

Bantu substratum interference in Mozambican Portuguese speech varieties

Simone Ashby and Sílvia Barbosa

Abstract

Bantu phonological constraints exert varying types of interference in the varieties of Portuguese spoken in Mozambique. While many of these processes – e.g. voicing dissimilation, consonant sequence simplification, and denasalization – are relatively well known among linguists specialized in the study of Mozambican Portuguese dialects, they have received only passing mention in the literature. Moreover, it is unclear from the available data how Bantu substratum influences characterize the Portuguese of L1 speakers and those who acquired the language at a very young age. The present work is based on the analysis of phonetic data from five native Portuguese speaking informants (three L1 speakers and two L2 speakers) from various locations in Mozambique, and whose language backgrounds represent Changana, Chope, GiTonga, Makhuwa, and Nyungwe. Evidence is presented showing the effects of local Bantu phonologies on informants's spoken Portuguese varieties for the aforementioned phenomena, plus additional processes that correspond with the literature on Bantu linguistics, such as devoiced aspirated word-final consonants, 'whistled' fricatives, and homorganic nasal epenthesis. We conclude that investigations of substratum phonetic and phonological interference are important for the long-term monitoring of a contact situation, and help to elucidate the principles governing how segments are realized in indigenizing languages.

Keywords: Bantu substratum interference, Mozambique, Portuguese speech varieties, contact language effects, indigenizing languages

1. Introduction

This article looks at the processes that account for differences between spoken varieties of Mozambican Portuguese (MP) and the European Portuguese standard (EP), and which involve the interference of local Bantu phonological features, constraints and phenomena. The spoken Portuguese of five informants from different urban centers in Mozambique was examined using auditory perception and instrumental analysis. Informants represent a variety of linguistic backgrounds and bilingual or multilingual capabilities. Three of the five informants reported Portuguese as their first language, despite being raised by parents with a Bantu language as their mother tongue, while the remaining two informants reported learning Portuguese as young children. In terms of substrate languages, informants reported a native or near-native understanding of one or more of the following Bantu languages: Chope and GiTonga, of the Inhambane Group; Makhuwa, of the Makhuwa Group; Nyungwe, of the Senga-Sena Group; and Changana and Tshwa, of the Tswa-Ronga Group. Fluency in English was further reported by two of the informants.

From the informant population of this study alone, one begins to apprehend the degree of multilingualism that characterizes Mozambique's diverse language topography. Indeed, Mozambique is a country where multiple languages are in contact with one another, and where social context and factors such as age, gender, education and occupation govern the manner and extent to which they are used (Firmino 1995, 2000, 2002; Gonçalves 1996, 2010; Stroud & Gonçalves 1997; Stroud 2007). Like other former Portuguese colonies in Africa – such as Angola and Guinea-Bissau – it is also a country where indigenized varieties of the vehicular Portuguese language are emerging among growing L1 and L2 speaker populations. By examining the effects of local Bantu substrate languages on a selection of MP *topolectal* varieties by both L1 and L2 speakers, this study aims to contribute to the understanding of phonetic variation in MP, and to document current findings for the long-term monitoring of language contact effects in Mozambique. As such, this study is concerned with describing the linguistic outcomes of a prolonged contact between superstrate and substrate languages as viewed from a sociohistorical perspective. It is hoped that the present study may, like others before it (e.g. Duarte *et al.* 1999; Firmino 2002; Gonçalves 1990; Gonçalves 2010; Machungo 2000), serve in raising the profile of MP, and indigenizing language varieties in general, as important for the advancement of linguistic knowledge and as varieties worthy of study in their own right.

The processes submitted for analysis in this article were observed in varying combinations for the different informants studied. Many have a foundation in the literature on Bantu phonetics and phonology, including the monothongization of diphthongs, vowel and consonant sequence simplification, glide epenthesis, syncope, apocope, apheresis, voicing dissimilation, the realization of word-final aspirated devoiced consonants, 'whistled' fricatives, palatalization, continuantization, homorganic nasal epenthesis, and denasalization. Collectively, these processes may be interpreted, along with existing lexical, morphological, and syntactic studies, as helping to define the current language situation in Mozambique. Processes such as homorganic nasal epenthesis – which has been the subject of considerable attention

in the literature on Bantu linguistics, but which has received no notice in studies of MP – also highlight the interplay between transfer from the substrate, superstratal influence, and universal preference laws in governing how segments are realized in the context of Portuguese nasal vowels.

2. Sociohistorical background

Mozambique extends along the Indian Ocean, from its northern border with Tanzania to the country's southwest reaches, bordering Swaziland and South Africa. The interior is made up of horizontally striated river valley settlements that extend from the much larger urban areas that dot the coast. At the time of writing, the population of Mozambique was estimated at over 22 million, with 37% of the population residing in cities (CIA Factbook). The capital city of Maputo is located in the country's southernmost tip, an area that is integrally connected with South Africa in terms of a shared economic structure and communications network (Newitt 2002:186).

In the dawn of Mozambique's independence from Portugal, the Education Minister at the time, Graça Machel, defended the government's selection of Portuguese as the official language of Mozambique in an address to the 1st National Seminar on the 'Teaching of Portuguese' in 1979: "The need to fight the oppressor called for an intransigent struggle against tribalism and regionalism. It was this necessity for unity that dictated to us that the only common language – the language which had been used to oppress – should assume a new dimension" (Lopes 1999:104).

Since that time, Mozambique has undergone a steady language shift by decree. To this day, Portuguese remains the official language of Mozambique, where it is spoken as a lingua franca by 50.37% of the population, an additional 10.7% of which regard Portuguese as their native language. Portuguese is used in all official administrative, governmental, and judicial communications. It is the language of instruction in Mozambican schools and the Eduardo Mondlane universities, and it is used by the vast majority of Mozambican media outlets. While the Bantu stratum constitutes the majority languages of Mozambique and, according to the most recent census in 2007, the native tongues of 90% of the population, Portuguese is viewed as "the language of science, knowledge, and power", and that which holds the most promise for obtaining employment and enhancing one's upward social mobility (Da Conceição 1999:15).

The disproportionately elevated status of Portuguese has its roots in the colonial period, when a policy of assimilation and controlled access to the Portuguese language offered opportunities for a more respected position in society for oneself and one's children (Gonçalves 2010:33; Stroud 2007:509). Access to Portuguese was also sanctioned by the Catholic church, which, to a limited extent, led to an early instantiation of indigenized popular varieties (Stroud 2007:509-510). Then, as now, Portuguese was largely restricted to the major urban areas and the capital Maputo in particular, while rural Mozambicans had very little or no contact with the Portuguese language, and relied on local Bantu languages for communication outside the family.

According to the 2007 census, an estimated 24 Bantu languages are spoken in Mozambique. The dominant languages and corresponding percentages of native speakers are: Makhuwa (26.3%), Changana (11.4%), Lomwe (7.9%), Sena (7%), and Chuabo (6.3%). These and the less dominant languages of Mozambique constitute four zones and eight major language groupings, as delineated by Guthrie (1967-1971): Zone G-G40–Swahili; Zone P-P20–Yao, along with P23–Makonde, and P30–Makhuwa (plus P32–Lomwe, P34–Chwabo); Zone N-N30–Nyanja, along with N40–Senga-Sena; and Zone S-S10–Shona, along with S50–Tswa-Ronga, and S60–Copi (Lopes 1999:87; Maho 2003).

Since independence, a number of ideological arguments have been raised in favor of adopting an active stance on the promotion and development of African languages in Mozambique. Official overtures in this direction have, however, been few and far between, despite regular broadcasts in dominant Bantu languages by Radio Mozambique. More recent factors have intervened to ameliorate the situation, such as the involvement of churches and NGOs in promoting the use and instruction of Bantu languages, and limited television programming in languages such as Changana (Stroud 2007:516-517). Stroud (2007:520) also cites the resurgence of informal trade activities in recent years as having reinforced the use of African languages within larger regions and across borders.

Nevertheless, Mozambique remains caught in a pre-independence cycle, whereby Portuguese is evaluated as suiting more formal types of communication between educated urbanites, while local Bantu languages generally connote a less sophisticated, more rural medium of communication. In urban centers, some bilingual parents are opting to raise their children solely in Portuguese, perhaps with the hope of propelling them towards a brighter future, and thereby ushering in a new generation of L1 speakers. With a greater preponderance of young people speaking Portuguese, and males in particular, who also tend to be fluent in a greater number of languages, “urbanity and mobility translate into a gendered and age differentiated access to linguistic resources, introducing social stratifications around multilingualism [not found] in the rural areas” (Stroud 2007:521).

In terms of MP language models, lawmakers and educators determined soon after independence that the teaching of Portuguese in schools should aim towards EP. However, in subsequent years, “practice showed that such an idealistic goal was not achievable, and even no longer desired because it lacked the marks of an emerging national identity” (Lopes 1999:123). Since then, Mozambique has exercised what Lopes (1999:123) describes as a “*laissez-faire* policy” concerning the normativization and standardization of Portuguese. Authors such as Da Conceição (1999), Gonçalves (2010; 1996), Lopes (1979; 1999), and Stroud (2007) observe a general cognizance among Mozambicans that there is a ‘correct’ form of spoken and written Portuguese. Many Mozambicans continue to look to EP as the standard model and target variety, while others, like the community of Polana Cimento in Maputo, demonstrate a high regard for their own variety of Portuguese (Stroud 2007:522). External influence comes not only by way of Lisbon, but also via the metropolises of Rio de Janeiro and São Paulo in the form of soap operas and other Brazilian programmes that have become popular among Mozambicans with access to a television or the internet.

