



CÁTEDRA DE PORTUGUÊS
Língua Segunda e Estrangeira

CADERNOS
DE PESQUISA
#02

NEOLOGISMOS
DO PORTUGUÊS
DE
MOÇAMBIQUE

ORGANIZAÇÃO
INÊS MACHUNGO



CÁTEDRA DE PORTUGUÊS
Língua Segunda e Estrangeira

CADERNO DE PESQUISA Nº 2
NEOLOGISMOS DO PORTUGUÊS DE MOÇAMBIQUE

Organização

Inês Machungo

Maputo

2016

O projecto “Cadernos de Pesquisa” nasceu da necessidade de se criar um veículo económico e rápido, destinado a divulgar os produtos da investigação realizada pela equipa da Cátedra de Português Língua Segunda e Estrangeira, ou resultantes de projectos de investigação apoiados pela Cátedra.

Este projecto pretende servir como indicador da vitalidade da pesquisa sobre o português língua não materna, em várias dimensões.

Cada número dos “Cadernos de Pesquisa” é dedicado a uma área específica, sob orientação de organizador(es) especializado(s).

Conselho Editorial

Perpétua Gonçalves (coordenação geral)

Inês Machungo

Conceição Siopa

Capa: Carla Pacheco Faria

Impressão: Imprensa Universitária - Maputo, Moçambique

Tiragem: 300 exemplares

Endereço Institucional:

Faculdade de Letras e Ciências Sociais - Universidade Eduardo Mondlane

Campus Universitário Principal - Maputo, Moçambique

Endereço electrónico: catedra.uem.ic@gmail.com

Esta edição foi financiada pelo Camões - Instituto da Cooperação e da Língua - Portugal.

ÍNDICE

APRESENTAÇÃO	4
NEOLOGISMOS DO PORTUGUÊS DE MOÇAMBIQUE: A CONSTRUÇÃO DE UMA BASE DE DADOS	5

APRESENTAÇÃO

No Caderno de Pesquisa nº 2, "Neologismos do Português de Moçambique", apresentam-se produções neológicas que ocorrem no Português de Moçambique, recolhidas e armazenadas no Observatório de Neologismos do Português de Moçambique (ONPM), uma base de dados aberta e de acesso livre, disponível em www.catedraportugues.uem.mz. Pretende-se com este Caderno disponibilizar, de forma sistematizada, material lexical que pode constituir fonte de pesquisa no domínio da linguística aplicada, em particular no ensino do português, na pesquisa em lexicologia/lexicografia e em sociolinguística, entre outros.

Na primeira parte deste Caderno indicam-se os objectivos do ONPM, e aspectos relacionados com as opções metodológicas inerentes a este tipo de pesquisa. Na segunda parte, apresenta-se a base de dados, constituída por 714 entradas lexicais.

A pesquisa que culminou com a constituição da base de dados dos neologismos do Português de Moçambique foi desenvolvida em diferentes fases e, em cada uma delas, foi crucial o apoio de membros da equipa da Cátedra e de outros colaboradores.

Agradeço à Margarita Correia, Judit Freixa e Vítor Lindegaard pela discussão frutuosa em torno de diferentes aspectos teóricos e metodológicos inerentes à pesquisa. Agradeço também o apoio de Ezra Nhampoca, Marta Siteo, Amina Carimo e Castigo Massango na recolha e digitalização dos dados.

Inês Machungo

NEOLOGISMOS DO PORTUGUÊS DE MOÇAMBIQUE: A CONSTRUÇÃO DE UMA BASE DE DADOS

1. Introdução

A língua portuguesa em Moçambique, é língua segunda para cerca de 40% dos falantes, que exibem diferentes graus de proficiência linguística. Nos últimos 20 anos, o número de falantes de português como língua segunda ou como língua nativa expandiu-se consideravelmente, com reflexos notórios no seu desenvolvimento a todos os níveis e, particularmente, no domínio lexical: assistimos, todos os dias, ao aparecimento de palavras novas na língua corrente e nas línguas de especialidade.

Esta forma de criação lexical, ou seja, o recurso ao uso de palavras novas, tem motivações de carácter linguístico e/ou extralinguístico:

(i) O português é uma língua herdada de uma cultura exógena que não possui vocabulário para expressar determinadas realidades etnossocioculturais locais;

(ii) A competência e repertório lexical em português são reduzidos e, como tal, os falantes criam palavras novas para expressar conceitos para os quais já existem palavras;

(iii) Os falantes, embora não os dominando, conhecem os processos morfológicos de formação de palavras em português;

(iv) O substrato linguístico bantu fornece material lexical que permite criar novas formas lexicais e integrá-las no repertório lexical do Português;

Estes são os factores que enformam a grande criatividade neológica no seio dos falantes do Português de Moçambique (PM), e é neste contexto que se criou, no sítio da Cátedra de Português Língua Segunda e Estrangeira, um Observatório dos Neologismos do Português de Moçambique (ONPM).

