



Universidade Federal da Bahia
Instituto de Letras
Programa de Pós-Graduação em Língua e Cultura
Rua Barão de Jeremoabo, nº147 - CEP: 40170-290 - Campus Universitário Ondina Salvador-BA
Tel.: (71)3283 - 6256 – Site: <http://www.ppgll.ufba.br> - E-mail: pgletba@ufba.br



ISAMAR NEIVA DE SANTANA

VOCABULÁRIO DIALETAL BAIANO

V.1

SALVADOR
2017

ISAMAR NEIVA DE SANTANA

**VOCABULÁRIO DIALETAL BAIANO
V.1**

Tese apresentada ao Programa de Pós-Graduação em Língua e Cultura da Universidade Federal da Bahia, como requisito parcial para obtenção do grau de Doutor em Letras (Língua e Cultura).

Área de concentração: Linguística Histórica
Orientador: Prof. Dr. Américo Venâncio Lopes Machado Filho

**SALVADOR
2017**

Modelo de ficha catalográfica fornecido pelo Sistema Universitário de Bibliotecas da UFBA para ser confeccionada pelo autor

NEIVA, Isamar
VOCABULÁRIO DIALETAL BAIANO / Isamar NEIVA. -- Salvador,
2017.
vol. 1. 270 f. : il

Orientador: AMÉRICO VENÂNCIO LOPES MACHADO FILHO.
Tese (Doutorado - PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA E
CULTURA) -- Universidade Federal da Bahia, ILUFBA, 2017.

1. Lexicografia. 2. Variação lexical. 3. Dialectologia. 4.
Dicionário Dialectal Brasileiro (DDB). 5. Bahia . I. MACHADO
FILHO, AMÉRICO VENÂNCIO LOPES. II. Título.

À minha mãe, pela dedicação de uma vida inteira a cuidar de mim com amor incondicional. É por você, Mãe.

À memória de Dil, minha tia-mãe, pelo legado de fé, conhecimento e amor que me deixou e do qual jamais esquecerei. Por ter me ensinado através de seus versos que “os espinhos podem curar mãos calejadas” até que estes mesmos “brotem flores”.
“Viva o Amor.”

AGRADECIMENTOS

Agradeço a Deus, meu Mestre e Senhor, Companheiro e Amigo, por ser meu escudo, refúgio e o sentido de minha existência.

Ao meu orientador, o professor doutor Américo Venâncio Lopes Machado Filho, por me fazer enxergar horizontes onde e quando eu só via neblina. E à sua linda família, Silvana, Lis e o “caçula” Nuno, pelo imensurável carinho de sempre.

Ao meu orientador em Paris, o professor doutor Salah Mejri. *Je vous remerci pour tout.*

À professora e amiga Inês Sfar, *Merci infiniment, pour tout.*

À Banca Examinadora, presidida pela professora doutora Célia Telles e as professoras doutoras Suzana Cardoso, Aparecida Isquerdo, Marcela Paim e Cândida Seabra, pelas contribuições que dadas ao trabalho.

À Banca do Exame de Qualificação, os professores doutores Emílio Pagotto e Silvana Ribeiro pelas contribuições dadas ao desenvolvimento da tese.

Às professoras Jacyra Mota e Suzana Cardoso, minhas mães acadêmicas, pelo apoio nas decisões mais difíceis e pela oportunidade de pertencer à Família ALiB. À Marcela Paim, pelo cuidado em todo o processo de Doutorado-Sanduíche.

À coordenadora do Programa de Pós-Graduação em Língua e Cultura – a professora Alícia Duhá Lose, pela atenção. E aos “meninos” da Secretaria da Pós-Graduação – Ricardo, Thiago e Cris – pela disposição e cuidado. Vocês valem ouro!

À FAPESB – Fundação de Amparo à Pesquisa do Estado da Bahia – e à CAPES-COFECUB – durante o doutorado-sanduíche – pelo investimento com o qual foi possível aprofundar o conteúdo acadêmico.

A minha família pela proteção e cuidado. Por encherem a minha casa de riso quando vazios e distâncias se fizeram presentes. Por fazerem os quatro meses e meio – período do meu doutorado-sanduíche, em Paris – passarem mais rápido e menos sofrível para minhas “veinhas” – mãe e tia. Em especial, a

Lis, Anne, Caren e Carol pela presença marcante e efetiva desde o meu primeiro vestibular até a reta final da conclusão desta tese.

Aos professores do PROFICI - PROEMF, em especial Angelo, Priscila e Ingrid, pelo conhecimento adquirido. Merci! Aos colegas e amigos, pelo conhecimento partilhado. Merci!

Ao Grupo Nêmesis, pelo sentido de grupo.

Aos amigos brasileiros que ganhei em Paris, Sônia, Fran, Ju Vinuto, Dani, Bruna Lago e a amiga mexicana Belém. *Mes amis, vous sont fantastics.*

À professora Ana Maria Cortez por me mostrar a Paris que poucos conhecem e a Ana Maravilhosa, que poucos ainda descobriram!!!

À Maison du Brésil, onde residi em Paris, pelo respeito e dedicação com que tratam todos os seus residentes.

À IBNJ Paris, especialmente ao pastor Diogo e pastora Natasha. À IADEP, em especial ao pastor Evilásio e irmã Celina. E ao “Harpinha” pelo apoio e carinho, em todos os momentos.

A todos que, direta ou indiretamente, fazem parte dessa trajetória, obrigada!

Em situação de poço, a água equivale a uma palavra em situação dicionária: isolada, estanque no poço dela mesma, e porque assim estanque, estancada; e mais: porque assim estancada, muda e muda porque com nenhuma comunica, porque cortou-se a sintaxe desse rio, o fio de água por que ele discorria.

(MELO NETO, 1999, p. 350-351)

Un dictionnaire commencerait à partir du moment où il ne donnerait plus les sens, mais les besognes des mots.

(BATAILLE, Documents, 1968)

RESUMO

NEIVA, Isamar. **Vocabulário Dialeto Baiano**. 2017. 270 p. Tese de Doutorado (Programa de Pós-graduação em Língua e Cultura). Universidade Federal da Bahia.

A evidência do fato de que, mesmo com os avanços nos estudos lexicais, os trabalhos de pesquisa de cunho lexicográfico, em especial no Brasil, ainda não privilegiam satisfatoriamente a variação, configurou-se um ponto fundamental para a construção desta tese, o *Vocabulário Dialeto Baiano* – primeiro produto do *Projeto Dicionário Dialeto Brasileiro* (DDB), o qual se desenvolve no âmbito do *Projeto Atlas Lingüístico do Brasil* (Projeto ALiB). A tese registra, lexicograficamente, itens lexicais em variação, presentes no *corpus* constituído de 92 inquéritos realizados em 22 localidades do estado da Bahia e obtidos como respostas a 202 questões concernentes a 14 áreas conceituais – *Acidentes geográficos; Fenômenos atmosféricos; Astros e tempo; Atividades agropastoris; Fauna; Corpo humano; Ciclos da vida; Convívio e comportamento social; Religião e crenças; Jogos e diversões infantis; Habitação; Alimentação e cozinha; Vestuário e acessórios e Vida urbana*. A tese é apresentada em três volumes. O Volume I, além de considerações preliminares e finais, consta de capítulos que versam sobre os pressupostos teórico-metodológicos adotados, explicitam os procedimentos metodológicos aplicados na pesquisa e apresentam resultados da análise dos dados. Os volumes II e III destinam-se ao registro lexicográfico dos dados obtidos na pesquisa e se constituem de textos pré e pós-dicionarísticos que ressaltam a relevância do projeto para o conhecimento da realidade plurilingüística brasileira no que tange ao léxico em uso na fala utentes do português brasileiro contemporâneo. No Volume II – versão fascicular – apresentam-se os fascículos temáticos concernentes às 14 áreas do QSL ALiB e relações possíveis. A versão geral do *Vocabulário Dialeto Baiano* encontra-se no Volume III.

Palavras-chave: Variação lexical. Lexicografia Variacional. Vocabulário dialeto baiano.

RÉSUMÉ

NEIVA, Isamar. **Vocabulaire Dialectal Bahianais**. 2017. 270 pages. Thèse de Doctorat (Programme de Post-Doctorat en Langue et Culture). Université Fédérale de Bahia.

Malgré les études lexicales élaborées, il est évident que les travaux en lexicologie spécialement au Brésil, n'ont pas aboutis d'une façon satisfaisante l'étude de la variation. Ceci constitue un point fondamental pour l'élaboration de cette thèse, le *Vocabulaire Dialectal Bahianais* – premier produit du *Projet Dictionnaire Dialectal Brésilien* (DDB), celui-ci se développe dans le cadre du *Projet Atlas Linguistique du Brésil* (Projet ALIB). Cette thèse enregistre lexicalement, des items lexicaux en variation se trouvant dans le corpus qui font partie de 92 enquêtes réalisées dans 22 localités de l'état de Bahia et obtenus comme réponses à 202 questions relatives à 14 domaines conceptuels – *Accidents géographiques ; Phénomènes atmosphériques ; Astres et Temps ; Activités agro-pastorales ; Faune ; Corps Humain ; Cycles de la Vie ; Comportement et vie en Société ; Religion et croyances ; Jeux et amusements infantiles ; Habitation ; Alimentation et cuisine ; Habillement et accessoires ; Vie urbaine*. La thèse est présentée en trois volumes. Le Volume I contient des considérations préliminaires et finales. Les chapitres abordent les présupposés théorique-méthodologique adoptés, explicitent les procédés méthodologiques pris en compte dans la recherche et présentent les résultats de l'analyse des données. Les volumes II et III visent le registre lexicographique des données obtenues dans cette recherche et se constituent de textes pré et post-dictionnaristes qui mettent en évidence l'importance du projet qui vise la connaissance de la réalité plurilinguistique brésilienne en ce qui touche le lexique en usage en portugais brésilien contemporain. Dans le Volume II – version fasciculaire – sont présentés les fascicules thématiques relatifs aux 14 domaines du QSL ALIB et leurs possibles relations. La version générale du *Vocabulaire Dialectal Bahianais* se trouve dans le Volume III.

Mots-clé: Variation lexical. Lexicographie Variationnelle. Vocabulaire dialectal bahianais.

LISTA DE FIGURAS

FIGURA 1 -	Imagens do <i>Facebook</i>	31
FIGURA 2 -	Distribuição das unidades lexicais	38
FIGURA 3 -	A periodização das gramáticas portuguesas	45
FIGURA 4 -	Proposta de periodização da constituição do léxico do português - Portugal e Brasil	47
FIGURA 5 -	Teste de avaliação das variantes no <i>corpus</i>	52
FIGURA 6 -	De “Je ne sais pas” a “Ché pas”: exemplificação do processo de variação / mudança	53
FIGURA 7 -	Diagrama redutor dos níveis de análise da língua	77
FIGURA 8 -	Micro e macrolinguística	82
FIGURA 9 -	Taxonomia de obras lexicográficas	101
FIGURA 10 -	Diagrama relacional dos objetos teóricos dos estudos do léxico, em função do sistema linguístico e do grau de gramaticalidade ou lexicalidade	116
FIGURA 11 -	Excerto de <i>wordlist</i> do <i>corpus</i> por frequência e em ordem alfa	129
FIGURA 12 -	Colocações de NOVO no <i>corpus</i> analisado	130
FIGURA 13 -	Concordância de “mais novo” nos dados do ALiB, concernentes à Bahia	131
FIGURA 14 -	Chave de consulta atual do Vocabulário Dialetal Baiano	141
FIGURA 15 -	Legenda geolinguística do DDB	152
FIGURA 3 -	Diagrama do processo de elaboração da legenda geolinguística do Vocabulário Dialetal Baiano.....	152
FIGURA 17 -	Verbetes de ‘papagaio’ do Vocabulário Dialetal Baiano, relativo à Fauna	175
FIGURA 18 -	Verbetes de ‘papagaio’ do Vocabulário Dialetal Baiano, relativo à Brinquedos e diversões infantis	175
FIGURA 19 -	Verbetes de ‘papagaio’ do Vocabulário Dialetal Baiano, relativo à Convívio e comportamento social	176

LISTA DE QUADROS

QUADRO 1 -	Propostas de periodização para a história da Língua Portuguesa	44
QUADRO 2 -	Síntese das principais periodizações do português brasileiro, com ênfase nos fatos sócio-históricos	46
QUADRO 3 -	Formas divergentes no português	49
QUADRO 4 -	Síntese dos tipos de entrada de lexias herança latina no léxico português	49
QUADRO 5 -	Critérios de classificação dos Atlas linguísticos quanto à coleta dos dados	87
QUADRO 6 -	Critérios de classificação dos Atlas linguísticos quanto à apresentação dos dados	91
QUADRO 7 -	Estudos com base em dados do Projeto ALiB	97
QUADRO 8 -	Distribuição das localidades analisadas	126
QUADRO 9 -	Microestrutura preliminar do DDB, para verbetes plenos, em 2010	137
QUADRO 10 -	Microestrutura atual do verbete pleno constante do Vocabulário Dialetal Baiano	139
QUADRO 11 -	Metaplasmos	143
QUADRO 12 -	Confronto entre a pergunta do questionário ALiB e a proposta de semagrama	150
QUADRO 13 -	Questões constantes do APFB e do ALiB	

LISTA DE TABELAS

TABELA 1 -	Principais domínios em que a influência pré-românica	56
TABELA 2 -	Domínios dos germanismos	57
TABELA 3 -	Domínios da dominação árabe	58
TABELA 4 -	Domínios da influência além-mar	59
TABELA 5 -	Domínios da influência cultural europeia	60
TABELA 6 -	Síntese das influências do contato linguístico por domínios conceituais	61
TABELA 7 -	Domínios da influência indígena	62
TABELA 8 -	Domínios da influência dos imigrantes	66
TABELA 9 -	Influência francesa e inglesa no léxico do português	69
TABELA 10 -	Português europeu (PE) versus Português brasileiro (PB): Diferenças lexicais	71
TABELA 11 -	Léxico geral versus Léxico especializado: algumas características	79
TABELA 12 -	A produtividade da lexia 'base' na língua portuguesa: principais acepções	80
TABELA 13 -	A produtividade da lexia 'base' na língua portuguesa: acepções por rubricas	80
TABELA 14 -	A produtividade da lexia 'base' na língua portuguesa: formação de novas lexias	81
TABELA 15 -	Principais metaplasmos descritos em obras lexicográficas da segunda fase da geolinguística no Brasil: (AMARAL, 1920), (NASCENTES, 1922) e (MARROQUIM, 1934)	104

LISTA DE GRÁFICOS

Gráfico 1 -	Distribuição do comportamento léxico-variacional de 'Rio de pequena extensão'.....	159
Gráfico 2 -	Distribuição do comportamento léxico-variacional de 'Qualquer elemento de madeira que sirva de passagem de pessoas por cima de um rio pequeno'.....	161
Gráfico 3 -	Distribuição do comportamento léxico-variacional de 'Ponto de término de um rio ou de desaguamento em outro'	162
Gráfico 4 -	Distribuição do comportamento léxico-variacional de 'Fenômeno natural que implica no movimento em rotação espiral, ocasional / eventual, que forma um funil na água'	163
Gráfico 5 -	Distribuição do comportamento léxico-variacional de 'Fruta cítrica, pouco ácida, de cheiro forte, cuja casca se solta, facilmente, dos gomos, podendo ser descascada com as mãos'	164
Gráfico 6 -	Distribuição do comportamento léxico-variacional de 'Leguminosa de formato longitudinal e aspecto gômico, em cujo interior constam sementes de cor avermelhada usada na alimentação humana em diversas formas de cocção'	165
Gráfico 7 -	Distribuição do comportamento léxico-variacional de 'Tipo de flores brancas e amarelas com as quais se prepara uma infusão empregada como calmante e digestivo'	166
Gráfico 8 -	Distribuição do comportamento léxico-variacional de 'Cada um dos grupos de cachos de banana'	166
Gráfico 9 -	Distribuição do comportamento léxico-variacional de 'Inflorescência da bananeira'	167
Gráfico 10 -	Distribuição do comportamento léxico-variacional de 'Bananas que nascem grudadas'	169
Gráfico 11 -	Distribuição do comportamento léxico-variacional de 'Gêmeos'	170
Gráfico 12 -	Distribuição do comportamento léxico-variacional de 'Raiz comestível, de cor branca por dentro e coberta por uma casca marrom'	171
Gráfico 13 -	Distribuição do comportamento léxico-variacional de 'Raiz branca por dentro, coberta por uma casca marrom destinada, exclusivamente, à elaboração de farinha'	171
Gráfico 14 -	Distribuição do comportamento léxico-variacional de 'Veículo de propulsão humana, constituído de uma única roda dianteira e duas hastes na parte oposta usado para transporte de pequenas cargas'	173
Gráfico 15 -	Distribuição do comportamento léxico-variacional de 'Ave de penas coloridas que quando domesticado, pode emitir sons como se na linguagem humana'	174
Gráfico 16 -	Distribuição do comportamento léxico-variacional de 'Animal mamífero que, quando se sente ameaçado, emite fedor	176
Gráfico 17 -	Distribuição do comportamento léxico-variacional de 'Glândula	177

	mamária da vaca'	
Gráfico 18 -	Distribuição das variantes secundárias de 'peito'	178
Gráfico 19 -	Distribuição das variantes secundárias de 'úbere'	178
Gráfico 20 -	Distribuição do comportamento léxico-variacional de 'Patas dianteiras do cavalo'	179
Gráfico 21 -	Distribuição do comportamento léxico-variacional de 'Sanguessuga'	180
Gráfico 22 -	Distribuição do comportamento léxico-variacional de 'Calcanhar' ..	181
Gráfico 23 -	Distribuição do comportamento léxico-variacional de 'Cócegas'	182
Gráfico 24 -	Distribuição do comportamento léxico-variacional de 'Caçula'	183
Gráfico 25 -	Distribuição do comportamento léxico-variacional de 'menstruação'	186
Gráfico 26 -	Distribuição do comportamento léxico-variacional de 'menopausa'	186
Gráfico 27 -	Distribuição do comportamento léxico-variacional de 'veneziana'	190
Gráfico 28 -	Ausência de resposta para o referente 'veneziana'	191
Gráfico 29 -	Distribuição do comportamento léxico-variacional de 'cachaça'	192
Gráfico 30 -	Distribuição do comportamento léxico-variacional de 'rouge'	192
Gráfico 31 -	Estratificação social das variantes concernentes a 'rouge'	193
Gráfico 32 -	Distribuição do comportamento léxico-variacional de 'rouge'	196

SUMÁRIO

VOLUME 1

INTRODUÇÃO	18
PARTE I – FUNDAMENTAÇÃO TEÓRICA	23
Capítulo 1	
ENTRE PALAVRAS	24
1.1 O CARÁTER NOMEADOR E DINÂMICO DAS PALAVRAS	25
1.1.1 No tempo: a linha tênue entre arcaísmos e neologismos	33
1.2 O CARÁTER ESTRUTURAL DAS PALAVRAS	34
1.2.1 O problema da palavra ‘palavra’: revendo a discussão que deu muito “pano pra manga”	34
1.2.1.1 Unidades monolexicais	36
1.2.1.2 Unidades polilexicais	37
CONSIDERAÇÕES DO CAPÍTULO	39
Capítulo 2	
PELOS CAMINHOS DO LÉXICO PORTUGUÊS: BREVE PANORAMA	40
2.1 ANDANÇAS: PASSOS QUE FIZERAM, FAZEM E RECRIAM O CAMINHO	43
2.1.1 “Caminhando e cantando e seguindo a canção” e os passos: O arcabouço lexical do português	48
2.1.1.1 A herança latina	48
2.1.1.2 PORTUGAL	54
2.1.1.2.1 <u>Contato linguístico</u>	54
2.1.1.2.1.1 Influência pré-românica	54
2.1.1.2.1.2 Influência pós românica	56
2.1.1.2.1.3 Influência das línguas do além-mar	58
2.1.1.2.2 <u>Influência cultural</u>	60
2.1.1.3 BRASIL	61
2.1.1.3.1 <u>Contato linguístico</u>	62
2.1.1.3.1.1 O aluvião indígena	62
2.1.1.3.1.2 A presença africana	63
2.1.1.3.1.3 Base portuguesa	65
2.1.1.3.2 Influência dos imigrantes	66
2.1.1.3.3 <u>Influência cultural</u>	66
2.1.1.3.3.1 Estrangeirismos	66

2.2 O LÉXICO PORTUGUÊS EUROPEU E O LÉXICO PORTUGUÊS BRASILEIRO:	69
2.2.1 Áreas dialetais em Portugal e no Brasil	72
CONSIDERAÇÕES DO CAPÍTULO	74

Capítulo 3

OS ESTUDOS DO(S) LÉXICO(S) NO BRASIL	75
3.1 CARACTERIZAÇÃO DOS ESTUDOS (CIENTÍFICOS) DOS LÉXICOS	77
3.1.1 Tipos de estudos dos léxicos	81
3.1.1.1 <i>Os caminhos dos estudos dialetais no Brasil</i>	83
3.1.1.1.1 <u>Atlas linguísticos brasileiros: do APFB ao ALiB</u>	85
3.1.1.1.1.1 O Atlas Linguístico do Brasil	94
3.1.1.1.1.1.1 Estudos com base de dados do Projeto ALiB	97
3.1.1.2 <i>Lexicografia</i>	98
3.1.1.3 <i>Lexicultura</i>	107
3.1.2 Novos rumos para os estudos dos léxicos: interfaces	109

PARTE II 113

Capítulo 4

POR UMA LEXICOGRAFIA VARIACIONAL: “REVENDO RAZÕES” E CONCEITOS	114
3.1 PALAVRA, PALAVRINHA, PALAVRÃO, PALAVRA-CHAVE, PALAVREADO (E OUTRAS PALAVRAS): SOBRE O OBJETO TEÓRICO-OBSERVACIONAL ADOTADO	
3.2 A FINALIDADE DE UM VOCABULÁRIO DIALETAL: RAZÕES	118

PARTE III – METODOLOGIA

Capítulo 5

PROCEDIMENTOS METODOLÓGICOS PARA A ELABORAÇÃO DO VOLUME I 125

5.1 DA CONSTITUIÇÃO DO *CORPUS* AO TRATAMENTO DOS DADOS 125

5.1.1 As localidades: dos critérios de seleção aos aspectos socioculturais 125

5.1.2 Informantes: do conceito de indivíduo ao conceito de comunidade de fala 127

5.1.3 Métodos: da coleta de dados às estratégias de tratamento dos dados 127

5.1.3.1 Estratégias para o tratamento dos dados 129

5.1.3.1.1 Lematização dos dados e delimitação do corpus 129

5.1.3.2 Estratégias de documentação: a transcrição dos dados 132

5.1.3.2.1 Estratégias de arquivamento dos dados: base de dados 133

5.1.4 Critérios para análises dos dados 134

Capítulo 6

PROCEDIMENTOS METODOLÓGICOS PARA A ELABORAÇÃO DOS VOLUMES II E III (VOCABULÁRIO DIALETAL BAIANO) 135

6.1 MACROESTRUTURA 135

6.1.1 Os processos de constituição do verbete 142

6.1.1.1 Critérios para elaboração das entradas próprias 142

6.1.1.2 Sistema remissivo 145

6.1.1.3 Aspectos lexicográfico-culturais 148

6.1.1.3.1 Definição e formulação da pergunta do QSL 148

6.1.1.3.2 Abonação 151

6.1.1.3.3 Legenda geolinguística 152

6.1.1.3.5 Achegas enciclopédicas 153

6.2 GLOSSÁRIO DE TERMOS TÉCNICOS USADOS NO VOCABULÁRIO DIALETAL BAIANO 154

PARTE IV – ANÁLISE DOS DADOS

Capítulo 7

ANÁLISES DOS DADOS EM PERSPECTIVA DIALETOLÓGICA E/OU LEXICOLÓGICA

157

7.1 ANÁLISE DOS DADOS POR ÁREAS TEMÁTICAS

157

7.2 A VARIAÇÃO LEXICAL NAS MESORREGIÕES DA BAHIA

197

7.3 DO ATLAS PRÉVIO DOS FALARES BAIANOS AO VOCABULÁRIO DIALETAL BAIANO: A VARIAÇÃO ESPAÇO-TEMPORAL DAS LEXIAS DOCUMENTADAS

202

CONSIDERAÇÕES FINAIS

219

REFERÊNCIAS

224

APÊNDICES

ANEXOS

INTRODUÇÃO

INTRODUÇÃO

[...] a Lexicografia contemporânea tem caminhado no sentido de perscrutar, a passos largos e de forma incontestavelmente sólida, esse movediço alicerce que serve de base às variegadas gramáticas das línguas naturais: o léxico. (MACHADO FILHO, 2010, p. 51)

Com esse novo viés, um olhar mais acurado passa a ser admitido pelos estudos¹ que têm o léxico como objeto de análise, na tentativa promissora de transformá-lo em uma parte mais amplamente observável da língua, sobretudo no que concerne a seu processo de formação.

Entretanto, mesmo com os avanços nos estudos lexicais, grande parte dos dicionários, principalmente aqueles elaborados a partir da segunda metade do século XX, apresentam muitas deficiências estruturais, além de serem redundantes e pouco adequados para a consulta.

Tem-se observado ainda que, na percepção do usuário, a finalidade dos dicionários para a maioria da sociedade é similar à busca pelas Gramáticas – livros que representam o conjunto de regras gramaticais, segundo uma norma de uso, a norma-padrão, em geral – e consiste em dirimir dúvidas relativas, em geral, à grafia correta, ao significado e à busca por sinônimos, bem como ocorre no que tange à pronúncia, considerando o fato de muitos dicionários apresentarem noções de ortoépia – indicação normativa de uma pronúncia. Vê-se, ainda que

em relação ao registro da variação nos dicionários contemporâneos, publicados no Brasil, se refere meramente a marcas de uso, que normalmente refletem uma certa carga de preconceito em face do padrão ortográfico que neutraliza quaisquer outras atualizações linguísticas que se possam insinuar na nomenclatura. (MACHADO FILHO, 2010, p. 52)

Com vistas a contribuir para o entendimento de língua como instrumento social de comunicação diversificado, é, pois, oportuno considerar a emergência da continuidade e aprofundamento de realização de um estudo dialetológico – de cunho lexicográfico – ou lexicográfico – de cunho dialetológico.

Diante dos fatos emergentes, Machado Filho (2010) propôs uma interface entre a Dialetologia e a Lexicografia, que consiste na elaboração de um Dicionário Dialetal Brasileiro (DDB), com base em dados do Projeto Atlas

¹ Notícias acerca dos estudos dos léxicos são trazidas no *Capítulo 3* desta tese.

Linguístico do Brasil (Projeto ALiB) – a fim de sistematizar a variação lexical patente nas diversas regiões do país e, por conseguinte, construir um dicionário capaz de apreender os diversos falares da população brasileira e assegurar, minimamente, a difusão mais ampla do conhecimento cultural do país, no que tange aos usos lexicais.

Pautando-se, pois, na perspectiva apresentada, um ponto a ser observado e fundamental para a construção desta tese, o *Vocabulário Dialeto Baiano*, é o fato de que, mesmo com os avanços nos estudos lexicais, os trabalhos de pesquisa de cunho lexicográfico, em especial no Brasil, ainda não privilegiam satisfatoriamente a variação.

Por admitir que não há como ignorar a nítida diversidade geográfica em relação aos níveis da língua, sendo o Brasil formado e marcado pela diversidade sociocultural justificável por seu processo histórico, questionou-se: como o léxico se distribui na perspectiva da variação dialetal – e socioletal – no estado da Bahia, em função das áreas conceituais preconizadas no Projeto ALiB?

Pretendeu-se, assim, construir um Vocabulário Dialeto, a partir da documentação das respostas a 202 questões concernentes a 14 áreas conceituais obtidas através de gravação de 92 inquéritos realizados em 22 localidades do estado da Bahia, sendo, destas, 15 pontos documentados por Nascentes e (ou) selecionadas para a elaboração do Atlas Prévio dos Falares Baianos – o APFB.

Obedecendo à metodologia adotada pelo Projeto ALiB, de cuja base se extraíram os dados avaliados na execução do trabalho ora em desenvolvimento, o vocabulário proposto, bem como prevê o DDB, contemplou itens lexicais relativos a: Acidentes geográficos; Fenômenos atmosféricos; Astros e tempo; Atividades agropastoris; Fauna; Corpo humano; Ciclos da vida; Convívio e comportamento social; Religião e crenças; Jogos e diversões infantis; Habitação; Alimentação e cozinha; Vestuário e acessórios e Vida urbana.

Assim, com o propósito de contribuir com o DDB, desenvolveu-se esta pesquisa, considerando a história e o presente e, para além do produto final – o *Vocabulário Dialeto Baiano* –, tendo por os objetivos: i) analisar as formas documentadas, sob a perspectiva dialetológica, tendo em vista o registro de

lexias; ii) documentar, em perspectiva lexicográfica, as variantes obtidas no *corpus*; iii) propor e estabelecer uma metodologia que atenda à necessidade da divulgação expressiva da diversidade linguística com base na interface entre a Lexicologia e os estudos dialetais; iv) fornecer material de suporte para o processo de ensino e aprendizagem de língua portuguesa a partir do conhecimento da realidade linguística brasileira, e na diversidade concernente à variação semântico-lexical e dos seus dialetos em particular.

Para a documentação e dicionarização do léxico no português brasileiro contemporâneo, mais especificamente na Bahia, aventaram-se hipóteses considerando: i) o caráter e o estatuto da variação lexical; ii) as áreas conceituais investigadas; iii) as características sociais específicas.

No tocante ao primeiro item, pressupôs-se encontrar problemas, já preconizados por Fernandez² (1998), quanto ao estabelecimento do conceito de variante no esteio de Tarallo³ (1986) –, de modo a não poder ser aplicado a todas as lexias documentadas, sendo necessário, pois, o estabelecimento de traços que reflitam a variação em outros níveis de observação, além – e concomitantemente – do léxico.

Considerando a diversidade das 14 áreas conceituais investigadas, pressupôs-se que algumas formas obtidas com base nas questões concernentes a *Corpo humano, Ciclos da vida, Convívio e comportamento social, Religião e crenças, Vestuário e acessórios, e Vida urbana* se caracterizem como de maior monitoramento de fala e revelem tabus linguísticos.

Admitindo o fato de que algumas formas lexicais se caracterizam como peculiares a um grupo específico – gênero, faixa etária, escolaridade, e até mesmo em relação ao meio urbano ou rural –, pensou-se ser preciso

² “El estudio de la variación léxica se enfrenta a los mismos problemas que la variación gramatical. Entre esos problemas destaca, naturalmente, el establecimiento de equivalencias entre supuestas variantes, que tiene como trasfondo, muy especialmente em este nivel léxico-semántico, la larga profía sobre la existencia o la imposibilidad teórica de la sinonímia.” (FERNANDEZ, 1998, p. 28-29)

O estudo da variação lexical enfrenta os mesmos problemas que a variação gramatical. Entre tais problemas, é claro, o estabelecimento de equivalências entre supostas variantes, que tem como base, e especialmente no nível semântico-lexical, a longa tradição da existência ou a impossibilidade teórica de sinonímia. (FERNANDEZ, 1998, p. 28-29, **tradução nossa**)

³ Segundo Tarallo (1986, p. 8), “variantes linguísticas são diversas maneiras de se dizer a mesma coisa em um mesmo contexto e com o mesmo valor de verdade [...]”.

especificar, quando necessário, se se trata de um uso restrito a um determinado grupo social.

Importa a esta pesquisa não apenas investigar o comportamento da variação lexical em dicionários de língua portuguesa no Brasil, mas propor a sua inserção e adequação para o conhecimento e divulgação da realidade plurilinguística brasileira, no tocante ao léxico.

Pretende-se, pois, com o *Vocabulário Dialetal Baiano* e, por conseguinte com o DDB, atender à urgente ampliação do conhecimento da realidade plurilinguística brasileira, sobretudo no âmbito dialetal da variação léxica. Para tanto, foi preciso imergir nas profundas águas do Léxico, a fim de fazer emergir os elementos lexicais desconhecidos e (ou) desprestigiados socialmente e, por conseguinte, submergir os preconceitos sociolinguísticos.

No tocante à estrutura, com finalidade didática e de manuseio, a versão final da tese encontra-se dividida em três volumes.

O *Volume I* constitui-se de: i) Introdução; ii) Parte I – Fundamentação Teórica; Parte II – Metodologia; Parte III – Análise dos dados, Considerações finais.

A Parte I – Fundamentação Teórica – inclui os 3 primeiros capítulos, nos quais foi possível refletir acerca do léxico, avaliando seus caracteres, estabelecer um panorama sócio-histórico e fazer uma abordagem, ainda que sucinta, dos estudos lexicológicos e lexicográficos, no Brasil.

Na Parte II, no Capítulo 4 *Por uma lexicografia variacional: “revendo razões” e conceitos*, discutem-se i) a delimitação do objeto teórico-observacional; ii) o estabelecimento de um arcabouço teórico-metodológico constituído do diálogo com outras áreas as quais se fizeram relevantes à pesquisa, visando, à consolidação dos pressupostos da Lexicografia Variacional.

A Parte III – Metodologia – constitui-se dos capítulos 5 e 6, que apresentam, respectivamente, os procedimentos metodológicos usados para o desenvolvimento do Volume I – incluindo os métodos de análise de dados – e os que foram usados para a elaboração dos volumes II e III, as versões fascicular / temática e geral do *Vocabulário Dialetal Baiano*.

Na Parte IV – Análise dos dados – exibem-se o cotejo dos resultados obtidos na quantificação dos dados e na confecção dos verbetes de cada área temática com considerações tecidas sobre a abordagem do fenômeno estudado em dicionários língua portuguesa e, em que, também, se apresentam a argumentação que constitui os resultados do *corpus*, provenientes de uma análise qualitativa, a fim de justificar o estabelecimento das macro e microestruturas do vocabulário proposto.

Por fim, as *Considerações finais*.

No Volume II, o Vocabulário temático consta de um texto pré-dicionarístico, os fascículos temáticos concernentes às 14 áreas do QSL ALiB e relações possíveis e um texto pós-dicionarístico que ressaltam a relevância do projeto para o conhecimento da realidade plurilinguística brasileira no que tange ao léxico em uso na fala utentes do português brasileiro contemporâneo.

PARTE I – FUNDAMENTAÇÃO TEÓRICA

*Pello fio tiraràs o nouello e pello passado o que está por vir*¹.
(DELICADO, 1651)

Aqui, o fio com o qual se tece a Parte I destina-se à descrição de leituras acerca do léxico destina-se à reflexão com base na descrição de leituras acerca do léxico, seu panorama sócio-histórico e os estudos, no Brasil, a ele relacionados.

No Capítulo 1, *Entre palavras*, discutem-se os caracteres nomeador, dinâmico, sociocultural e estrutural das palavras, re(i)novando a reflexão acerca da impropriedade da palavra ‘palavra’ nos estudos lexicais.

Com vistas a mostrar que já no passado, na passagem do latim para o português, a variação lexical é importante e relevante para se pensar aspectos presentes no *corpus* documentado e analisado, sobretudo no que se refere às formas, no capítulo 2, *Pelos caminhos do léxico português*, desenvolve-se um sucinto panorama da constituição histórica do léxico português.

No capítulo 3, *Estudos do(s) léxico(s) no Brasil*, apresenta-se um panorama dos estudos lexicológicos e lexicográficos no Brasil, a fim de incluir o presente trabalho no âmbito desses estudos, de forma propositiva, com base na inovadora perspectiva da Lexicografia Variacional.

¹ Este é um dos adágios portugueses, do século XVII, compilados em na obra *Adagios portuguezes reduzidos a lugares communs*, de Antônio Delicado, em 1651 .

Capítulo 1

ENTRE PALAVRAS

Entre coisas e palavras – principalmente palavras – circulamos. A maioria delas não figura nos dicionários de há trinta anos, ou figura com outras acepções. A todo momento impõe-se tomar conhecimento de novas palavras e combinações de.

Você que me lê, preste atenção. Não deixe passar nenhuma palavra ou locução atual, pelo seu ouvido, sem registrá-la. Amanhã, pode precisar dela. E cuidado ao conversar com seu avô; talvez ele não entenda o que você diz. O malote, o cassete, o spray, o fuscão, o copião, a Vemaguet, a chacrete, o linóleo, o nylon, o nycron, o diafone, a informática, a dublagem, o sinteco, o telex... Existiam em 1940?

Ponha aí o computador, os anticoncepcionais, os mísseis, a motoneta, a Velosolex, o biquíni, o módulo lunar, o antibiótico, o enfarte, a acupuntura, a biônica, o acrílico, o tá legal, o apartheid, o som pop, a arte pop, as estruturas e a infraestrutura. Não esqueça também (seria imperdoável) o Terceiro Mundo, a descapitalização, o desenvolvimento, o unissex, o bandeirinha, o mass media, o lbope, a renda per capita, a mixagem.

Só? Não. Tem seu lugar ao sol a metalinguagem, o servomecanismo, as algias, a coca-cola, o superego, a Futurologia, a homeostasia, a Adecif, a Transamazônica, a Sudene, o Incra, a Unesco, o Isop, a OEA, e a ONU. Estão reclamando, porque não citei a conotação, o conglomerado, a diagramação, o ideologema, idioleto, o ICM, a IBM, o falou, as operações triangulares, o zoom, e a guitarra elétrica.

Olhe aí na fila – quem? Embreagem, defasagem, barra tensora, vela de ignição, engarrafamento, Detran, poliéster, filhotes de bonificação, letra imobiliária, conservacionismo, carnet da girafa, poluição. Fundos de investimento, e daí? Também os incentivos fiscais. Know-how. Barbeador elétrico de noventa microrranhuras. FenoliteBaquelite, LP e compacto. Alimentos super congelados. Viagens pelo crediário, Circuito fechado de TV Rodoviária. Argh! Pow! Click!

Não havia nada disso no Jornal do tempo do Venceslau Brás, ou mesmo, de Washington Luís. Algumas coisas começam a aparecer sob Getúlio Vargas. Hoje estão ali na esquina, para consumo geral. A enumeração caótica não é uma invenção crítica de Leo Spitzer. Está aí, na vida de todos os dias. Entre palavras circulamos, vivemos, morremos, e palavras somos, finalmente, mas com que significado?

(DRUMMOND DE ANDRADE, 1988, p. 1794-1795)

Se se pudesse considerar o texto supracitado uma epígrafe, certamente, haver-se-ia de contrariar o gênero textual. Mas, não por acaso, “Entre palavras”, título dessa prosa poética de Carlos Drummond de Andrade é também título deste primeiro capítulo.