Meanwhile, Portuguese in Mozambique has increasingly come to be regarded as a language under threat due to the strengthening of economic ties with South Africa and Mozambique's other Anglophone neighbors, its recent entry into the British Commonwealth, and economic and linguistic intervention from France (Da Conceição 1999:22). Given that Mozambique is one of the poorest countries in the world, and yet one that is fast adapting to the changing dynamics of the global political economy, the increasing role of languages such as English threaten to unseat Portuguese as the most viable and widely taught language in Mozambique.

3. Diglossia and substratum interference

So far, we have tried to establish the major historical, political, economic, and social circumstances that explain the current dominant position enjoyed by Portuguese in Mozambique, along with the curtailed role of a multiplicity of indigenous Bantu languages. The diglossic juxtaposition that defines this nation linguistically entails a further set of inter-system dynamics, which have implications for the lexicons, morphologies, grammars, and phonologies of the different languages in contact. In the nomenclature on diglossia, traditionally limited to the description of dialects of the same language, the superimposed 'prestige' variety is referred to as 'High' (H), and the 'non-prestige' varieties, whose use is restricted to informal exchanges, and which are excluded from the medium of instruction in schools, are labeled as 'Low' (L) (Ferguson 1959). Fishman (1967) later modified these constructs to include multilingual situations. In both types of contact situations, varieties or languages marked as H have been observed to exert more influence on those identified as L than the reverse situation. This dynamic, referred to by Thomason & Kaufman (1988) as borrowing, is documented in a variety of descriptions of Bantu languages, such as Changana/Tsonga (Gonçalves & Siteo 1999; Siteo 1991) and Nyungwe (Rego 2000). Conversely, the substratum interference exerted by local Bantu languages on MP has been the subject of a relatively large number of variationist and SLA studies devoted primarily to describing the lexical, morphological, syntactic and semantic features of MP (e.g. Gonçalves 1994, 2002a, 2002b; Gonçalves & Chimbutane 2004; Gonçalves & Siteo 1999; Lopes 1979).

Substratum interference in MP at the phonological and phonetic levels, on the other hand, remains relatively undocumented, with the exception of some brief remarks concerning oral production 'errors' and the more obvious features of MP spoken varieties. This seems a rather glaring omission in light of the fact that language contact situations, such as that in Mozambique, are known to have demonstrable phonological effects on the recipient language, in addition to the more well-studied dynamic of large-scale lexical borrowing that is common among substrate languages (Sankoff 2002:643). The lack of sufficient phonological and phonetic documentation also leaves bare the potential for more robust diachronic and synchronic accounts of the development, indigenization, variation, and change of MP in years to come.

4. Phonological descriptions of EP and the Bantu languages

4.1. EP

EP refers to the varieties of Portuguese spoken in Coimbra and Lisbon. In Portugal, these are the varieties heard most often in radio and television broadcasts. They also represent Portugal’s ‘prestige’ varieties and those deemed appropriate for transmission of the language in L2 settings (Mateus & d’Andrade 2000:4). The motivation for including such a description in this article lies in having a baseline for comparison in the identification of Bantu substratum interference in MP speech varieties. Thus, the processes described in Section 5 will, in many cases, refer indirectly to the present Section concerning the description of basic phonological and phonetic features of EP, as well as the subsequent Section, which aims to describe some of the major features of the Mozambican Bantu languages with which this study is concerned.

4.1.1. Consonants

Table 1 presents the consonantal inventory of EP, based on Mateus & d’Andrade (2000).

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular
Nasal	m			n		ɲ		
Plosive	p b		t d				k g	
Fricative		f v		s z	ʃ ʒ			
Lateral				l		ʎ	ʟ	
Flap				r				
Trill								ʀ¹

Table 1. EP consonants based on the description in Mateus & d’Andrade (2000)

EP consonants are distributed in the three basic positions: initial, medial, and final.

- (1) a. Word initial: all consonants except [ʎ, ɲ², r, ʟ]
- b. Word-medial (between vowels): all consonants, along with corresponding fricative allophones [β, ð, γ]
- c. Word-final: [ʟ, r, ʃ]

The realization of the voiceless alveolar fricative /s/ is realized as [ʒ] when followed by a voiced consonant, e.g. **le [ʒ] ma lesma** ‘slug’. When followed by a voiceless consonant, /s/ is realized as [ʃ], e.g. **ce [ʃ] ta cesta** ‘basket’. Alternatively, between vowels in cross-word contexts, /s/ becomes the voiced fricative [z], e.g. **casa [z] amarelas casas amarelas** ‘yellow houses’.

1. Mateus & d’Andrade (2000:11) submit that the uvular fricative [ʀ] shows signs of becoming dominant in the Lisbon area.
 2. The palatal lateral [ʎ] and palatal nasal [ɲ] occur in a very small number of word-initial contexts, which according to Mateus & d’Andrade (2000:11) consist of the pronoun ‘Ihe’ in the case of [ʎ], and an extremely limited number of rarely used loanwords in the case of both sounds.

The rhotics [r] and [ʀ] appear intervocally in word-medial position, e.g. **ca** [r] **o** *caro* ‘expensive’ and **ca** [ʀ] **o** *carro* ‘car’. [r] also occurs syllable- and word-finally, e.g. **po** [r] **co** *porco* ‘pig’ and **ma** [r] *mar* ‘sea’, as well as in consonant clusters, e.g. **p** [r] **ato** *prato* ‘plate’, but it is restricted from occurring word-initially or following a nasal consonant. The uvular trill [ʀ], on the other hand, never occurs in word-final position, and is restricted to word-initial position, e.g. [ʀ] **oma** *roma* ‘Rome’, and in positions preceded by a syllable-final consonant, e.g. **pal** [ʀ] **ar** *palrar* ‘gabble’, or nasal vowel, e.g. **hon** [ʀ] **ar** *honrar* ‘honor’.

The laterals [l] and [ɫ] are in complementary distribution, with [l] occurring word-initially, e.g. [l] **ata** *lata* ‘tin’ and in clusters, e.g. **p** [l] **anta** *planta* ‘planta’, and [ɫ] occurring syllable- and word-finally, e.g. **sa** [ɫ] **tar** *saltar* ‘jump’ and **ma** [ɫ] *mal* ‘evil’.

Lexical exceptions to the above rules must be memorized. The more common example concerns words with silent consonants that are still preserved in the orthography, e.g. [‘**batifmu**] *baptismo* ‘baptism’.

4.1.2. Vowels and glides

EP oral and nasal vowels are presented in Figure 1, based on the description of vowels in Mateus & d’Andrade (2000).

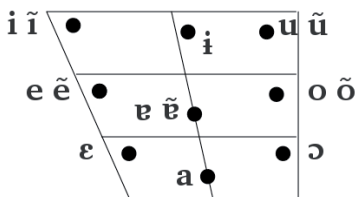


Figure 1. EP vowels based on the description in Mateus & d’Andrade (2000)

The distribution of stressed and unstressed vowels is described in (2).

- (2) a. Stressed oral vowels: [i, e, ε, v, a, ɔ, o, u]
- b. Unstressed pre-tonic oral vowels: [i, i, v, u]
- c. Unstressed post-tonic, non-final oral vowels: [i, i, v, u]
- d. Unstressed final vowels: [i, v, u]
- e. Stressed nasal vowels: [ĩ, ē, ẽ, õ, ũ]
- f. Pre-tonic nasal vowels: [ĩ, ẽ, ẽ, õ, ũ]

Each of the oral vowels can appear in a stressed position, with the exception of [i], which is realized as an allophone of /e/ and /ε/ between consonants. The phonetic realization of stressed [v] occurs as an allophone of the stressed front vowel /e/ before a palatal consonant or glide, e.g. **t** [vɐ] **a** *telha* ‘tile’, **l** [vɐ] **te** *leite* ‘milk’. /a/ is also realized as [v] when preceded by a nasal consonant, e.g. **c** [v] **ma** *cama* ‘bed’, **c** [v] **na** *cana* ‘reed’, and **c** [v] **nhão** *canhão* ‘cannon’. The correspondences between oral vocalic phonemes and their unstressed phonetic counterparts are presented in (3).

- (3) a. /a/ realized as [ɐ].
 b. /e, ɛ/ are realized as [i].
 c. /o, ɔ/ are realized as [u].
 d. /i, u/ have no allophones, but in connected speech are elided in word-final position followed by a consonant.
 e. The stressed vowels [i, e, ɛ, ɐ, a, ɔ, o, u] correspond with the more limited unstressed set [i, i, ɐ, u], involving centralization and raising.

Regular exceptions to the above rules are presented in (4).

- (4) a. Unstressed /a/ is preserved when followed by a velar lateral [ɫ] coda, e.g. **s**[aɫ] **gado** *salgado* ‘salty’ (*s[ɐɫ] **gado**).
 b. Unstressed /a/ is preserved when it occurs as a constituent in a falling diphthong, e.g. **g**[aj] **teiro** *gaitero* ‘piper’ (*g[ɛj] **teiro**).
 c. Unstressed /a/ is preserved when followed by evaluative suffixes initiated with [z], e.g. **devag**[a] **rzinho** *devagarzinho* ‘slowly’ (***devag**[ɐ] **rzinho**).
 d. Unstressed /o, ɔ/ are preserved word-initially, e.g. [o] **perário** *operário* ‘laborer’ (*[u] **perário**), and [ɔ] **rlando** ‘Orlando’ (*[u] **rlando**).
 e. Unstressed /e/ may be realized as one the free variants [e, i] word-initially, e.g. [e] **rnesto** and [i] **rnesto** *Ernesto* ‘Ernest’.

