde incide a marcação de número, e assumirem a categoria de sintagma na localização pós-nominal, na qual a marcação PL é dispensada por questões funcionalistas, que visam evitar a redundância (Kiparsky, 1972:195). Este aspecto foi especificamente confirmado no cruzamento entre as variáveis *ordem do item na cadeia dos constituintes flexionáveis* e *posição linear*, que permitiu concluir, também, que o Princípio do Processamento Paralelo, com zeros levando a zeros e marcas conduzindo a marcas (Scherre, 1988:208), não tem intervenção na marcação PL do SN do PA, já que a ausência de pluralização no elemento posterior ao item não flexionável que é inserido na cadeia sintagmática é, essencialmente, motivada por razões semântico-estruturais. Hierarquicamente, os determinantes pré-nucleares propiciam a inserção de marcas de acordo com a seguinte escala decrescente de favorecimento: indefinidos, demonstrativos, possessivos, quantificadores, artigos definidos e artigos indefinidos.

Relativamente à variável *marcas precedentes*, abstivemo-nos de a observar efectuando cruzamentos com as outras variáveis do tipo estrutural, visto entendermos que ela espelha melhores resultados se for tratada individualmente (cf. Guy, 1981a:178). Este aspecto seria confirmado ao efectuarmos amalgamações de factores para constituir

grupos de factores de acordo com os elaborados por outros autores (Scherre, 1988; Lopes, 2001; P. Andrade, 2003; Baxter, 2004, 2009) e observarmos o peso dos mesmos na inserção de marcas PL. De facto, a metodologia das amalgamações e cruzamento dos seus grupos de factores com outras variáveis independentes do tipo estrutural não permitem, quanto a nós, que se observem, com exactidão, questões inerentes à problemática da implementação (*actuation problem*) (Labov, 1972a, 1982; Weinreich, Labov & Herzog, 2006 [1968]), isto é, se determine o peso que certos factores detêm na pluralização dos elementos do SN, já que esses factores poderão estar a reflectir meras realizações idiolectais, que acabam viciando resultados que se pretendem comunitários. Relativamente ao estudo comparado do modo como o grupo de factores *marcas precedentes* intervém na marcação PL do item posterior de diferentes variedades parcialmente reestruturadas de português, o mesmo permitiu verificar não só que o PA é a variedade que se encontra no estágio mais baixo da aquisição da regra de concordância PL mas também que o português adquirido em situação de contacto por falantes de um CP manifesta os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Esta conclusão, por sua vez, reforça a probabilidade de, em todas as variedades africanas e brasileiras de português observadas, a CPL-var ser fortemente condicionada pela indicação do PL na primeira oportunidade (cf. Kiparsky, 1972:195), por norma o item colocado na posição imediatamente anterior ao núcleo do SN, devido a influências que ancoram nas línguas do substrato africano (cf. Guy, 1981a:301-302), ancestrais ou directas. Descarta-se assim, uma vez mais, a probabilidade de a CPL-var ser motivada por questões determinadas pelo Processamento Paralelo (Scherre, 1988:208), com zeros originando zeros e marcas levando a marcas.

No que concerne à variável *posição linear*, o seu cruzamento com a variável social *idade* foi determinante para se observar não só a forma como são inseridas as marcas de PL mas também compreender como se desenvolve, geracionalmente, a regra de concordância no SN. Confirmou-se o que já se observara acerca de a marcação PL ser do tipo funcional e determinada, no PA, por morfemas do tipo *singleton*. Deste modo, a pluralização é atribuída de modo parcial, desde o início da aquisição do PtgL2, ao elemento pré-nuclear adjacente, ou seja, à posição DET (Longobardi, 1994; Baxter,

2009), mesmo em sintagmas de configuração mais desenvolvida. A partir desta localização, o PL desenvolve-se, ainda com características de concordância parcial, através de morfemas dissociados, que são atribuídos aos elementos mais próximos do *fulcro*. A ordem de atribuição contempla os elementos do SDET, primeiro, e os itens à direita do núcleo, depois (cf. Fig. 1). Não obstante, o perfil da âncora imediatamente pré-nuclear é mantido geracionalmente. Estes aspectos foram determinados pelo facto de a ancestral ASL ter tido, como intervenientes, falantes de L1's africanas do grupo níger-congo atlântico que haviam já ultrapassado o período crítico para a aquisição, não podendo, por tal, adquirir os traços de PL das categorias funcionais (cf. Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003). Estes falantes foram responsáveis, depois, pela transmissão de DLP's deficitários que determinaram uma instanciação em cadeia de novos/errados parâmetros, os quais, por sua vez, impossibilitaram um reajustamento paramétrico, já que os novos/errados parâmetros foram tomados como modelos para uma nova reinstanciação paramétrica na L1 nativizada (cf. Gonçalves, 2004:235). Perdido os traços virtuais não especificados das categorias funcionais, as faixas etárias mais novas de Almojarife, sobretudo as menos escolarizadas, não conseguem suplantar a variação fixada, limitando-se a aplicar a pluralização redundante de modo esporádico, e apenas nos elementos pré-nucleares não imediatos, quando começam a elaborar SN's mais complexos. Ainda assim, vislumbra-se uma ligeira aquisição da regra de concordância, sobretudo por parte dos elementos mais escolarizados da comunidade, um aspecto que contraria o pressuposto de o apagamento da marca formal de PL poder ser justificado em termos de deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 1993, 2000, 2007a), já que esta determina a eliminação diacrónica do *-s* morfémico PL, ao passo que, o que sucede no PA é um situação diametralmente oposta, verificando-se a aquisição/inserção geracional do referido morfema.

Levando em conta a problemática da transição (*transition problem*) (Labov, 1972a, 1982; Weinreich, Labov & Herzog, 2006 [1968]), a fala de Almojarife fornece evidências para se concluir que, em termos comunitários, e de acordo com hierarquia da “processabilidade” (Plag, 2008a, 2008b), não foram adquiridas plenamente as regras da flexão contextual, instanciadas a partir do nível 4 do desenvolvimento da interlíngua

(Tabela 3.1). Como tal, a variação no PA encontrar-se-á entre os patamares 3 e 4 do desenvolvimento sintáctico proposto pela referida hierarquia, fase em que se deverá manter estabilizada, porque não se encontram reunidas condições socioeconómicas em São Tomé que façam prever um continuado desenvolvimento da aquisição da regra da concordância PL, que a leve a aproximar-se mais do padrão do PE.

As constatações fornecidas pelas variáveis do género estrutural acerca das motivações para a inserção de marcas PL no SN do PA, seu ponto inicial de aplicação e posterior desenvolvimento geracional da regra de concordância PL, foram fortalecidas e complementadas com a observação do comportamento das variáveis do tipo fonológico. De entre estas, apenas a variável *saliência fónica* apresenta algum peso na inserção da pluralização nos itens do SN, embora o Princípio da Saliência Fónica (Naro & Lemle, 1976:240-241; Lemle & Naro, 1977:47) não pareça reflectir-se no PA, já que o maior grau de concordância não incide sobre os itens que implicam uma saliência fónica mais perceptível na oposição singular/plural. Efectivamente, a tendência para o favorecimento da marcação é bastante acentuada apenas nos itens que finalizam em *-s* no singular, cujo PL se faz através da inserção de *-es*, ou seja, em elementos que evidenciam um grau ligeiro de distinção fónica singular/plural. Por outro lado, a inibição também se verifica mais nos elementos finalizados em *-ão* do que nos itens regulares, isto é, nos que apresentam maior distinção fónica singular/plural, talvez por apresentarem diferentes formas de pluralização (*-ão* > *-ãos*, *-ão* > *-ões*, *-ão* > *-ães*), o que leva os falantes a hesitarem na selecção daquela que será mais apropriada (Scherre, 1988:124). Lembremos aqui que Faria, Freitas & Miguel (2001:53) apontam também a irrelevância da saliência fónica na inserção de fricativas morfológicas em coda de sílaba, que determinam a realização da pluralização na fase precoce de aquisição do PEL1. Tal resulta da incapacidade de as crianças isolarem e interpretarem separadamente os morfemas a nível das palavras, um fenómeno apontado também na aquisição bilingue (Koehn, 1994:49) e ASL (DeGraff, 1999b:482) e que, ao que tudo indica, ocorre no PA.

Comparando os dados do PA com os de outras variedades africanas e brasileiras de português, verifica-se que o padrão de marcação do primeiro apresenta ligeiras diferenças relativamente ao de algumas das segundas, ficando tal a dever-se ao modo como alguns dos factores foram amalgamados por outros autores, o que originou

factores distintos dos constituídos por nós. Contudo, o factor *itens que finalizam em –s no singular* evidencia um comportamento diametralmente oposto ao das outras variedades, uma vez que inibe a marcação PL nestas. A inibição tem sido justificada com base na analogia que o /s/ final estabelece com o morfema –s, marca de PL, originando o fenómeno da assimilação regressiva (Braga, 1977:72).

O facto de ocorrer distanciamento na forma de marcar o PL em alguns itens produzidos por diferentes falantes de Almojarife, levou-nos a observar a relação estabelecida entre a variável linguística *saliência fónica* e a variável social *escolaridade*. Confirmou-se que quer o padrão de favorecimento da marcação (itens que terminam em –s no singular) quer o padrão de inibição (itens regulares) se mantêm constantes em todos os níveis de escolaridade, ou seja, em todas as gerações.

A variável *tonicidade* mostrou que os oxítonos e os monossílabos tónicos detêm a primazia na marcação PL, enquanto os paroxítonos, os proparoxítonos e os monossílabos átonos a inibem. Neste aspecto, o padrão de marcação do PA é idêntico ao de outras variedades de português, mantendo-se praticamente constante em todos os níveis de escolarização e faixas etárias. Esta particularidade evidencia que os falantes almojarifanos não possuem diferentes concordâncias motivadas pela mesma variável fonológica, o que seria confirmado posteriormente, com recurso ao cruzamento entre as variáveis *saliência* (amalgamação entre factores das variáveis linguísticas *saliência fónica* e *tonicidade*) e *escolaridade*. A constituição deste grupo de factores permitiu verificar também que a escala de favorecimento à marcação é liderada pelos oxítonos finalizados em –s no singular, seguindo-se os monossílabos átonos, os itens terminados em –r no singular e os oxítonos ou monossílabos tónicos. Como itens inibidores da marcação, aparecem os paroxítonos regulares e os nasais regulares. A confirmação do peso exercido pelos elementos terminados em –s no singular na inserção de marcas PL denota, uma vez mais, que o fenómeno da assimilação regressiva (Braga, 1977:72) não é acolhido no PA. Quanto ao elevado índice de marcação dos monossílabos átonos, deixa perceber que, na aquisição do PtgL2, os proclíticos dos substantivos terão sido interpretados como sílabas iniciais destes, ocorrendo não-aquisição quer da sua noção de definitude quer da informação PL do tipo sintáctico, tal como acontece nos CP's atlânticos e seus substratos do grupo níger-congo atlântico. Por seu lado, o perfil de

marcação PL no PA, motivado tanto pela *tonicidade* como pelo eixo *saliência*, é bastante semelhante aos de outras variedades africanas e brasileiras de português, confirmando, de novo, que o português adquirido em situação de contacto por falantes de um CP manifesta os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a). Paralelamente, verifica-se que o efeito da *saliência* na marcação PL do PA, ao longo do seu percurso geracional, ancora nos achados dos estudos de aquisição que defendem que o desenvolvimento da morfologia flexional se inicia nas oposições salientes (Bayley, 1994:167). Portanto, as oposições mais salientes são adquiridas precocemente, sendo a marcação favorecida pelas distinções morfofonológicas plural/singular mais acentuadas, como acontece com os oxítonos terminados em *-s* no singular, os plurais tónicos com mudança no radical e os plurais tónicos sem mudança no radical. Paralelamente, a geração mais idosa de Almojarife revela tendência a não pluralizar os elementos terminados em *-r* no singular, em virtude de não fazerem a distinção entre elementos vocálicos em final de palavra e o som [r], inexistente em línguas do grupo níger-congo atlântico, e sujeito a fenómeno de lambdacismo no santomense. Desta forma, a aquisição do PL das formas menos salientes e dos elementos terminados em *-r* no singular ocorre em período mais tardio, como nos revelam os valores dos elementos terminados em *-r* no singular e dos monossílabos átonos, que devem o favorecimento da pluralização à aquisição do artigo nas FE-2 e FE-1, e dos plurais átonos sem mudança no radical, que, geracionalmente, nunca atingem números que apontem para um favorecimento à inserção de marcas de PL.

A variável *contexto fonológico posterior*, apesar da sua fraca influência na marcação PL, permitiu não só complementar conclusões acerca do modo com as variáveis do tipo fonológico intervêm nesta mas também confirmar que o fenómeno da haplologia sintáctica (Scherre, 1988:244) não intervêm na inibição do PL no SN do PA. Os falantes deste tendem a propiciar mais a pluralização quando o *contexto fonológico seguinte* é um item vocálico do que quando é configurado por um elemento consonantal, embora ambos revelem pesos inibidores para a inserção de marcas PL. Ainda assim, verifica-se que as realizações da comunidade de Almojarife tendem para a busca do padrão silábico



universal CV. Observando o comportamento dos traços de fonte e de classe principal das consoantes, verificou-se também que o vozeamento e a sonoridade pouca peso exercem na pluralização, apesar de as consoantes não-soantes [–sonoras] e as surdas [–vozeadas][–sonoras] evidenciarem tendência para propiciarem mais a marcação do que as consoantes soantes [+sonoras] e sonoras ou vozeadas [+vozeadas][+sonoras]. Quanto às pausas, são o factor que mais beneficia a pluralização, com os valores da pausa final sobrepondo-se aos da pausa interna.

Analisando a influência das zonas, pontos e modo de articulação das consoantes na marcação PL, constatou-se que os itens não-coronais favorecem ligeiramente a marcação, enquanto os elementos coronais a inibem. Por outro lado, verificou-se também que quanto maior é a obstrução à passagem do ar no tracto bucal, mais acentuado é o propiciamento à inserção PL, o que permitiu confirmar a escala hierárquica de favorecimento à marcação evidenciado pelas consoantes: (i) palatais, com especial aclave para as velares e declive para as apicais; (ii) labiais; (iii) dentais, com aclave para as lábio-dentais e declive para as ápico-dentais; (iv) alveolares. Como as vogais são acusticamente realizadas sem intervenção do ponto apical, vão favorecer a inserção de marcas, apesar de serem produzidas sem obstrução à passagem de ar pelo tracto bucal. Quanto a intervenção da caixa de ressonância, é débil no processo de pluralização, tal como havia sido notado por Scherre (1988:255) para a fala carioca.

O cruzamento entre as variáveis *contexto fonológico posterior* e *saliência fónica* confirmou que as formas mais salientes, logo mais perceptíveis, detêm maior probabilidade de serem marcadas do que as formas menos salientes. Assim, o *contexto fonológico seguinte* tem tendência a propiciar menos a inserção PL em itens regulares, aumentando esse propiciamento quando se trata de realizações nasais regulares ou elementos terminados em *–s* e *–r*, com estes últimos evidenciando marcação categórica do PL, independentemente do tipo de realização fonológica posterior. Este aspecto confirmou, em definitivo, que os fenómenos da assimilação regressiva e da haplogia sintáctica não exercem peso na inibição da pluralização, conforme defendido por Braga (1977:72) e Scherre (1988:244), suspeitando-se que estes pressupostos tenham sido generalizados com base em ocorrências heterogéneas e mínimas. Assim, o apagamento da pluralização resultará da correlação entre os fenómenos de lambdacismo e de síncope

em busca da estrutura dissilábica CVCV do santomense e seus substratos africanos, por um lado, e do padrão das marcas precedentes, que configuram ocorrências consistentes e homogêneas, representadas pelas estruturas S\_ (item analisado na 2ª. posição, marca formal na 1ª. posição) e N\_ (item analisado na 2ª. posição, numeral na 1ª. posição). Desta forma, os falantes não inserem marcas no contexto fonológico posterior porque a marca de pluralização já foi semântica ou estruturalmente inserida no item anterior, realizando-se o padrão habitual de marcação PL no PA, ou seja, com inserção de marcas na primeira oportunidade, por norma no elemento pré-nuclear. A não-marcação nos outros elementos do SN acontece por questões funcionalistas (Kiparsky, 1972:195), evitando-se a redundância e estabelecendo-se o mesmo tipo de marcação estrutural verificado nos CP's e seus substratos africanos.

Conclui-se, então, que a motivação maior para a marcação/inibição PL dos itens do SN não é de carácter fonológico, resultando, essencialmente, de aspectos estruturais. Deste modo, a questão da fraca influência fonológica na marcação retira força aos pressupostos de as alterações nos padrões de concordância assentarem em processos de alteração fónica, históricos e internos ao próprio PE, como a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 2000), arguindo, em contrapartida, a favor da hipótese de influências das línguas africanas, levantada por Guy (1981a:3001-302), e que sugere que a variação ancora em um ou mais substratos.

Relativamente à variável *animacidade*, a única do tipo semântico por nós constituída, constatou-se a sua fraca influência na inserção de marcas de PL. No PA, ocorre tendência para se pluralizarem mais os itens de traço [+humano] do que os itens de traço [-humano]. Todavia, o facto de os adjectivos substantivados que eliminam o mecanismo morfossintáctico de função meramente gramatical para indicar o género, que passa a ser do tipo referencial, ocorrerem em posição pós-nuclear, leva a que os itens de traço [+animado] propiciem menos a marcação PL do que os elementos de traço [-animado]. Portanto, este tipo de inibição, que contribui para aplicação da CPL-var no SN do PA, é determinado mais por condicionalismos estruturais conotados à posição relativamente ao núcleo do SN do que por questões de carácter semântico. Por outro lado, o sistema de marcação PL estabelecido por esta variável para o PA apresenta um paralelismo evidente com outras variedades de português. Este aspecto vem também argumentar a

favor da hipótese de o português adquirido por falantes de um CP manifestar os mesmos padrões de variação do português adquirido quando a língua ancestral de substrato não é um CP (Figueiredo, 2008; 2009a).

Após análise de todas as variáveis linguísticas, constatou-se que os falantes em ASL, ao lidarem fonologicamente com a estrutura DET+N como um todo fonoprosódico na fase inicial da aquisição (DeGraff, 1999b:482), não só não percebem a segmentação entre as diferentes sequências morfológicas desta como também não estabelecem a correspondência entre as características formais e as formas morfofonológicas dos itens em questão (cf. Prévost & White, 1999 – MSIH). Desta forma, quer os classificadores nominais das L1's do grupo níger-congo atlântico quer o marcador temático *inen* do santomense poderão ter sido associados aos determinantes do português, concentrando-se as marcas de PL nestes, ou seja no primeiro elemento do SN, mais concretamente no item imediatamente pré-nuclear, em detrimento das demais pluralizações na cadeia sintagmática. Este aspecto revela não apenas que este elemento configura um *early system morpheme* gerado conjuntamente com os morfemas de conteúdo mas igualmente que ele é acompanhado por não-aquisição de material morfossintáctico nos outros itens do SN, e que se detecta quer nos CP's quer nas variedades reestruturadas de português. Como o estado FL0 da aquisição é composto apenas por conceitos semânticos, ou seja, caracteriza-se pela ausência sistemática de marcação de género e de número, os falantes em processo de aquisição do PtgL2, para além de lidarem inicialmente com as palavras como um todo, reconhecem a estrutura interna das formas morfológicamente complexas por comparação com outras formas, recorrendo a conexões lexicais (Koehn, 1994:35), ou seja, estabelecem ligações morfofonológicas entre a LA (o PE) e a sua L1, para elaborarem novas formações na interlíngua. Dando sequência a este entendimento, a inserção de pluralização na posição imediatamente pré-nuclear das gerações almorávida mais novas pode reflectir, então, o intrometimento de morfemas de conteúdo do português na configuração estrutural da L2 rudimentar dos falantes adultos em situação de aquisição do português, especialmente em estruturas nominais que transitaram geracionalmente via DLP's até à FE-1. Tais configurações nominais requeriam marcação PL imediatamente à esquerda, posição em que apareceria quer o classificador nominal de línguas do grupo níger-congo atlântico quer o marcador

temático *inen* do santomense, tendo pertencido às gerações seguintes a responsabilidade de reforçar este traço, assimilando-o como parte integrante no desenvolvimento da estrutura sintagmática quer do PA quer de outras variedades africanas e brasileiras de português. Assim, na gramática dos falantes do PtgL2 e dos falantes bilingues do PtgL1 influenciado por estruturas dos CP's atlânticos ou e de línguas africanas atlântico-ocidentais, a posição adjacente à esquerda do núcleo, isto é, a categoria DET, foi assumida como o ponto da marcação PL *singleton* (cf. Baxter, 2009:292). E no caso da pluralização dos outros elementos que compõem o SN, o que se nota é que a mesma tem apenas a função de orientar gramaticalmente a concordância, sendo esta possível com recurso aos *bridge late system morphemes*. Nesta conformidade, o suporte gramatical é tomado das línguas de substrato, enquanto os morfemas de conteúdo são, predominantemente, adquiridos do PE, sofrendo processos de gramaticalização, de modo a suprirem as necessidades da estrutura abstracta (P. Andrade, 2003:79). Por seu lado, os *early system morphemes* do superstrato foram absorvidos por essas línguas, enquanto os *late system morphemes* não transitaram para as línguas crioulas. O facto de os elementos pré-nucleares terem um papel activo na construção mental e estrutural da concordância relega os itens pós-nucleares para funções sintácticas de mera adjunção (Lucchesi, 2000a; Figueiredo, 2009a), ocorrendo a tendência para o núcleo e elementos pós-nucleares inibirem fortemente a marcação (Kiparsky, 1972:195; Castro & Ferrarinetto, 2007:75), já que a posterior activação dos *late system morphemes* é dispensada por questões funcionais, em virtude de a pluralização se revelar redundante. Na perspectiva da aquisição, a não-realização fonética da marca PL licenciada por {s} obedece a uma operação de identificação de conteúdo semântico-funcional (Cerqueira, 1994:125), que garante ao SDET, enquanto projecção máxima de N e SN, a possibilidade de projectar os traços categoriais e flexionais destes dois núcleos alargados, permitindo que os mesmos estabeleçam o pareamento de traços entre si. Por seu lado, a marca PL em DET vai permitir que a mesma não seja visível em N, ou seja, vai tornar possível o licenciamento de {s} nulo neste (Godinho, 2005:161). Paralelamente, como determinadas línguas não seleccionam traços [-interpretáveis] de género e número (Franceschina, 2002:76), como acontece em dialectos do grupo níger-congo atlântico e CP's atlânticos, os falantes adultos destes experimentam dificuldades na aquisição dos mesmos traços do português,

o que os leva a estabelecerem novos/errados parâmetros (Gonçalves, 2004:235), responsáveis pela produção da CPL-var a nível do SN no PtgL2 e no PtgL1 nativizado.

As questões referidas anteriormente permitem constatar que a posição pré-adjacente ao núcleo do SN e o princípio da redundância da pluralização actuam no processo de ASL, mas em momentos de distintos. Por outro lado, constata-se também que a L1 tem papel preponderante na inibição de marcas formais de PL. Face a estes aspectos, confirmaram-se os seguintes pressupostos: (i) existe relação entre posição e categoria gramatical, com influência na marcação PL, já que, em SN's de estrutura reduzida (2 ou 3 elementos), se estabelece um paralelo entre determinantes e primeira posição, substantivos e segunda posição e adjectivos e terceira posição (Guy, 1981b:179); (ii) não se detecta a intervenção do Princípio do Processamento Paralelo (Scherre, 1988:208) na inserção da pluralização no SN do PA, com marcas conduzindo a marcas e zeros levando a zeros; (iii) não se vislumbra a intervenção significativa de fenómenos fonológicos, como a assimilação regressiva (Braga, 1977:72), a haplologia sintáctica (Scherre, 1988:244) ou a deriva românica (Naro, 1981; Scherre, 1988; Naro & Scherre, 2000), no desenho da CPL-var do PA, sendo a motivação desta determinada, essencialmente, por questões de carácter estrutural; (iv) dificuldades de percepção do tipo fonológico na fase inicial da aquisição (cf. Faria Freitas & Miguel, 2001), visando a construção mental de concordância temática, levaram a que a categoria DET do português fosse interpretada como um afixo semântico apenso ao início do nome (Baxter, 2004:120), e não como categoria funcional que determina a construção a nível sintáctico; (v) tendo como premissa que a gramática da L1 é o ponto de partida para a ASL (Schwartz & Sprouse, 1996; Sprouse & Vance, 1999; Siegel, 2006; Sprouse, 2006), os falantes dos dialectos do grupo níger-congo atlântico, substratos do santomense, e dos CP's atlânticos, não seleccionam os traços [-interpretáveis] de número, inexistentes nas suas línguas, e experimentam dificuldades na aquisição dos mesmos traços do português, estabelecendo novos/errados parâmetros (cf. Gonçalves, 2004:235); (vi) no PA vai ocorrer marcação PL de acordo o sistema das línguas africanas (cf. Guy, 1981a:301-302), mais concretamente do grupo níger-congo atlântico (Baxter, 2009:292), nas quais a pluralização é controlada por um classificador imediatamente pré-nominal que estabelece paralelismos estruturais com a categoria DET das variedades africanas e

brasileiras de português reestruturado por transmissão linguística irregular; (vii) o sistema de marcação PL no SN do PA faz incidir o papel fulcral PL, principalmente, no elemento pré-nominal, com o desenvolvimento da concordância PL iniciando-se a partir da introdução pós-sintáctica de um PL *singleton*, visível no núcleo funcional DET, através de um sufixo que serve de âncora para o controlo da pluralização (cf. Baxter, 2009:292-293 – Fig. 1), mantendo-se os restantes elementos do SN inalterados, visto a informação de PL se tornar redundante, caso seja inserida nestes (Kiparsky, 1972:195; Castro & Ferrari-Neto, 2007:75); (viii) se não existir elemento visível em DET, o sufixo é atribuído ao núcleo complemento N; (ix) sendo a categoria DET a posição relevante para a interpretação da referência PL (cf. Costa & Silva, 2006b; Castro & Ferrari-Neto, 2007), a expressão de número é feita apenas neste elemento, com a identificação dessa informação envolvendo o processamento da concordância no SDET; (x) como a FL0 de bilingues apenas envolve o módulo pragmático da linguagem, não detendo o módulo do conhecimento gramático (Meisel, 1994a:92), não ocorre desenvolvimento normal deste (aquisição das categorias funcionais – Hipótese da Diferenciação da Linguagem), visto não haver reajustamento paramétrico (cf. Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), já que os novos/errados parâmetros foram tomados, em cadeia, como modelos para uma nova reinstanciação paramétrica na L1 nativizada (Gonçalves, 2004:235); (xi); o PL desenvolve-se com características de concordância parcial, através de morfemas dissociados, que são atribuídos aos elementos mais próximos do *fulcro*, contemplando os itens do SDET, primeiro, e os constituintes à direita do núcleo, posteriormente, mas mantendo, geracionalmente, o perfil da âncora na posição pré-adjacente ao núcleo do SN; (xii) o português adquirido em situação de contacto por falantes de um CP manifesta os mesmos padrões de variação do português adquirido numa situação de transmissão linguística irregular na qual a língua ancestral não é um CP (Figueiredo, 2008; 2009a); (xiii) de entre as variedades africanas e brasileiras de português analisadas, o PA é a que se encontra no estágio mais baixo da aquisição da regra de concordância PL, situando-se o seu grau de variação entre os patamares 3 e 4 do desenvolvimento sintáctico proposto pela hierarquia da “processabilidade” (Plag, 2008a,

2008b), já que não foram adquiridas, plenamente, as regras da flexão contextual, instanciadas a partir do nível 4 do desenvolvimento da interlíngua.

Quanto às quatro variáveis extralinguísticas constituídas, todas exercem influência sobre a aplicação da regra PL, com especial destaque para o grupo de factores *escolaridade* (Tabela 14.1). Os resultados gerais revelam-se consistentes com o quadro de variação apresentado pelos falantes de Almojarife, reflectindo, sincronicamente, aquele que terá sido o desenvolvimento diacrónico da regra de concordância de número da gramática da comunidade. Ao ser seleccionada pela ferramenta VARBRUL como a que mais peso exerce na configuração da CPL-var do SN do PA, a variável *escolaridade* confirmou que os anos de estudo têm influência na aquisição da regra de PL. Não obstante, os condicionalismos específicos que, nos períodos colonial e pós-colonial, afectaram o sistema educativo de São Tomé, em geral, e da comunidade de Almojarife, em particular, foram incontornáveis para determinar o modo como esta variável actua sobre a pluralização no SN. De facto, após a retirada definitiva dos colonos portugueses, as estruturas de ensino desarticularam-se e o corpo docente contratado, composto por nativos, terá contribuído também para estabilizar a CPL-var, visto deter um padrão de registo mais próximo do dos seus próprio alunos do que do padrão normativo do PE. Desta forma, o desenvolvimento da regra de concordância PL regista uma flutuação, evidenciada pela ausência de uma linha curvilínea ou de inclinação progressiva, a denotar que o PA não se encontra em estágio de mudança em curso, mas sim em situação de variação estável. Este aspecto determinou, então, que observássemos a variável *escolaridade* em articulação com as outras variáveis extralinguísticas.

O pressuposto de que os falantes com mais alto grau de escolaridade prestam mais atenção ao mecanismo da concordância comprovou-se no nosso estudo, já que são eles quem elabora SN's mais complexos e com maior grau de concordância entre os elementos destes. De facto, é a partir do momento em que os falantes começam a frequentar o ensino que se dá a maior expansão da aplicação da regra de concordância, uma vez que o ápice desta acontece na transição do estado de analfabetismo para o de escolarizado. Deste modo, mais do que uma determinação interna à própria língua, o maior índice de aplicação de marcas de PL aparece também conotado a factores de ordem social, como a *escolaridade*. Contudo, a desarticulação do sistema

socioeconómico e educativo do país retraiu a aquisição do desenvolvimento da regra de concordância PL, pelo que a aplicação da pluralização pelos falantes escolarizados de Almojarife continua a manifestar variação. A esta, não será também alheio o facto de ter ocorrido apagamento dos traços não-interpretáveis na *interlíngua* transmitida irregularmente como L1, o que impede o restabelecimento dos parâmetros que disponibilizariam as categorias funcionais (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clasher, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003) e provoca a instanciação de uma nova/errada parametrização (Gonçalves, 2004:235). Esta, por seu lado, é responsável pela fossilização suave (cf. Long, 2003:490) que determina que o perfil de concordância PL continue a ancorar a marcação no elemento em posição pré-adjacente ao núcleo do SN. Este tipo de fossilização, por sua vez, pode tornar-se permanente, caso os falantes não recebam *input* correctivo, isto é, não se alterem as condições do actual panorama de ensino em São Tomé.

Com recurso à variável *idade* foi possível verificar não só a existência de gramáticas distintas nas diferentes faixas etárias da comunidade mas também que a distribuição da regra de concordância PL pelas três faixas etárias de Almojarife apresenta um *continuum* que faz pressupor, à primeira vista, que o PA se encontra em estágio de mudança em curso, visto ser a faixa etária mais nova aquela que produz mais realizações a favorecerem moderadamente a concordância PL no SN. Contudo, situando-se o ápice da aquisição desta na transição da FE-2 para a FE-1, constatou-se que o mesmo condiz não só com o período em que ocorreu maior migração do meio rural para o ambiente urbano, levada essencialmente a cabo por indivíduos do sexo masculino, mas também com a fase em que as estruturas socioeconómicas da administração colonial, nomeadamente do ensino, estavam no seu apogeu. Como tal, existem dois tipos de falantes na FE-1: os mais velhos, que experienciaram um contacto mais directo com o PE, quando este fazia sentir de forma marcada o seu padrão; e os mais novos, afectados pela desarticulação das antigas estruturas socioeconómicas coloniais, nomeadamente do seu sistema de ensino e viário, o que impediu que eles pudessem dar continuidade à expansão da regra de concordância PL. Efectivamente, com a comunidade a ser jogada para um isolamento cada vez mais acentuado, os falantes mais jovens da FE-1 não podem aceder ao mesmo nível de ensino dos seus antecessores e, quando frequentam a



escola, contactam com docentes falantes do português santomense, isto é, que apresentam um padrão linguístico próximo do seu. Assim sendo, a variação estabilizou, levando o PA a conservar muitas das suas características mais criouliizantes.

Acerca da variável *estadia*, é possível concluir que a influência desta no favorecimento da marcação PL terá mais a ver com situações de migração, sobretudo levadas a cabo pela faixa etária intermédia de Almojarife, do que com ausências para fora do país, já que as últimas não implicaram um contacto directo com falantes do PE. Desta forma, o factor *ausência da comunidade* não contribuiu para que os falantes almojarifanos passassem a produzir padrões de concordância PL mais próximos dos do PE.

No que concerne à variável *sexo*, permitiu verificar que os homens lideram a inserção de marcas nas duas faixas etárias mais idosas de Almojarife, acontecendo uma inversão no padrão de propiciamento de marcas PL na geração mais nova, com as mulheres sendo responsáveis pela mesma. O comportamento da variável *sexo* é explicado em função de fenómenos sociolinguísticos diversos, tendo os homens, inicialmente, mantido um contacto mais directo com o exterior da comunidade, situação que se inverteu recentemente, dada a nova realidade socioeconómica de São Tomé. Sem postularmos pressupostos, por necessitarem de consistência científica que não foi possível observar neste trabalho, não deixa de ser curioso que, a comparação com o comportamento da variável *sexo* de outras variedades africanas e brasileiras de português nos tenha evidenciado a particularidade de, nas comunidades rurais, os homens revelarem um uso maior da regra de concordância em número, enquanto as mulheres apresentam um maior índice de variação. Em sentido contrário, nas comunidades urbanas, são as mulheres a liderar o uso das formas mais próximas do português padrão, cabendo aos homens apresentar o maior grau de variação.

A L1, da mesma forma que outras línguas que um indivíduo fala, influencia o novo sistema linguístico em aquisição, especialmente nas fases iniciais desta. Ao analisarmos a CPL-var no SN do PA, procurámos, no plano linguístico, observar a relação estabelecida entre factores derivados da estrutura da gramática, a fim de verificarmos duas características fundamentais da interlíngua: (i) até que ponto ela emerge como um reflexo imediato da estrutura da gramática da L1; (ii) até que aspecto ela reflecte

aproximação ao sistema da LA. Vistas estas questões, poderíamos passar a identificar os factores condicionadores, derivados da estrutura das gramáticas presentes na mente dos aprendentes do PA, buscando um entendimento acerca dos processos cognitivos e sociolinguísticos subjacentes à aquisição do PtgL2, por parte de falantes de uma L1 africana atlântico-ocidental ou de um CP atlântico, e do PtgL1 adquirido de modo defectivo, por parte de falantes de um CP atlântico. Em última instância, e comparando o perfil de marcação PL no SN do PA com o de outras variedades africanas e brasileiras de português, seria possível determinar se o português adquirido por falantes de um CP manifesta os mesmos padrões de variação do português adquirido por falantes de uma língua que não é um CP. Analisados e entendidos estes fenómenos, estariam detectadas as orientações para perceber as motivações que originam a CPL-var no SN do PA, encontrando-se, conseqüentemente, a resposta para determinar qual (ou quais) das quatro questões levantadas na discussão em torno destas motivações (cf. introdução do presente trabalho) justifica, adequadamente, a origem do fenómeno da CPL-var no SN do PB popular.

Das constatações gerais, foi possível vislumbrar que os almosarifanos apenas aplicam a pluralização em 51% dos itens analisados, enquanto a concordância plena de PL ocorre somente em 29% dos SN's que a requerem. Estes aspectos revelam que não há tendência para a fixação da regra de concordância PL no PA, encontrando-se esta língua em estágio de variação estável. A intervenção da posição pré-nuclear adjacente na fixação da pluralização e concentração da concordância espelha o papel fulcral do PL no classificador nominal na tipologia níger-congo e, portanto do substrato africano do crioulo santomense, por um lado, bem como o papel fulcral do PL no marcador temático *inen* dos nomes de traço [+humano] do santomense, por outro lado. Neste estudo, ao defendermos o acesso parcial à GU (Hawkins, 1993, 1998, 2001; Yip, Rutherford & Clashen, 1995; Hawkins & Chan, 1997; Franceschina, 2002, 2003), entendemos que os princípios invariáveis se mantêm acessíveis na ASL, mas também que os parâmetros não podem ser restabelecidos, visto as categorias funcionais deixarem de estar disponibilizadas depois do período crítico do desenvolvimento, ou seja, que o aprendente da L2 só tem acesso às categorias funcionais da sua L1. Deste modo, os traços de número são mais dificilmente adquiridos na ASL se não estiverem presentes no

conjunto de traços funcionais das L1's. Como o traço de número é seleccionado opcionalmente pelas línguas, em virtude de ser do tipo [-interpretável], os falantes nativos de línguas africanas atlântico-ocidentais ou CP's atlânticos, que não têm concordância sintáctica em número, quando aprendem o português, encontram obstáculos na aquisição do referido traço, visto o mesmo constituir um parâmetro deste idioma. Com base nestes pressupostos, é lícito inferir que a variação na concordância de número registada nas produções dos falantes almoxarifanos está relacionada com o facto de o traço de número não estar instanciado nem no santomense nem nos substratos deste. Como o perfil de marcação PL no SN do PA é idêntico ao das variedades africanas e brasileiras de português que tomámos para observações numa perspectiva comparada, conclui-se que o português adquirido por falantes de um CP manifesta os mesmos padrões de variação do português adquirido por falantes de uma língua que não é um CP, mas que teve no bojo da sua formação uma envolvência de transmissão linguística irregular em que intervieram línguas do substrato africano do grupo níger-congo atlântico. Desta forma, confirma-se também o pressuposto de Guy (1981a:301-302) acerca da influência das línguas africanas na CPL-var do SN de variedades brasileiras de português emergentes em situação de transmissão linguística irregular, em que ocorreu forte *input* de dados de estruturas crioulas influenciadas por estruturas africanas no processo de aquisição e reestruturação da L1 (Baxter & Lucchesi, 1997). Nesta conformidade, a explicação para a formação das línguas reestruturadas, para além de levar em conta as questões relacionadas com o tipo de aquisição e o desenvolvimento da L2 e L1 até ao estágio final de ILn ou língua-I, não pode também descartar a matriz sociolinguístico-histórica em que a aquisição se dá (Holm, 1998; 2004), assimilando as perspectivas da transmissão linguística irregular, com influências das L1's ancestrais (Winford, 2003a) e das línguas pidgin ou línguas crioulas.