São objectivos do ONPM:

- Fazer o levantamento dos neologismos que ocorrem no discurso quotidiano dos falantes do PM, na imprensa escrita e falada, e em outras produções escritas e orais devidamente identificadas;
- Descrever e analisar os recursos que os falantes do PM utilizam para actualizar o seu repertório lexical;
- Criar uma base de dados do léxico do PM disponível na *Internet*, com inclusão periódica das novas unidades lexicais;

- Contribuir para a actualização do léxico do português em dicionários de língua geral, em glossários e vocabulários;
- Estabelecer uma plataforma que facilite a permuta de informação com outros observatórios de neologia ou unidades similares.

2. Métodos e problemas

A abordagem do facto neológico, ou seja, os métodos usados na sua identificação são controversos. Foram muito criticados os critérios adoptados por Guilbert (1975), para quem, para fazer um “juízo neológico”, o investigador tem de possuir uma competência lexical de tal maneira vasta que lhe permita determinar se uma certa unidade lexical já foi ou não usada num discurso. Em estudos mais recentes sobre neologia, (Cabré 2004, Correia et al. 2004) o critério lexicográfico tem sido o mais usado; de acordo com este critério, para determinar sobre a neologicidade de um item lexical, deve-se verificar a sua atestação, ou não, num *corpus de exclusão* constituído por dicionários de língua ou outros *corpora* criados para o efeito. Dado que se trabalha com um volume enorme de dados, ferramentas adequadas (por ex: SEXTAN, WebCorp, entre outras) são construídas para a extracção automática ou semiautomática de neologismos.

No ONPM, assumiu-se que *neologismo* é uma unidade de léxico (palavra, lexia ou sintagma), cuja forma significante ou cuja relação significado-significante, caracterizada por um funcionamento efectivo num determinado modelo de comunicação, não se tinha realizado no estágio imediatamente anterior do código da língua (Rey 1976). Tomando como base esta perspectiva, foram identificados, no PM, os seguintes processos básicos de criação lexical:

(i) *neologismos formais*, que resultam de processos gramaticais de formação de palavras.

Ex: cabelaria [[cabelo] + [aria]] ‘cabeleireiro’

Neste grupo destacam-se os *empréstimos*, que resultam da importação de itens lexicais de outras línguas.

Ex: xitiqueiro [[xitique] + [eiro]] ‘membro de um grupo de xitique¹’

¹ *Xitique*, (*xitiki* nas línguas Ronga/Changana) é um tipo de organização em que os membros do grupo de ‘xitique’ contribuem com um certo valor monetário previamente estabelecido, cujo total é atribuído a cada membro, de forma rotativa.

(ii) *neologismos semânticos*, provenientes da atribuição de um significado novo a uma palavra existente.

Ex: chapa '*meio de transporte colectivo de passageiros*'

Na identificação dos neologismos, adoptou-se o critério lexicográfico, com as adaptações necessárias ao quadro sociolinguístico do PM. É sabido que classificar uma unidade como um neologismo tem sempre um carácter subjectivo, dependendo da intuição linguística de cada investigador. Contudo, no ONPM, foram seleccionadas como candidatas a neologismos, formas que não têm atestação lexicográfica, muito embora, sob o ponto de vista sociolinguístico e psicológico, algumas destas unidades nem sempre sejam sentidas pelos falantes como neológicas, nem como formas alógenas, isto é, provenientes de outras línguas. As formações por via de empréstimo, em particular às línguas bantu, não são vistas como formas externas ao português, ao invés, são frequentemente sentidas pelos falantes como componentes intrínsecas do acervo lexical do PM.

A adopção do critério lexicográfico permitiu, pois, identificar neologismos que para os falantes parecem estar perfeitamente integrados na língua, mas que não constam dos dicionários de língua (ex: *calamidade* 'roupa usada'). Permitiu igualmente detectar neologismos que, tendo tido uma grande vitalidade em determinado período histórico-social, caíram em desuso sem que tivessem sido dicionarizados. Estão neste caso, por exemplo, palavras como *xirico* 'rádio', *continuador* 'criança', *nação* 'capital'.