The glides [j, w] combine with vowels to form diphthongs. EP has 16 diphthongs, 11 of which are oral and five of which are nasal, i.e. [aj, ɛj, ej, ɔj, oj, uj, iw, ew, ɛw, aw, ɛ̃j, ẽj, õj, új, ẽw]. According to Mateus and d’Andrade (2000:50), these sounds are interpreted as underlying high vowels, and rising diphthongs alternate with hiatus realizations. In both careful and connected speech, glide formation occurs in the three basic word positions, e.g. **r**[ja] **lmente** *realmente* ‘really’, **san**[ja] **mento** *sanemento* ‘sanitation’, and **pass**[ja] **r** *passear* ‘walk’. The glide [w] also appears after /k, g/ in place of the high back vowel /u/ in words such as [kw] **idar** *cuidar* ‘care’ and **i**[gw] **al** *igual* ‘equal’.

4.1.3. Syllable constraints

In terms of syllable configurations, EP onsets can be word-initial or word-medial, and consist of a single consonant or consonant cluster of the combination {plosive, fricative} + {flap, lateral}, e.g. [pr] **aia** *igual* ‘beach’, **ca**[br] **a** *cabra* ‘goat’, [gl] **obo** *globo* ‘globe’, **de**[kl] **arar** *declarar* ‘declare’, [fr] **aco** *fraco* ‘weak’, **re**[fr] **ear** *refrear* ‘rein in’, [fl] **ores** *flores* ‘flowers’, and **con**[fl] **uente** *confluente* ‘tributary’. At the phonetic level, the elision of unstressed vowels is extremely common in EP, thereby resulting in any number of consonant sequences, such as [tʃfɔn] *telefone* ‘telephone’. In this last case, /l/ becomes syllabic, leaving an empty nucleus in the first syllable. Also, the final vowel is dropped, with the resyllabification of /n/ as a coda following the stressed vowel.

Syllabic rhymes can consist of a single vowel, e.g. t[u] *tu* ‘you’, or a diphthong, e.g. b[oj] *boi* ‘ox’, while codas are restricted to the phonemes /l, r, s/ and their respective phonetic realizations, e.g. ca[ɫ] **do** *caldo* ‘broth’, **pe**[r] **der** *perder* ‘lose’,

ca [ʃ] **ca** *casca* ‘bark’ or ‘peel’, and **ra** [ʒ] **gar** *rasgar* ‘rip’ or ‘tear’. Penultimate syllables tend to be the primary stress bearing units, unless where noted in the orthography through the use of diacritics, and with some additional exceptions.

4.2. Bantu languages

The following is an attempt to describe some of the more relevant features and processes that occur among Bantu languages, and is largely based on the phonological descriptions provided in Maddieson (2003) and Hyman (2003). Where possible, special emphasis will be given to the languages listed in Section 1 (Changana S53, Chope S61, GiTonga S62, Makhuwa P31, Nyungwe N43, and Tshwa S51), and which make up the substratum of the informants analyzed for this study. Here, the focus is skewed towards consonantal features and processes, due to the predominately non-vocalic nature of the substratum interference described in the following Section of this article.

4.2.1. Consonants

The consonantal inventory of Proto-Bantu (PB) is believed to have been relatively small, including /**p, b, t, d, k, g, m, n, ɲ, c, j**/. Given the fact that [**β, l, ɣ**] are realized in a number of modern Bantu languages, there is some speculation that the voiced phonemes /**b, d, g**/ were also realized in PB as [+continuant]. There is also some speculation concerning the specification of /**c, j**/, as /**c**/ is realized as [**s**] and /**j**/ can be realized as [**z, j, ɣ**] in many modern Bantu languages.

Modern Bantu languages tend to have consonantal systems that are considerably more complex. Changana, for example, which has close to two million speakers and is found in southern Mozambique, is reported to have an exceptionally large consonantal inventory, containing around 100 contrastive elements (Janson 1999:1). In addition to the sounds described above for PB, Changana includes: a set of stops with lateral releases, i.e. /**tl, dl**/; the corresponding voiceless aspirated stops /**ph, th, tlh, kh**/; a labialized set /**tw, tlw, kw, thw, dw, dlw, gw**/; and the prenasalized stops /**mp, nt, ntl, ŋk, mph, nth, ntlh, ŋkh, mb, nd, ndl, ŋg, ntw, ntlw, ŋkw, nthw, ntlhw, ŋkhw, ndw, ndlw, ŋgw**/. Janson & Engstrand (2001) further describe the affricates [**pʃ**] and [**ndʒh, nbvh**³], and Janson (1991) describes a set of prenasalized voiced aspirated stops [**mbh, ndh, ndlh, ŋgh, ŋgwh**], for which it is unclear whether this last set of sounds is phonologically contrastive.

Conversely, voiceless stops have taken the place of voiced prenasalized stops for a number of the dialects of Makhuwa, which has approximately eight million speakers and is found in the Mozambican province of Nampula. This language additionally features contrasting unaspirated and aspirated voiceless stops, i.e. /**p, ph, t, th, ʈ, ʈh, k, kh**/, instead of contrasting voiced and voiceless stops. Table 2 presents the consonantal inventory for the Ruvuma variety of Makhuwa, taken from Kisseberth (2003).

3. The specification of ‘[**nbvh**]’ in Janson & Engstrand (2001) should likely have been the homorganic [**mbvh**].

	Labial	Dental	Alveolar	Alveo- palatal	Velar ⁴
Unaspirated plosive	p	t	ʈ	c	k
Aspirated plosive	ph	th	ʈh		kh
Prenasalized	mb		nd	nj	ŋg
Voiceless fricative			s		h
Voiced fricative	v				
Nasal	m		n	ny	ŋ
Glide				y	w

Table 2. Ruvuma Makhuwa consonants (adapted from Kisseberth 2003:549)

Kisseberth (2003:550) describes a correspondence of the affricate /c/ with /s/, /z/, and /ð/ in other dialects of Makhuwa. He also cites the realization of /s/ as [ʃ] in some dialects, and gliding in cases of dissimilar vowel sequences. In terms of morphophonemic constraints, stems can have no more than one aspirated stop, and multiple voiceless coronal stops are assimilated in stems as all dental or all alveolar (Kisseberth 2003:550). The author further describes the morphophonemic phenomenon of nasal assimilation, whereby the vowel is elided in the prefix *ni-* before a coronal consonant, and a homorganic nasal is realized in the place of /n/.

Homorganic nasal consonants occur in virtually all of the Bantu languages, whether in the form of prenasalized consonants or, in some languages, across morphemes (Hyman 2003:49). According to Hyman (2003:50), “many have restrictions either on which N+C combinations are possible, or in where within the word structure NC may occur.” Additional phenomena observed in the context of the nasal + voiceless stop combination include aspiration and concomitant elision of either the N (e.g. in Swahili G40) or C (e.g. in the Southern Bantu language Shona S10). Lenition in N+C combinations can also occur in the combination nasal + fricative, whereby the nasal is deleted (e.g. in Yao P21), as well as causing continuantization of the C (e.g. in Ganda E15) (Hyman 2003: 50-51). A process of nasal simplification, like that which occurs under Meinhof’s Law, has also been attested, in which a sequence of noncontiguous N+C combinations results in the elision of the subsequent N (Hyman 2003:56).

Other processes involve the aspiration of consonants when followed by a high vowel, as attested in Hyman (2003:53) for languages such as Makhuwa and Doko C301, while in other languages, the consonant becomes an affricate or fricative (e.g. in Ngom B22b). For a relatively large number of languages, the high vowel context is also known to condition allophones of /d/, which is preserved as [d] before [i], but which is realized as [ɪ] or [r] before other vowels (Hyman 2003:54). Hyman (2003:56) further discusses the dissimilatory process that occurs under Dahl’s Law, in which a voiceless stop is realized as its voiced counterpart if followed in the next syllable by another voiceless consonant.

4. As noted in Kisseberth (2003:549), “‘h’ and ‘w’ both require a preceding moraic nasal to assimilate to a velar”.

Bantu languages are typically restricted to having just one of the following liquids /l, r, r/ (Maddieson 2003:25). As in EP, post-consonant glides are considered to be underlyingly vocalic. These sounds also condition some of the same processes described above for high vowel contexts derived from the tense vowels /i, u/. In other cases, [y] results in palatalization, whereas [w] may undergo velar fortition (Hyman 2003:55; Maddieson 2003:25). Certain Bantu languages are also well-known for a set of laryngeal productions that result in implosives and ejective stops and affricates, as well as a phenomenon known as depressor consonants, characteristic of the eastern and southern regions. And, of course, a wide variety of clicks characterize the Southern Bantu languages.

Additional articulatory phenomena of note include whistled fricatives, described in Shosted (2006) for Tshwa, and elsewhere for the Tshwa-Ronga S50 and Shona S10 groups (Bladon et al. 1987; Siteo 1996; Ladefoged and Maddieson 1996; Maddieson 2003). Tshwa, which has approximately 700,000 speakers and is found in southern Mozambique, features alveolar edge-tone whistled fricatives that are allophonic in nature, and which reveal a distinctive spectral envelope (Shosted 2006).