A interpretação do conjunto dos resultados obtidos a partir da análise das variáveis que constituímos para este estudo dá-nos uma ideia bastante clara do quadro da variação observado no SN do PA, excluindo a análise de cada variável como um parâmetro independente. Contudo, sentimos que a análise aprofundada desta variedade de português continua ainda por ser feita, a fim de a mesma poder continuar a dar o seu contributo não só para o estudo das línguas em contacto, em geral, mas também para

ajudar a identificar melhor os problemas subjacentes à aquisição das diferentes categorias linguísticas, em particular. Assim, num futuro próximo, e com recurso aos dados de que dispomos, pretendemos completar as observações acerca da CPL-var no SN do PA, efectuando a observação desta numa perspectiva sintagmática. Paralelamente, e na perspectiva da sociolinguística variacionista, efectuaremos também estudos sobre a aquisição e desenvolvimento da concordância de género e de estruturas predicativas. Por outro lado, observações sobre o padrão dos nomes nus, das construções que envolvem pronomes pessoais e suas funções sintácticas ou das estruturas com orações relativas, não poderão também ser desconsideradas, sob pena de deixarmos de contribuir com achados significativos para os estudos da aquisição da L2 e L1, e consequente desenvolvimento das suas estruturas linguísticas. Para se confirmar se o PA se encontra ainda em estágio de variação estável ou não, será necessário proceder à recolha, num futuro próximo, de mais dados sincrónicos, especialmente nas faixas etárias das crianças e jovens adolescentes.

Por fim gostaríamos de referir que o presente estudo, apesar de não ter tido como objectivo assuntos de natureza pedagógica, contribuiu também para que, enquanto docentes de PtgLE, passássemos a entender melhor as dificuldades dos nossos aprendentes na aquisição e uso do português e reforçássemos a convicção de que é absolutamente essencial identificar os problemas subjacentes à aquisição das diferentes categorias linguísticas de uma língua tipologicamente distinta da L1, a fim de se providenciar *input* apropriado aos aprendentes do PtgLE. O estudo realizado despertou-nos ainda mais o interesse pela problemática subjacente à ASL e, em particular, pela aquisição da morfossintaxe do PtgLE. Durante o tempo em que se desenvolveu o presente estudo, tivemos o privilégio de ensinar português a falantes nativos do chinês, sendo-nos dado a perceber que esta língua partilha algumas características com as línguas do ramo bantu. Nesta perspectiva, impõe-se também a recolha de dados de português a falantes nativos deste ramo, que permitirão estabelecer estudos comparados com o PtgL2 de falantes chineses, a fim de se observar se o padrão de aquisição e desenvolvimento das categorias linguísticas estabelece analogias entre ambos. Por outro lado, a análise comparativa da mesma recolha com os dados que se observaram neste trabalho permitirá encaixar mais um elo na cadeia da transmissão linguística que, em

séculos passados, se terá iniciado em terras africanas, para aportar, depois, no Brasil. Este aspecto, poderá confirmar, em definitivo, o pressuposto acerca da influência das línguas do substrato africano na CPL-var do SN de variedades brasileiras de português (Guy, 1981a:301-302), emergentes em situação de transmissão linguística irregular em que ocorreu forte *input* de dados de PtgL2 influenciado por estruturas de línguas africanas (Baxter & Lucchesi, 1997) do grupo níger-congo atlântico (Baxter, 2009:292) no processo de aquisição e reestruturação da L1.

<sup>161</sup> Embora não se tivesse analisado o fenómeno da CPL-var no SN do PA numa perspectiva sintagmática, achou-se importante obter dados sobre a concordância plena entre todos os elementos do SN, já que, observando esta em concomitância com as análises que iremos levar a cabo, poder-se-á tirar conclusões mais exactas acerca da possibilidade de a marcação PL ocorrer maioritariamente na primeira oportunidade, com maior incidência sobre o elemento em posição imediatamente anterior ao núcleo sintagmático. A confirmar-se este pressuposto, ganha forte sustentação a possibilidade de o sistema de pluralização no SN do PA ocorrer conforme o sistema de pluralização dos seus ancestrais substratos bantu (Guy, 1981a:301-302), deixando, por tal, de serem inseridas marcas nos restantes itens da cadeia sintagmática, em virtude de a mesma se revelar redundante (cf. Kiparsky, 1972; Castro & Ferrari-Neto, 2007:75).

<sup>162</sup> Na perspectiva da aquisição, sendo ambos os traços de género e número processados no âmbito do SDET, o facto de os aprendizes em ASL não interpretarem o traço de género implica que a categoria intermédia SNUM não foi totalmente adquirida. À partida, este aspecto iria também comprometer a interpretação do traço de número, determinando CPL-var. Contudo, como se referiu no ponto 1.4.2, para as diversas classes de nomes das línguas bantu, representadas por afixos que fazem a diferenciação entre singular e PL, as distinções sexuais não se revelam importantes (Welmers, 1973:159), um aspecto que transitou para os CP's e variedades africanas e brasileiras de português. Desta forma, a marcação de género pode restringir-se aos elementos nucleares, sendo o valor referencial do sexo fornecido com nomes adjectivados que acompanham o nome. A eliminação do mecanismo morfossintáctico de função meramente gramatical para a indicação do género, e consequente não utilização de elementos morfémicos do tipo sufixal para marcar este, vai determinar que a interpretação do traço de número não seja comprometida pela aquisição do género referencial, tornando o peso da variável *marcação de género* irrelevante no condicionamento à inserção das marcas de PL.

<sup>163</sup> Neste tipo de tratamento dos dados, faz-se apenas uma rodada dos mesmos, incluindo todos os factores. A rodada produz células que permitem observar as combinações de factores e grupos de factores, detectando-se, assim, as situações em que há anomalias nos dados. Outro dos usos deste tipo de rodada é aquele a que recorremos no caso em questão, isto é, comparar a frequência das variáveis dependentes, de forma a analisar-se o efeito das combinações em termos de peso de factores. Refira-se que a outra opção oferecida pelo VARBRUL concerne à rodada designada por *step up/step down*, que se distingue da opção *one-level* em virtude de testar a significância para cada grupo de factores. Para informações detalhadas sobre o processamento *one-level*, consulte-se Poplack & Tagliamonte (1998).

<sup>164</sup> O resultado de estatisticamente “não significativo” quer dizer que “*as diferenças eliminadas com a amalgamação não eram significativas e que, de fato, devem ser jogadas fora, eliminadas*” (Guy & Zilles, 2007:194), optando-se, então, pelo factor amalgamado.

<sup>165</sup> Relembremos que, quando um factor com trinta ou mais ocorrências regista uma frequência de 100%, o seu valor se sobrepõe ao efeito de qualquer outro contexto presente. Na terminologia de análise da ferramenta VARBRUL, este factor representa um *knockout* verdadeiro, que reflecte uma mudança consumada (pr. 1). Como tal, a metodologia aconselha que tais factores não sejam incluídos na análise variacionista, dado que os seus pesos não precisam de ser calculados. Ainda assim, “*devem ser relatados, e o seu papel em processos de mudança, especialização de significado ou de função merece ser discutido*” (Guy & Zilles, 2007:158).

<sup>166</sup> A própria autora admite, contudo, não ter observado adequadamente a “*diferença sistemática encontrada entre o efeito da presença ou ausência de marcas antes do terceiro, quarto ou quinto elementos, em SNs com marcas no seu primeiro ou segundo elemento*” (Scherre, 1988:170), pelo que se tornaria imperioso estudar esta variável “*em termos de oposições gerais e oposições mais específicas.*” (Scherre, 1978:91). Quanto às conclusões globais para estas ocorrências, são explicadas de acordo com o fenómeno da redundância. Ainda assim, Scherre (1988:170) admite que não conseguiu um princípio geral que desse plena conta do de todo o fenómeno da CPL-var.

<sup>167</sup> Como se pode confirmar pelo exemplo da Tabela 12.3 (“*noventa tal contos*” [ZECAH1]), as realizações que configuram a codificação NM\_ (item na 3ª. posição, antecedido de numeral sem –s na 1ª. posição e elemento sem pluralização na 2ª. posição) possuem, todas elas, a ocorrência “tal” em posição pré-nuclear. Como se sabe, o indefinido “tal” pode ocorrer em posição adnominal com função de determinante. Caso traduza intensidade ou grau elevado, quando combinado com substantivos abstractos ou massivos, apresenta valor adjectival e pode ser flexionado em número (p.e.: “*Nunca conheci ninguém com tais qualidades*”). O mesmo acontece quando “tal” indica intensidade, quantidade ou número superior, em correlação com uma frase consecutiva principiada pela conjunção “que” (p.e.: “*Fizeram tais exigências, que foi impossível aceitá-las*”). No entanto, os casos em que codificamos “tal” como item gramatical não flexionável em número reportam-se a ocorrências em combinação com numerais cardinais, indicando aproximação numérica (Casteleiro, 2001, vol. G-Z:3502). De facto, realizações do tipo “\**Tenho vinte e tais anos*” são agramaticais no português. Refira-se ainda que, relativamente à configuração do tipo NsM\_ (item em 3ª. posição, antecedido de numeral finalizado em –s na 1ª. posição e elemento sem pluralização na 2ª. posição), o nosso corpus não regista qualquer realização.

<sup>168</sup> Guy (1981a:180), ao contrário de P. Andrade (2003:112), não concorda com a hipótese funcionalista kiparskiana de que a informação semântica de pluralização encerrada pelos numerais tem tendência a ser retida na estrutura de superfície. Assim, Guy (1981a:180) aponta para a possibilidade de a influência destes elementos gramaticais na marcação reflectir automonitorização e autocorreção, ou seja, a probabilidade de o informante estar a adaptar as suas realizações ao discurso padrão do entrevistador.

<sup>169</sup> Ainda a propósito deste tipo de pluralizações, remetemos outras constatações para os capítulos 1 (ponto 1.2.2.1) e 2 (ponto 2.6.1) do presente trabalho, nos quais são tratadas questões da concordância PL dentro do SN do PCV e do PE, respectivamente.

<sup>170</sup> “*If you throw everything into one factor group, it can be termed the ‘kitchen sink effect’. While such a model might fit the data better, it will not tell you as much if it misses linguistically valid generalizations elsewhere.*” (Tagliamonte, 2006:157).

<sup>171</sup> Scherre (1988:228) não procedeu à codificação e análises do adjectivo em termos de subclasses semânticas, em virtude de ter detectado que, em termos percentuais, o adjectivo do tipo ‘avaliativo’ “*apresenta o mesmo grau de concordância de que os demais adjectivos pospostos*”.

<sup>172</sup> No artigo de Schiffrin (1981) é apresentado um estudo sobre o modo como o pretérito perfeito e o presente histórico servem para referir eventos passados no texto narrativo. Para tanto, a autora efectua uma análise quantitativa de setenta e três excertos de narrações orais, observando os seguintes aspectos: (i) constrangimentos que actuam ao nível do presente histórico e do pretérito perfeito; (ii) funções do presente histórico na narração; (iii) razões que determinam a significância do presente histórico. Independentemente de ter constatado que o presente histórico é usado para incrementar o impacto dramático da narração ou apresentar acontecimentos que o narrador considera mais relevantes, a autora mostra também que a narração é delimitada pelo discurso, no qual as variações formais e funcionais da gramática são observadas e controladas de forma sistemática. Deste modo, a organização da narrativa delimita as fronteiras dentro das quais o presente histórico ocorre, enquanto os vários constrangimentos estruturais e funcionais restringem (ou favorecem) a permuta entre os dois tempos verbais.

<sup>173</sup> Os exemplos apresentados na Tabela 12.7 para os factores 4 e 5 revelam amiúde, por exemplo, a ocorrência de itens antecedentes que possuem marca semântica de PL.

<sup>174</sup> No PB, o artigo definido “*é, na maioria dos casos, de emprego facultativo junto a possessivos em referência a nome expresso:*

Meu livro ou o meu livro.

*OBSERVAÇÃO: É obrigatório o artigo quando o possessivo é usado sem substantivo, em sentido próprio ou translato:*

*Bonita casa era a minha. Fazer das suas.*

*Mas sem artigo dizemos várias expressões, como de meu, de seu natural, linguagem com que traduzimos “os bens próprios de alguém” – a primeira – e “qualidades naturais” – a última:*

*Nunca tive de meu, outro bem maior.*

*‘Bernardes era como estas formosas de seu natural que se não cansam com alindamentos, a quem tudo fica bem’ [AC].*

*Dispensa ainda artigo o possessivo que entra em expressões com o valor de alguns:*

*Os Lusíadas têm suas dificuldades de interpretação.*

*Finalmente, na expressão de ato usual, que se pratica com frequência, o possessivo vem normalmente sem artigo:*

*Às oito toma seu café.” (Bechara, 1999 [1928]:157).*

<sup>175</sup> Tradução nossa.

<sup>176</sup> Scherre (1988:257-265) levou a cabo a observação da influência da função sintáctica sobre o número de marcas plurais no SN, em virtude de ter detectado um tipo especial de SN que, nos seus dados, se apresenta consistentemente menos marcado do que os outros: o SN que funciona como resumo de ideias anteriores (SN resumitivo ou fático), e que se pode apresentar como uma unidade parentética (p.e. *Passavam filmes sobre a vida de Jesus Cristo. Esses negócio assim!*). Para tanto, criou uma variável independente denominada *Função Sintáctica do SN*, com o objectivo de analisar o comportamento dos diferentes tipos de SN’s e confirmar se as funções sintácticas tradicionais exercem influência sobre o número de marcas formais de PL no SN. Um dos factores dessa variável inclui as estruturas sintagmáticas que a linguista classificou de *SN’s de Função Abortada*, isto é, as situações em que ocorrem interrupções no discurso, não sendo possível a identificação precisa da sua função sintáctica. Ora, este é precisamente o caso do exemplo avançado por Lopes (2001:180), em que a parte da realização que antecede a interrupção do discurso pode ser vista como uma estrutura sintagmática de função sintáctica abortada, composta por determinante e categoria substantivada (ex. *as mesmas... quer dizer... assim... PESSOAS que são*). Refira-se apenas que as conclusões de Scherre (1988:265) evidenciaram que as funções sintácticas tradicionais não exibem oposição evidente entre si e

---

que quer os SN's de funções consideradas essenciais e integrantes quer os de funções acessórias apresentaram índices de concordância bastante semelhantes. Não obstante, os SN's resumitivos inibem a marcação, enquanto os outros a favorecem. A autora acredita que o desfavorecimento dos primeiros se fica a dever à sua carga funcional e, pelo facto de ocorrerem normalmente à direita da oração, perdem a possibilidade de se apresentarem como ponto de referência do discurso.

<sup>177</sup> Ainda no âmbito do projecto “*Semi-creolization: testing the hypothesis against data from Portuguese-derived languages of São Tomé (Africa)*”, Alan N. Baxter recolheu dados de alguns falantes do PTL2, que ainda não foram trabalhados até à data presente.

<sup>178</sup> Tradução nossa.

<sup>179</sup> Segundo Stenzel (1994), também não ocorre marcação de caso na fase FL0 de aquisição bilingue das línguas alemã e francesa como L1's.

<sup>180</sup> Para Koehn (1994), as duas fases do processo de amadurecimento da aquisição das línguas das crianças bilingues ocorrem de acordo com as seguintes etapas: (i) ausência sistemática de género e número, até aos 2;4 anos; (ii) noções gramaticais de género e número, entre os 2;5 e os 5;0 anos. Müller (1994a), como referimos (ponto 2.6.4.2), pormenoriza que a noção gramatical de género acontece aos 2;0 anos, enquanto a de número se dá aos 2;4 anos).

<sup>181</sup> Os estudos de Meisel (1994a) sobre o modo como se processa a aquisição de finitude, concordância e tempo na gramática da fase FL0, e que o levaram a concluir que esta não possui categorias funcionais, fizeram também com que o autor postulasse que a categoria central a ser depreendida é o VERBO. Deste modo, as realizações frásicas circunscrevem-se ao SV nesta fase, uma vez que a ordenação entre os elementos que aparecem neste ainda não é dependente da forma morfológica nem é regulada por princípios gramaticais. Como tal, a sequência dos itens frásicos deriva da relação entre tema e rema, não obstante Meisel (1994a) admitir que se observam já padrões específicos para cada uma das línguas, alemão e francês, explicadas pela influência e frequência do *input* a que o falante está sujeito.

<sup>182</sup> O motivo que leva estes itens a não beneficiarem a marcação prende-se com o facto de revelarem alguma incerteza quanto ao modo de formarem o PL, conforme salientou Scherre (1988:124), ao referir-se às diferentes maneiras de se pluralizarem os elementos com singular em *-ão* (p.e. *ancião* > *anciãos/anciões/anciães*). A este propósito, note-se também que as duas formas de se pluralizarem os itens finalizados em *-il*, caso sejam oxítonos ou paroxítonos (p.e. *funil* > *funis*; *fácil* > *fáceis*), poderão levar os falantes a hesitarem na forma de inserir a marcação de número, um pouco à semelhança daquilo que sucede com os elementos em *-ão*.

<sup>183</sup> Informação prestada verbalmente por Anna Jon-And ao autor da presente pesquisa em Colónia, Alemanha, durante a realização do *Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLe)*, Universidade de Colónia, Alemanha, 11-15 de Agosto.

<sup>184</sup> Para o MRJ, a equipa do Projecto Censo estabeleceu três categorias de escolarização: antigo Ensino Primário (informantes com um a quatro anos de escolarização), antigo Ensino Ginásial (informantes com cinco a nove anos de escolarização) e antigo Ensino Colegial (informantes com nove a onze anos de escolarização). Quanto à NURC, os informantes foram escolarmente estratificados de acordo com as duas fases da recolha de dados: (i) década de 70, com informantes unicamente de Escolaridade Superior; (ii) década de 90, com informantes de Escolaridade Fundamental (um a cinco anos de escolarização), Escolaridade Média (2º. grau completo, isto é, 11 anos de escolarização) e Escolaridade Superior.

<sup>185</sup> Pesos relativos apresentados no MRJ para os diferentes grupos: Primário = pr. 0,48; Ginásial = pr. 0,27; Colegial = pr. 0,37 (Scherre, 1988:86).

<sup>186</sup> A consoante áptico-palatal surda [ʃ] não foi incluída nas amalgamações, visto ocorrer em apenas quatro dados.

<sup>187</sup> Para constatação dos exemplos a que nos referimos, veja-se Ferraz (1979:55).

<sup>188</sup> Para uma observação detalhada de exemplos de harmonização vocálica nos substratos africanos do santomense, consulte-se Ferraz (1979:49-51).

<sup>189</sup> “*Considera-se a ausência ou deficiência da monitoração da instituição escolar como o principal responsável pela variação da concordância*” (Lopes, 2001:106).

<sup>190</sup> Factores constituídos por Jon-And (2009:4) para o grupo de factores *escolaridade*: a) zero anos de escolaridade; b) 3-4 anos de escolaridade; c) 5-7 anos de escolaridade.



---

<sup>191</sup> Tradução nossa.

## Referências bibliográficas

### I. Bibliografia consultada e citada:

- Abney, Steve, 1987. *The english noun phrase in its sentential aspect*. Cambridge, MA: The MIT Press. Dissertação de doutoramento.
- Adams, Marianne Patalino. 1987. *Old french, null subjects, and verb second phenomena*. Oakland, CA: University of California. Dissertação de doutoramento.
- Adamson, H. Douglas. 1988. *Variation theory and second language acquisition*. Washington, DC: Georgetown University Press.
- Adamson, H. Douglas & Ceil Kovak. 1981. Variation theory and second language acquisition: an analysis of Schumman's data. In David Sankoff & Henrietta Cedergren (orgs.), *Variation omnibus*, 285-292. Edmonton: Linguistic Research Inc.
- Adendorff, Rajend. 2002. *Fanakalo - a pidgin in South Africa*. Language in South Africa. Cambridge, MA: Cambridge University Press.
- Adjemian, Christian. 1976. On the nature of interlanguage systems. *Language Learning*, 26(2). 297-320.
- Agheyisi, Rebecca Nogieru. 1986. *An edo-english dictionary*. Benin City: Ethiope Publishing Corporation.
- Alexandre, Nélia & Tjerk Hagemeyer. 2007. Bare-nouns and the nominal domain in Saotome. In Marlyse Baptista & Jacqueline Guéron (eds.), *Noun phrases in creole languages: a multi-faceted approach*, 37-59. Amsterdam/Philadelphia: John Benjamins.
- Almada, Maria Dulce de Oliveira. 1961. *Cabo Verde: contribuição para o estudo do dialecto falado no seu arquipélago*. Lisboa: Junta de Investigações do Ultramar.
- Amaral, Amadeu. 1920. *O dialeto caipira*. São Paulo: O Livro.
- Andersen, Roger W. 1980. Creolization as the acquisition of a second language as a first language. In Albert Valdman & Arnold Highfield (eds.), *Theoretical orientations in creole studies*, 273-299. San Diego: Academic Press.
- 1983a. *Pidginization and creolization as language acquisition*. Rowley, MA: Newbury House.
- 1983c. Transfer to somewhere. In Susan M. Gass & Larry Selinker (eds.), *Language transfer in language learning*, 177-201. Rowley, MA: Newbury House.
- Anderson, Stephen.R. 1982. Where's morphology? *Linguistic Inquiry* 13. 571-612.
- Andrade, António. 1999. Demonstrativos e [ana]fóricos em latim. *Ágora. Estudos Clássicos em Debate*, 1. 155-171
- Andrade, Leila Minatti. 2003. *Rupturas e contínuos da concordância nominal de número em textos orais de informantes de Tubarão (SC) e São Borja (RS)*. Santa Catarina: Universidade do Sul de Santa Catarina. Dissertação de mestrado.
- Andrade, Patrícia Ribeiro de. 2003. *Um fragmento da constituição sócio-histórica do português do Brasil – variação na concordância nominal de número em um dialeto afro-brasileiro*. Salvador: Universidade Federal da Bahia – Instituto de Letras. Dissertação de mestrado.
- Aronof, Mark. 1976. *Word-formation in generative grammar*. Linguistic Inquiry Monograph 1 Cambridge, MA: The MIT Press.
- Asher, Ronald E. (ed.). 1994. *The encyclopedia of language and linguistics*, vol. 1. Oxford/New York/Seoul/Tokyo: Pergamon Press.

- Azevedo, Páscoa F. Maria P. 2003. Interferência das línguas maternas de Moçambique no português. In Departamento de Língua e Cultura Portuguesa – Faculdade de Letras da Universidade de Lisboa (ed.). *1.ªs jornadas de língua e cultura portuguesa (18/19 de Junho de 2002)*, 91-102. Lisboa: Departamento de Língua e Cultura Portuguesa – Faculdade de Letras da Universidade de Lisboa.
- Bagno, Marcos. 2001. Mudança lingüística: um fenômeno onde toda prescrição é inútil. *Letras – Revista do Instituto de Letras (PUC Campinas)*, vol. 20 (1/2). 44-60.  
 –2002 (coord.). *Lingüística da norma*. São Paulo: Ed. Loyola.
- Baia, Maria de Fátima de Almeida. 2008. Estudo experimental sobre o formato prosódico inicial na aquisição do português brasileiro [em linha]. In *Estudos Lingüísticos*, 37 (2). 27-36. [Consult. 04 Abr. 2010]. Disponível em: [http://www.gel.org.br/estudoslinguisticos/volumes/37/EL\\_V37N2\\_03.pdf](http://www.gel.org.br/estudoslinguisticos/volumes/37/EL_V37N2_03.pdf).
- Baker, Mark C. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry*, 16. 373-415.  
 –1988. *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.  
 –2001. *The atoms of language: the mind's hidden rules of grammar*. New York, NY: Basic Books.  
 –2002. Building and merging, not checking: the nonexistence of (Aux)-SVO languages. *Linguistic Inquiry*, 33. 321-328.
- Baker, Philip. 1982. On the origins of the first mauritians and of the creole languages of their descendants: a refutation of Chaudenson's 'Bourbonnais' theory. In Philip Baker & Chris Corne (eds.), *Isle de France creole: affinities and origins*, 131-259. Ann Arbor: Karoma.  
 –1994. Creativity in creole genesis. In Dany Adone & Ingo Plag (eds.) *Creolization and language change*, 65-84. Tübingen: Max Niemeyer Verlag.
- Bakker, Peter. 2003. Pidgin inflectional morphology and its implication for creole morphology. In Geert Booji & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 3-33. Dordrecht: Kluwer Academic Publishers.
- Baldinger, Kurt. 1972 [1958]. *La formación de los dominios lingüísticos en la Península Ibérica*. Madrid: Gredos.
- Baptista, Marlyse. 2002. *The syntax of cape Verdean creole: the Sotavento varieties*. *Linguistik Aktuell/Linguistics Today*, vol. 54. Amsterdam/Philadelphia: John Benjamins.  
 –2003. Inflectional plural marking in pidgins and creoles: a comparative study. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 315-332. Tübingen: Max Niemeyer Verlag.  
 –2007. On the syntax and semantics of DP in cape verdean creole. In Marlyse Baptista & Jacqueline Guéron (eds.), *Noun phrases in creole languages: a multi-faceted approach*, 61-105. Amsterdam/Philadelphia: John Benjamins.
- Baptista, Marlyse, Heliana Mello & Miki Suzuki. 2007. Kabuverdianu, or cape verdean, and kriyol, or Guinea-Bissau (creole portuguese). In John Holm & Peter L. Patrick (eds.), *Comparative creole syntax: parallel outlines of 18 creole grammars*, 53-82. Westminster Creolistic Series 7. Plymouth, UK: Battlebridge Publications.
- Bardovi-Harlig, Kathleen. 2000. *Tense and aspect in second language acquisition: form, meaning and use*. Oxford: Blackwell.

- Barreña, Andoni. 1997. Desarrollo diferenciado de sistemas gramaticales en un niño vasco-español bilingüe. In Ana Teresa Pérez-Leroux & William R. Glass (eds.), *Contemporary perspectives on the acquisition of spanish*, 55-74. Somerville, MA: Cascadilla Press.
- Barreto, Manuel Saraiva. 1977. Aspectos da lusofonia em Moçambique: para uma lusofonia moçambicana. Algumas questões linguísticas e didáticas. In *Actas do primeiro encontro nacional para a investigação e ensino do português – 1976*, 529-548. Águeda: Grafilarte.
- Bates, Elizabeth *et alii*. 1982. Functional constraints on sentence processing: a crosslinguistic study. *Cognition*, 11. 245-299.
- Bates, Elizabeth & Judith C. Goodman. 1999. On the emergence of grammar from the lexicon. In Brian MacWhinney (ed.), *The emergence of language*, 29-79. Mahwah, NJ: Erlbaum.
- Battistella, Edwyn L. 1996. *The logic of markedness*. Oxford: Oxford University Press.
- Baxter, Alan Norman. 1988. *A grammar of kristang (Malaca creole portuguese)*. Pacific Linguistics, série B, 95. Camberra: Pacific Linguistics.
- 1992. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a criouliização prévia: um exemplo do Estado da Bahia. In Ernesto d’Andrade & Alain Kihm, *Actas do colóquio sobre crioulos de base lexical portuguesa*, 72-90. Lisboa: Colibri.
- 1997. Creole-like features in the verb system of an afro-brazilian variety of portuguese. In Arthur K. Spears & Donald Winford (orgs.), *Pidgins and creoles: structure and status*, 265-288. Amsterdam/Philadelphia: John Benjamins.
- 1998. O português vernáculo do Brasil: morfossintaxe. In Matthias Perl & Armin Schwegler (eds.), *América negra: panorâmica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*, 97-134. Frankfurt am Main: Vervuert/Madrid: Iberoamericana.
- 2002. Semicreolization? – The restructured portuguese of the tongas of São Tomé – a consequence of L1 acquisition in a special contact situation. *Journal of Portuguese Linguistics*, 1. 7-39.
- 2003. *A reestruturação do português numa situação de contacto africana*. III Congresso da ABRALIN – Associação Brasileira de Linguística. Universidade Federal do Rio de Janeiro, 13-15 de Março.
- 2004. The development of variable NP plural agreement in a restructured african variety of portuguese. In Geneviève Escure & Armin Schwegler (eds.), *Creoles, contact, and language change: linguistic and social implications*, vol. 27, 97-126. Amsterdam: John Benjamins.
- 2009. A concordância de número. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 269-293. Salvador: EDUFBA.
- Baxter, Alan Norman & Dante Lucchesi. 1997. A relevância dos processo de pidginização e criouliização na formação da língua portuguesa no Brasil. *Estudos Lingüísticos e Literários*, 19. 65-84.
- 1998. Un paso más hacia la definición del pasado criollo del dialecto afro-brasileiro de Helvecia (Bahia). In Klaus Zimmerman. (ed.), *Actas del congreso internacional sobre lenguas criollas de base española y portuguesa*, 119-141. Berlin: Instituto Ibero-Americano.

- Baxter, Alan Norman, Dante Lucchesi & Maximiliano Guimarães. 1997. Gender agreement as a “decreolizing” feature of afro-brazilian dialect of Helvecia. *Journal of Pidgin and Creole Languages*, 12(1).1-57.
- Baxter, Alan Norman & Norma da Silva Lopes. 2004b. *Variação no uso de artigos: uma comparação*. Estudos sobre a variação no português brasileiro, XX Jornada Nacional de Estudos Linguísticos, GELNE, Universidade Federal da Paraíba, João Pessoa.
- 2009. O artigo definido. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 319-329. Salvador: EDUFBA.
- Bayley, Robert. 1994. Interlanguage variation and the quantitative paradigm: past tense marking in chinese-english. In Elaine E. Tarone, Susan M. Gass & Andrew D. Cohen (eds.), *Research methodology in second-language acquisition*, 157-181. Hillsdale: Lawrence Erlbaum Associates.
- 1996. Competing constraints on variation in the speech of adult chinese learners of english», in Robert Bayley & Dennis R. Preston (eds.), *Second language acquisition and linguistic variation*, 97-120. Amsterdam: John Benjamins.
- Bayley, Robert & Dennis R. Preston (eds.). 1996. *Second language acquisition and linguistic variation*. Amsterdam/Philadelphia: John Benjamins.
- Bean, Martha & Constance Gergen. 1990. Individual variation in fossilized interlanguage performance. In Hartmut Burmeister & Patricia L. Rounds (eds.), *Variation in second language acquisition. Proceedings of the Tenth Second Language Research Forum*, vol. 1, 205-219. Eugene, OR: Department of Linguistics, University of Oregon.
- Beard, Robert. 1995. *Lexeme-morpheme base morphology*. Albany, N.Y.: State University of New York Press.
- Bechara, Evanildo. 1999 [1928]. *Moderna gramática portuguesa*, 37<sup>a</sup> ed. (revista e ampliada). Rio de Janeiro: Ed. Lucerna.
- Beebe, Leslie M. 1980. Sociolinguistic variation and style shifting in second language acquisition. *Language Learning*, 2. 433-448.
- 1988 (ed.). *Issues in second language acquisition: multiple perspectives*. New York, NY: Newbury House.
- Beck, Maria-Luise. 1998. L2 acquisition and obligatory head movement: english-speaking learners of german and the local impairment hypothesis. *Studies in Second Language Acquisition*, 20. 311-348.
- Béjar, Susana. 2003 *Phy-syntax. A theory of agreement*. Toronto: University of Toronto. Dissertação de doutoramento.
- Belazi, Heidi M., Edward J. Rubin & Almeida Jacqueline Toribio. 1994. Code switching and x-bar theory: the functional head constraint. *Linguistic Inquiry*, 25(2). 221–237.
- Belleti, Adriana. 1990. *Generalized verb movement*. Torino: Rosenberg & Sellier.
- Bender, Gerald. 1980. *Angola sob domínio português: mito e realidade*. Lisboa: Livraria Sá da Costa.
- Bentivoglio, Paola A. 1987. A variação nos estudos sintáticos. In *Estudos linguísticos XIV. Anais e Seminários do GEL*, 2-29. Campinas.
- Bentley, William Holman. -1967 [1887]. *Dictionary and grammar of the kongo language as spoken at San Salvador: the ancient capital of the old Kongo Empire, West Africa/compiled and prepared for the Baptist Mission on the Kongo River, West Africa, by the Rev. W. Holman Bentley*. London: Baptist Missionary Society and Trübner.

- Berg, Thomas & Ulrich Schade. 1992. The role of inhibition in a spreading-activation model of language production. *Journal of Psycholinguistic Research*, 21. 405-462.
- Berlinck, Rosane de Andrade. 1989. A construção V SN no português do Brasil: uma visão diacrônica do fenômeno da ordem. In Fernando Tarallo (org.), *Fotografias sociolinguísticas*, 95-112. Campinas: Pontes.
- Bernstein, Judy B. 1993. *Topics in the syntax of nominal structure across romance*. New York, NY: City University. Dissertação de doutoramento.
- Bhattacharjya, Dwijen. 2007. Nagamese (restructured assamese). In John Holm & Peter L. Patrick (eds.), *Comparative creole syntax: parallel outlines of 18 creole grammars*, 237-254. Westminster Creolistic Series 7. Plymouth, UK: Battlebridge Publications.
- Biase, Bruno di & Satomi Kawaguchi. 2002. Exploring the typological plausibility of processability theory: language development in italian second language and japanese second language. *Second Language Research*, 18. 274-302.
- Biberauer, Theresa & Ian Roberts. 2005. Changing EPP-parameters in the history of english: accounting for variation and change. *English Language and Linguistics*, 9(1). 1-42.
- Biberauer, Theresa & Marc Richards. 2006. True optionality: when the grammar doesn't mind. In Cedric Boeckx (ed.), *Minimalist theorizing*, 35-67. Amsterdam: John Benjamins.
- Bickerton, Derek. 1975. *Dynamics of a creole system*. Cambridge, MA: Cambridge University Press.
- 1981. *Roots of language*. Ann Arbor: Karoma.
  - 1983. Creole languages. *Scientific American*, 249(8). 116–122
  - 1984a. The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7. 173-222.
  - 1988. Creole languages and the bioprogram. In Frederick J. Newmeyer (ed.), *Linguistics: the Cambridge survey, vol. 2 – Linguistic theory: extensions and implications*, 267-284. Cambridge, MA: Cambridge University Press.
  - 1999. How to acquire language without positive evidence: what acquisitionists can learn from creoles. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 49-75. Cambridge, MA: The MIT Press.
- Bloomfield, Leonard. 1933. *Language*. New York, NY: Holt.
- Bobaljik, Jonathan David. 1995. *Morphosyntax: the syntax of verbal inflection*. Cambridge, MA: MIT Press. Distributed by MITWPL. Dissertação de doutoramento.
- Bonvini, Emílio. 2008. Línguas africanas e português falado no Brasil. In Fiorin, José Luiz & Margarida Petter (orgs.). *África no Brasil – a formação da língua portuguesa*, 15-73. São Paulo: Ed. Contexto.
- Booji, Geert. 1994. Against split morphology. In Geert Booji & Jaap van Marle (eds.), *Yearbook of morphology 1993*, 27-50. Dordrecht: Kluwer.
- 1995. Inherent vs. contextual inflection and the split morphology hypothesis. In Geert Booji & Jaap van Marle (eds.), *Yearbook of morphology 1994*, 1-16. Dordrecht: Kluwer.
- Borer, Hagit. 1984. *Parametric syntax*. Dordrecht: Foris.
- Borer, Hagit & Kenneth N. Wexler. 1987. The maturation of syntax. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*, 123-172. Dordrecht: Reidel.

- Bortoni-Ricardo, Stella Maris. 1994. Variação lingüística e actividades de letramento em sala de aula. *Revista Internacional de Língua Portuguesa*, 12. 82-94.
- Braga, Maria Luíza. 1977. *A concordância de número no sintagma nominal no triângulo mineiro*. Rio de Janeiro: Pontifícia Universidade Católica. Dissertação de mestrado.
- 1984. Tópico e ordem vocabular. *Boletim da ABRALIN*, 6. 174-188.
- 2003. Variáveis discursivas sob a perspectiva da teoria da variação. In, Maria Cecília Mollica & Maria Luíza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 101-116. São Paulo: Contexto.
- Braga, Maria Luíza & Maria Marta Pereira Scherre. 1976. A concordância de número no SN na área urbana do Rio de Janeiro. *Anais do 1º Encontro Nacional de Linguística*. 463-474. Rio de Janeiro: Pontifícia Universidade Católica.
- Branco, Maria Inês Castelo. 1984. *Pequeno curso de língua portuguesa*. Lisboa: Fundação Calouste Gulbenkian.
- Brásio, António. 1952/1953/1954. *Monumenta missionaria africana – África Ocidental*, vols. 1, 2, 3 e 4. Lisboa: Agência Geral do Ultramar.
- Bresnan, Joan. 1982. *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Brito, Ana Maria Barros. 1996. A ordem de palavras no sintagma nominal em português numa perspectiva de sintaxe comparada – um caso particular: os Ns deverbais eventivos. In Inês Duarte & Isabel Leiria (orgs.), *Actas do Congresso Internacional do Português*, vol. 1, 81-106. Lisboa: Ed. Cilibiri.
- 2003a. Categorias sintáticas. In Maria Helena Mira Mateus *et alii*, *Gramática da língua portuguesa*, 7ª ed., cap. 11, 323-432. Lisboa: Caminho – Coleção Universitária/Série Linguística.
- 2003b. Os possessivos em português numa perspectiva de sintaxe comparada [em linha]. *Revista da Faculdade de Letras “LÍNGUAS E LITERATURAS”*, II. 495-522. [Consult. 13 Nov. 2008]. Disponível em: <http://repositorio-aberto.up.pt/bitstream/10216/7875/2/3979.pdf>.
- Brown, H. Douglas 1980. *Principles of languages learning and teaching*. Englewood Cliffs, N. J.: Prentice Hall.
- 1994. *Teaching by principles*. New Jersey, NJ: Prentice Hall.
- Bruhn de Garavito, Joyce. 1994. *Acquisition of the spanish plural: what is taught and what must be aquired*. The fourteenth Cincinnati conference on romance languages and linguistics. Cincinnati, Ohio.
- 2003. The (dis)association between morphology and syntax: the case of L2 spanish. In Silvina Montrul & Francisco Ordóñez (eds.), *Theoretical linguistics and language development in hispanic languages*, 398-417. Somerville, MA: Cascadilla Press.
- 2005. *Acquisition of the spanish plural by french L1 speakers: the role of transfer* [em linha]. [Consult. 15 Nov. 2008]. Disponível em: [www.transatlantic.uwo.ca/general/intranet/pdf](http://www.transatlantic.uwo.ca/general/intranet/pdf).
- Bruhn de Garavito, Joyce & Cristina Atoche. 2006. Variability in contact spanish: implications for second language acquisition. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 87-111. Amsterdam/Philadelphia: John Benjamins.