Recentemente, esse mesmo texto fora utilizado em questão de prova⁵

⁵ A referida prova concerne ao Concurso AGERBA – 2016, realizada pela Agência Estadual de Regulação de Serviços Públicos de Energia, Transportes e Comunicações da Bahia (AGERBA) para o cargo de Especialista em Regulação, tendo como requisito nível superior.

de Língua Portuguesa de um concurso público, na qual se solicitava que, após “uma leitura atenta do texto”, o leitor inferisse “quanto ao potencial nomeador das palavras”. Foram apresentadas as opções,

- a) as expressões de uso técnico não podem ser entendidas por um falante comum.
- b) as tecnologias inventadas são consideradas inovações em qualquer época.
- c) uma vez construído um sentido para a palavra, ela não pode ter outros.
- d) as coisas sempre existiram, faltava ao homem apenas nomeá-las.
- e) as classificações econômicas foram culturalmente construídas.

O gabarito final considerou correta a alternativa ‘e’.

Ora, entre (tantas) palavras e tantas inferências possíveis constantes nessa prosa poética drummondiana, é de se indignar que, ainda hoje, o texto seja utilizado como “pretexto”, um mero subterfúgio usado, exclusivamente, para análise gramatical ou compreensão textual, em processos seletivos. No entanto, por mais esdrúxulas que possam parecer as alternativas da questão supracitada, tais servem, aqui, para a reflexão acerca da dinâmica lexical, pertinente a este trabalho.

O primeiro ponto que se destaca refere-se ao “potencial nomeador das palavras”, no enunciado da questão do referido concurso. Em termos linguísticos, o que seria esse “potencial nomeador”? Haveria palavras que não possuíssem o potencial nomeador? Se não, como identificá-lo nas palavras chamadas gramaticais?

Outro ponto curioso e relevante a este trabalho concerne à possibilidade de construção de sentidos para uma palavra e aos sentidos e significados que as palavras assumem. “Uma vez construído um sentido para a palavra, ela não pode ter outros”? Se sim, como explicar que “papagaio”, por exemplo, ora possa designar uma ave, ora possa ser o nome dado a um brinquedo, e ainda possa se referir à pessoa que fala demasiadamente? Até que ponto os dicionários atuais de língua portuguesa estariam atentos às acepções que as palavras possam ter, sobretudo nos rincões dialetais em que os sentidos possam ser inusitados?

Qual a relação entre as mudanças sócio-históricas que incluem os desenvolvimentos tecnológicos e o uso das palavras? Em que consistem as diferenças entre as expressões de uso técnico e as de uso de uso geral, a ponto de que não possam ser entendidas por um falante comum? Seria possível, em algum momento, um elemento de língua especializada se tornar geral, comum a qualquer

falante? Qual a dimensão da possibilidade de trânsito entre o mais técnico e o mais geral?

Considerando o fato de que o gabarito da já referida prova considera como correta a alternativa ‘e’, questiona-se se somente “as classificações econômicas são culturalmente construídas”. Haveria, então, palavras mais marcadas culturalmente do que outras?

Entre tantas perguntas, esse preâmbulo há de nortear as reflexões que se fazem neste capítulo acerca dos caracteres nomeador e dinâmico, sociocultural e estrutural das palavras. Busca-se, também, refletir acerca da (in)pertinência da palavra ‘palavra’ para os estudos linguísticos, mais especificamente, os lexicológicos.

1.1 O CARÁTER NOMEADOR E DINÂMICO DAS PALAVRAS

As coisas precedem suas denominações e podem até existir sem que as tenham.
(BASSETO, 2001, p. 74)

Muitos outros filósofos se envolveram no debate sobre se a relação entre as palavras e as coisas é puramente convencional ou a expressão da essência das coisas. Um deles, Pedro Abelardo (1079-1142), colocou o problema nos seguintes termos: Se todas as rosas do mundo desaparecessem, o nome “rosa” ainda assim continuaria tendo significado? Por trás dessa questão se esconde a secreta relação entre as palavras e as coisas, além da teimosa recusa da linguagem em ser mero veículo de expressão dos objetos ou das idéias dos sujeitos.
(SILVA, 2008)

Não aleatoriamente nem por convenção, ambas as epígrafes foram selecionadas.

Embora a primeira⁶ suceda, cronologicamente, à segunda, esta parece responder àquela. Afinal, tudo que faz parte da vida dos seres humanos tem um nome, em alguns casos, diversos nomes, de modo que, conforme “a língua, conforme a época ou grupos sociais, conforme o contexto apresente tal ou qual objetivo específico, vê-se dominar ora uma forma ora outra, ora uma variante ora outra.” (BAKHTIN, 2006, p. 150).

⁶ A citação se refere aos pressupostos do Método de *Wörten und Sachen* (“palavras e coisas”), fundado em 1909, por Rudolf Meringer e Hugo Schuchardt, os quais, de acordo com Basseto (2001, p. 76), “rompeu os limites estreitos dos neogramáticos” – que se restringiam ao aspecto fonético – “buscando o que há de vivo e não sujeito às leis cegas, buscando estabelecer a etimologia e até a biografia das palavras, deu o devido lugar à semântica e tornou os estudos filológicos mais objetivos”.

É bem verdade que o questionamento sobre as formas de representar e de significar as coisas e a ‘realidade’ sempre se fizeram presentes na história da humanidade. Aliás, é válido ressaltar que “foi através da linguagem que o processo civilizatório se desenvolveu, pela simples transmissão de tudo o que foi aprendido de geração para geração, cumulativamente” (FERNANDES, 2011, p. 161).

No âmbito da Linguística, a discussão circunda o signo e o significante, sua arbitrariedade⁷ e sua convencionalidade. Para Langacker (1975, p.100)

As imagens conceptuais são altamente estruturadas. [...] nossa concepção de bicicleta não consiste simplesmente na associação dos conceitos roda, pedal, direção, quadro, selim, etc., mas num arranjo destes elementos numa estrutura coerente e familiar. Sem a configuração, não temos mais do que uma pilha de peças. (LANGACKER, 1975, p. 100)

Para Greimas (1966, p. 5)

Le monde humain nous paraît se définir essentiellement comme le monde de la signification. Le monde ne peut être dit ‘humain’ que dans la mesure où il signifie quelque chose (GREIMAS, 1966, p. 5)⁸.

Já segundo Biderman (1998, p.11), a “nomeação da realidade pode ser considerada como a etapa primeira no percurso científico do espírito humano de conhecimento do universo”.

Biderman fala em nomeação. Greimas em significação. Coseriu faz a distinção entre ambas ao dizer que, em princípio

seulement les rapports de signification sont structurables ; les rapports de désignation ne le sont pas. La désignation concrète (d’un objet déterminé) est un fait de ‘discours’, tandis que la signification est un fait de ‘langue’ (technique du discours). Aussi les rapports de signification sont-ils constants (du point de vue synchronique), tandis que les rapports de désignation concrète sont inconstants (variables). En outre, la désignation peut être

⁷ Não raro se relembra a “buruca” – uma criação meramente ilustrativa que um professor utiliza em sala de aula para designar “mesa”, com o intuito de demonstrar a arbitrariedade do signo linguístico. Aliás, essa arbitrariedade é tão real que o mesmo professor descobriu, um dia, que um jornalista, também, ator e produtor que havia falecido se chamava “Buruca” – o jornalista Luiz Carlos Pulchério de Medeiros.

⁸ O mundo humano parece ser definido essencialmente como o mundo do significado. O mundo pode ser chamado de "humano" apenas na medida em que significa algo (GREIMAS, 1966, p. 5, tradução nossa)

métaphorique, tandis que la signification ne l'est pas, du point du vue synchronique et distinctif [...] (COSERIU apud GECKELER, 1971, p. 95)⁹

Diante disso, seria possível dizer que a significação é conceitual e a denominação, material.

Nesse sentido, parece ser mais coerente atribuir às palavras, em geral, como princípio, o caráter de representação e de significação. O potencial nomeador seria, pois, paramétrico, estipulado por um *continuum* entre 'mais nomeador' e 'menos nomeador'.

Haveria, então, palavras que não possuem o tal potencial nomeador, *i.e.*, palavras cujo traço de nomeação seria não marcado?

Parece que sim. A estas, a tradição gramatical tem chamado de palavras gramaticais, em oposição às palavras lexicais.

Entende-se por palavras gramaticais aquelas que, no plano estrutural, no plano sintático, se apresentam de forma subsidiária em oposição a outra categoria de palavras, as lexicais, as quais, por si só, podem constituir núcleos. Segundo Tesnière (1976, p. 56),

les mots constitutifs sont comme les pierres de construction de la phrase, tandis que les mots subsidiaires ne sont que le ciment qui sert à assurer plus de cohésion à leur agencement (TESNIÈRE, 1976, p. 56)¹⁰.

No plano semântico, as palavras constitutivas, lexicais, são consideradas "cheias", *i.e.*, palavras que estão marcadas pela sua carga e função semântica e representam os tijolos da linguagem. Já as gramaticais, subsidiárias, são consideradas "vazias", servindo apenas como cimento articulador do discurso, "dont le rôle est uniquement d'indiquer, de préciser ou de transformer la catégorie des mots pleins et de régler leurs rapports entre eux" (TESNIÈRE, 1976, p. 53).

⁹ Apenas as relações de significado são estruturáveis; as relações de designação não são. A designação concreta (de um determinado objeto) é um fato de "discurso", enquanto o significado é um fato da "linguagem". Assim, as relações de significado são constantes (do ponto de vista sincrônico), enquanto as relações de designação concreta são inconstantes (variáveis). Além disso, a designação pode ser metafórica, enquanto o significado, do ponto de vista sincrônico e distintivo, não é [...] (COSERIU apud GECKELER, 1971, p. 95, tradução nossa)

¹⁰ [...] as palavras constitutivas são como as pedras da construção da frase, enquanto as palavras subsidiárias são apenas o cimento que serve para garantir mais coesão (TESNIÈRE, 1976, p. 56, tradução nossa)

As preposições, por exemplo, são frequentemente classificadas como “palavras vazias”, estando seu significado atrelado às palavras “cheias” (substantivos, verbos, adjetivos e alguns advérbios), conforme afirma a tradição gramatical. No entanto, estudos mais recentes, como os realizados pela Ecolinguística, afirmam que as preposições não são desprovidas de significado. E nem poderiam ser, pois, ao fim ao cabo são morfemas.

Sob essa ótica, é válido relembrar, aqui, que, o sistema preposicional do português, ao menos em suas linhas gerais, proveio do sistema latino. Como língua eminentemente flexional, o latim apresentava uma estrutura sintética, cuja subordinação / ordenação sintática, era realizada através de seis casos: nominativo, genitivo, dativo, acusativo, ablativo e vocativo, os quais eram reforçados pelas preposições latinas. Talvez seja assim que o

falante de português sabe, por intuição, que a língua seleciona preposições específicas para funções determinadas. E, muitas vezes, estruturas sintagmáticas diferentes no seu conteúdo semântico apresentam uma tal identidade, no nível de superfície, que só a diacronia pode explicar a diferença. Em *veio de Roma, anel de ouro, casa de Pedro*, há um valor sintático-semântico que emana do todo sintagmático e que é, em cada caso, diferente, a despeito da mesma preposição *de*. É que, no fundo, há um ablativo de origem, um genitivo ou ablativo de matéria e um genitivo de posse. (GUIMARÃES, 1987, p. 68)

Tomando-se, por exemplo, a preposição "de", é possível notar que expressa significados e que contribui para a construção do significado¹¹. Em português, atualmente, Farenzena e Dalpian (2008) apresentam 17 usos distintos do "de", corroborando o que o senhor verbo Ser diz a Emília, em seu passeio pela casa das preposições, no País da Gramática:

Creio que a preposição DE é a mais importante. [...] Num concurso de utilidade, DE venceria. É como ali adiante a Conjunção E, que é a menor de todas. Tão econômica que até se escreve com uma letra só – e no entanto é uma danadinha de útil. (LOBATO, 2008, p. 68)

Mas, como nem só de ‘de’ sobrevive a língua,

¹¹ Em uma das aulas de Léxico, fora lido o texto de Drummond, usado como epígrafe deste capítulo. Em determinado momento, mais especificamente, ao final da frase “A todo momento impõe-se tomar conhecimento de novas palavras e combinações de.”, houve uma quebra na linearidade da leitura, que estava sendo feita por uma aluna de Letras, também graduada em Jornalismo. Tal “quebra” deveu-se ao seu estranhamento em relação a esse uso inusitado de *de*, feito pelo poeta. A preposição ‘de’ parecia pressupor complemento e este em elipse, tornou-se vago, nulo, para quem lia. No entanto, vale não esquecer que “um copo vazio está cheio de ar” (GIL, 1974)

[...] bem consideradas as coisas, não existe palavra que não seja indispensável. Sem os Nomes, de que valeríamos nós, Verbos? E sem Verbos, de que valeriam os Nomes? Todas as palavras ajudam-se umas às outras, e desse modo os homens conseguem exprimir todas as idéias que lhe passam pela cabeça. (LOBATO, 2008, p. 67-68)

Embora nem todas as palavras nomeiem, certamente, todas significam.

O “caráter nomeador” parece não ser identificável nas palavras chamadas gramaticais. No entanto, estas representam e significam, se e somente se, e quando manifestos na sintaxe, ou seja, na relação com as palavras precedentes e subsequentes.

Nas chamadas “palavras vazias” ou “palavras gramaticais”, a significação é manifesta intralinguisticamente, ou seja, na relação e combinação com outras palavras. Nas outras, as “palavras cheias”, também chamadas de lexicais, a significação ocorre extralinguisticamente, cujo referente conceitual e ou nocional existe na realidade.

A título de exemplificação, observe-se o diálogo, a seguir, extraído do *corpus*.

INQ. – Como é que você se referia às patas da frente do cavalo?

INF. – É porque eu nunca tive necessidade de... falar. É sério...

INQ. – Falar disso. É.

INF. – É porque tem coisas que a gente num... num... num fala e que num sabe.

(Inq. 093/06 (Salvador – BA) / Inf.: mulher, faixa etária 1, universitária)

É consabido, também, que nem sempre cabe às palavras – ao léxico – o papel denominativo na língua já que outras estratégias podem emanar de diferentes níveis no processo de construção do significado.

E o que dizer da ‘rebimboça da parafuseta’?, quando, embora o referente, a famigerada peça, não exista, o léxico, a designação, existe.

Dessa forma, considerando o fato de que até mesmo as palavras consideradas “cheias” podem não nomear e que, dentro de dada cultura, podem existir referentes que não são nomeados, talvez, seja mais coerente atribuir às palavras, em geral, como princípio, o caráter de representação e de significação.

Valendo-se de metalinguagem, (em) entre palavras, além de expor o caráter nomeador das palavras, Drummond denuncia o caráter efêmero, sugerindo que

exista uma certa dificuldade de apreender e acompanhar a constante (re)inovação do léxico.

Segundo Teles (1987, p. 20), “no seu processo criador, a língua se movimenta dentro de coordenadas geométricas e temporais, lançando o passado (arcaísmos) em face, do futuro (neologismo)”.

Recentemente, veicularam-se, em redes sociais¹², imagens que remontam a objetos que fizeram parte da vida e do senso-comum de uma geração e que foram substituídas por outros elementos, ao longo dos anos. Além das recordações das experiências de uma época, as imagens chamam a atenção pelo discurso incutido, por vezes, com equívocos conceituais e certa carga de estigma social. Observem-se, a seguir, alguns exemplos:

¹² Foi criada uma página no Facebook.



FIGURA 1 - Imagens do Facebook

Fonte: Adaptação de imagens do Facebook, elaborada pela autora.

Equivocadamente, nas imagens 3 e 4 da *Figura 1*, os objetos estão associados ao fator temporal, como sugere a frase “você está velho!”. No entanto, parecem refletir experiências de âmbito espacial – com traço mais rural.

A imagem 1, da mesma figura – com a inserção da formulação da pergunta – parece generalizar o fato que relaciona a caneta à “fita cassete” – para facilitar a “rebobinação⁹” quando se pretendia voltar a um ponto do áudio, de maneira a atestar a sua especificidade a um momento da história, com o agravante do desuso, provocado pela extinção do objeto.

O mesmo ocorre na imagem 2 da *Figura 1*, cuja frase atrai. Em “**Sou da época em que entregar fita sem rebobinar dava muita!**” (grifo nosso), além

⁹ A rebobinação manual da fita cassete era feita, mais facilmente, com a caneta esferográfica *Bic*, que a época era hexagonal e prendia-se aos ganchos constantes nos orifícios da fita cassete. Ao girar a caneta, fazendo movimentos circulares, a fita voltava ao ponto desejado, ou ao início.

de assumir-se como pertencente a um grupo de momento específico – “sou da época”, o autor faz menção a fatos que dificultam ou até mesmo impossibilitam a comunicação, já que com a extinção das videolocadoras, provenientes do advento e dos avanços das tecnologias, “entregar fita sem rebobinar”, tornou-se uma prática incompreensível às comunidades atuais, em geral.

Em mesma perspectiva, o artigo “Rebobinar, cair a ficha e virar o disco: expressões que ficaram no tempo”¹⁰, de Milton Jung (2014), traz a tona, de forma bastante curiosa, expressões que entraram em desuso, mostrando possíveis razões para que isso ocorresse, de modo a crer-se que as expressões “cair a ficha”, “virar o disco” e “rebobinar a fita” somente são compreendidas, atualmente, por quem conheceu os objetos e as ações a que aludiram ou de que se metaforizaram.

Empregava-se a expressão “caiu a ficha” quando alguém que demorasse a compreender algo, só depois de algum tempo, compreendesse ou percebesse algo de que se estivesse falando. A expressão provém do uso de fichas em telefones públicos, as quais “caíam” quando a chamada ou ligação era completada. Com o passar dos tempos e advento de tecnologias, as fichas foram trocadas por cartões, os “orelhões” passaram a conviver com a lexia telefone público, cujo referente tem perdido função e espaço, haja vista as aplicações e os aplicativos dos celulares – telefones móveis e *smartphones* – se multiplicarem, a cada dia.

Consequentemente, a expressão que tenderia, em função original, ao desaparecimento, pode ou não sobreviver na língua, pela função metafórica que assuma durante sua existência na língua.

Aliás, é válido ressaltar que o acesso a esses meios de comunicação decorre de mudanças socioeconômicas, socioculturais e tecnológicas.

A expressão “virar o disco”, usada como advertência a uma repetição, remonta ao vinil cujas faixas musicais se encontram em dois lados, diferente do CD. Segundo Jung (2014), como ouvir disco tem ganhado certo prestígio social, ultimamente, a expressão pode perdurar.

Em contrapartida, a expressão “rebobinar a fita” que se tornou sinônimo de refletir sobre algum acontecimento e é resultante de um processo

¹⁰ O texto é disponível em uma rede social do autor.

metafórico, não sobreviverá à extinção das locadoras, do videocassete e das fitas de vídeo de rolo.

Como se nota, em razão de mudanças socioculturais, expressões contemporâneas “virar o disco” e “rebobinar a fita” poderão, a depender de diversos fatores, ter (des)continuidades diferentes.

Esse fato instiga o pensamento de que gerações futuras e já as hodiernas em algumas comunidades poderão não compreender o sentido de lexias que entrarem em desuso com o passar do tempo e incita indagações como: O que condiciona a vitalidade de algumas lexias e condena outras? Quantas lexias podem já ter sido ou podem ser perdidas por falta de documentação, sobretudo quando o referente deixa de ser comum?

Esses questionamentos tornam propícia a reflexão acerca da dinâmica lexical, mais especificamente, no que tange à linha tênue que une ou separa o arcaísmo do neologismo; entre as mudanças linguísticas e as mudanças semânticas.

1.1.1 No tempo: *continuum* entre arcaísmos e neologismos

Como proscriver em absoluto o ARCAÍSMO, quando se recebem, aconselham e festejam os NEOLOGISMOS? [...]

Tôdas as gerações assistem ao reabrir de palavras antiquadas, que outra vez, ao influxo de novos tempos, rebentam de seu, espontâneas e belas [...] Com os ARCAÍSMOS a lei é a mesma que a respeito dos NEOLOGISMOS: usem-se “discretamente, quando necessários, ou úteis”. (BARBOSA *apud* CARDOSO E CUNHA, 1987, p. 198)

A proscrição de que trata Rui Barbosa em um de seus discursos acerca do padrão linguístico jurídico faz pensar na dificuldade metodológica em asseverar / atestar a falência e desaparecimento de uma ‘palavra’.

Não raro é possível ouvir: “faz tempo que não ouço essa palavra”. Aconteceu com a expressão *cutruvia* usada em uma obra da dramaturgia televisiva brasileira que focalizava os usos e costumes do Nordeste, mais especificamente de comunidades ribeirinhas. Tão logo ouvida, os adeptos de redes sociais e de novelas se movimentaram em comentários acerca desse uso. Desses, um, por exemplo, diz: “Cutruvia, nossa faz tempo que não ouço

essa palavra”. Dessa forma, o que para tal seria algo, possivelmente, em desuso, para outros tantos – por exemplo, a autora – seria um neologismo.

Esse fato, embora simplório, evidencia a relativização desses termos, que parecem ser estanques e estancados.

Em *O léxico furtado do passado, na história do futuro, de Antônio Vieira*, de Machado Filho e Oliveira (ora, no prelo) apresenta quatro motivos: i) a não dicionarização de dado item lexical; ii) a substituição deste por outro, embora esteja registrada lexicograficamente; iii) a não existência / desaparecimento do referente e iv) a mudança de sentido do item lexical.

1.2 O CARÁTER ESTRUTURAL DAS PALAVRAS

1.2.1 O problema da palavra ‘palavra’: revendo a discussão que deu muito “pano pra manga”

Teixeira (1991: 1316-1317) resume muito bem o problema: a palavra não é uma unidade fonética porque as línguas não realizam as suas construções em palavras separadas, mas em conjuntos de palavra que obedecem apenas a um acento principal: [unstrabálham / mazotrusnãw], para além de que duas unidades diferentes podem ser a mesma palavra (novo/nova) e duas unidades foneticamente iguais podem representar palavras diferentes [canto (verbo) / canto (subst.)]. Também não é uma unidade morfo-sintáctica uma vez que a sua função pode ser desempenhada por duas ou mais palavras: assim, em Penso, logo existo, a mesma função morfo-sintáctica da palavra logo pode ser executada por duas palavras (Penso, por conseguinte existo) ou por um sintagma (Penso, portanto isso é sinal de que existo). Por último, escreve Teixeira, «as palavras são usualmente portadoras de sentidos diversificados (polissemia) que actualizam conforme os contextos sintácticos e pragmáticos em que se inserem.» (idem, 1317), pelo que também não pode ser considerada como uma unidade semântica.

(SANROMÁN, 2001, p. 134, nota de rodapé)

Quanto ao primeiro texto epigrafado – Entre palavras –, destacam-se os tipos de palavras que o compõem. São ‘palavras’ simples, compostas, siglas, estrangeirismos, tecnoletos, gírias, etc.). Esse acervo suscita uma primeira indagação que concerne ao “famigerado” conceito de palavra.

Facilmente reconhecida por falantes em sua língua, o item lexical ‘palavra’ e seu conceito se configuram como assunto de uma das discussões.

A proposta de Basílio (1987), em *Teoria lexical*, embora não esclareça o conceito de palavra, viabiliza uma oportuna reflexão ao apresentar problemas nos eixos e pressupostos, destacam-se o eixo da pausa e o eixo da distinção entre as palavras.

Com base no eixo da pausa, o conceito de palavra como “qualquer sequência que ocorra entre espaços e/ou sinais de pontuação”, restringe-se à escrita. O eixo da distinção entre palavras pelos processos de flexão e derivação também apresenta dificuldades de entendimento, como se observa no seguinte fragmento do texto:

Por ora, podemos prosseguir com o conceito intuitivo que temos de palavra, que nos faz concordar que par/pares são duas formas, singular e plural, da mesma palavra, enquanto par/ímpar são duas palavras distintas. É bom lembrar, no entanto, que há casos em que, mesmo intuitivamente, as coisas não são tão claras. Vejam, por exemplo, o caso do particípio passado. Devemos considerar, digamos, perdido como uma forma do verbo perder ou como uma outra palavra? Essa equivale à questão de se devemos considerar o particípio passado como um caso de flexão ou um caso de derivação. (BASÍLIO, 1987, p. 12)

Em suma, nota-se que em ambos os eixos, o conceito de palavra constituiu um problema para gramáticos e linguistas.

Considerando, pois, “a impropriedade terminológica da palavra ‘palavra’” para os estudos do léxico, propõe-se, aqui, retomar questionamentos com vistas a uma reflexão que não se restrinja à nomenclatura empregada mas ao objeto do léxico.

Partindo do conceito de Léxico como “um sistema multidinâmico em que elementos menores e maiores do que se reconhece como “palavra” são o seu objeto teórico”, conforme Machado Filho (2014); considerando que cada nível de análise tem uma unidade de observação, um objeto teórico – o fone, da Fonética; o fonema, da Fonologia; o morfema, da Morfologia; a sentença, da Sintaxe; o discurso, do Discurso – e, que interfaces entre tais unidades são passíveis de realização, tem sido adotado o termo *lexia* como unidade do Léxico.

A *lexia*, termo e conceito cunhados por Pottier (1974), refere-se à unidade lexical memorizada, pertencente a uma categoria (forma do significado) ou a classes superiores. Segundo sua proposta, há quatro tipos de

lexias: i) simples, quando corresponde à “palavra tradicional”; ii) composta, se for resultante da integração semântica que se manifesta formalmente; iii) complexa, quando se constitui em uma sequência em vias de lexicalização e possui graus diferenciados e; iv) textual, quando alcançam o nível de enunciado ou de um texto, a exemplo dos hinos, preces, charadas, provérbios etc.

Ao considerar que uma lexia pode ser constituída por um lexema ou por um gramema ou, inclusive, por n lexemas e n gramemas, a proposta de Pottier (1974) parece superar a proposta de Martinet (1974) por abranger elementos que não são apenas lexicais ou gramaticais.

Quanto à lexicalidade, pois, em consonância a Sterkenburg (1983, p. 80), é possível considerar uma unidade lexical e, por conseguinte, lexicográfica como a lexicalização numa língua de um conceito ou de um objeto extralinguístico. Essa lexicalidade é garantida pelo processo de formação de palavras – derivacional, composicional e fraseológico.

Desta maneira, numa tentativa de atualização da taxonomia, são aqui, consideradas unidades lexicográficas tanto unidades monolexicais (lexias simples, segundo ótica de Pottier) como as combinações de unidades lexicais, aqui chamadas de unidades polilexicais (lexias compostas e complexas).

1.2.1.1 Unidades monolexicais

Sob a perspectiva da forma do léxico, são consideradas unidades monolexicais as lexias constituídas por morfemas flexionais ou formadores (derivacionais ou composicionais) subsequentes à base (raiz ou radical).

Na tradição lexicográfica, as unidades lexicais constituídas por derivação afixal – que sejam por acréscimo ou por redução, de que a derivação regressiva é exemplo, à base lexical – são, comumente, inventariadas na macroestrutura dos dicionários, outras, porém, têm sido alvo de exclusão.

Diferente da derivação, o processo de composição pode formar unidades monolexicais e polilexicais.

Apenas se enquadram no paradigma das unidades monolexicais as lexias compostas por meio da aglutinação – quando a estrutura fonológica é

modificada e, por conseguinte, há perda de segmentos gráficos, como ocorre em *aguardente*.

Inversamente, a justaposição – quando os elementos integrantes não sofrem mudanças fonológicas, como ocorre em *beija-flor*, *passatempo*, *plano-piloto* – se insere nos processos formativos de unidades polilexicais, sobre os quais se passa a falar, na sequência.

1.2.1.2 Unidades polilexicais

Uma característica que se invoca na delimitação de unidades polilexicais (colocações, compostos, expressões idiomáticas) e os sintagmas livres é fixação. Segundo Santos (2009, p. 40), a fixação é um

conceito escalar, pois o grau de fixação está relacionado com o número de elementos fixos de uma expressão. Quanto maior for o número de elementos fixos de uma expressão, mais elevado será o grau de fixação. Têm-se, assim, expressões “totalmente fixas”, “meramente fixas”, “semi-fixas” etc. (SANTOS, 2009, p. 40).

Para o reconhecimento das lexias e identificação dos diferentes graus de fixação, Pottier (1974) propõe alguns testes, dentre os quais se destacam o da comutabilidade, *i.e.*, de substituição de um dos elementos da unidade e o da inserção, *i.e.*, a inclusão de um elemento novo na unidade.

Como exemplificação, toma-se, aqui, a expressão idiomática *Maria vai com as outras*.

No que tange à comutação, nota-se que, mesmo podendo ser entendida como uma sentença – objeto da sintaxe –, caso seja interpretada como ‘alguém que se chama Maria se locomove para a algum lugar, com outras pessoas do mesmo gênero’, há a possibilidade de substituição por um elemento menor da unidade do léxico, por exemplo, *influenciável*.

O teste da inserção mostra que essa forma encontra-se cristalizada, ou seja, com alto grau de fixação na língua, de modo que a possibilidade de inclusão de outro elemento e manutenção do sentido é nula.

A idiomaticidade também pode ser vista como um fenômeno escalar, pois se observa uma variação que vai do mais opaco ao quase transparente. Vieira (2009) admite que a observação da opacidade de uma unidade lexical

pode contribuir para distinguir os compostos da colocação, haja vista o fato de que quanto mais opacos sintática e semanticamente se revelem ser os compostos, mais prototípicos são. As colocações, diferentemente, quanto menos opacas sintática e semanticamente, mais prototípicas são.

Para Mejri (2008, p.201),

Les collocations ne sont pas des séquences figées: ce sont des combinaisons libres qui connaissent une attraction lexicale qui favorise dans le discours l'emploi d'unités appropriées les unes aux autres ; les séquences complètement figées refusent toute variation de quelque nature qu'elle soit. (MEJRI, 2008, p. 201)

Se se estabelecido um *continuum* do 'mais livre' ao 'mais restrito', as unidades polilexicais por fraseologia ou, mais sinteticamente, unidades fraseológicas, as unidades constituídas de verbo suporte, as colocações e as sentenças fixas, as quais, segundo Mejri (2008) podem ser variáveis.

Em síntese, as unidades lexicais poder-se-iam ser classificadas da seguinte maneira:

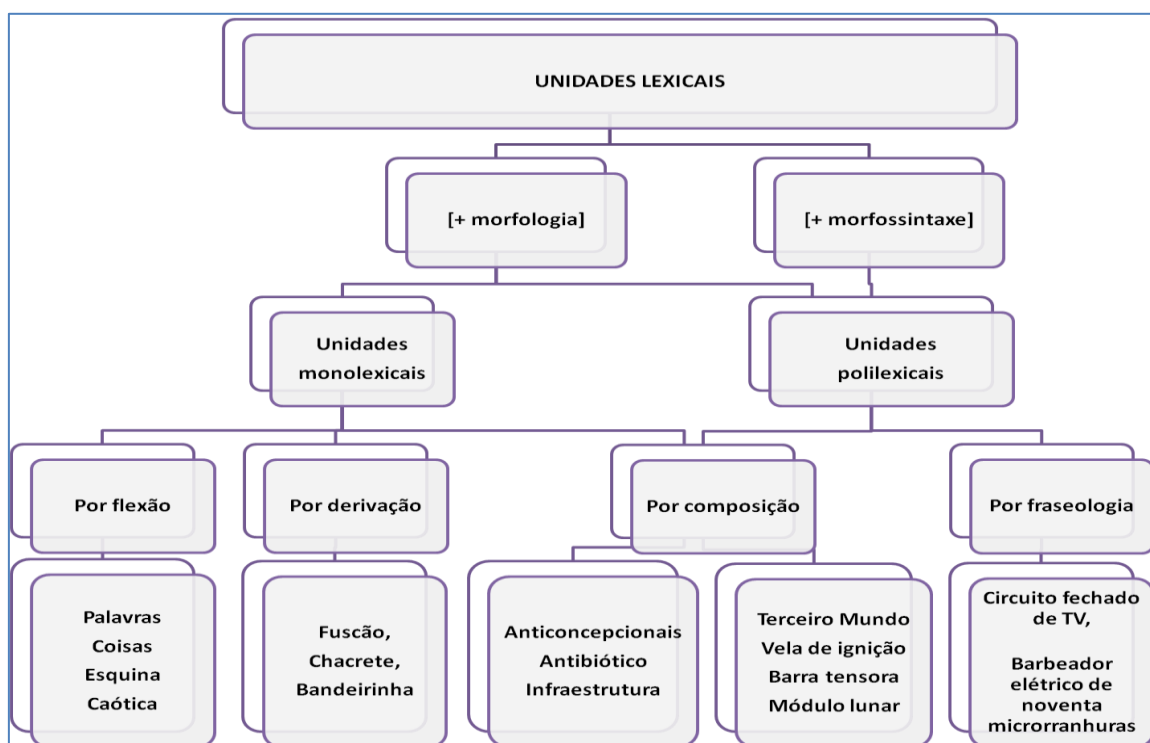


FIGURA 2 – Distribuição das unidades lexicais
Fonte: Esquema elaborado pela autora desta tese (2017)

É, pois, tendo como objeto de análise uma vasta gama de tipos distintos de lexias – as ditas palavras simples, as construídas por processos de derivação e (ou) de composição, os fraseologismos – frasesmas ou culturemas –, dentre outras – que este trabalho se desenvolve.

Capítulo 2

PELOS CAMINHOS DO LÉXICO PORTUGUÊS

Você não sabe o quanto eu caminhei
Pra chegar até aqui
(FARIAS; LAZÃO; GARRIDO, 1998)

Era um caminho que de tão velho, minha filha,
já nem mais sabia aonde ia...
Era um caminho
velhinho,
perdido...
Não havia traços
de passos no dia
em que por acaso o descobri:
pedras e urzes iam cobrindo tudo.
O caminho agonizava, morria
sozinho...
Eu vi...
Porque são os passos que fazem os caminhos!
(QUINTANA, 1989)

Estudar (o léxico de) uma língua, através de seus registros, é como reabrir um caminho¹², uma “estrada”, uma “picada”, uma “trilha”, uma “variante”, uma “vereda” ou mesmo possivelmente uma “vareda”.

Sem registro, sem estudos e sem ferramentas adequadas que viabilizem a passagem, “pedras” e “urzes” podem sufocar a língua, sobretudo quando já não há mais quem trafegue por ela, quem a use. É o que ocorre com as línguas ágrafas. Foi o que aconteceu com muitas línguas indígenas que existiam no Brasil, antes da “descoberta”.

Estima-se que, no Brasil de 517 anos passados, mais de 1.500 línguas indígenas eram faladas. Extintas, de forma brusca e gradativa, o país conta, atualmente, com cerca de 200, *i.e.*, apenas 13% subexistem. Fundado em 2013, o *Instituto Aryon Rodrigues* é um dos principais centros de pesquisas linguísticas, as quais têm sido realizadas na tentativa de preservar esse patrimônio linguístico e cultural do Brasil.

¹² Todos os itens lexicais seguintes foram extraídos do *corpus* analisado, mais especificamente, no que concerne às respostas obtidas com a formulação da questão do QSL 062 ALiB: *O que se abre com o facão, foice para passar por um mato fechado?*

Em contrapartida, as línguas não ágrafas, também denominadas “línguas de cultura”, são mais facilmente estudadas. E esse fato reforça a relevância dos registros linguísticos.

Graças aos registros em língua portuguesa, dos quais se tem notícia, e dos estudos linguísticos, mais especificamente, os que se inserem no ramo da Linguística Histórica, é consabido que, ao longo dos tempos, os caminhos do léxico português foram trilhados e (re)feitos, passo a passo, com pés de muitas gentes, que além de imprimir suas marcas, ditaram a direção ou direções que teriam levado a língua ao seu estado atual, reafirmando, pois, que “os homens fazem a língua, e não a língua os homens”, como prenunciava Fernão de Oliveira (2000, cap. IV)

Mas, a língua é mais que um caminho. É uma prática social que reflete a cultura de um povo. É História. E como toda História, fundamentada em fatos encenados e contracenados por personagens que atuam como protagonistas e (ou) coadjuvantes.

Assim como a reconstrução de uma narrativa, por vezes, pode ser evidenciada ao longo da obra por indicações esparsas dos anos e pelo desvendamento das trajetórias (des)conexas das personagens, no tocante à língua, é possível admitir que, ao escutá-la e (ou) escrutá-la, comportamentos inusitados possam ser identificados.

Curioso pode ser, então, observar o comportamento de dois itens lexicais, relacionados ao teatro: *coxia* e *bastidores*, os quais também são diretamente associadas a um pano de fundo de uma realidade aparente. Ao serem levantadas as suas acepções, ter-se-á o seguinte:

coxia

TEAT nos teatros, corredor em torno do palco, não visível ao público; bastidores. ETIM *it. corsia* 'corrente de água, corredor us. como passagem ao longo de espaços laterais ocupados' (HOUAISS, 2009)

bastidores

TEAT corredores que contornam a cena, no palco, fora das vistas dos espectadores; caixa do palco, coxias. ETIM rad. do part. *bastido* (v. bastir 'bastear') + *-or* (HOUAISS, 2009)

Considerando, por exemplo, o étimo italiano *corsia*, de *coxia*, podem ser inferidas questões intra e extralinguísticas: i) alteração fônica que resulta em mudança metaplásmica; ii) extensão de sentido, de modo que o ‘corredor us. como

passagem ao longo de espaços laterais ocupados’ ganha sentido do ‘corredor em torno do palco’; iii) algum tipo de situação de contato linguístico, este motivado por um contexto sócio-histórico específico, configurando-se “neologia semântica”, conforme Guilbert (1973).

Como se pode observar, há remissão recíproca tanto para *coxia* quanto para *bastidores*. Assim, no tocante às relações semânticas possíveis, considerando a sinonímia, poder-se-ia prever que um dos itens lexicais seja mais ou menos usado em dada comunidade, em função do prestígio de um em detrimento a outro e prever, ainda, indicadores diastráticos, diatópicos, diafásicos para isso.

Mas, não só. Até que ponto teriam sido efetivamente sinônimos? Não teria havido traços sêmicos próprios a cada um desses itens que os distanciam em algum grau? Se houve, como abrir essa cortina?

Sob a perspectiva de Genette (1984, p. 33) de que estudar a sequência temporal de uma narrativa é

confrontar a ordem de disposição dos acontecimentos ou segmentos temporais na história, na medida em que é indicada explicitamente pela própria narrativa ou pode ser inferida deste ou daquele indício indireto (GENETTE, 1984, p.33)

e com vistas a desenvolver um sucinto panorama da constituição histórica do léxico português, estabelece-se, aqui, um trânsito, não linear, passado-presente.

No esteio da ideia de Tarallo (1990, p. 23), de que

em qualquer fase da história da língua que investiguemos, iremos encontrar formas residuais do passado mais remoto contracenando com formas inovadoras de um futuro. (TARALLO, 1990, p. 23)

acredita-se que o meio – o presente – seja um lugar estratégico da estrada, propício para enxergar com nitidez as pegadas e passos da história do léxico português, interagindo sobre ele, de maneira a vislumbrar no horizonte próximo e pretérito a realidade sociolinguística atual.