4.2.2. Vowels

Proto-Bantu vowels include /i, i, e, u, u, o, a/. Most modern Bantu languages, such as Nyungwe, which has approximately 400,000 speakers and is found in Central Mozambique, have five-vowel (5V) inventories containing the phonemes /i, e, u, o, a/. Others have expanded their inventories, such as the southern Sotho-Tswana Group S30, which features contrasts between /e, o/ and their tensed counterparts /e, o/. Unlike EP, contrastive nasalized vowels are largely unattested, with the exception of some languages in western Mozambique, which are surmised by Maddieson (2003:23) to be the result of nasal consonant harmony extending across syllables.

Like EP, many of the 5V Bantu languages constrain the distribution of vowels within the word and stem. A set of rigid constraints also governs which vowels can occur together, thus resulting in a variety of reduction and assimilation processes. It has been shown for Punu B43 that vowels in stem-initial position are resistant to such processes, while post-stem vowels are more mutable (Hyman 2003:46), a phenomenon which is further described in Section 4.2.3. According to Hyman (2003:46), vowel height harmony is particularly common among Bantu languages and involves the assimilation of /i, u/ with a preceding mid vowel. Other types of harmony occur, such as those involving advanced tongue root (ATR), or rounding. A wide range of compensatory vowel lengthening processes also occur, e.g. as the result of contiguous vowels, in contexts preceding nasal + consonant sequences, and in penultimate syllables (Hyman 2003:48). Conversely, Hyman (2003:49) describes the phenomenon of vowel shortening that can occur, for example, at the end of a prosodic unit, or among any of the vowels preceding the ante-penultimate position in a phrase.

4.2.3. Syllable constraints and dynamics

PB syllables were constrained to the following shapes: CV(V) and V, and N, the latter two of which “were most likely limited to prefixes” (Hyman 2003:43). When analyzed as discrete consonants, the nasal (N) combines with consonants (C) in forming the type of NC sequences described in Section 4.2.1, and “the only consonant cluster in PB” (Hyman 2003:43). Hyman (2003:44) also describes syllabic nasals, which result from elision of the vowel in *mV-* prefixes.

Languages such as Ruwund L53 tend to elide word-final vowels, thereby resulting in words that end with a resyllabified consonant coda. A similar dynamic is described for Basáa A43, which features both processes of apocope and syncope. As might be expected, closed syllables are also found in a number of lexical borrowings. In general, however, the modern Bantu languages favor open syllables, perhaps in large part owing to their agglutinative morphology (Hyman 2003:44). This is true, for example, of Changana, which only allows syllables of the shape CV and V (Janson 1999:1).

4.2.4. Tone

It is widely known among linguists that the vast majority of languages in the Bantu family have a two-tone (H, L) system. Tone distinctions serve a variety of functions and mark affixes and stems, as well as phrasal and syntactic constituents (Nurse & Philippson 2003:8). For a full description of Bantu tone, see Kisseberth & Odden (2003).

5. Data collection and analysis

5.1. Data collection

The informants selected for this study were recorded in Lisbon, Portugal, the basic details of which are provided in Table 3. Two are female and three are male, all of Mozambican nationality, and ranging in age from 19 to 42. As indicated in Section 1, the two informants from Maputo (009 and 013) and the informant from Nampula (010) consider Portuguese to be their mother tongue, whereas the informants from Inhambane (017) and Tete (020) reported learning Portuguese as young children. At the time of recording, Informants 010 and 017 had been residing in Lisbon for a respective total of five and eight years. The remaining Informants 009, 013, and 020 had arrived in Portugal within a period of one week and one year. Given these circumstances, and the fact that the data elicitation were conducted by researchers from Lisbon, dialectal accommodation, or “adjustments in pronunciation and other aspects of linguistic behavior in terms of a drive to approximate one’s language to that of one’s interlocutor” (Trudgill 1983:143), should be considered a potential factor affecting the dialect of origin of some of the more long-term Lisbon residents. However, as the following analysis and attitudinal questionnaire responses (see Appendix) demonstrate, data for Informant 010, and to a lesser extent Informant 017, still show a rather dramatic influence by the substrate, and identification with the dialect of origin.

Informant ID	009	010	013	017	020
Age	30	42	19	35	22
Sex	male	male	female	male	female
Nationality	Mozambican	Mozambican	Mozambican	Mozambican	Mozambican
Place of origin	Maputo	Nampula	Maputo	Inhambane	Tete
Education level	Bachelor's degree	Bachelor's degree	High school diploma	Bachelor's degree course in progress	Bachelor's degree
Time spent in Lisbon	2 weeks	5 years	1 year	8 years	2 months
Languages spoken fluently	Portuguese (L1), Gi'Tonga	Portuguese (L1), English	Portuguese (L1), English	Chope (L1), Changana, Portuguese	Nyungwe (L1), Portuguese
Mother's L1	Changana	Makhuwa	Gi'Tonga	Chope	Nyungwe
Father's L1	Gi'Tonga	Makhuwa	Changana	Chope	N/A

Table 3. Informant details

Materials for the elicitation of read speech are based on those established in Rodrigues (2003), with the inclusion of a small set of additional words and phrases deemed necessary for capturing other relevant contexts. Audio recordings and stimulus prompts were controlled by an investigator, who remained seated in the same room as the informant, albeit in the periphery and not directly in front of the informant. For the read speech elicitation task, informants were asked to read the individual phrases and sentences projected in front of them on PowerPoint slides. Once this task was completed, the elicitation of spontaneous speech data was conducted in the form of an oral questionnaire for obtaining general speaker information and attitudinal data. Recordings were performed using a Marantz digital voice recorder, with a microphone positioned on the table in front of the informant. Later, the roughly 90-minute-long digital audio files were separated into recording blocks.

Corpus-based accent models were then developed through the assessment of segmental data by trained phoneticians, who use Praat (Boersma & Weenink 2010) to identify and label target segments, based on a combination of auditory judgment and waveform and spectrogram analysis. Each accent model underwent a separate pass by a total of three phoneticians until agreement was reached concerning the complete set of data points described.

5.2. Analysis

The main objective of this article is to provide qualitative descriptions of Bantu substratum interference in MP at the segmental level, and to open the way for future studies aimed at describing indigenizing Luso-African speech varieties. We acknowledge that quantitative studies involving informant populations that have remained in their birthplaces are needed to develop an understanding of how topolectal spoken varieties differ. However, such aims are beyond the scope and pragmatic constraints of the current study. Rather, in describing the following phenomena, we will attempt to link our observations with material from the literature on Bantu language phonologies (described in Section 4.2), and where possible, contemplate the motivation for certain processes in terms of governing principles – i.e. substratal interference, superstratal influence, and universal preference laws.

The MP examples presented in the following subsections are drawn from the read speech portion of the corpus. Where unclear from the orthography, the neighboring context is also transcribed. The symbols ‘.’ and ‘#’ respectively denote syllable and word boundaries. Readers should note that the following examples (5) through (16) include the relevant informant ID for a given pronunciation pattern, and are not necessarily meant to describe the speech patterns of all of the informants and varieties considered for this study.

5.2.1. Monothongization of diphthongs

The examples in (5) present a sample of words that feature diphthongs in EP, and their monothongization by Informants 009, 010, 017, 020, and to a lesser extent 013 (Maputo). One of the more noticeable characteristics of the MP varieties, compared

with EP, is the expression of full vowels in the context of the stressed phoneme /e/. These and other vowels are further realized as non-nasal in a variety of environments (see also Section 5.2.11). In terms of the diphthongs evident in EP, one observes the simplification of these vowel shapes in numerous word positions, and in multiple rising and falling formations. Of particular note is the resyllabification that occurs respectively in (5h) and (5j) for Informants 017 (Inhambane) and 009 (Maputo), and the restructuring of segments in (5n) by Informant 020 (Tete). In general, all of the examples in (5) appear to reflect a certain degree of faithfulness to local Bantu vowel inventories and phonotactics. Given that these alternate relatively freely with diphthongs in similar contexts, we also see, of course, the influence of the superstrate, or ‘source’ variety – both of which might suggest the occasional mitigation of universal preference laws for articulatory ease. Indeed, the monothongization of /ej/ also occurs in some Brazilian Portuguese varieties.

(5) EP	MP	Informant
a. S [ɛj] al <i>Seixal</i> (toponym)	s [ɛj] al	010, 017 ⁵
b. Ald [ɛj] a <i>Galega Aldeia Galega</i> (toponym)	Ald [ɛ] a <i>Galega</i>	020
c. ont [ɛj] <i>ontem</i> ‘yesterday’	ont [ɛ]	009, 010, 017, 020
d. b [ɛj] -me-quer <i>bem-me-quer</i> ‘forget-me-not’	b [ɛ] -me-quer	009, 010 ⁶
e. gara [ʒɛj] <i>garagens</i> ‘garages’	gara [ʒɛ]	010 ⁷ , 017, 020
f. materi [aj] <i>materiais</i> ‘materials’	materi [a]	010
g. [o] tocentos <i>oitocentos</i> ‘eight hundred’	[o] tocentos	017, 020
h. ap [o] o <i>apoio</i> ‘support’	ap [o.] o	017
i. investigaç [õj] <i>investigações</i> ‘investigations’	investigaç [on]	010
j. az [uj] <i>azuis</i> ‘blue’ (pl.)	az [u.ɨ]	009
k. [ew] ropeu <i>uropeu</i> ‘European’	[e] ropeu	020
l. r [ew] nião <i>reunião</i> ‘meeting’	r [u] nião	009, 010, 013, 017, 020
m. religi [ɛw] <i>religião</i> ‘religion’	religi [õ]	009
n. núc [lju] s <i>núcleo</i> ‘nucleus’	núc [ilu] s	020
o. diár [ju] s <i>diários</i> ‘dailies’	diár [o] s	013, 017, 020

5.2.2. Vowel sequence simplification and glide epenthesis

Vowel sequence simplification is illustrated in (6) for a selection of hiatus contexts both within and across word boundaries for Informants 009, 010, 013, 017, and 020. In many of the examples, one observes a tendency – characteristic of all the MP varieties examined in this study – to elide unstressed word-final vowels and syllables (see also Section 5.2.3), and preserve the stressed vowel.