- Bruhn de Garavito, Joyce & Lydia White. 2002. The L2 acquisition of spanish DPs. The status of grammatical features. In Ana Teresa Pérez-Leroux & Juana M. Licerias, (eds.), *The acquisition of spanish morphology: the L1/L2 connection*, 153-176. Dordrecht: Kluwer.
- Brustad, Kristen E. 2000. *The syntax of spoken arabic: a comprehensive study of Moroccan, Egyptian, Syrian and Kuwaiti dialects*. Washington, DC: Georgetown University Press.
- Bruyn, Adrienne. 2002. The structure of the Surinamese creoles. In Eithne B. Carlin & Jacques Arends (eds.), *Atlas of the languages of the Suriname*. Leiden: KITVL Press 2000.
- Burzio, Luigi. 1986. *Italian syntax*. Dordrecht: Reidel.
- Cabral, Lisender Augusto Vicente. 2005a. *Complementos verbais preposicionados do português de Angola*, vol.1. Universidade de Lisboa. Dissertação de mestrado.  
 — 2005b. *Complementos verbais preposicionados do português de Angola - Corpus*, vol.2. Universidade de Lisboa. Dissertação de mestrado.
- Cain, Jacquelin, Márcia Weber-Olsen & Rosslyn Smith. 1987. Acquisition strategies in a first and second danguage: are they the same? *Journal of Child Language*, 14. 333-352.
- Callou, Dinah Maria Isensee. 1998. Um estudo em tempo real em dialeto rural brasileiro: questões morfossintáticas. In Sybille Große & Klaus Zimmermann (eds.), *"Substandard" e mudança no português do Brasil*, 255-272. Frankfurt am Main: TFM.
- Câmara Jr., Joaquim Mattoso. 1975. *Estrutura da língua portuguesa*, 5ª. ed.. Petrópolis: Vozes.
- Campos, Maria Henriqueta Costa & Maria Francisca Xavier. 1991. *Sintaxe e semântica do português*. Lisboa: Unv. Aberta.
- Campos, Hector. 1997. On subject extraction and the antiagreement effect in romance. *Linguistic Inquiry*, 28. 92-119.
- Cardoso, Ana Josefa Gomes. 2005. *O papel da língua materna na aquisição de uma segunda língua: o caso da língua caboverdiana (breve abordagem gramatical)* [em linha]. Lisboa: Centro de Estudos Multiculturais. Pós-Graduação em Português Língua Não-Materna. [Consult. 12 Dez. 2008]. Disponível em: [http://www.multiculturas.com/textos/lingua\\_caboverdiana\\_Ana-Josefa.pdf](http://www.multiculturas.com/textos/lingua_caboverdiana_Ana-Josefa.pdf).
- Cardoso, Boaventura. 1980. *O fogo da fala – Exercícios de estilo*. Lisboa: Ed. 70.
- Cardoso, Eduardo Augusto. 1989. *O crioulo da Ilha de S. Nicolau, de Cabo Verde*. Ministério da Educação. Instituto de Cultura e Língua Portuguesa (Portugal); Instituto Cabo-Verdiano do Livro (Cabo Verde). Lisboa: INCM.
- Carnie, Andrew. 2007[2002]. *Syntax. A generative introduction*, 2<sup>nd</sup> ed. Oxford/Malden, MA.: Blackwell.
- Carreira, António. 1979. Portuguese research on the slave trade. In UNESCO (ed.), *The african slave trade from the fifteenth to the nineteenth century. Reports and papers of the meeting of experts organized by UNESCO at Port-on-Prince, Haiti, 1978*, 251-264.
- Carstens, Vicky. 1991. *The morphology and syntax of determiner phrases in kiswahili*. Los Angeles, CA: University of California. Dissertação de doutoramento.  
 —2000. Concord in minimalist theory. *Linguistic Inquiry*, 31(2). 319-355.



- Carvalho, Hebe Macedo de. 1997a. *Concordância nominal: uma análise variacionista*. João Pessoa: UFPB. Dissertação de mestrado.
- Carvalho, Maria José Albarran. 1991. *Aspectos sintático-semânticos dos verbos locativos no português oral do Maputo*. Lisboa: Ministério da Educação, Instituto de Cultura e Língua Portuguesa.
- Casteleiro, João Malaca. 1981. *Sintaxe transformacional do adjetivo – regência das construções completivas*. Lisboa: Instituto de Investigação Científica.
- 2001 (coord.). *Dicionário da língua portuguesa contemporânea da academia das ciências de Lisboa*, vols. A-F e G-Z. Academia das Ciências de Lisboa & Fundação Calouste Gulbenkian. Lisboa: Ed. Verbo.
- Castilho, Ataliba Teixeira de. 1997a. A gramaticalização. *Estudos Lingüísticos e Literários*, 19. 25-64.
- Castro, Ana. 2004. O paralelismo entre o DP e a frase e o estatuto dos adjetivos. In *Actas do XIX Encontro Nacional da Associação Portuguesa de Linguística*, 101-112 [em linha]. Lisboa: APL. [Consult. 07 Fev. 2010]. Disponível em: <http://www.apl.org.pt/docs/actas-19-encontro-apl-2003.pdf>.
- 2007 O processamento da concordância de número interna ao DP por crianças de 2 anos falantes de português europeu. In Maria Lobo & Maria Antónia Coutinho (eds.), *XXII Encontro da Associação Portuguesa de Linguística; textos seleccionados*, 211-221. Lisboa: Ed. Colibri.
- Castro, Ana & Fernanda Pratas. 2006. Capeverdean DP-internal number agreement: additional arguments for a distributed morphology approach. In João Costa & Maria Cristina Figueiredo Silva (eds.), *Studies on agreement*, 11-24. Amsterdam/Philadelphia: John Benjamins.
- Castro, Ana & João Costa. 2003. Weak forms as X<sup>0</sup>: prenominal possessives and preverbal adverbs in portuguese. In Ana Teresa Pérez-Leroux & Yves Roberge (eds.), *Romance linguistics: theory and acquisition*, 95-110. Amsterdam/Philadelphia: John Benjamins.
- Castro, Ana & José Ferrari-Neto. 2007. Um estudo contrastivo do PE e do PB com relação à identificação de informação de número no DP. *Letras de Hoje*, 42(1). 65-76.
- Cedergren, Henrietta J. 1973. *The interplay of social and linguistic factors in Panamá*. Cornell University. Dissertação de doutoramento.
- Cedergren, Henrietta J. & David Sankoff. 1974. Variable rules: performance as a statistical reflection of competence. *Language*, 50. 333-355.
- Cerqueira, Vicente Cruz. 1993. *Primeiro o demonstrativo; depois o artigo: considerações sobre aquisição de categorias funcionais*. Campinas: Unicamp/São Paulo.
- 1994. Que faz uma criança com a marca plural? Um estudo da aquisição da concordância em português. In Mora, Jacyra & Vera Rolemberg (orgs.), *Actas do Primeiro Congresso Internacional da Associação Brasileira de Lingüística*, 119-126. UFBA. Salvador: ABRALIN – FINEP – UFBA.
- Chambers, J. K. 2002 [1995]. *Sociolinguistic theory: linguistic variation and its social significance (language in society)*, 2<sup>a</sup>. ed. Oxford/Cambridge, MA: Wiley-Blackwell.
- Chambers, J. K. & Peter Trudgill. 1980. *Dialectology*. Cambridge, MA: Cambridge University Press.

- Chambers, J. K., Peter Trudgill & Natalie Schilling-Estes (eds.). 2003 [2002]. *The handbook of language variation and change*. Oxford/Cambridge, MA: Blackwell.
- Chatelain, Heli. 2001. *Folk-tales of Angola: fifty tales, with ki-mbundu text literal english translation*. Honolulu, Hawaii: University Press of the Pacific.
- Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics*, 6. 339-405.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- 1965. *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
- 1972. *Language and mind*. New York, NY: Harcourt Brace Jovanovich.
- 1981a. *Lectures on government and binding*. Dordrecht: Foris.
- 1981b. Principles and parameters in syntactic theory. In Norbert Hornstein & David Lightfoot (eds.), *Explanation in linguistics*, 32-75. London: Longman Group.
- 1981c. Markedness and core grammar. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of markedness in generative grammar: proceedings of the 1979 GLOW Conference*, 123-146. Pisa: Scuola Normale Superiore di Pisa.
- 1986a. *Knowledge of language: its nature, origin and use*. New York, NY: Praeger.
- 1986b. *Barriers*. Cambridge, MA: The MIT Press.
- 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building*, 20. 1-52. Cambridge, MA: The MIT Press.
- 1996 [1995]. *The minimalist program*, 2<sup>nd</sup> ed. Cambridge, MA: The MIT Press.
- 1998. Minimalist inquiries: the framework. *MIT Occasional Papers in Linguistics*, 15. Republicado em 2000, in Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by step: essays in syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: The MIT Press.
- 1999a. *O programa minimalista*. Tradução portuguesa, apresentação e notas de Eduardo Paiva Raposo. Lisboa: Ed. Caminho.
- 2001a. Beyond explanatory adequacy. *Occasional Papers in Linguistics*, 20. Cambridge, MA: The MIT Press.
- 2001b. Derivation by fase. In Michael Kenstowicz (ed.), *Ken Hale: A life in linguistics* [Current studies in linguistics, 36], 1-52. Cambridge, MA: The MIT Press.
- Chomsky, Noam & Howard Lasnik. 1999 [1993]. The theory of principles and parameters. In Joachim Jacobs *et alii* (orgs.), *Syntax: international handbook of contemporary research*, 51-95. Berlin: De Gruyter.
- Cinque, Guglielmo. 1994. On the evidence of partial N-movement in the romance DP. In Guglielmo Cinque *et alii* (eds.), *Paths towards Universal Grammar. Studies in Honor of Richard S. Kayne*, 85-110. Washington, DC: Georgetown University Press.
- Clahsen, Harald & Pieter Muysken. 1986. The availability of universal grammar to adult and child learners: a study of the acquisition of german word order. *Second Language Research*, 2(2). 93–119.
- Clements, Joseph Clancy. 2009. *The linguistic legacy of spanish and portuguese: colonial expansion and language change*. New York, NY: Cambridge University Press.
- Clermont, Jean & Henrietta J. Cedergren. 1979. Les R de ma mère sont perdus dans l’air. In Pierrette Thibault (ed.), *Le français parlé: études sociolinguistiques*, 13-28. Edmonton: Linguistic Research.
- Clyne, Michael G. 1980. Triggering and language processing. *Canadian Journal of Psychology*, 34. 400-406.

- Coelho, Francisco Adolpho. 1880/1881. Os dialectos românicos ou neo-latinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa*, 2. 129-196.
- 1882. Os dialectos românicos ou neo-latinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa*, 3. 451-478.
- 1886. Os dialectos românicos ou neo-latinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa*, 6. 705-755.
- Cole, Desmond T. 1955. *An introduction to tswana grammar*. Cape Town: Longmans, Green.
- Cook, Vivian James. 1993. *Linguistic and second language acquisition*. New York, NY: Palgrave.
- 1994. The metaphor of access to universal grammar in L2 learning. In Nick C. Ellis (ed.), *Implicit and explicit learning of language*, 477-502. London: Academic Press.
- 1996. *Second language learning and language teaching*. London: Hodder Education.
- Cook, Vivian James & Mark Newson. 2007 [1988]. *Chomsky's universal grammar: an introduction*, 3rd ed. rev. Oxford: Blackwell Publishing.
- Corder, Stephen Pit. 1967. The significance of learners' errors. *International Review of Applied Linguistics*, 5(4). 161-170.
- 1971. Idiosyncratic dialects and error analysis. *IRAL*, 9 (2). 147-160.
- 1973. *Introducing applied linguistics*. Harmondsworth/Baltimore: Penguin Education.
- 1974. Error analysis. In J. P. B. Allen & Stephen Pit Corder (eds.), *Techniques in applied linguistics (the Edinburgh course in applied linguistics)*, vol. 3. 122-154. London: Oxford University Press.
- 1981b. *Error analysis and interlanguage*. London: Oxford University Press.
- Cornips, Leonie & Aafke Hulk. 2006. External and internal factors in bilingual and bidialectal language development: grammatical gender of the dutch definite determiner. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 355–377. Amsterdam: John Benjamins.
- Corrêa, Letícia Maria Sicuro. 2001. Uma hipótese para a identificação do género gramatical com particular referência para o português. *Letras de Hoje*, 125. 289-296.
- Corrêa, Letícia Maria Sicuro, Maria Cristina Lobo Name & José Ferrari-Neto. 2004. O processamento de informação de interface na aquisição de gênero e de número. *Letras de Hoje*, 39(3). 123-137.
- Corrêa, Letícia Maria Sicuro, Marina Augusto & José Ferrari-Neto. 2005. *The early processing of number agreement in the DP: evidence from the acquisition of brazilian portuguese*. 30th Boston University Conference on Language Development. Boston University, EUA, 04-06 de Novembro.
- Correia, Susana. 2005. A aquisição da rima em português europeu: ditongos e consoantes em final de sílaba [em linha]. In Inês Duarte & Isabel Leiria (orgs.), *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística*, 483-493. Lisboa: Colibri/APL. [Consult. 04 Abr. 2010]. Disponível em: <http://www.apl.org.pt/docs/actas-20-encontro-apl-2004.pdf>.
- Costa, Ana Luísa & João Costa. 2001. *O que é um advérbio?* Lisboa: Edições Colibri & Associação de Professores de Português.
- Costa, João. 1996. Adverb positioning and V movement in english: some new evidence. *Studia Linguistica*, 50. 22-34.

- Costa, João & Ana Lúcia Santos. 2004 [2003]. *A falar como os bebés. O desenvolvimento linguístico das crianças*, 2ª. ed. Lisboa: Ed. Caminho.
- Costa, João & Maria Cristina Figueiredo Silva. 2006a. Introduction. In João Costa & Maria Cristina Figueiredo Silva (eds.), *Studies on agreement*, 1-9. Amsterdam/Philadelphia: John Benjamins.
- 2006b. Nominal and verbal agreement in portuguese: an argument for distributed morphology. In João Costa & Maria Cristina Figueiredo Silva (eds.), *Studies on agreement*, 25-46. Amsterdam/Philadelphia: John Benjamins.
- Couto, Mia. 2006. *O outro pé da sereia*. Lisboa: Ed. Caminho.
- Crisma, Paola. 1997. *L'articolo nella prosainglese antica e la teoria degli articoli nulli*. Padua: University of Padua. Dissertação de doutoramento.
- Croft, William. 2000. *Explaining language change: an evolutionary approach*. Harlow, Essex: Longman Group.
- Crowley, Terry. 2008. Pidgin and creole morphology. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creoles studies*, 74-97. Oxford: Wiley-Blackwell.
- CTI de São Tomé. 1968a. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicossocial N.º. 1/68, Jan. 1968*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 3, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1968b. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicossocial N.º. 2/68, Abr. 1968*. Caixa n.º. 1, proc. AHM/DIV/208/01/07, doc. 3A, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1970. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicológica N.º. 03/70, 03 Out. 1970*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 4C, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1973. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicológica N.º. 01/73, 31 Mar. 1973*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 7A, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Cunha, Celso & Lindley Cintra. 1997 [1984]. *Nova gramática do português contemporâneo*, 13ª ed. Lisboa: Ed. João Sá da Costa.
- DeCamp, David. 1971. The study of pidgin and creole languages. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 13-43. Cambridge, MA: Cambridge University Press.
- DeGraff, Michel. 1997. Verb syntax in creolization (and beyond). In Liliane M. V. Haegeman, *The new comparative syntax*, 64-94. London: Longman Group.
- 1999a. Creolization, language change and language acquisition: a prolegomenon. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*. 1-46. Cambridge, MA: The MIT Press.
- 1999b. Creolization, language change and language acquisition: an epilogue. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 473-543. Cambridge, MA: The MIT Press.

- 2001a. Morphology in creole genesis: linguistics and ideology. In Michael J. Kenstowicz (ed.), *Ken Hale: a life in a language*, 53-121. Cambridge, MA: The MIT Press.
- Delgado-Martins, Maria Raquel. 1996. Relação fonética/fonologia: a propósito do sistema vocálico do português. In Inês Duarte & Isabel Leiria (orgs.), *Actas do Congresso Internacional sobre o Português*, vol. 1, 311-325 Lisboa: Ed. Colibri/Associação Portuguesa de Linguística.
- DeWaele, Jean-Marc & Daniel Véronique. 2001. Gender assignment and gender agreement in advanced french interlanguage: a cross-sectional study. *Bilingualism: Language and Cognition*, 4. 275-97.
- Dickerson, Lonna J. 1974. *Internal and external patterning of phonological variability in the speech of japanese learners of english: towards a theory of second language acquisition*. Urbana-Champaign: University of Illinois. Dissertação de doutoramento.
- 1975. The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, 9. 401-407.
- Dickerson, Wayne B. 1976. The psycholinguistic unity of language learning and language change. *Language Learning*, 26. 215-231.
- Duarte, Inês & Fátima Oliveira. 2003. Referência nominal. In Maria Helena Mira Mateus *et alii*, *Gramática da língua portuguesa*, 7ª ed., cap. 8, 205-242. Lisboa: Caminho – Coleção Universitária/Série Linguística.
- Duarte, Maria Eugênia L. 1993. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In Ian Roberts & Mary Kato (orgs.), *Português brasileiro: uma viagem diacrônica*, 107-128. Campinas: Editora da Unicamp.
- 1999. A sociolinguística paramétrica: perspectivas. In Dermeval da Hora & Elizabeth Christiano (orgs.), *Estudos linguísticos: realidade brasileira*, 107-114. João Pessoa: Idéia.
- Dufour, Robert & Judith F. Kroll. 1995. Matching words to concepts in two languages: a test of the concept mediation model of bilingual representation. *Memory and Cognition*, 23(2). 116-180.
- Dulay, Heidi, Marina Burt & Stephen Krashen. 1982. *Language two*. New York, NY: Oxford University Press.
- Elia, Sílvio. 1979. *A unidade linguística do Brasil*. Rio de Janeiro: Padrão.
- Ellis, Rod. 1985. *Understanding second language acquisition*. Oxford: Oxford University Press.
- 1988. Are classroom and naturalistic acquisition the same? *Studies in Second Language Acquisition*, 11. 305-328.
- 1994. *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, Nick C. & Alan Beaton. 1995. Psycholinguistic determinants of foreign language vocabulary learning. In Birgit Harley (ed.), *Lexical issues in language learning*, 107-165. The Best of Language Learning Series. Amsterdam: JohnBenjamins [Michigan: Ann Arbor].
- Ellis, Nick C. & Richard Schmidt. 1997. Morphology and longer distance dependencies: laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, 19. 145-171.
- Embick, David. 1997. *Voice and interface of syntax*. Philadelphia, PA: University of Pennsylvania. Dissertação de doutoramento.

- 2000. Features, syntax, and categories in the latin perfect. *Linguistic Inquiry*, 31(2). 185-230.
- Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry*, 32 (4). 555-595.
- Emmerich, Charlotte. 1984. *A língua de contato no Alto Xingu – origem, forma e função*. Universidade Federal do Rio de Janeiro. Dissertação de doutoramento.
- Enç, Muvet. 1991. The semantics of specificity. *Linguistic Inquiry*, 22. 1-25.
- Esquadrão de Polícia Militar N°. 2222. 1962. *História da Unidade – Companhia de Polícia Militar N°. 2222. Comando Territorial Independente de São Tomé e Príncipe. Relatório do Esquadrão de Polícia Militar N°. 2222, Jun. 1962*. Caixa n°. 1, proc. PT AHM/DIV/208/01/16, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Faria, Isabel Hub, Maria João Freitas & M Miguel. 1998. Functional categories in early acquisition of european portuguese. In Antonella Sorace, Caroline Heycock & Richard Shillcock (orgs.), *Proceedings of the Gala 97. Conference on Language Acquisition, Edinburgh, April 4-6*, 115-120. Edinburgh: University of Edinburgh.
- 2001. Interaction between prosody and morphosyntax: plurals with codas in the acquisition of european portuguese. In Barbara Höehle & Jürgen Weissenborn (eds.), *Approaches to bootstrapping: phonological, lexical, syntactic and neurophysiological aspects of early language acquisition*, 45-57. Amsterdam: John Benjamins.
- Ferguson, Charles A. 1971. Absence of copola and the notion of simplicity: a study of normal speech, baby talk, foreigner talk, and pidgins. In Dell Hymes (ed.), *Pidginization and creolization of languages*, 141-150. Cambridge, MA: Cambridge University Press.
- Fernandes, Marisa. 1996. *Concordância nominal na Região Sul*. Florianópolis: Universidade Federal de Santa Catarina. Dissertação de mestrado.
- Fernandes-Boëchat, Márcia. 1989. *Utilization of the natural learning sequence within communicative syllabus implementation*. Newcastle: University of Newcastle. Dissertação de doutoramento.
- Fernandéz-García, Marisol. 1999. Patterns in gender agreement in the speech of second language learners. In Javier Gutiérrez-Rexach & Fernando Martínez-Gil (eds.), *Advances in hispanic linguistics: papers from the 2<sup>nd</sup> Hispanic Linguistics Symposium*, 3-15. Sommerville, MA: Cascaduilla Press.
- Ferrari-Neto, José. 2003. *Reconhecimento do número gramatical e processamento da concordância de número no sintagma determinante na aquisição do português brasileiro*. Rio de Janeiro: Pontífica Universidade Católica do Rio de Janeiro – Departamento de Letras. Dissertação de mestrado.
- Ferrari-Neto, José, Letícia Maria Sicuro Corrêa & Marina R. A. Augusto. 2005. O processamento da informação de interface na aquisição do sistema de número gramatical no DP em português brasileiro. In Lúcia Maria Pinheiro Lobato et alii (orgs.), *Anais do IV Congresso Internacional da ABRALIN*, 1075-1082. Associação Brasileira de Lingüística: Universidade de Brasília [Consult. 31 Jan. 2010]. Disponível em:  
<http://www.abralin.org/publicacao/abralin2005.pdf>.
- Ferraz, Luiz Ivens. 1975. The origin and development of four creoles in the Gulf of Guinea. Reimpressão do *African Studies Quarterly Journal*. Witwatersrand University Press, Johannesburg. Pietermaritzburg, Natal: The Natal Witness (Pty).

- 1979. *The creole of São Tomé*. Reimpressão do *African Studies*, 37 (1/2). Witwatersrand University Press, Johannesburg. Pietermaritzburg, Natal: The Natal Witness (Pty).
- Ferronha, António Luís Alves. 1989. O comércio português de escravos nos séculos XV e XVI. In Luís de Albuquerque (dir.), *Portugal no mundo – povoamento e colonização do Reino de Portugal; início dos descobrimentos marítimos portugueses; o avanço no Atlântico*, vol. 1, 315-324. Lisboa: Publ. Alfa.
- Figueiredo, Carlos Filipe Guimarães. 2003. O português em Angola (algumas ocorrências em contexto literário). In Departamento de Língua e Cultura Portuguesa - Faculdade de Letras da Universidade de Lisboa (ed.), *1<sup>a</sup>.s jornadas de língua e cultura portuguesa (18/19 de Junho de 2002)*, 113-128. Lisboa: Departamento de Língua e Cultura Portuguesa – Faculdade de Letras da Universidade de Lisboa.
- 2008. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife (São Tomé). *Papia – Revista Brasileira de Estudos Crioulos e Similares*, 18. 23-43.
- 2009a. A configuração do SN do português reestruturado da comunidade de Almojarife – São Tomé [em linha]. *RCBLPE – Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, 1(1). 28-55. [Consult. 10 Jun. 2009]. Disponível em: <http://rcblpe.sftw.umac.mo/doc/A%20Configura%C3%A7ao%20do%20SN%20Plural%20do%20Portugu%C3%AAs%20Reestruturado%20da%20Comunidade%20de%20Almojarife%20-20Sao%20Tom%C3%A9.pdf>.
- 2009b. *Concordância plural variável no sintagma nominal do português reestruturado de Almojarife: motivações morfofonológicas*. Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade de Colónia, Alemanha, 11-15 de Agosto.
- 2010. *Paralelismos morfossintáticos em variedades bantu, crioulos atlânticos de base portuguesa e variedades africanas e brasileiras de português: transferência ou trajecto universal de aquisição?* 10<sup>ème</sup> Colloque International de l'Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Llacan, Campus CNRS de Villejuif, Paris, France, 01-03 de Julho.
- Firmino, Gregório. 1995. *Revisiting the 'language question' in post-colonial Africa: the case of portuguese and indigenous languages in Mozambique*. Ann Arbor: UMI Company.
- Flynn, Suzanne. 1989. The role of the head-initial/head-final parameter in the acquisition of english relative clauses by adult spanish and japanese speakers. In Susan M. Gass & Jacquelyn Schachter (eds.), *Linguistic perspectives on second language acquisition*, 89-108. Cambridge Applied Linguistic Series. New York, NY: Cambridge University Press.
- Franseschina, Florencia. 2002. Case and  $\Phi$ -feature agreement in advanced L2 spanish grammars. *EUROSLA Yearbook*, 2. 71-86. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- 2005. *Fossilized second language grammars – the acquisition of grammatical gender*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

- Freitas, Maria João & Matilde Miguel. 1998. Prosodic and syntactic interaction: the acquisition of NP functional projections in european portuguese. In Tina Cambier-Langeveld, Anikó Lipták & Michael Redford (eds.), *Proceedings of ConSOLE6*, 27-44. Leiden: SOLE.
- Fromkin, Victoria & Robert Rodman. 1993 [1974]. *Introdução à linguagem*. Coimbra: Liv. Almedina.
- Gaies, Stephen J. 1977. The nature of linguistic input in formal language learning: linguistic and communicative strategies in ESL teachers' classroom language. In H. Douglas Brown, Carlos Yorio, & Ruth Crymes (eds.), *On TESOL '77 - Teaching and learning english as a second language: trends in research and practice*, 204-212. Washington: TESOL.
- 1979. Linguistic input in first and second language learning.. In Fred R. Eckman & Ashley J. Hastings (eds.), *Studies in First and Second Language Acquisition*. Rowley, MA: Newbury House.
- Galves, Charlotte Marie Chambelland. 1993. O enfraquecimento da concordância no português brasileiro. In Mary Kato & Ian Roberts (orgs.), *Português brasileiro. Uma viagem diacrônica*, 387-403. Campinas: Editora de Campinas.
- Gann, Lewis H. & Peter Duignan, 1981. Introduction. In Lewis H. Gann & Peter Duignan (eds.), *Colonialism in Africa 1870-1960*, 1-26. London: CUP.
- Gärtner, Eberhard. 1989. Remarques sur la syntaxe du portugais en Angola e au Mozambique. In Jean-Michel Massa & Matthias Perl (eds.), *La langue portugaise en Afrique*, 29-54. Rennes: Université de Haute Bretagne.
- 2003. *Particularidades morfo-sintáticas do português não-padrão do Brasil, de Angola e de Moçambique*. Comunicação apresentada no âmbito do mestrado em Linguística Geral da FLUC. Coimbra: Universidade de Coimbra.
- Gass, Susan. M. 1988. Second language acquisition and linguistic theory: the role of language transfer. In Suzanne Flynn & Wayne O'Neil (eds.). *Linguistic theory in second language acquisition*, 384-403. Dordrecht: Kluwer Academic Publishers.
- Gass, Susan M. & Larry Selinker. 2001. *Second language acquisition: an introductory course*. Mahwah, NJ: Lawrence Earlbaum Associates, Publishers.
- Genesee, Fred. 1989. Early bilingual development: one language or two? *Journal of Child Language*, 16. 161-179.
- Gilbert, Glenn G. 1983. Transfer in second language acquisition. In Roger W. Andersen (org.), *Pidginization and creolization as language acquisition*, 207-213. Rowley, MA: Newbury House.
- Givón, Talmy. 1979. *On understanding grammar*. New York, NY: Academic Press.
- 1984. Universals of discourse structure and second language acquisition. In William E. Rutherford (ed.), *Language universals and second language acquisition*, 109-136. Amsterdam: John Benjamins.
- 1998. The functional approach to grammar. In Michael Tomasello (ed.), *The new psychology of language: cognitive and functional approaches to language structure*, 41-66. Mahwah, NJ: Lawrence Erlbaum.
- 2001. *Syntax*. Amsterdam: John Benjamins.
- Godinho, Ana Paula Batista Marques Cleto de Oliveira. 2005. *A aquisição da concordância de plural no sintagma nominal por aprendentes chineses de português língua estrangeira*. Universidade de Lisboa, Faculdade de Letras. Dissertação de doutoramento.



- 2008. *A variação da concordância de número e género no sintagma nominal na interlíngua de indivíduos adultos chineses aprendentes de português língua não materna*. Encontro sobre Português como Língua Não Materna – APL (Associação Portuguesa de Linguística). Faculdade de Letras da Universidade de Lisboa, Portugal, 11-12 de Abril.
- 2009. *Second language acquisition of portuguese noun phrase agreement by adult cantonese native speakers*. University of Macau Linguistics Seminar. University of Macau, Faculty of Social Sciences and Humanities, Department of Portuguese, Macau, 16 de Dezembro.
- Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- 1999. The emergence of the semantics of argument structure constructions. In Brian MacWhinney (ed.), *The emergence of language*, 197-212.. Mahwah, NJ: Lawrence Erlbaum.
- Gonçalves, Perpétua. 1997. Tipologia de “erros” do português oral de Maputo: um primeiro diagnóstico. In Christopher Stroud & Perpétua Gonçalves (orgs.), *Panorama do português oral de Maputo. Volume II: a construção de um banco de “erros”*, 35-67. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- 2000. A génese de línguas formadas em contextos multilingues: uma abordagem paramétrica. In Ernesto d’ Andrade, Maria Antónia Coelho da Mota & Dulce Pereira (orgs.), *Crioulos de base portuguesa (Actas do workshop sobre crioulos de base lexical portuguesa. FLUL, 29 e 30 de Abril de 1999)*, 247-257. Braga: Associação Portuguesa de Linguística.
- 2004. Towards a unified vision of classes of language acquisition and change: arguments from the genesis of mozambican african languages. *Journal of Pidgin and Creole Languages*, 19(2). 225-259.
- Gonçalves, Perpétua *et alii*. 1998. Estruturas gramaticais do português: problemas e exercícios. In Perpétua Gonçalves & Christopher Stroud (orgs.), *Panorama do português oral de Maputo –Volume III: estruturas gramaticais do português – problemas e aplicações. Cadernos de pesquisa No. 27 – Moçambique*, 36-159. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- Gonçalves, Perpétua & Christopher Stroud (orgs.). 1998. *Panorama do português oral de Maputo –Volume III: estruturas gramaticais do português – problemas e aplicações. Cadernos de pesquisa N°. 27 – Moçambique*, 36-159. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- Good, Jeff. 2003. Morphosyntactic tone raising in saramacan: the reanalysis of substrate phonology as a tonal morphology. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 105-134. Tübingen: Max Niemeyer Verlag.
- Grainger, Jonathan & Cecile Beauvillain. 1988. Associative priming in bilinguals: some limits of interlingual facilitation effects. *Canadian Journal of Psychology*, 42. 56-60.
- Green, David W. 1986. Control, activation and resource: a framework and a model for the control of speech in bilinguals. *Brain and Language*, 27. 210-223.
- Gregg, Kevin R. 1990. The variable competence model of second language acquisition, and why it isn’t. *Applied Linguistics*, 11(4). 364-383.
- Grimshaw, Jane Barbara. 2005 [1991]. *Words and structure*. Stanford: CSLI Publications.

- Grosjean, François. 1982. *Life with two languages: an introduction to bilingualism*. Cambridge, MA: Harvard University Press.
- 1985. The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development*, 6. 467-477.
- Gryner, Helena. 1990. *A variação de tempo-modo e conexão nas orações condicionais do português*. Rio de Janeiro: Faculdade de Letras - Universidade Federal do Rio de Janeiro. Dissertação de doutoramento.
- Güldemann, Tom & Tjerk Hagemeijer. 2006. Negation in the Gulf of Guinea creoles: typological and historical perspectives. *Congresso Anual da ACBPPE – Associação de Crioulos de Base Lexical Portuguesa e Espanhola*. Universidade de Coimbra, Portugal, 28-30 de Junho.
- Guy, Gregory Riordan. 1981a. *Linguistic variation in brazilian portuguese*. University of Pennsylvania. Dissertação de doutoramento.
- 1981b. Parallel variability in american dialects of spanish and portuguese. In David Sankoff & Henrietta Cedergren (orgs.), *Variation omnibus*, 85-95. Edmonton: Linguistic Research Inc.
- 1986. *Saliency and the direction of syntactic change*. University of Sydney, Cornell University. mimeo.
- 1988. Advanced Varbrul analysis. In Kathleen Ferrara *et alii* (eds), *Linguistic change and contact*, 124-136. Austin, Texas: Dept. of Linguistics, University of Texas in Austin.
- 1989. On the nature and origins of popular brazilian portuguese. *Estudios sobre Español de America y Linguística Afroamericana*, 83. 227-245. Bogotá: Instituto Caro Y Cuervo.
- 2005. A questão da criouliização no português do Brasil. In Ana Maria Stahl Zilles (org.), *Estudos de variação lingüística no Brasil e no Cone Sul*, 15-62. Porto Alegre: Editora da Universidade Federal do Rio Grande do Sul.
- Guy, Gregory Riordan & Ana Zilles. 2007. *Sociolingüística quantitativa – instrumental de análise*. Parábola Editorial: São Paulo, SP.
- Hagège, Claude. 1996. *L'enfant aux deux langues*. Paris: Ed. Odile Jacob.
- Hagemeijer, Tjerk. 1999. As ilhas de Babel: A criouliização no Golfo da Guiné. *Revista Camões*, 6. 74-88.
- 2000. Verbos e gramaticalização em são-tomense. In Ernesto d' Andrade, Maria Antónia Coelho da Mota & Dulce Pereira (orgs.), *Crioulos de base portuguesa (Actas do workshop sobre crioulos de base lexical portuguesa. FLUL, 29 e 30 de Abril de 1999)*, 111-126. Braga: Associação Portuguesa de Linguística.
- 2007. *Clause structure in santome*. Universidade de Lisboa, Faculdade de Letras. Dissertação de doutoramento.
- 2009. As língua de São Tomé e Príncipe [em linha]. *RCBLPE – Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, 1(1). 2-27. [Consult. 10 Fev. 2010]. Disponível em:  
<http://rcblpe.sftw.umac.mo/doc/As%20L%C3%ADnguas%20de%20S.%20Tom%C3%A9%20e%20Pr%C3%ADncipe.pdf>.
- Haiman, John. 1983. Icon and economic motivation. *Language*, 59. 781-819.
- Håkansson, Gisela, Manfred Pienemann & Susan Sayehli. 2002. Transfer and tipological proximity in the context of L2 processing. *Second Language Research*, 18(3). 250-273.

- Hale, Kenneth. 1988. Linguistic theory: generative grammar. In Suzanne Flynn & Wayne O'Neil (eds.), *Linguistic theory in second language acquisition*, 26-33. Dordrecht: Kluwer Academic Publishers.
- Halle, Morris. 2000. Distributed morphology: impoverishment and fission. In Jacqueline Lecarme, Jean Lowenstamm & Ur Shlonsky (eds.), *Research in afroasiatic grammar: papers from the Third Conference on Afroasiatic Languages (Sophia Antipolis, 1996)*, 125-151. Amsterdam/Philadelphia: John Benjamins.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building, 20*, 111-176. Cambridge, MA: The MIT Press.
- Hamers, Josiane F. & Michel H. A. Blanc. 2000 [1989]. *Bilinguality and bilingualism*. Cambridge, MA: Cambridge University Press.
- Han, ZhaoHong. 1998. *Fossilization. An investigation into advanced L2 learning of a typologically distant language*. Birbeck College, University of London. Dissertação de doutoramento.
- Hancock, Ian F. 1969. The Malaca creoles and their language. *Afrasian*, 3. 38-45.
- Harris, James W. 1982. Nonconcatenative morphology and spanish plurals. *Journal of Linguistics Research*, 1(1). 15-31.
- Hawkins, Roger. 1993. The selective availability of Universal Grammar in second language acquisition: a specifier-head/head-complement developmental asymmetry. *Transactions of the Philological Society*, 91. 215-245.
- 1998. *The inaccessibility of formal features of functional categories in second language acquisition*. Pacific Second Language Research Forum (PacSlrf), Tóquio.
- 2001. *Second language syntax: a generative introduction*. Oxford: Blackwell.
- Hawkins, Roger & Cecilia Yuet-hung Chan. 1997. The partial availability of universal grammar in second language acquisition: the failed functional features hypothesis. *Second Language Research*, 13(3). 187-226.
- Henriques, Isabel Castro. 1989. O ciclo do açúcar em São Tomé nos séculos XV e XVI. In Luís de Albuquerque (dir.), *Portugal no mundo – povoamento e colonização do Reino de Portugal; início dos descobrimentos marítimos portugueses; o avanço no Atlântico*, vol. 1, 264-280. Lisboa: Publ. Alfa.
- Higgs, Theodore V. & Ray Clifford. 1982. The push toward communication. In Theodore V. Higgs (ed.), *Curriculum, competence, and the foreign language teacher*, 57-79. ACTFL Foreign Language Education Series. Lincolnwood, IL: National Textbook.
- Höhle, Barbara. *et alii*. 2002. The origins of syntactic categorization for lexical elements: the role of determiners. In João Costa & Maria João Freitas (eds.), *Proceedings of the GALA 2001 Conference on Language Acquisition*, 106-111. Lisboa: Associação Portuguesa de Linguística.
- Holm, John. 1987. Creole influence on popular brazilian portuguese. In Glenn G. Gilbert (ed.), *Pidgin and creole language: essays in memory of John E. Reinecke*, 406-429. Honolulu: University of Hawaii Press.
- 1988/1989. *Pidgins and creoles*, vols. I & II. Cambridge, MA: Cambridge University Press.
- 1992. Popular brazilian portuguese: a semi-creole. In Ernesto d'Andrade & Alain Kihm (orgs.), *Actas do colóquio sobre crioulos de base lexical portuguesa*. Lisboa: Ed. Colibiri.

- 1998. *The study of semi-creoles in the 21<sup>st</sup> century*. Symposium on Pidgin and Creole Linguistics in the 21<sup>st</sup> Century: Essays at Mullenium's End. Society for Pidgin and Creole Linguistics, New York, January.
- Hopper, Paul J. 1987. Emergent grammar. In *Papers of the 13<sup>th</sup> Annual Meeting, Berkley Linguistic Society*, 139-157. Berkley: Berkley Linguistic Society.
- 1988. Emergent grammar and a priori grammar postulate. In Deborah Tannen (ed.), *Linguistics in context*, 117-134. Norwood, NJ: Ablex.
- 1998. Emergent grammar. In Michael Tomasello (ed.), *The new psychology of language: cognitive and functional approaches to language structure*, 155-175. Mahwah, NJ: Lawrence Erlbaum.
- Hopper, Paul J. & Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge, MA: Cambridge University Press.
- Hornstein, Norbert & David Lightfoot (eds.). 1981. *Explanation in linguistics: the logical problem of language acquisition*. London: Longman Group.
- Horvarth, Barbara. M. 1985. *Variation in Australian English: the sociolects of Sydney*. Cambridge, MA: Cambridge University Press.
- Houaiss, Antônio. 1985. *O português no Brasil*. Rio de Janeiro: Unibrade.
- Huang, Chen-The James. 1982. Move WH in a language without WH movement. *Linguistic Review*, 1. 41-80.
- Hudson, Joyce. 1983. *Grammatical and semantic aspects of Fitzroy Valley kriol*. SIL/AAB: Darwin.
- Hyams, Nina. 1987. The theory of parameters and syntactic development. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*, 1-22. Dordrecht: Reidel.
- Hyltenstam, Kenneth & Niclas Abrahamsson. 2003. Maturational constraints in SLA. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 539-588. Oxford: Blackwell.
- Hymes, Dell. 1971a (org.). *Pidginization and creolization of languages*. Cambridge, MA: Cambridge University Press.
- Ingram, David. 1989. *First language acquisition*. Cambridge, MA: Cambridge University Press.
- Instituto Nacional de Estatística de Moçambique. 2010. *Estatísticas de Moçambique: línguas* [em linha]. [Consult. 03 Fev. 2010]. Disponível em: [http://www.ine.gov.mz/censos\\_dir/recenseamento\\_geral/estudos\\_analise/lingua](http://www.ine.gov.mz/censos_dir/recenseamento_geral/estudos_analise/lingua).
- Instituto Nacional de Estatística de São Tomé e Príncipe. 2003a. *RGPH 2001: estado e estrutura da população de São Tomé e Príncipe* [em linha]. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe [Consult. 10 Fev. 2010]. Disponível em: [http://www.ine.st/files\\_pdf/Estrut\\_pop.pdf](http://www.ine.st/files_pdf/Estrut_pop.pdf).
- 2003b. *RGPH 2001: população de crianças e adolescentes em São Tomé e Príncipe* [em linha]. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe [Consult. 18 Fev. 2010]. Disponível em: [http://www.ine.st/files\\_pdf/CriancasAdol.pdf](http://www.ine.st/files_pdf/CriancasAdol.pdf).
- 2003/2005. *III recenseamento geral da população e da habitação de 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.