Esquadrinhá-la-emos, pois!

2.1 ANDANÇAS: PASSOS QUE FIZERAM, FAZEM E RECRIAM O CAMINHO

Se se voltar o olhar para os palcos da história, há de se observar que os resultados dos contatos linguístico-culturais estiveram diretamente associados ao período de permanência e poder que detiveram aqueles que impuseram relações de domínio e de dependência. E é assim que todo esse cenário se reflete no léxico, em que

se gravam – e não raro, pirogravam – as designações que rotulam as mudanças encadeadoras dos caminhos e dos descaminhos da humanidade, além de comporem o cenário da revelação tanto da realidade quanto dos fatos culturais que permearam sua história.

(BARCELOS, 2000, p. 142)

Isso pode ser observado em diversos momentos da história, os quais já foram descritos por vários autores em perspectivas distintas. Dentre tais, Furlan (2006) afirma que a transformação do latim em línguas ibéricas, mais especificamente, no tocante ao português, ocorreu a partir de três fatos cruciais: i) a conquista e a perda romana das regiões conquistadas na Península Ibérica; ii) o domínio árabe e consequente luta contra o Islã por parte dos cristãos que buscavam a Reconquista de suas terras; iii) a formação do reino de Portugal, pautada no mito do Quinto Império e na estratégia de crescimento fundamentada na expansão marítima.

Para Galves (2007),

Uma boa síntese da questão da periodização do português europeu, tal como ela aparece em obras clássicas sobre o assunto, pode ser achada em Mattos e Silva (1991, 1994). Para todos os autores considerados, o século 16 representa um marco incontornável na história da língua. Porém, há divergências sobre a sub-divisão, e a denominação, dos períodos que respectivamente precedem e seguem esse marco. Alguns autores, como Leite de Vasconcelos, Lima Coutinho, Matoso Câmara rotulam simplesmente a fase da língua que vai dos primeiros documentos remanescentes ao século 16 de português “arcaico” ou “antigo”. Outros subdividem esse período em duas fases, com uma linha divisória se situando entre a metade e o fim do século 14 (1350, fim da produção da lírica galegoportuguesa, e 1385, avento da Dinastia de Avis). Antes, temos o “português trovadoresco” (Silva Neto, Haug, Carolina Michaelis de Vasconcelos), o “galego-português” (Pilar Vasquez Cuesta, Paul Teyssier), ou ainda simplesmente o “português antigo” (Lindley Cintra, Ivo Castro). As duas primeiras denominações fazem referência à língua dos trovadores, remanescente da época em que o português e o galego eram um só idioma. A terceira, de maneira interessante para a discussão que seguirá, reduz a

essa sub-divisão a noção de antigo, usada por outros para um período bem maior. Para a segunda subdivisão, que abrange o período compreendido entre a segunda metade do século 14 e meados do século 16, as denominações variam mais. Os autores que tomavam a lírica trovadoresca como definitiva do período anterior, caracterizam este como da “prosa nacional”. Todos os outros parecem tratá-lo como um período essencialmente de transição. Lindley Cintra e Ivo Castro o chamam de “português médio”. Pilar Vasquez Cuesta vai mais longe, ao definir esse período em referência ao seguinte: “português pré-clássico”. É também o que faz Paul Teyssier, quando caracteriza o período que vai do fim da lírica galego-portuguesa até os Lusíadas como de “formação do português clássico”. Nesses termos, não se trata do fim de uma época mas da gestação de uma outra.

Castro (1988, p. 12) sumariza a questão da periodização com o quadro, ora reproduzido.

ÉPOCA	Leite de Vasconcelos	S. Silva Neto	Pilar Vázquez Cuesta	Lindley Cintra
até s. IX (882)	pré-histórico	pré-histórico	pré-literário	pré-literário
até +/- 1200 (1214-1216)	proto-histórico	proto-histórico		
até 1385 / 1420	português arcaico	trovadoresco	gal.-português	port. antigo
até 1536 / 1550		port. comum	port. pré-class.	port. médio
até s. XVIII	português moderno	português moderno	port. clássico	port. clássico
até s. XIX / XX			port. moderno	port. moderno

QUADRO 1 - Propostas de periodização para a história da Língua Portuguesa
Fonte: Castro (1988), com adaptações.

Com novo viés, Galves (2007) e Galves, Namiuti e Paixão de Sousa (2005) propõem uma periodização pautada da competição de gramáticas, esquematizando-a de maneira a contrapor à visão tradicionalista, de modo que “a linha superior corresponde às periodizações tradicionais, e a linha inferior à periodização” defendida pela autora, na qual “as zonas em cinza correspondem aos períodos de competição de gramáticas”, como mostra a Figura, a seguir:

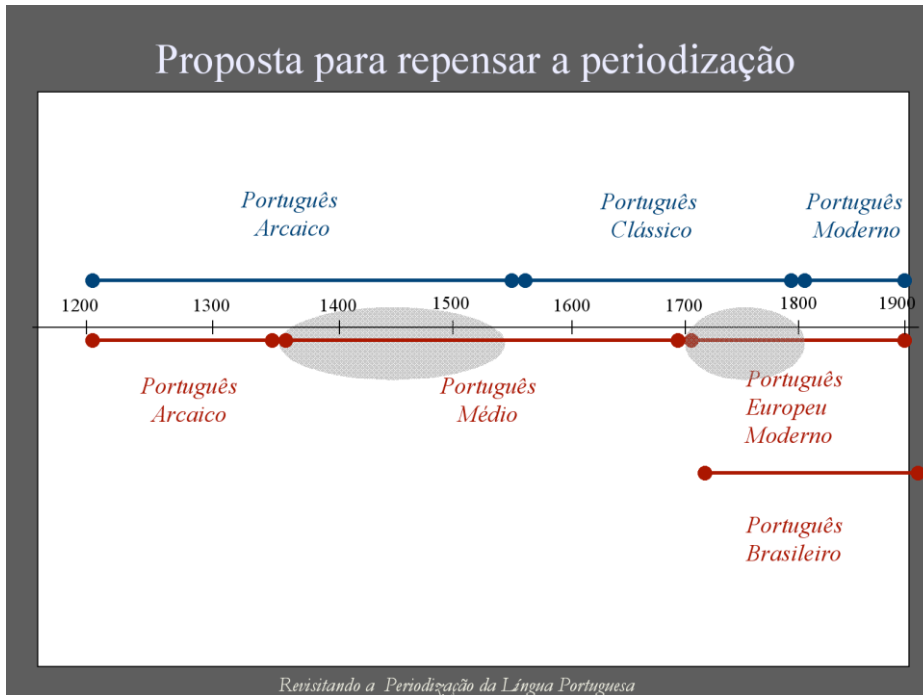


FIGURA 3 - A periodização das gramáticas portuguesas
 Extraído de: Galves, Namiuti e Paixão de Sousa (2005)

A proposta se revela interessante por inserir o português brasileiro. Porém, como se observa, há várias periodizações linguísticas na história do português quase todas centradas em questões de ordem gramatical.

No que tange à formação do português brasileiro, existem, atualmente, várias propostas de periodização. Dessas, podem-se citar cinco – (SILVA NETO, 1986 [1950]); (TEYSSIER, 1997); (PESSOA, 1997); (LOBO, 2003) e (LUCCHESI, 2017, no prelo) –, as quais levam em consideração fatos sócio-históricos relevantes como mostra o Quadro 2.

	SILVA NETO (1986 [1950])	TEYSSIER (1997)	PESSOA (1997)	LOBO (2003)	LUCCHESI (2017, no prelo)
1ª fase	1532-1654 (do início da colonização à expulsão dos holandeses)	1552-1808 (o período colonial até a chegada de D. João VI)	1534 a 1750 (multilinguismo, descoberta do ouro)	Séc. XVI-1850 Multilinguismo generalizado Não-urbanização; Não-escolarização; Não-estandardização linguística	<1500 – tupinização
2ª fase	1654 a 1808 (da expulsão dos holandeses à chegada da Corte ao Brasil)	1808-1822 (da relutância à independência do Brasil)	(3 subfases): 1ª subfase: 1750 a 1808 2ª subfase: 1808 a 1850 (fim do tráfico) 3ª subfase: 1850 a 1922	a partir de 1850 Multilinguismo localizado, Urbanização; Escolarização; Estandarização linguística	1532-1694 multilinguismo generalizado (do início da colonização à destruição do Quilombo)
3ª fase	a partir de 1808	a partir de 1822	a partir de 1922	---	1ª subfase 1694-1930 – homogeneização linguística 2ª subfase >1930 – nivelamento linguístico e urbanização

QUADRO 2 - Síntese das principais periodizações do português brasileiro, com ênfase nos fatos sócio-históricos

Fonte: Elaborado pela autora desta tese (2017).

Não obstante possam equívocos¹³, as periodizações se revelam interessantes por considerar a sócio-história.

Como se observa, há várias periodizações linguísticas na história do português, quer seja o europeu quer seja o brasileiro, quase todas centradas em fatos que revelam situações de contato linguístico-cultural as quais possam ter influência em questões de ordem gramatical.

No tocante ao léxico, entretanto, não há uma periodização.

Considerando, pois, a ausência – e carência – de uma periodização com ênfase na formação do léxico português e admitindo as influências dos contatos

¹³ Dentre os equívocos destacam-se: i) a falta de coerência na caracterização de cada fase e o caráter preconceituoso com que trata os povos das línguas do contato, na proposta de Silva Neto (1986); ii) a incipiência e a minimização do contingente negro, na proposta de Teyssier (1997); a falta de coerência científica em considerar, concomitantemente, a história da língua falada e a história da língua literária e a falta de veracidade histórica na apresentação do multilinguismo apenas para o período de 1534 a 1750, na proposta de Pessoa (1997).

linguístico-culturais e das mudanças sócio-históricas “nessa longa estrada da vida” da língua portuguesa, estabelece-se uma possível proposta de periodização da estruturação do léxico português, sob a perspectiva de que seja “útil, periodizar quando, na base da periodização estão mudanças estruturais, globais e profundas” (LOBO, 2003). Observe-se o esquema a seguir:

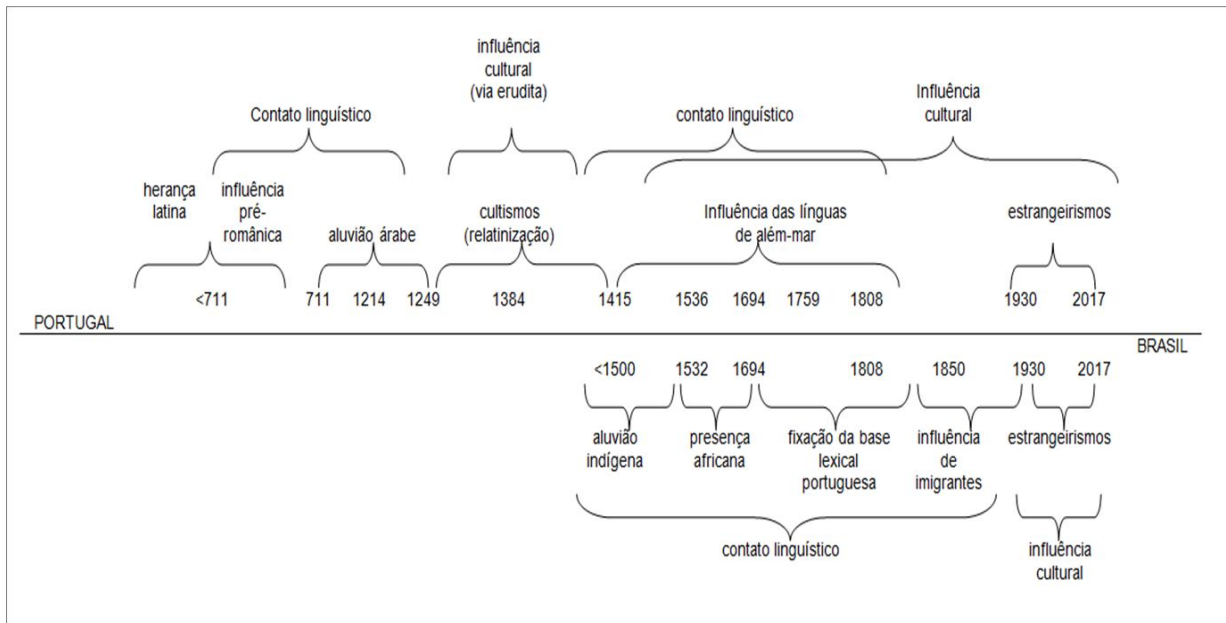


FIGURA 4 - Proposta de periodização da constituição do léxico do português - Portugal e Brasil
Fonte: Elaborado pela autora desta tese (2017).

Muito embora não se pretenda enfatizar fases para a formação do léxico português, tomou-se, para tanto, como base da descrição, duas situações propícias para a formação do léxico, o contato linguístico e as influências culturais, em duas perspectivas temporais – antes e depois do Brasil.

Os cortes epistemológicos do panorama proposto para a constituição do léxico do português visam, essencialmente, a destacar as principais influências linguísticas na formação do léxico português, das quais é possível distinguir os tipos de contribuição. As datas se justificam, pois, por representarem momentos importantes para a história sociolinguística e histórico-cultural do léxico português.

2.1.1 “Caminhando e cantando e seguindo a canção” e os passos: O arcabouço lexical do português¹⁴

Seguindo os passos que fizeram os caminhos da língua portuguesa, é possível encontrar rastros de i) herança latina ii) influências de línguas em contato e iii) influências culturais que, também, contribuíram para a formação do arcabouço lexical e difusão do léxico português –, estando nesta última o que hoje chamamos de estrangeirismos.

2.1.1.1 A herança latina

É consabido que a variação lexical era presente no latim, esteio principal da construção lexical do português. e que a difusão desse latim marcado pela variação léxico-semântica na Península Ibérica não ocorreu de maneira homogênea nas várias regiões em virtude de fatores como a diversidade da época da romanização que foi mais profunda ao sul do que ao norte, tendo por consequência a derivação do latim vulgar aos falares regionais.

No território que hoje é Portugal, o falar regional foi considerado como possuidor do caráter similar ao da língua latina e apresentado como “espelho da língua latina”, ou, caetaneando, seja a “flor do Lácio, sambódromo Lusamérica, latim em pó”.

A formação e diferenciação / caracterização das línguas românicas, ocorreu em níveis da língua, dentre tais se destacam as alterações fonético-fonológicas, alterações morfológicas e alterações sintáticas.

Castilho (ano), apresenta um quadro em que se destacam as seguintes diferenças: i) quanto à vogal postônica, manutenção na România Oriental – Latim Vulgar tegula > Italiano tegola, “telha” – e perda na Ocidental, como em Latim tegula > tegla > Portugues telha; ii) manutenção das consoantes surdas intervocalicas, como em Latim focu > Italiano fuocu e sonorização das consoantes surdas intervocalicas, como em fogo, no português; iii) assimilação do grupo consonantal octo > oito e semivocalização da primeira consoante do grupo *ct*, como em oito; dentre outras tantas alterações fônicas, mórficas e sintáticas.

¹⁴ Apresenta-se, aqui, uma síntese da estruturação do léxico por campos conceituais com base em Piel (1989), Vasconcelos (s/d) e Mattos e Silva (2008)

Quanto às alterações no português, processos metaplásmicos operaram fortemente sobre os dados fônicos latinos, a ponto de permitir a existência de formas divergentes ou convergentes no português. Veja-se, por exemplo, o que registra Machado Filho (2014), sobre isso.

TIPO DE DIVERGENTE	SUBDIVISÃO	ÉTIMO	FORMAS EM PORTUGUÊS
Hereditária	Fonético-históricas	<i>legitimu-</i>	lídimo
			lindo
de empréstimo de	acumulação popular (<i>mot populaire</i>) e erudita (<i>mot savant</i>)	<i>intregu-</i>	inteiro
			íntegro

QUADRO 3 - Formas divergentes no português

Fonte: Machado Filho (2014), com adaptações feitas pela autora desta tese

Como se pode observar dos extratos apresentados no *Quadro 3*, “inteiro” e “íntegro”, embora provenham do mesmo étimo *intregu-*, não possuem, hoje, a mesma carga sêmica. O mesmo ocorre com as lexias provenientes *legitimu-*, “lindo”, “lídimo”, com o agravante de que exista, também, em português, a forma “legítimo”, que pode ser interpretada como sinônimo de “lídimo”, sendo esta, aparentemente, de uso mais erudito.

De forma sintética, podem-se considerar as seguintes fontes:

	Palavras Populares	Semi-eruditas	Eruditas
Entrada no léxico português	em épocas muito remotas > transformações próprias à língua popular	nos primórdios da época literária > leves alterações ao entrarem no circuito da linguagem comum:	provindas diretamente do Latim Clássico, a partir do século XIV e no século XV, com o Renascimento
exemplo	<i>macula</i> > <i>macla</i> > <i>malha</i>	<i>humanitate</i> > <i>humanidade</i>	<i>flamma</i> > <i>flama</i> ; <i>auscultare</i> > <i>auscultar</i> ; <i>solitariu</i> > <i>solitário</i>

QUADRO 4 – Síntese dos tipos de entrada de lexias herança latina no léxico português

Fonte: Elaborado pela autora desta tese (2017)

Como se observa, as mudanças nos exemplos apresentados envolvem processos metaplásticos, os quais se configuram relevantes, inclusive, para se pensar o português brasileiro, sobretudo se fatores históricos do passado forem interpretados com vistas a entender o presente e (ou) se se partir de fatos e dados do presente entrevendo, pois, o passado.

No tocante às “palavras populares”, Vasconcelos (s/d), na *Lição IX*, resenha os principais processos formativos, a derivação verbal, a derivação imprópria e a sufixação, e em relação a esta admite que

[...] o povo é muito produtivo nesse campo e emprega processos pitorescos, expressivos, ora cingindo-se às leis e regras por que se rege a língua culta, [...] mas alargando o seu campo de acção; ora contrariando-as instintivamente, para provocar certos efeitos de contraste. (VASCONCELOS, s/d, p. 80)

Ainda sobre a derivação, a autora ressalta a influência prosódica dos vocábulos e dos sufixos no destino dos itens lexicais. À página 66, a autora explica:

Há sufixos átonos que são de grande produtividade. Não na linguagem culta, mas na boca do vulgo.

Êste serve-se seguramente de processos ancestrais, mas por serem il-
leterários, não podem ser documentados com textos.

[...]

Dêsses tipos conservam-se alguns pelo motivo fonético indicado. E êsses exemplares populares e semi-eruditos, tradicionais ou tornados tradicionais pelo muito uso que devem ao significado, são tantos que promoveram outros análogos.

[...]

E o vulgo o que fêz? Encostando-se a todos êsses, substituiu o sufixo menos sonoro *-ero* de algumas palavras por *-aro*; p. ex. em *númaro*, *mísaro*; e também *-oro* de *fósforo*, que deu *fósforo* (a par de *fosfo*, *forfo*, *forfro*)

(VASCONCELOS, s/d, p. 66)

O excerto supracitado enfatiza a relação dos vocábulos provenientes de processos, em que a prosódia e a fonética são marcantes e determinantes, com os fatores sociais, mais especificamente, o letramento. Além disso, ressalta-se a analogia como agente desencadeador de outros vocábulos presentes na língua, até hoje, em algumas comunidades.

Outra questão relevante em relação ao “excurso prosódico”, citado pela autora, se refere à acentuação, já que a transformação das proparoxítonas em paroxítonas se caracteriza uma tendência comum às mudanças pelas quais passou o latim vulgar e o romance de Portugal.

“A tendência oposta é unicamente própria de semi-letrados” (VASCONCELOS, s/d, p. 62), que usam *erúdoito*, *rúbrica* e *púdico*, em lugar de *erudito*, *rubrica* e *pudico*. Aliás, *rúbrica* e *rubrica* convivem, hodiernamente, de modo

que, inclusive, a variante não-padrão parece ter mais prestígio social do que a canônica. Este fato remonta a ideia da autonomia do sujeito em relação à língua.

Acredita-se que, para além da “tendência”, existem outros processos metaplásticos envolvidos, os quais são fundamentais para o estudo da formação do léxico português.

Dentre vários exemplos de como os metaplasmos podem interferir em mudanças lexicais, destaca-se, aqui, o verbete concernente a *jeolhos* presente à página 284 do *Dicionário do Português Arcaico*, de Machado Filho (2013), em que, além de apresentar outras duas variantes – *geolho* e *giolho*, evidenciando o alteamento da vogal média anterior em posição pré-tônica –, instiga-se o questionamento e o interesse de conhecer como a forma arcaica se tornou *joelho*.

Independente do tipo de divergência – se “hereditária” ou se “de empréstimo” – e de suas subdivisões possíveis, fato é que as formas em português se distanciam no que tange à significação ou até mesmo à pragmática, por motivações sociais e (ou) fonético-históricas.

Tomando por analogia o exemplo, dado em sala de aula por Machado Filho, de que se em uma medida composta por 50% de água e 50% de álcool, fosse inserido 50% a mais de água, o álcool perderia sua essência e a água a sua composição, o que as tornaria, mutuamente, irreconhecíveis, é possível dizer que traços de itens lexicais se modifiquem, ao longo dos tempos, devido a variáveis infinitas. No entanto, se os processos de mudança fossem registrados, oportunamente, seria possível recuperar a identificação das formas em cada momento, atestando, mais coerentemente que, não há como mensurar perdas ou ganhos, e sim apenas momentos do item lexical na história.

Ora, se para a história da língua, mais especificamente do léxico, as alterações fônicas são de suma importância, por que, em perspectiva sincrônica, não se registram dados que as reflitam e que se configuram léxico de dada comunidade?

Sobre a força das normas para a conformação de características próprias de determinadas comunidades de fala, sobretudo as populares na história, Vasconcelos (s/d) já assegurava algumas tendências metaplásticas no português. Diz a autora:

O povo, pronunciando *fosfo*, *liro*, *Ciro*, *Emila*, *famila*, *Antoino*, *paito*, em lugar de *pátio*, *Antônio*, *família*, *Emília*, *círio*, *lírio*, *fósfor* ou *fósforo*, continua na

mesma via. *Cambra, numbro, combro, por câmara, número, cômodo, que são senão efeitos dessa mesma tendência?* (VASCONCELOS, s/d, p. 62)

Com olhar no passado para interpretar o presente e, por conseguinte, resgatar elementos marginais da língua, em uma breve investigação, destacam-se duas variantes de dois dos itens citados pela autora: *fósforo* e *família*, que se encontram nos resultados do QFF – Questionário Fonético-Fonológico, parte integrante do Questionário ALiB 2001 (Comitê), *corpus* com o qual se desenvolveu esta tese¹⁵. Observe-se:

N	Word	Freq.	N	Word	Freq.	%
4.221	FORMULOU	11	3.886	FALSO	3	
4.222	FORNECE	1	3.887	FALTA	9	
4.223	FORNO	3	3.888	FALTÀ	1	
4.224	FORQUIA	1	3.889	FALTANDO	1	
4.225	FORQUILHA	2	3.890	FALTAR	1	
4.226	FORRÔ	1	3.891	FALTAVA	1	
4.227	FORTALECI	1	3.892	FALTO	1	
4.228	FORTALEZA	1	3.893	FALTÔ	6	
4.229	FORTE	70	3.894	FALTOU	4	
4.230	FORTEMENTE	1	3.895	FALU	1	
4.231	FORTES	2	3.896	FAMÍ	1	
4.232	FORTUNA	2	3.897	FAMIA	2	
4.233	FOSCO	3	3.898	FAMILHA	19	
4.234	FÓSCO	1	3.899	FAMÍLHA	3	
4.235	FÓSFO	1	3.900	FAMÍLIA	47	0,02
4.236	FÓSFORO	20	3.901	FAMÍLIAS	1	
4.237	FÓSFOROS	1	3.902	FAMINTA	1	
4.238	FÓSFORU	3	3.903	FAMOSA	2	
4.239	FÓSFURO	1	3.904	FAMOSO	1	
4.240	FÓSFURU	1	3.905	FANHÊ	1	
4.241	FOSSE	38	3.906	FANHO	1	
4.242	FOTO	1	3.907	FANHOSA	1	
4.243	FRACA	5	3.908	FANHOSO	5	
4.244	FRACÊS	3	3.909	FANHOSU	3	
4.245	FRACO	10	3.910	FANTA	2	
4.246	FRÁGIL	1	3.911	FANTAHMA	4	
4.247	FRAGRA	1	3.912	FANTASIA	2	
4.248	FRANCÊS	10	3.913	FANTASMA	11	
4.249	FRANCISCO	4	3.914	FANTÁSTICO	2	
4.250	FRANGO	3	3.915	FARÁ	1	
4.251	FRANQUI	1	3.916	FARPA	1	

FIGURA 5 – Teste de avaliação das variantes no *corpus*

Fonte: Elaborado pela autora (2017)

O confronto permitiu a observação de que: i) para *fósforo*, além do alteamento da vogal média posterior em posição pós-tônica, como ocorre em *fósforo*, documentam-se *fósco*, *fósfo* e *fósfo*, esta última citada por Vasconcelos (s/d), no tocante ao momento da formação do léxico do português e; ii) embora a variante

¹⁵ Embora somente os dados obtidos do QSL, os demais questionários foram utilizados como forma de recuperar informações que fossem importantes.

famila, citada por Vasconcelos (s/d) não tenha sido registrada no *corpus* em análise, se observa a iotização em *famia* e até mesmo a supressão do a final, como ocorre em *fami*.

Vale ressaltar que as observações e os exemplos são citados com a finalidade de evidenciar a importância do registro das formas hoje estigmatizadas ou desconhecidas do público em geral para a história da língua.

Atualmente, na França, é comum escutar, em contextos de informalidade, a expressão “ché pas” – [ʃepa] – para “je ne sais pas”. Ambas as formas coexistem no francês contemporâneo com os traços [+formal] para “je ne sais pas” e [+informal] para a variante “ché pas” – [ʃepa] –, bastante usual entre jovens e em situações / contextos informais.

Por enquanto, porém, “je ne sais pas” e “ché pas” – [ʃepa] – são variantes de registro (formal e informal) Esse fato pode ser observado como um exemplo de mudança linguística.

É válido lembrar que, historicamente, a marca de negação em francês é o “ne”. Hoje, porém, tem sido, às vezes, apagado. É comum escutar, ainda: “sucre ou pas”.

O esquema, a seguir, mostra o processo de variação atual.

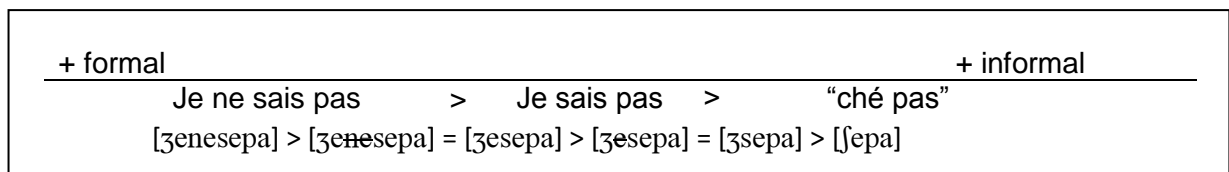


FIGURA 6 – De “Je ne sais pas” a “Ché pas”: exemplificação do processo de variação / mudança
Fonte: Elaborado pela autora (2017)

Diante desses fatos, é importante pensar como a documentação de uma forma em determinado momento da história pode servir para (ter a função de) salvaguardar a história da língua.

Seguindo os passos que fizeram os caminhos da língua portuguesa, é possível encontrar rastros das influências de contato linguístico e (ou) de contato cultural. Toda a discussão, a seguir, será feita com base em campos conceituais, com vistas ao interesse deste trabalho.

Para melhor visualização dos dados, opta-se por apresentá-las em duas perspectivas espaciais: em Portugal e no Brasil.

2.1.1.2 PORTUGAL

2.1.1.2.1 Contato linguístico

Os fatos apresentados visam ressaltar a importância do contato linguístico, inclusive na formação do léxico português

2.1.1.2.1.1 Influência pré-românica

A influência pré-românica, segundo Vasconcelos¹⁶ (s/d, p. 284), provém de “elementos estranhos que já tinham sido acolhidos pelos próprios Romanos, e foram por eles passados ao português”. Dentre essas “vozes sobreviventes de um grande naufrágio” – para valer-se da metáfora utilizada por Piel (1989 [1976], p.11) –, destacam-se as ibéricas, celtas, gregas, fenícias e orientais.

Piel (1989) e Mattos e Silva (2008) concordam na ideia de que os idiomas pré-latinos deixaram como legado lexical uma série de termos especiais relacionados aos conceitos topográficos, espécies de vegetais e noções de ambiente rústico.

Sobre a influência ibérica, Vasconcelos (s/d, p. 289) considera que os “mais importantes e numerosos, e quasi os únicos persistentes são nomes próprios, tanto topográficos, de rios montes, lugares, como de pessoas e sobretudo de divindades.” A autora destaca, também, a escassez de vocábulos fenícios, afirmando que “Quasi nada é seguramente fenício” (p. 293).

À página 295, a autora atribui à introdução do cristianismo – com os textos bíblicos (nas traduções chamadas *Itala* e *Vulgata*) – a presença de elementos gregos e orientais, mais especificamente, os “vocábulos *hebraicos*, internalizados pelo cristianismo – pela Bíblia”. Nesse ponto, talvez fosse mais adequado dizer que a inserção desses vocábulos no léxico português tenha ocorrido por influência cultural e não por contato linguístico.

Em relação à influência celta, à página 292, a autora distingue duas camadas de vocábulos: i) uma, “anterior às conquistas, entrou directamente em Roma, vindo do Norte da Itália, da Gália Cisalpina, desde o século IV”, da qual o português obteve relativos ao vestuário como, por exemplo, *saia* (sagum) e *bragas* (braccas) e;

¹⁶

ii) outra, mais abundante, proveniente da conquista da Gália por César, à qual pertencem elementos relacionados *mina, duna, roca*, dentre outros. No entanto, ressalta a autora que ambas “nacionalizaram-se e popularizaram-se tão completamente que entraram no património comum das línguas românicas como se fôsem verdadeiramente latinas”, de modo a ser imprudente admiti-las como tipicamente celtas.

Postura similar tem Piel (1989, p.11) ao afirmar que os elementos celtas e de idiomas anônimos “têm suas raízes em línguas aparentadas, possivelmente, com o basco ou com um substrato mediterrâneo”.

É válido destacar que a difusão do latim na Península Ibérica não ocorreu de maneira homogênea nas várias regiões em virtude de fatores como a diversidade da época da romanização que foi mais profunda ao sul do que ao norte, tendo por consequência a derivação do latim vulgar aos falares regionais. Por isso, segundo Piel (1989, p. 11),

a repartição geográfica dos referidos pré-romanismos é bastante desigual. Enquanto uns abrangem vastíssimas zonas da Península, prolongando-se a sua área mesmo muito além dos Pirinéus, há outros que se apresentam arrincoados em certas zonas galaico-portuguesas. O fundo linguístico de onde procedem não é menos variado. (PIEL, 1989, p.11)

Nesse sentido, é possível aventar que a heterogeneidade, o tipo e o período de romanização possam ter influenciado na inserção de elementos pré-românicos no léxico português. Outro fator relevante que inviabiliza a exatidão do reconhecimento das influências de línguas pré-românicas refere-se à ausência de documentação dessas.

Em síntese, no tange aos campos conceituais a que pertecem os elementos pré-românicos, observa-se certa divergência entre os autores, de modo que enquanto Vasconcelos (s/d) enfatiza a influência do contexto religioso, do cristianismo, Piel (1989) centra-se nos conceitos topográficos. Considerando as informações apresentadas, pode-se propor a seguinte síntese:

Tabela 1 – Principais domínios em que a influência pré-românica

	Domínios conceituais	Exemplos de vocábulos
POR CONTATO LINGUÍSTICO	conceitos topográficos	<i>barranco, lapa, morro, mouta,</i>
	espécies de vegetais espontâneas	<i>tojo, carrasco, chaparro, mato</i>
	ambiente rústico	<i>seara, broa, bezerro, bruxa</i>
	divindades	<i>Messias, Satanás</i>
POR INFLUÊNCIA CULTURAL	vocábulos hebraicos internacionalizados pelo cristianismo	<i>amém, sábado, páscoa</i>

Fonte: Elaborada pela autora (2017)

2.1.1.2.1.2 Influência pós românica

É consensual que as influências pós-românicas se restrinjam à presença germânica e árabe.

Quanto à contribuição germânica para o léxico português, Vasconcelos (s/d, p. 296) considera que seja tão numerosa quanto valiosa. Segundo a autora, subsistem no português, cerca de trezentos termos.

Em contrapartida, para Piel (1989, p. 13) a presença germânica não é tão marcante. Embora a distinga em i) os dialetos germânicos, ii) os visigotismos, oriundos do idioma dos visigodos, do ramo oriental e iii) os germanismos “de segunda mão”, galicismos de origem frâncica, o autor destaca, a dificuldade, em alguns casos, de definir se se trata de um visigotismo ou de um elemento frâncico, trazido simultaneamente com outros galicismos.

Segundo Piel (1989), os exemplos mais seguros de visigotismos do português referem-se a atividades militares, conceitos jurídicos e objectos caseiros, assim como dois nomes de aves (“mejengra ‘*chapim*’ e laverca ‘*cotovia*’”). Quanto aos dialetos germânicos, afirma que “só poucos foram transmitidos às línguas românicas, como CARPA ‘carpa’ e COFEA ‘coifa’” (p.13).

Segundo Mattos e Silva (2008, p. 86)

Coloma Leal, abordando os empréstimos germânicos numa perspectiva sócio-linguística (sic), considera que, já antes das invasões germânicas, numerosos germanismos se introduziram no latim, pelos intercâmbios que se estabeleceram nas fronteiras, nas guarnições da retaguarda, em que os soldados, sobretudo os chefes, eram romanos, enquanto as mulheres eram da região dominada pelos Romanos. Com base nisso, conclui a autora que os primeiros germanismos pertencem a duas áreas lexicais, aparentemente contraditórias: a dos assuntos domésticos e a das questões militares.

De acordo com a assertiva de Coloma Leal, corroborada por Mattos e Silva (2008), pode-se pressupor que a divergência entre Vasconcelos e Piel esteja relacionada à datação, ao contexto sócio-histórico da influência germânica.

Observe-se, a seguir, os campos conceituais dos germanismos.

Tabela 2 – Domínios dos germanismos

	domínios conceituais	exemplos
DIALETOS GERMÂNICOS	---	apenas CARPA 'carpa' e COFEA 'coifa'.
VISIGOTISMOS	atividades militares, conceitos jurídicos e objetos caseiros	<i>aleive, bando, espeto, espora, espia, escanção, luva, roca</i>
	nomes de aves (manifestamente de agouro)	<i>mejengra 'chapim' e laverca 'cotovia'.</i>
GERMANISMOS “DE SEGUNDA MÃO”	relacionados ao período feudal	<i>bastir, elmo, estala, guerra; guardar, guarnecer, jardim, rico, sala, trégua, venda 'fita'</i>

Fonte: Elaborada pela autora (2017)

Das “vozes não-latinas”, a influência árabe, indubitavelmente, é a que mais se destaca.

Segundo Piel (1969), que chega a denominar de “aluvião lexical árabe”, a presença do léxico dessa língua no português, diferentemente dos demais elementos cujas influências no léxico se restringem a aspectos específicos, abrange todos os setores da vida material (designações de cargos; termos castrenses: de administração; de plantas cultivadas e silvestres; de profissões e indústrias; de unidades de medidas; de animais; de produtos agrícolas; de arquitetura; de ciências exatas).

Ainda sobre a influência árabe, Machado Filho (ano), no texto *Léxico de étimo árabe em uso no período arcaico*, compova a “indelebilidade da história e da sócio-história na composição das línguas naturais”, ao depreender com base nos dados do corpus analisado que “alguns elementos ainda se encontram em plena produtividade no português brasileiro contemporâneo. O autor diz, ainda, que:

Depois de 764 anos do último momento em que esteve presente na Península Ibérica, convivendo com portugueses e moçárabes, a força da cultura árabe poderia, por mais estranho que pudesse parecer, **afagar fulanos, achacar refeces** ou **mesquinhos, até açoitar sandios**, prendê-los com **algemas**, ou enfeitar de **aljôfar o forro**, para que a liberdade reluzisse do passado. Alguns dialetos do português ainda podem, talvez, utilizar a **nora** para retirar água de um poço, que uma prospecção dialectológica, na dimensão do Projeto ALiB, pudesse confirmar ou negar.

Por se tratar de uma investigação voltada para o levantamento da presença de elementos árabes, em uso no português arcaico, um falante do português brasileiro, talvez, não reconhecesse muitas delas, entretanto, poderiam identificar outras formas sem reconhecer a origem árabe.

Considerando que as contribuições da dominação árabe são refletidas em diferentes aspectos da vida quotidiana, pode-se inferir que na língua também tenha havido influências no tocante aos mesmos aspectos que se constituam campos conceituais. Observem-se, a seguir, algumas influências.

Tabela 3 – Domínios da dominação árabe

	domínios conceituais	ítems lexicais
ALUVIÃO ÁRABE	designações de cargos e dignidades:	<i>alcaide, alferes, almoxarife;</i>
	termos castrenses	<i>arraial, arrebate, alcácer, alcáçova, atalaia</i>
	de administração	<i>aldeia, arrabalde, alfoz, alfândega, alvará, almoeda</i>
	de plantas cultivadas e silvestres:	<i>arroz, algodão, alcachofra, cenoura, laranja, alfarroba,</i>
	de profissões e indústrias:	<i>alfaiate, alveitar, almocreve, alvanel, algoz, azenha, atafona,</i>
	de unidades de medida:	<i>almude, arrâtel, alqueire,</i>
	de animais:	<i>atum, alcatraz, alforreca,</i>
	de particularidades topográficas:	<i>albufeira, alverca, algar, lezíria,</i>
	de artigos de luxo e instrumentos de música:	<i>almofada, alcatifa, marfim, alfinete, adufe, rabeca, anafil,</i>
	de produtos agrícolas e industriais:	<i>azeite, álcool, alcatrão,</i>
	da vida pastoril:	<i>zagal, alfeire, rês, tabefe,</i>
	de arquitetura:	<i>aljube, chafariz, açoteia,</i>
	das ciências exatas:	<i>algarismo, álgebra, cifra, auge,</i>

Fonte: Elaborada pela autora desta tese (2017)

2.1.1.2.1.3 Influência das línguas do além-mar

Designam-se, aqui, influências de além-mar aquelas provenientes da expansão marítima e colonização portuguesas, iniciadas em 1415, as quais são designadas por Piel (1989) e Vasconcelos (s/d) como elementos exóticos¹⁷.