Algumas das formas do PM aqui consideradas como neológicas encontram-se registadas nos poucos produtos de carácter lexicográfico existentes (cf. Lopes et al. 2002, Dias 2002). Contudo, entendeu-se que a sua inclusão nesta base de dados é útil, tendo em vista a elaboração, no futuro, de dicionários de língua.

Com esta abordagem pretende-se incluir no acervo lexical do ONPM, todas as unidades lexicais novas que são parte do repertório lexical dos falantes, tornando-o num filtro para a elaboração de vocabulários, dicionários e outros recursos lexicográficos do PM.

3. Constituição do ONPM

3.1 O corpus de extracção

Assumindo que os *media* são por excelência um veículo de divulgação de neologismos (Alves 2007) o *corpus* de extracção, isto é, o conjunto de textos de onde foram extraídos os neologismos incluídos no ONPM é constituído por:

(i) Jornais de larga difusão, editados em Maputo e Beira ("Notícias", "O País", "Diário de Moçambique", "Savana", "Zambeze") que foram despojados de forma sistemática, em particular as secções de opinião pública, entrevistas, política, desporto e economia;

(ii) Jornais de circulação mais restrita ("A Verdade", "Mediafax") bem como periódicos editados em outras regiões do país, de frequência irregular ("Horizonte", "Wamphula") que foram despojados esporadicamente;

(iii) Textos escritos de diferentes origens, tais como: revistas universitárias, desportivas, textos propagandísticos que aparecem no mercado com certa regularidade. A inclusão deste tipo de textos deve-se ao facto de neles se encontrarem neologismos espontâneos que não foram submetidos aos filtros de correcção por que os periódicos de larga difusão passam;

(iv) Produção oral espontânea, pois é na oralidade que se encontra um grande número de formas neológicas (Rádio, TV ou outras formas de comunicação oral).

As técnicas de extracção automática e semiautomática de neologismos foram preteridas pela de extracção manual, pelas razões que a seguir se apresentam:

(i) Nos poucos jornais disponíveis por via electrónica, há rubricas da versão impressa, importantes para a recolha de neologismos, que não constam da versão electrónica do jornal;

(ii) O despojamento automático requer a construção de ferramentas electrónicas para o seu manuseamento, de custo elevado e de que ainda não dispomos;

(iii) O despojamento automático não permite a identificação de neologismos semânticos e sintagmáticos. Por exemplo, neologismos como *cinzentinho* ('polícia de protecção') ou *casa de mãe espera* ('casas construídas junto às maternidades, em zonas rurais, onde as parturientes aguardam para dar à luz'), não são identificáveis por processos automatizados de detecção de neologismos.

A desvantagem da utilização sistemática do despojamento manual é o facto de ser um processo muito moroso e oneroso, por envolver um grande número de investigadores e muitas horas de trabalho. Além disso, no caso do ONPM, grande parte dos investigadores são jovens e, como tal, têm um conhecimento restrito do léxico da língua portuguesa; por esta razão, itens lexicais que fazem parte do léxico da língua portuguesa, são amiúde seleccionados como candidatos a neologismos; em contrapartida, formas que são locais, por serem frequentes, não são sentidas como neológicas.

A extracção automática e/ou semiautomática, a que, pelas razões acima apontadas, não se teve acesso na construção da base de dados do ONPM, tem a vantagem de sendo processos mais rápidos de manuseamento dos dados, reduzirem os custos, e tem “ainda como mais-valia clara a objectividade, a sistematicidade, a exaustividade, a minimização de erros, e a possibilidade de reutilização dos resultados” (Correia et al. 2004).

3.2 O *corpus* de exclusão

Trabalhos mais recentes de neologia utilizam um ***corpus lexicográfico de exclusão*** ou seja, um conjunto de dicionários de língua e outras bases de dados lexicais, usados para determinar se uma unidade lexical é neológica sob o ponto de vista lexicográfico. Por outras palavras, pelo critério lexicográfico, excluem-se todas as unidades que se encontrem registadas em dicionários gerais de língua e outros *corpora* lexicais. A natureza do corpus de exclusão é, pois, determinante para os resultados da pesquisa neológica (Correia & Antunes, 2010). A inexistência de dicionários de língua do PM tem implicações na metodologia adoptada na identificação das unidades neológicas do PM. Assim, e tendo em conta que a variante moçambicana do português ainda não está padronizada, e que, a nível oficial, se optou por usar como referência a norma padrão europeia, o corpus de exclusão do ONPM é constituído por dicionários de língua do português europeu e outras bases de dados lexicais (Vocabulário Ortográfico do Português - VOP). A grande influência que a variante brasileira do português tem em Moçambique ditou que fossem também incluídos dicionários do Português do Brasil, estando o corpus de exclusão assim constituído:

- *Dicionário da Língua Portuguesa Contemporânea* - Academia de Ciências de Lisboa e Editorial Verbo, 2001;
- *Grande Dicionário da Língua Portuguesa* - Porto Editora, 2004;
- *Dicionário Houaiss da Língua Portuguesa* - Temas e Debates (Versão Portuguesa), 2003.
- *Dicionário Priberam da Língua Portuguesa*. [Em linha], 2008 - 2016.
- *Vocabulário Ortográfico do Português - (VOP)* - ILTEC, 2010

À abordagem lexicográfica na identificação de neologismos e tomando em consideração critérios utilizados por outros observatórios de neologismos (por exemplo, OBNEO²), foram impostas algumas restrições. Por exemplo, não são considerados neologismos os aumentativos, os diminutivos, os superlativos, os advérbios em *-mente*, quer constem ou

² Observatori di Neologia do Institut Universitari de Lingüística Aplicada da Universidade Pompeu Fabra em Espanha.

não do corpus de exclusão, por, potencialmente, serem derivados ilimitados. Foram também excluídas as palavras que contêm o prefixo *ex-* quando se junta a radicais simples ou derivados que façam referência a profissões ou relações pessoais, entre outros aspectos (Cabré 2003) (ex: *ex-irmão*). Contudo, nos diminutivos e aumentativos, tomou-se em consideração os casos em que as palavras são usadas com significados diferentes dos da palavra derivante (por exemplo, *girinho* recarga com tempo de antena para comunicação telefónica via celular).

4. A construção da base de dados dos candidatos a neologismos

Para a inserção, manuseamento e armazenamento dos candidatos a neologismos foi construído um programa informatizado inspirado em outros desenhados para idênticos fins, em particular os programas usados pelo Observatori de Neologia do IULA (Institut Universitari de Lingüística Aplicada de la Universitat Pompeu Fabra, Espanha), e pelo Observatório de Neologia do Português do ILTEC (Instituto de Linguística Teórica e Computacional, Portugal). A ficha neológica contém os seguintes campos: neologismo, categoria gramatical, marca tipográfica, tipo de neologismo, contexto, nota, localização do neologismo e autor³.

Veja-se abaixo, exemplos de fichas de registo de neologismos criadas e usadas no ONPM:

The image shows a web form for registering a neologism. At the top, there are navigation links: 'Pesquisa', 'Novo neologismo' (highlighted), 'Utilizadores', and 'Sair'. On the right, it says 'Na sessão como imachungo'. The form is divided into three main sections on the left: 'Neologismo' with a text input field, 'Contexto' with a larger text area, and 'Notas' with another text area. On the right, the 'Categorização' section contains several dropdown menus: 'Categoria gramatical' (selected: 'adj (adjectivo)'), 'Tipo de neologismo' (selected: 'FABR (formal por abreviação)'), 'Tipo de importação' (selected: '- (-)'), and 'Fonte' (selected: 'DM (Diário de Moçambique)'). There is also a field for 'Outra fonte' and a date field for 'Data da fonte' (16-06-2016). Below this is the 'Marcas tipográficas' section with four checkboxes: 'Aspas', 'Parêntesis', 'Itálico', and 'Negrito'. A 'Guardar' button is located at the bottom right of the form.

Figura 1 - Ficha para registo de neologismo

³ Para visualizar fichas completas poderá abrir a ligação http://www.catedraportugues.uem.mz/?__target__=observatorio&page=busca&id=820

[« Voltar à pesquisa](#)

814. **butique-inclina**

Contexto

"Marta gosta de roupas da <butique-inclina>"

Notas

=mercado informal, caracterizado pelo facto de os produtos à venda serem expostos no chão, obrigando os compradores a se curvarem para poderem apreciá-los.

Categorização

Categoria gramatical: **substantivo fem. sing. (f)**
 Tipo de neologismo: **formal por composição (FCOM)**
 Tipo de importação: - (-)
 Fonte: **Televisão (TV)**
 Data da fonte: **11-05-2010**

Marcas tipográficas

Aspas: **Não** Itálico: **Não**
 Parêntesis: **Não** Negrito: **Não**

 [Imprimir esta página](#)