- Inverno, Liliana Cristina Coragem. 2004. Português vernáculo do Brasil e português vernáculo de Angola: reestruturação parcial vs. mudança linguística. In Mauro Fernández, Manuel Fernández-Ferreiro & Nancy Vázquez Veiga (eds.), *Los criollos de base ibérica: ACBLPE*, 201-213. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
- 2005. *Angola's transition to vernacular portuguese*. Universidade de Lisboa - Faculdade de Letras. Dissertação de mestrado.
- 2008. *The noun phrase in angolan vernacular portuguese: evidence from substrate influence*. Congresso Anual da ACBPLE – Associação de Crioulos de Base Lexical Portuguesa e Espanhola. Universidade de Lisboa, Portugal, 03-05 de Julho.
- 2009. A transição de Angola para o português vernáculo: estudo morfossintático do sintagma nominal [em linha]. *Angola do Outro Lado do Tempo...* [Consult. 02 Maio 2009]. Disponível em:  
<http://tudosobreangola.blogspot.com/2009/04/transicao-de-angola-para-o-portugues.html>.
- Jacobson, Roman. 1963 [1959]. *Essais de linguistique générale – les fondations du langage*. Paris: Les Éditions de Minuit.
- 1966 [1932]. Zur struktur des russischen verbums [On the structure of russian verbs]. In Eric P. Hamp Martin Joos, Fred W. Householder & Robert Austerlitz (eds.), *Readings in linguistics II*, 22-30. Chicago. IL: University of Chicago Press.
- Jaeggli, Osvaldo. 1984. Subject extraction and the null subject parameter. *NELS*, 14. 132-153.
- Jaeggli, Osvaldo & Kenneth J. Safir (eds.). 1989. *The null subject parameter (studies in natural language and linguistic theory)*. Dordrecht/Boston: Kluwer Academic Publishers.
- Jake, Janice L. & Carol Myers-Scotton. 1998. *How to build a creole: splitting and recombining lexical structure*. Annual Meeting of the Society for Pidgin/Creole Linguistics, New York City, 9<sup>th</sup>-10<sup>th</sup> January.
- James, Winford. 2003. The role of tone and rhythm structure in the organization of grammatical morphemes in tobagonian. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 165-192. Tübingen: Max Niemeyer Verlag.
- Jin, Young-Sun & I. Fischler. 1990. Effects of concreteness on cross-language priming of lexical decision. *Perceptual & Motor Skills*, 70. 1139-1154.
- Johnson, Daniel Ezra. s.d. *RBRUL home page* [em linha]. [Consult. 18 Maio 2008]. Disponível em:  
[http://www.ling.upenn.edu/~johnson4/Rbrul\\_manual.html](http://www.ling.upenn.edu/~johnson4/Rbrul_manual.html).
- Jon-And, Anna. 2008. *Concordância de número no sintagma nominal do português de Moçambique num contexto comparativo*. Congresso Anual da ACBPLE – Associação de Crioulos de Base Lexical Portuguesa e Espanhola. Universidade de Lisboa, Portugal, 03-05 de Julho.
- 2009. *Concordância de número no sintagma nominal do português L2 falado em Cabo Verde*. Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade de Colónia, Alemanha, 11-15 de Agosto.

- Joshi, Aravind K. 1985. How much context-sensitivity is necessary for assigning structural descriptions: tree adjoining grammars. In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language parsing: theoretical, computational, and psychological perspectives*, 206-250. Cambridge, MA: Cambridge University Press.
- Katamba, Francis. 1993. *Morphology*. New York, NY: St. Martin's Press.
- Kato, Mary Aizawa (org.). 1995. Sintaxe e aquisição na teoria de Princípios e Parâmetros. *Letras de Hoje*, 30(4).57-73.
- 1999. Os frutos de um projecto herético: parâmetros na variação intralingüística. In Dermeval da Hora & Elizabeth Christiano (orgs.), *Estudos lingüísticos: realidade brasileira*, 95-196. João Pessoa: Idéia.
- Kato, Mary Aizawa & Fernando Tarallo. 1989. Harmonia trans-sistêmica: variação inter e intra-lingüística. *Predição*, 6. 1-41.
- 1996. *Gramática do português falado: convergências*, vol. V. Campinas: Unicamp/São Paulo.
- Kawaguchi, Satomi. 2005. Argument structure and syntactic development in Japanese as a second language. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 253-298. Amsterdam/Philadelphia: John Benjamins.
- Kayne, Richard S. 1989. Facets of past participle agreement in Romance. In Paola Benincà (ed.), *Dialectal variation and the theory of grammar*, 85-103. Dordrecht: Foris.
- 2000. *Parameters and universals*. Oxford: Oxford University Press.
- Kegl, Judy, Ann Senghas & Marie Coppola. 1999. Creation through contact: sign language emergence and sign language change in Nicaragua. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 179-237. Cambridge, MA: The MIT Press.
- Kemp, William. 1979. L'histoire récente de ce que, qu'est-ce que et qu'osque à Montréal. In Pierrette Thibault (ed.), *Le français parlé: études sociolinguistiques*, 53-74. Edmonton: Linguistic Research.
- 1981. Major sociolinguistic patterns in Montréal French. In David Sankoff & Henrietta Cedergren (org.), *Variation omnibus*, 3-16. Edmonton: Linguistic Research Inc.
- Kihm, Alain. 1980. *Aspects d'une syntaxe historique: études sur le créole portugais de Guiné-Bissau*. Paris: Univ. de Paris (III). Dissertação do 3º Ciclo.
- 1994. *Kriyol syntax: the Portuguese-based creole language of Guinea-Bissau*. Amsterdam: John Benjamins.
- 2003. Inflectional categories in creole languages. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 333-363. Tübingen: Max Niemeyer Verlag.
- 2008. Creoles, markedness, and default settings: an appraisal. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creoles studies*, 411-439. Oxford: Wiley-Blackwell.
- 2010. *The fall and rise of inflectional morphology in Portuguese-related pidgin-creoles: a paradigm function morphology approach*. 10ème Colloque International de l' Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Llacan, Campus CNRS de Villejuif, Paris, France. 1-3 de Julho.
- Kiparsky, Paul. 1972. Explanation in phonology. In Stanley Peters (ed.), *Goals of linguistic theory*, 189-225. New Jersey, NJ: Prentice Hall.

- Klein, Thomas B. 2006a. Creole phonology typology: phoneme inventory size, vowel quality distinctions and stop consonant series. In Parth Bath & Ingo Plag (eds.), *The structure of creole words: segmental, syllabic and morphological aspects*, 3-21. Tübingen: Max Niemeyer Verlag.
- 2006b. Diversity and complexity in the typology of syllables in creole languages [em linha]. *Thomas B. Klein's Research Page*. [Consult. 27 Dez. 2009]. Disponível em: <http://personal.georgiasouthern.edu/~tklein/research.html>.
- Klein, Wolfgang & Clive Perdue. 1992. *Utterance structure: developing grammars again*. Amsterdam: John Benjamins.
- 1997. The basic variety (or: couldn't natural languages be much simpler?). *Second Language Research*, 30(4). 301-347.
- Klein-Andreu, Flora. 1983. Grammar in style: spanish adjective placement. In Flora Klein-Andreu (org.), *Discourse perspectives on syntax*, 143-179. New York, NY: Academic Press.
- Kleinman, Howard. 1978. The strategy of avoidance in adult second language acquisition. In William C. Ritchie (ed.), *Second language acquisition research: issues and implications*, 157-174. New York, NY: Academic Press.
- Koehn, Caroline. 1994. The acquisition of gender and number morphology within NP. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 29-51. Amsterdam: John Benjamins Publishing.
- Kolb, Bryan & Ian Q. Whishaw. 2003. *Fundamentals of human neuropsychology*. New York, NY: Worth Publishers.
- Köppe, Regina. 1994a. The DUFDE project. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 15-28. Amsterdam: John Benjamins Publishing.
- 1994b. NP-movement and subject raising. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 209-234. Amsterdam: John Benjamins Publishing.
- Kouwenberg, Silvia. 1992. From OV to VO. Linguistic negotiation in the development of berbice dutch creole. *Lingua*, 88. 263-299.
- 1994. *A grammar of berbice dutch creole*. Berlin/New York, NY: De Gruyter.
- Kouwenberg, Silvia & Peter Patrick (eds.). 2003. Reconsidering the role of second language acquisition in pidginization and creolization. Special issue of *Studies in Second Language Acquisition*, 25(2), 175-306.
- Krashen, Stephen. D. 1982. *Principles and practice in second language acquisition*. Oxford: Pergamon.
- 1985. *The input hypothesis: issues and implications*. New York, NY: Longman Group.
- 1988. *Second language acquisition and second language learning*. UK: Prentice Hall International.
- Krashen, Stephen. D. & Tracy Terrell, 1983. *The natural approach: language acquisition in the classroom*. London: Prentice Hall Europe.
- Kroch, Anthony. 1989a. Function and grammar in the history of english: periphrastic *do*. In Ralph W. Fasold & Deborah Schiffrin (eds.), *Language change and variation (Amsterdam studies in the theory and history of linguistic science, series IV: current issues in linguistic theory)*, 133-172. Amsterdam: John Benjamins.

- 1989b. Reflexes of grammar in patterns of language change. *Language, Variation and Change*, 3. 199-244.
- 2001. Syntactic change. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 699-729. Oxford/Malden, MA: Blackwell.
- 2004. *If at first you don't succeed: imperfect language learning and its implications for language change*. 2004 UNC-Chapel Hill Spring Linguistics Colloquium. University of North Carolina at Chapel Hill, Chapel Hill, NC, March 27<sup>th</sup>.
- Kroch, Anthony & Anne Taylor. 1997. Verb movement in old and middle english: dialect variation and language contact. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 298-325. Cambridge, MA: Cambridge University Press.
- Kroll, Judith F. & Alexandra Scholl. 1992. Lexical and conceptual memory in fluent and nonfluent bilinguals. In Richard Jackson Harris (ed.), *Cognitive processing in bilinguals*, 191-204. North-Holland: Elsevier Science Publishers.
- Kuno, Susumu. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry*, 5. 117-136.
- Laberge, Suzanne. 1977. *Étude de la variation des pronoms définis et indéfinis dans le français parlé à Montréal*. Université de Montréal. Dissertação de Doutorado.
- Labov, William. 1966. *The social stratification of english in New York City*. Washington DC: Center for Applied Linguistics.
- 1969. Contraction, deletion and inherent variability of the english copula. *Language*, 45. 715-762.
- 1972a. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- 1972b. *Language in the inner city. Studies in the black english vernacular*. Philadelphia: University of Pennsylvania Press.
- 1974. Estágios na aquisição do inglês *standard*. In Fonseca, Maria Stella V. & Moema F. Neves (orgs.), *Sociolingüística*, 49-85. Rio de Janeiro: Eldorado Tijuca.
- 1978. *Sociolinguistic patterns*. Oxford: Blackwell.
- 1981. What can be learned about change in progress from synchrony descriptions. In David Sankoff & Henrietta Cedergren (org.), *Variation omnibus*, 177-199. Edmonton: Linguistic Research Inc.
- 1982. Building on empirical foundations. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Perspectives on historical linguistics*, 17-92. Amsterdam/Philadelphia: John Benjamins.
- 1983. *Modelos sociolingüísticos*. Trad. José Miguel Marinas Herreras. Madrid: Ediciones Cátedra.
- 1994. *Principles of linguistic change. Volume 1: internal factors*. Oxford: Blackwell.
- Ladgeway, Adam. 2000. *A comparative syntax of the dialects of Southern Italy: a minimalist approach*. Oxford: Blackwell.
- Lado, Robert. 1957. *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor, Michigan: University of Michigan Press.
- Laman, Karl Edvard. 1936. *Dictionnaire kikongo-français avec une étude phonétique décrivant les dialectes les plus importants de la langue dite kikongo*. Brussels.
- Langacker, Ronald. 1987. *Foundations of cognitive grammar I*. Stanford, CA: Stanford University Press.
- Larsen-Freeman, Diane & Michael Long. 1991. *An introduction to second language acquisition research*. London: Longman.



- Lavandera, Beatriz R. 1984. *Variación y significado*. Buenos Aires: Hachette.
- Lefebvre, Claire. 1981. Variation in plural marking: the case of cuzco quecua. In David Sankoff & Henrietta Cedergren (eds.), *Variation omnibus*, 125-133. Edmonton: Linguistic Research Inc.
- 1986. Relexification in creole genesis revisited: the case of haitian creole. In Pieter Muysken & Norval Smith (eds.), *Substrata versus universals in creole genesis*, 15(1), 231-258. Amsterdam: John Benjamins.
- 1998. *Creole genesis and the acquisition of grammar: the case of haitian creole*. Cambridge, MA: Cambridge University Press.
- 2003. The emergence of productive morphology in creole languages: the case of Haitian creole. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 35-80. Tübingen: Max Niemeyer Verlag.
- Lefebvre, Claire & Anne-Marie Brousseau. 2002. A grammar of fongbe. Berlin: Mouton de Gruyter.
- Lefebvre, Claire, Lydia White & Christine Jourdan (eds). 2006a. *L2 acquisition and creole genesis: dialogues*. Amsterdam/Philadelphia: John Benjamins.
- 2006b. Introduction. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 1-14. Amsterdam/Philadelphia: John Benjamins.
- Leiria, Isabel. 1996. Aquisição de uma língua não materna. Um exemplo: o aspecto verbal. In Isabel Hub Faria *et alii* (coords.), *Introdução à linguística geral e portuguesa*, 71-84. Lisboa: Ed. Caminho.
- Lemle, Miriam & Anthony Julius Naro. 1977. *Competências básicas do português. Relatório final de pesquisa apresentado às instituições patrocinadoras*. Rio de Janeiro: Fundação MOBREAL e Fundação Ford.
- Lenneberg, Eric. 1967. *Biological foundations of language*. New York, NY: Wiley.
- Leung, Yan-kit Ingrid. 2003. Failed Features versus Full Transfer/Full Access in the acquisition of a third language: evidence from tense and agreement. In Juana M. Liceras *et alii* (eds.), *Proceedings of the 6<sup>th</sup> generative approaches to second language acquisition conference (GASLA 2002)*, 199-207. Somerville, MA: Cascadilla Proceedings Project.
- Levelt, Willem J. M. 1989. *Speaking: from intention to articulation*. Cambridge, MA: The MIT Press.
- Liceras, Juana M., Lourdes Díaz & Caroline Mongeon. 2000. N-drop and determiners in native and non-native spanish: more on the role of morphology in the acquisition of syntactic knowledge [em linha]. *Círculo de Lingüística Aplicada a la Comunicación*, 3. 34-62. [Consult. 19 Mar. 2010]. Disponível em: <http://www.ucm.es/info/circulo/no3/liceras.pdf>.
- Lightbown, Patsy M. & Nina Spada. 2010 [2006]. *How languages are learned*. 3<sup>rd</sup> ed. Oxford: Oxford University Press.
- Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge, MA: Cambridge University Press.
- 1991a. *How to set parameters: arguments from language change*. Cambridge, MA: The MIT Press.
- 1999a. Creoles and cues. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 431-452. Cambridge, MA: The MIT Press.

- 1999b. *The development of a language: acquisition, change and evolution*. Oxford: Blackwell.
- Lima, Carlos Henrique da Rocha. 2002 [1957]. *Gramática normativa da língua portuguesa*. 42ª ed.. Rio de Janeiro: José Olympio Ed.
- Lléo, Conxita 1997. Filler syllables, proto-articles and early prosodic constraints in Spanish and German. *Language Acquisition: knowledge, representations and processing*. Proceedings of GALA'97.
- 1998. Proto-articles in the acquisition of Spanish: interface between phonology and syntax. In Ray Fabri, Albert Ortman & Teresa Parodi (eds.), *Models of inflexion*, 175-195. Tübingen: Max Niemeyer Verlag.
- 2001. The interface of phonology and syntax: the emergence of the article in the early acquisition of Spanish and German. In Jürgen Weissenborn & Barbara Höhle (eds.), *Approaches to bootstrapping. Phonological, lexical, and neurophysiological aspects of early language acquisition*, viii. 23-44. Amsterdam: John Benjamins.
- Long, Michael H. 2003. Stabilization and fossilization in interlanguage development. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 487-536. Oxford: Blackwell.
- Longobardi, Giuseppe. 1994. Reference and proper names. *Inquiry*, 25(4). 609–665.
- Lopes, Norma da Silva. 2001. *Tópicos de concordância*. Salvador: Universidade Federal da Bahia. Dissertação de doutoramento.
- López Meirama, Belén. 1997. *La posición del sujeto en la cláusula monoactancial en español*. *Lalia, Series Maior*, 7. Santiago de Compostela: Universidade de Santiago de Compostela.
- López-Ornat, Susana 1997. What lies between a pre-grammatical and a grammatical representation? Evidence on nominal and verbal form-function mapping in Spanish from 1;7 to 2;1. In Ana Teresa Pérez-Leroux & William R. Glass (eds.), *Contemporary perspectives on the acquisition of Spanish*, 3-20. Somerville MA: Cascadia Press.
- Lorenzino, Gerardo. 1998. *The angolar creole portuguese of São Tomé: its grammar and sociolinguistic history*. New York, NY: The City University of New York. Dissertação de Doutoramento.
- Lucchesi, Dante. s.d. As duas grandes vertentes da história sociolingüística do Brasil [em linha]. In *Projecto vertentes do português rural do Estado da Bahia*. [Consult. 31 Maio 2009]. Disponível em: <http://www.vertentes.ufba.br/asduasvertentes.htm>.
- 1993. The article systems of Cape-Verde and São Tomé creole portuguese: general principles and specific factors. *Journal of Pidgin and Creole Languages*, 8(1). 81-108.
- 1994. Variação e norma: elementos para uma caracterização sociolingüística do português do Brasil. *RILP – Revista Internacional de Língua Portuguesa*, 12. 17-28.
- 1996. Variação, mudança e norma: a questão brasileira. In Suzana Alice Marcelino Cardoso. (org.), *Diversidade lingüística e ensino*, 69-80. Salvador: EDUFBA.
- 1998a. A constituição histórica do português brasileiro como um processo bipolarizado: tendências actuais de mudança nas normas culta e popular. In Sybille Große, & Klaus Zimmermann (eds.), *“Substandard” e mudança no português do Brasil*, 73-100. Frankfurt am main: TFM.

- 1999. A variação na concordância de gênero em dialectos despidginizantes e descrioulizantes do português do Brasil. In Klaus Zimmermann (ed.), *Lenguas criollas de base lexical española y portuguesa*, 477-502. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
  - 2000a. *A variação na concordância de gênero em uma comunidade de fala afro-brasileira: novos elementos sobre a formação do português popular do Brasil*. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Faculdade de Letras. Dissertação de doutoramento.
  - 2001. As duas grandes vertentes da história sociolinguística do Brasil (1500-2000). *D.E.L.T.A.: Documentação de Estudos em Lingüística Teórica e Aplicada*, vol. 17(1). 97-130.
  - 2003. O conceito de transmissão lingüística irregular e o processo de formação do português do Brasil. In Cláudia Roncarati & Jussara Abraçado (orgs.). *Português brasileiro: contacto lingüístico, heterogeneidade e história*, 272-284. Rio de Janeiro: 7Letras.
  - 2004a. Contacto entre línguas e variação paramétrica: o sujeito nulo no português afro-brasileiro. *Lingua(gem)*, 2(1). 63-92.
  - 2004b. *Sistema, mudança e linguagem*. São Paulo: Parábola.
  - 2007. Alterações no quadro dos pronomes pessoais e na aplicação da regra de concordância verbal nas normas culta e popular como evidência da polarização sociolinguística do Brasil e da relevância histórica do contato entre línguas. *Dialnet*, 19. 52-87.
  - 2008a. Aspectos gramaticais do português brasileiro afectados pelo contacto entre línguas: uma visão de conjunto. In Cláudia Roncarati & Jussara Abraçado (orgs.), *Português brasileiro II: contacto lingüístico, heterogeneidade e história*, 366-390. Niterói: Ed. UFF.
  - 2009a. História do contacto entre línguas no Brasil. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 41-73. Salvador: EDUFBA.
- Lucchesi, Dante & Alan Norman Baxter. 2006. Processos de crioulização na história sociolinguística do Brasil. In Suzana Alice Marcelino Cardoso, Jacyra Andrade Mota & Rosa Virgínia Mattos e Silva (orgs.), *Quinhentos anos de história lingüística do Brasil*, 1ª. ed., vol. 1, 163-218. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia.
- 2009. A transmissão lingüística irregular. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 101-124. Salvador: EDUFBA.
- Lucchesi, Dante & Alzira Macedo. 1997. A variação na concordância de gênero no português de contacto do Alto Xingu. *Papia – Revista Brasileira de Estudos Crioulos e Similares*, 9. 20-36.
- Lucchesi, Dante & Ilza Ribeiro. 2009. Teorias da estrutura e da mudança lingüísticas e o contacto entre línguas. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 125-153. Salvador: EDUFBA.
- Lucchesi, Dante & Silvana Araújo. 2010. A teoria da variação lingüística [em linha]. In *Vertentes do português popular do Estado da Bahia*. [Consult. 20 Out. 2010]. Disponível em:  
<http://www.vertentes.ufba.br/a-teoria-da-variacao-linguistica>.

- Luís, Ana R. 2008. Tense marking and inflectional morphology in indo-portuguese creoles. In Susanne Michaelis (ed.), *Roots of creole structures: weighting the contribution of substrates and superstrates*, 83-121. Creole Language Library. Amsterdam/Philadelphia: John Benjamins.
- Lumsden, John. S. 1987. *Syntactic features: parametric variation in the history of english*. MIT. Cambridge, MA: MIT Press. Distributed by MITWPL. Dissertação de doutoramento.
- Magalhães, Telma Moreira Vianna. 2004. Valorando traços de concordância dentro do DP. *D.E.L.T.A.*, vol. 20(1). 149-170.
- Maling, Joan M. 1976. Notes on quantifier-postposing. *Linguistic inquiry*, 7. 708-718.
- Mansouri, Fethi. 2005. Agreement morphology in arabic as a second language: typological features and their processing implications. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 117-153. Amsterdam/Philadelphia: John Benjamins.
- Marques, Irene Maria Guerra. 1985 [1983]. Algumas considerações sobre a problemática linguística em Angola. In Instituto de Cultura e Língua Portuguesa (ed.), *Actas do Congresso sobre a Situação Actual da Língua Portuguesa no Mundo*, 205-224. Lisboa: Instituto de Cultura e Língua Portuguesa.
- Marrero, Victoria & Carmen Aguirre, 2003. Plural acquisition and development in spanish. In Silvina Montrul & Francisco Ordóñez (eds.), *Linguistic theory and language development in hispanic languages: papers from the 5th Hispanic Linguistics Symposium and the 4th Conference on the Acquisition of Spanish and Portuguese*, 275-296. Somerville, MA: Cascadilla Press.
- Marroquim, Mário. 1945 [1934]. *A língua do Nordeste (Alagoas e Pernambuco)*, 2ª. ed. Rio de Janeiro: Nacional.
- Martinet, André. 1955. *Économie des changements phonétiques*. Berna: Francke.
- Martins, Custódio Cavaco. 2007. *A aquisição das noções de tempo e aspecto para aprendentes chineses de português como língua estrangeira*. Macau: Faculdade de Ciências Sociais e Humanas da Universidade de Macau. Dissertação de doutoramento.
- Mateus, Dalila Cabrita. 1999. *A luta pela independência: a formação das elites fundadoras da FRELIMO, MPLA e PAIGC*. Mem Martins: Editorial Inquérito.
- Mateus, Maria Helena Mira & Esperança Cardeira. 2007. *Norma e variação*. Coleção “O Essencial sobre Língua Portuguesa”. Lisboa: Caminho.
- Mateus, Maria Helena Mira & Alina Villalva. 2006. *O essencial sobre língua portuguesa. Linguística*. Coleção “O Essencial sobre Língua Portuguesa”. Lisboa: Caminho.
- Mather, Patrick-André. 2006. Second language acquisition and creolization – same (i-) processes, different (e-) results. *Journal of Pidgin and Creole Languages*, 21(2), IV. 231-274.
- May, Robert. 1985. *Logical form: its structure and derivation*. Cambridge, MA: The MIT Press.
- McClelland, James L. & David E. Rumelhardt. 1981. An interactive-activation model of context effects in letter perception, Part 1: an account of basic findings. *Psychological Review*, 88. 375-405.

- McCowen, Lisa Griebeling & Scott M. Alvord. 2006. Mi mamá es bonito: acquisition of spanish gender by native english speakers [em linha]. In Carol A. Klee & Timothy L. Face (eds.), *Selected proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 161-169. Somerville, MA: Cascadilla Proceedings Project. [Consult. 19 Mar. 2010]. Disponível em: <http://www.lingref.com/cpp/casp/7/paper1283.pdf>.
- McLaughlin, Barry. 1987. *Theories of second language learning*. London: Edward Arnold.
- 1989. Babes and bathwaters: how to teach vocabulary [em linha]. *Working Papers of the Bilingual Research Group*. Santa Cruz: University of California. [Consult. 29 Nov. 2009]. Disponível em: <http://people.ucsc.edu/~mclaugh/pelagius.html>.
- 1990. “Conscious” vs. “unconscious” learning. *TESOL Quarterly*, 24, 617-634.
- McLaughlin, Barry & Roberto Heredia. 1996. Information approaches to research on second language acquisition and use. In Tej K. Bathia & William C. Ritchie (eds.), *Handbook of second language acquisition*, 213-228. London: Academic Press.
- Meillet, Paul Antoine. 1906. L'état actuel des études de linguistique générale. (Leitura à Introdução do Curso de Gramática Comparativa no “Collège de France”). [Reimpresso in Meillet, Antoine. 1958. *Linguistique historique et linguistique générale*, vol. I, 1-18. Paris: Honoré Champion].
- Meisel, Jürgen M. 1994a. Getting FAT: finiteness, agreement and tense in early grammars. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 89-129. Amsterdam: John Benjamins Publishing.
- 1994b (ed.). *Bilingual first language acquisition: french and german grammatical development*. Amsterdam: John Benjamins Publishing.
- 1997. The acquisition of the syntax of negation in french and german: contrasting first and second language development. *Second Language Research*, 13, 227-263.
- Mello, Heliana Ribeiro de. 1997. *The genesis and development of brazilian vernacular portuguese*. Ann Arbor: U.M.I.
- Melzian, Hans Joachim. 1937. *A concise dictionary of the bini language of southern Nigeria*. London: Kegan Paul, Trench Trubner & Co.
- Mendes, Beatriz Correia. 1985. *Contributo para o estudo da língua portuguesa em Angola*. Centro de Linguística da Faculdade de Letras da Univ. de Lisboa.
- Mendonça, Renato de. 1936. *O português do Brasil: origens – evolução – tendências*. Rio de Janeiro: Civilização Brasileira, S.A.
- Milroy, Lesley & Pieter Muysken. 1995. Introduction: code-switching and bilingualism research. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: cross-disciplinary perspectives on code-switching*, 1-14. New York, NY: Cambridge University Press.
- Mingas, Amélia A. 2000. *Interferência do kimbundu no português falado em Lwanda*. Porto: Campo de Letras Ed.
- Mollica, Maria Cecília. 1977. *Estuda da cópia nas construções relativas em português*. Rio de Janeiro: Pontifícia Universidade Católica. Dissertação de mestrado.
- 1984. Alguns fatores da pausa entre verbo e sujeito. *Boletim da ABRALIN*, 6, 141-158.

- 2003a. Fundamentação teórica: conceituação e delimitação. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 9-14. São Paulo: Contexto.
- 2003b. Relevância das variáveis não linguísticas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 26-31. São Paulo: Contexto.
- Monteiro, Clóvis do Rego. 1933. *A linguagem dos cantadores. Segundo textos coligidos e publicados por Leonardo Mota*. Rio de Janeiro. Dissertação apresentada, em concurso, à Congregação do Colégio Pedro II.
- Montrul, Silvina A. 2004. *The acquisition of spanish: morphosyntactic development in monolingual and bilingual L1 acquisition and adult L2 acquisition*. Amsterdam: John Benjamins.
- 2006. Incomplete acquisition in bilingualism as an instance of language change. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 379-400. Amsterdam/Philadelphia: John Benjamins.
- 2008. *Incomplete acquisition in bilingualism – re-examining the age factor*. Amsterdam/Philadelphia: John Benjamins.
- Morais, Maria Aparecida Corrêa Ribeiro Torres. 1993. Aspectos diacrônicos do movimento do verbo, estrutura da frase e caso nominativo no português do Brasil. In Ian Roberts & Mary Kato (orgs.), *Português brasileiro: uma viagem diacrônica*, 263-306. Campinas: Editora da Unicamp.
- Moreno, Albertina & António Tuzine. 1998. Distribuição social de variáveis linguísticas no português oral de Maputo. In Christopher Stroud & Perpétua Gonçalves (orgs.), *Panorama do português oral de Maputo. Volume II: a construção de um banco de “erros”*, 68-89. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- Mota, Maria Antónia Coelho da. 1996. Línguas em contacto. In Isabel Hub Faria *et alii* (coords), *Introdução à linguística geral e portuguesa*, 505-533. Lisboa: Ed. Caminho.
- Mufwene, Salikoko S. 1999. On the language bioprogram hypothesis: hints for tazie. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 95-127. Cambridge, MA: The MIT Press.
- 2001. *The ecology of language evolution*. Cambridge, MA: Cambridge University Press.
- Mühlhäusler, Peter. 1981. The development of the category of number in tok pisin. In Pieter Muysken (ed.), *Generative studies in creole languages*, 35-84. Dordrecht, Holland: Foris.
- 1986. *Pidgin and creole linguistics*. Oxford: Blackwell.
- Müller, Natascha. 1994a. Gender and number agreement within DP. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 53-88. Amsterdam: John Benjamins Publishing.
- 1994b. Parameters cannot be reset: evidence from the development of COMP. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 235-269. Amsterdam: John Benjamins Publishing.
- Mussa, Alberto Baeta Neves. 1991a. *O papel das línguas africanas na história do português do Brasil*. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Dissertação de mestrado.

- 1991b. Breve nota sobre a concordância nominal de número no português do Brasil. *RILP – Revista Internacional de Língua Portuguesa*, 5/6. 72-77. Associação das Universidades de Língua Portuguesa.
- Muysken, Pieter. 1981a. Creole tense/mode/aspect systems: the unmarked case. In Pieter Muysken (ed.), *Generative studies on creole languages*, 181-199. Dordrecht: Floris.
- 1981b. Quechua causatives and logical form: a case study in markedness. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of markedness in generative grammar: proceedings of the 1979 GLOW Conference*, 443-473. Pisa: Scuola Normale Superiore di Pisa.
- 1988. Are creoles a special type of language?. In Frederick J. Newmeyer, *Language: psycholinguistic and biological aspects*. Linguistics: The Cambridge Survey, vol. 3, 285-301. Cambridge, MA: Cambridge University Press.
- 1995. Code-switching and grammatical theory. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: cross-disciplinary perspectives on code-switching*, 177-198. New York, NY: Cambridge University Press.
- 2001. The origin of creole languages: the perspective of the second language learning. In Norval Smith & Tonjes Veenstra (eds.), *Creolization and contact*, 157-173. Amsterdam/Philadelphia: John Benjamins.
- Myers-Scotton, Carol. 1997. *Dueling languages: grammatical structure in codeswitching*. New York, NY: Oxford University Press.
- 2001. Implications of abstract grammatical structure: two targets in creole formation. *Journal of Pidgin and Creole Languages*, 16. 217-273.
- 2002. *Contact linguistics: bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, Carol & Janice L. Jake. 2000a. Four types of morpheme: evidence from aphasia, code switching and second language acquisition. *Linguistics*, 38. 1053-1100.
- 2000b. Testing the 4-M model: an introduction. *International Journal of Bilingualism*, 38. 1-8.
- Name, Maria Cristina Lobo & Leticia Maria Sicuro Corrêa. 2003. Delimitação perceptual de uma classe correspondente à categoria funcional D: evidências da aquisição do português. *Fórum Lingüístico*, 3(1). 55-88.
- Naro, Anthony Julius. 1981. The social and structural dimensions of a syntactic change. *Language*, 57. 63-98.
- 2003a. Modelos quantitativos e tratamento estatístico. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 15-25. São Paulo: Contexto.
- Naro, Anthony Julius & Maria Marta Pereira Scherre. 1993. Sobre as origens do português popular do Brasil. *D.E.L.T.A.*, vol. 9, n°. especial. 437-454.
- 2000. Variable concord in portuguese: the situation in Brazil and Portugal. In John McWhorter (ed.), *Language change and language contact in pidgins and creoles*, 235-256. Amsterdam: John Benjamins.
- Naro, Anthony Julius & Miriam Lemle. 1976. Syntactic diffusion. In Sandord B. Steever *et alii* (eds.), *Papers from the parasession on dyachronic syntax*, 221-241. Chicago, IL: Chicago Linguistic Society.
- Negreiros, Almada. 1895. *O dialecto de S. Thomé. Vocabulario*.

- Nemser, William. 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics*, 9. 115-123.
- Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology: with a new foreword*. Studies in Generative Grammar, 28. Berlin: Mouton de Gruyter.
- Nina, Terezinha de Jesus Carvalho. 1980. *Concordância nominal/verbal do analfabeto da micro-região Bragantina*. Porto Alegre: Pontifícia Universidade Católica. Dissertação de mestrado.
- Noyer, Rolf. 1997. *Features, positions and affixes in autonomous morphological structure*. New York, N.Y.: Garland.
- 2006. *Distributed morphology* [em linha], 734-737. Elsevier. [Consult. 22 Nov. 2009]. Disponível em:  
<http://www.scribd.com/doc/13160480/Noyer-Distributed-Morphology>.
- Nunes, Ana Margarida Belém. 2009. *Voz e emoção em português europeu*. Aveiro: Universidade de Aveiro. Dissertação de Doutoramento.
- Ogot, Bethwell Allan. 1999. The Kongo kingdom and its neighbours. In Bethwell Allan Ogot (ed.), *A general history of Africa: Africa from the sixteenth to the eighteenth centuries*, 273-289. California: University of California Press.
- Oliveira, Ana Maria Roza. 2002. Acesso ao léxico e alternância de línguas em bilingues. *Educação & Comunicação*, 7. 86-101.
- Oliveira, Fátima. 2003. Tempo e aspecto. In Maria Helena Mira Mateus *et alii*, *Gramática da língua portuguesa*, 7ª ed., cap. 6, 127-178. Lisboa: Caminho – Coleção Universitária/Série Linguística.
- Oliveira, Marco António de & Maria Luíza Braga. 1997. On focusing sentences in brazilian portuguese. In Gregory Riordan Guy *et alii* (eds.), *Towards a social science of language*, vol. 2, 207-221. Amsterdam/Philadelphia: John Benjamins.
- Oliveira, Mário António Fernandes. 1990. *Reler África*. Coimbra: Centro de Estudos Africanos, Instituto de Antropologia, Universidade de Coimbra.
- Oller, D. Kimbrough. 1974. *Toward a theory of phonological process in first and second language learning*. Western Conference of Linguistics 1974. Seattle, Washington, USA.
- Omena, Nelize Pires de. 1978. Pronome pessoal de 3ª pessoa: suas formas variantes em função acusativa. Rio de Janeiro: Pontifícia Universidade Católica. Dissertação de mestrado.
- Paiva, Maria da Conceição A. de. 1991. Ordenação das cláusulas causais: forma e função. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Dissertação de doutoramento.
- Paiva, Maria da Conceição A. de & Maria Marta Pereira Scherre. 1999. Retrospectiva sociolingüística: contribuições do PEUL. *D.E.L.T.A.*, vol. 15, nº. especial. 201-232.
- Paolillo, John C. s.d. *R-Varb home page* [em linha]. [Consult. 14 Abr. 2008]. Disponível em:  
<http://ella.slis.indiana.edu/~paolillo/projects/varbrul/rvarb/>.
- 2002. *Analyzing linguistic variation: statistical models and methods*. Stanford, California: CSLI Publications – Center for the Study of Language and Information, Leland Stanford Junior University.
- Paradis, Michel. 1981. Neurolinguistic organization of a bilingual's two languages. In James E. Copeland & Philip W. Davis (eds.), *The Seventh LACUS Forum*, 486-494. Columbia, SC: Horn Beam Press.