5. For Informant 017, this syllable is realized with a highly intensity fricative [s] and empty nucleus.

6. This vowel is realized as [ɛ] by Informant 010.

7. The final consonant /s/ is also elided by this informant.

(6) <u>EP</u>	<u>MP</u>	<u>Informant</u>
a. m [ej.u] <i>meio</i> ‘means’	m [ej*]	013, 020
b. contrib [u.i] <i>contribui</i> ‘contributes’	contrib [u*]	010 ⁸
c. fal [u#a] lto <i>falo alto</i> ‘I speak loudly’	fal [*#a] lto	010, 020
d. J [u.ẽw] <i>João</i> ‘John’	J [ãw]	009, 013, 017, 020 ⁹
e. fal [u#oɔ] e <i>falo hoje</i> ‘I speak today’	fal [*#oɔ] e	010, 020

The examples in (7) show examples of glide epenthesis observed for Informants 009, 010, 013, and 017, whereas very little gliding overall was observed for Informant 020 (Tete). Again, the phenomenon of glide epenthesis is resonant of some Mozambican Bantu language phonologies, e.g. Makhuwa, and the often rigid constraints governing vowel co-occurrence. Glide epenthesis also occurs in the Shona variety of Zimbabwean English (Kadenge 2009:161-163), as well as in some Brazilian Portuguese varieties.

(7) <u>EP</u>	<u>MP</u>	<u>Informant</u>
a. r [i.a] <i>ria</i> ‘lagoon’	r [i.ja]	009, 017
b. pass [ja] mos <i>passeamos</i> ‘we walked’	pass [i.ja] mos	009
c. ró [zju] <i>róseo</i> ‘rosy’	ró [zi.ju]	009, 010, 017
d. pass [ej.u] <i>passeio</i> ‘footpath’	pass [ej.jo/u]	009, 013, 017
e. m [ej.ɔ] r <i>maior</i> ‘bigger’	m [aj.jɔ/o] r	010
f. Corr [oj.uj] <i>Corroios</i> (toponym)	Corr [oj.juj]	009, 017

5.2.3. Syncope, apocope, and apheresis

Examples of syncope, apocope, and apheresis for Informants 009, 010, 013, 017, and 020 are provided in (8). Again, we see the elision of word-final vowels and syllables, along with the loss of unstressed word-initial and word-medial syllables, and the frequent resyllabification and occasional reordering of segments. Echoes of substrate constraints appear evident in the elision of word-final vowels, resulting in words that end with a resyllabified consonant coda, e.g. in (8f) and (8h). Separately, in (8c), we see evidence that syncope of the penultimate pre-tonic syllable takes place after the evaluation of the preceding sibilant /s/ as a non-candidate for palatalization. And in (8h), we see what is essentially the same phenomenon, i.e. precedence of the palatalization constraint with a subsequent shuffling of syllable boundaries. The example in (8e), on the other hand, which was produced at the end of an intonational phrase, features the realization of a non-palatal coda [s], perhaps suggesting substrate interference of the nasal + consonant variety, similar to that which was described in Section 4.2.1.

The reader will also recall that in Section 4.1.3, a very similar dynamic was described for non-standard EP, whereby unstressed vowels – and particularly those which are realized as reduced vowels – frequently undergo a variable form of elision. Thus, it is difficult to tease apart the different principles behind the

8. The final vowel in this case of Informant 010 is followed by the low intensity sibilant [ʃ].

9. For Informant 020, the name *João* is realized with a denasalized vowel sequence and final nasal consonant, i.e. J[awn].

phenomena of syncope, apocope, and apheresis for speakers of MP. Clearly, the superstrate constraint governing the respective non-palatalization or palatalization of sibilants in onset and coda position has been faithfully adhered to in (8c) and (8h), despite the reorganization of syllables. It is also evident that the universal preference constraint for articulatory ease is in effect. However, given the degree of weakening phenomena observed for MP, and the fact that many of these words and phrases were produced in isolation, it appears that this type of constraint may be slightly less universal than is generally supposed and derives some of its motivation from the substrate. Alternatively, the wide-scale nature of the weakening observed could be the result of an overgeneralization of non-standard EP tendencies.

(8)	<u>EP</u>	<u>MP</u>	<u>Informant</u>
a.	[luke lize'sēw] <i>localização</i> 'location'	[loke *zə'sēw]	010
b.	[pri 'gar# rili zi'ēw] <i>pregar religião</i> 'preach religion'	[pre 'gar#* li zi'ō]	009 ¹⁰
c.	[munisi 'paɫ] <i>municipal</i> 'municipal'	[munis *'paɫ]	010
d.	[frigu 'rifiku] <i>frigorífico</i> 'refrigerator'	[frigo 'ri*fu]	020
e.	['ēti] <i>antes</i> 'before'	['an *s]	010
f.	[ɔr 'lēdu] <i>Orlando</i> 'Orlando'	[ɔr 'land*]	017
g.	[fu 'zimuɟ] <i>fugimos</i> 'we flee'	[fu 'zim*]	020
h.	[u # vi 'tidu] <i>o vestido</i> 'the dress'	[uv #*'jid*]	013

5.2.4. Consonant sequence simplification

As illustrated in (9), consonant sequence simplification occurs among all of the informants, and in the utterances of Informants 017 (Inhambane) and 020 (Tete), in particular. One observes epenthesis in (9a), (9b), and (9c), consonant assimilation in (9c), a reordering of the segments in (9h), and elision within the clusters [**ps**], [**ʃt**], [**ʒv**], [**ʃf**], [**fr**], and [**rb**] in (9c), (9d), (9f), (9g), and (9h). The elision in (9g) is unique among the other cases in that it occurs in the stressed syllable, indicating what appears to be a relatively strong degree of influence by the substrate. Informants displayed variable tendencies in the simplification of syllables containing liquids and faithfulness to the substratal constraint for open syllables, e.g. the realization of *hemoglobina* 'hemoglobin' as **hemo** [**gəl**] **bina**, and *intervém* 'intervenes' as **in** [**trə**] **vém** by Informant 013 (Maputo). Informant 020, on the other hand, showed some occasional restructuring of sequences involving nasal consonants, e.g. in the realization of *hifen* 'hyphen' as **hi** [**fne**], and revealed a more frequent form of substrate interference overall in the realization of open syllable formations. The example *hifen* further shows the interpretation by Informant 020 of the word ending *-en* as a sequence of two phonemes, i.e. /**en**/, indicating perhaps that phonological nasal vowels do not exist in this variety of MP.¹¹

10. While Informant 009 produces a trilled [**r**] in word-final position, the authors observe a tendency by some MP speakers to drop the final /r/, as is common in varieties of Brazilian Portuguese.

11. Evidence of this sort is further exemplified in the surfacing of non-homorganic nasal segments, e.g. **onte** [**m**] *ontem* 'yesterday' (observed across informants), and **regíme** [**n**] *regimen* 'regimen' (observed for Informant 013).

(9) EP	MP	Informant
a. E [dg] ar <i>Edgar</i> 'Edgar'	E [di/eg] ar	017, 020
b. o [mn] isciente <i>omnisciente</i> 'omniscient'	o [mən] isciente	010, 013, 017, 020
c. eru [ps] ão <i>erupção</i> 'eruption'	eru [pis] ão, eru [ts] ão, eru [*s]ão ¹²	009, 010, 020
d. inve [ʃti] gações <i>investigações</i> 'investigations'	inve [ʃ*i] gações	013
e. mai [ʒ#v] ale <i>mais vale</i> 'is better'	mai [ʃ#v] ale	020
f. tr [ēʃf] ormação <i>transformação</i> 'transformation'	tr [a*f] ormação	020
g. [fr] escas <i>frescas</i> 'fresh' (pl.)	[f*] escas	020
h. pertu [rb] ação <i>perturbação</i> 'disturbance'	pertu [*b] ação, per [tru] bação ¹³	010 017, 020

5.2.5. Voicing dissimilation and juxtaposition

The examples in (10) illustrate long-distance voicing dissimilation and voicing juxtaposition for Informant 010 (Nampula). Readers should note that this phenomenon was exclusively observed in the speech of Informant 010, who comes from a Makhuwa background, and does not describe the other varieties and informants considered for this study. Interestingly, both long-distance voicing dissimilation and voicing juxtaposition surface with surprising regularity in the utterances of Informant 010, offsetting the fact that this speaker has maintained a relatively long-term residence in Lisbon. Realizations such as **obce** [gat*] *obcecado* 'obsessed' summon memory of the Makhuwa constraint restricting the number of aspirated stops in stems to just one instance, despite the fact that in this case it is applied across an inflected word. We also see clear evidence of Dahl's law in (10a), and a tendency to realize voiced stops as voiceless in (10d), (10e), and (10h). Examples (10f) and (10g) show a more elaborated version of Dahl's law, as extended to Portuguese, in the juxtaposition of voiced and voiceless consonants. Here, we see evidence of strong substratal interference restricting the domain of voiced stops in non-contiguous, multi-stop sequences to non-final syllables – likely stemming from a constraint for the restriction of aspirated stops to word-final position.