¹⁷ Embora se compreenda as perspectiva dos autores, levando em consideração a época em que os textos foram publicados, optou-se por não utilizar o termo “exótico” haja vista a carga semântica pejorativa que possui, atualmente.

Segundo Piel (1989, p.15), essas vozes se infiltram no léxico português, seja por via literária (cronistas), seja por via oral (convivência com povos indígenas; comércio), no século XVI, “aumentando assim consideravelmente o caudal de termos orientais recebidos anteriormente por intermédio do árabe”.

Quanto a essa classe de elementos, sem explicações pormenorizadas, às páginas 319 a 322, Vasconcelos (s/d) apenas apresenta amostras “dos mais conhecidos, empregados positivamente na linguagem comum dos cultos e do vulgo” e assevera que “[...] centenas de brasileirismos [...] registrados em dicionários portugueses modernos, por eles estarem destinados aos dois países, não são empregados vulgarmente em Portugal.” (p. 320)

Mattos e Silva (2008) cita o trabalho de Denise Dias, cuja síntese se apresenta na tabela seguinte:

Tabela 4 – Domínios da influência além-mar

	domínios conceituais	itens lexicais	atualização dos itens lexicais
CARTA DE CAMINHA	indumentárias e adereços	<i>sombreiro de penas de aves; cabeleira de penas de ave; carapuça de penas</i>	elementos que se distinguem pelas funções e pelos formatos
		<i>ramal grande de continhas</i>	não é mais usado
		<i>pedra verde roim</i>	pedra usada nos lábios dos índios como sinal de autoridade
	armas e instrumentos	<i>arco e setas</i>	arco e flechas
		<i>cornobuzina</i>	instrumentos de sopro feitos de chifres de animais
		<i>pedras como cunhas metidas em um pau entre duas talas muito bem coladas</i>	machado de pedra
	moradia e transporte	<i>almadia</i>	jangada
		<i>choupaninha de rama verde</i>	ocas
	coisas do mar	<i>botelho</i>	sargaço
		<i>fura-buchos</i>	vira-bucho
	coisas da terra	<i>inhame</i>	mandioca
		<i>palmas/palmitos</i>	palmeiras ou coqueiros
		<i>ouriços verdes</i>	urucum
	fauna	<i>papagaios vermelhos</i>	araras vermelhas
		<i>papagaios pequeninos</i>	espécies de aves similares (periquito, cuiuba...)

Fonte: Elaborada pela autora (2017)

2.1.1.2.2 Influência cultural

Quanto aos “empréstimos culturais”, designação adotada por Mattos e Silva (2008) para as entradas lexicais que “decorrem não contacto de populações de línguas distintas, em um mesmo território”, mas provenientes de contatos constituídos com base em intercâmbios culturais, Vasconcelos (s/d) e Piel (1989) convergem na ideia de que os mais significativos são os franceses e italianos.

A influência civilizadora francesa no léxico português pode ser caracterizada em dois momentos: i) nos séculos XI, XII e XIII, com os primeiros contatos dinásticos e ao período feudal, nos quais se destacam vocábulos relacionados ao comércio e ii) nos séculos XIV e XV, com as influências culturais e literárias, a exemplo da moda que se reflete em vocábulos relacionados à indumentária; da culinária; dos termos relacionados à música, dentre outros.

Quanto à contribuição do italiano, Piel (1989, p. 15) ressalta o papel histórico-cultural que a Itália desempenhou o qual se propagou por quase toda Europa. De modo geral, para o autor, a influência concerne a termos de civilização e terminologias artísticas. O autor ressalta, porém, que nem “todos os italianismos do português foram importados directamente, tendo às vezes servido de intermediários o espanhol ou o francês”.

Em contrapartida, Vasconcelos (s/d, p. 317) destaca e atribui à contribuição italiana os vocábulos relativos à arte militar e à navegação, com a seguinte assertiva: “[...] se os Normandos e Neerlandeses foram mestres dos Portugueses para a navegação no oceano Atlântico, os Italianos o foram para o Mediterrâneo.”

Observem-se algumas contribuições:

Tabela 5 – Domínios da influência cultural europeia

	domínios conceituais		ítems lexicais
Franceses Provençais	lírica dos provençais		<i>trovar, trovador, refrão, segrel, jogral.</i>
	indumentária (reflexo da moda francesa)		<i>boné, chapéu, colete, blusa, paletó;</i>
	culinária		<i>croquete, filete, fricassé, puré</i>
	termos de música		<i>charamela, fabordão, oboé.</i>
	comunicações		<i>comboio, trem, cais, bilhete</i>
Italiano	arte	de literatura	<i>soneto, burlesco</i>
		de música	<i>piano, contralto, ópera</i>
		de arquitetura	<i>balcão, fachada</i>
		de pintura	<i>aguarela, pitoresco</i>
Espanhol	ambiente cortês		<i>cavalheiro, lhano, airoso</i>
	noções militares		<i>cabecilha, caudilho, guerrilha</i>

Fonte: Elaborada pela autora desta tese (2017)

Com vistas a sintetizar as contribuições para a formação do léxico português, propõe-se um novo parâmetros, explicitado na tabela abaixo.

Tabela 6 - Síntese das influências do contato linguístico por domínios conceituais

domínios conceituais	vozes do contato (palavras de empréstimos)				
	pré-românica	germânica	árabe	européia	não europeia expansão marítima
cargos e divindades	+	-	+	-	-
arte (música, literatura)	-	-		+	-
termos castrenses	-	-	+	-	-
administração	-	-	+	-	-
de plantas cultivadas e silvestres:	-	-	+	-	+
profissões e indústrias:	-	-	+	-	-
unidades de medida:	-	-	+	-	-
animais:	-	-	+	-	-
contexto religioso	+	-	-	-	-
culinária		-	-	+	+
particularidades topográficas	+	-	+	-	-
artigos de luxo e instrumentos de música	-	-	+	-	-
produtos agrícolas e industriais:	-	-	+	-	+
vida pastoril:	-	-	+	-	-
arquitetura:	-	-	+	+	-
ciências exatas:	-	-	+	-	-
atividades militares	-	+	-	+	-

Fonte: Elaborada pela autora desta tese (2017)

2.1.1.3 BRASIL

Outro cenário a ser interpretado no que tange à formação do léxico do português brasileiro é o das “vozes” indígenas e africanas, quase sempre amordaçadas pela história, em sua integridade, conquanto mais evidentes no léxico nacional, já que

é sobretudo quando se trata de identificar objetos e noções próprios à realidade brasileira, ao clima, à flora, à fauna, às tradições locais, aos costumes, à cultura popular, à vida social que o “brasileiro” manifesta a sua criatividade vocabular e fraseológica. E, para isso, tem recorrido freqüentemente às duas fontes postas à sua disposição pelas duas populações com as quais os portugueses se misturaram no solo brasileiro: as línguas dos indígenas (em primeiro lugar, o tupi) e as línguas dos escravos negros (TEYSSIER, 1997, p.72).

Dessa forma, “o que viria a se transformar no português brasileiro passa, a partir dos primeiros contatos, a vivenciar os efeitos plurais que a sócio-história impõe a todas as línguas naturais” (MACHADO FILHO, no prelo, tradução)

2.1.1.3.1 Contato linguístico

2.1.1.3.1.1 O aluvião indígena

One can say that Portuguese keeps, today, something around a thousand inherited lexical items from Arabic, inherited during the Islamic rule on the Iberian peninsula process, which occurred between 711 and 1249 in Portugal, and around five hundred inherited words from Germanic languages. But it would not be possible to count numerically, in the same manner, the effective contribution of Brazilian indigenous languages, without the risk of great uncertainty until a systematic and historical research in language dictionaries and vast corpora is performed, although it is suspected that it exceeds the Arab heritage, at least in the passive vocabulary (MACHADO FILHO, no prelo).

Machado Filho (no prelo) considera ainda ser possível percorrer todo o alfabeto na etimologia tupi na língua portuguesa. Essa evidência é também notada por Teyssier (1997, p. 71) ao considerar que seja “por todo esse vocabulário tipicamente brasileiro, [...] que a velha língua geral como que sobrevive.”

Entretanto, no que tange à contribuição indígena pode-se atestar que esta se concentre fortemente na toponímia, na fauna e na flora.

Tabela 7 – Domínios da influência indígena

		campos conceituais	itens lexicais
ALUVIÃO INDÍGENA	toponímia		<i>Aracaju, Guanabara, Carioca, Tijuca),</i>
	flora	uma série de nomes de árvores	<i>peroba, canjarana, caroba, imbuia, jacarandá, araticum, ipê, cipó</i>
		frutas	<i>pitanga, maracujá, jaboticaba, caju. abacaxi</i>
	fauna	mamíferos	<i>capivara, quati, tatu, sagüi;</i>
		serpentes	<i>caninana; sucuri;</i>
		peixes	<i>acará piranha</i>
		pássaros	<i>araponga urubu curiango curió sabiá.</i>
	personagens espectrais e inquietantes		<i>saci e caipora.</i>
	alimentação		<i>mingau, moqueca.</i>
idiomatismos		<i>andar na pindaíba estar de tocaia</i>	

Fonte: Elaborada pela autora desta tese (2017)

É válido ressaltar que as línguas indígenas foram gramaticizadas por portugueses, na mesma época em que surgiram as primeiras gramáticas de língua

portuguesa. A gramática de Anchieta é exemplo disso. E embora existam estudos sobre tupinismos, urge, como para Mattos e Silva (2008, p. xx) “que se faça um vocabulário geral das inovações lexicais que se incorporam ao português a partir da segunda metade do século XV, e avançando pela rica documentação escrita no século XVI”.

2.1.1.3.1.2 A presença africana

Regarding the contributions of African languages for the composition of the Portuguese lexical inventory, it must be considered initially that speakers of a single language family - the *níger-congo family* – from the *congo-cordofaniano* branch would have come to Brazil throughout the slavery period, except for a small group of speakers of the Islamized Hausa language, another linguistic branch - the Afro-Asian- would come in the nineteenth century to a few urban centers without large lexical impacts for southern Portuguese on this side of the globe (MACHADO FILHO, no prelo).

No tocante às línguas africanas trazidas ao Brasil, ressalta Teyssier que, com efeito, os escravos pertenciam às mais variadas etnias. Entretanto, duas línguas africanas tiveram um papel particularmente importante no Brasil: o iorubá, que está na base de um vocabulário próprio à Bahia concernente às cerimônias de religiões de base africana ou à cozinha afro-brasileira, e o quimbundo, que legou ao Brasil um vocabulário mais geral, como se pôde observar na citação acima.

Teyssier (1997, p. 71-72) ressalta, ainda, alguns “problemas complexos” levantados pelo vocabulário brasileiro de origem africana.

Certas palavras passaram diretamente da África a Portugal, sem transitar pelo Brasil, e foram, posteriormente, introduzidas no país pelos portugueses. É o caso de inhambe, palavra africana que se encontra sob a forma espanhola ñame no Diário de Cristóvão Colombo (1492) e sob a forma portuguesa na carta em que Pêro Vaz de Caminha, em 1500, dá notícia do descobrimento do Brasil ao rei D. Manuel. Se nos ativermos ao vocabulário indubitavelmente introduzido pelos escravos transportados ao Brasil, verificaremos que ele é de origem diversa (TEYSSIER, 1997, p. 71-72).

Segundo Tinhorão (1988), com efeito, os escravos pertenciam às mais variadas etnias. Entretanto, duas línguas africanas tiveram um papel particularmente importante no Brasil: o iorubá (falado atualmente na Nigéria) e o quimbundo (falado em Angola).

O *quimbundo* chegou ao Brasil um vocabulário mais geral, quase sempre integrado à língua comum (ex.: caçula, cafuné, molambo moleque). Muitas vezes esse vocabulário evoca o universo das plantações de cana-de-açúcar (ex.: banguê),

com os escravos, seu modo de vida e suas danças (ex.: senzala, mocambo, maxixe, samba). Já o iorubá está na base de um vocabulário próprio à Bahia, relativo às cerimônias do candomblé (por ex.: orixá) ou à cozinha afro-brasileira (ex.: vatapá, abará, acará, acarajé).

Ainda segundo Tinhorão (1988), à página 377, a

constância das referências a uma língua de negros, em quase cinco séculos de história do teatro em Portugal, leva a imaginar que, se tal forma corrompida de falar o português de mistura com termos africanos chegou a construir quase um dialecto na metrópole, alguma consequência deverá ter resultado de tal intercâmbio linguístico. E, na verdade, embora a sintaxe portuguesa continuasse inatingida, pelo facto de as alterações da fala se terem circunscrito sempre à fonética e à morfologia da língua de empréstimo, o léxico não deixaria de acusar, afinal, exemplos resultantes de tão longa troca cultural (TINHORÃO, 1988, p. 377).

Embora a citação apresente diversos equívocos, faz-se relevante refletir acerca da influência das línguas africanas para a construção do léxico do português, a ponto de “quase construir” um dialeto da metrópole. Registre-se, entretanto, uma nota de repúdio ao preconceito expresso pelo autor em “forma corrompida de falar o português de mistura com termos africanos”.

Ainda na mesma citação, o autor alega que a sintaxe portuguesa continuava inatingida “pelo facto de as alterações da fala se terem circunscrito sempre à fonética e à morfologia da língua de empréstimo”, mas que o léxico já acusava exemplos resultantes “de tão longa troca cultural”. Questiona-se: Até que ponto as alterações fonéticas e morfológicas interferiram no léxico?

Às páginas 382 a 392, são listadas, por Tinhorão (1988), 97 palavras de origem africana presentes no português em Portugal. Destas, 39 são do Quimbundo, 10 do Ambundo. Ressaltam-se *batuque*, *cupata* e *fumo*, respectivamente, do landim, da língua geral africana e do congôês. As demais são listadas sem a identificação etimológica.

Um fato que chama atenção é o registro de formas alterações ortográficas como ocorre em *gimbo* e *jimbo*, sendo esta a forma canônica e oriunda do quimbundo *njimba*; e alterações fônicas como ocorre em *enxacôco* para *xacôco* – por processo prótese – e *malufo* e *marafô* para *marufo* – por lambdacismo e (...), respectivamente.

Uma questão ainda se revela relevante: como saber se as palavras de origem de línguas africanas foram trazidas diretamente para o Brasil ou já estavam na língua portuguesa? Como se procederia lexicograficamente?

2.1.1.3.1.3 Base portuguesa

Mattoso Câmara Jr. (1979) desenvolve uma interessante análise da constituição do léxico português com base em cinco campos conceituais¹⁸: i) mundo físico; ii) partes do corpo humano; iii) parentesco; iv) tempo (transcurso) e v) tempo (climático).

À página 199, assevera o autor:

A importância lingüística está aqui ligada à estrutura fundamental da cultura a que a língua serve e às designações das coisas e atividades humanas básicas. Os campos semânticos (sic) referentes a essas coisas e atividades, dentro da estruturação cultural, são preenchidos por palavras nucleares da língua, decorrentes do acervo mais íntimo e primitivo; as contribuições por empréstimo e as substituições lexicais posteriores são aí altamente expressivas. (CÂMARA JR., 1979, p. 199-200)

Não obstante o equívoco do autor, em não distinguir campos semânticos de campos conceituais – com os quais trabalha, de fato –, a citação se faz pertinente e a análise interessante a esta pesquisa, justamente, em função das áreas temáticas com que se desenvolve o presente estudo.

A lista de itens lexicais do autor – termos portugueses que se enquadram no critério considerado “palavras primitivas e de uso corrente” – compreende 152 lexias, dentre as quais: i) 53 são relacionadas ao mundo físico; ii) 47 às partes do corpo humano; iii) 21 concernem ao parentesco; iv) 16 são de tempo (transcurso) e v) 15 de tempo (climático).

2.1.1.3.1.4 Influência dos imigrantes

Ao pesquisar sobre as influências de línguas de imigrantes para a composição do léxico do português brasileiro, destacaram-se línguas: a italiana, a japonesa. É possível observar que influências tão específicas resultaram da própria

¹⁸ O autor chama de campos semânticos, que se confunde, no entanto, com o que se convencionou chamar de conceitual e lexical, em Lexicografia.

troca cultural entre brasileiros e imigrantes destes países, e as palavras trazidas por eles refletem a sua influência não só na língua mas na cultura brasileira, em geral.

Tabela 8 – Domínios da influência dos imigrantes

	domínios conceituais	itens lexicais
ITALIANO	culinária italiana	Lasanha: lasagna; Nhoque: gnocchi; Espaguete: spaghetti; Talharim: taglierini; Mussarela: mozzarella
	termos científicos e médicos	escorbuto, fontanela e malária
	áreas militar e naval	atacar, brigada, canhão, coronel, batalhão, esquadrão, mosquete, casamata, fosso, reduto, soldado, infantaria, florete, avaria, carena, mercante, piloto, fragata, galera, bússola,
	arquitetura, costura, teatro e às artes plásticas	aquarela, esboço, desenho, esfumar, têmpera, paleta, e magenta; baeta, brocado, feltro e tafetá; balcão, baldaquim, cenário, entalhe, maquete, mezanino, pérgula, pilastra,
	serviços financeiros ¹⁹	banco, concordata, crédito, débito, bancarrota, empresário
JAPONÊS	esportes	judô, jiu-jitsu, caratê, kendo, sumo, gateball
	frutas	caqui, a maçã Fuji, poncã (tangerina)

Fonte: Elaborada pela autora (2017)

2.1.1.3.2 Influência cultural

2.1.1.3.2.1 Estrangeirismos

Com o advento da globalização, os estrangeirismos no cenário nacional se tornaram cada vez mais frequentes e geraram, há alguns anos, nos finais do século XX, uma polêmica desencadeada pelo Projeto de Lei 1.676 de 1999, de autoria do então deputado Aldo Rebelo, do PCdoB, que objetiva tratar da “proteção, defesa e uso da língua portuguesa” no Brasil. No **Art. 4º** consta que:

Todo e qualquer uso de palavra ou expressão em língua estrangeira, ressalvados os casos excepcionados nesta lei e na sua regulamentação, será considerado lesivo ao patrimônio cultural brasileiro, punível na forma da lei.

[...]

Parágrafo único. Para efeito do que dispõe o caput deste artigo, considerar-se-á:

I - prática abusiva, se a palavra ou expressão em língua estrangeira tiver **equivalente**²⁰ em língua portuguesa;

II - prática enganosa, se a palavra ou expressão em língua estrangeira puder induzir qualquer pessoa, física ou jurídica, a erro ou ilusão de qualquer espécie;

¹⁹ Os termos relacionados aos serviços financeiros, amplamente disseminados em todo mundo, decorrem da institucionalização desses, em Florença, no séc. XV.

²⁰ (grifo nosso)

III - prática danosa ao patrimônio cultural, se a palavra ou expressão em língua estrangeira puder, de algum modo, descaracterizar qualquer elemento da cultura brasileira.

Sobre o fator “equivalência”, citado no Projeto de Lei supracitado, é consabido que todas as línguas humanas têm diferentes lacunas no estoque léxico ou vocabulário, as quais podem ser preenchidas por empréstimos linguísticos. Ocorre, entretanto, que essa questão não pode ser observada estritamente do ponto de vista linguístico, mas e sobretudo com viés sociocultural. Observe o que diz Schmitz (2000, p. 44-45), acerca da proposta de estabelecimento de uma grafia aportuguesada de determinadas vocábulos, que embora sejam comuns socialmente, pertencem a um grupo específico.

Os praticantes das atividades *karaokê* e *karatê* aceitariam respectivamente as grafias *caraoquê* e *caratê*. As palavras de origem hebraica, *shloishim* e *matzeiva*, seriam reconhecíveis por parte dos que precisam desses vocábulos se a grafia das mesmas fossem aportuguesadas? A vontade de apagar diferenças, impor uma uniformização e controlar a produção escrita por via de multas parece uma postura xenofóbica e antidemocrática. (SCHMITZ, 2000, p. 44-45)

A descrição detalhada das influências das vozes do contato, das palavras de empréstimos demonstra a fraqueza dos argumentos de Rebelo (1999) para criação do Projeto de Lei 1.676, antes visto. Mais ainda, mostra como indiretamente da fonte de que se importe o item lexical, o sistema linguístico da língua receptora fará as devidas adequações fônicas e mórficas para a aceitação e uso disseminado.

É possível admitir, pois, que a inserção de palavras estrangeiras não comprometem o idioma, inversamente, o enriquecem e o desenvolvem cultural e técnico-cientificamente, já que, sempre assim, e deverá ser assim, a construção do espólio lexical de uma língua.

Estudos recentes têm mostrado que um fator relevante para a diferença entre as variedades da língua portuguesa é a presença de estrangeirismos no português brasileiro.

Com base em resultados do estudo de caso sobre o uso e o sucesso de estrangeirismos que designam conceitos de futebol e de vestuário nas duas variedades nacionais do português ao longo dos últimos 60 anos e em diferentes variedades estilísticas, Soares da Silva (2015) demonstra a relevância dos estrangeirismos, na diferenciação do português brasileiro e do português europeu.

Para tanto, os estrangeirismos no vocabulário de futebol são distribuídos por dois conjuntos: anglicismos e todos os estrangeirismos (incluindo espanholismos, italianismos e galicismos). Os estrangeirismos no vocabulário do vestuário são distribuídos por três conjuntos: galicismos, anglicismos e todos os estrangeirismos.

No que tange ao vocabulário de futebol, os resultados atestaram que a influência dos estrangeirismos, principalmente, os anglicismos, é claramente maior no português brasileiro do que no português europeu.

Se por um lado, o português brasileiro tende à importação direta de estrangeirismos, à adaptação ou ao decalque, por outra via, o português europeu é mais suscetível à substituição dos estrangeirismos por termos vernáculos. É o que ocorre, por exemplo, com “guarda-redes” e “grande penalidade” para “goleiro” e “*pênalty*”, respectivamente.

No tocante ao léxico do vestuário,

verifica-se uma diminuição de francesismos e um claro aumento de anglicismos, por razões socioculturais bem conhecidas, designadamente a diminuição da influência do francês e o inevitável aumento da influência do inglês. No conjunto de todos os estrangeirismos no vocabulário do vestuário, temos um aumento no PE e uma situação de estabilidade no PB. A influência de estrangeirismos continua a ser maior no PB, embora a diferença entre as duas variedades seja menos acentuada (SILVA, 2015, p. 88).

Como se observa, a seguir, os itens lexicais provenientes de ambos difundiram-se por diferentes campos conceituais.

Tabela 9 – Influência francesa e inglesa no léxico do português

itens lexicais	
FRANCÊS	balé, bijuteria, bufê, butique, cachecol, cachepô, chance, dossiê, garçom, grand prix, maître, ménage, omeleta ~ omelete, revanche, réveillon, sabotagem
INGLÊS	tobaby-doll, babylook, bangalô, banguê-banguê, bonde, box, checape, check-in, drope, ferry-boat, futebol, gay, hall, hambúrguer, hobby, jeans, jipe, miss, piquenique, náilon, o.k., tanque, tíquete, w.c.,

Fonte: Elaborada pela autora (2017)

2.2 O LÉXICO PORTUGUÊS EUROPEU E O LÉXICO PORTUGUÊS BRASILEIRO: VEREDAS

A diferença do léxico do português europeu e do português brasileiro é amplamente difundida e já reconhecida popularmente. Recentemente, um texto curioso e anedótico circulou em redes sociais:

Acordei atrasado hoje, abri o **guarda-fatos** e não achava meu **fato** na **cruzeta**. **Tomei uma bica**, mas a **chavena** estava muito quente. Na rua cumprimentei o **almeida** que estava a **varrer a berma**. Peguei o **autocarro** que estava cheio, uma **bicha** enorme atrás de mim. Tinha um **zare** perturbando a todos. Cheguei ao **sítio de trabalho**, estavam em greve, um **ula-ula** total. Um calor, estava com a **justa** toda molhada já e uma **harpa** de morrer. Comprei uma **gasosa** pra matar a sede (NEPOMUCENO, 2008).

Para um utente/falante comum do português brasileiro, leigo dos usos do português europeu, o entendimento do texto supracitado, talvez seja comprometido ou sido prejudicado em função de alguns itens lexicais, mais especificamente os negritados. A reescrita do texto, com a substituição do correspondente semântico do item lexical, o tornaria claro, como se pode observar:

Acordei atrasado hoje, abri o **guarda-roupa** e não achava meu **terno** no **cabide**. **Tomei um cafezinho**, mas a **xícara** estava muito quente. Na rua cumprimentei o **gari** que estava **varrendo** o **acostamento**. Peguei o **ônibus** que estava **cheio**, uma **fila** enorme atrás de mim. Tinha um **bêbado** perturbando a todos. Cheguei ao **lugar de trabalho**, estavam em greve, um **confusão** total. Um calor, estava com a **camisa** toda molhada já e uma **fome** de morrer. Comprei um **refrigerante** pra matar a sede.

E, como poderia ser esse texto, em “bom em baianês”?

Acordei atrasado hoje, abri o **guarda-roupa** e não achava meu **terno** no **cabide**. **Tomei um cafezinho**, mas a **xícara** estava **pelando**. Na rua, cumprimentei o **gari** que estava a varrer o **meio-fio**, **passeio**. Peguei o **buzu** que estava **lotado**, uma **fila** enorme atrás de mim. Tinha um **bêbado cachaceiro** perturbando a todos. Cheguei ao **lugar de trabalho**, estavam em greve, um **muvuca** total. Um calor, estava com a **camisa** toda molhada já e uma **fome** de morrer. Comprei um **refri** pra matar a sede

Como se observa, alguns os itens lexicais *guarda-roupa*, *terno*, *cabide*, *cafezinho*, *xícara*, *fila*, *camisa*, *fome* se manteriam mas, outros como *acostamento*, *ônibus cheio*, *bêbado*, *correria* e *refrigerante* certamente, seriam alterados.

Os três itens lexicais que seriam modificados fazem parte do *corpus* desta pesquisa e, de fato, foi possível evidenciar, pela ausência ou baixa frequência de

uso dos mesmos, que o baiano já opta, hoje, por variantes menos comuns ao português brasileiro *standard*.

Segundo Biderman (2001), o léxico tem papel fundamental na estrutura e funcionamento da língua “porque refere os conceitos lingüísticos e extralingüísticos da cultura e da sociedade; por essa razão são bem grandes as diferenças lexicais entre o PE e o PB.”

Ainda sob a ótica de Biderman (1996) as diferenças entre os léxico português e brasileiro “situam-se em torno de umas cem palavras”, o que não é muito num total de 2.217. A autora afirma ainda que os:

itens lexicais usuais em Portugal e não no Brasil podem ser de dois tipos: 1. os significantes usados no Brasil são diferentes, ou seja, usamos um termo diverso; 2. o signo total não se usa no Brasil (ou é raro entre nós) em virtude das peculiaridades do universo físico e cultural português por oposição ao Brasil. (BIDERMAN, 1996, p. 41)

Com base na citação supracitada, é possível admitir que as peculiaridades da cultura e do meio ambiente nos quais o português brasileiro desenvolveu sua identidade linguístico-cultural geraram um vocabulário distinto do que se usa em Portugal.

O quadro, a seguir, parece aludir de maneira elucidativa ao que é dito.

Tabela 10 – Português europeu (PE) versus Português brasileiro (PB): Diferenças lexicais

domínios conceituais	itens lexicais		
	PE	PB	
		correspondente semântico	correspondente lexical (alteração semântica)
cidade (urbano)	café	bar	---
	sinalete	sinaleiro, sinal, semáforo	---
	fumo	fumaça	---
	montra	vitrine	---
meios de transporte	autocarro	ônibus	---
	camioneta		---
	comboio eléctrico	trem	---
	metropolitano (metro)	Bonde (em desuso) metrô	metropolitana (usado para região)
	mota/moto	moto	
vestuário	camisola	camiseta,	camisola (roupa feminina para dormir)
	cueca(s)	calcinha(s) feminina(s)	cueca (peça íntima masculina)
	calções	bermuda	---
	fato	terno	fato (acontecimento; bucho)
	fato de banho	maiô	---
	gabardine	capa-dechuva	---
	peúga	meia de homem ou soquete	---
aldeia e os trabalhos do campo	aldeão	roceiro, homem do campo, da roça	---
	apanha	colheita,	---
	ceifa	colheita, roçado,	---
	fonte	mina, minadeira,	fonte (local onde brota água do solo)
	monda	poda	---
	quinta	sítio, chácara,	---
	lareira	pedra do fogão de lenha	---
higiene pessoal	casa de banho	banheiro	---
	desodorizante	desodorante,	---
	pasta dentríoca	pasta de dentes	---
	duche	chuveiro	---

Fonte: Elaborada pela autora desta tese (2017).

A pesquisa de Biderman revelou, ainda hábitos alimentares diferentes, determinados pela frequência de uso, de modo que enquanto “vinho”, “peixe” e “sopa” aparecem entre as palavras mais frequentes nas pesquisa do português fundamental, inversamente, “arroz”, “feijão” e “carne” se revelam palavras fundamentais na variedade brasileira da língua e, aliás, no cardápio cotidiano do brasileiro, no prato do dia e do dia-a-dia do brasileiro. Acrescenta a autora:

Se acrescentarmos a esses vocábulos, a farinha (de mandioca) e algumas variedades de carne como a carne-seca, também chamada de carne-de-sol no Nordeste do Brasil, teremos outras palavras fundamentais no PB por constituírem a base alimentar de grande maioria dos brasileiros. (BIDERMAN, 2001, p. 971)

Muito embora a ampla e profunda bifurcação lexical, sobretudo se se “colocar mais farinha”, não é menos verdade que uma parte substancial do patrimônio lexical herdado do PE continua vigente no PB.

2.2.1 As diferenças léxico-dialetais do português brasileiro e do português europeu

2.2.1.1 Áreas dialetais em Portugal e no Brasil

A classificação dialetal do português europeu, adotada pela maioria dos estudos linguísticos é a proposta de Lindley Cintra (1971), na qual apresenta três dialetos distintos, caracterizados foneticamente.

Entretanto, para além das características fonéticas, depois de recolher, em 67 localidades espalhadas por Portugal, várias designações acerca uma série de ações, objetos ou seres relacionados com o ambiente da criação de gado, Lindley Cintra (1962) procedeu à sua análise e propôs três tipos de divisão do território português em áreas lexicais.

- a) O *Norte conservador*, com preponderância de vocábulos de origem latina e *sul inovador*, com palavras análogas de proveniência arábica.
- b) *Litoral conservador*, com maior número de arcaísmos, isolados e conservados nas zonas extremas e *interior inovador*, com predomínio de criações novas, vindas gradualmente de Espanha
- c) *Noroeste conservador*, com preponderância de tipos arcaicos latinos – *zonas de transição*, em forma de faixas orientadas do Nordeste ao Sudeste – *centro e sul inovadores*, sendo mais expostos às influências estrangeiras.

No que tange à proposta de divisão dialetal para o território brasileiro, parece sobrelevar-se a proposta de Antenor Nascentes. Embora diversas pesquisas estejam sendo desenvolvidas no intuito de comprovar se “tinha Nascentes razão”,

parece não haver, ainda, uma proposta de divisão do território brasileiro em áreas lexicais. Quiçá se se “colocar mais farinha”.

Machado Filho e Neiva (no prelo) acreditam que

macaxeira, *aipim* e *mandioca* poderiam ser utilizados para caracterizar isoléxicas nacionais, dividindo o país em três grandes áreas de predominância lexical. A *macaxeira* situar-se-ia no Norte e no Nordeste, a *mandioca* ocuparia parte das regiões Sudeste e seria característica do Centro-Oeste, enquanto o *aipim*, seria uma variante do Sul, em distribuição pelo litoral do Sudeste e parcialmente do Nordeste, tendo Porto Alegre, Florianópolis e Salvador, como expoentes isolexicais. (MACHADO FILHO; NEIVA, no prelo)

CONSIDERAÇÕES DO CAPÍTULO

“Pra chegar até aqui”, um longo caminho e sinuoso foi trilhado.

Este capítulo buscou refletir / refazer os passos que fizeram o caminho, inclusive mostrando como surgiu a bifurcação que constituíram as veredas da língua portuguesa: o português europeu e o português brasileiro. Para tanto, centrou-se no objeto observacional da pesquisa, de forma a tornar perceptível como o objeto teórico desta pesquisa se comporta, na história.

Uma primeira motivação para a construção deste capítulo parte do interesse em observar o processo de constituição do léxico português e, a partir daí, evidenciar as razões sócio-históricas e linguístico-culturais que fizeram do português brasileiro ser como é, diversificado e diverso do português europeu. Com o desenvolvimento do capítulo, tornou-se ainda mais pertinente evidenciar que e como o contato linguístico e as influências culturais, incluindo os estrangeirismos podem ter influenciado na diferenciação do léxico do PB e do PE.

O capítulo mostrou, ainda, que em relação à formação do léxico português, em Portugal, há uma vasta produção rica em detalhes e que, em contrapartida, no que tange à formação do léxico do português brasileiro – não obstante as muitas propostas de periodização –, ainda falta literatura.

Existem estudos específicos acerca de brasileirismos e estrangeirismos, tupinismos e africanismos mas estes não levam em consideração a formação do português brasileiro. Inversamente, há propostas acerca da formação do léxico português no Brasil mas sem a acuidade de especificar as línguas.

A síntese da estruturação do léxico com base em campos conceituais que permeou todo o capítulo foi assim realizada com vistas ao interesse da pesquisa que se concentra em áreas temáticas. Com base no que fora apreendido, a análise dos dados do *corpus* utilizado na pesquisa para o desenvolvimento do Vocabulário proposto mostrou que alguns dos itens lexicais citados nesse capítulo fazem parte do léxico do português brasileiro contemporâneo, mais especificamente, do léxico baiano.

Capítulo 3 OS ESTUDOS DO(S) LÉXICO(S) NO BRASIL

Na *Introdução* desta tese, foi dito que o léxico tem sido estudado pela Lexicografia contemporânea com um novo viés. Aqui, são trazidas notícias acerca dos léxicos e seus respectivos estudos.

Considerando ser este um trabalho de cunho lexicográfico, para iniciar a reflexão, buscou-se observar como os dicionários de língua atuais definem 'léxico'.

Em uma primeira consulta ao Aulete (2011), encontraram-se sete acepções:

lé.xi.co [cs]

a.

1. Ling. Próprio das palavras ou referente a elas; LEXICAL

sm.

2. Ling. O repertório de palavras de uma língua ou de um texto; VOCABULÁRIO

3. P.ext. Obra de compilação de uma parte (reduzida ou extensa) dos vocábulos de uma língua e seus significados; DICIONÁRIO

4. Dicionário de antigas línguas clássicas.

5. P.ext. Relação de palavras us. por um autor ou por uma escola ou movimento literário.

6. Ling. Conjunto dos lexemas da língua (proposto por Saussure), oposto ao conjunto de vocábulos.

7. P.ext. Gram. Componente da gramática internalizada de um falante que abarca todo o seu conhecimento das palavras (esp. sua pronúncia, significação e emprego numa sentença).

[F.: Do gr. *leksikós*, e, *on*.]

(AULETE, 2011, p. 853)

Talvez por estarem em maiúsculas, para além das definições, chamaram a atenção a ênfase dada às “palavras” ‘lexical’, ‘vocabulário’ e ‘dicionário’. Como consulente curioso, parte-se em busca da definição dada a estas palavras.

Encontra-se:

le.xi.cal [cs]

a2g.

1. Ling. Ver [léxico](#) .

2. Ling. Ref. a palavra; VOCABULAR

3. Ling. Que simboliza a delimitação do ambiente biossocial feita por certa língua (morfema lexical).

[Pl.: -cais.]

[F.: *léxi(o)* + *-al*.]

(AULETE, 2011, p. 853)

Como se nota, a primeira acepção é uma remissão – “ação ou resultado de encaminhar a outro ponto ou lugar”. No caso, considerando a consulta anterior, feita a “léxico”, não seria um encaminhamento a outro ponto mas, um retorno.

Continuando a saga, ‘vocabulário’ é a próxima parada obrigatória, em que se encontram 10 acepções:

vo.ca.bu.lá.ri.o

sm.

1. Conjunto dos vocábulos de uma língua; LÉXICO
2. Conjunto dos termos característicos de uma atividade ou campo do conhecimento (vocabulário do futebol; vocabulário psicanalítico); NOMENCLATURA; TERMINOLOGIA
3. Livro que contém o conjunto desses termos, em ordem alfabética e com as respectivas definições (vocabulário de filosofia; vocabulário de informática); DICIONÁRIO; GLOSSÁRIO
4. Relação em ordem alfabética e com as respectivas definições desses termos ou de termos pouco comuns ou regionais etc., us. em uma obra, e que constitui um apêndice dela; ELUCIDÁRIO; GLOSSÁRIO
5. Conjunto das palavras us. em determinada época ou estágio da língua: *o vocabulário medieval português.*
6. Conjunto das palavras us. por um autor em sua obra, por uma faixa etária, um grupo social ou de interesses etc. (vocabulário camoniano; vocabulário jovem; vocabulário dos sambistas)
7. Conjunto das palavras que uma pessoa conhece: *Ele tem bom vocabulário, lê muito.*
8. Dicionário de particularidades das palavras, que não inclui necessariamente definição de acepções (vocabulário ortográfico; vocabulário etimológico)
9. Inf. Conjunto dos símbolos us. para codificação de operações e instruções numa linguagem de programação
10. P.ext. Conjunto de palavras referentes a um tema: "O meu amigo (...) achou que Silvestre algumas vezes abusava do vocabulário dos eufemismos." (Camilo Castelo Branco, *Coração, cabeça, estômago*)
[F.: Do lat. medv. *vocabularium*.]
(AULETE, 2011, p. 1495)

Encontram-se, também, aqui, alguns equívocos lexicológicos e lexicográficos. Dicionário, vocabulário e glossário não são sinônimos. Segundo Vilela (1994, p. 13-14),

o léxico é o conjunto das palavras fundamentais, das palavras ideias duma língua; o vocabulário é o conjunto dos vocábulos realmente existentes num determinado lugar e num determinado tempo, tempo e lugar ocupados por uma comunidade linguística; **o léxico é o geral, o social e o essencial; o vocabulário é o particular, o individual e o acessório.** Há ainda uma outra perspectiva, a de ‘coleção de unidades’, em que o vocabulário se opõe a dicionário e glossário: **o dicionário é a recolha ordenada dos vocábulos duma língua, o vocabulário é a recolha de um sector determinado duma língua e o glossário é o vocabulário difícil de um autor, de uma escola ou de uma época.** (VILELA, 1994, p.13-14)

Em uma rápida busca à versão eletrônica Houaiss (2009), foram encontradas as acepções similares. Desta feita, porém, chama a atenção, não mais a definição ou as acepções, a lista de 20 unidades lexicais com a mesma raiz (radical) *lex_* (lexema, lexemático, lexia, lexical, lexicalização, lexicalizar, léxico, lexicogênico, lexicografar, lexicografia, lexicográfico, lexicógrafo, lexicologia, lexicológico, lexicólogo, léxicon, lexicogênico, lexiologia, lexiológico, lexis).