- 1985. Bilingualism. In Torsten Husen & T. Neville Potlethwaite (eds.), *The International Encyclopedia of Education: Research and Studies*, vol. 1, 489-493. Oxford: Pergamon Press.
- 1987. *The assessment of bilingual aphasia*. Hilldale, NJ: Lawrence Erlbaum Associates.
- 2001 [1989]. Bilingual and polyglot aphasia. In Rita Sloan Berndt (ed.), *Handbook of neuropsychology*, 2<sup>nd</sup> ed., 69-91. Amsterdam: Elsevier.
- Parreira, Adriano. 1989. Primórdios da presença militar portuguesa em Angola. O tráfico de escravos: 1483-1643. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do Ocidente; origem e desenvolvimento da colonização*, vol. 2, 214-236. Lisboa: Publ. Alfa.
- Peck Jr., Stephen Madry. 1988. *Tense, aspect and mood in Guinea Casamane portuguese creole*. Ann Arbor: University Microfilms. Dissertação de Doutoramento.
- Pereira, Dulce. 2002/2004. Crioulos de base portuguesa [em linha]. In *História da língua portuguesa, geografia da língua portuguesa*. [Consult. 03 Dez. 2007]. Disponível em:  
<http://www.instituto-camoes.pt/cvc/hlp/geografia/crioulosdebaseport.html>.
- 2006. *O essencial sobre língua portuguesa. Crioulos de base portuguesa*. Coleção “O Essencial sobre Língua Portuguesa”. Lisboa: Caminho.
- Pereira, Dulce, Eva Arim & Nuno Carvalho. 2006. Crioulo de Cabo Verde. In ILTEC, *Diversidade linguística na escola portuguesa. Projecto Diversidade Linguística na Escola Portugues (ILTEC)* [em linha], edição em cd-rom. Lisboa: ILTEC, Ministério da Educação & Fundação Calouste Gulbenkian. [Consult. 06 Fev. 2010]. Disponível em:  
[http://www.iltec.pt/divling/\\_pdfs/linguas\\_crioulo\\_cv.pdf](http://www.iltec.pt/divling/_pdfs/linguas_crioulo_cv.pdf).
- Pérez-Pereira, Miguel. 1989. The acquisition of morphemes: some evidence from Spanish. *Journal of Psycholinguistics Research*, 18. 289-311.
- 1991. The acquisition of gender. What spanish speakers tell us. *Journal of Child Language*, 18. 571-590.
- Pfaff, Carol W. 1979. Constraints on language mixing: intrasentential code-switching and borrowing in Spanish/English. *Language*, 55. 291-318.
- Picallo, M. Carme 1991. *Nominals and nominalization in catalan*. Probus, 3(3). 279-316.
- Pienemann, Manfred. 1998. *Language processing and second language development: processability theory*. Amsterdam: John Benjamins.
- 2000. Psycholinguistic mechanism in the development of English as a second language. In Ingo Plag & Klaus P. Schneider (eds.), *Language use, language acquisition and language history: (mostly) empirical studies in honour of Rüdiger Zimmermann*, 99-118. Trier: Wissenschaftlicher Verlag Trier.
- 2005a (ed.). *Cross-linguistic aspects of processability theory*. Amsterdam/Philadelphia: John Benjamins.
- 2005b. An introduction to processability theory. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 1-60. Amsterdam/Philadelphia: John Benjamins.
- Pienemann, Manfred, Bruno di Biase & Satomi Kawaguchi. 2005. Extending processability theory. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 199-251. Amsterdam/Philadelphia: John Benjamins.

- Pienemann, Manfred & Gisela Håkansson. 2007. Full transfer vs. developmentally moderated transfer: a reply to Bohnacker. *Second Language Research*, 23(4). 485-493.
- Pienemann, Manfred & Malcom Johnson, 1987. Factors influencing the development of language proficiency. In David Nunan (ed.), *Applying second language acquisition research*, 45-141. National Curriculum Resource Center, Adult Migrant Education Program.
- Pienemann, Manfred *et alii*. 2005a. Processing constraints on L1 transfer. In Judith F. Kroll & Annette M. B. DeGroot (eds.), *Handbook of bilingualism: psycholinguistic approaches*, 128-153. New York: Oxford University Press.
- 2005b. Processability, typological distance and L1 transfer. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 85-116. Amsterdam/Philadelphia: John Benjamins.
- Pina, Hernández. 1984. *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna*. Madrid: Siglo XXI.
- Pinker, Steven. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- 1991. Rules of language. *Science*, 253. 530-535.
- Pintzuk, Susan. 1988. *VARBRUL programs*. Philadelphia, PA: University of Pennsylvania Department of Linguistics.
- Plag, Ingo. 2008a. Creoles as interlanguages: inflectional morphology. *Journal of Pidgin and Creole Languages*, 23(1). 114-135.
- 2008b. Creoles as interlanguages: syntactic structures. *Journal of Pidgin and Creole Languages*, 23(2). 307-328.
- 2008c. Creoles as interlanguages: phonology. *Journal of Pidgin and Creole Languages*, 24(1). 121-140.
- 2008d. Creoles as interlanguages: word-formation. *Journal of Pidgin and Creole Languages*, 24(2). 339-362.
- Pollock, Jean-Yves, 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20. 365-424.
- Ponte, Vanessa Maria Lôbo. 1979. *A concordância nominal de uma comunidade de Porto Alegre*. Rio Grande do Sul: Pontifícia Universidade Católica. Dissertação de mestrado.
- Poplack, Shana. 1980a. The notion of the plural in Puerto Rico spanish: competing constraints on /s/ deletion. In William Labov (ed.), *Location language in time and space*, 55-67. Filadélfia: University of Pennsylvania.
- 1980b. Sometimes I'll start a sentence in english and y termino en español: towards a typology of code-switching. *Linguistics*, 18, 581-618.
- 1981. Mortal phonemes as plural morphemes. In David Sankoff & Henrietta Cedergren (eds.), *Variation omnibus*, 59-71. Canadá: Linguistic Research Inc.
- 1989. The care and handling of a megacorpus: the Ottawa-Hull french project. In Ralph Fasold & Deborah Schiffrin (eds.), *Language change and variation*, 411-444. Amsterdam/Philadelphia: John Benjamins.
- 1992 [1979]. *Function and process in a variable phonology*. Ann Arbor: U.M.I. Dissertation Services. Dissertação de doutoramento.

- Poplack, Shana & Sali A. Tagliamonte. 1998. Nothing in context: variation, grammaticization and past time marking in nigerian pidgin english. In Philip Baker & Anand Suya (eds.), *Changing meanings, changing functions: papers relating to grammaticalization in contact languages*, 71-94. Westminister: University of Westminister Press.
- Potter, Mary C. *et alii*. 1984. Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23. 23-38.
- Preston, Dennis R. 1996a. Variationist perspectives on second language acquisition. In Robert Bayley & Dennis R. Preston (eds.), *Second language acquisition and linguistic variation*, 1-45. Amsterdam/Philadelphia: John Benjamins.
- 1996b. Variationist linguistic and second language acquisition. In Tej Bathia & William Ritchie (eds.), *Handbok of second language acquisition*, 229-265. London: Academic Press.
- Prévost, Philippe & Lydia White. 1999. Accounting for morphological variability in second language acquisition: truncation or missing inflection?. In Marc-Ariel Friedmann & Luigi Rizzi (eds.). *The acquisition of syntax*, 202–235. London: Longman Group.
- Ramos, Jânia. 1999. “Sociolingüística paramétrica” ou “Variação paramétrica”. In Dermeval da Hora & Elizabeth Christiano (orgs.), *Estudos lingüísticos: realidade brasileira*, 83-93. João Pessoa: Idéia.
- Raposo, Eduardo Paiva. 1992. *Teoria da gramática. A faculdade da linguagem*. Coleção Universitária, Série Linguística. Lisboa: Ed. Caminho.
- Révah, Israel Salvator. 1959. Comment et jusqu’à quel point les parlers brésiliens permettent-ils de reconstituer le système phonétique des parlers portugais des XVIème – XVIIème siècles? In *Actas do III Colóquio Internacionais de Estudos Brasileiros* (1957), vol. 1, 273-291. Lisboa: [s.n.].
- Ribeiro, Gregório José. 1875. *Ilmo e exmo. Senr. Marquez de Sá da Bandeira, 26 Nov. 1875*. Caixa n°. 1, São Tomé e Príncipe, doc. s.n. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Ribeiro, João. 1897. *Diccionario grammatical*. Rio de Janeiro/São Paulo: Livraria Clássica/Francisco Alves.
- Ribeiro, Sílvia Isabel do Rosário. 2009. *Compostos nominais em português: as estruturas VN, NN, NprepN e NA*. LINCUM Studies in Romance Linguistics, 64. München: Lincom GmbH.
- Rice, Mabel L. & Kenneth N Wexler. 1996. Towards tense as a clinical marker of specific language impairment in english-speaking children. *Journal of Speech and Hearing Research*, 39. 1239-1257.
- Richards, Jack C. 1971. A non-contrastive approach to error analysis. *English Language Teaching*, 25. 204-219.
- Ritter, Elizabeth. 1991. Two funcional categories in noun phrases: evidence from modern hebrew. In Susan Rothstein (ed.), *Syntax and semantics, 25: perspectives in phrase structure*, 37-62.
- Rizzi, Luigi. 1982. *Issues in italian syntax*. Dordrecht: Foris.
- 1999. Broadening the empirical basis of universal grammar models: a commentary. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 453-472. Cambridge, MA: The MIT Press.

- Roberts, Ian. 1993. *Verb and diachronic syntax: a comparative history of english and french*. Dordrecht: Floris.
- 1997. Creoles, markedness and the language bioprogram hypothesis. *Estudos Lingüísticos e Literários*, 19. 11-24.
- 1999. Verb movement and markedness. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 287-327. Cambridge, MA: The MIT Press.
- 2007. *Diachronic syntax*. Oxford: Oxford University Press.
- Rodrigues, Ada Natal. 1974. *O dialecto caipira na região da Piracicaba*. São Paulo: Ática.
- Rodrigues, José Honório. 1985. A vitória da língua portuguesa no Brasil colonial. In José Honório Rodrigues Humanidades (ed.), *História viva*, 11-48. São Paulo: Global.
- Romaine, Suzanne. 1989. *Bilingualism*. Oxford: Basil Blackwell Publishers.
- 2003. Variation. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 409-434. Oxford: Blackwell.
- Rossi, Maria Aparecida Garcia Lopes. 1993. Estudo diacrônico sobre as interrogativas do português do Brasil. In Ian Roberts & Mary Kato (orgs.), *Português brasileiro: uma viagem diacrônica*, 263-306. Campinas: Editora da Unicamp.
- Rougé, Jean-Louis. 1992. Les langues des tongas. In Ernesto d'Andrade & Alain Kihm (eds.), *Actas do Colóquio Internacional sobre Línguas Crioulas de Base Bortuguesa*, 171-175. Universidade de Lisboa: Ed. Colibiri.
- Rousseau, Pascale & David Sankoff. 1978. Advances in variable rule methodology. In David Sankoff (ed.), *Linguistic variation: models and methods*, 57-69. New York, NY: Academic Press.
- Rumelhardt, David E. & James L. McClelland. 1982. An interactive-activation model of context effects in letter perception, Part 2: the contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89. 60-94.
- 1986. On learning the past tenses of english verbs. In *MIT Press Computational Models of Cognition and Perception Series: parallel distributed processing: explorations in the microstructure of cognition, vol. 2: psychological and biological models*, 216-271. Cambridge, MA: MIT Press.
- Sabino, Robin. 1983. Plural marking in the Virgin Islands english creole in the St. Thomas-St. John Community. *Penn Review of Linguistics*, 7. 3-11.
- Salles, Heloisa Maria Moreira Lima & Maria Aparecida Curupaná da R. de Mello. s.d. *Adjetivos em -vel: formação e produtividade* [em linha]. [Consult. 20 Mar. 2010]. Disponível em:  
[http://www.ufpe.br/pgletras/Investigacoes/Volumes/Vol.18.N.2\\_2005\\_ARTIGOSWEB/HeloisaSalles-MariaAparecidaMello\\_ADJETIVOS-EM-VEL\\_Vol18-N2\\_Art11.pdf](http://www.ufpe.br/pgletras/Investigacoes/Volumes/Vol.18.N.2_2005_ARTIGOSWEB/HeloisaSalles-MariaAparecidaMello_ADJETIVOS-EM-VEL_Vol18-N2_Art11.pdf).
- Sankoff, David & Gillian Sankoff. 1973. Sample survey methods and computer assisted analysis in the study of grammatical variation. In Regna Darnell (ed.), *Canadian languages in their social context*, 7-64. Edmonton, Alberta: Linguistic Research.
- Sankoff, David & Shana Poplack. 1980. A formal grammar for code-switching. *Working in Papers*, 8. 1-55.

- Sankoff, David, Sali A. Tagliamonte & Eric Smith. 2005. *Goldvarb X: a variable rule application for Macintosh and Windows* [em linha]. Toronto: Dept. of Linguistics, University of Toronto. [Consult. 12 Fev. 2007]. Disponível em: [http://individual.utoronto.ca/tagliamonte/Goldvarb/GV\\_index.htm](http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm).
- Sankoff, Gillian, William Kemp & Henrietta J. Cedergren. 1978. The syntax of *ce que*, *qu'est-ce que* variation and its social correlates. In Roger W. Shuy & J. Fishing (eds.), *Dimension of variability and competence*. Washington, D.C.: Georgetown University Press.
- 2002 (ed.). *História geral de Cabo Verde*, vol. III. Lisboa: Instituto de Investigação Científica Tropical; Centro de Estudos de História e Cartografia Antiga/Praia: Instituto Nacional de Investigação, Promoção e Património de Cabo Verde.
- Sapir, Edward. 1971 [1921]. *A linguagem: introdução ao estudo da fala*. Rio de Janeiro: Acadêmica.
- Schachter, Jacqueline. 1974. An error in error analysis. *Language Learning*, 24. 205-214.
- 1989. Testing a proposed universal. In Susan Gass & Jacqueline Schachter (Eds.), *Linguistic perspectives on second language acquisition*, 73-88. Cambridge, MA: Cambridge University Press.
- Schang, Emmanuelle. 2000. *L'émergence des créoles portugais du Golfe de Guinée*. Nancy: Université Nancy 2, Presses Universitaires du Septentrion. Dissertação de doutoramento.
- Scherre, Maria Marta Pereira. 1978. *A regra de concordância de número no sintagma nominal em português*. Rio de Janeiro: Pontifícia Universidade Católica. Dissertação de mestrado.
- 1988. *Reanálise da concordância nominal em português*. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Faculdade de Letras. Dissertação de doutoramento.
- 1994. Aspectos da concordância de número no português do Brasil. *RILP – Revista Internacional de Língua Portuguesa*, 12. 37-49.
- 1998a [1996]. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In Giselle Machline de Oliveira e Silva & Maria Marta Pereira Scherre (orgs.), *Padrões sociolinguísticos*. 2ª. ed., 85-117. Rio de Janeiro: Ed. Tempo Brasileiro.
- 2001. Phrase-level parallelism effect on noun phrase number agreement. *Language Variation and Change*, 13. 91-107.
- Scherre, Maria Marta Pereira & Anthony Julius Naro. 2003. Análise quantitativa e tópicos de interpretação do Varbrul. In Maria C. Mollica & Maria Luíza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 147-177. São Paulo: Contexto.
- Schiffirin, Deborah. 1981. Tense variation in narrative. *Language*. 57 (1). 5-62.
- Schiffirin, Richard. M. & Walter Schneider. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84. 127-190.
- Schneider, Walter & Richard. M. Schiffirin. 1977. Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review*, 84. 1-66.

- Schuchardt, Hugo Ernst Mario. 1888. Beiträge zur Kenntnis des kriolischen romanischen I. allgemeineres über das negerportugiesische. *Zeitschrift für Romanische Philologie*, 12. 242-254.
- 1980 [1909]. Die lingua franca. In *Zeitschrift für Romanische Philologie*, 33. Traduzido in Glenn G. Gilbert (ed.), *Pidgin and creole languages: selected essays by Hugo Schuchardt*, 65-88. Cambridge, MA: Cambridge University Press.
- Schumann, John H. 1978. *The pidginization process. A model for second language acquisition*. Rowley, MA: Newbury House.
- Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research*, 12. 40-72.
- Schwegler, Armin. 1991. Negation in palenquero: synchrony. *Journal of Pidgin and Creole Languages*, 6(2). 165-214.
- Scovel, Tom. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19(3 & 4). 245-253.
- Seabra, José Augusto. 2000. A descoberta do *outro* na carta de Pêro Vaz de Caminha [em linha]. *Camões – Revista de Letras e Culturas Lusófonas*, 8. [Consult. 11 Jun. 2009]. Disponível em:  
<http://www.instituto-camoes.pt/revista/descbroutro.htm>.
- Seliger, Herbert. 1988. Psycholinguistic issues in second language acquisition. In Leslie M. Beebe (ed.), *Issues in second language acquisition: multiple perspectives*, 17-40. New York, NY: Newbury House.
- Selinker, Larry. 1972. Interlanguage. *International Review of Applied Linguistics*, 10. 209-231.
- 1992. *Rediscovering interlanguage*. London: Longman Group.
- 1996. *Fossilization: what we think we know*. London: Longman Group.
- Selinker, Larry & Usha Lakshmanan. 1993. Language transfer and fossilization: the multiple effects principle. In Susan M. Gass & Larry Selinker (eds.), *Language transfer in language learning*, 195-216. Amsterdam: John Benjamins.
- Selinker, Larry & ZhaoHong Han. 2001. Fossilization: moving the concept into empirical longitudinal study. In C. Elder *et alii* (eds.), *Experimenting with uncertainty: studies in language testing*, 276-291. Cambridge, MA: Cambridge University Press.
- Siegel, Jeff. 2004a. Morphological simplicity in pidgins and creoles. *Journal of Pidgin and Creole Languages*. 19(1). 139-162.
- 2004b. Morphological elaboration. *Journal of Pidgin and Creole Languages*. 19(2). 333-362.
- 2006. Links between SLA and creole studies: past and present. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 15-46. Amsterdam/Philadelphia: John Benjamins.
- 2008. *The emergence of pidgin and creole languages*. Oxford/New York, NY: Oxford University Press.
- Sigurðsson, Halldór Ármann. 2004. Agree and agreement: evidence from Germanic. In Werner Abraham (ed.), *Focus on Germanic Typology – Studia Typologica*, 6. 101-156. Berlin: Akademie Verlag GmbH.
- Silva, Giselle Machline de Oliveira e. 1982. *Estudo da regularidade na variável do possessivo no português do Rio de Janeiro*. Rio de Janeiro: UFRJ - Universidade Federal do Rio de Janeiro, Faculdade de Letras. Dissertação de doutoramento.

- Silva, Giselle Machline de Oliveira e & Maria Marta Pereira Scherre (orgs.). 1998 [1996]. *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado no Rio de Janeiro*, 2ª. ed. Rio de Janeiro: Ed. Tempo Brasileiro.
- Silva, Jorge Augusto Alves da. 2005. *A Concordância verbal de terceira pessoa do plural no português popular do Brasil: um panorama sociolinguístico de três comunidades do interior do Estado da Bahia*. UFBA. Dissertação de doutoramento.
- Silva, Tomé Varela da. 1998. Kiriolu: spedju di nos alma. *Kultura*, 2. 109-121.
- Silva, Vera Lúcia Paredes da. 1988. *Cartas cariocas: a variação do sujeito na escrita informal*. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Dissertação de doutoramento.
- 2003. Relevâncias das variáveis lingüísticas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 67-71. São Paulo: Contexto.
- Silva, Vítor Miguel Duarte. 2010. *A aquisição da flexão de género e número no Predicativo: um estudo com aprendentes chineses de português língua não materna*. Universidade de Macau – Faculdade de Ciências Sociais e Humanas, Departamento de Português. Dissertação de mestrado.
- Silva Neto, Serafim da. 1986 [1957]. *História da língua portuguesa*, 1ª ed. Rio de Janeiro: Presença/INL.
- Simões, Luciene Juliano. 2004. Concordância nominal de número: questões de variação e aprendizagem. *ANPOLL- Boletim Informativo*, vol. 32(1).
- Sims, William R. 1989. Fossilization and learning strategies in second language acquisition. *Minne TESOL Journal*, 7. 61-72.
- Slobin, Dan I. 1973. Cognitive prerequisites for the development of grammar. In Charles A. Ferguson & Dan I. Slobin, (eds.), *Studies of child language development*, 175-208. New York, NY: Holt, Rinehart & Winston.
- 1985. Crosslinguistic evidence for the language-making capacity. In Dan I. Slobin (ed.), *The crosslinguistic study of language acquisition*, vol. 2. *Theoretical issues*, 1157-1256. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Smith, Norval S. H. 2006. Very rapid creolization in the frame of restricted motivation hypothesis. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 49-65. Amsterdam/Philadelphia: John Benjamins.
- Snyder, William, Ann Senghas & Kelly Inman. 2001. Agreement morphology and the acquisition of N-drop in spanish. *Language Acquisition*, 9. 157-173.
- Soares, Carla. 1998. *As categorias funcionais no processo de aquisição do português europeu: estudo longitudinal da produção espontânea de uma criança de 1;2.0 aos 2;2.17 anos*. Lisboa: Universidade de Lisboa. Dissertação de mestrado.
- Soares, Carlos & François Grosjean. 1984. Bilinguals in a monolingual and a bilingual speech mode: the effect on lexical access. *Memory and Cognition*, 12. 380-386.
- Sokolik, Margaret E. & Michael E. Smith. 1992. Assignment of gender to french nouns in primary and secondary language: a connectionist model. *Second Language Research*, 8(1). 39-58.
- Spencer, Andrew. 1991. *Morphological theory: an introduction to word structure in generative grammar*. Blackwell Textbooks in Linguistics. Oxford: Basil Blackwell.

- Sprouse, Rex A. 2006. Full transfer and relexification? Second language acquisition and creole genesis. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 169-181. Amsterdam/Philadelphia: John Benjamins.
- Sprouse, Rex A. & Barbara S. Vance. 1999. An explanation for the decline of null pronouns in certain Germanic and Romance languages. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 257-284. Cambridge, MA: The MIT Press.
- Sridhar, S. N. & Kamal K. Sridhar. 1980. The syntax of psycholinguistics of bilingual code mixing. *Canadian Journal of Psychology*, 34. 407-416.
- Stenson, Nancy. 1974. Induced errors. In John Schumann & Nancy Stenson (eds.), *New frontiers of second language learning*, 179-191. Rowley, MA: Newbury House.
- Stenzel, Achim. 1994. Case assignment and functional categories in bilingual children: routes of development and implications for linguistic theory. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: French and German grammatical development*, 160-208. Amsterdam: John Benjamins Publishing.
- Sterlin, Marie-Denise. 1998. Les différentes caractéristiques de *pou* en créole haïtien. *Travaux de Recherche sur le Créole Haïtien*, 3. 1-34. Université du Québec à Montréal.
- Stolz, Thomas. 1989. Kreolische morphologie. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 42. 56-70.
- Stump, Gregory T. 2001. *Inflectional morphology*. Cambridge, MA: Cambridge University Press.
- Suárez Fernández, Mercedes. 2008. Sobre el orden de constituyentes en la lengua medieval: la posición del sujeto y el orden básico en castellano alfonsí [em linha]. *Cahiers d'Études Hispaniques Médiévales*, 31. 263-310. [Consult. 01 Abr. 2010]. Disponível em:  
[http://www.persee.fr/web/revues/home/prescript/article/cehm\\_0396-045\\_2008\\_num\\_31\\_1\\_1874](http://www.persee.fr/web/revues/home/prescript/article/cehm_0396-045_2008_num_31_1_1874).
- Tagliamonte, Sali. A. 2006. *Analysing sociolinguistic variation*. Cambridge, MA: Cambridge University Press.
- Tarallo, Fernando. 1986. *A pesquisa sociolingüística*. São Paulo: Ática.
- 1987. Por uma sociolingüística românica paramétrica: fonologia e sintaxe. *Cadernos de Lingüística e Teoria da Literatura*, 7(13). 51-83.
- 1991. Reflexões sobre o conceito de mudança lingüística. *Organon*, 1(1). 11-22.
- 1996. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. In Ian Roberts & Mary Kato, *Português brasileiro: uma viagem diacrônica*, 69-106. Campinas, Ed. da Unicamp.
- Tarallo, Fernando & Tânia Alkmin. 1987. *Falares crioulos – línguas em contacto*. São Paulo: Ed. Ática.
- Tarone, Elaine. 1978. The phonology of interlanguage. In Jack C. Richards (ed.), *Understanding second and foreign language learning*, 15-33. Rowley, MA: Newbury House.
- 1982. Sitemacity and attention in interlanguage. *Language Learning*, 32. 69-82.
- 1983. On the variability of interlanguage systems. *Applied Linguistics*, 4. 143-163.



- 1995. Some influences on the syllable structure of interlanguage phonology. In Fred R. Eckman *et alii* (eds.), *Second language acquisition theory and pedagogy*, 232-247. New Jersey, NJ: Lawrence Erlbaum Associates.
- Tarone, Elaine, Uli Frauenfelder & Larry Selinker. 1976. Systematicity variability and stability/instability in interlanguage systems. In H. Douglas Brown (ed.), *Papers in second language acquisition*, [Language Learning Special Issue (4)], 93-134. Ann Arbor: University of Michigan Press.
- Terrell, Tracy D. 1979. Final /s/ in Cuban Spanish. *Hispania*, 62. 599-612.
- Tenreiro, Francisco. 1961. *A ilha de São Tomé*. Memórias da Junta de Investigações do Ultramar, 24. 2ª. ser. Lisboa: Junta de Investigações do Ultramar.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Tieppo, Daniela Ávila. 2003. *A variação na concordância de número numa comunidade rural afro-brasileira: uma abordagem sintagmática* [em linha]. [Consult. 12 Fev. 2009]. Disponível em:  
<http://www.vertentes.ufba.br/tieppo.doc>.
- Tomás, Gil *et alii*. 2002. The peopling of São Tomé (Gulf of Guinea): origins of slave settlers and admixture with the Portuguese. *Human Biology*, 74. 397-411.
- Tovar, Antonio. 1977. *Einführung in the sprachgeschichte der iberischen halbinsel*. Tübingen: Verlag Gunter Narr.
- Unidade da Companhia de Artilharia 3376. 1973a. *História da Unidade da Companhia de Artilharia 3376. Relatório da 2ª. Repartição (Quartel General), estudo da situação n.º. 1/73*. Caixa n.º. 1, proc. PT AHM/DIV/12/08/01/21, doc. 1, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1973b. *História da Unidade da Companhia de Artilharia 3376. Relatório da 2ª. Repartição (Quartel General), estudo da situação n.º. 1/73*. Caixa n.º. 1, proc. PT AHM/DIV/12/08/01/21, doc. 10, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Vänänen, Veikko. 1981 [1957]. *Introduction au latin vulgaire*. Paris: Ed. Klincksieck.
- Vainikka, Anne. 1993/1994. Case in the development of English syntax. *Language Acquisition*, 3. 257-325.
- Valente, Padre José Francisco (C. S. Sp.). 1964. *Gramática umbundu – a língua do centro de Angola*. Lisboa: Junta de Investigações do Ultramar.
- Valette, Rebecca M. 1991. Proficiency and the prevention of fossilization – An editorial. *The Modern Language Journal*, 75(3). 325-336.
- Valois, Daniel. 1991. *The internal syntax of DP*. Los Angeles, CA: University of California. Dissertação de doutoramento.
- Vance, Barbara S. 1989. *Null subjects and syntactic change in medieval French*. Ithaca, NY: Cornell University Microfilms. Dissertação de doutoramento.
- Veado, Rosa Maria Assis. 1982. *Comportamento lingüístico do dialeto rural*. Belo Horizonte: UFMG/PROED.
- Veiga, Margareth. 2002a. *Interferências do kimbundu no português de Luanda*. Trabalho de licenciatura apresentado no âmbito da disciplina de Sintaxe e Semântica do Português. Universidade de Lisboa, Faculdade de Letras – Departamento de Língua e Cultura Portuguesa.

- Veenstra, Tonjes. 1994. The acquisition of functional categories: the creole way. In Dany Adone & Ingo Plag (eds.) *Creolization and language change*, 99-115. Tübingen: Max Niemeyer Verlag.
- 1996. *Serial verbs in saramaccan*. The Hague: Holland Academic Graphics.
- 2003. What verbal morphology can tell us about creole genesis: the case of french-related creoles. In Ingo Plag (ed.), *Phonology and morphology of creole languages*, 293-314. Tübingen: Max Niemeyer Verlag.
- Venâncio, José Carlos. 1996. *A economia de Luanda e Hinterland no século XVIII: um estudo de sociologia histórica*. Lisboa: Ed. Estampa.
- Vicente, Elena da Silva Guerra. 2006. *O quantificador flutuante todos no português brasileiro e no inglês: uma abordagem gerativa*. Brasília: Universidade de Brasília. Dissertação de doutoramento.
- Vigil, Neddy A. & John W. Oller. 1976. Rule fossilization: a tentative model. *Language Learning*, 26(2). 281-295.
- Vihman, Marilyn May. 1985. Language differentiation by the bilingual infant. *Journal of Child Language*, 12(2). 297-324.
- Vilela, Mário. 1999. A língua portuguesa em África: tendências e factos. *Africana Studia*, 1. 175-195.
- Wagner, Nalgis de Fátima. 2001. Concordância nominal: uma análise de textos orais em situações formais e informais com dados caboclos [em linha]. *Revista Virtual*, 8 – Contestado e Educação – Universidade do Contestado. [Consult. 13 Mar. 2007]. Disponível em:  
[www.cdr.unc.br/PG/.../RevistaVirtual.htm](http://www.cdr.unc.br/PG/.../RevistaVirtual.htm).
- Wei, Li. 1996. *Variation in the acquisition of morpheme types in the interlanguages of chinese and japanese learners of english as second language*. Colombia, SC: University of South Carolina. Dissertação de doutoramento.
- Weiner, Judith & William Labov, 1983. Constraints on the agentless passive. *Journal of Linguistics*, 19. 29-58.
- Weinreich, Uriel, William Labov & Marvin I. Herzog. 2006 [1968]. *Fundamentos empíricos para uma teoria da mudança lingüística*. [Trad. de Marcos Bagno], 2ª. ed. São Paulo: Parábola Editorial.
- Welmers, William. 1973. *African languages structures*. Berkeley: University of California Press.
- White, Lydia. 1989. *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.
- 1990. Second language acquisition and universal grammar. *Studies in Second Language Acquisition*, 12. 121-133.
- 1993. Universal grammar: is it a new name for old problems? In Susan M. Gass & Larry Selinker (eds.), *Language transfer in language learning*, 217-232. Amsterdam: John Benjamins.
- 2000. Second language acquisition: from initial to final state. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 130-155. Oxford: Blackwell Publishers.
- 2003. *Second language acquisition and universal grammar*. Cambridge, MA: Cambridge University Press.
- White, Lydia et alii. 2004. Gender and number agreement in nonnative spanish. *Applied Psycholinguistics*, 25. 105-133.

- Williams, Edwin B. 1986 [1938]. *Do latim ao português*, 4<sup>a</sup> ed. Rio de Janeiro: Ed. Tempo Brasileiro.
- Winford, Donald. 2003a. *An introduction to contact languages*. Blackwell Publishers.
- Wolfram, Walt. 1985. Variability in tense marking: a case for the obvious. *Language Learning*, 35. 229-253.
- 1989. Systematic variability in second-language tense marking. In Miriam R. Eisenstein (ed.), *The dynamic interlanguage: empirical studies in second language variation*, 187-197. New York, NY: Plenum Press.
- Wolfson, Nessa. 1976. Speech events and natural speech: some implications for sociolinguistic methodology. *Language in Society*, 5. 189-209.
- Wood, Johanna. 2003. *Definiteness and number: determiner phrase and number phrase in the history of english*. Arizona State University. Dissertação de doutoramento.
- Wunderlich, Dieter. 1996. Minimalist morphology: the role of paradigms. In: Geert Booij & Jaap. van Marle (eds.), *Yearbook of morphology 1995*, 93-114. Dordrecht: Kluwer Academic Publishers.
- Xueping Wei. 2008. Implication of IL fossilization in second language acquisition [em linha]. *English Language Teaching*, 1(1). 127-131. [Consult. 26 Nov. 2009]. Disponível em:  
[http://docs.google.com/viewer?a=v&q=cache:N8YbDix86X0J:www.ccsenet.org/journal/index.php/elt/article/view/523/504+selinker+fossilization:+what+we+think+we+know&hl=pt-PT&gl=pt&pid=bl&srcid=ADGEESgC2Rr9zjciPKmDNYqRrW2Q9bkiI3Scw9PmiIjuYNVKGByBZkI2AEwoMUhPaLh-MfA85iJ80ZUsBvMbjOWTYBbQOeCexbBKUeodNsLRMe\\_gWybhDnQYLTm\\_OUgwExjLYzEwhaQ9n&sig=AHIEtBR-VpSkCOB4VioqMopusbP583mH2Q](http://docs.google.com/viewer?a=v&q=cache:N8YbDix86X0J:www.ccsenet.org/journal/index.php/elt/article/view/523/504+selinker+fossilization:+what+we+think+we+know&hl=pt-PT&gl=pt&pid=bl&srcid=ADGEESgC2Rr9zjciPKmDNYqRrW2Q9bkiI3Scw9PmiIjuYNVKGByBZkI2AEwoMUhPaLh-MfA85iJ80ZUsBvMbjOWTYBbQOeCexbBKUeodNsLRMe_gWybhDnQYLTm_OUgwExjLYzEwhaQ9n&sig=AHIEtBR-VpSkCOB4VioqMopusbP583mH2Q).
- Young, Richard. 1991. *Variation in interlanguage morphology*. New York, NY: Peter Lang Publishing, Inc.
- 1993. Functional constraints on variation in interlanguage morphology. *Applied Linguistics*, 14(1). 76-97.
- Young, Richard & Robert Bayley. 1996. VARBRUL analysis for second language acquisition Research. In Robert Bayley & Dennis R. Preston (eds.), *Second language acquisition and linguistic variation*, 253-306. Amsterdam/Philadelphia: John Benjamins.
- Yip, Virginia, William Rutherford & Harald Clahsen (eds.). 1995. Interlanguage and learnability: from chinese to english. *Language acquisition & language disorders*, vol. 11. Amsterdam: John Benjamins.
- Zau, Domingos Gabriel Ndele. 2001. *Desvios sintáticos no português falado em Luanda*. Trabalho de licenciatura apresentado no âmbito da disciplina de Sintaxe e Semântica do Português. Universidade de Lisboa, Faculdade de Letras – Departamento de Língua e Cultura Portuguesa.
- Zhang, Yanyin. 2005. Processing and formal instruction in the L2 acquisition of five chinese grammatical morphemes. In Manfred Pienemann (ed.), *Cross-linguistic aspects of processability theory*, 155-177. Amsterdam/Philadelphia: John Benjamins.
- Zuengler, Jane. 1989. Assessing an interaction-based paradigm: how accommodative should we be? In Miriam R. Eisenstein (ed.), *The dynamic interlanguage: empirical studies in second language variation*, 49-67. New York, NY: Plenum Press.

## II. Bibliografia pertinente consultada:

- Abaurre, Maria Bernardete M. & Ângela C. S. Rodrigues (orgs.). 2002. *Gramática do português falado: novos estudos descritivos*, vol. VIII. Campinas: Unicamp/São Paulo.
- Aboh, Enoch Oladé. 2006. The role of the syntax-semantics interference in language transfer. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 221-252. Amsterdam/Philadelphia: John Benjamins.
- Afonso, A. Martins. s.d. *Curso de história da civilização portuguesa*, 7ª. ed. Porto: Porto Ed.
- Agheyisi, Rebecca Nogieru. 1990. *A grammar of edo*. Unesco.  
—1991. The tense systems of nigerian languages and english. In Okon E. Essien (ed.), *Afrikanistische arbeitspapiere*, 27. 11-42.
- Albert, Martin L. & Loraine K. Obler. 1978. *The bilingual brain*. New York, NY: Academic Press.
- Albuquerque, Luís de. 1989a. Os primeiros contactos com os povos da Guiné. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do Ocidente; origem e desenvolvimento da Colonização*, vol. 2, 82-98. Lisboa: Publ. Alfa.  
—1989b. A colonização de São Tomé e Príncipe: os capitães do século XV. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do ocidente; origem e desenvolvimento da colonização*, vol. 2, 171-197. Lisboa: Publ. Alfa.  
—1989c. Balanço da expansão e da colonização portuguesas. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do ocidente; origem e desenvolvimento da colonização*, vol. 2, 277-316. Lisboa: Publ. Alfa.
- Alexiadou, Artemis & Elena Agnostopoulou. 1996. Symmetries, asymmetries and the role of agreement. 19<sup>th</sup> GLOW Colloquium. Athens, Greece, 17.4.1996. [*GLOW Newsletter*, 36. 12-13].  
—1998. Parametrizing agr: word order, verb movement and EPP checking. *Natural Language and Linguistic Theory*, 16(3). 491-539.
- Ali, M. Said. 1971 [1921-1923]. *Gramática histórica da língua portuguesa*, 7ª. ed. Rio de Janeiro: Ed. Melhoramentos.
- Almeida, Eugénio Luís da Costa. 1991. *São Tomé e Príncipe – notas para um estudo sócio-político* [em linha]. [Consult. 14 Jun. 2006]. Disponível em: <http://elcalmeida.home.sapo.pt/Naopublicados/STPrincipe.htm>.
- Almendra, Maria Ana & José Nunes de Figueiredo. 1997. *Compêndio de gramática latina*. Porto: Porto Ed.
- Amaral, Ilídio do. 1964. *Santiago de Cabo Verde. A terra e os homens*. Lisboa: Junta de Investigações do Ultramar.
- Andersen, Roger W. 1983b. Introduction: A language acquisition interpretation of pidginization and creolization. In Roger W. Andersen (ed.), *Pidginization and creolization as language acquisition*, 1-56. Rowley, MA: Newbury House.
- Andrade, Ernesto d'. 1977. *Aspects de la phonologie (générative) du portugais*. Lisboa: INIC.  
—1994. *Temas de fonologia*. Lisboa: Ed. Colibri.
- Andrade, Ernesto d' & Alain Kihm (orgs.). 1992. *Actas do Colóquio Internacional sobre Línguas Crioulas de Base Bortuguesa*. Lisboa: Ed. Colibri.

- Andrade, Ernesto d', Maria Antónia Coelho da Mota & Dulce Pereira (orgs.). 2000. *Crioulos de base portuguesa (Actas do workshop sobre crioulos de base lexical portuguesa. FLUL, 29 e 30 de Abril de 1999)*. Braga: Associação Portuguesa de Linguística.
- Andrade, Maria José de Souza. 1988. *A mão de obra escrava em Salvador 1811-1860*. São Paulo: Ed. Corrupio.
- Areas, Eduardo Kenedy. s.d. Fronteiras nebulosas: sintaxe e discurso [em linha]. [Consult. 09 Mar. 2010]. Disponível em: <http://www.filologia.org.br/cluerj-sg/anais/iii/completos%5Cmesas%5CM%2014%5Ceduardo%20Kenedy%20Areas.pdf>.
- Arends, Jacques. 2001. Simple grammars, complex languages. *Linguistic Typology*, 5(2/3). 180-182.
- Arranja, Álvaro. 1993. A escravatura na crónica da Guiné. *Histórias*, Ano XV, 168. 62-71.
- Atkinson, David, Mercedes Bengochea & Sandi Michele de Oliveira. 2010. Sociolinguistics in Spain and Portugal. In Martin J. Ball (ed.), *The Routledge handbook of sociolinguistics around the world*, 341-358. London/New York, NY: Routledge.
- Atkinson, Martin. 1992. *Children's syntax: an introduction to principles and parameters theory*. Oxford/Cambridge, MA: Blackwell Publishing.
- Baker, Carl Lee. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry*, 10(4). 533-581.
- Baker, Mark C. 2003. *Lexical categories: verbs, nouns and adjectives*. Cambridge, MA: Cambridge University Press.
- Baker, Philip. 1991. Causes and effects. *Journal of Pidgins and Creole Languages*, 6(2). 267-278.
- 1997. Directionality in pidginization and creolization. In Arthur K. Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 91-109. Amsterdam/Philadelphia: John Benjamins.
- 2000. Theories of creolization and the degree and nature of restructuring. In Ingrid Neumann-Holzschuh & Edgar W. Schneider. (eds.), *Degrees of restructuring in creole languages*. Creole Languages Library (CLL), vol. 22, 41-63. University of Regensburg. Amsterdam/Philadelphia: John Benjamins.
- Bakker, Peter. 1995. Pidgins. In Jacques Arends, Pieter Muysken & Norval Smith (eds.), *Pidgins and creoles: an introduction*, 25-39. Amsterdam/Philadelphia: John Benjamins.
- Ballog-Wen-Mewuda, Joseph Bato'ora. 1989a. Os entrepostos móveis e as relações com os povos circunvizinhos. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do ocidente; origem e desenvolvimento da colonização*, vol. 2, 99-111. Lisboa: Publ. Alfa.
- 1989b. A instalação de fortalezas na costa africana. Os casos de Arguim e da Mina. Comércio e contactos culturais In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do ocidente; origem e desenvolvimento da colonização*, vol. 2, 137-149. Lisboa: Publ. Alfa.
- Baltin, Mark & Chris Collins (eds.). 2000. *The handbook of contemporary syntactic theory*. Oxford/Malden, MA: Blackwell Publishers, Ltd (edition for sale in the mainland territory of the People's Republic of China only).