12. The word *erupção* is realized by Informants 009 and 020 respectively as **eru [*s] ão** and **eru [ts] ão**.

13. The word *perturbação* is realized as **per [tru] bação** for Informant 010.

(10)	<u>EP</u>	<u>MP</u>	<u>Informant</u>
a.	com [p] uto <i>computo</i> ‘computation’	com [b] uto	010
b.	on [t] em <i>ontem</i> ‘yesterday’	on [d] em	010
c.	obce [k] ado <i>obcecado</i> ‘obsessed’	obce [gat^{h*}] a	010
d.	arren [d] a <i>arrenda</i> ‘leases, rents’	arren [t^h] a	010
e.	á [g] ua <i>água</i> ‘water’	á [k^h] ua	010
f.	[p] ren [d] as <i>prendas</i> ‘gifts’	[b] ren [t^h] as	010
g.	Ou [t] u [b] ro <i>Outubro</i> ‘October’	Ou [d] u [p^h] ro	010
h.	[gw] ar [d] a-sol <i>guarda-sol</i> ‘parasol’	[kw] ar [t^h] a-sol	010

5.2.6. Aspirated devoiced consonants

Leading from the analysis in the previous subsection, one observes frequent consonant devoicing (both partial and full) and elongated periods of aspiration in word-final position by Informants 009, 010, 013, 017, and 020. In all of the examples presented in (11), aspiration replaces the final vowel – which we find in respective standard and non-standard EP varieties as reduced or elided. Aspiration occurs in the high vowel context described in Hyman (2003:53) for Makhuwa and Doko, as well as in non-high contexts, and alternates somewhat freely with maintenance of the following word-final vowel. Thus, again, we see the interference of the substrate, which is more faithfully adhered to by some speakers, e.g. Informants 010 (Nampula) and 020 (Tete).

(11)	<u>EP</u> ¹⁴	<u>MP</u>	<u>Informant</u>
a.	be [bo] <i>bebo</i> ‘I drink’	be [p^h]	010
b.	ver [di] <i>verde</i> ‘green’	ver [t^h]	009, 010, 020
c.	rebuça [du] <i>rebuçado</i> ‘sweet’	rebuça [t^h]	009, 010, 013 ¹⁵ , 020
d.	Aldeia Gale [gɐ] <i>Aldeia Galega</i> (toponym)	Aldeia Gale [k^h]	010
e.	objecti [vu] <i>objectivo</i> ‘objective’	objecti [t^h]	009, 010, 013, 017, 020 ¹⁶
f.	quin [zi] <i>quinze</i> ‘fifteen’	quin [s^h]	009, 013, 017

5.2.7. Whistled fricatives

Another phenomenon observed across informants involves the realization of elongated, high acoustic energy fricatives that commonly result in a whistled articulation. Note that in the examples in (12), whistled fricatives are denoted with the superscript extension ‘^{wh}’. Like Shosted’s (2006) findings for Tshwa, the whistled fricatives in MP are produced allophonically, and vary significantly from their non-whistled counterparts in terms of their spectral features. This is illustrated in the spectrogram image in Figure 2 for **guarda [ʃ]** *civi*[ʃ^{wh}] *guardas civis* ‘civil guard’, which reveals a strong and protracted concentration of energy above 3000 Hz for the whistled articulation.

14. Recall that in EP, the voiced stops /b, d, g/ may also be realized as [β, ð, ɣ] in intervocalic contexts.

15. For informant 013, this sound is unaspirated.

16. For Informant 020, this sound is partially devoiced.

(12) <u>EP</u>	<u>MP</u>	<u>Informant</u>
a. [s] implório <i>simplório</i> ‘simpleton’	[s ^{wh}] implório	009, 010, 020
b. bá [ʃ] cula <i>báscula</i> ‘weighing scale’	bá [ʃ ^{wh}] cula	009, 010, 013
c. guardas civi [ʃ] <i>guardas civis</i> ‘civil guard’	guardas civi [ʃ ^{wh}]	010, 017 ¹⁷ , 020
d. parabén [ʃ] <i>parabéns</i> ‘congratulations’	parabén [ʃ ^{wh}]	010, 017
‘congratulations’		

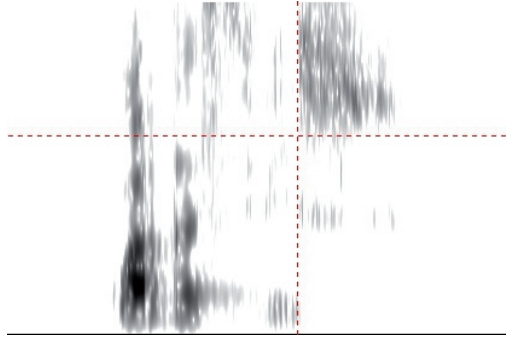


Figure 2. Spectrogram image of *guardas civis* (Informant 010)

5.2.8. Palatalization

In (13), examples are presented showing palatalization in different environments. This phenomenon was extremely common in the speech of Informant 020 (Tete), for whom palatalization occurs both progressively and regressively, affecting vowels and consonants alike. Note that in (13c), we also see evidence of a stop realized as an affricate in the context of a following high vowel, such as that described in Hyman (2003:53), and which is also a signature feature of most Brazilian Portuguese varieties.

(13) <u>EP</u>	<u>MP</u>	<u>Informant</u>
a. f [ejɜ] uma <i>fez uma</i> ‘did/made one’	f [ejɜ]	010, 017, 020
b. [ɛɲ] o <i>anho</i> ‘lamb’	[ajɲ] o	017, 020
c. o [ɜ#d] ias <i>os dias</i> ‘the days’	o [ɜ#dɜ] ias	020

5.2.9. Continuants

As in (10) and (11), the examples in (14) pertain almost exclusively to Informant 010 (Nampula). Here, we focus attention on this speaker’s realization of the stops /t, d/ as +continuant, i.e. [θ, ð], before +high and -high vowels, in both intervocalic environments and at the start of an intonational phrase, i.e. in (14a) and (14b). Example (14f) shows concurrent voicing dissimilation, whereby /d/ was realized as the voiceless labiodental fricative [θ], much like the phenomena described in

17. For Informant 017, the whistled production concerns the first word in the phrase, i.e. *guarda* [ʃ^{wh}] *guardas* ‘guard’.

Section 5.2.5. This example is also reflective of the assimilation of multiple voiceless coronal stops as all dental or all alveolar, as described in Kisseberth (2003:550). As witnessed in this and other examples, a more EP-like variety of continuantization of the voiced stops /b, d, g/ as [β, ð, ɣ] can be observed for Informants 009 and 017 (Inhambane).

The varieties of continuantization observed for Informant “010” are very consistent, and can quite easily be identified as resulting from interference by the substrate, which, in this case, is the particular variety (or varieties) of Makhuwa spoken by this informant’s parents. It is more difficult to tease apart the occasional continuantization observed for Informant 017, which is strictly limited to the [β, ð, ɣ] variety characteristic of EP. The fact that this speaker produces both stops and fricatives in intervocalic contexts could derive from the substrate or the superstrate, also perhaps reflecting the intervention, at times, of a more universal principle governing articulatory ease.

(14) <u>EP</u>	<u>MP</u>	<u>Informant</u>
a. [t] ira [t] eimas <i>tira teimas</i> ‘dictionary’ (colloq.)	[θ] ira [θ] eimas	010
b. [t] emos ... <i>Temos</i> ... ‘We have ...’	[θ] emos	010
c. [d] á-me dá-me ‘give me’	[ð] á-me	010
d. espiga [d/ð] íssima <i>espigadíssima</i> ‘spindliest’	espiga [ð] íssima	010
e. Quinta da Trinda [d/ð] e <i>Quinta da Trindade</i> (toponym)	Quinta da Trinda [ð] e	017
f. be [b/β] i [d/ð] as <i>bebidas</i> ‘beverages’	be [β] i [θ] as	010

5.2.10. Homorganic nasal epenthesis

Homorganic nasal epenthesis occurs across informants in “EP contexts” involving a nasal vowel followed by a consonant, as illustrated in (15). For the majority of informants and utterances where this phenomenon is realized, the conditioning environment is further restricted to the EP context of nasal vowel + stop, but see example (15f). Figure 3 shows waveform and spectrogram images for *trombone* ‘trombone’ by Informant 017 (Inhambane). Dotted lines mark the boundaries of the nasal tail [m], which are marked by the abrupt disappearance of the second formant on the left-hand side and a subsequent stop closure on the right-hand side. The non-homorganic nasal consonant [n] is shown as a means of revealing its spectral similarities with the nasal consonant in focus.

As described in Section 4.2.1, the homorganic nasal context is particularly relevant to the Bantu language family, both in the context of prenasalized consonants and across morpheme boundaries. The surfacing of nasal tails has also been described for varieties of Brazilian Portuguese (e.g. Lovatto et al. 2007). With respect to influencing factors, we see a significant degree of interference by the substrate for all MP varieties concerned, wherein the following consonant, and perhaps also Portuguese orthography, determines the articulatory specification

of the nasal phoneme which precedes it. If the somewhat controversial theory of phonological nasal vowels holds, here is one of the better examples showing the indigenization of MP in terms of a separate phonological construct that interprets such segments as a sequence of two phonemes instead of one. Of course, homorganic nasals also provide an optimal means of bridging the preceding vowel with the following consonant, thus reinforcing the regularity with which this process is applied both within and across informants. Thus, again we see the coalescence of substrate, superstrate, and universal preference laws in the phonetic transform.