A partir dessas unidades lexicais, pois, este capítulo é construído visando preencher algumas lacunas como, por exemplo, a caracterização dos léxicos e, por conseguinte, de seus estudos.

3.1 CARACTERIZAÇÃO DOS ESTUDOS (CIENTÍFICOS) DOS LÉXICOS

As definições de léxico encontradas nos dicionários, já citadas, se revelam insuficientes e, por vezes, inadequadas ou equivocadas. Nada obstante, para uma reflexão acurada e relevante ao objetivo desta tese, toma-se como ponto de partida a definição de léxico encontrada em ambos os dicionários como “repertório de palavras de uma língua”. Mas, que “palavras”? E como o léxico é caracterizado cientificamente?

O esquema, a seguir, proposto por Aitchison (1978, p.14) e adaptado por Nunes (2010, p. 88) revela o descaso com que muitos estudos linguísticos, ao longo dos tempos, tratou o léxico como nível de análise da língua.

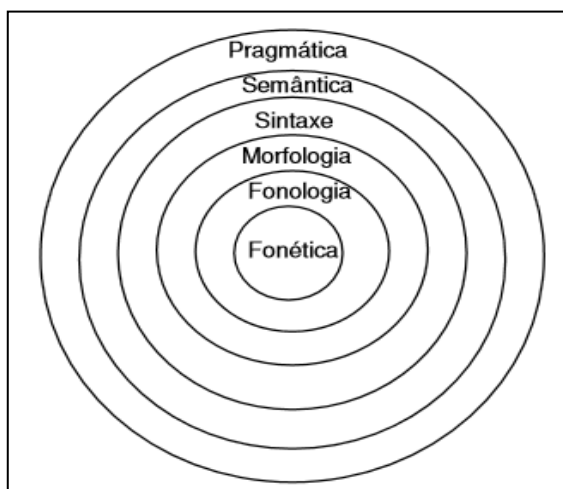


FIGURA 7 – Diagrama redutor dos níveis de análise da língua

Fonte: Aitchison (1978) *apud* Nunes (2010).

Ao observar o esquema, talvez, a indignação se traduzisse em uma pergunta: “Onde está o Wally?”

Para além de pensar a relevância da entrada dos estrangeirismos e dos decalques na língua, sobretudo se se pensar na dinâmica lexical, de como expressões passam a pertencer ao léxico comum, em determinado momento, podendo dispersar-se, evadir-se ou fixar-se na língua, a frase sugerida só pode ser compreendida por aqueles que, por experiência ou por conhecimento prévio, enciclopédico e de mundo, tiverem acesso à história do personagem. Eis o léxico “Wally” dentre os estudos linguísticos: um rosto perdido na multidão dos níveis de análise da língua.

Lewis (1993) além de cunhar o termo *Lexical Approach* – Abordagem Lexical –, coloca em xeque uma das mais sólidas bases do ensino de língua, sobretudo de língua estrangeira: a Gramática.

Em toda a sua obra permeia-se a ideia de que “a língua consiste de léxico gramaticalizado, não de gramática lexicalizada”, ou seja, de que a base de uma língua não é sua a estrutura gramatical mas o seu Léxico, leia-se vocabulário.

Segundo Machado Filho (2010, p. 51), o léxico é um “movediço alicerce que serve de base às variegadas gramáticas das línguas naturais” sob o qual “a Lexicografia contemporânea tem caminhado no sentido de perscrutar, a passos largos e de forma incontestavelmente sólida”.

Assevera Vilela (1994, p. 14) que no léxico “refletem mais clara e imediatamente todas as mudanças ou inovações políticas, econômicas, sociais, culturais ou científicas” de modo a poder ser considerado, pois, como “o subsistema da língua mais dinâmico” ou, para Machado Filho (2014), “um sistema multidinâmico”.

Já admitia Fernão de Oliveira que as dições,

[...] todas ellas ou são geraes a todos, como *Deos, pão, vinho, ceo e terra*, ou são particulares: e esta particularidade ou se faz antre officios e tratos, como os cavaleiros que têm huns vocabolos e os lavradores outros, e os cortesãos outros, e os religiosos outros, e os mecanicos outros, e os mercadores outros; ou também se faz em terras esta particularidade, porque os da Beira têm hũas falas e os d’Alentejo outras. E os homens da Estremadura são diferentes dos d’Antre Douro e Minho, porque assi como os tempos, assi também as terras criam diversas condições e conceitos. E o velho, como tem o entender mais firme com o que mais sabe, também suas falas são de peso e as do mancebo mais leves.

(Oliveira, 2000 [1536], cap. XXXVIII, p. 131 [52])

Numa tentativa de atualização da nomenclatura utilizada por Fernão de Oliveira, é possível considerar dois principais tipos de léxicos: o léxico geral – as ditas dições gerais, comuns a todos – e o léxico especializado – constituído das “dições particulares” as quais diferem quanto à especificidade das profissões – léxico especializado com o uso de termos técnicos, jargões profissionais –, à diatopia, à diacronia – que propiciam os arcaísmos e neologismos – e à diageracionalidade.

Em breve síntese, observe-se a tabela, a seguir:

Tabela 11 – Léxico geral versus Léxico especializado: algumas características

Léxico geral	Léxico específico / especializado
Lexia	Termo
Inclui, além das “palavras lexicais”, as unidades mínimas significativas da língua, as quais se costuma chamar de “palavras vazias” ou “palavras gramaticais”	Uma lexia específica dentro de um domínio conceitual específico
Significado / acepções	Significação, um conceito, uma noção precisa
Aquisição: por frequência de uso, aprendida inconscientemente. Usa-se uma lexia para descrever um objeto (material ou imaterial) do mundo Uso geral, comum a todos os falantes de uma língua.	Aquisição: conscientemente. Usa-se um termo para nomear uma entidade Uso específico de um grupo científico, terminólogos, Uso especialistas na terminologia de um campo

Fonte: Elaborada pela autora desta tese (2017).

É válido ressaltar, entretanto, que, assim como é tudo relacionado à vida humana, os léxicos também não são estanques, inversamente, fazem parte de um *continuum*, de modo a tornar possível o trânsito entre o geral e o específico e entre domínios distintos. Por outra via, entre domínios distintos um mesmo termo pode ter várias noções de significado.

Para melhor elucidar, tome-se como exemplo, a lexia ‘base’ registrada na versão eletrônica do Dicionário Houaiss (2009), em que se apresentam 38 usos distintos, dos quais 6 são acepções de “parte inferior de alguma coisa, considerada como seu suporte”

Tabela 12 – A produtividade da lexia ‘base’ na língua portuguesa: principais acepções

processo formativo	acepção	exemplo
Derivação: por extensão de sentido	numa superfície, primeira camada de massa ou tinta, sobre a qual se aplicam outras, destinadas a dar acabamento	<i>b. para maquiagem para pintura de portas</i>
Derivação: por metonímia.	o produto apropriado para fazer essa base	
Derivação: por metáfora	parte ou aspecto essencial de alguma coisa; princípio, origem	<i>estabelecer as b. de um acordo</i>
Derivação: por extensão de sentido.	conjunto de conhecimentos, fatos, dados, de que se dispõe para opinar, acusar etc.	<i>você não tem b. para dizer isso</i>
Derivação: por extensão de sentido	ingrediente principal de uma mistura	<i>o cacau é a b. do chocolate</i>
Derivação: por analogia	parte de um órgão pela qual ele se liga a um outro mais central na estrutura do organismo	<i>b. do polegar, da língua etc.</i>

Fonte: Houaiss (2009). A tabela foi elaborada pela autora desta tese (2017).

O Houaiss, traz, ainda, 12 rubricas em que se utiliza ‘base’ como termo, como léxico especializado.

Tabela 13 – A produtividade da lexia ‘base’ na língua portuguesa: acepções por rubricas

rubrica	acepção
construção	m.q. alicerce ('maciço de alvenaria')
arquitetura	numa coluna, a parte inferior, sobre a qual descansa diretamente o fuste; pé, pedestal
farmácia	ingrediente ger. inativo de uma preparação farmacêutica que serve como veículo de um ou mais princípios ativos; excipiente Ex.: os <i>emolientes</i> são as <i>b. das pomadas</i>
geometria	lado de um polígono ou face de um sólido que se considera embaixo
gramática gerativa	parte do componente sintático que define as estruturas básicas das orações de uma língua, sendo constituída do léxico e de categorias sintáticas e regras de reescrita (sintagmáticas), que expandem as categorias sintáticas em sintagmas, e introduzem, nos sintagmas, os itens lexicais
matemática	razão entre a unidade de uma determinada ordem e a unidade da ordem imediatamente inferior, num sistema de numeração (o sistema decimal, p.ex., utiliza a base 10)
	número que, numa potência, representa o fator que é multiplicado por si mesmo
artes gráficas	m.q. bloco de gaveta
música	nota tônica
petrologia	pasta amorfa ou cristalina que envolve os fenocristais das rochas porfiríticas; massa fundamental, pasta
química	substância ou íon capaz de reagir com um próton (íon hidrogênio) para formar um novo composto; substância que reage com (ou neutraliza) ácidos para formar sal e água
topografia	distância criteriosamente medida no terreno para a determinação topográfica e geodésica de outras distâncias

Fonte: Houaiss (2009). A tabela foi elaborada pela autora desta tese (2017).

Além dessas, ‘base’ também é usada na formação de compostos e fraseologismos, como ocorre em:

Tabela 14 - A produtividade da lexia 'base' na língua portuguesa: formação de novas lexias

lexia	rubrica	acepção
b. aérea	termo aeronáutico	base militar em que se concentram aeronaves, equipamentos e contingentes de uma força aérea
b. avançada	termo militar.	base militar provisória, que se localiza em área avançada do campo de operações e tem a função de apoiar as unidades envolvidas nas operações em curso
b. conjugada	química.	produto da reação de um ácido com um solvente juntamente com o ácido conjugado correspondente
b. de dados	informática.	m.q. banco de dados
b. de lançamento	astronáutica, termo militar.	instalação ou conjunto de instalações de onde são lançados mísseis e/ou foguetes espaciais
b. de operações	termo militar.	acampamento militar onde são planejadas as ações ofensivas e para onde os soldados retornam, caso a missão fracasse
b. do crânio	anatomia geral.	face inferior do crânio caracterizada por vários forâmens (sendo o mais importante o forâmen magno) e áreas delgadas de osso
b. espacial	astronáutica.	tipo de satélite artificial, de grande porte e amplos recursos, que serve de apoio à navegação no espaço; base orbital
b. forte	química.	base com alta tendência à dissociação
b. fraca	química.	base com baixa tendência à dissociação
b. militar	termo militar.	lugar de concentração de tropas e de meios materiais para a condução de operações militares
b. monetária	economia.	quantidade de dinheiro em papel-moeda existente num país
b. naval	termo de marinha.	base militar da marinha de guerra, compreendendo um conjunto de estabelecimentos, instalações e serviços estrategicamente localizados, para poder fornecer apoio a forças navais
b. nitrogenada	bioquímica, genética.	cada uma das cinco moléculas que compõem os diferentes <i>nucleotídios</i> , a saber: citosina, guanina, adenina, timina, uracila
b. orbital	astronáutica.	m.q. base espacial
b. ortogonal	álgebra.	base de um espaço vetorial cujos vetores são ortogonais
b. ortonormal	álgebra	base ortogonal cujos vetores são unitários
b. vetorial	matemática, álgebra.	conjunto de vetores linearmente independentes que gera um dado espaço vetorial

Fonte: Houaiss (2009). A tabela foi elaborada pela autora desta tese (2017).

Ainda, além das 18 lexias supracitadas, nota-se a incorporação de 'base' à gramática da língua, *i.e.*, gramaticalizada de modo a constituir conectivos, como ocorre em: i) 'de base' (ex.: material de base para análise); ii) 'com base em', 'à base de' (ex.: novas técnicas elaboradas à base das recentes descobertas) e estas com outras acepções e usos.

E, inclusive, 'base' se configura como expressão idiomática, 'tremar nas base' – ter muito medo ou receio; intimidar-se, apavorar-se –, que no referido dicionário (HOUAISS, 2009) é apresentada regionalismo: Brasil .

Como se pode observar, para além da produtividade de ‘base’ na língua portuguesa, parece ser bastante claro que e como o léxico se divide em léxico geral e léxico especializado. No entanto, é preocupação deste trabalho salientar o léxico variacional, que, talvez, por assim ser, seja também estigmatizado como alguns de seus utentes.

Por ora, apresentam-se os estudos que têm sido feitos, no Brasil, com vistas à observação do léxico variacional, quer seja especializado quer seja geral, quer transitório entre ambas as peculiaridades.

3.1.1 Tipos de estudos dos léxicos

Atualmente, no Brasil, dentre as vertentes teóricas da linguística contemporânea que têm adotado o estudo do léxico, destacam-se a Lexicologia e a Lexicografia.

Em relação à Lexicologia, Biderman (2001, p. 16) admite que possua “como objetivos básicos de estudo e análise a palavra, a categorização lexical e a estruturação do léxico”. Sob mesma perspectiva, Pontes (2009, p. 18) considera ser “a disciplina responsável pelo estudo das palavras de uma língua, em discursos individuais e coletivos”. Segundo Krieger (2006, p. 159), entretanto, a definição consensual de Lexicologia, como “o estudo científico do léxico”, “não é muito esclarecedora em si mesma”.

Em *História concisa da Linguística*, Weedwood (2002, p. 11) propõe o seguinte esquema:

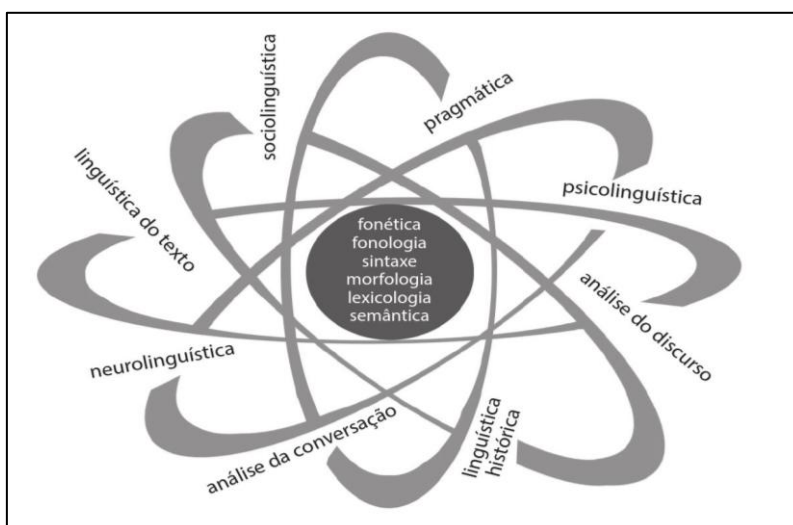


Figura 1: Micro e macrolinguística
Fonte: (WEEDWOOD, 2002, p. 11)

Embora não se pretenda discorrer, nesta tese, acerca da estruturação das ciências – micro ou macro – no âmbito da linguística, é relevante, ao menos, propor uma reflexão acerca da proposta de hierarquização das ciências em função dos seus respectivos objetos, sobretudo no que tange à microlinguística, onde está situada a lexicologia.

No cerne da Lexicologia, outra questão também relevante refere-se à possibilidade de estabelecimento de fronteiras e interfaces com outras áreas e disciplinas resultando estudos dialetológicos, sociolinguísticos, etnolinguísticos, além dos que se centram na Semântica e na Morfologia lexical.

Dentre os estudos que abordam a variação lexical, destacam-se, aqui, os realizados no âmbito da Dialectologia e da Lexicografia, com vistas à temática da tese: a elaboração de um produto dialetológico de cunho lexicográfico ou lexicográfico de cunho dialetológico, resultante de uma proposta de interface entre ambas as áreas.

3.1.1.1 *Os caminhos dos estudos dialetais no Brasil*

Adeus, Itaparica do meu coração, adeus, raízes que restarão somente num muro despencado ou outro, no gorgueio aflito de um sabiá sobrevivente, no adro de alguma igreja venerável por milagre preservada, **na fala, daqui a pouco perdida, de meus conterrâneos da contracosta.** (RIBEIRO, 2010²⁴, grifo nosso)

Chama a atenção no excerto supracitado a ideia de que a fala pode se perder. Eis a preocupação em documentar a fala, a fim de que estudos posteriores revelem as raízes. Muito embora a “Ilha” não faça parte do *corpus* analisado para a construção desta tese, o excerto se mostra relevante por aludir à importância da pesquisa dialetal, sobretudo como resgate ou preservação de raízes linguísticas.

Aliás, essa mesma constatação já havia sido feita por Vasconcelos (1987, p. 175).

É urgente, pois, recolher e estudar os materiais dialetológicos, antes que os dialetos deixem de ser falamos. Os principais e os mais relevantes não tardarão em desaparecer e muito em breve passaremos a conhecer deles apenas alguns fragmentos (VASCONCELOS, 1987, p. 175)

²⁴ O artigo “Adeus, Itaparica”, de João Ubaldo Ribeiro, foi publicado no espaço de Opinião, do Jornal A Tarde, em 22.01.2010.

De acordo com Cardoso (2010, p. 15), “a Dialetologia é um ramo de estudos linguísticos que tem por tarefa identificar, descrever e situar os diferentes usos em que uma língua se diversifica, conforme a sua distribuição espacial, sociocultural e cronológica”. Para tanto, tem, como pressuposto metodológico, a Geolinguística ou Geografia Linguística, cujo nascimento

como disciplina autônoma está associado à elaboração do ‘Atlas linguistique de la France’ (ALF) (1902-1910) de J. Gilliéron e E. Edmont, que embora surja na seqüência de outros trabalhos desta natureza, é o primeiro atlas linguístico a orientar-se pelos critérios mais rigorosos que esta disciplina veio a optar. A Geografia Linguística é, desde então, entendida como o estudo cartográfico dos dialetos (FERREIRA, 1996, p. 484)

e, a partir da qual, além da descrição e mapeamento da variação, um aspecto essencial das relações entre vida social e cultural do homem em seu ambiente natural é manifesto.

De acordo com Cardoso (2010, p. 68),

o começo da geolinguística está, assim, marcado pela busca da realidade nacional, entendida como a descrição linguística de área que, geográfica ou politicamente, se reveste de unidade.

No âmbito dos estudos geolinguísticos, no Brasil, a mais conhecida proposta de ordenação dos estudos dialetais é a de Antenor Nascentes (1953), que além de elucidar e instigar outros estudos a respeito do conhecimento da realidade propôs a delimitação de áreas dialetais.

Cardoso e Ferreira (1994) apresentam a história da dialetologia brasileira dividida em etapas ou fases²⁵: i) as primeira e segunda fases – caracterizadas, sobretudo, pela produção e publicação de vocabulários, glossários, dicionários e léxicos regionais, dos quais as obras *O dialeto caipira* (AMARAL, 1920) e *O linguajar carioca* (NASCENTES, 1922); ii) as terceira e quarta fases, caracterizadas, respectivamente pelo início da consolidação dos trabalhos em Geolinguística no Brasil, com a publicação do primeiro trabalho em geografia linguística realizado no Brasil – *O Atlas Prévio dos Falares Baianos* – APFB (1963) e ampliação do número

²⁵ Antenor Nascentes (1952) havia apresentado as fases dos estudos dialetais do Brasil, subdivididas em duas, as quais foram reformuladas por Cardoso e Ferreira em 1994.

de atlas linguísticos publicados ou em elaboração e pela incorporação de novas dimensões²⁶ e fusão com outras áreas afins ao trabalho e à pesquisa dialetológica.

Observa-se, pois, que os trabalhos lexicográficos foram ao longo dos tempos perdido espaço para as construção de atlas linguísticos e de outros trabalhos lexicológicos. No entanto, é válido ressaltar a importância de ambos para os estudos da variação lexical, mais especificamente, a variação diatópica do léxico.

Considerando a extrema importância de um atlas linguístico para a pesquisa linguística variacionista, pelo enorme acervo que fornecem, além de salvaguardar a memória sociolinguística de um povo, este pode ser considerado um instrumento para as políticas linguísticas, sobretudo no que tange ao ensino-aprendizagem.

Na história da Dialetologia brasileira, publicações de atlas regionais e estaduais foram fundamentais para a consolidação dessa área e para a elaboração do Atlas Linguístico do Brasil. Convém, então, rever o caminho.

3.1.1.1.1 Atlas linguísticos brasileiros: do APFB ao ALiB

A necessidade de aprofundar o conhecimento de uma dada região, proporcionado pelos atlas nacionais, motivou o aparecimento de atlas regionais que, como a própria denominação explicita, se destinam ao exame de áreas menores, buscando detalhar o conhecimento de regiões específicas, fazendo com que o que se deveria ter constituído na primeira geração de atlas linguísticos – a produção de atlas regionais –, viesse a aparecer num segundo momento e quando já circulavam os atlas nacionais. (CARDOSO, 2010, p. 68)

Como se sabe, a efetivação da geolinguística no Brasil ocorre com a publicação do *Atlas Prévio dos Falares Baianos (APFB)*, desenvolvido por Nelson Rossi *et al*, em 1963, tendo como principal objetivo depreender a variação diatópica em 50 localidades do estado. Desde então até o presente momento, foram desenvolvidos 16 Atlas Linguísticos regionais e estaduais, dentre os quais se incluem os trabalhos organizados e apresentados como tese.

Mais de cinquenta anos depois, em 2014, publica-se o *Atlas Linguístico do Brasil (ALiB)* coordenado por integrantes do projeto anterior, com uma vasta rede de pontos constituída de 250 localidades.

²⁶ Dialetologia monodimensional ou tradicional sendo substituída por visão mais ampliada dialetologia pluridimensional, que incorpora aspectos diastráticos, diagenéricos, diageracionais e diafásicos.

Embora os critérios para a seleção dos informantes para a elaboração de atlas linguísticos, tenham anteriormente seguido os ditames da Dialectologia tradicional, uma das características da Geolinguística atual é pautar-se na observação de parâmetros sociais em concomitância com a distribuição espacial, buscando perceber e revelar variáveis sociais, dentre as quais se destacam a escolaridade, o sexo e a faixa etária. Assim, os atlas que contemplam apenas a distribuição espacial são denominados monodimensionais e os que apresentam parâmetros sociais juntamente com a distribuição diatópica, de pluridimensionais.

Uma possível identificação dos Atlas pauta-se nos seguintes critérios concernentes à metodologia e à apresentação dos dados. No âmbito das informações metodológicas, destacam-se a seleção de informantes e a constituição de uma rede de pontos / localidades.

Os critérios para a seleção dos informantes para a elaboração de Atlas Linguísticos podem seguir os da Dialectologia Tradicional, apresentados por Nascentes, Silva Neto e Pedro Caruso: a) ter idade entre 30 e 60 anos; b) ser analfabeto ou semi-analfabetizado; c) ter nascido ou ter vivido, pelo menos, três quartos de sua vida na localidade. Há, porém, Atlas que contemplam, hoje, além da diatopia, alguns parâmetros sociais. Neste caso, apenas o item 'c', supracitado, ainda prevalece.

O estabelecimento de redes de pontos tem como objetivo depreender a variação diatópica levando em consideração a densidade demográfica de cada região e a distribuição espacial das localidades, com os ajustes necessários, considerando-se, para tanto, essencialmente, a importância das localidades na área estudada, sociolinguisticamente representativas e os limites interestaduais e internacionais. O Quadro-resumo, a seguir, mostra os critérios adotados pelos Atlas brasileiros publicados até a presente data.

QUADRO 7 – CRITÉRIOS DE CLASSIFICAÇÃO DOS ATLAS LINGÜÍSTICOS QUANTO À COLETA DOS DADOS

ATLAS BRASILEIROS (publicados)	REDE DE PONTOS / LOCALIDADES	PERFIL DOS INFORMANTES	NÚMERO DE QUESTÕES	NATUREZA DAS QUESTÕES ÁREAS TEMÁTICAS	OBSERVAÇÕES
APFB (1963)	50 localidades, em 16 zonas fisiográficas do Estado	100 informantes (57 mulheres e 43 homens), analfabetos ou semi-analfabetos, com idade variando entre 25 e 60 anos. * Seleção não sistemática dos informantes	182 perguntas	terra, vegetais, homem, animais.	Aplicação de teste de reconhecimento ou teste de identificação.
EALMG (1977)	116 localidades, <i>in loco</i> , e 302 pontos, por correspondência.	83 informantes, homens e mulheres, analfabetos, com idade entre 30 e 50 anos. * Seleção não sistemática dos informantes	415 perguntas	terra e folgedos infantis.	
ALPB (1984)	25 localidades base e 3 localidades satélite para cada base, as quais não constam da cartografia	107 informantes (três a dez por localidade), homens e mulheres, com faixa etária variando entre 30 a 75 anos e grau de instrução de analfabetos até 4ª série do Ensino Fundamental. * Seleção não sistemática dos informantes	parte 1 (geral) 289 questões parte 2 (específico) 588 questões	terra, homem, família, habitação e utensílios domésticos, aves e animais, plantação e atividades sociais perguntas concernentes aos cinco principais produtos agrícolas do estado: mandioca, cana-deaçúcar, agave, algodão e abacaxi.	
ALS I (1987) *foi concluído em 1973	15 localidades	30 informantes: (um homem e uma mulher, por localidade), com idade entre 35 e 53 anos. destes: 21 analfabetos, 1 alfabetizados e 8 “semianalfabetos”	686 questões 181 retiradas do Extrato de Questionário aplicado para o APFB e 505 selecionados dos questionários	terra, vegetais, homem, animais.	Constitui-se em um passo a mais para o mapeamento linguístico da área dos falares baianos. A equipe de dialetólogos da UFBA pode, em Sergipe, aperfeiçoar os instrumentos metodológicos utilizados na recolha do APFB.

ATLAS BRASILEIROS (publicados)	REDE DE PONTOS / LOCALIDADES	PERFIL DOS INFORMANTES	NÚMERO DE QUESTÕES	NATUREZA DAS QUESTÕES ÁREAS TEMÁTICAS	OBSERVAÇÕES
ALPR (1994)	65 localidades, que contemplavam todas as 24 microrregiões fisiográficas paranaenses.	130 informantes rigorosamente distribuídos por ambos os sexos, com idade entre 27 e 62 anos e escolaridade variando entre analfabetos e de primário completo.	325 questões	<i>terra</i> (natureza, fenômenos atmosféricos, astros e tempo/ flora/ plantas medicinais/ fauna) e <i>homem</i> (partes do corpo/ funções/ doenças/ vestuário e calçados/ agricultura/ instrumentos agrícolas/ brinquedos e jogos)	Além da documentação cartográfica da variação lexical e da variação fonética e a delimitação de isoglossas apresenta também um glossário.
ALERS (2002)	275 localidades, sendo 100 no Paraná, 80 em Santa Catarina e 95 no Rio Grande do Sul.	os informantes tem idade entre 28 e 58 anos e pouca escolaridade, sendo 2 por localidade nas áreas rurais e 3 nas áreas urbanas.	735 questões:	50 questões no Questionário Fonético -fonológico (QFF), 75 no Questionário Morfossintático (QMS) 610 no Questionário Semântico-lexical.	o primeiro a contemplar uma região político-geográfica do país (atlas regional). Abrange aspectos linguísticos e culturais referentes aos três estados da região sul do Brasil.
ALISPA (2004)	10 localidades	40 informantes, sendo 4 informantes por cidade, estratificados por sexo, idade e escolaridade até a 4ª série.	159 perguntas	de natureza fonético-fonológica	o primeiro atlas falante (sonoro) brasileiro. É apresentado apenas em versão CDROOM. É o primeiro atlas com dados da região norte do
ALS II (2005)	15 localidades	30 informantes: (um homem e uma mulher, por localidade), com idade entre 35 e 53 anos. destes: 21 analfabetos, 1 alfabetizados e 8 "semianalfabetos"	686 questões 181 retiradas do Extrato de Questionário aplicado para o APFB e 505 selecionados dos questionários preliminares.	homem	Possui resultados cartográficos numa perspectiva pluridimensional e tem comentários às cartas. É um atlas de 2ª geração.

ATLAS BRASILEIROS (publicados)	REDE DE PONTOS / LOCALIDADES	PERFIL DOS INFORMANTES	NÚMERO DE QUESTÕES	NATUREZA DAS QUESTÕES ÁREAS TEMÁTICAS	OBSERVAÇÕES
ALMS (2007)	32 localidades	128 informantes: quatro por localidade, pertencentes a ambos os sexos, de escolaridade variando entre analfabetos e pessoas com fundamental	557 questões.	dois tipos de questionários: Fonético-fonológico e lexical	É o primeiro atlas que apresenta dados da região centro-oeste do país. É composto de cartas fonéticas, semântico-lexicais e morfossintáticas.
(ALECE) – 2010	70 localidades	280 informantes: quatro por localidade, pertencentes a ambos os sexos, de escolaridade variando entre analfabetos e pessoas com 1º grau completo e idade variando entre 35 e 53 anos.	306 questões	(semântico-lexical).	É composto de 2 volumes: um introdutório e o outro com cartogramas (lexicais e fonéticos) e glossário.

Fonte: Elaborado pela autora desta tese (2017)

Pode-se observar que a rede de pontos de cada um dos Atlas varia de acordo com a demografia do estado.

No que tange ao perfil dos informantes, nota-se que ao longo dos tempos, passou-se a priorizar a sistematização. O ALS I é o primeiro Atlas a fazer isso. O rigor dos critérios é consolidado com a publicação do ALPR (1994).

Quanto ao tipo de coleta de dados, todos utilizaram questionários. A quantidade e a natureza das questões variam de acordo com os interesses de observação. Para além das questões de natureza léxico-semântica o ALERS (2002), o ALISPA (2004), O ALMS (2007) e o ALECE (2010) apresentam questões elaboradas para fins de observação da variação fonética.

Quanto à apresentação dos dados e à elaboração das cartas linguísticas, os Atlas podem variar de acordo com os critérios adotados. Dentre tais, destacam-se os critérios quanto ao tipo de informação que apresentam, à natureza dos dados que registram, ao nível de estudo da língua que apresentam e à representação dos dados. A presença de glossário também é um critério relevante, sobretudo, a este estudo.

O quadro, a seguir, mostra como cada um dos Atlas brasileiros publicados adota os critérios supracitados.

ATLAS BRASILEIROS (publicados)	CRITÉRIOS DE IDENTIFICAÇÃO DAS CARTAS LINGÜÍSTICAS			PRESENÇA DE GLOSSÁRIO
	Total de cartas / Tipos de cartas	Apresentação das variantes (lexias)	Informações interpretativas	
APFB (1963)	209 / 154 fonéticas e léxicas e 44 cartas resumo.	transcritos no interior da própria carta ou com legendas e símbolos,	Algumas cartas apresentam dados etnográficos, inclusive com ilustrações. Na parte inferior esquerda das cartas há comentários informativos sobre a	não apresenta glossário.
EALMG (1977)	78, 5 de identifica- ção, 21 léxicas, 24 fonéticas, 3 isófonas e 25 cartas isoléxicas.	símbolos no interior do mapa; e legendas na parte extrema esquerda das cartas constam os vocábulos de frequência mínima e, na direita, os vocábulos de alta frequência,	As cartas não trazem comentários informativos nem qualquer tipo de análise do material coletado.	No final do volume há um glossário com 139 verbetes dicionarizados com acepção diferente ou não dicionarizados.
ALPB (1984)	O Atlas é composto de cartas léxicas e cartas fonéticas, intercaladas.	.	As cartas não trazem comentários informativos. O segundo volume traz uma descrição detalhada da metodologia utilizada, os dados histórico-geográficos, geo-econômicos e sócio-culturais das localidades, a ficha dos informantes, os informantes por localidade, a análise das formas e estruturas lingüísticas encontradas sob os aspectos fonético-fonológicos e morfossintáticos..	O atlas apresenta, ainda, um glossário, com 363 verbetes dicionarizados em sentido diferente do uso geral ou não dicionarizados. Cada verbete vem com a transcrição fonética da realização mais freqüente na região e a indicação, entre parênteses, do número da carta onde se encontra
ALS I (1987) *foi concluído em 1973	180 cartas, sendo 11 introdutórias e 169 cartas léxicas	transcrição pormenorizada e numerosos dados etnográficos, tendo em vista a quantidade de notas que acompanham as cartas.	Em cada carta há a remissão à carta correspondente no Atlas Prévio dos Falares Baianos. Há no ALS, ainda, uma série de cartas conjuntas Bahia- Sergipe, com dados da Bahia, não apresentados no APFB.	não apresenta glossário.

<p>ALPR (1994)</p>	<p>cartas de identificação – seis; cartas léxicas – 92; cartas fonéticas – 70; cartas isoléxicas – 19; cartas isófonas – 10; cartas anexas com a distribuição geográfica do povoamento do</p>	<p>Os termos nas cartas léxicas vêm com símbolos coloridos, no interior da carta, com as legendas do lado direito superior, fora da carta, e com as transcrições fonéticas no interior da carta, no caso das cartas fonéticas.</p>	<p>No verso de cada carta há notas explicativas e de análise do material coletado.</p>	
<p>ALERS (2002)</p>	<p>O Atlas contém um total de 176 cartas, sendo 70 de fonética e fonologia, 104 de morfossintaxe e duas cartas auxiliares (microrregiões homogêneas e rede de pontos)</p>		<p>As cartas estão assim organizadas: na parte superior à esquerda vêm o item do questionário e a legenda da carta; na parte superior à direita vem o número da carta. Na parte inferior à direita vêm os gráficos de frequência. Nas páginas da direita vêm as cartas e nas páginas da esquerda vêm o item que constitui a carta, o número de ordem, o símbolo, a variante, a frequência e o ponto.</p>	
<p>ALISPA (2004)</p>	<p>Contém 600 cartas lingüísticas</p>	<p>o Menu a) Entrevistas; b) Informantes; c) Realização; d) Análise Acústica; e) Palavras. Através dos quais pode-se fazer busca para cada item do questionário, além da possibilidade de uma busca automática da variação fonética, em certos contextos, e sua relação com os fatores sociolingüísticos. Pode-se, também, ouvir as entrevistas integrais dos 40 informantes. O CD contém, ainda, alguns dos trabalhos já realizados como Dissertação de Mestrado, utilizando o corpus do ALISPA.</p>		

ALS II (2005)	108, das quais, três são de identificação.	As cartas semântico-lexicais são compostas com legendas coloridas emarcas diferenciadoras da realização masculina e feminina.	Na parte inferior direita há histogramas com percentuais de ocorrência masculina e feminina. No verso estão as variantes fonéticas dos itens lexicais, seguidas de notas e comentários.	
ALMS (2007)				
ALECE (2010)	246 cartas léxicas			

QUADRO 6 - Critérios de classificação dos Atlas linguísticos quanto à apresentação dos dados
 Fonte: Elaborado pela autora desta tese (2017)

3.1.1.1.1.1 O Atlas Linguístico do Brasil

No que se refere ao ALiB – a concretização de um sonho tão sonhado por muitos –, é importante ressaltar, ainda, que se trata de um Projeto de cunho nacional cujo objetivo é o mapeamento do território brasileiro no que concerne aos fenômenos linguísticos, iniciado em 1996, muito embora o sonho e anseio de muitos pesquisadores de construção de um atlas linguístico brasileiro tenham emergido em momento anterior, cuja primeira manifestação em prol de sua elaboração remonta ao ano de 1952:

[...] a intenção de elaborar o atlas lingüístico do Brasil que toma forma de lei através do Decreto 30.643, de 20 de março de 1952, cujo Art.3º., assenta como principal finalidade da Comissão de Filologia da Casa de Rui Barbosa a 'elaboração do atlas lingüístico do Brasil'. Tal determinação foi regulamentada pela Portaria nº536, de 26 de maio de 1952, que, ao baixar instruções referentes à execução do decreto de criação do Centro de Pesquisas Casa de Rui Barbosa, estabeleceu como finalidade principal, entre as pesquisas a serem planejadas, a própria elaboração do atlas lingüístico do Brasil. (FERREIRA; CARDOSO, 1994, p. 44)

Entre os anos de 2001 e 2013, o Projeto ALiB investigou os usos linguísticos em 250 localidades, capitais e municípios de estados brasileiros, selecionadas de acordo com critérios demográficos, sócio-históricos e culturais, a partir da aplicação de questionário específico, o *Questionário ALiB 2001* (COMITÊ, 2001).

No que tange à seleção dos entrevistados, ressalta-se que o Projeto ALiB documentou em cada localidade, através do método de gravação do tipo laboviano, a fala de indivíduos nascidos e residentes na capital, de cujos pais são oriundos, e que possuem pouca mobilidade geográfica.

Os critérios para a rigorosa constituição do perfil dos informantes visaram a apreender, além da diatopia, três variáveis sociais: a variável diagenérica, a diageracional e a diastrática. Para tanto, em todos os 250 pontos, realizou-se a seleção equitativa dos informantes no que tange ao gênero/sexo e às faixas etárias selecionadas – faixa 1, de 18 a 30 anos, e faixa 2, de 50 a 65 anos.

No que tange, no entanto, ao fator escolaridade – fundamental e universitário –, considerando o fato de que a maioria das localidades interioranas não contava, no início das realizações de inquéritos com a presença de faculdades e ou universidades, apenas se documentou a fala de universitários em capitais.

Após a delimitação do número de informantes, os inquéritos foram realizados com base em um questionário específico, o *Questionário ALiB 2001*, em que, além de um texto para leitura, são formuladas perguntas onomasiológicas e semasiológicas que contemplam: i) os níveis da língua em três questionários distintos – o Fonético-Fonológico (QFF), o Semântico-Lexical (QSL) e o Morfossintático (QMS) –, além das Questões de Prosódia; ii) as relações pragmático-discursivas; iii) a interação mais espontânea, a partir de temas semidirigidos e; iv) a concepção do falante a respeito da língua.

Dentre os questionários que contemplam os níveis da língua, o QSL consta de 202 questões divididas em 14 áreas temáticas, a saber: Acidentes geográficos; Fenômenos atmosféricos; Astros e tempo; Atividades agropastoris; Fauna; Corpo humano; Ciclos da vida; Convívio e comportamento social; Religião e crenças; Jogos e diversões infantis; Habitação; Alimentação e cozinha; Vestuário e acessórios e Vida urbana.

Lançados em 2014, os volumes I e II do ALiB (CARDOSO *et al*, 2014) materializam o sonho e o anseio dos pesquisadores, reacendidos por ocasião do Seminário Nacional Caminhos e Perspectivas para a Geolinguística no Brasil, realizado em Salvador, na Universidade Federal da Bahia, em 1996.

Os volumes mostram resultados da pesquisa realizada em 25 capitais do país, exceto Brasília e Tocantins, que não se incluem à pesquisa haja vista os critérios de constituição da rede de pontos.