- Baptista, Marlyse. 2005. New directions in pidgin and creole studies. *Annual Review of Anthropology*, 34. 33-42.
- Barbeiro, Luís Filipe. 1986. *Estrutura silábica do português. O papel da sílaba na análise dos processos fonológicos e fonéticos*. Lisboa: Faculdade de Letras da Universidade de Lisboa. Dissertação de mestrado.
- Barbosa, Pilar. 1995. *Null subjects*. Cambridge, MA: The MIT Press. Dissertação de doutoramento.
- 2002. *A propriedade do sujeito nulo e o princípio da projecção alargado* Lisboa: Ed. Colibri.
- Barlow, Michael. 2000. *Monoconpro 2.0*. Houston: Athelstan.
- Barras, Claude. 2002. *Transcriber* [online]. [Consult. 23 Fev. 2008]. Disponível em: <http://www.etca.fr/CTA/gip/Projets/Transcriber/>.
- Barros, João de. 1971 [1540]. *Gramática da língua portuguesa; cartinha, gramática, diálogo em louvor da nossa linguagem e diálogo da viciosa vergonha*. Reprod. facsim., leit., introd. e anot. por Maria Leonor Carvalhão Buescu. Lisboa: Faculdade de Letras.
- Bastide, Roger. 1980 [1959]. *Brasil, terra de contrastes*, 10<sup>a</sup>. ed. Prefácio de Paulo Duarte. São Paulo: Difusão Européia do Livro.
- Baxter, Alan Norman. 1995. Transmissão geracional irregular na história do português brasileiro – divergências nas vertentes afro-brasileiras. *RILP – Revista Internacional de Língua Portuguesa*, 14. Associação das Universidades de Língua Portuguesa. 72-90.
- 1996. Línguas pidgin e crioulas. In Isabel Hub Faria *et alii* (coords), *Introdução à linguística geral e portuguesa*, 535-549. Lisboa: Ed. Caminho.
- Baxter, Alan Norman & Norma da Silva Lopes. 2004a. A concordância do SN plural no português afrobrasileiro do século XIX. III Encuentro de la Asociación de Criollos de Base Léxica Portuguesa y Española, Universidade da Coruña, Galícia, Espanha, 26-27 de Junho de 2003.
- 2005. *O artigo definido em variação com zero no SN de referência específica: reestruturação em três variedades de português*. XIV Congresso da ALFAL – Asociación de Linguística e Filología de América Latina. Monterrey, 17-21 de Outubro.
- 2006. Bare definite reference NPs in an afro-brazilian portuguese dialect. *Revista Internacional de Linguística Iberoamericana*, 4(1). 55-69.
- Belleti, Adriana. 2000. Agreement projections. In Mark Baltin & Chris Collins (eds.). *The handbook of contemporary syntactic theory*, 484-510. Oxford/Malden, MA: Blackwell (edition for sale in the mainland territory of the People's Republic of China only).
- Bentley, William Holman. 1895. *Appendix to the dictionary and grammar of the kongo language*. London: Baptist Missionary Society and Kegan Paul, Trench Trübner.
- Bentzen, Kristine. 2003. V-to-I movement in the absence of morphological cues: evidence from adult and child northern norwegian. In Anne Dahl, Peter Svenonius & Marit R. Westgaard (eds.), *Proceedings of the 19<sup>th</sup> scandinavian conference of linguistics – Acquisition*, vol. 31(3), 573-588.

- Bernstein, Judy B. 2000. The DP hypothesis: identifying clausal properties in the nominal domination. In Mark Baltin & Chris Collins (eds.), *The handbook of contemporary syntactic theory*, 536-561. Oxford/Malden, MA: Blackwell Publishers, Ltd (edition for sale in the mainland territory of the People's Republic of China only).
- Berwick, Robert. C. 1982. *Locality principles and the acquisition of syntactic knowledge*. Cambridge, MA: The MIT Press. Dissertação de doutoramento.
- Berwick, Robert. C. & Amy S. Weinberg. 1984. *The grammatical basis of linguistic performance: language use and acquisition*. Cambridge, MA: The MIT Press.
- Bhatt, Parth & Ingo Plag (eds.). 2006. *The structure of creole words: segmental, syllabic and morphological aspects*. Tübingen: Max Niemeyer Verlag.
- Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge, MA: Cambridge University Press.
- Bickerton, Derek. 1977. Pidginization and creolization: language acquisition and language universal. In Albert Valdman. (ed.), *Pidgin and creole linguistics*, 49-69. Bloomington, IN: Indiana University Press.
- 1984b. Creole still is king: response to commentary. *Behavioral and Brain Sciences*, 7. 212-218.
- 1986a. Beyond roots: progress or regress? *Journal of Pidgin and Creole Languages*, 1. 135-140.
- 1986b. Beyond roots: the five-year test. *Journal of Pidgin and Creole Languages*, 1. 225-232.
- 1989. The lexical learning hypothesis and the pidgin-creole cycle. In Martin Pütz & René Dirven (eds.), *Wheels within wheels: papers of the Duisburg symposium on pidgin and creole languages*, 11-31. Frankfurt: Verlag Peter Lang.
- 1991. On the supposed 'gradualness' of creole development, *Journal of Pidgin and Creole Languages*, 6. 25-58.
- 1992. The sociohistorical matrix of creolization. *Journal of Pidgin and Creole Languages*, 7. 307-318.
- 2003. *Refuting the bioprogram is easy...* Society for pidgin and creole linguistics meeting. University of Hawai'i, Honolulu, 14-17 de Agosto.
- Biderman, Maria Tereza Camargo. 2001. *Teoria linguística – linguística quantitativa e computacional*. São Paulo: Martins Fontes.
- Bloomfield, Leonard. 1933. *Language*. New York, NY: Holt.
- Boléo, Manuel de Paiva. 1943. *Brasileirismos (problemas de método)*. Coimbra: Coimbra Ed.
- 1955. *Unidade e variedade da língua portuguesa*. Lisboa: Faculdade de Letras da Universidade de Lisboa.
- 1974. *Estudos de linguística portuguesa e românica*. Coimbra: Faculdade de Letras da Universidade de Coimbra.
- Bona, Alessandra Herlin de. 2007. *O papel do feedback corretivo na reversão da fossilização em aprendizes adultos*. Pelotas, RS: Faculdade de Letras da Universidade Católica de Pelotas – Dissertação de mestrado.
- Bonvini, Emílio. 1994. Angola: language situation. In R. E. Asher (ed.), *The encyclopedia of language and linguistics*, 127-128. New York/Oxford-Seoul/Tokyo: Pergamon Press.

- 2000. La langue des “pretos velhos” au Brésil: un créole à base portugaise d’origine africaine? *Bulletin de la Société de Linguistique de Paris*, t. XCV, fasc. 1, 389-416.
- Boretzky, Norbert. 1993. The concept of rule, rule borrowing and substrate influence in creole languages. In Salikoko S. Mufwene (ed.), *Africanisms in afro-american language varieties*, 74-92. Atenas: University of Georgia Press.
- Bourdier, Pierre. 1977. L’économie des échanges linguistiques. *Langue Française*, 34, 17-34.
- Braun, Maria. 2007. *Word-formation and creolization: the case of early sranan*. Tübingen: Max Niemeyer Verlag.
- Bresnan, Joan. 1972. *The theory of complementation in english syntax*. Cambridge, MA: The MIT Press. Dissertação de doutoramento.
- 2001. *Lexical-functional syntax*. Malden, MA: Blackwell.
- Brito, Ana Maria Barros. 1991 [1988]. *A sintaxe das orações relativas em português*. Faculdade de Letras da Universidade do Porto. Lisboa: INIC. Dissertação de doutoramento.
- Brito, António de Paula. 1887. Dialectos crioulos-portuguezes. Apontamentos para a grammatica que se fala na Ilha de S. Thiago de Cabo Verde. *Boletim da Sociedade de Geografia de Lisboa*, 7, 611-669.
- Bruser, Martina & André dos Reis Santos. 2002. *O dicionário do crioulo da Ilha de Santiago (Cabo Verde)* [dir. por Jürgen Lang]. Tübingen: Gunter Narr Verlag.
- Bruyn, Adrienne. 1996. On identifying instances of grammaticalization in creole languages. In Philip Baker & Anand Suya (eds.), *Changing meanings, changing functions: papers relating to grammaticalization in contact languages*, 29-46. London: University of Westminster Press.
- Bruyn, Adrienne, Pieter Muysken & Maaike Verrips. 1999. Double-object constructions in the creole languages: Development and acquisition. In Michel DeGraff (ed.), *Language creation and language change. creolization, diachrony and development*, 329-373. Cambridge, MA: The MIT Press.
- Buescu, Maria Leonor Carvalhão. 1984. *Historiografia da língua portuguesa - século XVI*, 1ª ed.. Coleção «Nova Universidade», Linguística. Lisboa: Livraria Sá da Costa Ed.
- Bull, Benjamim Pinto. 1989. *O crioulo da Guiné-Bissau, filosofia e sabedoria*. Lisboa/Bissau: ICALP/INEP.
- Bybee, Joan L. & Dan I. Slobin. 1982. Why small children cannot change language on their own. In Anders Ahlqvist (ed.), *Papers from the 5<sup>th</sup> International Conference on Historical Linguistics*, 29-37. Amsterdam: John Benjamins.
- Cadbury, William A. 1910. *Os serviços de S. Thomé*. Lisboa: Livraria Bertrand.
- Cafezeiro, Edwaldo Machado. 1981. *A metafonía portuguesa: aspectos sincrónicos e diacrónicos*. Policopiado. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Dissertação de doutoramento.
- Caldeira, Arlindo. 2002. A escravatura africana vista da América: Alonso de Sandoval e o tráfico de escravos em Angola no início do Século XVII. *Africana Studia: Revista Internacional de Estudos Africanos*, 5, 47-74.
- Câmara Jr., Joaquim Mattoso. 1965. *Introdução às línguas indígenas brasileiras*. Rio de Janeiro: Ao Livro Técnico.
- 1972. Línguas européias de ultramar: o português do Brasil. In Joaquim Mattoso Câmara Jr., *Dispersos*, 71-87. Rio de Janeiro: Fundação Getúlio Vargas.



- 1976. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.
- 1978. *Dicionário de linguística e gramática*, 8ª ed.. Lisboa: Ed. Vozes.
- Cançado, Márcia. 2006. O quantificador TUDO no PB – The quantifier “tudo” in Brazilian Portuguese. *Revista Letras*, 70. 157-182. Curitiba: Ed. UFPR.
- Carreira, Esperança. 2006. *O essencial sobre língua portuguesa. História do português*. Coleção “O Essencial sobre Língua Portuguesa”. Lisboa: Caminho.
- Carreira, António. 1983. *Cabo Verde. Formação e extinção de uma sociedade escravocrata*. Lisboa: ICL.
- Carroll, F. W. 1980. Neurolinguistic processing of a second language: experimental evidence. In Robin C. Scarcella & Stephen. D. Krashen (eds.), *Research in second language acquisition*, 81-88. Rowley, MA: Newbury House.
- 1997b. A influência da variável classe e posição em relação ao núcleo na concordância nominal de número. In Dermeval da Hora (org.), *Diversidade lingüística no Brasil*, 141-157. João Pessoa: Idéia.
- 1999. Concordância nominal de número: um fenômeno variável. In Maria Denilda Moura (org.), *Os múltiplos usos da língua*, 540-543. Maceió: EDUFAL, 1999.
- 2004. A influência da saliência fônica na concordância nominal falada em João Pessoa. In Dermeval da Hora (org.), *Estudos sociolingüísticos: perfil de uma comunidade*, 247-258. João Pessoa: UFPB.
- Carvalho, José G. Herculano de. 1979. *Teoria da linguagem*, vol. I. Coimbra: Ed. Atlântida.
- 1981. Deux langues creoles: le criol du CapVert e le forró de São Tomé. *Biblos*, 57. 1-15.
- Carvalho, Raimunda Coelho de. 1997. *A concordância de número no sintagma nominal na fala urbana de Rio Branco*. Campinas: Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. Dissertação de mestrado.
- Casteleiro, João Malaca, Américo Meira & José Pascoal. 1988. *Nível limiar – para o ensino/aprendizagem do português como língua segunda/língua estrangeira*. Lisboa: Ministério da Educação; ICALP.
- Castilho, Ataliba Teixeira de. 1993, (org.). *Gramática do português falado: as abordagens*, vol. III. Campinas: Unicamp/São Paulo.
- 1997b [1991, 2ª ed ], (org.). *Gramática do português falado: a ordem*, vol. I. 3ª ed. Campinas: Unicamp/São Paulo.
- Castilho, Ataliba Teixeira de & Margarida Basílio (orgs.). 1996. *Gramática do português falado: estudos descritivos*, vol. IV. Campinas: Unicamp/São Paulo.
- Castro, Ana. 2006. *On possessives in portuguese*. Lisboa: Universidade de Lisboa/Paris: Université Paris-8. Dissertação de doutoramento.
- Castro, Ivo. 2001 [1991]. *Curso de história da língua portuguesa*, 1ª ed., 2ª impressão. Lisboa: Universidade Aberta.
- Cenoz, Jasone & Fred Genesee (eds.). 2001. *Trends in bilingual acquisition*. Amsterdam: John Benjamins.
- Cerqueira, Vicente Cruz. 1996. *A sintaxe do possessivo no português brasileiro*. Campinas: Unicamp/São Paulo. Dissertação de doutoramento.
- Cervelló, Joseph Sánchez. 1999. São Tomé e Príncipe, 1953. A matança de Batepá. *História*. 26-37.
- Chatelain, Heli. 1988/1989. *Gramática elementar de kimbundu ou língua de Angola*. Genève: Typ. de Charles Schuchardt.

- Chaudenson, Robert. 2001. *Creolization of language and culture (revised in collaboration with Salikoko S. Mufwene)*. Londres: Routledge.
- Chevalier, Augusto. 1910. *L'île de San-Thomé*. Lisboa: [s.n.]. Cota: Cxa. 102-4. Sociedade de Geografia de Lisboa.
- Chomsky, Noam. 1959. A review of B. F. Skinner's "Verbal behaviour". *Language*, 35. 26-58.
- 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in english transformational grammar*, 184-221. Waltham, MA: Ginn and Company.
- 1973. Conditions on transformations. In Stephen R. Andersen & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232-286. New York, NY: Academic Press. (232-286).
- 1975. *Aspectos da teoria da sintaxe*. Coimbra: Arménio Amado.
- 1977a. *Reflexões sobre a linguagem*. Lisboa: Ed. 70.
- 1977b. On wh-movement. In Peter William Culicover, Thomas Wason & Adrian Akmajian (eds.), *Formal syntax*, 71-132. Nova Iorque: Academic Press.
- 1994. *O conhecimento da língua: sua natureza, origem e uso*. Lisboa: Caminho.
- 1999b. Derivation by phase. *MIT Occasional Papers in Linguistics*, 18. Cambridge, MA: The MIT Press.
- 2000. Minimalist inquiries: the framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: essays on minimalist syntax in honor of Howard Lasnik*, 89-155. Cambridge, MA: The MIT Press.
- Chomsky, Noam & Howard Lasnik. 1977. Filters and control. In *Linguistic Inquiry*, 8(3). 425-504.
- Cintra, Luís Filipe Lindley. 1970. Os ditongos decrescentes *ou* e *ei*: esquema de um estudo sincrónico e diacrónico. *Anais do 1º simpósio de filologia românica (Rio de Janeiro, 1958)*. 115-134.
- Clahsen, Harald. 1984. The acquisition of german word order: a test case for cognitive approaches to L2 development. In Roger W. Andersen (ed.), *Second languages: a cross-linguistic perspective*, 219-242. Rowley, MA: Newbury House.
- Clements, Joseph Clancy. 1996. *Genesis of a language: formation & development of korlay portuguese*. Amsterdam: John Benjamins.
- 2003. The tense-aspect system in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition*, 25(2). 245-281.
- Coelho, Francisco Adolpho. 1868. *A lingua portugueza: phonologia, etymologia, morphologia e syntaxe/por F. Adolpho Coelho*. Coimbra: Imprensa da Universidade.
- Comissão Temporária da Escravatura. 1925. A escravatura e a sociedade das nações. *Boletim da Agência Geral das Colónias*, ano I, 4. 24-55.
- Comrie, Bernard. 1981. *Language universals and linguistics typology*. London: Blackwell.
- 1988. Linguistic typology. In Frederick J. Newmeyer (ed.), *Linguistics: the Cambridge survey*, vol. 1, 447-461. Cambridge, MA: Cambridge University Press.
- Confortin, Helena. 2001. Atitudes lingüísticas de falantes bilíngües. *Letras – Revista do Instituto de Letras (PUC Campinas)*, vol. 20(1/2). 123-135.
- Connell, Bruce & David Zeitlyn. 2010. Sociolinguistics studies of West and Central Africa. In Martin J. Ball (ed.), *The Routledge handbook of sociolinguistics around the world*, 203-215. London/New York, NY: Routledge.

- Cook, Vivian James. 1988. *Chomsky's universal grammar: an introduction*. Oxford: Blackwell Publishers.
- Corder, Stephen Pit. 1978. Error analysis, interlanguage and second language acquisition. In Valerie Kinsella, (ed.), *Language teaching & linguistics: surveys*, 60-78. Cambridge, MA: Cambridge University Press.
- 1981a. Formal simplicity and functional simplification in second language acquisition. In Roger W. Andersen (ed.), *New dimensions in second language research*, 146-152. Rowley, MA: Newbury House.
- Corrêa, Letícia Maria Sicuro. 1999. Aquisição da linguagem: uma retrospectiva dos últimos trinta anos [em linha]. *D.E.L.T.A.: Documentação de Estudos em Lingüística Teórica e Aplicada*, vol. 15. [Consult. 12 Maio 2009]. Disponível em: [http://www.scielo.br/scielo.php?pid=S0102-4501999000300014&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0102-4501999000300014&script=sci_arttext).
- 2006. Conciliando processamento lingüístico e teoria de língua no estudo da aquisição da linguagem: habilidades discriminatórias de bebês, categorias funcionais e a disponibilidade de um sistema computacional lingüístico In Letícia Maria Sicuro Corrêa (ed.), *Aquisição da linguagem e problemas do desenvolvimento lingüístico*, 21-78. Rio de Janeiro: Ed. da PUC.
- Coseriu, Eugenio. 1979. *Sincronia, diacronia e história*. Rio de Janeiro: Presença.
- 1997. Scrambling in european portuguese. In Benjamin Bruening (org.), *Proceedings of SCIL 8, MIT Working Papers in Linguistics*, 95-114. Cambridge, MA: MITWPL.
- 1998. *Word order variation: a constraint-based approach*. Leiden:Universiteit Leiden. Dissertação de doutoramento.
- 2001a. VOS in portuguese: arguments against an analysis in terms of remnant movement. In Artemis Alexiadou et alii (eds.), *The dimensions of movement: from features to remnants*, 69-90. Amsterdam/Philadelphia: John Benjamins.
- 2001b. Postverbal subjects and agreement in unaccusative contexts in European Portuguese. *The Linguistic Review*, 18. 1-17
- 2002. A multifactorial approach to adverb placement: assumptions, facts and problems [em linha]. In *Lingua*, 114(6), *Special Issue on Adverbs*. 711-753. [Consult. 09 Mar. 2010]. Disponível em: [http://www.sciencedirect.com/science?\\_ob=GatewayURL&\\_origin=CONTENTS&\\_method=citationSearch&\\_piikey=S0024384103000494&\\_version=1&md5=643e7f67cf30febf794d75ac9f59c940](http://www.sciencedirect.com/science?_ob=GatewayURL&_origin=CONTENTS&_method=citationSearch&_piikey=S0024384103000494&_version=1&md5=643e7f67cf30febf794d75ac9f59c940).
- 2004. *Subject position and interfaces. The case of european portuguese*. Berlin: Mouton de Gruyter.
- Costa, João et alii. 2001. Concordância com *a gente*: um problema para a teoria de verificação de traços. [em linha]. [Consult. 25 Mar. 2009]. Disponível em: [http://www.clul.ul.pt/equipa/spereira/Costa\\_et\\_al01.pdf](http://www.clul.ul.pt/equipa/spereira/Costa_et_al01.pdf).
- Costa, João & Maria João Freitas. 2001. Morphological and/or prosodic place holders [em linha]. *Encontro do projecto Statistical Physics, Pattern Identification and Language Change*. Bielefeld: ZIF. [Consult. 17 Mar. 2010]. Disponível em: <http://www.uni-bielefeld.de/ZIF/FG/2000Complexity/costa.pdf>.
- Costa, Joaquim Vieira Botelho da & Custódio José Duarte. 1886. O crioulo de Cabo Verde. Breves estudos sobre o crioulo das Ilhas de Cabo Verde oferecidos ao Dr. Hugo Schuchardt. *Boletim de Geografia de Lisboa*, 2. 235-328. Lisboa.

- Coulmas, Florian (ed.). 2000 [1997]. *The handbook of sociolinguistics*. Malden, MA: Blackwell.
- Coutinho, Ismael de Lima. 1969. *Gramática histórica*. Rio de Janeiro: Acadêmica.
- Crystal, David. 1987. *The Cambridge encyclopedia of language*. Cambridge, MA: Cambridge University Press.
- 1991 [1971]. *A linguística*, 2ª ed.. Lisboa: Publ. D. Quixote.
- 1993. *Enciclopedia del lenguaje de la Universidad de Cambridge*. Madrid: Taurus.
- CTI de São Tomé. 1966. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicossocial N.º. 1/66, Jan. 1966*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 2, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1972a. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicológica N.º. 02/72, 14 Out. 1972*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 6A, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1972b. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicológica N.º. 03/72, 03 Jul. 1972*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 6B, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- 1972c. *Relatórios periódicos de acção psicossocial. Relatório periódico de acção psicológica N.º. 04/72, 31 Dez. 1972*. Caixa n.º. 1, proc. PT AHM/DIV/208/01/07, doc. 6C, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Cunha, Celso & Lindley Cintra. 2003 [1984]. *Nova gramática do português contemporâneo*, 3ª ed. revista, nova apresentação, 6ª impressão. Rio de Janeiro: Ed. Nova Fronteira.
- De Houwer, Annick (ed.). 1998. Bilingual acquisition [special issue]. *International Journal of Bilingualism*, 2(3).
- DeGraff, Michel. 1994. The morphology-syntax interface in creolization (and diachrony). *Studies in the Linguistic Sciences*, 24(2). 115-132.
- Direcção de Estatística de São Tomé e Príncipe. 1987. *1.º recenseamento geral da população e da habitação, 1981*. São Tomé: Direcção de Estatística de São Tomé e Príncipe.
- 1996. *II recenseamento geral da população e da habitação de 1991: seminário de disseminação dos resultados*. São Tomé: Direcção de Estatística de São Tomé e Príncipe.
- Duarte, Inês. 1987. *A construção de topicalização na gramática do português: regência, ligação e condições sobre movimento*. Universidade de Lisboa. Dissertação de doutoramento.
- Duarte, Inês & Matilde Miguel (orgs.). *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística: gramática e variação*, vol. III. Lisboa: Faculdade de Letras da Universidade de Lisboa.
- Dubois, Jean *et alii*. 1973. *Dicionário de linguística*. Paris: Lib. Larousse.
- Dunn, Ernst F. 1968. *An introduction to bini*. East Lansing: African Studies Center, Michigan State University.

- Embick, David & Rolf Noyer. 2007. Distributed morphology and the syntax/morphology interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*. Oxford/New York, NY: Oxford University Press.
- Emmerich, Charolte. 1992. O Português de contacto no Parque Indígena do Xingu, Mato Grosso, Brasil Central. *Estudos Linguísticos e Literários*, 13. 57-90.
- Ernst, Thomas. 2002. *The syntax of adjuncts*. Cambridge, MA: Cambridge University Press.
- Estado Maior do Exército, 2ª. Repartição, 3ª. Direcção Geral. s.d. *A defesa da colónia de São Tomé e Príncipe, 1939-1941*. Caixa nº. 1, proc. PT AHM/DIV/2/08/01/10, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Eubank, Lynn. 1996. Negation in early german-english interlanguage: more valueless features in the L2 initial stage. *Second Language Research*, 12. 73-106.
- Eynde, Frank van. 2006. NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics*, 42. 139-186.
- Færch, Claus & Grabielle Kasper. 1987. Perspectives on language transfer. *Applied Linguistics*, 8(2). 111-136.
- Fall, Yoro K. 1989. Escravatura, servidão e reconquista. In Luís de Albuquerque (dir.), *Portugal no mundo – povoamento e colonização do Reino de Portugal; início dos descobrimentos marítimos portugueses; o avanço no Atlântico*, vol. 1, 301-314. Lisboa: Publ. Alfa.
- Faria, Isabel Hub. 1983. *Para a análise da variação socio-semântica*, 2 vols. Lisboa: Universidade de Lisboa. Dissertação de doutoramento.
- Faria, Isabel Hub et alii (coords.). 1996. *Introdução à linguística geral e portuguesa*. Lisboa: Ed. Caminho.
- Farias, Jair Gomes de. 2006. Variação entre *a*, *para* e *em* no português brasileiro e no português europeu: algumas notas [em linha]. *Letras de Hoje*, 41(1). 213-234. [Consult. 14 Mar. 2008]. Disponível em: <http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/viewFile/589/420>.
- Felix, Sascha W. 1977. Interference, interlanguage and related issues. In Carol Molony, Helmut Zobl & Wilfried Stolting (eds.), *German in contact with other languages*, 237-258. Kronberg: Scriptor Verlag.
- Ferguson, Charles A. 1959. Diglossia. *World*, 15. 325-340.  
—1991. Diglossia revisited. *Southwest Journal of Linguistics*, 10(1). 214-234.
- Ferraz, Luiz Ivens. 1974. A linguistic appraisal of angolar. *Memoriam António Jorge Dias*, vol. 2. 177-186. Lisboa: Instituto de Alta Cultura/Junta de Investigações do Ultramar.  
—1978. The creole of São Tomé. *African Studies*, 37. 3-68 & 234-284.
- Ferraz, Luiz Ivens & Marius F. Valkhoff. 1975. A comparative study of são-tomense and cabo-verdiano creole. In Marius F. Valkhoff (ed.), *Miscelânea luso-africana*. 15-39.
- Ferreira, Fernanda. 2009. Marcadores de plural no português brasileiro e no crioulo cabo-verdiano. In Ana M. Carvalho (org.), *Português em contacto*, 107-130. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
- Fiengo, Robert Wilson. 1974. *Semantic conditions on surface structure*. Cambridge, MA: The MIT Press. Dissertação de doutoramento.

- Filho, Jomson Teixeira da Silva. 2010. Aquisição do número gramatical na concordância interna ao DP no português brasileiro [em linha]. *WebArtigos.com*. [Consult. 27 Abr. 2010]. Disponível em:  
<http://www.webartigos.com/articles/44041/1/aquisicao-do-numero-gramatical-na-concordancia-interna-ao-DP-no-portugues-brasileiro-nominal/pagina1.html>.
- Finnemann, Michael D. 1992. Learning agreement in the noun phrase: the strategies of three first-year spanish students. *International Review of Applied Linguistics in Language Teaching*, 30. 121-36.
- Fiorin, José Luiz & Margarida Petter (orgs.). 2008. *África no Brasil – a formação da língua portuguesa*. São Paulo: Ed. Contexto.
- Fontes, Carlos de Sousa Ferreira. 2008. *Estudo do léxico do são-tomense, um crioulo de base lexical portuguesa*. Congresso Anual da ACBPLe – Associação de Crioulos de Base Lexical Portuguesa e Espanhola. Universidade de Lisboa, Portugal, 03-05 de Julho.
- Fransceschina, Florencia. 2001. Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17. 213-247.
- 2003. Parameterized functional features and SLA. In Juana M. Liceras *et alii* (eds.), *Proceedings of the 6<sup>th</sup> generative approaches to second language acquisition conference (GASLA 2002)*, 97-105. Somerville, MA: Cascadilla Proceedings Project.
- Freitas, Maria João. 1997. *Aquisição da estrutura silábica do português europeu*. Lisboa: Universidade de Lisboa. Dissertação de doutoramento.
- 1998. Estatutos das consoantes que fecham sílabas no português europeu: evidência dos dados da aquisição. *Actas do XIV encontro nacional da Associação Portuguesa de Linguística*. 541-556. Braga: Associação Portuguesa de Linguística.
- Galloway, Linda M. & Robin C. Scarcella. 1982. Cerebral organization in adult second language acquisition: is the right hemisphere more involved? *Brain and Language*, 16. 56-60.
- Galves, Charlotte Marie Chambelland. 1996. Princípios, parâmetros e aquisição da linguagem. *Cadernos de Estudos Linguísticos*, 29. 137-152.
- 2008. O papel das línguas africanas na formação do português brasileiro: (mais) pistas para uma nova agenda de pesquisa. *Gragoatá* (UFF), 24. 145-164.
- Gardner-Chloros, Pénélope. 1983. Code-switching: approches principales et perspectives. *La Linguistique*, 19(2). 21-53.
- Gass, Susan. M. 1983. Language transfer and universal grammar relations. In Susan M. Gass & Larry Selinker (eds.), *Language transfer in language learning*, 69-82. Rowley, MA: Newbury House.
- 1987. The Resolution of conflicts among competing systems: a bidirectional perspective. *Applied Psycholinguistics*, 8. 329-350.
- Gass, Susan. M. & Jacquelyn Schachter (eds.). 1989. *Linguistic perspectives on second language acquisition*. Cambridge, MA: Cambridge University Press.
- Genesee, Fred. 2001. Bilingual first language acquisition: exploring the limits of the language faculty [em linha]. *Annual Review of Applied Linguistics*, 21. 153-168. [Consult. 12 Nov. 2008]. Disponível em:  
<http://journals.cambridge.org/download.php?file=%2FAPL%2FAPL21%2FS026719050000095a.pdf&code=2726eda4f5cb5e0f3d961c531b3262e2>.

- Genesee, Fred & Elena Nicoladis. 2005. Bilingual first language acquisition [em linha]. In Erika Hoff & Marilyn Shatz (eds.), *Handbook of language development*, 324-340. Oxford: Blackwell [Consult. 09 Ag. 2008]. Disponível em: <http://www.psych.mcgill.ca/perpg/fac/genesee/HDBK%20BFLA%20FINAL.pdf>.
- Gess, Randall & Julia Rogers Herschensohn. 2001. Shifting the DP parameter: A study of anglophone french L2 learners. In Caroline R. Wiltschire & Joaquim Camps (eds.), *Romance syntax, semantics and their L2 acquisition*, 105-119. Amsterdam: John Benjamins.
- Gibson, Edward & Kenneth N. Wexler. 1994. Triggers. *Linguistic Inquiry*, 25. 407-454.
- Gilbert, Glenn G. 1980, (ed.). *Pidgin and creole languages: selected essays by Hugo Schuchardt*. Cambridge, MA: Cambridge University Press.
- Giorgi, Alessandra & Giuseppe Longobardi. 1991. *The syntax of noun phrases*. Cambridge, MA: Cambridge University Press.
- Gleitman, Lila R. & Eric Wanner. 1982. Language acquisition: the state of the art. In Eric Wanner & Gleitman, Lila R. (eds.), *Language acquisition: the state of the art*, 319-346 Cambridge, MA: Cambridge University Press.
- Godoy, Luísa Andrade Gomes. 2005. *A palavra TUDO como quantificador universal puro no português brasileiro* [em linha]. Belo Horizonte: Faculdade de Letras da UFMG. Monografia apresentada ao Colegiado de Graduação da Faculdade de Letras da Universidade Federal de Minas Gerais como requisito parcial para obtenção do título de Bacharel em Letras [Consult. 04 Maio 2009]. Disponível em: <http://www.letras.ufmg.br/marciacancado/Monografia%20Luisa.pdf>.
- Goldin-Meadow, Susan & Carolyn Mylander. 1999. Beyond the input given: the child's role in the acquisition of language. *Language*, 66. 323-355.
- Gomes, Cristina Abreu & Cláudia Nívia Roncarati de Souza. 2003. Variáveis fonológicas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 73-80. São Paulo: Contexto.
- Gomes, Nataniel dos Santos. 2005. Observações sobre os clíticos [em linha]. In CiFEFiL, Almanaque CiFEFiL, edição em cd-rom. [Consult. 31 Dez. 2009]. Disponível em: <http://www.filologia.org.br/revista/artigo/7%2820%2908.htm>.
- Gonçalves, Anabela; et alii. 1996. *Quatro estudos em sintaxe do português (uma abordagem segundo a teoria dos princípios e parâmetros)*. Lisboa: Ed. Colibri. – “Coleção Estudos Linguísticos”.
- Gonçalves, Manuel da Luz & Lelia Lomba de Andrade. 2003. *Pa nu papia kriolu*. Boston: M & L Enterprises.
- Gonçalves, Perpétua. 1996. Aspectos da sintaxe do português de Moçambique. In Isabel Hub Faria et alii (coords), *Introdução à linguística geral e portuguesa*, 313-322. Lisboa: Ed. Caminho.
- González, Neide Therezinha Maia. 1994. *Cadê o pronome? – O gato comeu*. São Paulo: Universidade de São Paulo – Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas. Dissertação de doutoramento.
- Gordon Jr., Raymond G. (ed.). 2005. Languages of São Tomé e Príncipe [em linha]. *Ethnologue: Languages of the World*, 15ª ed. Dallas, Tex.: SIL International. [Consult. 03 Dez. 2007]. Disponível em: [http://www.ethnologue.com/show\\_country.asp?name=ST](http://www.ethnologue.com/show_country.asp?name=ST).

- Governo de Angola. 1971. *Anuário turístico: Angola*, 4. Lisboa: Editorial de Publicações Turísticas.
- Greenberg, Joseph Harold. 1996. Language universals, with special reference to feature hierarchies. *Janua Linguarum, Series Minor*, 59. The Hague: Mouton.
- Grosjean, François. 1995. A psycholinguistic approach to code-switching: the recognition of guest words by bilinguals. In Lesley Milroy & Pieter Muysken (eds.), *One speaker, two languages: cross-disciplinary perspectives on code-switching*, 259-275. New York, NY: Cambridge University Press.
- Gryner, Helena & Nelize Pires de Omena. 2003. A interferência das variáveis semânticas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 89-100. São Paulo: Contexto.
- Guerreiro, Manuel Viegas. 1997. *Povo, povos e cultura (Portugal – Angola – Moçambique)*. Lisboa: Ed. Colibri.
- Guinness, Grattan. 1882. *Grammar of the Congo language as spoken two hundred years ago, translated from the latin of Brusciotto*. London: Hodder & Stoughton.
- Guthrie, Malcolm. 1948. *The classification of the african languages*. London: Oxford Univ. Press.
- Guy, Gregory Riordan. 1980. Contextual conditioning in variable lexical phonology. *Language Variation and Change*, 3. 223-239.
- 1993. The quantitative analysis of linguistic variation. In Dennis R. Preston, (ed), *American dialect research*, 223-249. Amsterdam/Philadelphia: John Benjamins.
- Gregory Riordan Guy et alii (eds.). 1996. *Towards a social science of language*, vol. 1. Variation and change in language and society, 199-220. Amsterdam/Philadelphia: John Benjamins.
- Haegeman, Liliane M. V. & Jacqueline Guéron. 1999. *English grammar: a generative perspective*. Oxford: Blackwell.
- Hagemer, Tjerk. 2001. *Os crioulos do Golfo da Guiné: perspectivas sócio-históricas [em linha] [Consult. 02 Out. 2008]*. Disponível em: [www.unb.br/il/liv/crioul/coloq.htm](http://www.unb.br/il/liv/crioul/coloq.htm).
- Hall, Robert. 1966. *Pidgin and creole languages*. Ithaca, NY: Cornell University Press.
- Halle, Morris. 1984. Fonética. In Ruggiero Romano (dir.), *Enciclopédia Einaudi*, vol. 2 – “Linguagem - Enunciação”, 132-155. Lisboa: INCM.
- Halle, Morris & Alec Marantz. 1994. Some key features of distributed morphology. In *MITWPL 21: Papers on phonology and morphology*, 275-288. Cambridge, MA: The MIT Press.
- Hammarberg, Björn. 1997. Conditions on transfer on phonology. In Allan R. James & Jonathan Leather (eds.), *Second language speech? Structure and process*, 161-180. Berlin: Mouton de Gruyter.
- Han, ZhaoHong. 1998. 2004. *Fossilization in adult second language acquisition*. Clevedon: Multilingual Matters.
- Harley, Heidi & Rolf Noyer. 1999. State-of-the-article: distributed morphology. *Glott*, 4(4). 3-9.
- Harris, James W. 1991. The exponence of gender in spanish. *Linguistic Inquiry*, 22. 27-62.
- Harris, John. 1984. Syntactic variation and dialect divergence. *Journal of Linguistics*, 20. 303-327.