(15) EP	MP	Informant
a. s [ĩ] plório <i>simplório</i> ‘simpleton’	s [í/im] plório	009, 010, 013, 020
b. [ẽ] belezam [ẽ̃ to] <i>embelezamento</i> ‘embezzlement’	[ẽm/em] belezam [ẽn/en] to	009, 010, 013, 017, 020
c. [ĩ] tervém <i>intevém</i> ‘intervenes’	[ín/in] tervém	009, 010, 013, 017, 020
d. eng [ẽ] drou <i>engendrou</i> ‘dreamed up’	eng [ẽn/en] drou ¹⁸	009, 010, 013, 017, 020
e. [ẽ̃] gústia <i>angústia</i> ‘anguish’	[ãŋ/aŋ] gústia	009, 010, 013, 017, 020
f. seixal [ẽ̃] se <i>seixalense</i> ‘of Seixal’	seixal [ẽn/en] se	013, 020

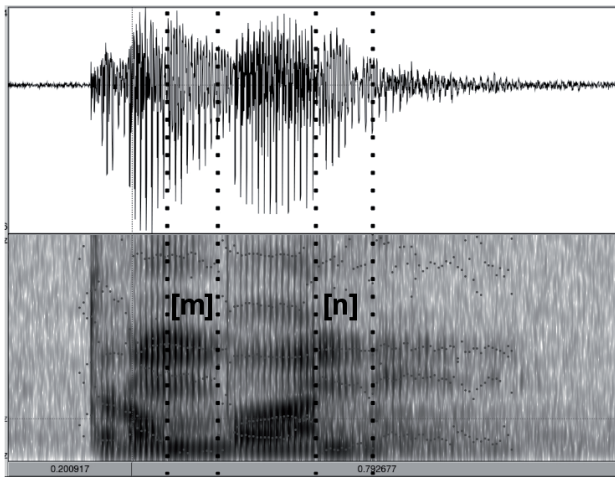


Figure 3. Waveform and spectrogram image of *trombone* (Informant 017)

18. Note that the initial sequence in *engendrou* is largely excluded from this phenomenon, given that the ‘g’ in this case is soft.

5.2.11. Denasalization

As indicated in Section 5.2.1, denasalization occurs across informants, and in the utterances of Informants 010 (Nampula) and 017 (Inhambane), in particular. As the examples in (16) illustrate, this process may be tied to certain high-frequency words, such as *ontem* (16c), and compound constructions, such as that presented in (16d). Alternatively, in (16b), we see possible evidence of the type of lenition described by Hyman (2003:50-51) for nasal + fricative combinations, while elision of the nasal in (16a) – and for some informants (16c) – appears to be conditioned by a substrate constraint for open syllables. Again, the fact that these denasalized articulations vary somewhat freely with more EP-like nasalized pronunciations indicates the interaction of substrate, superstrate, and universal preference laws – along with the possible influence of frozen lexical forms – in guiding how a sound is realized.

(16)	<u>EP</u>	<u>MP</u>	<u>Informant</u>
a.	Lat [i] <i>Latim</i> ‘Latin’	Lat [i]	010, 017
b.	[ẽ] <i>ngendrou engendrou</i> ‘dreamed up’	[e] <i>ngendrou</i>	010, 017
c.	ont [ẽj] <i>ontem</i> ‘yesterday’	ont [e/em] ¹⁹	009, 010, 013, 017, 020
d.	b [ẽj] -me-quer <i>bem-me-quer</i> ‘forget-me-not’	b [ẽ/e] -me-quer	009, 010, 013, 017, 020

6. Conclusion

The different MP varieties studied exhibited varying forms and degrees of Bantu substratum interference, including the monothongization of diphthongs, vowel and consonant sequence simplification, glide epenthesis, syncope, apocope, apheresis, the voicing dissimilation and juxtaposition of consonants, aspirated devoiced consonants in word-final position, whistled fricatives, palatalization, continuantization, homorganic nasal epenthesis, and denasalization. Where relevant, evidence from the literature on Bantu language phonological features was cited, along with an indication of the particular substrates in focus. Individually and overall, these descriptions offer testimony of the phenomena that characterize different spoken varieties of MP, and contribute toward a more balanced understanding of the linguistic variation that sets MP apart from other varieties of Portuguese. It is further hoped that these data will serve as a window in time for subsequent synchronic and diachronic studies aimed at describing the development, indigenization, variation, and change of MP in years to come.

It was shown that the above processes occur in the utterances of both L1 and L2 Portuguese speakers. Phenomena such as the voicing dissimilation and juxtaposition of consonants, aspirated devoiced consonants in word-final position, and continuantization were indeed more consistent and pervasive in the speech of certain L1 informants. Faithful adherence by the L1 Informant 010 (Nampula) was demonstrated for some of the more well-documented features of Makhuwa. These include: a rigid restriction on the number of aspirated stops allowed in a stem

19. See footnote 9.

(or word, in this case), Dahl's law dissimilation dynamics and extended versions thereof for restricting aspirated stops to word-final position, and the assimilation of multiple voiceless coronal stops as all dental or all alveolar. Such evidence is particularly interesting in light of the limited faculty reported by Informant 010 for speaking Makhuwa (the native language of his parents), and the fact that this speaker has spent the past five years living in Lisbon. One possible explanation may have to do with the status of Makhuwa as Mozambique's most widely spoken indigenous language. Perhaps coming from a language community of eight million speakers leads to a more 'validated' and codified form of substratum interference. Of course, the fact that Informant 010 is a native speaker of Portuguese calls into question the appropriateness of terms like 'substrate', thus signalling the need for an updated nomenclature and set of practices concerning the study of indigenizing language varieties.

Finally, because interference by the superstrate and substrate is no more static than the language varieties they influence, we devoted additional attention to the contemplation of how these two systems interact, both with one another and with universal preference laws for conditioning the phonetic transform in indigenizing languages. For the majority of processes described above, variable forms of interaction were observed for each of these constraint bundles in guiding how a sound is realized. Thus, much like the findings in Alber & Plag (1999) for creole development, it would appear that all three play a fundamental and occasionally overlapping role in shaping the indigenized language.

Acknowledgements

The authors gratefully acknowledge the full support of the Fundação para a Ciência e a Tecnologia (PTDC/CLE-LIN/100335/2008) for this research.

References

- Alber, B. & I. Plag 1999. 'Epenthesis, deletion and the emergence of the optimal syllable in Creole: the case of Saranan', *Lingua* 111:811-840.
- Bladon, A., C. Clark & K. Mickey 1987. 'Production and perception of sibilant fricatives: Shona data', *Journal of the International Phonetic Association* 17:39-65.
- Boersma, P. & D. Weeink 2010. Praat: Doing Phonetics by Computer, computer program, v. 5.1.43. Retrieved from <http://www.praat.org> 4 August 2010.
- CIA. The World Factbook—Mozambique. Retrieved from <https://www.cia.gov/library/publications/the-world-factbook/geos/mz.html> 10 December 2010.
- Da Conceição, M. 1999. 'A brief look at the sociolinguistics of Ronga and other languages spoken in Mozambique', *Working Papers in Linguistics* 16:9-30.
- Duarte, I., A. Gonçalves, M. Miguel & M. A. Mota 1999. 'Não cheguei de aprender nada – Áreas de variação e tendências de mudança no Português de Moçambique'. In A. C. Lopes & C. Martins (eds), *Actas do XIV Encontro Nacional da Associação Portuguesa de Linguística* 1, 477-493. Lisbon: Associação Portuguesa de Linguística.