O Volume 1, Introdução, consta de seis capítulos que recontam a “história” do Atlas Linguístico do Brasil, apresentando a trajetória da geolinguística no Brasil, a metodologia adotada inclusive no que tange à constituição da rede de pontos, a adoção e aplicação de questionário, a seleção dos informantes e à apresentação cartográfica dos resultados, além de uma discussão sobre os problemas metodológicos encontrados durante o desenvolvimento da pesquisa.

O Volume 2 do *Atlas Linguístico do Brasil* (ALiB), publicado em 2014, constitui-se de 159 cartas linguísticas que refletem a variação fonética (em 46 cartas), morfológica (em 7 cartas) e lexical (em 106 cartas) as quais engatam uma nova perspectiva quanto à observação de fatos linguísticos que poderá confirmar ou refutar a proposta de Nascentes, já citada.

Uma das características notáveis consiste na presença de dados estatísticos usados como complemento da cartografia dos resultados em diferentes estratégias e tipos de gráficos.

Para as cartas fonéticas, adotou-se o tipo “coluna” de duas cores que podem representar: i) aspectos linguísticos como a presença / ausência da realização de um fenômeno fônico e (ou) a identificação dessa variação em contextos distintos ou opostos; ii) aspectos sociolinguísticos, com base em uma perspectiva sócio-fonética, como a porcentagem de realizações quanto ao sexo, à faixa etária e (ou) à escolaridade. As porcentagens relacionadas a esses parâmetros geralmente são predefinidas por cinco intervalos de valores fixos.

Já para as cartas léxicas e morfossintáticas, optou-se pelo gráfico “pizza”, com esquema de cores que determinam as porcentagens de uso das variantes.

É válido ressaltar que a legibilidade das cartas linguísticas se configurou um ponto conflitante, sobretudo, em razão dos requisitos básicos impostos pela semiologia gráfica de que a densidade da informação e as diferentes representações adotadas para projetar a simultaneidade de dados estratificados possam interferir no rápido, facilitado e prático processamento da informação e, por conseguinte, da sua interpretação.

Essa preocupação sobre o mapeamento das respostas no Atlas caracterizado por sua multidimensionalidade e sobre a complexidade da leitura de dados foi claramente expressa por Telles e Ribeiro (2014, p. 123) em um dos capítulos constantes do volume 1 de ALiB, momento em que as autoras atestam que a representação pluridimensional, em muitos casos, não justificavam “o risco de uma representação de pouca legibilidade” nem, tampouco, um esforço a mais na interpretação a ser feita pelo leitor.

O Atlas consta ainda de Notas e Índice remissivo que fazem referência muito sintética à motivação léxico-semântica, as quais podem servir para de contributo para o estudo semântico-motivacional, além de consolidarem a perspectiva de uma análise da estratificação social do léxico da comunidade e do repertório de indivíduos.

Conquanto atlas linguísticos caminhem na direção do conhecimento e da valorização da diversidade linguística, seu trabalho tem como produto final a elaboração de cartas linguísticas cuja consulta sugere a informação de outras áreas e cujo acesso requererá conhecimento mais específico. Ressalta-se, pois, sua

utilidade para os lexicógrafos, por exemplo, os quais podem aumentar suas bases de dados a partir das pesquisas dialetais, semasiológica ou onomasiologicamente.

Todas essas observações foram importantes para o estabelecimento da metodologia adotada pelo Vocabulário Dialetal Baiano.

3.1.1.1.1.1 Estudos com base de dados do Projeto ALiB

No que tange aos estudos lexicais, mais especificamente os que tomam por base as áreas temáticas do QSL do Projeto ALiB, tornou-se relevante averiguar o que tem sido feito em perspectiva lexicológica e lexicográfica. Para tanto, como fonte de pesquisa e investigação, foram observadas as informações de trabalhos realizados constantes no site ALiB²⁷ e os que se ao Grupo Nêmesis, mais especificamente, os trabalhos que têm sido desenvolvidos com vistas a contribuir para o Dicionário Dialetal Brasileiro.

Evidentemente, muitos outros trabalhos têm sido realizados com o *corpus* do Projeto ALiB, mas, por questões óbvias de economia de tempo, apenas citam-se, aqui, os que se desenvolveram com base nos dados da Bahia e que foram listados no site do Projeto ALiB.

ÁREA TEMÁTICA	TÍTULO DO TRABALHO	AUTOR (ANO DE PUBLICAÇÃO)
Fenômenos atmosféricos e astros e tempo	O léxico nosso de cada dia na Bahia e no Paraná: acidentes geográficos, fenômenos atmosféricos, astros e tempo	Oliveira (2007)
Ciclos da vida	Menstruação na Bahia: um estudo em dois tempos distintos	Santos (2013)
	Aspectos do léxico relacionado ao ciclo da vida na Bahia e no amazonas	Sousa (em andamento)
	Denominações referentes ao nascimento na bahia e no amazonas	Sousa (andamento)
Religião e crenças	Religiões e crenças na Bahia: aspectos do léxico espelhados nos dados do Projeto ALiB	Oliveira (2016)
jogos e diversões infantis	Brinquedos e brincadeiras infantis na área do falar baiano	Ribeiro (2012)
Alimentação e cozinha	A culinária no <i>corpus</i> do projeto ALiB: revisitando o falar baiano	Silva Santos (em andamento)

QUADRO 7 – Estudos com base em dados do Projeto ALiB

Fonte: Elaborado pela autora desta tese (2017)

²⁷ As informações se referem à consulta feita até dezembro de 2016.

3.1.1.2 Lexicografia

Sobre a tradição dicionarística, Verdelho (2003, p. 417), em *Dicionários: testemunhos da memória linguística*, afirma ter assumido “a tarefa de recolher, registrar e disponibilizar as palavras conhecidas ou testemunhas que afloraram ao discurso”, constituindo um acervo da memória coletiva de uma sociedade através do léxico que por sua vez, segundo Seabra (2006, p. 7) “encontra-se arraigado à história, tradição e costumes de um povo, estando, por isso, em constante processo de expansão, alteração e contração”.

Admitido que essa visão mais clara acerca do léxico tenha possibilitado o aprimoramento do trabalho lexicográfico, Correia (2008, p. 8) destaca a relevância

de dicionários de pendor predominantemente descritivo (e não normativo), dando conta do uso real que é feito da língua, visando a descrição do vocabulário da chamada língua corrente, fornecendo informação pertinente não apenas para a função de compreensão linguística, mas também, e de forma significativa, informação pertinente para a função de produção e para o alargamento do vocabulário do consulente (CORREIA, 2008, p. 8)

Ocorre que, infelizmente, no que tange à língua portuguesa, o desenvolvimento da lexicografia baseada em *corpus* e da lexicografia de aprendizagem, em Portugal e no Brasil, é recente e ainda incipiente.

Isquerdo (2008), num percurso histórico acerca dos estudos lexicográficos, no Brasil, apresenta quatro sincronias relacionadas à história social do país:

- a) séculos XVI, XVII e XVIII, em que se destacam os dicionários jesuítas, em que há registro do universo vocabular da época, e os relatos de viagens, nos quais evidenciam-se a impressão do colonizador, a descrição da fauna e da flora e a contribuição para identificação dos “nomes dos Brasil” ou “coisas do Brasil”.
- b) século XIX, em que se destacam a inserção de brasileirismos aos dicionários gerais e a produção de dicionários de brasileirismos, motivadas

por espírito nacionalista surgem produções lexicográficas sobre brasileirismos que tinham a pretensão de registrar e fixar a norma brasileira, obras que ora tinham como propósito descrever a norma nacional (brasileirismos) em oposição à europeia, ora buscavam registrar vocabulários regionais. (ISQUERDO, 2008, p. 122)

c) século XX, momento de

transição entre a produção lexicográfica centrada na questão dos brasileirismos, na língua brasileira, e uma lexicografia voltada para a língua portuguesa do Brasil, obras com ampla nomenclatura que buscam dar guarida o melhor possível do léxico do português brasileiro, com caráter mais autônomo, já que a produção do século anterior, em sua maioria, tinha como propósito complementar as obras produzidas além-mar. (ISQUERDO, 2008, p. 128)

d) século XXI, em que se apresentam tendências da Lexicografia contemporânea, como o lançamento de dicionários gerais da língua de grande porte e de versões eletrônicas e *online* de dicionários gerais da língua e a produção de dicionário com nomenclatura extraída de *corpora* textuais.

No que tange à Lexicografia, a questão é controversa. Em *Dicionários na teoria e na prática*, na seção intitulada “A Lexicografia deve ser vista como técnica ou como ciência?” há um debate entre Seabra e Welker.

Segundo Welker, as atividades da lexicografia resultam produtos diferentes e, portanto, os tipos de Lexicografia – a Lexicografia prática ou “arte de compor dicionários” e a Lexicografia teórica ou metalexigrafia – e o estatuto de técnica ou de ciência seriam também distintos, de modo que a primeira seria apenas uma técnica enquanto e a segunda – a metalexigrafia – poderia ser considerada, de fato, ciência.

Inversamente, Seabra (2011, p.30) admite a complementaridade de ambas as vertentes, de modo que a lexicografia pode ser entendida como “uma disciplina linguística de caráter científico que contempla os aspectos teóricos e práticos da elaboração de um dicionário” que se insere, portanto, “domínio da linguística aplicada”.

Há, porém, no âmbito de ambas as disciplinas – a Lexicologia e a Lexicografia –, um amplo espectro investigativo que pode ser caracterizado de acordo com a perspectiva – diacrônica ou sincrônica – e com o objeto – o léxico geral, o léxico variacionista e o léxico especializado.

No que concerne, especificamente, aos estudos lexicográficos, Miranda (2014, p. 215-231), após discutir acerca das inadequações de classificação das obras lexicográficas nas taxionomias existentes, estabelece uma proposta

taxonômica ancorada em cinco critérios linguísticos: i) o número de línguas consideradas; ii) a distinção entre “discurso livre” e “discurso repetido”; iii) a distinção entre ênfase informativa no plano do significante e ênfase informativa no plano do significado; iv) a distinção entre perspectiva semasiológica e perspectiva onomasiológica e; v) a distinção entre uma representação do léxico diassistemicamente inclusivo e uma representação do léxico diassistemicamente restritivo. O autor vale-se, ainda, da distinção entre obras lexicográficas para falantes nativos de uma língua e obras lexicográficas para falantes não nativos de uma língua, como um critério funcional.

Os critérios são explicitados em um esquema, a seguir reproduzido, em que se é possível localizar os produtos lexicográficos elaborados de acordo com a perspectiva adotada.

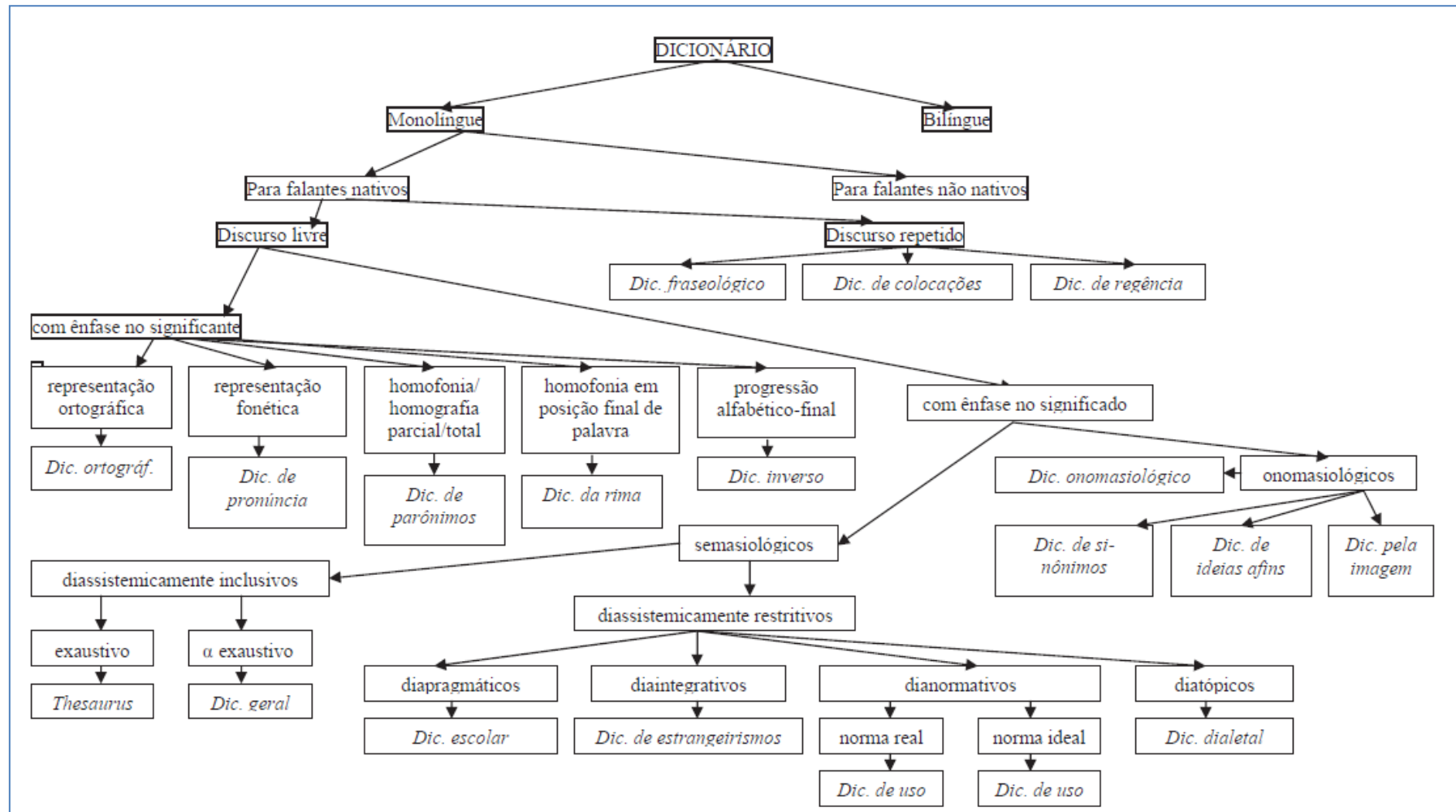


FIGURA 9 - Taxonomia de obras lexicográficas

Fonte: (MIRANDA, 2014, p. 228)

Com base na proposta taxonômica de Miranda (2014), é possível categorizar ramos da Lexicografia conforme o objeto e objetivo que adotam: i) a Lexicografia Pedagógica – cujos produtos são voltados ao ensino de língua materna ou estrangeira – ; ii) a Lexicografia Computacional – que visa à elaboração de dicionários eletrônicos – ; a Lexicografia Aplicada – cujo objeto de estudo são as crenças, dificuldades e estratégias sobre o uso do dicionário – ; a Lexicografia Regional – para o estudo dos regionalismos léxicos. O diálogo entre as vertentes é possível e relevante.

Com vistas ao cerne deste trabalho, os comentários, aqui, restringem-se a esta última categoria – Lexicografia regional – , a qual, segundo Lara (2007, p. 101), é subdividida em duas linhas de investigação: i) da presença dos regionalismos nos dicionários gerais e ii) da apropriação dos regionalismos como objeto exclusivo de estudo, *i.e.*, para a elaboração de “vocabulários dialetais ou dicionários de regionalismos”.

3.1.1.2.1 *Regionalismos em dicionários de língua ou dicionários gerais*

Sobre a discussão da configuração dos regionalismos, versa Isquierdo (2006, p. 14) que

no âmbito de uma língua implica considerar a noção de norma regional e popular, já que esses fatos linguísticos situam-se na esfera da variação lexical de natureza diatópica, ou seja, a variação que se processa no eixo horizontal ou espacial. (ISQUERDO, 2006, p. 14)

A inserção desses usos em dicionários gerais, no Brasil, porém, não é feita de maneira satisfatória.

Em um trabalho recente (no prelo), apresentado no IV Congresso de Dialectologia e Sociolinguística, a autora desta tese observou que das 134 lexias analisadas e computadas, 57 são dicionarizadas, 22 dicionarizadas com outro sentido e 55 não estão dicionarizadas. A análise foi feita em quatro dicionários aprovados pelo PNLD-2012: Dicionários– Bechara (2011), Borba (2011), Aulete (2011) e Houaiss (2011).

Observou-se, ainda, que enquanto Bechara (2011) e Borba (2011) registram menor número de variantes, apenas cinco e sete das 14 formas documentadas, o

Aulete (2008) e o Houaiss (2009) se caracterizam como mais completos, por registrarem 9 das 10 lexias, respectivamente. E vale ressaltar, ainda, que o Houaiss (2009) apresenta uma série de outras variantes.

Todos os quatro dicionários analisados alegam registrar regionalismos e estrangeirismos mas não explicam, nem especificam em seus verbetes. Aipim e macaxeira, por exemplo, são caracterizadas como regionalismos e caçula como brasileirismo no Houaiss e no Aulete. No entanto, não há especificação da região do Brasil. Talvez seja esse um dos problemas mais graves presente em todos os dicionários analisados.

Assim, embora não seja possível demarcar, com precisão, a dimensão da inserção da variação dialetal, com os dados analisados, notou-se certo descuido com a descrição diatópica.

Para além da questão da dimensão da inserção da variação linguística nos dicionários, foram observados alguns problemas estruturais que dizem respeito à remissão e ao registro etimológico.

3.1.1.2.2 *Dicionários (vocabulários e glossários) de regionalismos*

Na perspectiva da lexicografia regional²⁸ do século XX e XXI, *O dialeto caipira* (1920), de Amadeu Amaral, *O linguajar carioca* (1922), de Antenor Nascentes e *A língua do Nordeste* (1934), de Mário Marroquim, se destacam como obras das primeira e segunda fases da geolinguística no Brasil, haja vista marcarem

a produção de trabalhos de cunho monográfico voltada para a observação de uma área determinada, buscando descrever os fenômenos que a caracterizam não só do ponto de vista semântico-lexical, mas também do fonético-fonológico e morfossintático. Nota-se a existência da preocupação com uma metodologia de abordagem dos fenômenos orientada para o exame da realidade observada in loco e considerada nos seus diferentes aspectos. (CARDOSO, 2010, p. 134)

Com base na citação supracitada, listam-se os principais metaplasmos encontrados nas três obras ora destacadas.

²⁸ Um levantamento historiográfico de obras que possibilitem a variação lexical no Brasil, a partir do século XX até o século XXI, em especial vocabulários dialetais, tem sido feita por Ivan Pedro Nascimento, dentro do Grupo de Pesquisa Nêmesis.

Tabela 15 – Principais metaplasmos descritos em obras lexicográficas da segunda fase da geolinguística no Brasil: (AMARAL, 1920), (NASCENTES, 1922) e (MARROQUIM, 1934)

principais metaplasmos	Obras lexicográficas da segunda fase da geolinguística no Brasil		
	<i>O dialeto caipira</i> (1920) Amadeu Amaral	<i>O linguajar carioca</i> (1922) Antenor Nascentes	<i>A língua do Nordeste</i> (1934) Mário Marroquim
prótese	+	+	+
epêntese	+	+	-
paragoge	+	-	+
aférese	+	+	+
síncope	+	+	-
apócope	+	+	+
metátese	+	+	+
assimilação	+	+	+
dissimilação	+	+	+
nasalização	+	-	+
palatalização	+	-	+
despalatalização	+	-	-

Fonte: Elaborado pela autora desta tese (2017)

Em suas obras, os autores tecem considerações acerca do léxico documentado. Em *O Dialeto caipira*, Amaral (1920, p.55) pondera que o

vocabulário do dialeto é, naturalmente, bastante restrito, de acordo com a simplicidade de vida e de espírito, e, portanto com as exíguas necessidades de expressão dos que o falam. Esse vocabulário é formado, em parte: a) de elementos oriundos do português usado pelo primitivo colonizador, muitos dos quais se arcaizaram na língua culta; b) de termos provenientes das línguas indígenas; c) de vocábulos importados de outras línguas, por via indireta; d) de vocábulos formados no próprio seio do dialeto. (AMARAL, 1920, p. 55)

O autor ressalta, ainda, a grande tendência para a criação, por derivação e por composição, de muitos neologismos.

Dessa mesma ideia comunga Nascentes (1922), acerca do linguajar carioca, asseverando que o dialeto “tem dado provas de grande vitalidade quer por composição, quer por derivação”, além de formações teratológicas, caracterizadas pela etimologia popular e pela derivação regressiva.

Em outras considerações, Nascentes (1953, p. 181) enfatiza o prestígio do falar do Rio de Janeiro.

A principal característica do léxico carioca é, se assim nos podemos exprimir, o seu cosmopolitismo. Com efeito, capital e mais importante cidade do Brasil, o Rio de Janeiro exerce sôbre o resto do país uma força centrípeta que acarreta para o vocabulário carioca termos oriundos de todos os Estados. Ao lado desta força existe a contrária, que espalha pelo país inteiro os neologismos cariocas, como se deu com o verbo *avacalhar* e com a palavra *paredro*, por exemplo. (NASCENTES, 1953, p. 181, grifos do autor)

Ao longo dos tempos, publicaram-se diversas outras obras, de mesma categoria.²⁹ Opta-se, entretanto, por restringir os comentários a quatro obras: *O léxico rural: glossário, comentários* (CARDOSO, FERREIRA, 2000); *Dicionário das Expressões Idiomáticas Mais Usadas no Brasil: organização onomasiológica* (RIVA, 2013); *Dicionário do Nordeste* (NAVARRO, 2013); *Dicionário do dialeto rural no Vale do Jequitinhonha Minas Gerais* (ANTUNES, 2013).

O léxico rural: glossário, comentários (CARDOSO, FERREIRA, 2000)

A obra reúne todas as unidades lexicais cartografadas no *Atlas Prévio dos Falares Baianos-APFB* (1963) e no *Atlas Linguístico de Sergipe-ALS* (1987) e registra, lexicograficamente, ampla complexidade lexical e fonética revelada. Apresenta abonações. Inserem-se, ao final da obra, comentários que se referem a fatos fonéticos e a questões léxico-semânticas.

Consta de 880 verbetes.

Cada verbete tem como entrada a forma ocorrente na região, independentemente de encontra-se ela registrada ou não nos dicionários gerais da língua, seguida da indicação BA, Bahia, e/ou SE, Sergipe, ou BA,SE (cartas conjuntas) com a informação, entre parênteses, da carta ou cartas em que ocorre. ((CARDOSO, FERREIRA, 2000, APRESENTAÇÃO)

²⁹ *Glossário de vocábulos brasileiros, tanto dos derivados como daqueles a origem é ignorada* (1883-1884) de visconde de Beaurepaire-Rohan; *Apostilas ao dicionário de vocábulos brasileiros* (1912), de P. Carlos Teschauer; *Dicionário de brasileirismos* (1912), de Rodolfo Garcia; *Glossário paraense* (1905), de Vicente Chermont de Miranda; *O idioma do hodierno Portugal comparado com o do Brasil* (1879), de Paranhos da Silva; *Vocabulário Sul-Rio-Grandense* (1935) de Luiz Carlos Moraes; *Vocabulário Amazônico* (1942) de Amando Mendes; *Novo Dicionário da Gíria Brasileira* (1956) de Manuel Viotti; *Dicionário de Termos Populares (Registrados no Ceará)* (1959) de Florival Serraine; *Vocabulário de Têrmos Populares e Gíria da Paraíba* (1959) de Leon Clerot; *Dinâmica de uma linguagem: o falar de Alagoas* (1976) de Paulino Santiago, Potiguar; *gíria rio-grandense* (1980) de Raimundo Nonato; *Dicionário de língua popular da Amazônia* (1985) de Paulo Jacob.

Dicionário das Expressões Idiomáticas Mais Usadas no Brasil: organização onomasiológica (RIVA, 2013)

Em perspectiva onomasiológica, a obra registra expressões idiomáticas caracterizadas como mais usadas no Brasil com base em 400 conceitos que permitem a visualização de analogias existentes entre cada expressão idiomática, que sejam quando agrupadas dentro de um mesmo conceito quer sejam no cotejo entre tais. Apresenta recortes de contextos nos quais as expressões idiomáticas estão em usos reais.

Consta de cerca de 1500 expressões idiomáticas (ou idiomatismos) com alta frequência de uso no Brasil.

Dicionário do Nordeste (NAVARRO, 2013)

A obra é fruto de vinte e um anos de minuciosa pesquisa. Foi elaborado com vistas a atender a necessidade de “traduzir” para os colegas – jornalistas do estado de São Paulo– certos termos normalmente empregados pelo autor no cotidiano das redações paulistanas.

Consta de mais de dez mil verbetes e expressões usadas em todos os estados da região.

As três obras foram importantes a esta pesquisa pelas características que apresentam e pela contribuição lexical e lexicográfica que apresentam. É relevante ressaltar, entretanto que, dessas, apenas Cardoso e Ferreira (2000) documentam registros de fala com os quais os Atlas – APFB e ALS – foram elaborados. Navarro (2013) vale-se de obras literárias para a construção do Dicionário. Já Riva (2013) registra, sistematicamente, em todos verbetes, os contextos das expressões idiomáticas, em usos reais da escrita jornalística, mais especificamente, de textos veiculados *on line*.

A estratégia de organização onomasiológica utilizada por Riva (2013) se revela bastante interessante, haja vista ser este trabalho realizado com base em áreas temáticas.

3.1.1.3 Lexicultura

A Lexicultura surgiu na França dos finais do século XX, momento em que o tema da identidade motivou um intenso debate na sociedade francesa, marcada por dois conflitos devastadores: i) a queda de um extenso império colonial, momentos de forte instabilidade social e política, e ii) a perda do estatuto de grande potência, pelo menos nos termos em que o era em 1900.

Mas, apesar de múltiplas tribulações, a França tornou-se portadora de um colossal legado histórico e cultural. Nas mais diversas áreas, o país foi palco de movimentos importantes, tendo gerado figuras de relevo que ficarão para sempre associadas à história.

Diante esse cenário, em 1987, Robert Galisson – autor do “Dicionário das marcas do cotidiano”, que relaciona as principais marcas francesas que influenciam a língua e cultura e o primeiro estudioso a falar e criar uma metodologia para o ensino da segunda língua –, cunhou o neologismo terminológico, o conceito de Lexicultura. Para esse autor, o conceito de estudo lexical, de semântica e da abordagem morfológica não é suficiente para a aprendizagem de vocabulário, sobretudo por não dar conta do entendimento de algumas construções e de sentidos que possam ter em determinados contextos.

A Lexicultura, mais amplamente divulgada por Galisson (1991, 1999, 1999b), admite que algumas lexias são mais mobilizadas pela sua ‘carga cultural compartilhada’ (CCC) – valor acrescentado ao sentido referencial da lexia, que é conhecido e compartilhado entre os membros pertencentes a uma cultura e constitui fator de aproximação e de reconhecimento mútuo – do que pelo seu significado.

Mollard-Desfour (1998, 2000, 2001, 2004) admite que as lexias com carga cultural compartilhada (CCC) são as “palavras que transmitem a cultura particular de sociedade, seus costumes, crenças e que são portadores de códigos culturais lexicalizados compartilhados por uma grande parte ou a totalidade de uma comunidade no mesmo período e no mesmo local geográfico.”

Os fenômenos lexiculturais, os que designam os elementos de cultura presentes no léxico, quer no subsistema da língua corrente, quer nos subsistemas das línguas de especialidade, estão presentes nas macro e micro-estruturas do dicionário de especialidade, instrumento indispensável aos especialistas das

diferentes áreas do conhecimento, aos tradutores e ao ensino-aprendizagem da língua materna e das línguas estrangeiras para fins específicos.

Essa área tem, hoje, maior visibilidade no âmbito dos estudos de língua estrangeira, na confecção de dicionários bilingues, ensino de língua para estrangeiros.

Essa área tem, hoje, maior visibilidade no âmbito dos estudos de língua estrangeira, na confecção de dicionários bilingues, ensino de língua para estrangeiros.

A 'didactologia' é uma disciplina em que se encaixa a ação de Galisson desde os anos 80, com vistas a aproximar expectativas e necessidades dos assuntos específicos vigentes relacionados à sociedade em que se inserem alunos e professores. Essa disciplina tem como objeto de estudo o acesso aberto a línguas e culturas em um lugar de emancipação da nativa, e conhecimento / reconhecimento no exterior.

Em seu texto *Didactologie: de l'éducation aux langues-cultures à l'éducation par les langues-cultures* (Galisson 2002, p. 497-510) assevera:

Rien ne justifie en effet cette ségrégation. Pas plus au plan épistémologique qu'au plan éthique. Au plan épistémologique, savoir et savoir-faire s'inscrivent naturellement dans le même continuum et quand l'un fait défaut, il appelle l'autre, il ne le repousse pas. Au plan éthique, le concept théorie/pratique hiérarchise injustement les acteurs de ce double procès supposé : les acteurs de terrain, dits praticiens, frustrés de responsabilités et cependant jugés sur leurs actes, sont ancillarisés, colonisés par les acteurs de cabinet, dits théoriciens, eux-mêmes à l'abri de toute évaluation sociale.

De fato, nada justifica a segregação. Mas, não é isso que se vê por cá, do outro lado do Atlântico quando se trata de preconceito linguístico. Como esquecer da polêmica do livro didático, entre maio e junho de 2011, em que se previa a variação linguística no âmbito do nível morfossintático?

TSITSA (2004), no artigo intitulado *Le français de l'art culinaire un projet dictionnaire* observa o "Francês para Fins Específicos" ou seja, os empregos franceses em contextos "textual ou social, científica, profissional e técnica, visando expor o projeto de um dicionário bilíngüe, híbrido e "intersticial" da lexiculture, quase herética por design e porque não diz respeito a uma especialidade classificada "nobre". A autora acredita que tendo o dicionário como decodificação ou solução de problemas, permitiria o acesso e aprendizagem mais funcional, inclusive, da língua.

Mas, que tipo de dicionário conseguiria tal feito, se o que se tem visto é uma política normatizadora impositora e cerceadora de usos linguísticos?

No âmbito da Lexicografia, quiçá, a sensatez tão almejada esteja ganhando adeptos, os quais já há muito se preocupavam em garantir a legitimação dos usos linguísticos, acreditando que isso

é o que um grande educador faz. Mas um tolo faz exatamente o contrário. Se vir que a vida lhe deu um limão, renuncia e diz: "Estou derrotado. É o destino. Não tive sorte". Depois, põe-se a queixar-se amargamente do mundo e entrega-se a uma orgia de autopiedade. O homem sensato, porém, ao receber um limão, diz: "Que lição posso tirar deste infortúnio? Como poderei melhorar minha situação? Como poderei transformar este limão numa limonada?" (CARNEGIE, 2012, p. 201)

Considerando a proposta de TSITSA (2004), é possível admitir que a perspectiva lexicultural se aplique não somente ao âmbito dos estudos em língua estrangeira mas também ao estudo de língua materna, sobretudo se se considerar a diversidade sociocultural da realidade linguística brasileira e o caráter pragmático no uso das lexias.

3.1.2 Novos rumos para os estudos dos léxicos: interfaces

“Mais importante do que escolher novos caminho é descobrir novos horizontes.”

A iniciativa audaciosa e pensada na necessidade de um maior conhecimento da realidade linguística do Brasil, desenvolve-se no Projeto Atlas Linguístico do Brasil (Projeto ALiB), de cunho nacional que visa ao mapeamento do território brasileiro no que concerne aos fenômenos linguísticos e, associado a ele, o Projeto Dicionário Dialetal Brasileiro (Projeto DDB) – “obra de verve coletiva e interinstitucional que envolverá diversos especialistas, quer na área da dialectologia, quer nas áreas da lexicografia e das ciências da informação, do Brasil e da França” (MACHADO FILHO, 2010, p. 67), com vistas ao desenvolvimento de um dicionário que possa assegurar a difusão mais ampla do conhecimento da realidade linguística brasileira, no âmbito dos estudos lexicográficos.

Analogamente à elaboração de atlas regionais para a concretização do Atlas Linguístico do Brasil se justificou pela amplitude geográfica do país, considerando os fatos atinentes ao Dicionário Dialetal Brasileiro e visando a contribuir para sua completude, tem-se elaborado glossários temáticos e vocabulários regionais, inclusive esta tese, *Vocabulário Dialetal Baiano*.

Admite-se a pertinência das ideias lexicoculturais para estudo da variação no português brasileiro, inclusive na perspectiva lexicográfica como tem sido feito na lexicografia variacional, inclusive, para o desenvolvimento desta pesquisa e construção do Vocabulário Dialetal Baiano, que contribuirá para o DDB. Talvez, assim, com a divulgação de certas lexias, particularmente as mais representativas lexicoculturalmente, seria possível ter acesso a uma cultura que nos é pouco ou não conhecida.

Encontram-se, em andamento, também, trabalhos concernentes a outras regiões do País: i) a tese de doutorado intitulada *Vocabulário do Centro-Oeste brasileiro*, em desenvolvimento, por Daniela Silva Costa, da Universidade Federal do Mato Grosso do Sul, sob a orientação da professora Aparecida Negri Isquierdo; ii) a dissertação de mestrado *Vocabulário dialetal da região Norte*, por Cemary Correia de Sousa, da Universidade Federal da Bahia, sob orientação do professor doutor Américo Venâncio Lopes Machado Filho.

No que concerne ao Nordeste brasileiro, para além da tese *Vocabulário Dialetal Baiano*, já citada, foram desenvolvidos glossários temáticos, desenvolvidos no âmbito da Iniciação Científica, dentre os quais se destacam: i) *Ciclos da vida em Sergipe: glossário temático com base em dados do Projeto ALiB e A variação lexical em capitais do Nordeste brasileiro: fascículo sobre Convívio e Comportamento Social*, ambos os trabalhos, por Ivan Pedro Santos Nascimento; ii) *Glossário temático do convívio e comportamento social no Maranhão: dados do ALiB*, por Jéssica Farias dos Santos; iii) *O estado do Ceará numa perspectiva sociocomportamental: glossário temático com base em dados do Projeto ALiB*, por Cemary Correia de Sousa.

Por fim, considerando as reflexões acerca dos estudos léxicos, no esteio de Verdelho (2003), de que os

dicionários acompanham o percurso diacrónico da língua, dão conta da evolução e cristalização das formas das palavras, e constituem, por isso, fontes privilegiadas para o estudo da variação e fixação fonética, e podem

eventualmente fundamentar observações de âmbito sociolinguístico. Os dicionários portugueses documentam, por exemplo, a progressiva alternância e concorrência entre formas como “cousa / coisa”, “dous / dois”, “louro / loiro” (VERDELHO, 2003, p. 418),

pretendeu-se, nesta tese, ressaltar a importância do registro e descrição sociodialetal, em perspectiva lexicográfica, sobretudo, no que tange às alterações de caráter fônico no português brasileiro contemporâneo, identificando-as, sistematicamente, e relacionando-as aos processos históricos de mudança ocorridos na língua portuguesa no seu processo formativo.

PARTE II

**POR UMA LEXICOGRAFIA VARIACIONAL:
“REVENDO RAZÕES” E CONCEITOS**

Capítulo 4

POR UMA LEXICOGRAFIA HISTÓRICO-VARIACIONAL: “REVENDO RAZÕES” E CONCEITOS

Mergulhar em uma pesquisa requer recursos e vontade. Assim como a jovem mariposa do conto¹ atribuído a Rubem Alves não se contenta em voar em torno das lâmpadas, mas se apaixona por uma estrela, o pesquisador deve sentir o imprescindível desejo de voar mais alto, tendo o cuidado de não se tornar, porém, uma espécie de Ícaro, sobretudo no que concerne à utilização de material adequado ao alvo que se propõe atingir. (NEIVA, 2012, p. 70)

Faz-se necessário a um “mergulhador”, sobretudo ao que executa a arte do mergulho científico – conduzido como parte de uma pesquisa científica ou de uma atividade educacional – o estudo apurado do mar onde se lançará, a fim de estar ciente de que precisa descobrir se sob os corais há outras vidas a serem observadas ou se permanecerá a olhar exclusivamente para a beleza inerte cuja vida parece de pedra. De igual forma, no tocante aos estudos linguísticos, faz-se necessário que o pesquisador tenha consciência dos percalços que possam surgir e que esteja consciente de que os mergulhos em mares “nunca (e até mesmo os já que foram) antes navegados” requerem olhares e perspectivas diferentes.

A epígrafe supracitada faz referência à introdução do capítulo Metodologia da Dissertação de Mestrado da autora desta Tese e é aqui utilizada a fim de ratificar a ideia e o ideal da autora acerca do comportamento e postura imprescindível a um pesquisador: sair da superfície, quer seja com voos mais altos quer seja com mergulhos mais aprofundados, a fim de atingir a conscientização social da realidade plurilinguística brasileira.

Para tanto, assim como nas peripécias do pesquisador/mergulhador não se limitam ao mar ou à incessante busca por descobertas se incluem a delimitação do objeto de estudo, para a escolha do caminho e das ferramentas utilizadas, neste trabalho consideraram-se o interesse pela descrição da realidade linguística brasileira sob a ótica da diversidade –, as hipóteses e os objetivos específicos adotados no projeto delineado, *a priori*.

¹ Embora a autoria desse conto seja atribuída a Rubem Alves, ressalta-se que foi encontrado na obra virtual Coletâneas de Histórias, de Wagner Luis Marques.

A experiência de mergulho não é um requisito básico à prática, mas, certamente, possibilita ao pesquisador maior habilidade em suas descobertas e também em suas escolhas, sobretudo no momento de executar a tarefa mais árdua: decidir qual caminho trilhar e quais ferramentas utilizar, haja vista que nem sempre o caminho mais curto ou o já conhecido é o mais seguro.

No que tange ao registro da variação lexical relacionadas aos níveis da língua, – com os quais se podem perceber peculiaridades e similaridades nos hábitos, costumes e maneira de falar de dadas localidades, e, por conseguinte, determinar áreas dialetais – atualmente, têm sido feito em estudos sincrônicos, porém “cada um no seu quadrado” ou “cada macaco no seu galho”.

Paris, 2010. Eis que o professor doutor Machado Filho, inicialmente, cogita para seu pós-doutoramento, a elaboração de um dicionário especializado, tendo o léxico do futebol como tema. Ao longo do período, a ideia é substituída pela da construção de um dicionário dialetal brasileiro, tendo como fonte de dados, o *corpus* do Projeto Atlas Linguístico do Brasil. Surgia, assim, com vistas a uma interface entre a Lexicografia e a Dialectologia, mais uma modalidade de mergulho no Léxico: a Lexicografia Histórico-Variacional.

Desde o momento em que se adjetivou a área, tornou-se incongruente manter os conceitos, os quais já estavam muito bem já assentados por ambas as áreas. Para construir a “ponte” entre tais, foi preciso, pois, encontrar um *locus* e um *topus* e admitir que o objeto da Lexicografia Variacional distingue, obviamente, do objeto de ambas as áreas, haja vista o prisma observacional, sobre o qual se passa a tratar.