- Heckler, Evaldo, Sebald Back & Egon Massing. 1984/1985. *Dicionário morfológico da língua portuguesa (Vols. 1,2,3,4, e 5)*. São Leopoldo (Rio Grande do Sul): UNISINOS (Universidade do Vale do Rio dos Sinos).
- Heine, Bernd & Derek Nurse (eds.). 2000. *African languages: an introduction*. Cambridge, MA: Cambridge University Press.
- Heine, Bernd & Mechtild Reh. 1984. *Grammaticalization and reanalysis in african languages*. Hamburg: H. Helmut Buske.
- Heine, Bernd & Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language*, 27. 529-572.
- Helms-Park, Rena. 2003. Transfer in SLA and creoles: the implications of causative serial verbs in the interlanguage of vietnamese ESL learners. *Studies in Second Language Acquisition*, 25. 211-244.
- Heringer, Hans Jürgen & José Pinto de Lima. 1987. *Palavra puxa palavra – comunicação e gramática dependencial*. Identidade: Série Língua Portuguesa. Lisboa: Instituto de Cultura e Língua Portuguesa – Ministério da Educação e Cultura.
- Henanz, M. Loisa & José M. Brucart. 1987. *La syntaxis. Principios teoricos. La oracion simple*. Barcelona: Editorial Critica.
- Hesseling, Dirk Christiaan. 1979 [1897]. Het hollandsch in Zuid Afrika. *De Gids*, 61. 138-162; reimpresso como «Dutch in South Africa». In Tom L. Markley & Paul T. Roberge (eds. e trads.), *On the origin and formation of creoles: a miscellany of articles*. Ann Arbor: Karoma.
- Hodges, Tony & Malyn Newitt. 1988. *São Tomé e Príncipe: from plantation colony to microstate*. Boulder: Westview Press.
- Hoff, Erika & Marilyn Shatz (eds.). 2006. *Handbook of language development* [em linha]. Oxford: Blackwell. [Consult. 27 Set. 2008]. Disponível em:  
[http://books.google.pt/books?id=PkSYF8pM3pIC&pg=PA368&lpg=PA368&dq=shatz+bilingual+handbook+of+language+development&source=bl&ots=SOWLQYZfjE&sig=ptNpITjEoUnH313Bju3pJ7qL-lo&hl=pt.PT&ei=1bWLSrBF5yu6gP4n\\_C2Cg&sa=X&oi=book\\_result&resnum=1&ct=result](http://books.google.pt/books?id=PkSYF8pM3pIC&pg=PA368&lpg=PA368&dq=shatz+bilingual+handbook+of+language+development&source=bl&ots=SOWLQYZfjE&sig=ptNpITjEoUnH313Bju3pJ7qL-lo&hl=pt.PT&ei=1bWLSrBF5yu6gP4n_C2Cg&sa=X&oi=book_result&resnum=1&ct=result)
- Holm, John. 2000a. Semi-creolization: problems in the development of theory. In Ingrid Neumann-Holzschuh & Edgar W. Schneider. (eds.), *Degrees of restructuring in creole languages*. Creole Languages Library (CLL), vol. 22, 19-40. University of Regensburg. Amsterdam/Philadelphia: John Benjamins.
- 2000b. *An introduction to pidgins and creoles*. Cambridge Textbooks in Linguistics. Cambridge, MA: Cambridge University Press.
- 2004. *Languages in contact – the partial restructuring of vernaculars*. Cambridge, MA: Cambridge University Press.
- Holm, John, Gerardo A. Lorenzino & Heliana R. de Mello. 2000. Differing degrees of restructuring in two vernaculars: Caribbean Spanish and brazilian portuguese. In Ernesto d' Andrade, Maria Antónia Coelho da Mota & Dulce Pereira (orgs.), *Crioulos de base portuguesa (Actas do workshop sobre crioulos de base lexical portuguesa. FLUL, 29 e 30 de Abril de 1999)*, 201-222. Braga: Associação Portuguesa de Linguística.
- Holm, John & Peter L. Patrick (eds.). 2007. *Comparative creole syntax: parallel outlines of 18 creole grammars*. Westminster Creolistic Series 7. Plymouth, UK: Battlebridge Publications.

- Holmes, Janet. 2001 [1992]. *An introduction to sociolinguistics*, 2<sup>nd</sup> ed. Essex: Longman/Pearson Education Ltd.
- Holzman, Mathilda. 1997. *The language of children*, 2<sup>nd</sup> ed. Oxford: Blackwell.
- Huber, Joseph. 1986 [1933]. *Gramática do português antigo*. Trad. port. de Maria Manuela Gouveia Delille. Lisboa: Fundação Calouste Gulbenkian.
- Hudson, Carla L. & Elissa L. Newport. 1999. Creolization: could adults really have done it all?. In Annabel Greenhill, Heather Littlefield & Cheryl Tano (eds.), *Proceedings of the 23<sup>rd</sup> annual Boston University conference on language development*, vol. 1, 265-276. Somerville, MA: Cascadilla Press.
- 2005. Regularizing unpredictable variation: the roles of adult and child learners in language formation and change. *Language Learning and Development*, 1. 151-195.
- Hudson, Richard A. 1996 [1980]. *Sociolinguistics*, 2<sup>nd</sup> ed. Cambridge, MA: Cambridge University Press.
- Hulk, Aafke & Leonie Cornips. 2006. Between 2L1 – and child L2 acquisition: an experimental study and bilingual dutch. In Conxita Lléo (ed.), *Interfaces in multilingualism: acquisition representation and processing*, vol.4, 115-138. Amsterdam: John Benjamins.
- Humboldt, Friedrich Wilhelm von. 1859. *De l'origine des formes grammaticales et de leur influencesur le développement des idées*. Traduction de Alfred Tonnellé. Paris: A. Franck.
- Hyams, Nina. 1992. *Null subjects in child language and the implications of cross-linguistic variation*. Cornell Symposium on Language Acquisition and Linguistic Theory. Cross-linguistic Perspectives. Ithaca, NY, Abril 1992.
- Hyams, Nina & Kenneth N. Wexler. 1993. On the grammatical basis of null subjects in child language. *Linguistic Inquiry*, 24(3). 421-459.
- Hymes, Dell. 1971b. *On communicative competence*. Philadelphia: University of Pennsylvania Press.
- Ijaz, I. Helene. 1986. Linguistic and cognitiv determinants of lexical acquisition in a second language. *Language Learning*, 36. 401-451.
- Ilari, Rodolfo (org.). 1996 [1992]. *Gramática do português falado: níveis de análise lingüística*, vol. II, 3<sup>a</sup> ed. Campinas: Unicamp/São Paulo.
- Instituto Camões & Centro de Linguística da Universidade de Lisboa. 2001. *Português falado: documentos autênticos*, edição em cd-rom.
- Instituto Nacional de Estatística de São Tomé e Príncipe. 2003c. *Migrações: III recenseamento geral da população e da habitação de 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- 2005a. *Dinâmica natural da população em São Tomé e Príncipe: III recenseamento geral da população e da habitação de 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- 2005b. *A mulher em São Tomé e Príncipe: III recenseamento geral da população e da habitação de 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- Isensee, Dinah Maria Montenegro. 1964. *O falar de Mato Grosso (Bahia): fonêmica, aspectos da morfo-sintaxe e do léxico*. Dissertação de mestrado.
- Jackobson, Roman. 1968. *Child language, aphasia and phonological universals*. The Hague: Mouton.

- Jarvis, Scott & Terence Odlin. 2000. Morphological type, spatial reference and language transfer. *Studies in Second Language Acquisition*, 22. 535-556.
- Jeroslow, Elizabeth Helen McKinney. 1974. *Rural cearense portuguese: a study of one variety of nonstandard brazilian speech*. Dissertação de mestrado.
- Jiang, Nan. 2000. Lexical representation and development in a second language. *Applied Linguistics*, 21(1). 47-77.
- Jordens, Peter. 1994. Acquiring german and french in a bilingual setting. In Jürgen M. Meisel (ed.), *Bilingual first language acquisition: french and german grammatical development*, 3-14. Amsterdam: John Benjamins Publishing.
- Joseph, Brian. D. 2001. Is there such a thing as grammaticalization?. *Language Sciences*, 23(2/3). 163-186.
- Joseph, John Earl. 1987. *The rise of language standards and standard languages*. Londres: Frances Pinter.
- Jourdan, Christine & Roger M. Keesing. 1997. From fisin to pijin: creolization in process in the Solomon Islands. *Language in Society*, 26. 401-420.
- Kean, M. L. 1986. Core issues in transfer. In Eric Kellerman & Michael Sharwood Smith (eds.), *Crosslinguistic influence in second language acquisition*, 80-90. Oxford: Pergamon.
- Keesing, Roger. M. 1991. Substrates, calquing and grammaticalization in melanesian pidgin. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization, vol. 1: Focus on theoretical and methodological issues*, 315-342. Amsterdam/Philadelphia: John Benjamins.
- Kellerman, Eric. 1978. Giving learners a break: native language institutions as a source of predictions about transferability. *Working Papers on Bilingualism*, 36. 59-92.
- 1995. Crosslinguistic influence: transfer to nowhere?. *Annual Review of Applied Linguistics*, 15. 125-150.
- Kempen, Gerard & Edward Hoenkamp. 1987. An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11. 201-258.
- Kerswill, Paul. 1996. Children, adolescents and language change. *Language Variation and Change*, 8. 177-202.
- Kerswill, Paul & Anne Williams. 2000. Creating a new town koiné: children and language change in Milton Keynes. *Language in Society*, 29(1). 65-115.
- King, Ruth Elizabeth. 2000. *The lexical basis of grammatical borrowing: a Prince Edward Islands french case study*. Amsterdam: John Benjamins.
- Klein, Elaine C. & Gita Martohardjono. 1999. Investigating second language grammars: some conceptual and methodological issues in generative SLA research. In Elaine C. Klein & Gita Martohardjono (eds.), *The development of second language grammars: a generative approach*, 3-34. Amsterdam: John Benjamins.
- Koch, Ingedore Grunfeld Villaça. 1997 (org.). *Gramática do português falado: desenvolvimentos*, vol. VI. Campinas: Unicamp/São Paulo.
- Kouwenberg, Silvia. 2003 (ed.). *Twice as meaningful. Reduplication in pidgins, creoles and other contact languages*. London: Battlebridge.
- Kouwenberg, Silvia & John Victor Singler (eds.). 2008. *The handbook of pidgin and creoles studies*. Oxford: Wiley-Blackwell.
- Krashen, Stephen. D. 1973. Lateralization, language learning and the critical period: some new evidence. *Language Learning*, 23. 63-74.

- Labov, William. 1971. The notion of system in creole languages. In Dell Hymes (org.), *Pidginization and creolization of languages*, 447-472. Cambridge, MA: Cambridge University Press.
- Lardier, Donna. 1998. Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research*, 14(4). 359-375.
- Lassagne, François. 2004. Créole, la naissance d'une langue. *Science & Vie – Hors Série*, 227. 78-91.
- Lausberg, Heinrich. 1981. *Linguística românica*, 2<sup>a</sup> ed. Lisboa: Fundação Calouste Gulbenkian.
- Lefebvre, Claire. 2000. What do creole studies have to offer to mainstream linguistics?. *Journal of Pidgins and Creole Languages*, 15(1). 127-153.
- Lefebvre, Claire & John S. Lumsden. 1989. Les langues créoles et la théorie linguistique. *Revue Canadienne de Linguistique*, 34. 319-337.
- Lehmann, Christian. 1982. Thoughts on grammaticalization. A programmatic sketch, vol. I. *Arbeiten des Kölner Universalien-Projekts*, 48. Köln: Universität zu Köln, Institut für Sprachwissenschaft.
- Lennon, Paul. 1991. Error elimination and error fossilization: a study of an advanced learner in the L2 community. *International Review of Applied Linguistics*, 93/94. 129-151.
- Lessa-de-Oliveira, Adriana Stella Cardoso. 2005. Aquisição da linguagem e variação lingüística. *Estudos Lingüísticos*, XXXIV. 409-414.
- Lewis, M. Paul (ed.). 2009. *Ethnologue: languages of the world*, 16<sup>th</sup> ed. [em linha]. Dallas, TEX.: SIL International. [Consult. 24 Agst. 2010]. Disponível em: <http://www.ethnologue.com/>.
- Liceras, Juana M. *et alii*. 2006. L2 acquisition as a process of creolization: insights from child and adult code-mixing. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 113-144. Amsterdam/Philadelphia: John Benjamins.
- Lightfoot, David. 1989. The child's trigger experience: degree-0 learnability. *Behavioral and Brain Sciences*, 12(2). Cambridge, MA: Cambridge University Press.
- 1991b. *How to set parameters: degree-0 learnability*. Cambridge, MA: The MIT Press.
- 1993. Uma ciência da história? *D.E.L.T.A.*, vol. 2(9). 275-294.
- Lipski, John M. 2000. Bozal spanish: restructuring or creolization?. In Ingrid Neumann-Holzschuh & Edgar W. Schneider. (eds.), *Degrees of restructuring in creole languages*. Creole Languages Library (CLL), vol. 22, 437-468. University of Regensburg. Amsterdam/Philadelphia: John Benjamins.
- Longobardi, Giuseppe. 2000. The structure of DP's: some principles, parameters and problems. In Mark Baltin & Chris Collins (eds.). *The handbook of contemporary syntactic theory*, 562-603. Oxford/Malden, MA: Blackwell Publishers, Ltd (edition for sale in the mainland territory of the People's Republic of China only).
- Lopes, David. 1936. *A expansão da língua portuguesa no Oriente durante os séculos XVI, XVII e XVIII*. Barcelos: Portucalense Ed. Lda.

- Lopes, Marília. 1989. A exploração económica da Guiné e de Cabo Verde nos séculos XV e XVI. In Luís de Albuquerque (dir.), *Portugal no mundo – povoamento e colonização do Reino de Portugal; início dos descobrimentos marítimos portugueses; o avanço no Atlântico*, vol. 1, 250-263. Lisboa: Publ. Alfa.
- Lopes, Norma & Alan Norman Baxter, 2009. *O desenvolvimento de regras de concordância variável em variedades de Português a partir de modelos-estímulos (inputs) diferentes*. ROSAE - I Congresso Internacional de Linguística Histórica, Salvador, Bahia, Brasil, 26-29 Julho, 2009.
- López-Escartin, Nuria. 1993. *Données de base sur la population: São Tomé e Príncipe*. Paris: Centre Français sur la Population et le Développement.
- Lucchesi, Dante. 1998b. *Sistema, mudança e linguagem*. Lisboa: Colibri.
- 2000b. *Reanálise da variação na concordância de gênero em um dialeto afro-brasileiro*. XVIII Jornada de Estudos Linguísticos do GELNE. Salvador: Universidade Federal da Bahia.
- 2008b. A concordância nominal em estruturas passivas e de predicativo do sujeito em comunidades rurais afro-brasileiras isoladas no contexto da história sociolinguística do Brasil. In Sebastião Votre & Cláudia Roncarati (orgs.), *Anthony Julius Naro e a lingüística no Brasil: uma homenagem acadêmica*, 1<sup>a</sup>. ed., vol. 1, 148-168. Rio de Janeiro: 7 Letras.
- 2009b. A metodologia. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds.), *O português afro-brasileiro*, 155-164. Salvador: EDUFBA.
- Lucchesi, Dante, Alan Norman Baxter & Ilza Ribeiro (eds.). 2009. *O português afro-brasileiro*. Salvador: EDUFBA.
- Lucchesi *et alii*. 2009. O português afro-brasileiro: as comunidades analisadas. In Dante Lucchesi, Alan Norman Baxter & Ilza Ribeiro (eds), *O português afro-brasileiro*, 75-100. Salvador: EDUFBA.
- Lucchesi, Dante & Silvana Araújo. s.d. A sociolinguística variacionista: fundamentos teóricos e metodológicos [em linha]. In *Projecto vertentes do português rural do Estado da Bahia*. [Consult. 09 Nov. 2008]. Disponível em: <http://www.vertentes.ufba.br/socio.htm>.
- Lüdi, Georges & Bernard Py. 1986. *Être bilingue*. Berne: Peter Lang.
- Lumsden, John. S. 1999a. Language acquisition and creolization. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 129-157. Cambridge, MA: The MIT Press.
- 1999b. The role of relexification in creole genesis. *Journal of Pidgin and Creole Languages*, 14. 225-258.
- Lust, Barbara, Margarita Suaner & John Whitman (eds.). 1994. *Syntactic theory and first language acquisition: cross-linguistic perspectives*, vol. 2. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Lyons, John. 1999. *Definiteness*. Cambridge, MA: Cambridge University Press.
- Macambira, José Rebouças. 1999. *A estrutura morfo-sintática do português: aplicação do estruturalismo linguístico*. São Paulo: Ed. Pioneira.
- MacDonald, Marguerite G. 1988. *Fossilization and an emerging social dialect*. *Lenguas Modernas*, 15. 115-124.
- Macedo, Alzira Vertheim Tavares de. 2003. Linguagem e contexto. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 60-66. São Paulo: Contexto.

- Machado, Augusto Reis (ed.). s.d. *Viagem de Lisboa à Ilha de São Tomé escrita por um piloto português (século XVI)*. Lisboa: Portugália Ed.
- Maho, Jouni Filip. 2003. A referencial classification of the bantu languages: an update of Guthrie's referential system. In Derek Nurse & Gérard Philippson (eds), *The bantu languages*. Routledge language family series, 4, 639-651. London & New York: Routledge.
- Maia, Padre António da Silva (Missionário Secular da Arquidiocese de Luanda – Angola). 1964a. *Dicionário complementar português-kimbundu-kikongo (Línguas nativas do centro e norte de Angola)*, 1ª. ed. Cucujães: Tipografia das Missões – Edição e propriedade do autor.
- 1964b. *Lições de gramática de quimbundu*. Cucujães: Escola Tipográfica das Missões.
- Maianga, José. 1980. A luta dos escravos em São Tomé no Século XVI. *África, Literatura, Arte e Cultura*, 9. 437-443.
- Mantero, Francisco. 1954 [1910]. *A mão d'obra em São Tomé e Príncipe* [versão facsimile]. Lisboa: Empresa Nacional de Publicidade.
- Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis et alli (eds.), *University of Pennsylvania Working Papers in Linguistics*, vol. 4(2), Proceedings of the 21st Annual Penn Linguistics Colloquium, 201-225. Philadelphia, PA: University of Pennsylvania.
- Marques, Maria Emília Ricardo. 2003. *Português, língua segunda*. Lisboa: Universidade Aberta.
- Marques de Barros, Cónego Marcelino. 1887. O guinéense. *Revista Lusitana*, vol. 174/181. 271-302.
- Marquilhas, Rita. 1996. Mudança linguística. In Isabel Hub Faria et alii (coords.), *Introdução à linguística geral e portuguesa*, 563-588. Lisboa: Ed. Caminho.
- Martinet, André. 1974. *A linguística sincrónica*. Coleção “Biblioteca Tempo Universitário”. Rio de Janeiro: Ed. Tempo Brasileiro.
- Martins, Ana Maria. 2000a. Polarity itens in romance: underspecification and lexical change. In Susan Pintzuk, George Tsoulas & Anthony Warner (eds.), *Diachronic syntax: models and mechanisms*, 232-248. Oxford/New York, NY: Oxford University Press.
- Martins, Cristina dos Santos Pereira. 1997. Bilinguismo e manifestações verbais bilingues. Uma breve sinopse teórica. *Revista Portuguesa de Filologia*, 21. 63-125.
- Master, Peter, John H. Schumann & Margaret E. Sokolik. 1989. The experimental creation of a pidgin language. *Journal of Pidgin and Creole Languages*, 4(1). 37-63.
- Mateus, Maria Helena Mira et alii. 1990. *Fonética, fonologia e morfologia do português*. Lisboa: Unv. Aberta.
- 2003 [1982]. *Gramática da língua portuguesa*, 7ª ed. Lisboa: Caminho – Coleção Universitária/Série Linguística.
- Mateus, Maria Helena Mira & Fernanda Bacelar do Nascimento (orgs.). 2005. *A língua portuguesa em mudança*. Lisboa: Ed. Caminho.
- Mather, Patrick-André. 2000. Creole genesis: evidence from West African L2 french. In Dicky Gilbers, John Nerbonne & Jos Schaecken (eds.), *Languages in contact*, 247-261. Amsterdam/Atlanta, GA: Rodopi.

- 2001a. Revue of Lefebvre (1998), creole genesis and the acquisition grammar: The case of the haitian creole. *Studies in Language*, 25(1). 125-137.
- Maurer, Phillippe. 1995. *L'angolar. Un créole afro-portugais parlé à São Tomé*. Hamburgo: Buske.
- McEnery, Tony, Richard Xiao & Yukio Tono. 2006. *Corpus-based language studies. An advanced resource book*. London/New York, NY: Routledge.
- McWorther, John H. 1997. *Towards a new model of creole genesis*. New York, NY: Peter Lang.
- 1998. Identifying the creole prototype: vindicating a typological class. *Language*, 74(4). 788-818.
- 2002. The world's simplest grammars are creole grammars. *Linguistic Typology*, 5(2/3). 125-166.
- Meillet, Paul Antoine. 1912. L'évolution des formes grammaticales. *Scientia (Revista di Scienza)*, 6(12). 384-400. [Reimpresso in Meillet, Antoine. 1921. *Linguistique historique et linguistique générale*, 130-141. Paris: Honoré Champion].
- Meisel, Jürgen M. 1983. Transfer as a second-language strategy. *Language & Communications*, 3(1). 11-46.
- 1994c. Code-switching in young bilingual children: the acquisition of grammatical constraints. *Studies in Second Language Acquisition*, 16. 413-441.
- Mello, Heliana Ribeiro de et alii. 1998. O português vernáculo do Brasil. In Matthias Perl & Armin Schwegler (eds.), *América negra: panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*, 71-173. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
- 2002. Semicreolization? – The restructured portuguese of the tongas of São Tomé – a consequence of L1 acquisition in a special contact situation. *Journal of Portuguese Linguistics*, 1. 7-39.
- Mello, Heloísa Augusta Brito de. 2003. Atitudes lingüísticas de adolescentes americano-brasileiros de uma comunidade bilíngüe no interior de Góias. *Letras*, vol. 22, 1/2. 85-114. Campinas, SP: Centro de Linguagem e Comunicação/Pontifícia Universidade Católica.
- Melo, Gladstone Chaves de. 1946. *A língua do Brasil*. Rio de Janeiro: Agir.
- Mendonça, Renato de. 1933. *A influência africana no português do Brasil*. Rio de Janeiro: Sauer.
- Mesquitela, Teresa. 1987. A cultura afro-brasileira. *África Hoje*, A.3, 23. 39-41.
- Mesthrie, Rajend et alii. 2009 [2002]. *Introducing sociolinguistics*, 2<sup>nd</sup> ed. Edinburgh: Edinburgh University Press.
- Migge, Bettina. 1998. Substrate influence in creole formation: the origin of give-type serial verb constructions in the surinamese plantation creole. *Journal of Pidgin and Creole Languages*, 13(2). 215-265.
- Milroy, Lesley & Matthew Gordon. 2003. *Sociolinguistics: method and interpretation*. Oxford/Malden, MA: Blackwell.
- Ministério do Exército de Portugal. 1969. *Carta de contratados. Ministério do Exército – Rep. Gab. Ministro, 2<sup>a</sup>. Secção, 24. Jan. 1969*. Caixa n.º. 1 proc. PT AHM/FO/006/J/1415/579/146, doc. 927, São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).

- Móia, Telmo Lopes. 1992a. Proposta de revisão da ‘elevação de objecto’ no quadro da teoria da regência e da ligação [em linha]. *Actas do VII encontro da Associação Portuguesa de Linguística (Lisboa 1991)*, 257-270. Lisboa: Associação Portuguesa de Linguística. [Consult. 24 Fev. 2008]. Disponível em: [http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia\\_apl1991.pdf](http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia_apl1991.pdf).
- 1992b. Sobre classes semânticas de adjetivos [em linha]. *Cadernos de Semântica*, 7. Faculdade de Letras da Universidade de Lisboa. [Consult. 24 Fev. 2008]. Disponível em: [http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia\\_adjectivos1992.pdf](http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia_adjectivos1992.pdf).
- 1992c. *A sintaxe das orações relativas sem antecedente expresso do português*. Lisboa: Faculdade de Letras da Universidade de Lisboa. Dissertação de mestrado.
- 1993a. Aspectos da modificação das estruturas nominais [em linha]. *Discursos. Estudos de Língua e Cultura Portuguesa*, 4 - Semântica das estruturas nominais, 37-63. Coimbra: Universidade Aberta. [Consult. 12 Dez. 2009]. Disponível em: [http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia\\_discursos4.pdf](http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia_discursos4.pdf).
- 1993b. Sobre o lugar dos demonstrativos na arquitectura semântica do sintagma nominal [em linha]. *Cadernos de Semântica*, 11. 2ª. ed. revista. Faculdade de Letras da Universidade de Lisboa. [Consult. 24 Fev. 2008]. Disponível em: [http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia\\_demonstrativos1993.pdf](http://www.fl.ul.pt/dlgr/pessoais/tmoia/tmoia_demonstrativos1993.pdf).
- Móia, Telmo Lopes & Ana Teresa Alves. 2004. Differences between european and brazilian portuguese in the use of temporal adverbials. *Journal of Portuguese Linguistics*, 3.1. 37-67.
- Moita, Irisalva. 1989. O Congo – primeira tentativa de uma colonização continental (1484-1510). In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do Ocidente; origem e desenvolvimento da colonização*, vol. 2, 198-214. Lisboa: Publ. Alfa.
- Mollica, María Cecília & Maria Luiza Braga (orgs.). 2003. *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto.
- Morais-Barbosa, Jorge (ed.). 1967. *Estudos linguísticos crioulos*. Lisboa: Academia Internacional da Cultura Portuguesa/Tip. Silvas, Lda. [Reedição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa].
- 1983. *Études de phonologie portugaise*, 2ª. ed. Évora: Universidade de Évora.
- Mufwene, Salikoko S. 1986. Les langues créoles peuvent-elles être définies sans allusion à leur histoire?. *Etudes Créoles*, 9. 135-150.
- 1990. Transfer and substrate hypothesis in creolistics. *Studies in Second Language Acquisition*, 12. 1-23.
- 2002. Competition and selection in language evolution. *Selection* 3, 45-56.
- Mühlhäusler, Peter. 1974. Pidginization and simplification of language. *Pacific Linguistics*, 26, series B. Canberra: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.
- 1980. Structural expansion and the process of creolization. In Albert Valdman & Arnold Highfield, A. (orgs.), *Theoretical orientations in creole studies*, 19-55. New York, NY: Academic Press.



- Müller, Ana Lúcia de Paula. 2002. Nomes nus e o parâmetro nominal no português brasileiro [em linha]. *Revista Letras* (Curitiba), 58. 331-344. [Consult. 26 Dez. 2009]. Disponível em:  
[http://74.125.153.132/search?q=cache:QtXbQ5HY7FIJ:www.fflch.usp.br/dl/anamuller/pdf/nus.pdf+nomes+nus&cd=1&hl=pt-PT&ct=clnk&gl=pt&lr=lang\\_pt](http://74.125.153.132/search?q=cache:QtXbQ5HY7FIJ:www.fflch.usp.br/dl/anamuller/pdf/nus.pdf+nomes+nus&cd=1&hl=pt-PT&ct=clnk&gl=pt&lr=lang_pt).
- 2003. A semântica do sintagma nominal. In Ana Lúcia Müller, Esmeralda V. Negrão & Maria José Foltran (orgs.), *Semântica formal*, 1ª. ed., 61-74. São Paulo: Ed. Contexto.
- Nakuma, Constancio K. 1998. A new theoretical account of “fossilization”: implications for L2 attrition research. *International Review of Applied Linguistics*, (36)3. 247-256.
- Naro, Anthony Julius. 1978. A study on the origins of pidginization. *Language*, 54(2). 314-347.
- 2003b. O dinamismo das línguas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 43-50. São Paulo: Contexto.
- Naro, Anthony Julius & Maria Marta Pereira Scherre. 1996. Contact with media and linguistic variation. In Jennifer Arnold *et alii* (eds.), *Sociolinguistic: variation data, theory, and analysis*, 223-228. Stanford: CSLI Publications.
2003. O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In Cláudia Roncarati & Jussara Abraçado (orgs.), *Português brasileiro: contacto linguístico, heterogeneidade e história*, 285-302. Rio de Janeiro: 7Letras.
- 2007a. *Origens do português brasileiro*. São Paulo: Parábola.
- 2007b. Concordância variável em português: a situação no Brasil e em Portugal. In Anthony Julius Naro & Maria Marta Pereira Scherre (orgs.), *Garimpo das origens do português brasileiro*, 49-69. (Lingua[gem], 20). São Paulo: Parábola Ed.
- Nascentes, Antenor. 1953. *O linguajar carioca*, 2ª. ed. Rio de Janeiro: Organização.
- Negreiros, Almada. s.d. *História ethnographico da Ilha de S. Thomé*. Lisboa: Antica Casa Bertrand-José Bastos.
- 1928. Etnografia de São Tomé e outros elementos linguísticos. *Anuário Comercial, Industrial e Agrícola da Província de São Tomé e Príncipe*.
- Neto, Serafim da Silva. 1960. *A língua portuguesa no Brasil. Problemas*. Rio de Janeiro: Acadêmica.
- 1976 [1950]. *Introdução ao estudo da língua portuguesa no Brasil*, 3ª. ed. Rio de Janeiro: Presença.
- Neumann-Holzschuh, Ingrid & Edgar W. Schneider (eds.). 2000. *Degrees of restructuring in creole languages* – Creole Languages Library (CLL), vol. 22. University of Regensburg. Amsterdam/Philadelphia: John Benjamins.
- Neves, C. Agostinho das, J. Manuel Flores & A. Teodoro de Matos. 1989. A repressão contra os escravos de São Tomé (1595) e a guerra em Ceilão (1587-1617). In Luís de Albuquerque (dir.), *Portugal no mundo – a decadência do império português; a recuperação possível; o imperialismo português no Brasil nos séculos XVIII-XIX*, vol. 5, 100-112. Lisboa: Publ. Alfa.
- Neves, Maria Helena de Moura. 1999 (org.). *Gramática do português falado: novos estudos*, vol. VII. Campinas: Unicamp/São Paulo.

- 2000 [1999]. *Gramática de usos do português*, 4ª reimpressão. São Paulo: Ed. UNESP.
- Newmeyer, Frederick. J. 1983. *Grammatical theory: its limits and its possibilities*. Chicago, IL: University of Chicago Press.
- Newport, Elissa L. 1999. Reduced input in the acquisition of signed languages: contributions to the study of creolization. In Michel DeGraff (ed.), *Language creation and language change. Creolization, diachrony and development*, 161-178. Cambridge, MA: The MIT Press.
- Nogueira, Rodrigo de Sá. 1950/1951. As línguas bantas e o português: temas de linguística banta. *Estudos Coloniais: Revista de Escola Superior*, 2.
- 1952. Temas de linguística banta: dos elementos prefixativos bantos. *Estudos Coloniais, Estudos da Escola Superior Colonial*, 3. 5-38.
- Nunes, José Joaquim. 1989 [1919]. *Compêndio de gramática histórica portuguesa (fonética e fonologia)*, 9ª ed.. Lisboa: Clássica Ed..
- Odlin, Terence. 1989. *Language transfer*. Cambridge, MA: Cambridge University Press.
- Ogliari, Maria Marlene. 2003. Contato, diglossia e bilingüismo: situações linguísticas gestadas em Prudentópolis-PR [em linha]. *Anais do V Encontro do Círculo de Estudos Linguísticos do Sul - CELSUL, Universidade Federal do Paraná – UFPR, Curitiba - PR, 17 e 18 de Outubro de 2002*. 1075-1082. [Consult. 21 Nov. 2010]. Disponível em:  
<http://www.celsul.org.br/Encontros/05/pdf/149.pdf>.
- Oliveira, Dercir Pedro. 2008. A variação linguística no Brasil. In Cláudia Roncarati & Jussara Abraçado (orgs.), *Português brasileiro II: contacto linguístico, heterogeneidade e história*, 93-100. Niterói: Ed. UFF.
- Oliveira, Fernão de. 1975 [1536]. *A gramática da linguagem portuguesa* [Leit., introd. e anot. por Maria Leonor Carvalhão Buescu]. Lisboa: INCM.
- Oliveira e Silva, Giselle Machline. 1996. Realização facultativa do artigo definido diante de possessivo e de patronímico. In Giselle Machline Oliveira e Silva & Maria Marta Pereira Scherre (orgs.), *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*, 119-145. Rio de Janeiro: Tempo Brasileiro.
- Omena, Nelize Pires de & Maria Eugênia Lamoglia Duarte. 2003. Variáveis morfossintáticas. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 81-88. São Paulo: Contexto.
- Paiva, Maria da Conceição A. de. 2003a. A variável gênero/sexo. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 33-42. São Paulo: Contexto.
- 2003b. Transcrição de dados linguísticos. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 135-146. São Paulo: Contexto.
- Paiva, Maria da Conceição A. de & Maria Eugênia Lamoglia Duarte. 2003. Mudança linguística: observações no tempo real. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 179-190. São Paulo: Contexto.
- Paradis, Michel. 1977. Bilingualism and aphasia. In Haiganoosh Whitaker & Harry A. Whitaker, (eds.), *Studies in neurolinguistics*, vol. 3, 65-122. New York, NY: Academic Press.

- 1978 (ed.). *Aspects of bilingualism*. Columbia, South Carolina: Hornbeam Press.
- 1983. *Readings on aphasia in bilinguals and polyglots*. Montreal: Didier.
- 1993. Linguistic, psycholinguistic and neurolinguistic aspects of “interference” in bilingual speakers: the activation threshold hypothesis. *International Journal of Psycholinguistics*, 2. 133-145.
- 1997. The cognitive neuropsychology of bilingualism. In Annette M. B. DeGroot & Judith F. Kroll (eds.), *Tutorials in bilingualism: psycholinguistic perspectives*, 331-354. Mahwah, NJ: Lawrence Erlbaum Associates.
- Parkvall, Mikael. 2000. The alleged creole past of Brazilian vernacular Portuguese. In Ernesto d’ Andrade, Maria Antónia Coelho da Mota & Dulce Pereira (orgs.), *Crioulos de base portuguesa (Actas do workshop sobre crioulos de base lexical portuguesa. FLUL, 29 e 30 de Abril de 1999)*, 223-246. Braga: Associação Portuguesa de Linguística.
- Partee, Barbara Hall, Alice G. B. ter Meulen & Robert Eugene Wall. 1990. *Mathematical methods in linguistics*. Dordrecht: Kluwer Academic Publishers.
- Patkowski, Mark. 1980. The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30(2). 449-472.
- Paulston, Christina Bratt & G. Richard Tucker (eds.). 2006 [2003]. *Sociolinguistics: the essential readings*. Oxford/Cambridge, MA: Blackwell.
- Peet Jr., William. 1978. *Relativization in a creole continuum*. Manoa: University of Hawaii. Dissertação de doutoramento.
- Percegon, Marcélia Silva. 2005. *A fossilização no processo de aquisição de segunda língua* [em linha]. Curitiba: Universidade Federal do Paraná. Dissertação apresentada como requisito parcial para obtenção do grau de Mestre. [Consult. 02 Abr. 2009]. Disponível em:  
[http://www.diaadiaeducacao.pr.gov.br/diaadia/diadia/arquivos/File/conteudo/artigos\\_teses/Ingles/percegon.pdf](http://www.diaadiaeducacao.pr.gov.br/diaadia/diadia/arquivos/File/conteudo/artigos_teses/Ingles/percegon.pdf).
- Perdue, Clive. 1993a, (ed.). *Adult language acquisition: cross-linguistic perspective. Vol. 1: field methods*. Cambridge, MA: Cambridge University Press.
- 1993b. *Adult language acquisition: cross-linguistic perspective. Vol. 2: the results*. Cambridge, MA: Cambridge University Press.
- Pereira, Dulce. 1996. Crioulo de Cabo Verde. In Isabel Hub Faria *et alii* (coords.), *Introdução à linguística geral e portuguesa*, 551-559. Lisboa: Ed. Caminho.
- 2004. Contacto de línguas e aquisição de uma língua não materna. *Projecto “vamos conversar na escola – nu bem papia na skola”*. Lisboa: Escola Superior João de Deus.
- Pereira, Maria Ângela Botelho. 1984. *Gênero e número em português. Estudo das relações forma-sentido na gramática*. Rio de Janeiro: UFRJ. Dissertação de mestrado.
- Petter, Margarida Maria Taddoni. 2009. Traços morfossintáticos comuns às variedades angolana, brasileira e moçambicana de português. *Papia – Revista Brasileira de Estudos Crioulos e Similares*, 19. 201-220.
- Pfaff, Carol W. 1992. The issue of grammaticalization in early German second language acquisition. *Studies in Second Language Acquisition*, 14. 273-296.
- Pinker, Steven. 1994. *The language instinct: the new science of language and mind*. Londres: Penguin.

- 2002. *O instinto da linguagem: como a mente cria a linguagem*. São Paulo: Ed. Martins Fontes.
- Pintzuck, Susan, George Soulas & Anthony Warner (orgs.). 2000. *Diachronic syntax, models and mechanisms*. Oxford University Press.
- Plag, Ingo. 2001. The nature of derivational morphology in creoles and non-creoles. *Journal of Pidgin and Creole Languages*, 16(1). 153-160.
- 2003a (ed.). *Phonology and morphology of creole languages*. Tübingen: Max Niemeyer Verlag.
- 2003b. Introduction: the morphology of creole languages. In Geert Booji & Jaap van Marle (eds.), *Yearbook of morphology 2002*, 1-2. Dordrecht: Kluwer Academic Publishers.
- 2005. Morphology in pidgins and creoles. In Keith Brown (ed.), *Encyclopedia of language and linguistics*, 2<sup>nd</sup> ed., vol. 8, 304-308. Oxford: Elsevier.
- Ploae-Hanganu, Mariana. s.d. *Le créole portugais de l’Afrique - sa base portugaise*, vols. 1 e 2. Bucareste: Instituto de Linguística da Universidade de Bucareste. Dissertação de doutoramento.
- Pontes, Jerónimo Xavier de Sousa. 2005. *História da educação em São Tomé e Príncipe*. Lisboa: Universidade de Lisboa. Dissertação de mestrado.
- Poplack, Shana. 1990. Variation theory and language contact: concepts, methods and data. In Dennis R. Preston (ed.), *American dialect research – celebrating the 100<sup>th</sup> anniversary of the American Dialect Society 1889-1989*, 251-286. Amsterdam: John Benjamins.
- Pratas, Fernanda. 2002. *O sistema pronominal do caboverdiano (de Santiago) – questões de gramática*. Lisboa: FCSH – Universidade de Lisboa. Dissertação de mestrado.
- Quelhas, António Antunes. 1965. *A colonização portuguesa em S.Tomé e Príncipe no século XIX*. Lisboa: Universidade Técnica.
- Quint, Nicolas. 2008. A realização do sujeito em português do Brasil: deriva versus criouliização. In José Luiz Fiorin & Margarida Petter (orgs.), *África no Brasil – a formação da língua portuguesa*, 75-88. São Paulo: Ed. Contexto.
- Quint-Abrial, Nicolas. 1998. *O dicionário cabo-verdiano-português de Santiago*. Lisboa: Verbalis.
- Radford, Andrew. 1997. *Syntax: a minimalist introduction*. Cambridge, MA: Cambridge University Press.
- 2000. *Children in search of perfection: towards a minimalist model of acquisition*[em linha]. [Consult. 27 Mar. 2008]. Disponível em: <http://privatewww.essex.ac.uk/~radford/PapersPublications/perfection.htm>.
- Ramat, Anna Giacalone. 1992. Grammaticalization processes on language transfer. *Applied Linguistics*, 8(2). 111-136.
- Ramchand, Gillian & Charles Reiss (eds.). 2007. *The Oxford handbook of linguistic interfaces*. Oxford/New York, NY: Oxford University Press.
- Raposo, Eduardo Paiva. 1973. Sobre a forma “o” em português. *Boletim de Filologia*, tomo XXII, fascículos 3 e 4. 361-415.
- 1978. *Introdução à gramática generativa: sintaxe do português*. Lisboa: Moraes Ed.
- Reinecke, John. 1937. *Marginal languages: a sociological survey of the creole languages and trade jargons*. Ann Arbor: University Microfilms International. Dissertação de doutoramento, Yale University.