- Ferguson, C. A. 1959. 'Diglossia', *Word* 15:325-340.
- Firmino, G. 1995. 'O caso do Português e das Línguas Indígenas de Moçambique', *Revista Internacional de Língua Portuguesa* 13:33-43.
- , 2000. *Situação Linguística de Moçambique*. Maputo: Instituto Nacional de Estatística .
- , 2002. *A Questão Linguística na África pós-colonial: O Caso do Português e das línguas autóctones em Moçambique*. Maputo: Promédia.
- Fishman, J. A. 1967. 'Bilingualism with and without diglossia; diglossia with and without bilingualism', *Journal of Social Issues* 23,2:29-38.
- Gonçalves, P. 1990. 'A construção de uma gramática do Português em Moçambique: Aspectos da estrutura argumental dos verbos'. Doctoral thesis, Universidade de Lisboa, Lisbon.
- , 1994. 'Movimentos sintácticos no Português de Moçambique'. In *Momentos da Crítica Literária VIII, Campina Grande*, 471-479.
- Gonçalves, P. 1996. *Português em Moçambique: Uma Variedade em Formação*. Maputo: Livraria Universitária & Faculdade de Letras da UEM.
- , 1997. 'Tipologia de erros do Português oral de Maputo'. In C. Stroud & P. Gonçalves (eds), *Panorama do Português Oral de Maputo. Vol III: A Construção de um Banco de Erros. Cadernos de Pesquisa* 24:37-70. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- , 2002a. 'The role of ambiguity in second language change: the case of Mozambican African Portuguese', *Second Language Research* 18:325-347.
- , 2002b. 'Os pronomes pessoais reflexivos não-argumentais no Português de Moçambique'. In S. Grosse & A. Schonberger (eds), *Ex Oriente Lux: Festschrift für Eberhard Gärner zu seinem 60. Geburtstag*, 231-239. Frankfurt am Main: Valentia.
- , 2010. *A génese do Português de Moçambique*. Lisbon: Imprensa Nacional - Casa da Moeda.
- Gonçalves, P. & F. Chimbutane 2004. 'O papel das línguas Bantu na Génese do Português de Moçambique: o comportamento sintáctico de constituintes locativos e direccionais', *Papia* 14:7-30.
- , 2009. 'Assimetrias da mudança linguística em situação de contacto de línguas: o caso do português e das línguas bantu em Moçambique'. In A. M. Carvalho (ed.), *Português em Contato*, 31-52. Madrid: Iberoamericana.
- Gonçalves, P. & B. Siteo 1999. 'Mudança linguística em situação de contacto de línguas: o caso do Changana e do Português', *Travessias* 1:73-86.
- Guthrie, M. 1967-1971. *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages*. Vols. 1-4. Farnborough: Gregg Press.
- Hyman, L. M. 2003. 'Segmental Phonology'. In D. Nurse and G. Philippson (eds), *The Bantu Languages*, 42-58. New York: Routledge.
- Janson, T. 1999. 'Prenasalized voiced aspirated stops in Changana/Tsonga'. Paper presented at Fonetik 99, Göteborg University.
- Janson, T. & O. Engstrand 2001. 'Some unusual sounds in Changana', Lund University Department of Linguistics. *Working Papers* 49:74-77.
- Kadenge, M. 2009. 'African English: the indigenization of English vowels by

- Zimbabwean native Shona speakers', *The Journal of Pan African Studies* 3:1:156-173.
- Kisseberth, C. 2003. 'Makhuwa (P30)'. In D. Nurse and G. Philippson (eds), *The Bantu Languages*, 546-565. New York: Routledge.
- Kisseberth, C. & D. Odden 2003. 'Tone'. In D. Nurse and G. Philippson (eds), *The Bantu Languages*, , 59-70. New York: Routledge.
- Ladefoged, P. & I. Maddieson 1996. *The Sounds of the World's Languages*. Oxford: Blackwell Publishing Ltd.
- Lopes, A. J. 1979. 'Mozambican Portuguese Words and Expressions'. In *Longman English Dictionary for Portuguese Speakers*. Harlow: Longman ELT.
- Lopes, A. J. 1999. 'The Language Situation in Mozambique'. In Robert B. Kaplan & Richard B. Baldauf (eds), *Language planning in Malawi, Mozambique, and the Philippines*. Tonawanda, New York: Multilingual Matters Ltd.
- Lovatto, L., A. Amelot, L. Crevier-Buchman, P. Basset & J. Vaissière 2007. 'A fiberoptic analysis of nasal vowels in Brazilian Portuguese', *Proceedings of ICPHS*, Saarbrücken, 549-552.
- Machungo, I. 2000. Neologisms in Mozambican Portuguese: A morphosemantic study. Doctoral thesis, University of Ghana, Legon.
- Maddieson, I. 2003. 'The sounds of the Bantu languages'. In D Nurse and G. Philippson (eds), *The Bantu Languages*, 15-41. New York: Routledge.
- Maho, J. 2003. 'A classification of the Bantu languages: an update of Guthrie's referential system'. In D. Nurse and G. Philippson (eds.), *The Bantu Languages*, 639-651. New York: Routledge.
- Mateus, M. H. & E. d'Andrade 2000. *The Phonology of Portuguese*. New York: Oxford University Press.
- Newitt, M. 2002. 'Mozambique.' In P. Chabal (ed.), *A History of Postcolonial Lusophone Africa*, 185-235. London: Hurts & Company.
- Nurse, D. & G. Philippson 2003. 'Introduction'. In D. Nurse and G. Philippson (eds), *The Bantu Languages*, 1-12. New York: Routledge.
- Rego, S. 2000. Contributo para a Constituição de um Corpus de Portuguesismos em Nyungwe. Master's thesis, Universidade de Lisboa, Lisbon.
- Rodrigues, M. C. M. 2003. *Lisboa e Braga: Fonologia e Variação*. Lisbon: Fundação Calouste Gulbenkian.
- Sankoff, G. 2002. 'Linguistic outcomes of language contact'. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Oxford: Blackwell Publishing Ltd.
- Shosted, R. K. 2006. 'Just put your lips together and blow? Whistled fricatives in Southern Bantu', *Proceedings of ISSP 2006: 7th International Seminar on Speech Production*, Belo Horizonte, 561-572.
- Sitoe, B. 1991. 'Empréstimos lexicais do Português no Tsonga', *Revista Internacional de Língua Portuguesa* 5/6:106-113.
- Sitoe, B. 1996. *Dicionário Changana-Português*. Maputo: Instituto Nacional do Desenvolvimento da Educação.
- Stroud, C. 2007. 'Multilingualism in ex-colonial countries'. In P. Auer & Li Wei (eds), *Handbook on Multilingualism and Multilingual Communication*, 509-538. Berlin: Mouton de Gruyter.

- Stroud, C. & P. Gonçalves 1997. *Panorama do Português Oral de Maputo. Volume 1: Objectivos e Métodos*. Maputo: INDE.
- Thomason, S. & T. Kaufman 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Trudgill, P. 1983. *On Dialect*. New York: Basil Blackwell Publishing Ltd.

Appendix

Informant ID	009	010	013	017	020
What is your accent?	Mozambican.	African.	Mozambican.	Mozambican.	Nyungwe ²⁰ .
Do you like your accent?	Yes, because it's a source of identity.	Yes, because it's how I identify myself.	I love it because it connects me to where I came from and where I grew up.	Yes, I love it.	Yes, I like my dialect.
Which accent do you like most?	I'm indifferent. For me, there is no better or worse.	Lisbon, because it's very soft and rhythmic.	Mine, because I think it's lovely.	Mine!	I like the accent of Lisbon because they speak Portuguese very well here, and I would like to learn more Portuguese.
Which accent do you like least?	It doesn't exist.	Alentejano ²¹ , because I can't understand it.	The accent of São Tomé, because the Rs are not pronounced.	Brazilian.	The accent of Guinea, because when they speak, I can't understand anything.
If you had to choose another accent, what would it be, and why?	It doesn't exist.	Lisbon, because it's easier to understand.	Brazilian, because it's very distinct and different from the others.	Madeiraense ²² , because it seems joyful.	The Portuguese of Lisbon, because I like the Portuguese.
Have you ever had a situation where other people do not understand what you say?	No, others understand me.	Yes, I had situations working in a call center where I was misunderstood.	I think that people understand me, but sometimes I use MP words derived from our dialect, and people ask what I meant by that. I do not feel bad about it.	Yes I had situations like this, and I felt embarrassed.	Yes, I had situations like this when my friends couldn't understand me.
What opinion do you think others have about your accent?	I don't know.	I don't think they consider it bad.	I think people do not care about it, and even like it. I don't know anyone who doesn't.	Many people think that all Africans have the same accent.	I don't know.
Which is the accent you hear most on television and the radio?	Southern Mozambique.	Several, but I hear the accents of Lisbon and Porto most.	Lisbon.	Lisbon.	The Portuguese of Lisbon.

Table 4. Attitudinal questionnaire and informant responses

20. It is unclear whether Informant 020 was describing her native language accent of Nyungwe or the Nyungwe variety of Portuguese.
 21. The accent of Alentejo, in the south-central region of Portugal.
 22. The accent of Madeira.

Authors' addresses

Simone Ashby
Instituto de Linguística Teórica e Computacional (ILTEC)
Avenida Elias Garcia, 147-5° Dto.
1050-099 Lisboa
Portugal
ashbyhanna@gmail.com

Sílvia Barbosa
Instituto de Linguística Teórica e Computacional (ILTEC)
Avenida Elias Garcia, 147-5° Dto.
1050-099 Lisboa
Portugal
silvia@iltec.pt

Résumé

Les contraintes phonologiques bantu exercent divers types d'interférences dans les variétés de portugais parlées au Mozambique. Alors que nombre de ces processus – tels que dissimilation du voisement, simplification des séquences consonantiques et dénasalisation – sont relativement bien connus des linguistes spécialisés dans l'étude des dialectes portugais mozambicains, la littérature ne fait que les mentionner.

En outre, les données disponibles ne permettent pas de cerner clairement la façon dont les influences du substrat bantu caractérisent le portugais des locuteurs L1 et ceux qui ont acquis la langue à un très jeune âge. Le présent travail est basé sur l'analyse des données phonétiques acquises auprès de cinq informateurs locuteurs natifs portugais (trois locuteurs L1 et deux L2) provenant de divers lieux du Mozambique, et dont les langues sont le changana, le chope, le gitonga, le makhua et le nyungwe. Nous apportons les preuves des effets des phonologies bantu locales sur les variétés de portugais parlées par les informateurs pour les phénomènes mentionnés plus haut, et pour d'autres processus concordant avec la littérature relative à la linguistique bantu tels que les consonnes de fin de mot aspirées dévoisées, les fricatives « sifflées », et les épenthèses nasales homorganiques. Nous concluons que les études effectuées sur l'interférence phonétique et phonologique du substrat sont importantes pour le suivi à long terme d'une situation de contact, et aide à élucider les principes qui gouvernent la segmentation dans les langues locales.