4.1 PALAVRA, PALAVRINHA, PALAVRÃO, PALAVRA-CHAVE, PALAVREADO (E OUTRAS PALAVRAS): SOBRE O OBJETO TEÓRICO ADOTADO

Da semia
Do sema, do semema, do semantosema
Do lexema
Do classema, do mema, do sentema,
Libera nos, Domine
(DRUMMOND DE ANDRADE, 1988, p. 862)

No capítulo 1, foram discutidos os caracteres nomeador, dinâmico, sociocultural e estrutural das palavras, re(i)novando a reflexão acerca da impropriedade da palavra ‘palavra’ nos estudos lexicais.

Até então, estava assentado para a Lexicografia contemporânea o conceito de *lexia*. No entanto, considerando os dados documentados no *corpus*, o termo *lexia* não pareceu também adequado à Lexicografia Histórico-Variacional, que se pretende fazer, sobretudo no que tange à relação do léxico com os demais níveis da língua.

Não adiantava apenas “exorcizar”, era preciso preencher o lugar que ficou. Não bastava apenas substituir o nome *lexia* por outro, fazia-se necessário que o conceito recobrisse a forma. Aliás, até mesmo o conceito de variante, também já assentado para os estudos variacionistas parecia não recobrir de maneira satisfatória.

Uma das razões porque o termo *lexia* não é satisfatório para designar o objeto da Lexicografia Variacional que se tem desenvolvido consiste no fato de que se pretendeu inserir, para além dos itens lexicais que fazem parte do léxico comum ou léxico geral e (ou) dos termos os quais pertencem ao léxico especializado, a especialidade das normas, de comunidades linguísticas, caracterizadas sociodialetoalmente.

Em *Do conceito de nomia para os estudos do léxico em perspectiva variacional e histórica*, Machado Filho (no prelo) afirma que, embora seja

clara a distinção entre esses dois objetos teóricos, o léxico que pode ser denominado de “sociodialetoal” ou “léxico das normas”, por sua vez, caracteriza-se pela relação que estabelece a partir do nível de abstração normal e sua efetiva materialização na fala. (MACHADO FILHO, no prelo)

Diante disso, considerando o fato de que a proposta do Dicionário Dialetoal Brasileiro e, conseqüentemente, deste *Vocabulário Dialetoal Baiano* é trabalhar com o léxico das normas, mais especificamente, o léxico sociodialetoal, optou-se por adotar o conceito de “nomia” cunhado por Machado Filho (2016)²,

(por apropriação e reversão de seu antônimo, anomia, registrado como padrão em Sociologia) para se referir às unidades do léxico fortemente caracterizadoras de normas sociodialetoais e objetos de estudo da lexicografia e da lexicologia variacional e histórica, com vistas a contribuir para os avanços teóricos das pesquisas na área dos estudos do léxico. (MACHADO FILHO, no prelo)

² Comunicação intitulada *Do conceito de Nomia para os estudos do Léxico em perspectiva variacional e histórica*, apresentada no IV Congresso Internacional de Dialetoal e Sociolinguística (IV CIDS), na Université Paris Sorbonne, em 2016.

Didaticamente, o autor apresenta um esquema diagramático bastante elucidativo, reproduzido, a seguir:

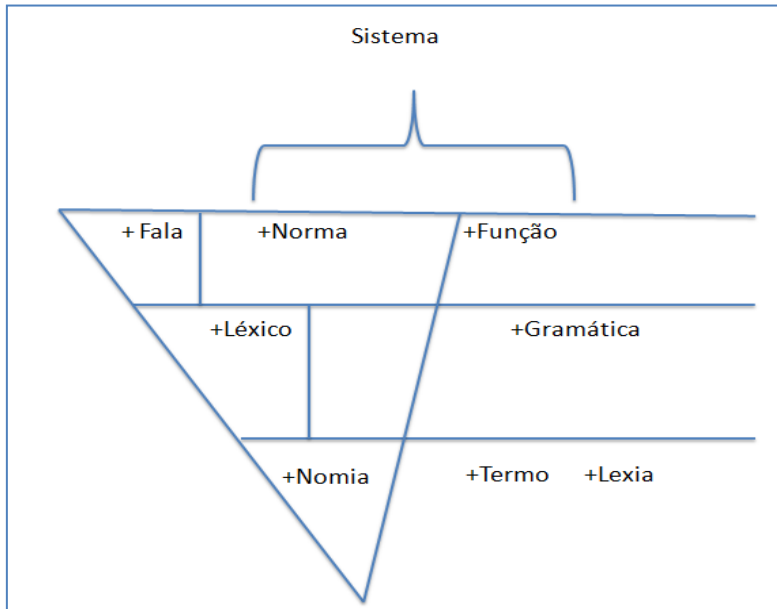


FIGURA 10 - Diagrama relacional dos objetos teóricos dos estudos do léxico, em função do sistema linguístico e do grau de gramaticalidade ou lexicalidade.

Fonte: (MACHADO FILHO, no prelo)

‘Nomia’, vale ressaltar, pode ser caracterizado como um desdobramento do conceito revisitado de variante lexical proposto, também, pelo autor (MACHADO FILHO, 2014) para classificar as unidades lexicais distintas

de se representar, em um mesmo contexto, um mesmo valor significativo ou funcional, independentemente de as alterações na forma terem origem fonética, fonológica, morfológica, sintática ou discursiva (MACHADO FILHO, 2014, p.273)

que sejam

fortemente caracterizadoras de normas sociodialetais e objetos de estudo da lexicografia e da lexicologia variacional e histórica, com vistas a contribuir para os avanços teóricos das pesquisas na área dos estudos do léxico.(MACHADO FILHO, no prelo)

Essa caracterização sociodialetoal de que fala o autor, pode ser entendida, também, como a Carga Cultural Compartilhada, com a qual a trabalha a Lexicultura.

Tomando como exemplo as oposições “chegou”, “chegaro” e “chegaram”, observa-se que do ponto de vista do padrão linguístico “chegou” e “chegaro” são marcadas negativamente e, portanto, se opõem a “chegaram”. Em contrapartida, se considerada a presença do morfema de marcação de plural, “chegaram” e “chegaro” são marcadas positivamente e se opõem a “chegou”, marcada negativamente.

Para melhor elucidação, observe-se o esquema, a seguir:

ponto de vista do padrão linguístico	(-)	(-)	(+)
	chegou	chegaro	chegaram
presença do morfema de marcação de plural	(-)	(+)	(+)

FIGURA – Exemplificação da delimitação de variáveis para um estudo variacionista, conforme parâmetros adotados.

Fonte: Elaborado pela autora desta tese (2017).

Como se percebe, a quantidade de variantes depende do parâmetro adotado, pois, é o pesquisador que constrói o objeto de análise bem como as variáveis.

Ou seja, que define o objeto teórico é o olhar.

Nesse sentido, sob a perspectiva da Lexicografia Variacional, “aimpim” ou “impim” se caracterizam como léxico de dada comunidade, haja vista se trabalhar com o valor semântico da unidade lexical. Ressalta-se, ainda, que aipim – forma canônica – é também considerada nomia já que também se constitui o léxico de uma comunidade. Ocorre que esta forma lexical, dicionarizada, é mais aceita socialmente do que as demais que são estigmatizadas.

As alterações ao longo dos tempos, na história do léxico português, vistas no capítulo 2, parecem justificar a utilização de registros que poderiam ser considerados, sob dada ótica, como registros fônicos de uma mesma unidade lexical.

Como se sabe, a mudança fônica não é abrupta, mas gradual, de modo que se propaga na estrutura da língua e na estrutura da sociedade, e que, assim sendo, “abroba” ou “abobra” podem ser consideradas léxico de dada comunidade, na sincronia, e pode se tornar léxico geral. No entanto,

o importante agora não é mais esquadrihar o passado para vermos como chegamos nele, mas nos determos na contemporaneidade para ver o quanto deste passado se prolongou no nosso presente e quanto ainda somos a partir dele. (PINHEIRO, 2000, p. 108)

E, assim, salvaguardar uma parte da história da língua.

Sob essa perspectiva,

o DDB objetiva conjugar de forma abrangente língua e cultura, no sentido de permitir o conhecimento mais abrangente possível da relação dialetal que se estabelece, através do léxico, nas comunidades que o utilizam. (MACHADO FILHO, 2010)

4.2 A FINALIDADE DE UM VOCABULÁRIO DIALETAL: RAZÕES

Em compensação outro velho amigo, desta vez o P. Gomes, me telefona para falar sobre palavras:

___ Existe a palavra **iniciante**?

___ Claro que existe.

___ Então olha se tem no dicionário.

Olho no “Aurélio”. Não tem.

___ Existe **telefonista**? ___ insiste ele.

___ Já vem você. Vai me dizer que não tem.

___ Então olha.

Não tem. Que diabo de brincadeira é essa? Coisas do P. Gomes, que só ele descobre. Com essa me deixou meio **desbundado**.

Olho se tem o verbo **desbundar**. Tem. (F, 15-05-85: 10, c. 3-6)

(SABINO, *apud* ALVES, 2002, p. 85)

O excerto, supracitado, da crônica de Fernando Sabino em sua coluna “Dito e Feito” do Jornal *O Globo*, citado por Alves (2002, p. 85) incita alguns questionamentos sobre o pensamento coletivo acerca da relação entre o léxico e a lexicografia brasileira: i) a (não) inserção de um item lexical em dicionário como sinônimo de sua (não) existência e; ii) o Aurélio como sinônimo de dicionário, fato que dá título a um artigo de Biderman (2000).

Sobre o primeiro aspecto, causa preocupação – e não raro indignação – perceber que, por vezes, veicula-se, midiaticamente, e é incutida na sociedade, a ideia de que a existência de uma “palavra” dependa de seu registro em um dicionário.

A crônica de Sabino remonta ao ano de 1985. Mais de trinta anos depois, “desbundadamente”, o vocábulo **iniciante** continua sem ser inserido ao *Dicionário Aurélio*. No entanto, muitos “iniciantes” surgiram, deixaram de ser, nesse intervalo de

tempo, e, ainda se pode notar tendência a atender a demanda do público em crescimento, sobretudo no que tange aos vários manuais para iniciantes, de áreas diversas.

Em 1990, nove anos antes do Projeto de Lei, já citado no capítulo 2, em combate aos estrangeirismos, também no âmbito da política nacional – que, aliás, não tem rendido apenas polêmicas linguísticas, até hoje – o ex-ministro do Trabalho e Previdência Social do Governo Collor, o sindicalista Antônio Rogério Magri, ao responder a um repórter que o questionara se o salário também seria reduzido, com a frase: “O salário do trabalhador é **imexível**”³, gerou grande polêmica sobre a “existência” do item lexical, que ainda é repercutida, inclusive.

Polêmicas à parte, fato é que os dicionários mais amplamente conhecidos de um público comum – como são os dicionários Aurélio e Houaiss – parecem ser “imexíveis” quanto à inserção de um verbete para **imexível**, que somente é dicionarizada por Aulete (2011).

Diferente de **iniciante**, **imexível** não parece ter sido alvo de aceitação social, haja vista não ser comum ouvir-se, exceto quando em referência ao fato supracitado, com certa carga de estigma.

Amorim (2012), inicia sua dissertação de mestrado, sobre as construções causais com *por causa que*, com a seguinte assertiva:

A revista *Veja*, em edição de 11 de agosto de 2010, publicou artigo criticando o desempenho linguístico dos candidatos à Presidência da República no primeiro debate exibido na TV, naquele ano. Sob um enfoque carregado de preconceito linguístico, tal artigo apresentou uma seção contendo os “10 erros de português que acabam com qualquer entrevista de emprego”. Um desses erros refere-se ao uso do conector causal *por causa que*: “‘por causa que’ é mais do que errado – **nem sequer existe**” (VEJA, 2010, p. 99) (AMORIM, 2012, p. 15)

Embora se encontre dicionarizada, o conector causal é considerado pela Revista como erro e como “inexistente” na língua, o que além de demonstrar preconceito remonta à falta de conhecimento linguístico.

Fatos como os citados, aqui, poderiam gerar questionamentos acerca do que determinaria a existência de uma “palavra”, se a sua dicionarização, se o seu prestígio ou mesmo se a colocação – recorrência em contextos linguísticos.

³ (grifo nosso)

Pretende-se, entretanto, para além das reflexões possíveis, evidenciar que na língua, sobretudo no léxico nada é “imexível”, de modo que as palavras existem “por causa de” seus os utentes, e (ou), como considera Barcelos (2000), que a

necessidade de comunicar-se e se fazer entender obriga os falantes a se lançarem ao uso de unidades lexicais para, com elas, enunciarem seus pensamentos por meio de entidades vocabulares que, nem sempre, estiveram ou estão disponíveis para o seu uso, impondo-se, então, a urgência de criá-las ou evocá-las no fragor do ato expressivo. (BARCELOS, 2000, p. 142)

Embora a dicionarização e o prestígio que possa adquirir conseqüentemente não estejam, diretamente, relacionados à existência de dado uso, ressalta-se a importância do registro lexicográfico de usos lexicais que hoje são consideradas “erro” ou até mesmo os que não possuem pouca expressividade quanto à veiculação. Acredita-se que dicionarizar um uso que é estigmatizado possa ser um grande avanço para a diminuição dos estigmas e preconceitos sociolinguísticos.

Outra reflexão oportuna se refere à ideia equivocada de que se “está no Aurélio, como é comumente denominado nosso mais usado dicionário, a unidade lexical já é considerada integrada à língua portuguesa falada no Brasil.” (ALVES, 1994, p. 43), acreditando que uma obra dicionarística possa representar a “verdade”, o “certo”, no que concerne ao léxico.

Biderman (2000), em *Aurélio: Sinônimo de dicionário?*, não obstante o destaque dentre os seus concorrentes que se revelaram inferiores no que tange à atualização e no que tange ao conteúdo, apresenta impropriedades quanto à nomenclatura, à microestrutura, à definição, a ordenação das acepções em palavras polissêmicas.

Talvez, tenham sido essas as mesmas razões que levaram o *PNLD Dicionários 2012*, a não aprovar o Aurélio, de tipo 4 – aqueles que, por seu porte, formato e objetivos, se aproximam dos que se dirigem ao público geral, embora se destinem, prioritariamente, a alunos do ensino médio.

Além das duas questões que deram início a esta reflexão em torno da perspectiva lexicográfica, tende-se a uma terceira que se refere ao atendimento à demanda dos consulentes e os efeitos na sociedade.

Embora em âmbitos diferentes quanto às modalidades da língua, a “famigerada” frase do personagem do conto *Famigerado*, de Guimarães Rosa (1988,

p. 13), “Vosmecê agora me faça a boa obra de querer me ensinar o que é mesmo que é: *fasmisgerado... faz-megerado... falmisgeraldo... familias-gerado...?*”, parece aludir aos consulentes de um dicionário de língua portuguesa, no Brasil, no que tange aos objetivos mais aparentes: saber “o que é” – a definição –, e qual o “uso correto”, normalmente, no que tange à norma-padrão “o uso correto” da palavra.

Em *Com direito à palavra: dicionários em sala de aula*, às páginas 16 e 17, são apresentados alguns dos “serviços” que podem ser prestados por um dicionário, dentre os quais são considerados “mais importantes”:

- tirar dúvidas sobre a escrita de uma palavra (ortografia);
- esclarecer os significados de termos desconhecidos (definições, acepções);
- precisar outros usos de uma palavra já conhecida (definições, acepções);
- desvendar relações de forma e de conteúdo entre palavras (sinonímia, antonímia, homonímia etc.);
- informar a respeito das coisas designadas pelas palavras registradas (informações sobre o inventor dos balões a gás e o contexto de época, num verbete como balão);
- indicar o domínio, ou seja, o campo do conhecimento ou a esfera de atividade a que a palavra está mais intimamente relacionada; essa informação é particularmente importante quando uma mesma palavra assume sentidos distintos (ou acepções) em diferentes domínios, como planta, em biologia e em arquitetura;
- dar informações sobre as funções gramaticais da palavra, como sua classificação e características morfossintáticas (descrição gramatical);
- indicar os contextos mais típicos de uso do vocábulo e, portanto, os valores sociais e/ou afetivos a ele associados (níveis de linguagem; estilo);
- assinalar, quando é o caso, o caráter regional de uma palavra (informação dialetológica);
- descrever a pronúncia culta de termos do português (ortoépia) e a pronúncia aproximada de empréstimos não aportuguesados;
- prestar informações sobre a história da palavra na língua (datação; indicação de arcaísmos e de expressões em desuso)
- revelar a origem de um vocábulo (etimologia).

Como se observa, as contribuições sugeridas não se restringem ao âmbito linguístico e podem atender a alguns dos interesses dos consulentes.

Muito embora se perceba certo aprimoramento na produção dos dicionários brasileiros no tocante ao método e aplicação de tecnologias, ao conteúdo com base em *corpora*, à proposta e ao viés lexicográfico, as obras ainda apresentam muitas deficiências estruturais, sobretudo no que tange à variação linguística, além de

serem redundantes, pouco adequados para a consulta e, por vezes, se mostrarem preconceituosos.

Como exemplo dos equívocos presentes em obras lexicográficas ainda veiculadas, no Brasil, observem-se, a seguir, acepções de dois verbetes arrolados no *Grande Dicionário Sacconi* (2009) e em Aulete (2011)

periguete (pe) *s.f.(a) Gíria* 1. Mulher que frequenta bares noturnos apenas para arrumar companhia; putinha da noite; piranha noturna; pistoleira: *toda periguete adora ter a marca da calcinha do biquíni aparecendo*. 2. Mulher fuxiqueira, bisbilhoteira: *cuidado com sua vizinha, que ela é a maior periguete do bairro!* 3. Mulher que se veste de forma pretensamente elegante e age de modo espalhafatoso; perua. (SACCONI, 2009)

menino *interj.* 4. Indica admiração e equivale a *Minha nossa!* ou a *Deus me acuda! Menino! Ele brinca de boneca!!!* (SACCONI, 2009)

curandeiro 1 Que trata pessoas doentes por meio de rezas e feitiçarias. 2 Que trata doentes sem ser formado em medicina, muitas vezes com métodos que incluem rezas, magias e beberagens **sm** 3 Aquele a quem se atribui a capacidade de curar por meio de rezas e feitiçarias. 4 Médico que se vale desses meio. 5 *Pej.* Médico de má qualidade e de formação duvidosa (AULETE, 2011)

Se os “dicionários são monumentos do mais valioso patrimônio de uma comunidade, são um testemunho privilegiado da memória” (VERDELHO, 2003, p. 413) sócio-histórica e linguístico-cultural, e refletem “o conjunto dos usos sociais da língua” torna-se inadmissível aceitar acepções, definições e verbetes como os citados, os quais além não serem coerentes à prática lexicográfica, ratificam preconceitos sociais, inclusive no que tange à variação linguística.

Entretanto, mesmo com os avanços nos estudos lexicais, grande parte dos dicionários, principalmente aqueles elaborados a partir da segunda metade do século XX, apresentam muitas deficiências estruturais, além de serem redundantes e pouco adequados para a consulta.

Tem-se observado ainda que, na percepção do usuário, a finalidade dos Dicionários para a maioria da sociedade é similar à busca pelas Gramáticas – livros que representam o conjunto de regras gramaticais, segundo uma norma de uso, a norma-padrão, em geral – e consiste em dirimir dúvidas relativas, em geral, à grafia correta e à busca por sinônimos, bem como ocorre no que tange à pronúncia, haja vista o fato de muitos dicionários apresentarem noções de ortoepia – indicação normativa de uma pronúncia. Vê-se, ainda que

em relação ao registro da variação nos dicionários contemporâneos, publicados no Brasil, se refere meramente a marcas de uso, que normalmente refletem uma certa carga de preconceito em face do padrão ortográfico que neutraliza quaisquer outras atualizações linguísticas que se possam insinuar na nomenclatura. (MACHADO FILHO, 2010, p. 52)

Tais fatos revelam grandes preocupações que motivam à consolidação da Lexicografia Variacional, com a elaboração de glossários, vocabulários e dicionários dialetais, atendendo a inquietações que se referem à legitimação dos usos linguísticos, e promovendo a interface entre a Dialetologia e a Lexicografia, sobretudo se se considerar a assertiva de Cardoso (2011, p. 329)

Os atlas linguísticos, na sua grande maioria, focalizam o léxico e a fonética/fonologia da língua considerada. Há um volume muito grande de lexias, arroladas em cada atlas. O aproveitamento desses dados lexicais pelos lexicógrafos, porém, não tem sido proporcional ao que os atlas oferecem, pensando-se, por exemplo, na variedade lexical do português. Onde está a dificuldade em assumir-se uma postura dessa natureza? (CARDOSO, 2011, p. 329)

Foi assumindo, pois, a postura sugerida por Cardoso (2011), com vistas a contribuir para o DDB, a fim de sanar “a ausência de uma metodologia de explicitação e divulgação de dados dos atlas linguísticos mais eficaz e mais transparente constitui-se no grande obstáculo à junção desses dois campos, desses dois interesses – o geolinguístico e o lexicográfico” (CARDOSO, 2011, p. 329) – um dos aspectos comentados pela autora como causa, razão sem motivo para a não utilização dos dados dialetais em obras lexicográficas – que esta pesquisa foi elaborada.

PARTE III

METODOLOGIA

Capítulo 5

PROCEDIMENTOS METODOLÓGICOS PARA A ELABORAÇÃO DO VOLUME I

5.1 DA CONSTITUIÇÃO DO *CORPUS* AO TRATAMENTO DOS DADOS

Obviamente, é possível estudar a língua sem o uso de *corpus* como objeto observacional mas, considerando os pressupostos teórico-metodológicos sob os quais se fundamenta esta pesquisa, torna-se relevante o trabalho com dados reais e concretos dos usos da língua, sobretudo por admitir a necessidade da não caricatura dos fatos linguísticos, como ocorre, inclusive, em manuais e livros didáticos.

Considerando os fatos, optou-se por trabalhar a variação no domínio da fala, visando a estabelecer, lexicograficamente, um panorama geodialetal e sociocultural, sobretudo por admitir a diversidade étnico-cultural brasileira e os processos sócio-históricos relevantes para a constituição do léxico do português brasileiro, especificamente no que tange ao estado da Bahia.

Sob esta perspectiva, para a elaboração do vocabulário proposto, o *corpus* utilizado compôs-se de itens lexicais, concernentes a 14 áreas temáticas, extraídos de 92 inquéritos realizados em 22 localidades do estado da Bahia pelo Projeto Atlas Linguístico do Brasil – o Projeto ALiB.

Os procedimentos concernentes ao tratamento e análises dos dados foram feitos conforme a metodologia adotada pelo referido projeto. Porém, ao longo do desenvolvimento do trabalho lexicográfico em perspectiva variacional, requereram-se outros critérios que somente puderam ser delimitados, a partir da análise prévia, embora cuidadosa, dos dados, alterando, pois, em alguns aspectos, o planejamento inicial.

É importante dizer que no momento em que se iniciou esta pesquisa, em 2012, os dados da Bahia já haviam sido coletados em pesquisa de campo¹ e os inquéritos já haviam sido previamente transcritos². Para esta pesquisa, pois, optou-

¹ Particpei de algumas coleta de dados como bolsista IC, como auxiliar em 3 localidades.

² O ALiB conta com três momentos: a transcrição prévia com base na audição dos inquéritos e duas revisões das transcrições com respectivas audições. De posse das transcrições, foi necessário fazer as revisões.

se por realizar uma revisão das transcrições ou uma nova transcrição com outros critérios dos quais se tratará, *a posteriori*.

Apresenta-se, a seguir, uma breve reflexão decorrente da descrição detalhada de alguns aspectos concernentes às localidades, aos informantes e ao método utilizado na constituição e tratamento do *corpus*.

5.1.1 As localidades: dos critérios de seleção aos aspectos socioculturais

Em cada estado brasileiro, o ALiB selecionou localidades conforme critérios demográficos, históricos e culturais, considerando a extensão do estado e da região e a natureza de seu povoamento.

No que tange à Bahia, foram selecionadas 22 localidades das quais seis coincidem com as localidades listadas por Nascentes (1958) como sugestão para investigação futura e elaboração de Atlas: Barra, Caetité, Carinhanha, Jacobina, Jeremoabo e Vitória da Conquista. Essas localidades e mais outras três (Itaberaba, Santa Cruz Cabrália e Santana) também se registram no APFB.

As 13 demais localidades – Juazeiro, Euclides da Cunha, Irecê, Barreiras, Alagoinhas, Seabra, Santo Amaro, Salvador, Valença, Jequié, Ilhéus, Itapetinga e Caravelas – pertencem apenas ao ALiB.

O quadro mostra a distribuição das localidades pontos do ALiB, no estado da Bahia, em relação à coincidência com os pontos sugeridos por Nascentes e as localidades documentadas no APFB.

LOCALIDADE	COINCIDÊNCIA		LOCALIDADE	COINCIDÊNCIA	
	com APFB	com Nascentes		com APFB	com Nascentes
Barra	+	+	Euclides da Cunha	–	–
Caetité	+	+	Irecê	–	–
Carinhanha	+	+	Barreiras	–	–
Jacobina	+	+	Alagoinhas	–	–
Jeremoabo	+	+	Seabra	–	–
Vitória da Conquista	+	+	Santo Amaro	–	–
Itaberaba	–	+	Valença	–	–
Santa Cruz Cabrália	–	+	Jequié	–	–
Santana	–	+	Ilhéus	–	–
Salvador	–	–	Itapetinga	–	–
Juazeiro	–	–	Caravelas	–	–

QUADRO 7 – Distribuição das localidades analisadas

Fonte: Elaborado pela autora (2017)

As sete mesorregiões baianas foram contempladas na seleção das localidades. Considerando esse fato, tornou-se necessário estudar aspectos das mesorregiões relacionados à presença de rios, às atividades agrícolas, dentre outros, sobretudo em função das áreas temáticas e das perguntas com as quais as lexias foram documentadas.

As informações foram confrontadas na análise dos dados, de modo que a ausência de respostas ou a diversidade nos usos pode ser justificada pela presença ou ausência dos aspectos estudados na localidade ou na região, como fatores condicionantes.

5.1.2 Informantes: do conceito de indivíduo ao conceito de comunidade de fala

Com vistas à observação das variáveis sociais relacionadas ao gênero/sexo e à faixa etária, em cada localidade, o ALiB registrou a fala de 4 indivíduos, equiparadamente, homens e mulheres com idades entre 18 a 30 e 50 a 65 anos. Nas capitais, o ALiB buscou, ainda, investigar o fator escolaridade, em dois níveis fundamental e universitário, realizando, pois, entrevistas com 8 indivíduos.

Desta forma, o ALiB contou com a colaboração de 96 indivíduos que nasceram e quem residem na localidade, cujos pais são também oriundos, e que possuem pouca mobilidade geográfica.

Embora se tenha controlado essas informações na constituição da base de dados, esta pesquisa não objetivou desenvolver análises das variáveis sociais. I

Em contrapartida, toda a pesquisa foi pautada no conceito de comunidade de fala adotado.

5.1.3 Métodos: da coleta de dados às estratégias de tratamento dos dados

A coleta de dados do ALiB foram realizadas através do método de gravação de áudio do tipo laboviano, ou seja, com base em um questionário específico, o Questionário ALiB 2001, em que são formuladas perguntas onomasiológicas e semasiológicas que contemplam: i) as relações pragmático-discursivas; ii) a interação mais espontânea, a partir de temas semidirigidos; iii) a concepção do falante a respeito da língua e; iv) os níveis da língua – o Questionário Fonético-

Fonológico (QFF), o Questionário Morfossintático (QMS) e o Questionário Semântico-Lexical (QSL).

Esta pesquisa se restringiu à observação das respostas dos informantes às perguntas constantes do QSL, que se constitui de 202 questões subdividas em 14 áreas temáticas: *Acidentes geográficos; Fenômenos atmosféricos; Astros e tempo; Atividades agropastoris; Fauna; Corpo humano; Ciclos da vida; Convívio e comportamento social; Religião e crenças; Jogos e diversões infantis; Habitação; Alimentação e cozinha; Vestuário e acessórios e Vida urbana.*

É válido destacar que os pesquisadores em Dialetoologia e Sociolinguística têm utilizado gravadores de voz como forma de arquivar a fala dos informantes para que, a partir da audição e transcrição dos enunciados *à posteriori*, possam extrair o máximo de informações linguísticas e socioculturais intrínsecas e (ou) correlacionadas às formas linguísticas empregadas. Para tanto, a elaboração e uso de questionário são importantes para o direcionamento da pesquisa e requer do pesquisador habilidade e cuidado no que tange à fidedignidade dos dados, sobretudo, no que se refere ao fato de não causar indução da resposta.

Assim sendo, importa ao pesquisador, sobretudo àquele que opte por realizar entrevistas em gravação por questionário, ser dinâmico e perspicaz em sua abordagem, buscando, portanto, extrair do informante não apenas o item lexical a que deseja pesquisar, mas também os contextos linguístico e social.

Caruso (1998) aponta algumas dificuldades encontradas no trabalho de campo, sobretudo no que tange à realização da entrevista que se vale do método de gravação da fala dos informantes, cujos critérios já foram mencionados. Admite a percepção de que durante a entrevista algumas questões apresentavam problemas, sobretudo as questões em que se previa a variação léxico-semântica.

Essa preocupação foi percebida em grande parte dos inquéritos analisados, os quais propiciaram a realização de um estudo adequado da variação léxica no âmbito da oralidade. Há que ressaltar, porém, casos que deixaram dúvidas quanto à validação da resposta, inviabilizando saber se trata de uma variante ou não. Para tanto, procedeu-se à elaboração de estratégias para o tratamento dos dados.

5.1.3.1 Estratégias para o tratamento dos dados

Dentre as estratégias adotadas nesta pesquisa para o tratamento dos dados, destacam-se: i) a lematização dos dados e delimitação do *corpus*; ii)

5.1.3.1.1 Lematização dos dados e delimitação do *corpus*

A fim de compor o vocabulário, procedeu-se à delimitação das variantes com base na análise do comportamento de todas as unidades lexicais e na lematização das unidades.

Com a utilização do *software WordSmith 4.0* – um fragmentador de texto utilizado pela Oxford University Press na composição dos resultados de suas pesquisas lexicográficas e adotado para a composição do DDB –, foi possível obter, *a priori*, os signos lemativos – todas as formas lexicais presentes em uma lista de palavras de um dado *corpus* (PNLD, 2012) –, por frequência e por ordem alfabética.

N	Word	Freq	%	Texts	%lemmas/Set
1	IN	3 551	6,77	12	100,00
2	E	2 126	4,06	12	100,00
3	INQ	2 006	3,83	12	100,00
4	#	1 547	3,71	12	100,00
5	QUE	1 396	2,66	12	100,00
6	DE	1 269	2,42	12	100,00
7	A	1 243	2,37	12	100,00
8	O	1 169	2,23	12	100,00
9	NÃO	980	1,87	12	100,00
10	EU	881	1,68	12	100,00
11	E	829	1,58	12	100,00
12	TEM	720	1,37	12	100,00
13	NÉ	663	1,26	12	100,00
14	CHAMA	544	1,04	12	100,00
15	UM	477	0,91	12	100,00
16	ASSIM	449	0,86	12	100,00
17	UMA	443	0,85	12	100,00
18	Á	413	0,79	12	100,00
19	NLIM	394	0,75	12	100,00
20	AQUI	362	0,69	12	100,00
21	OUTRO	346	0,66	12	100,00
22	SE	343	0,65	12	100,00
23	COMO	335	0,64	12	100,00
24	PRA	321	0,61	12	100,00
25	VOCÊ	305	0,58	10	83,33
26	JEITO	281	0,54	11	91,67
27	NO	276	0,53	12	100,00
28	MAIS	268	0,51	12	100,00

Word	Freq	%	Texts	%lemmas/Set
546	BLOCO	1	1	8,33
547	BLUSA	1	1	8,33
548	BO	2	2	16,67
549	BOA	13	0,02	5 41,67
550	BOAZINHA	1	1	8,33
551	BOCA	3	2	16,67
552	BOCADO	3	1	8,33
553	BODE	3	2	16,67
554	BODEGA	1	1	8,33
555	BOI	12	0,02	5 41,67
556	BÓI	1	1	8,33
557	BOIADA	1	1	8,33
558	BOLA	3	2	16,67
559	BOLADO	1	1	8,33
560	BOLDO	1	1	8,33
561	BOLHINHAS	1	1	8,33
562	BOLNHAS	2	2	16,67
563	BOLO	3	1	8,33
564	BOLSO	9	0,02	4 33,33
565	BOM	26	0,05	9 75,00
566	BOMBA	1	1	8,33
567	BOMBOM	7	0,01	2 16,67
568	BOMBOMZINHO	1	1	8,33
569	BONECA	3	1	8,33
570	BONITA	1	1	8,33
571	BONITAS	1	1	8,33
572	BONITO	7	0,01	7 58,33
573	BONZINHO	1	1	8,33

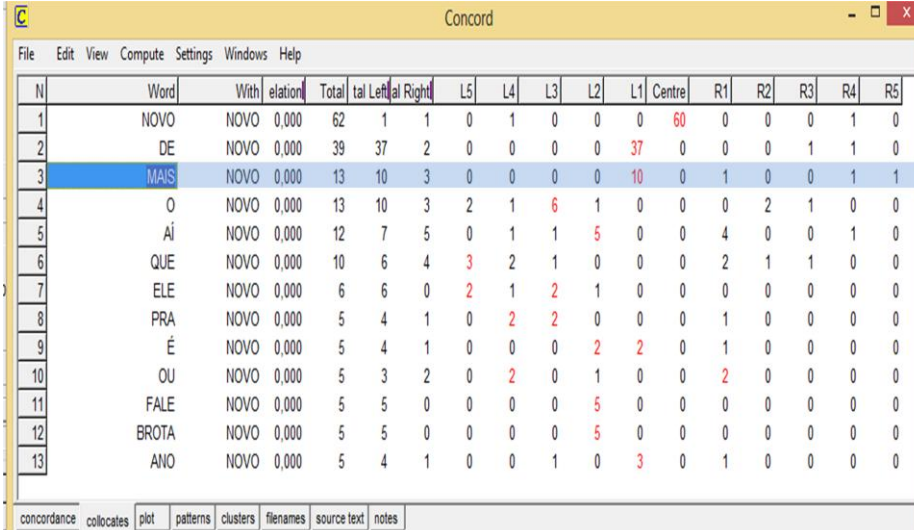
FIGURA 11: Excerto de *wordlist* do *corpus* por frequência e em ordem alfa

Para tanto, tornou-se necessário modificar a extensão do arquivo original – as transcrições do ALiB – em DOC para a versão TXT, tendo em vista que o processamento dos dados no programa de fragmentação de texto é melhor executado dessa maneira, sem, porém, abdicar da identificação sociodialetal.

Como processo de teste, para uma melhor visualização e observação futura, foram feitos três tipos de arquivamentos dos dados em versão TXT: i) inquérito completo de cada um dos informantes; ii) por informante + localidade + questionário; iii) por informante + localidade + área temática.

Considerando a relevância das possibilidades oferecidas, após a extração dos signos lexicais, por frequência e por ordem alfabética, utilizou-se o concordanciador – ferramenta que permite a observação do comportamento das unidades lexicais, sobretudo no que tange à composição de lexias complexas.

Como forma de exemplificação, observe-se, a seguir, uma amostra realizada para a análise-piloto, em que a colocação de “mais novo” permite a interpretação de ser esta uma variante de “caçula”.



The screenshot shows the Concord software window with a concordance table. The table has columns for N, Word, With, relation, Total, tal, Left, al, Right, L5, L4, L3, L2, L1, Centre, R1, R2, R3, R4, R5. The rows list various words and their co-occurrence with 'NOVO' across different positions in a sentence.

N	Word	With	relation	Total	tal	Left	al	Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	NOVO	NOVO	0,000	62	1	1	0	1	0	0	0	0	0	60	0	0	0	1	0
2	DE	NOVO	0,000	39	37	2	0	0	0	0	0	0	37	0	0	0	1	1	0
3	MAIS	NOVO	0,000	13	10	3	0	0	0	0	0	0	10	0	1	0	0	1	1
4	O	NOVO	0,000	13	10	3	2	1	6	1	0	0	0	0	2	1	0	0	0
5	ÁI	NOVO	0,000	12	7	5	0	1	1	5	0	0	4	0	0	0	1	0	0
6	QUE	NOVO	0,000	10	6	4	3	2	1	0	0	0	2	1	1	0	0	0	0
7	ELE	NOVO	0,000	6	6	0	2	1	2	1	0	0	0	0	0	0	0	0	0
8	PRA	NOVO	0,000	5	4	1	0	2	2	0	0	0	1	0	0	0	0	0	0
9	É	NOVO	0,000	5	4	1	0	0	2	2	0	2	2	0	1	0	0	0	0
10	OU	NOVO	0,000	5	3	2	0	2	0	1	0	0	2	0	0	0	0	0	0
11	FALE	NOVO	0,000	5	5	0	0	0	0	5	0	0	0	0	0	0	0	0	0
12	BROTA	NOVO	0,000	5	5	0	0	0	0	5	0	0	0	0	0	0	0	0	0
13	ANO	NOVO	0,000	5	4	1	0	0	1	0	3	0	1	0	0	0	0	0	0

FIGURA 12: Colocações de NOVO no *corpus* analisado.

A colocação – coocorrência de itens lexicais em contextos morfossintáticos específicos e semanticamente transparentes – pode ser percebida pela ferramenta utilizada e como se nota nas lexias documentadas “de” e “mais” ocorrem junto a NOVO com mais frequência. Junto a “novo”, “de”, obviamente, forma um lexia gramatical, as quais são mais frequentes na língua.

A segunda lexia mais frequente, “mais”, coocorre junto a “novo” e forma ora uma construção sintática ora uma lexia. Em uma análise prévia, a partir da concordância propiciada pelo *software*, puderam ser feitas algumas observações.