- Rickford, John R. 1987. *Dimensions of a creole continuum*. Stanford: Stanford University Press.
- Rio-Torto, Graça Maria. 1998. *Morfologia derivacional. Teoria e aplicação ao português*. Porto: Porto Ed.
- Roberts, Ian. 2000. Head movement. In Mark Baltin & Chris Collins (eds.). *The handbook of contemporary syntactic theory*, 113-147. Oxford/Malden, MA: Blackwell Publishers, Ltd (edition for sale in the mainland territory of the People's Republic of China only).
- Roeper, Thomas & Williams, Edwin. 1987. *Parameter setting*, 123-172. Dordrecht: Reidel.
- Rojas, Juan Pedro. 2006. *Processo de fossilização na interlíngua de hispanofalantes aprendizes de português no Brasil: acomodação consentida?* [em linha]. Brasília: Universidade de Brasília – Instituto de Letras. Dissertação de mestrado. [Consult. 31 Mar. 2009]. Disponível em:  
[http://bdt.d.bce.unb.br/tesdesimplificado/tde\\_busca/arquivo.php?codArquivo=641](http://bdt.d.bce.unb.br/tesdesimplificado/tde_busca/arquivo.php?codArquivo=641).
- Roncarati, Cláudia & Jussara Abraçado (orgs.). 2003. *Português brasileiro: contacto lingüístico, heterogeneidade e história*. Rio de Janeiro: 7Letras.
- 2008 (orgs.). *Português brasileiro II: contacto lingüístico, heterogeneidade e história*. Niterói: Ed. UFF.
- Rougé, Jean-Louis. 2004. *Dictionnaire étymologique des créoles portugais d'Afrique*. Paris: Eds. Karthala.
- 2008. A inexistência de crioulo no Brasil. In José Luiz Fiorin & Margarida Petter (orgs.), *África no Brasil – a formação da língua portuguesa*, 63-73. São Paulo: Ed. Contexto.
- 2010. *Créolisation et productivité morphologique: les créoles portugais d'Afrique*. 10ème Colloque International de l' Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLe). Llacan, Campus CNRS de Villejuif, Paris, France. 1-3 de Julho.
- Russel-Wood, A. J. R. 1992. *A world on the move: the portuguese in Africa, Asia and America 1415-1808*. New York, NY: St. Martin's Press.
- Salvaterra, Hélder. 2004a. *Projeções demográficas de São Tomé e Príncipe no horizonte 2025: actualização na base dos resultados do censo 2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- 2004b. *Resultados globais : III recenseamento geral da população e da habitação: RGPH-2001*. São Tomé: Instituto Nacional de Estatística de São Tomé e Príncipe.
- Sambú, Malam. 1999. *Ditadura e problemas sociais na África lusófona*. Macau: Edição do autor.
- Sánchez, Liliana. 2006. Bilingual grammars and creoles: similarities between functional convergence and morphological elaboration. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: dialogues*, 277-294. Amsterdam/Philadelphia: John Benjamins.
- Sankoff, David. 1978. *Linguistic variation: models and methods*. New York, NY: Academic Press.
- 1988. *Variable rules*. In Ulrich Ammon, Norbert Dittmar & Klaus Mattheier (eds.), *Sociolinguistics – an international handbook of the science of language and society*, 984-997. Berlin/New York, NY: Walter de Gruyter.

- 1991. Using the future to explain the past. In Francis Byrne & Thorn Huebner (eds.), *Development and structures of creole languages*, 61-74. Amsterdam: John Benjamins.
- Sankoff, David & Pierrette Thibault. 1981. Weak complementarity: tense and aspect in Montreal french. In B. B. Johns & D. R. Strong (eds.), *Syntactic change*, vol. 25, 205-215. Ann Arbor: University of Michigan.
- Sankoff, David & William Labov. 1979. On the Uses of Variable Rules. *Language in Society*, 8(2). 189-222.
- Santos, Catarina Madeira. 1996. A formação das estruturas fundiárias e a territorialização das tensões sociais: São Tomé, primeira metade do século XVI. *Africana Studia : Revista Internacional de Estudos Africanos*, 54/55. 51-91.
- Santos, Maria Emília Madeira. 1989. Os primeiros “lançados” na costa da Guiné: aventureiros e comerciantes. In Luís de Albuquerque (dir.), *Portugal no mundo – as zonas de influência do Ocidente; origem e desenvolvimento da colonização*, vol. 2, 125-136. Lisboa: Publ. Alfa.
- Santos, Maria Emília Madeira *et alii*. 1991a (eds.). *História geral de Cabo Verde*, vol. I. Lisboa: IICT/Praia: Direcção Geral do Património Cultural de Cabo Verde.
- 1991b (eds.). *História geral de Cabo Verde*, vol. II. Lisboa: Centro de Estudos de História e Cartografia Antiga; IICT/Praia: Instituto Nacional da Cultura de Cabo Verde.
- Santos, Raimundo Enedino dos. 2009. *The behavior of clitics in the portuguese of the Tongas as a result of linguistic contact*. WOCAL6 - World Congress of African Languages. Colónia, Alemanha, 17-21 de Agosto.
- Scantamburlo, Luigi. 1999. *Dicionário do guineense*, vol. I: *introdução e notas gramaticais*. Lisboa: Ed. Colibri e FASPEBI.
- 2002. *Dicionário do guineense: dicionário guineense-português. Disionariu guinensi-purtuguis*, vol. 2. Bissau: FASPEBI.
- Scherre, Maria Marta Pereira. 1996. Pressupostos teóricos e suporte quantitativo. In Giselle Machline de Oliveira e Silva & Maria Marta Pereira Scherre (orgs.), *Padrões sociolinguísticos*, 37-50. Rio de Janeiro: Ed. Tempo Brasileiro.
- 1998b. Paralelismo lingüístico. *Estudos da Linguagem*, 7. 29-59.
- Scherre, Maria Marta Pereira & Anthony Julius Naro. 2000. *Garimpando as origens estruturais do português brasileiro* [em linha]. Congresso Internacional – 500 anos de Língua Portuguesa no Brasil. Universidade de Évora, Évora, Portugal, 8-13 Maio. [Consult. 25 Jan. 2010]. Disponível em: [http://www.marcosbagno.com.br/conteudo/arquivos/for\\_scherre\\_naro.htm](http://www.marcosbagno.com.br/conteudo/arquivos/for_scherre_naro.htm).
- Schmitt, Cristina & Alan Munn. 1999. Against the nominal mapping parameter: Bare nouns in brazilian portuguese. In Pious Tamanji, Masako Hirotoni & Nancy Hall (orgs.), *Proceedings of the North-Eastern Linguistics Society (NELS) 29*, 339-355. Amherst, MA: GLSA, University of Massachusetts.
- Schuchardt, Hugo Ernst Mario. 1979. On creole portuguese. In T. L. Markey (ed. e trad.), *The ethnography of variation – selected writings of pidgins and creoles: Hugo Schuchardt*, 59-72. Ann Arbor: Karoma.
- Science & Vie . 2004. Découvertes: du langage aux langues. *Hors-série*, 227.
- Sebba, Mark. 1997. *Contact languages: pidgins and creoles*. New York, NY: St. Martin's Press.

- Sequeira, Francisco Júlio Martins. s.d. *Gramática histórica da língua portuguesa*. Lisboa: Liv. Popular.
- Selinker, Larry & John Lamendella. 1978. Two perspectives on fossilization in interlanguage learning. *Interlanguage Studies Bulletin*, 3. 143-191.
- 1979. The role of extrinsic feedback in interlanguage fossilization: a discussion of “rule fossilization”. A tentative model. *Language Learning*, 29(2). 363-375.
- Sheen, Ronald. 1987. The importance of negative transfer in the speech of near bilinguals. In Dietrich Nehls (ed.), *Interlanguage studies*, 43-57. Heidelberg: Julius Groos Verlag.
- Siegel, Jeff. 1997. Mixing, leveling and pidgin/creole development. In Arthur K. Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 111-149. Amsterdam/Philadelphia: John Benjamins.
- 1999. Transfer constraints and substrate influence in melanesian pidgin. *Journal of Pidgin and Creole Languages*, 14. 1-44.
- 2000. Substrate influence in Hawaií creole English. *Language in Society*, 29. 197-236.
- 2003. Substrate influence in creoles and the role of transfer in second language acquisition. *Studies in Second Language Acquisition*, 25(2). 185-209.
- Silva, Baltasar Lopes da. 1984. *O dialecto crioulo de Cabo Verde*. Lisboa: Imprensa Nacional.
- Silva, Giselle Machline de Oliveira e. 2003. Coleta de dados. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 117-133. São Paulo: Contexto.
- Silva, Jaline Pinto da. 2001. Contribuições africanas nos falares do Brasil. In José Pereira da Silva (ed.), *História da língua portuguesa: cadernos da pós-graduação em língua portuguesa*, 1 [em linha]. [Consult. 24 Jan. 2009]. Disponível em: [http://www.filologia.org.br/monografias/historia\\_da\\_lingua\\_portuguesa.html](http://www.filologia.org.br/monografias/historia_da_lingua_portuguesa.html).
- Silva, Jorge Augusto Alves da. 2003. *A concordância verbal no português afro-brasileiro: um estudo sociolinguístico de três comunidades rurais do Estado da Bahia*. UFBA. Dissertação de mestrado.
- Silva, Rosa Virgínia Mattos e. 1989. *Estruturas trecentistas. Para uma gramática do português arcaico*. Lisboa: INCM.
- 1991. *O português arcaico. Fonologia*. São Paulo: Contexto, Ed. da Universidade Federal da Bahia.
- 1994. *O Português arcaico. Morfologia e sintaxe*. São Paulo: Contexto, Ed. da Universidade Federal da Bahia.
- 2005 [1996]. *Contradições no ensino de português: a língua que se fala X a língua que se ensina. Repensando a língua portuguesa*. São Paulo: Ed. Contexto.
- Silva-Corvalán, Carmen. 1986. On the problem of meaning in sociolinguistic studies of syntactic variation. In Dieter Kastovsky & Aleksander Szwedek (eds.), *Linguistic across historical and geographical boundaries: in honour of Jacek Fidiak. Vol. 1: Linguistic theory and historical linguistics*, 111-123. The Hague: Mouton de Gruyter.
- 1989. *Sociolinguística. Teoria y análisis*. Madrid: Ed. Alhambra.
- Silva Neto, Serafim da. 1950. Falares crioulos. *Brasília*, 5. 1-28.
- 1963 [1950]. *Introdução ao estudo da língua portuguesa no Brasil*, 2ª. ed. Rio de Janeiro: Departamento de Imprensa Nacional.

- Simioni, Leonor. 2007. A concordância de número no DP: propostas minimalistas [em linha]. *Estudos Lingüísticos*, XXXVI(1). 117-125. [Consult. 18 Fev. 2010]. Disponível em:  
<http://www.gel.org.br/estudoslinguisticos/edicoesanteriores/4publica-estudos-2007/sistema06/12.PDF>.
- Singler, John Victor. 1988. The homogeneity of the substrate as a factor in pidgin/creole genesis. *Language*, 64. 27-51.
- 1990. On the use of sociohistorical criteria in the comparison of creoles. *Linguistics*, 28. 645-669.
- 1992. Nativization and pidgin/creole genesis: a reply to Bickerton. *Journal of Pidgin and Creole Languages*, 7. 319-333.
- 1993. African influence upon afro-american language varieties: a consideration of sociohistorical factors. In Salikoko S. Mufwene (ed.), *Africanisms in afro-american language varieties*, 235-253. Atenas: University of Georgia Press.
- 1996. Theories of creole genesis, sociohistorical considerations and evaluation of evidence: the case of haitian creole and the relexification hypothesis. *Journal of Pidgin and Creole Languages*, 11(2). 185-230.
- 2006. Children and creole genesis. *Journal of Pidgin and Creole Languages*, 21(1). 157-173.
- Slobin, Dan I. 1977. Language change in childhood and history. In John Theodore MacNamara (ed.), *Language learning and thought*, 185-214. New York, NY: Academic Press.
- Smith, Marilyn Chapnik. 1997. How do bilinguals access lexical information? In Annette M. B. DeGroot & Judith F. Kroll (eds.), *Tutorials in bilingualism: psycholinguistic perspectives*, 145-168. Mahwah, NJ: Lawrence Erlbaum Associates.
- Smith, Norval S. H. 2008. Creole phonology. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creoles studies*, 98-129. Oxford: Wiley-Blackwell.
- Sorace, Antonella. 1999. Initial stages, end-stages and residual optionality in L2 acquisition. In Annabel Greenhill, Heather Littlefield & Cheryl Tano (eds.), *Proceedings of the 23<sup>rd</sup> annual Boston University Conference on language development*, vol. 2, 666-674. Sommerville, MA: Cascadilla.
- Sousa, Antônio Rômulo Bezerra de et alii. 2007. A teoria inatista de aquisição da linguagem [em linha]. *Revista Virtual Partes*, ano V, 17-02-2007. [Consult. 04 Nov. 2009]. Disponível em:  
<http://www.partes.com.br/educacao/ateoriainatista.asp>.
- Souza, Antônio Carlos Santana de. s.d. *A constituição da identidade e dos territórios afro-brasileiros em Mato Grosso do Sul: estudos sócio-etnolinguísticos* [em linha]. [Consult. 09 Mar. 2009]. Disponível em:  
[http://www.fflch.usp.br/dl/gela/downloads/Souza/SOUZA\\_identidade\\_territorios\\_afro.pdf](http://www.fflch.usp.br/dl/gela/downloads/Souza/SOUZA_identidade_territorios_afro.pdf).
- Spencer, Andrew & Arnold M. Zwicky. 1998. *The handbook of morphology*. Oxford: Basil Blackwell.
- Stroud, Christopher & Perpétua Gonçalves (orgs.). 1997. *Panorama do português oral de Maputo – Volume II: a construção de um banco de “erros”*. Maputo: Instituto Nacional do Desenvolvimento da Educação.



- Stump, Gregory T. 2008. *Inflectional morphology: a theory of paradigm structure*. Cambridge, MA: Cambridge University Press.
- Szabolcsi, Anna. 1994. The noun phrase. In Ferenc Kiefer & Katalin É. Kiss (eds.), *Syntax and semantics: the syntactic structure of hungarian*, 27. 179-274. San Diego: Academic Press.
- Szendrői, Kriszta. 2001. *Focus and the syntax-phonology interface*. London: University College London. Dissertação de doutoramento.
- Tamburelli, Marco. The importance of paradigm formation in bilingual acquisition: evidence from italian [em linha]. In M. Pearce & N. Topinzi (eds.), *UCL working papers in linguistics*, 17. 27-57. [Consult. 06 Fev. 2009]. Disponível em: <http://www.langsci.ucl.ac.uk/linguistics/publications/WPL/05papers/tamburelli.pdf>.
- 2007. *The role of lexical acquisition in simultaneous bilingualism* [em linha]. University of London. Dissertação de doutoramento. [Consult. 15 Dez. 2008]. Disponível em: [http://www.phon.ucl.ac.uk/home/marco/thesis/m\\_tamburelli\\_chapter\\_one\\_emb.pdf](http://www.phon.ucl.ac.uk/home/marco/thesis/m_tamburelli_chapter_one_emb.pdf).
- Tarallo, Fernando. 1993. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In Ian Roberts & Mary Kato (orgs.), *Português brasileiro: uma viagem diacrônica*, 35-68. Campinas: Editora da Unicamp.
- Tarallo, Fernando & Tânia Alkmin. 1996. Turning different at the end of the century: 19<sup>th</sup> century brazilian portuguese. In Gregory Riordan Guy et alii (eds.), *Towards a social science of language*, vol. 1. Variation and change in language and society, 199-220. Amsterdam/Philadelphia: John Benjamins.
- Tavares, Cónego José Lourenço. 1942. *Das línguas e dialectos bantu de Angola*. Lisboa: Congresso Colonial.
- Tavares, Miguel Sousa. 2003. *Equador*. Alfragide: Oficina do Livro.
- Teyssier, Paul. 1993 [1980]. *História da língua portuguesa*. Lisboa: Ed. Sá da Costa.
- Thep-Ackrapong, Tipa. 1990. *Fossilization: a case study of practical and theoretical parameters*. Normal, Ill.: Illinois State University. Dissertação de doutoramento.
- Thomason, Sarah Grey. 1997. A typology of contact languages. In Arthur K. Spears & Donald Winford (eds.), *The structure and status of pidgins and creoles*, 71-88. Amsterdam: John Benjamins.
- Thompson, Robert Wallace. 1961. A note on some possible affinities between the creole dialects of the old world and those of the new. In Robert Brock Le Page (ed.), *Proceedings of the conference on creole language studies*, 107-113. London: Macmillan.
- Todd, Loreto. 1974. *Pidgins and creoles*. London: Routledge and Kegan Paul.
- Todorov, Tzvetan & Oswald Ducrot. 1977. *Dicionário das ciências da linguagem*, 4<sup>a</sup>. ed. Lisboa: Publicações D. Quixote.
- Tomás, Maria Isabel Gonçalves. 1999. *Os espaços do crioulo*. Lisboa: Comissão Nacional para as Comemorações dos Descobrimentos Portugueses.
- Tonioli, Selma & Vanda Bartalini Baruffaldi. 2007. Sociolingüística: uso e norma na fala urbana. *Revista da Pós-graduação, UNIFIEO*, 1(2).101-111.
- Traugott, Elizabeth Closs. 1977. Natural semantax: its rolle in the study of second language acquisition. In Stephen Pit Corder & Eddie Roulet (eds.), *The notions of simplification, interlanguage and pidgins and their relation to second language pedagogy*, 132-162. Geneva: Droz.

- Trudgill, Peter. 2002. *Sociolinguistic variation and change*. Washington, DC: Georgetown University Press.
- Unidade da Companhia de Artilharia 3376. 1973c. *História da Unidade da Companhia de Artilharia 3376. Anexo "B" ao estudo de situação de informações n° 1/73*. Caixa n° 1, proc. PT AHM/DIV/12/08/01/21, doc. s.n., São Tomé e Príncipe. Lisboa: AHM – Secção de Investigação, Leitura e Divulgação (SILD).
- Vainikka, Anne & Martha Young-Scholten. 1994. Direct access to x'-theory: Evidence from korean and turkish adults learning german. In Teun Hoekstra & Bonnie D. Schwartz (eds.), *Language acquisition studies in generative grammar*, 265-316. Amsterdam: John Benjamins.
- 1996a. Gradual development of L2 phrase structure. *Second Language Research*, 12. 7-39.
- 1996b. The early stages in adult L2 Syntax: additional evidence from romance speakers. *Second Language Research*, 12. 140-176.
- Valdman, Albert. 1978. *Le créole: structure, statut et origine*. Paris: Klincksieck.
- Van Coetsen, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Patten, Bill. 1996. *Input processing and grammar instruction in second language acquisition*. Norwood, N.J.: Ablex.
- Varela, Soledad. 1992. *Fundamentos de morfología*. Madrid: Sintesis.
- Veiga, Manuel. 1982. *Diskrison strutural di lingua kabuverdianu*. Prai: Institutu Kabuverdianu ddi Livru.
- Veiga, Margareth. 2002b. *A situação actual do kimbundo: até que ponto o kimbundo existe como língua?* Trabalho de licenciatura apresentado no âmbito da disciplina de Seminário. Universidade de Lisboa, Faculdade de Letras – Departamento de Língua e Cultura Portuguesa.
- Véronique, Daniel. 1994. Naturalistic adult acquisition of french as L2 and french-based creole genesis compared: insights into creolization and language change?. In Dany Adone & Ingo Plag (eds.), *Creolization and language change (Linguistische Arbeiten, 317)*, 117-137. Tübingen: Max Niemeyer Verlag.
- Vigário, Marina, Sónia Frota & Fernando Martins. 2007. Frequência de unidades e padrões fonológicos no português europeu e no português do Brasil: uma prospectiva. *Workshop Domínios: Prosódia e Sintaxe*. Universidade Estadual de Campinas, Brasil. 16-17 Abril, 2007.
- Villalva, Alina. 2000. *Estruturas morfológicas. Unidades e hierarquias nas palavras do português*. Lisboa: Fundação Calouste Gulbenkian & FCT.
- 2003. Estrutura morfológica básica. In Maria Helena Mira Mateus et alii, *Gramática da língua portuguesa*, 7ª ed., cap. 22, 917-938. Lisboa: Caminho – Coleção Universitária/Série Linguística.
- Votre, Sebastião Josué. 2003. Relevância da variável escolaridade. In, Maria Cecília Mollica & Maria Luiza Braga (orgs.), *Introdução à sociolinguística: o tratamento da variação*, 51-57. São Paulo: Contexto.
- Weinreich, Uriel. 1968 [1953]. *Languages in contact: findings and problems*. The Hague: Mouton.
- Westphal, Ernst Oswald Johannes. 1958. An introductory comparative study of negation in bantu. *Mitteilungen des Instituts für Orientforschung*, 6(2). 284-320.

- Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: an introduction*, 11-42. Cambridge, MA: Cambridge University Press.
- Winford, Donald. 2000. "Intermediate" creoles and degrees of change in creole formation: the case of bajan. In Ingrid Neumann-Holzschuh & Edgar W. Schneider. (eds.), *Degrees of restructuring in creole languages*. Creole Languages Library (CLL), vol. 22, 215-246. University of Regensburg. Amsterdam/Philadelphia: John Benjamins.
- 2002. Creoles in the context of contact linguistics. In Glenn G. Gilbert (ed.), *Pidgins and creoles in the twenty-first century*, 287-354. New York, NY: Peter Lang.
- 2003b. Contact-induced changes: classification and process. *Ohio State University Working Papers in Linguistics*, 57. 129-150.
- Wolfson, Nessa. 1991. The linguistic variable: fact and fantasy. *American Speech*, 66(1). 22-32.
- 1999. *Lenguas criollas de base lexical española y portuguesa*. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
- Xavier, Maria Francisca & Maria Helena Mira Mateus (orgs.). 1990. *Dicionário de termos linguísticos* - vol. 1. Lisboa: Ed. Cosmos.
- 1992 (orgs.). *Dicionário de termos linguísticos* - vol. 2. Lisboa: Ed. Cosmos.
- Yayun, Ancy Sun. 2008. Input processing in second language acquisition: a discussion of four input processing models [em linha]. Working Papers. *TESOL & Applied Linguistics*, vol. 8(1). Teachers College, Columbia University. [Consult. 06 Nov. 2009]. Disponível em:  
<http://journals.tc-library.org/index.php/tesol/article/viewFile/359/260>.
- Young, Richard. 1989. Ends and means: methods for the study of interlanguage variation. In Susan M. Gass *et alii* (eds.), *Variation in SLA: psycholinguistic issues*, 65-90. Clevedon: Multilingual Matters.
- Zimmermann, Klaus (ed.). 1999. *Lenguas criollas de base lexical española y portuguesa*. Frankfurt-am-Main: Vervuert/Madrid: Iberoamericana.
- Zobl, Helmut. 1980. The formal and development selectivity of L1 influence on L2 acquisition. *Language Learning*, 30. 43-57.

### III. Sítios pertinentes consultados:

Abeokuta Web Pages: Yorùbà from English

– <http://www.abeokuta.org/yoruba.htm> [Consult. 24 Maio 2010].

Angola

– <http://www.cpires.com/africa.html> [Consult. 07 Maio 2008].

Angola do Outro Lado do Tempo

– <http://tudosobreangola.blogspot.com/> [Consult. 02 Maio 2009].

Associação Brasileira de Estudos Crioulos e Similares

– <http://www.unb.br/il/liv/crioul/textos/abecs.html>. [Consult. 06 Jun. 2005].

Associação de Crioulos de Base Lexical Portuguesa e Espanhola

– <http://www.umac.mo/fsh/dp/acblpe/index.html>. [Consult. 03 Jan. 2005].

Associação Portuguesa de Linguística

– [http://www.apl.org.pt/f\\_index.htm](http://www.apl.org.pt/f_index.htm). [Consult. 30 Mar. 2005].

Biblioteca Nacional de Portugal

– <http://www.bnd.pt/memorias/lingua/lingua.html>. [Consult. 04 Abr. 2005].

Centro de Linguística da Universidade de Lisboa

– <http://www.clul.ul.pt/>. [Consult. 24 Fev. 2008].

Ciberdúvidas da Língua Portuguesa

– <http://ciberduvidas.sapo.pt/ficha.htm>. [Consult. 13 Abr. 2005].

Creolica – Revue du Groupe Européen de Recherches en Langues Créoles

– <http://www.creolica.net/>. [Consult. 19 Fev. 2006].

Directoluso – Pensar a Lusofonia: Labirinto da Cultura

– <http://directoluso.blogspot.com>. [Consult. 07 Maio 2008].

– <http://directoluso.blogspot.com/2007/05/histria-da-tv-stp.html>. [Consult. 24 Fev. 2008].

Distributed Morphology: Frequently Asked Questions List

– <http://www.ling.upenn.edu/~rnoyer/dm/>. [Consult. 22 Nov. 2009].

Ethnologue – Languages of the World

– <http://www.ethnologue.com/>. [Consult. 14 Jun. 2006].

Geocities

– <http://www.geocities.com/kimbunduhp/adverbios.htm>. [Consult. 07 Fev. 2005].

Ikuska Libros, S. L.

– [www.ikuska.com/Africa/Paises/santo\\_tome.htm](http://www.ikuska.com/Africa/Paises/santo_tome.htm). [Consult. 23 Maio 2006].

Instituto Camões – Centro Virtual Camões

– 2001. Mapa dos crioulos de base portuguesa [em linha]. *Tempo da língua*. [Consult. 31 Maio 2005]. Disponível em:

<http://www.instituto-camoes.pt/cvc/tempolingua/03.html>.

– 2002/2004. Mapa I: geografia do português e dos crioulos de base portuguesa [em linha]. *História da Língua portuguesa, geografia da língua portuguesa*. [Consult. 31 Maio 2005]. Disponível em:

<http://www.instituto-camoes.pt/cvc/hlp/index.html>.

– 2002/2004. Dialectos portugueses [em linha]. *História da língua portuguesa, geografia da língua portuguesa*. [Consult. 31 Maio 2005]. Disponível em:

<http://www.instituto-camoes.pt/cvc/hlp/geografia/dialectosportugueses.pdf>.

- 2002/2004. Gramática histórica do português [em linha]. *História da língua portuguesa*. [Consult. 12 Jan. 2006]. Disponível em:  
<http://www.instituto-camoes.pt/cvc/hlp/gramhist/index.html>.

#### Jornal de São Tomé e Príncipe

- <http://www.jornal.st/directorio.php?page=173>. [Consult. 12 Dez. 2009].

#### Kimbundo HP

- <http://www.linguakimbundu.com/index3.html>. [Consult. 15 Nov. 2007].

#### Literatura Brasileira - Textos Literários em Meio Eletrônico

- *A carta*, de Pêro Vaz de Caminha [em linha]. Edição de base: *Carta a El Rei D. Manuel*, Dominus, São Paulo, 1963. [Consult. 11 Jun. 2009]. Disponível em:  
<http://www.cce.ufsc.br/~nupill/literatura/carta.html>.

#### Mongabay.com

- 2004/2007. 2005 Population estimates for cities in São Tomé and Príncipe [em linha]. *Population figures*. [Consult. 18 Maio 2006]. Disponível em:  
[www.mongabay.com/igapo/2005\\_world\\_city\\_populations/S%e3o\\_Tom%e9\\_and\\_Pr%edncipe.html](http://www.mongabay.com/igapo/2005_world_city_populations/S%e3o_Tom%e9_and_Pr%edncipe.html).

#### Nação Ovimbundu

- Gramática do umbundu. [Consult. 08 Dez. 2009]. Disponível em:  
<http://www.ovimbundu.org/Educacao/Rudimentos-de-Gramatica-da-Lingua-Umbundu.html>.

#### Pedrabika – Kau di Papia

- <http://pedrabika.blogspot.com/2006/09/papia-di-blog.html>. [Consult. 27 Set. 2006].

#### Projeto NURC-RJ (Projeto da Norma Urbana Oral Culta do Rio de Janeiro)

- <http://www.letras.ufrj.br/nurc-rj/>. [Consult. 24 Fev. 2008].

#### Projeto Vertentes do Português Rural do Estado da Bahia

- <http://www.vertentes.ufba.br/>.
- *Chave de transcrição do projeto vertentes*. [em linha]. [Consult. 30 Maio 2005]. Disponível em:  
<http://www.vertentes.ufba.br/chave.htm>.
- *Conhecer a história sociolinguística da Bahia e do Brasil*. [em linha]. [Consult. 15 Fev. 2009]. Disponível em:  
<http://www.vertentes.ufba.br/justificativa.htm>.
- *O conceito de transmissão linguística irregular* [em linha]. [Consult. 12 Nov. 2008]. Disponível em:  
<http://www.vertentes.ufba.br/conceito.htm>.
- *O paradoxo do observador e a técnica da entrevista sociolinguística*. [em linha]. [Consult. 09 Nov. 2008]. Disponível em:  
<http://www.vertentes.ufba.br/paradoxo.htm>.

#### Quimera Editores

- <http://www.quimera-editores.com/catalogo/teatro/vicente.html>. [Consult. 08 Jan. 2006].
- <http://www.quimera-editores.com/vicente/pdf/Clerigo.pdf>. [Consult. 08 Jan. 2006].
- <http://www.quimera-editores.com/vicente/pdf/Fragua.pdf>. [Consult. 08 Jan. 2006].
- <http://www.quimera-editores.com/vicente/pdf/Nau.pdf>. [Consult. 08 Jan. 2006].

#### República Democrática de São Tomé e Príncipe – Instituto Nacional de Estatística:

- <http://www.ine.st/>. [Consult. 06 Jun. 2006].

Surrey Morphology Group Agreement Web Site

- <http://www.surrey.ac.uk/LIS/SMG/projects/agreement/agreement.html>. [Consult. 02 Mar. 2010].

The Ohio State University – Department of Linguistics, Études Créoles

- <http://www.ling.ohio-state.edu/research/jpcl>. [Consult. 02 Jan. 2006].

Tudo de Bom! – Gramática Yoruba

- <http://vidademacumbeiro.blogspot.com/2008/12/gramtica-yoruba.html> [Consult. 24 Maio 2010].

UNESCO – The Courier

- [s.d.]. Winners and losers [em linha]. *Languages: conflict or coexistence?*. [Consult. 01 Jun. 2006]. Disponível em:  
[http://www.unesco.org/courier/2000\\_04/uk/doss03.htm#top](http://www.unesco.org/courier/2000_04/uk/doss03.htm#top).

Wikipedia

- 2007. *Portuguese creole*. [Consult. 02 Abr. 2006]. Disponível em:  
[http://pt.wikipedia.org/wiki/Portuguese\\_creole](http://pt.wikipedia.org/wiki/Portuguese_creole).
- 2007. *São Tomé e Príncipe*. [Consult. 02 Abr. 2006]. Disponível em:  
[http://pt.wikipedia.org/wiki/S%C3%A3o\\_Tom%C3%A9\\_e\\_Pr%C3%ADncipe](http://pt.wikipedia.org/wiki/S%C3%A3o_Tom%C3%A9_e_Pr%C3%ADncipe).
- 2007. *Grammar*. [Consult. 02 Abr. 2006]. Disponível em:  
<http://en.wikipedia.org/wiki/Grammar>.



## *Curriculum vitae do autor*

### **Carlos Filipe Guimarães Figueiredo**

#### **Formação académica:**

*Universidade de Lisboa, Faculdade de Letras*

- Licenciatura em Língua e Cultura Portuguesa (Língua Estrangeira) – 2003.

#### **Categoria profissional:**

*Universidade de Macau, Faculdade de Ciências Sociais e Humanas, Departamento de Português*

- Assistente eventual (Senior instructor).

#### **Cargos adicionais:**

- Coordenador dos Cursos de Português Língua Estrangeira – 2009 à presente data.
- Coordenador do Segundo Ano da Licenciatura em Estudos Portugueses para Estrangeiros – 2006 à presente data.
- Representante eleito dos Instrutores Leitores para o Conselho Académico da Universidade de Macau – 2009 à presente data.

#### **Experiência profissional na área do ensino:**

*Universidade de Macau* – 2003 à presente data.

- Português Língua Estrangeira.
- Escrita (Português Língua Estrangeira).
- Escrita Prática com Propósito Profissional (Português Língua Estrangeira).
- Cultura Portuguesa (Português Língua Estrangeira).
- Fonética (Português Língua Estrangeira).
- Gramática Portuguesa (Português Língua Estrangeira).
- Conversação (Português Língua Estrangeira).
- Leitura (Português Língua Estrangeira).
- Prática em Laboratório de Língua – Iniciação (Português Língua Estrangeira).
- Laboratório de Língua: Desenvolvimento da Produção e Competência Orais (Português Língua Estrangeira).
- Laboratório de Língua: Desenvolvimento da Compreensão Auditiva (Português Língua Estrangeira).
- Português com Objectivos Jurídicos (Português Língua Estrangeira).
- Língua e Cultura Portuguesas para Professores Chineses de Chinês Língua Estrangeira em Países de Língua Portuguesa (Português Língua Estrangeira).

*Escola Superior das Forças de Segurança de Macau* – 2009 à presente data.

- Português Língua Estrangeira para Oficiais do Ramo Policial.

*IPOR – Instituto Português do Oriente* – 2003-2004.

- Português Língua Estrangeira.



*Universidade de Lisboa – Departamento de Língua e Cultura Portuguesa – 2003.*

- Português Língua Estrangeira.

**Outras actividades académicas:**

- *Comité Selectivo dos Estudantes da Licenciatura em Estudos Portugueses para Estrangeiros (1º Ano) do Departamento de Português – Universidade de Macau.*
  - Membro selectivo.
- *Associação dos Estudantes da Licenciatura em Estudos Portugueses para Estrangeiros da Universidade de Macau.*
  - Fundador e supervisor coordenador.

**Associação:**

- *ACBLPE – Associação dos Crioulos de Base Lexical Portuguesa e Espanhola.*
  - Secretário da Mesa da Assembleia-geral.

**Áreas de investigação:**

- Sociolinguística;
- Reestruturação do português em situações de contacto;
- Aquisição de primeira língua;
- Aquisição de segunda língua;
- Aquisição do português por comunidades crioulofonas.

**Projectos de pesquisa:**

- *Dissertação de doutoramento:*
  - A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almojarife (São Tomé).
    - Investigador principal.
- *CIELA - Centro de Investigação de Estudos Luso-Asiáticos da Universidade de Macau:*
  - Português reestruturado por aquisição de língua em situação de contacto: aquisição geracional, tipologia linguística e análise gramatical.
    - Investigador principal.
  - Línguas crioulas e semi-crioulas de base portuguesa: tipologia, análise gramatical e risco de extinção.
    - Co-investigador.

**Publicações:**

*Artigos:*

- 2009. A configuração do SN do português reestruturado da comunidade de Almojarife, São Tomé. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola*, 1(1). 28-55.
- 2008. A concordância variável no sintagma nominal plural do português reestruturado de Almojarife (São Tomé). *PAPIA* 18. 23-43.

- 2003. O português em Angola: algumas ocorrências em contexto literário. In Departamento de Língua e Cultura Portuguesa – Faculdade de Letras da Universidade de Lisboa (ed.), *1<sup>as</sup> Jornadas de Língua e Cultura Portuguesa (18/19 de Junho de 2002)*, 113-128. Lisboa: Departamento de Língua e Cultura Portuguesa – Faculdade de Letras da Universidade de Lisboa.

### **Comunicações:**

- 2010. *Paralelismos morfossintáticos em variedades bantu, crioulos atlânticos de base portuguesa e variedades africanas e brasileiras de português: transferência ou trajecto universal de aquisição?* 10<sup>ème</sup> Colloque International de l'Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Llacan, Campus CNRS de Villejuif, Paris, France, 1-3 de Julho.
- 2009. *Concordância plural variável no sintagma nominal do português reestruturado de Almojarife: motivações morfofonológicas.* Joint Summer Meeting of the Society of Pidgin and Creole Linguistics (SPCL) and the Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade de Colónia, Alemanha, 11-15 de Agosto.
- 2008. *Encaixamento linguístico da configuração do sintagma nominal plural do português reestruturado da comunidade de Almojarife (São Tomé).* Congresso Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade Nova de Lisboa, Portugal, 03-05 de Julho.
- 2006. *O sintagma nominal no português reestruturado de Almojarife (São Tomé).* Congresso Anual da Associação de Crioulos de Base Lexical Portuguesa e Espanhola (ACBPLE). Universidade de Coimbra, Portugal, 26-28 de Junho.
- 2003. *A arte da Reconquista: subsídios para um percurso do multicultural ao intercultural no ensino/aprendizagem do português (LE).* 2<sup>as</sup> Jornadas de Língua e Cultura Portuguesa. Universidade de Lisboa, Portugal, 30-31 de Maio.
- 2002. *O português em Angola: algumas ocorrências em contexto literário.* 1<sup>as</sup> Jornadas de Língua e Cultura Portuguesa. Universidade de Lisboa, Portugal, 18-19 de Junho.
- 2001. *Mais línguas, mais Europa: celebrar a diversidade linguística e cultural da Europa.* Colóquio Ano Europeu das Línguas 2001. Universidade de Lisboa, 25-26 de Janeiro.

### **Revisões de texto (livros):**

- 2010. *Ditama – Dicionário Temático de Macau*, vol. 1. Macau: Ábaco-Cushing.

### **Menções honrosas:**

- Bolsa de Mérito Excepcional (10 Melhores Estudantes de Portugal), 1999-2000, 2000-2001, 2001-2002, 2002-2003.
- CAIPS (Menção Honrosa concedida pela Embaixada do Canadá em Portugal), 1994.

### **Certificados diversos:**

*Instituto Camões, Lisboa*

- Curso de Leitores de Língua e Cultura Portuguesas para a Ásia (Português Língua Estrangeira) – 2003.

*CIVEC - Centro de Formação Profissional, Lisboa*  
- Desenvolvimento das Aptidões de Chefia – 1999.

*IFILP – Instituto de Formação e Investigação da Língua Portuguesa*  
- Língua Portuguesa com Objectivos de Pesquisa Científica – 1999.

*EGOR – Recursos Humanos, SA, Lisboa*  
- Desenvolvimento da Gestão de Recursos Humanos e Relações Inter-profissionais –  
1994.

*Instituto Americano de Lisboa, Lisboa*  
- Inglês Língua Estrangeira – 1985.

*IATA – Instituto de Aperfeiçoamento Técnico Acelerado, Lisboa*  
- Desenho de Infra-estruturas para Construção Civil – 1981.