Figura 2 - Ficha de neologismo

5. Apresentação da base de dados do *Caderno de Neologismos*

Na presente base de dados, cada entrada lexical contém informação da categoria gramatical e fonte do neologismo. O item lexical é apresentado dentro do contexto linguístico de onde foi extraído, e nos casos em que o contexto não seja esclarecedor, adicionou-se um equivalente ou informação que permita a apreensão do significado do item lexical. A inclusão do contexto e equivalente revela-se particularmente importante no caso dos empréstimos e dos neologismos semânticos. Assim, neste *Caderno*, uma entrada lexical tem o seguinte formato:

carrinha unidose
s.f.
o Hospital Provincial de Tete introduziu a chamada <Carrinha Unidose>. Nesta carrinha são colocadas quantidades de medicamentos exactas a serem administradas aos doentes

FCOM Notícias 2011

A ortografia das palavras, em particular a dos empréstimos às línguas bantu, revelam uma certa variação, resultante de não haver ainda uma escrita estandardizada⁴ para este tipo de palavras (por exemplo, *tchovar/txovar/chovar* ou *xitique/chitique/xitiki*); assim, optou-

⁴ O Vocabulário Ortográfico Comum da Língua Portuguesa (VOC), de que o Vocabulário Ortográfico do Português de Moçambique (VOMOLP) é parte, cujo objectivo é o de estabelecer e oficializar normas ortográficas comuns para as diferentes variedades do Português, é uma ferramenta indispensável na resolução deste tipo de constrangimentos.

se pelo uso da forma que apresenta maior número de ocorrências nos *corpora* consultados.

A base de dados do ONPM é uma base aberta que está em constante renovação e actualização. A metodologia adoptada na sua construção, permite que a informação nela contida possa ser utilizada em pesquisa linguística sobre o PM, e para apoio ao ensino do português. Ela constitui também uma fonte importante na obtenção de dados que alimentam o Vocabulário Ortográfico do PM e outros recursos lexicográficos que venham a ser elaborados.

Referências:

Alves, Ieda. (2007). *Neologismo. Criação Lexical*. 3^a. ed. São Paulo: Ática.

Cabré, Teresa. (2004). Metodología del trabajo en neología: criterios, materiales y procesos. *Observatori di Neologia, papers de l'IULA, sèrie Monografies*, 9. Barcelona: Universitat Pompeu Fabra, Institut Universitari de Linguística Aplicada.

Correia, Margarita et al. (2004). O Observatório de Neologia do Português – ONP: criação e apresentação. *Actas do XX Encontro da Associação Portuguesa de Linguística*, Lisboa: APL, pp. 471- 482.

Correia, Margarita e Antunes, Mafalda. (2010). A constituição do corpus de exclusão - problemas e efeitos. *Puente - Revista de Traducción, Interpretación y Terminología*, 9: 33-48.

Dias, Hildizina. (2002). *Minidicionário de Moçambicanismos*. Maputo: Edição da Autora.

Guilbert, Louis. (1975). *La Créativité Lexicale*. Paris: Librairie Larouse.

Lopes, et al. (2002). *Moçambicanismos- Para um Léxico de Usos do Português Moçambicano*. Maputo: UEM - Livraria Universitária.

Rey, Alain. (1976). Néologisme: un pseudo-concept?. *Cahiers de Lexicologie*, 28: 3-17.

ETIQUETAS PARA A CLASSIFICAÇÃO DOS NEOLOGISMOS

1. O campo *categoria gramatical* contém os seguintes códigos:

adj	adjectivo
adj pl	adjectivo plural
adv	advérbio
art	artigo
conj	conjunção
f	feminino
f pl	feminino plural
interj	interjeição
loc	locução
m	masculino
m pl	masculino plural
prep	preposição
pron	pronome
v intr	verbo intransitivo
v pron	verbo pronominal
v trans	verbo transitivo

2. O campo *“tipo de neologismo”* contém os seguintes códigos:

EMP	empréstimo
Efr	empréstimo do francês
Eing	empréstimo do inglês
Elb	empréstimo de língua bantu
EAfr	empréstimo adaptado do francês
EAing	empréstimo adaptado do inglês
EAlb	empréstimo adaptado de língua bantu
FABR	formal/ abreviação
FACR	formal/ acronímia
FCOMP	formal/composição
FCULT	formal/composição culta
FCONV	formal/conversão
FPRSU	formal/prefixação e sufixação
FLEX	formal/lexicalização
FPRE	formal/prefixação
FSIG	formal/siglação
FSINT	formal/sintagmática

FSUF	formal/sufixação
FVAR	formal/variação
O	outro
SEM	semântico
SINT	sintático

3. Campo *fontes*:

DM	Jornal Diário de Moçambique
JD	Jornal Domingo
JH	Jornal Horizonte
JN	Jornal Notícias
JS	Jornal Savana
JZ	Jornal Zambeze
RD	Rádio
TV	Televisão
OR	Oral
O	Outras

Símbolos

≈ equivalente a