N	Concordance	Set	Tag	Word #	t. #	os. #	os. #	os. #	os. #	File	%	
1	E quem é que chama, o povo mais novo ou o povo mais velho? INF. – É o			1.240	171	7%	0	1%	0	1%	a 090 04_qls.txt	53%
2	o povo mais antigo ou o povo mais novo ? INF. – Esse pessoal qui não tem			2.087	412	0%	0	8%	0	8%	a 100 01_qls.txt	50%
3	– Tem outro jeito. INF. – Filho mais novo . [□□□□□□□□□□□□□□] (132)			2.939	667	0%	0	4%	0	4%	a 100 04_qls.txt	56%
4	dá esse nome? INF. – Inquanto ele tá novo , mais depois que já crece é o			1.653	301	0%	0	9%	0	9%	s 102 03_qls.txt	39%
5	INQ. – O senhor chama assim mais novo ? INF. – Era, até hoje ainda chamo			7.043	998	0%	0	7%	0	7%	a 083 03_qls.txt	76%
6	(121) INF. – Olha, quando eu era mais novo eu ouvia falar em boi □□□□□□.			5.412	773	7%	0	9%	0	9%	a 083 03_qls.txt	58%
7	nome, esse criado, apareceu mais novo , era dorido □□□□□□□□□□			2.470	336	1%	0	7%	0	7%	a 083 04_qls.txt	66%
8	é o mais velho e esse o...? INF. – Mais novo □□□□□□□□□□ (132) INQ. –			2.110	428	7%	0	1%	0	1%	a 084 02_qls.txt	61%
9	peessoa mais velha. Que os mais novo não. INQ. – Já chamam logo de			2.627	543	3%	0	0%	0	0%	e 096 04_qls.txt	79%
10	□□□□□□□□ hoje é mais novo . Agora naquele tempo era assim.			149	34	0%	0	4%	0	4%	e 095 03_qls.txt	4%
11	(131) INF. – Fala filho mais novo , caçulo □□□□□□□□□□ (132) INF.			1.840	259	0%	0	5%	0	5%	a 083 02_qls.txt	64%
12	(131) INF. – Filho mais novo □□□□□□□□□□□□□□ INQ.			1.580	260	1%	0	3%	0	3%	s 087 02_qls.txt	63%

FIGURA 13: Concordância de “mais novo” nos dados do ALiB, concernentes à Bahia

Nas linhas 1 e 6, tem-se uma construção sintática em que “mais” é intensificador do adjetivo “novo”, de modo que o sintagma se configura qualificador do nome. Em contrapartida, nas linhas 3, 11 e 12, embora “mais” continue exercendo sobre “novo” a intensificação, nota-se que se trata de uma lexia identificada pela colocação, a qual pode ser comutada com a lexia “caçulo”.

A colocação foi imprescindível para a avaliação do estatuto de lexias, ou seja, se se tratava uma construção sintática ou se se trata de uma lexia, quando um composto se cristaliza de modo que o significado dos componentes se perde em função de um novo significado que se aplica ao composto.

O uso das ferramentas do *software WordSmith 4.0* foi crucial para o esclarecimento quanto à lexicalização e à identificação dos diferentes graus de fixação de estruturas sintáticas por propiciar a realização dos testes sugeridos por Pottier (1972), dos quais se falou no Capítulo 1 – a comutabilidade e a inclusão de um elemento na unidade.

As estruturas que não se configuraram como lexicalizadas foram analisadas separadamente e não foram inclusas ao Vocabulário.

Em casos de dúvidas quanto à validação da resposta, no que tange ao valor de variante, tornou-se necessário avaliar a formulação da pergunta feita pelo inquiridor e, “persistindo os sintomas”, averiguar se o inquiridor insiste na pergunta a fim de que o informante confirme a resposta. O critério de frequência também foi utilizado para determinar a validação da resposta.

Em casos de ocorrência única, após a tentativa de avaliação da formulação da pergunta e da insistência do inquiridor na confirmação da resposta pelo

informante, não podendo, pois, precisar se seriam ou não variantes, as lexias foram desconsideradas da análise e não inclusas ao Vocabulário.

5.1.3.2 *Estratégias de documentação: a transcrição dos dados*

Segundo Paiva (2003, p. 135), transcrever os dados extraídos dos inquéritos linguísticos que simulam eventos de fala é “[...] transpor o discurso falado, de forma mais fiel possível, para registros gráficos mais permanentes, necessidade que decorre do fato de que não conseguimos estudar o oral através do próprio oral”. Admitindo a necessidade de que se busque captar não apenas a forma linguística, mas também a real intenção do falante, faz-se necessário que essa transcrição seja rigorosa e o mais inequívoca possível, respeitando assim, as hesitações e as marcas da oralidade.

No campo das transcrições de inquéritos linguísticos gravados e executados com base em questionários recomenda-se o estabelecimento do texto através da chamada transcrição grafemática. É o que faz o Projeto ALiB.

No entanto, é válido ressaltar que os critérios de transcrição se diferenciam a depender do interesse de cada do pesquisador. No caso desta pesquisa, muito embora se tenha conhecimento dos critérios de transcrição grafemática do ALiB, foi preciso estabelecer outros critérios, haja vista o objetivo lexicográfico.

Optou-se, pois, pela transcrição grafemática conforme a orientação do Projeto ALiB, apenas para fatos fônicos que, lexicograficamente, deveriam ser representados pela relevância cultural e sociodialetal que apresentam, dentre os quais se destacam:

- i) o apagamento de consoantes finais, como em lápi (lápiz);
- ii) o apagamento de sílabas, como em sabo (sábado);
- iii) acréscimo de sílabas, como em avoar (voar);
- iv) as metáteses, como em protuguês (português);
- v) a iotização ou despalatalização da lateral, como em mulé / muié (mulher) e fio (filho);
- vi) a semivocalização do /x/, como em arma / auma;
- vii) redução de ditongos, como em fêra (feira);
- viii) outros casos (não previstos), como coresma, redemunho, causo.

Não foram considerados, grafematicamente, os fatos fônicos que não representem caracterização sociodialetal, ou seja, fenômenos que são amplamente difundidos, já sendo considerados caracterizadores do português brasileiro geral. Dentre tais, se destacam:

- i) o alteamento das vogais átonas finais (*menino, rude* etc...);
- ii) a ditongação em contextos específicos (*mas > mais; arroz > arrois*);
- iii) a iotização do nh (*amanhã > amãã*)

Para esses casos, optou-se por inserir a transcrição fonética e observações, as quais são válidas também para marcar formas que não puderam ser representadas graficamente em função da inexistência de um correspondente gráfico do fonema em questão.

A exemplo disso, citam-se dois fenômenos bastante significativos que caracterizam áreas dialetais no Nordeste e, inclusive, dentro da Bahia:

- i) a oscilação do / t / e / d /
- ii) as vogais médias pretônicas

5.1.3.2.1 Estratégias de arquivamento dos dados: base de dados

A priori, para o armazenamento dos dados, utilizou-se o programa *Microsoft Office Excel*. Com o desenvolvimento da pesquisa, optou-se por utilizar o programa *Microsoft Office Access*, por considerá-lo mais funcional para a análise dos dados. As variantes foram arquivadas, por área temática, com as informações sociodialetais e o contexto em que se registraram na sentença.

5.1.4 Análises dos dados

Tendo, pois, constituído o *corpus* e delimitado as variantes, foram executados outros procedimentos metodológicos, visando à maior profundidade das análises. As estratégias de arquivamento dos dados contribuíram para a análise quantitativa dos dados com o auxílio dos gráficos gerados pelo próprio programa *Microsoft Office Access*.

É válido ressaltar que pesquisas variacionistas possuem outros métodos e instrumentos de análises quantitativas, a exemplo do *Goldvarb 2001* etc. Embora se

tenha conhecimento da existência e instrumentalização dos programas pelos trabalhos anteriores, julgou-se não ser relevante para a pesquisa, considerando não serem esses o foco e o objeto.

A análise quantitativa dos dados, do modo como foi realizada, foi satisfatória pois permitiu a seleção dos lemas principais, haja vista ser a frequência, um dos critérios adotados.

No primeiro momento de observação dos dados, buscou-se centrar não apenas no caráter quantitativo dos usos, mas, e sobretudo, no comportamento sociocultural das lexias documentadas, a partir dos contextos em que os atos de fala dos informantes se manifestaram. Para tanto, a avaliação da formulação da pergunta foi fundamental, sobretudo para refletir acerca da definição, posteriormente, elaborada e empregada no verbete do vocabulário.

Tendo por objetivo contribuir para a descrição dos fenômenos linguísticos a fim que se possa conhecer a realidade brasileira e, tendo como produto final a construção de um vocabulário dialetal, esta pesquisa restringiu-se à observação da diatopia. A análise quantitativa dos dados foi realizada, pois, com base nas mesorregiões da Bahia.

Foram investigadas possíveis condicionamentos extralinguísticos, admitindo que os usos linguísticos podem ser condicionados por fatores internos e externos, sendo o indivíduo social também responsável pelo processo e mecanismo de mutabilidade linguística, de modo a ser igualmente responsável pelas alterações.

Capítulo 6

PROCEDIMENTOS METODOLÓGICOS PARA A ELABORAÇÃO DOS VOLUMES II E III (VOCABULÁRIO DIALETAL BAIANO, GERAL E POR FASCÍCULOS)

6.1 MACROESTRUTURA

[...] o **DDB** deve privilegiar um sistema de remissão bastante eficiente e complexo, capaz de dar conta de todas as co-referências possíveis, incluindo-se aí o sistema de identificação isoglósica da unidade lexical sobre o que adiante se referirá na composição microestrutural a ser sugerida. [...] Note-se que possíveis realizações variacionais do tipo <abroba> ou <abobra>, relacionadas ao padrão *abóbora*, independentemente de não serem ortograficamente reconhecidas, devem ocupar sua posição alfabética nessa lista de palavras. (MACHADO FILHO, 2010, p. 62)

O *Vocabulário Dialectal Baiano* obedece às bases metodológicas do DDB, cuja macroestrutura preliminar, elaborada a partir da interpretação-piloto do léxico presente em alguns inquéritos realizados no estado da Bahia, foi apresentada no WorkALiB 2013, e já se encontra publicada (Machado Filho; Neiva, 2013)

É válido ressaltar que as primeiras ideias de aproveitamento, em perspectiva lexicográfica, dos dados da variação lexical documentadas pelo Projeto ALiB para a construção do DDB foram, antes, publicadas por Machado Filho (2010) no número 41, da Revista Estudos Linguísticos e Literários, periódico científico dos programas de Pós-Graduação em Língua e Cultura, e Literatura e Cultura, da Universidade Federal da Bahia.

Como se pode depreender da citação ora epigrafada, a macroestrutura do DDB objetiva permitir que o consulente, especializado ou leigo em linguística, reconheça, imediatamente, as variáveis sociodialetais registradas pelo ALiB, para cada uma das lexias em variação.

A nomenclatura – conjunto de verbetes em uma obra dicionarística –, compõe-se de todas as variantes lexicais registradas com entrada em ordem estritamente alfabética e se estrutura da seguinte maneira: i) página em orientação da vertical, em duas colunas, tamanho A4, margem conforme a ABNT; ii) entradas deslocadas à esquerda e texto com recuo de 10; justificado, espaçamento simples; iii) fonte do verbete em fonte *Arial Narrow*, tamanho 9, com negritos, itálicos e

sublinhados, quando preciso, obedecendo aos critérios adotados para a microestrutura.

O Vocabulário conta, ainda, com dois tipos de verbetes: i) os verbetes plenos, em que o lema principal serve de base para a remissão para outras entradas, e; ii) os verbetes remissivos, cujas entradas se referem aos lemas secundário ou múltiplo e às variantes vocabulares dos lema principal.

No tocante à microestrutura, Welker (2004, p. 107), citando Rey-Debove (1971, p. 21) ressalta a padronização necessária em todos os verbetes mas também admite a não existência dos mesmos tipos de informação para todos os lemas. Nesse sentido, é possível atestar que os verbetes assumam uma microestrutura própria dentro de uma padronização preestabelecida.

Com vistas a propiciar um acesso rápido e capaz de condensar informações de cunho linguístico, lexicultural e sociodialetal das lexias, permitindo, inclusive, a identificação de possíveis relações lexicais e semânticas entre as lexias registradas e os registros fônicos a serem exibidos grafematicamente, estabeleceu-se, *a priori*, a seguinte microestrutura, expressa **no quadro**, a seguir:

ITENS	INDICADORES	
	TIPOGRÁFICOS	NÃO TIPOGRÁFICOS
lema principal	redondo, letra minúscula, negrito	seguido de traço
[transcrição fonética]	em <i>SildoulosIPA</i> ,	entre colchetes, sílaba tônica marcada com sinal ' , seguida de traço
classificação gramatical	redondo, minúsculo, conforme lista de abreviaturas,	seguida de ponto
(étimo, origem ou processo de formação)	étimo ou origem e (ou) processos formativos em <i>itálico</i> ,	entre parênteses, encerrado(a) por ponto, precedido(a) do sinal <, sigla convencional para língua de origem (para o étimo), processos formativos marcados com sinal +
<i>Áreas temáticas do QSL:</i>	redondo, minúsculo, <i>itálico</i>	encerrado por ponto
'Definição ou acepções'	em minúsculas, em paráfrase lexicográfica,	entre aspas simples, encerrada(s) por ponto
Descrição sociodialeto do lema principal	redondo, iniciais das localidades em maiúsculo,	seguinte à cabeça do verbete, entre chaves, com nome da localidade e a caracterização sociolinguística do informante representada por numeral, separados por barra (/) e encerrada por ponto
lemas secundários	em letra minúscula e negrito, em redondo quando a forma for dicionarizada, em <i>itálico</i> quando for variante não dicionarizada,	precedido(s) de til (~), sendo a última variante encerrada por ponto.
Descrição sociodialeto do lema secundário	redondo, iniciais das localidades em maiúsculo,	entre chaves, encerradas por ponto, como o nome da localidade e a caracterização sociolinguística do informante representada por numeral.
Remissões lexicais	em minúscula, em redondo quando a forma for dicionarizada, em <i>itálico</i> quando for variante não dicionarizada	precedidas do sinal → e encerradas por ponto.
Legenda geolinguística	---	<p>Diagrama de uma legenda geolinguística com uma estrutura hierárquica de caixas contendo as letras E, O, V, S, F, C, S, R, M, N, E.</p>

QUADRO 9 – Microestrutura preliminar do DDB, para verbetes plenos, em 2010.

Fonte: Elaborado pela autora desta tese (2017).

Com o desenvolvimento da pesquisa, a microestrutura foi alterada em alguns aspectos:

- i) A 'definição' sucedia a 'área temática' e passou a precedê-la.
- ii) Cogitava-se, *a priori*, a possibilidade de inserir acepções, quando preciso. Optou-se, na versão atual, por considerar homonímia, identificado-a pela marca de sobrescrito, no verbete pleno e no verbete remissivo com o emprego de numerais arábicos.
- iii) Considerando admitir o tipo de definição lexicográfica, acrescentou-se o item 'formulação da pergunta', que é apresentado após a 'área

temática', em *itálico*, com o número da questão do QSL com a qual a lexia foi documentada, seguida da formulação da pergunta como consta no Questionário ALiB.

- iv) Escolha de indicadores tipográficos e não-tipográficos para caracterizar a 'abonação' e 'marcas de uso', itens previstos na microestrutura preliminar de 2010.
- v) Exclusão do item 'transcrição fonética'. Embora esse item, inversamente à ortoépia, não possua caráter normativo, mas, estritamente, descritivo, optou-se por excluir, a fim de evitar interpretações equivocadas.
- vi) Substituição da descrição sociodialetal – nome da localidade com a indicação do número do informante – e da descrição diatópica – número / ponto do ALiB, seguido do nome da localidade – de todas as entradas, em todos os verbetes, por um 'marcador para consulta', apensado a este trabalho.

A Lexicografia contemporânea tem recorrido a recursos e estratégias proativas para resolver problemas de indicadores, os quais nem sempre, são suficientes dentro de um verbete. Dentre as estratégias, destaca-se a inserção de textos exógenos para o auxílio da decodificação da informação.

Ciente da existência e êxito de decisões metodológicas dessa natureza, para a decodificação do item 'legenda geolinguística', optou-se pela elaboração de um 'marcador de consulta', em que se apresenta a legenda geolinguística explicada: i) número e nome de cada localidade; ii) identificação sociodialetal, no que tange ao gênero e à faixa etária – 1 = homem, jovem; 2 = mulher, jovem; 3 = homem, idoso; 4 = mulher, idoso.

Constam, ainda, no 'marcador', a 'chave de consulta' exemplificada com um verbete pleno do *corpus* e 'alguns indicadores não-tipográficos'.

Com vistas a melhor esclarecimento e visualização da proposta, apresentam-se, a seguir, a microestrutura atual, expressa no quadro, por ordem de apresentação de cada item no verbete e a chave de consulta.

ITENS		DESCRIÇÃO E (OU) EXEMPLIFICAÇÃO		INDICADORES	
				TIPOGRÁFICOS	NÃO TIPOGRÁFICOS
1. entrada unidade lexical que é objeto das informações seguintes	padrão / normal	Ex.: aipim		redondo, letra minúscula, negrito,	---
	com indicação de homônimo	quando a lexia possui homônimo e ambos são documentados no <i>corpus</i> Ex.: mandioca¹ mandioca²		redondo, letra minúscula, negrito, com indicação do homônimo por algarismos arábicos sobrescritos	---
	com indicação de lema múltiplo	quando houver variação de cunho mórfico ou morfosintático Ex.: corcunda(o)		redondo, letra minúscula, negrito, com morfema que caracteriza a alteração, em negrito e redondo,	entre parênteses
	entrada falsa	quando a forma padrão não for documentada no <i>corpus</i> . Ex.: [calcanhar]; te[r] neném		redondo, letra minúscula, negrito,	morfema ou lexia, entre colchetes
2. Definição lexicográfica'		definição em que apenas o hiperônimo características específicas compõem a codificação da informação da lexia Ex.: 'Instrumento para projetar, feito de forquilha e material elástico, utilizado, também, em algumas comunidades como brinquedo'		em minúsculas,	entre aspas simples
3. Formulação da pergunta do QSL ALiB		Ex.: <i>Como se chama o brinquedo feito de uma forquilha e duas tiras de borracha, que os meninos usam para matar passarinho.</i>		em <i>itálico</i> , com o número da questão do QSL com a qual a lexia foi documentada,	seguido de ponto.
7 variantes	7.1 lemas secundários	toda variante lexical que se caracterize como registro gráfico-fônico do lema principal. Ex.: ~ corcunda ~ cocunda ~ cacunda ~ caucunda		em letra minúscula e negrito, em redondo	precedido(s) de ~ (til)
	7.2 lemas múltiplos	toda variante lexical que se caracterize por alteração de cunho mórfico / morfosintático		com morfema que caracteriza a alteração, em negrito e redondo,	entre parênteses
		indicação da existência do lema / lexia após a entrada registro da lexia Ex.: ≈ corcund(o)		em minúsculo, negrito e sublinhado	precedido por ≈ (duplo til)
	7.3 variantes vocabulares	lexias formadas por radicais distintos ao lema principal ou cuja base lexical difere do lema principal Ex.: ≅ gimbrundo		em minúsculo, negrito,	precedido pelo símbolo ≅

ITENS	DESCRIÇÃO E (OU) EXEMPLIFICAÇÃO	INDICADORES	
		TIPOGRÁFICOS	NÃO TIPOGRÁFICOS
8	quando a lexia possuir antônimos ou outro tipo de relação semântica Ex.: → gimbra	em redondo, negrito, com indicação do homônimo por algarismos arábicos sobrescritos	precedidas do sinal →
	quando a lexia possuir um homônimo também documentado no <i>corpus</i> Ex.: → mandioca²	em redondo, negrito, com indicação do homônimo por algarismos arábicos sobrescritos	
9	Abonação exemplo extraído do <i>corpus</i> que serve para ilustrar o emprego do item lexical. Ex.: <i>É diferente o cheiro, mais ativo. Laranja-cravo é mais ativo o cheiro do que a pocã</i>	em negrito, itálico, em cor cinza	---
10	Legenda geolinguística representação lexicográfica das localidades estudadas divididas nas cinco mesorregiões da Bahia	---	
11	Achega enciclopédicas item facultativo em que se apresentam observações pertinentes acerca da lexia de entrada. Dentre tais informações, a referência ao registro da lexia no APFB.	tipo de marca, em redondo e negrito,	antecedida por ♦

Quadro 10 – Microestrutura atual do verbete pleno constante do *Vocabulário Dialetal Baiano*

Fonte: Elaborado pela autora desta tese (2017).

A microestrutura adotada para verbetes plenos é constituída, pois, de 8 itens obrigatórios – entrada do lema principal; classificação gramatical; registro etimológico ou de processo formativo; definição lexicográfica; área temática do QSL; as variantes que podem se constituir como lemas secundários e (ou) múltiplos e variantes vocabulares; a abonação e; legenda geolinguística – e três itens opcionais: as remissões semânticas, a marca de uso e as achagas de verbete. Cada item é apresentado por indicadores tipográficos e não-tipográficos específicos.

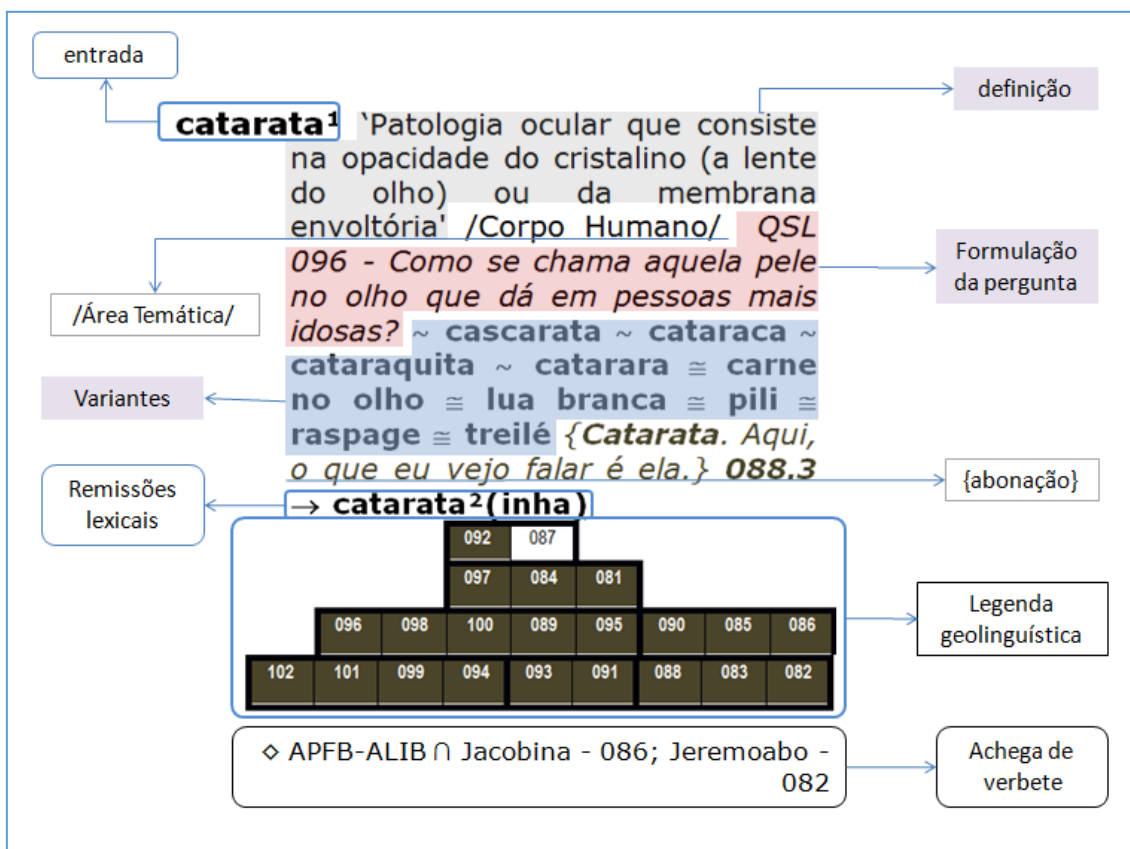


FIGURA 14: Chave de consulta atual do *Vocabulário Dialectal Baiano*
Fonte: Elaborada pela autora desta tese (2017)

Para constituir essa microestrutura e, por conseguinte, a nomenclatura – apresentada em colunas –, ao longo do desenvolvimento do trabalho lexicográfico em perspectiva variacional, requereram-se critérios que somente puderam ser delimitados, a partir da análise prévia, embora cuidadosa, dos dados, alterando, pois, em alguns aspectos, o projeto inicial.

6.1.1 Os processos de constituição do verbete

Considerando o projeto inaugural de registro de todas as variantes lexicais, de posse dos dados, em uma análise prévia, percebeu-se que algumas lexias permitiriam a discussão sobre os critérios para a apresentação de cada um dos itens, no vocabulário proposto, obedecendo aos processos de constituição do verbete:

- i) critérios para elaboração das entradas próprias;
- ii) sistema remissivo, em que se discutem a apresentação do verbete pleno quanto à possível inserção de lemas secundários e múltiplos, outras variantes e remissões semânticas e a apresentação de cada um desses itens quando em entrada própria;
- iii) aspectos linguístico-estruturais, em que se discute a apresentação da classificação gramatical e dos registros etimológicos da lexia;
- iv) aspectos lexicográfico-culturais, em que se problematizam a definição, a inserção da marca de uso, da abonação e da formulação da pergunta e da legenda geolinguística;

Os processos passam a ser discutidos, pormenorizadamente, na sequência.

6.1.1.1 Critérios para elaboração das entradas próprias

A delimitação do *corpus* tornou-se necessária em virtude da não validação de algumas respostas, por não representarem, lexicalmente, o conceito proposto. Desta forma, as estruturas sintáticas obtidas como respostas às questões foram analisadas em um capítulo da tese, mas não foram inseridas no Vocabulário. Para tanto, o conceito de léxico adotado foi fundamental.

Para a representação gráfica das lexias que se apresentam como entradas dos verbetes no Vocabulário foram estabelecidos dois principais critérios: i) o critério gráfico-linguístico – se, e somente se, houver alteração da representação gráfica dos fonemas – e ii) o critério sociodialetal – se, e somente se, houver caracterização sociodialetal do fenômeno, ou seja, quando o fenômeno não for amplamente difundido, ou seja, não seja um fenômeno caracterizador do português brasileiro geral.

Observem-se, a seguir, como os fatos fônicos de uma lexia foram considerados e apresentados no verbete.

FENOMENOS	EXEMPLOS	Representação gráfica	Forma de inclusão no verbete
Aférese do es no verbo estar	(ta, tão, tou, tava, tiver, tavam)	representação gráfica	grafema
Apócope do l e do r final	(sal > sá; senhor > senhô)	representação gráfica	entrada falsa
vocalização da lateral	calcanhar > caucanhar	representação gráfica	entrada falsa
Africada t e d	(tia e dia)	transcrição fonética	achega enciclopédica
Alteamento das vogais átonas finais	e > i; o > u	representação gráfica	grafema
Vogais médias pretônicas		transcrição fonética	achega enciclopédica
Ditongação da adversativa	mas; (mais)	transcrição fonética	achega enciclopédica
lotização do nh	(amanhã > amãã)	transcrição fonética	achega enciclopédica
Aspiração da fricativa s	(mesmo > mermo)	representação gráfica	grafema
Aspiração da lateral l	(volta > vorta)	representação gráfica	grafema
Contração das preposições de e em com os indefinidos um e uma	(dum, num, duma, numa)	representação gráfica	grafema
Contração da preposição para, seguida ou não de artigo e dos advérbios aqui e acolá	(pra, pro, praqui, pracolá)	representação gráfica	grafema
Contração do se e do me com a palavra embora	(simbora, m'imbora)	representação gráfica	grafema
lotização do lh	(velho > veio)	representação gráfica	grafema
Palatalização ou despalatalização do lh	Mulher > mulé	representação gráfica	grafema
Nasalização do r final na forma infinitiva do verbo vir	(vim)	representação gráfica	grafema
Perda da nasalização final	(comeram, comero)	representação gráfica	grafema
Síncope do r de grupos consonantais	(alpendre > alpende)	representação gráfica	grafema
Desnasalização das vogais postônicas		representação gráfica	grafema
Contração das proparoxítonas em paroxítonas		representação gráfica	grafema
Contração da preposição para com os pronomes pessoais eu, ele, ela:	(p'eu, p'ele, p'ela, pra eu, pra ele, pra ela)	representação gráfica	grafema
Síncope do d no grupo nd:	(passando > passano)	representação gráfica	grafema

QUADRO 11 – Metaplasmos

Fonte: Elaborado pela autora desta tese (2017)

É válido ressaltar que embora todos os fenômenos fônicos e os metaplásticos listados façam parte da história da língua portuguesa e continuem em variação no português brasileiro, apenas foram representados graficamente na entrada de verbete – quando possível – aqueles que se caracterizam sociodialetalmente como “coisa da Bahia” ou mesmo do Nordeste.

No entanto, era pertinente incluí-los, de alguma maneira, no verbete. Para tanto, os fenômenos que não puderam ser representados graficamente, em função da inexistência de um correspondente gráfico do fonema em questão, foram marcados em achegas enciclopédica – um item facultativo que compõe o verbete – em que se apresentam a transcrição fonética e observações acerca do fenômeno em questão, do qual se tratará, *a posteriori*.

Em casos em que a representação gráfica da lexia se distancia do padrão ortográfico, tornou-se necessário a inserção do padrão, mesmo que este não tenha ocorrido no *corpus*. Para tanto, utilizou-se a estratégia cunhada por Machado Filho como “falsa entrada”, as quais

indicam que, embora a lexia pesquisada não esteja atestada na forma gráfica que se encontra patente entre indicadores estruturais, especificamente pelos colchetes, a sua correspondente histórica estaria devidamente lematizada no dicionário, conquanto em forma morfológica de plural, não-canônica, portanto, em função dos dados, que exemplarmente aqui só teria ocorrido com essa configuração linguística no *corpus*.

(MACHADO FILHO, 2010)

Ora um leigo da existência de “çapatos” na história da língua portuguesa, seria ele capaz de supor encontrar em dicionário a lexia grafada com ç na inicial? É provável que jamais.

Desta forma, admitindo o fato de que um dicionário deva atender a um público não restrito – embora se compreenda os objetivos de cada consulente e de cada dicionário – tendo sido documentada a lexia “çapatos”, em corpora do português arcaico, a inserção do verbete constituído pela entrada [sapatos] e indicação remissiva a “çapatos” parece ser mais que pertinente para o reconhecimento da lexia pelo consulente, caso contrário, não seria possível encontrar “sapatos”. O consulente seria, pois, “descalçado” de conhecimento.

Machado Filho (ano) propõe que a nomenclatura de um dicionário histórico

deveria idealmente comportar não apenas toda a variação detectada nos *corpora*, mas, também, fomentar uma estratégia de "falsas entradas" em português moderno – somente quando estritamente necessárias – devidamente sinalizadas, contudo, com indicadores estruturais, tipográficos e não-tipográficos, como elementos facilitadores de consulta, isto é, nos casos especiais em que a alfabetação pudesse ser comprometida. (MACHADO FILHO, ano, página)

Sob mesma perspectiva, a estratégia de “entrada falsa”, usada pelo autor em obra lexicográfica de cunho histórico-variacional, pode ser também aplicada à *Lexicografia Variacional*, em perspectiva sincrônica, sobretudo nos casos em que haja comprometimento do entendimento da lexia e (ou) do respeito à ortografia padrão atual.

É válido ressaltar, pois, que a proposta do Vocabulário é de inclusão e não de exclusão. Ou seja, a dicionarização de todas as lexias presentes no *corpus* ocorreu de modo a registrar os usos desconhecidos ou estigmatizados em concomitância ao culto, ao padrão, não assumindo, pois, uma postura partidária, panfletária e excludente.

6.1.1.2 Sistema remissivo

Em *Lexicografia Histórica e questões de método*, no tocante à confecção de um dicionário histórico do português, Machado Filho (2012) considera que ser preciso

[...] privilegiar, para além desse procedimento antes sugerido de lematização, um sistema de remissão, de alguma forma perdulário, que possa arcar com grande parte da exuberância gráfica existente, evitando com isso que não se deixe de permitir ao público-alvo uma consulta rápida e eficaz às unidades léxicas de seu interesse. [...]

(MACHADO FILHO, 2012, p. 382)

Assim, com ênfase na remissão exaustiva dos dados, a lematização das lexias documentadas foi realizada visando a permitir, quando possível, a inserção, no verbete pleno – constituído pelo lema principal –, de lemas secundários, lemas múltiplos, variantes vocabulares e remissões semânticas, os quais possuem entrada própria e remissiva ao lema principal e se caracterizam como variantes lexicais.

Como critério para a seleção do *lema principal*, com o qual se constitui a entrada do verbete pleno, admitiu-se, *a priori*, a dicionarização da lexia, ou seja, sua

forma padrão e, considerando a possibilidade de não identificação do dado, no *corpus*, optou-se pelo critério da frequência.

Nos verbetes são apresentados, também, lemas secundários, lemas múltiplos, as variantes vocabulares e uma remissão semântica.

Por *lema secundário* entende-se, aqui, toda variante lexical que se caracterize como registro gráfico-fônico do lema principal. No verbete pleno, os possíveis *lemas secundários* se posicionam na sequência da questão do QSL, precedidos de til (~), em letra minúscula e negrito, em redondo quando a forma for dicionarizada, em itálico quando for variante não dicionarizada, sendo a última variante encerrada por ponto.

No que tange aos *lemas múltiplos* – “particularidades de flexão ou derivação que porventura provoquem alomorfia lexemática” (PNLD Dicionários, 2012) ou particularidades que denunciam alguma modificação de natureza mórfica –, optou-se por i) indicar a existência, a partir da inserção do morfema que caracteriza a alteração, entre parênteses, após a entrada do lema principal; ii) situá-lo, na sequência da descrição diatópica dos possíveis lemas secundários, em minúsculo, negrito e sublinhado, precedido pelo sinal (≈).

Designam-se *variantes vocabulares* as lexias formadas por radicais distintos ao lema principal ou cuja base lexical difere do lema principal. No verbete pleno, as lexias são precedidas do sinal de til (~), em minúsculo, negrito e redondo, quando a forma for dicionarizada e em itálico quando for variante não dicionarizada, encerradas por ponto.

Os *lemas secundários e múltiplos*, para além de serem apresentados como variantes no verbete principal, possuem entrada própria e são apresentados em verbete-remissivo – microestruturalmente esvaziado –, em que se registram para além da entrada, a remissão ao lema principal e a legenda geolinguística correspondente a seu emprego espacial e social.

Considerando a relevância da informação de cunho etimológico às *variantes vocabulares*, optou-se por inserir o item nos verbetes cujas entradas se constituam de lexias desse grupo de variantes.

Para visualização e melhor compreensão do que é dito, observem-se, a seguir, modelos de verbetes remissivos constituídos de lemas secundários, múltiplos e variantes vocabulares.

O sistema remissivo opera também na perspectiva de indicar possíveis relações entre a entrada e outras lexias que possuem relações semânticas existentes no *corpus*. No verbete pleno, são os últimos itens apresentados, antes da legenda geolinguística, e aparecem precedidos do sinal → , em minúscula, em redondo quando a forma for dicionarizada e em itálico quando for variante não dicionarizada, encerradas por ponto.

Ainda em relação às entradas, mais especificamente no que tange à relação entrada-acepção – aqui considerados como homonímia –, foram feitos outros testes com base no *corpus*. Em uma amostra constituída por dados obtidos como respostas à questão 39 do QSL, “Como se chamam as frutas menores que a laranja, que se descascam com a mão, e normalmente, deixam cheiro na mão?” (COMITÉ..., 2001), constantes da área temática *Atividades agropastoris*, obteve-se, dentre outras, a realização de *pocã*, como mostram os excertos, a seguir:

Excerto 1:

Pocã. [...] tangerina [...] A diferença é só o nome. Mas a fruta é a merma.

Inq. 096/4 (Caetité – BA) / Mulher, faixa 2, nível fundamental)

Excerto 2:

É diferente o cheiro, mais ativo. Laranja-cravo é mais ativo o cheiro do que a pocã

Inq. 095/3 – (Jequié – BA) / Homem, faixa 2, nível fundamental)

Como se pode observar, no Excerto 2, o informante caracteriza *pocã* com traços que distinguem da *laranja-cravo* que é uma variante de *tangerina*. Esse fato promoveu a consideração, em perspectiva lexicográfica, dessa lexia, não como uma variante, mas como elemento homônimo, merecendo, portanto, tratamento diferenciado, ou seja, indicações sobrescritas – *pocã*¹ : *pocã*² –, no caso de elaboração de vocabulários, glossários ou dicionários, como se prevê o DDB e por conseguinte, o vocabulário proposto.

Tendo sido eleita a variante *tangerina* como lema principal pelos critérios já comentados, na macroestrutura do Vocabulário, o item **pocã**¹ é remissivo a **tangerina** configurando-se variante – possuem os mesmos traços – e outro – **pocã**² – é lema principal, cujos traços, para dado informante, distinguem essa fruta da tangerina – a casca mais grossa e o tamanho.

Para esses casos, optou-se pela inserção na nomenclatura de maneira que o lema principal se apresenta em redondo, letra minúscula, negrito, com algarismos

arábicos sobrescritos. Em casos como o de **pocã**², a abonação, acompanhada da identificação sociodialetoal – registro da fala do informante – se mostra imprescindível para a construção do vocabulário proposto, sobretudo por se considerar o fator sociocultural.

6.1.1.3 Aspectos lexicográfico-culturais

Discutem-se, aqui, aspectos lexicográfico-culturais presentes no verbete do Vocabulário: i) a relação entre a definição a formulação da pergunta; ii) a inserção de marcas de uso; iii) a relevância da abonação; iv) a elaboração de uma legenda geolinguística e v) achegas enciclopédicas.

6.1.1.3.1 Definição e formulação da pergunta do QSL

Para construir um dicionário é preciso técnica e nomenclatura.
(MACHADO FILHO)

[...] arte suprema, em lexicografia, é a da *definição*.
(IMBIS *apud* WELKER, 2004, p. 117)

Machado Filho, em disciplinas e orientações acerca dos estudos do léxico e do trabalho lexicográfico, tem alertado para a importância da definição em quaisquer sejam as áreas e tipos de trabalhos científicos, haja vista ser esta a codificação da informação semântica disponível.

Nos trabalhos sobre definição, costumam-se distinguir as definições quanto ao tipo: i) sinonímica, quando elaborada por alegados equivalentes sêmicos; ii) lógica, quando consiste em um equação sêmica incontestada do *definiendum* ou do que deve ser definido pelo *definiens* e; iii) lexicográfica, quando se constitui de uma equação sêmica “de compromisso”, em que opõem o *genus proximum* e a *diferentia specifica*.

Para este trabalho, adotou-se o tipo de definição lexicográfica e para tanto, foram testadas, *a priori*, como forma de treinamento do raciocínio lógico e propício para a construção, algumas estratégias como a de campos lexicais e a de semagrama.

A estratégia de campos lexicais, utilizada para auxiliar na composição do item, é fundamentada no modelo estruturalista, com base na ideia coseriana e na Lexemática. A partir da estratégia de campos lexicais, pode-se obter uma definição lexicográfica, partindo, prioritariamente, do *genus proximum*.

As informações semânticas identificadas são agrupadas em paráfrases definitórias, que comumente obedecem ao padrão dos dicionários franceses e se caracterizam pela categoria gramatical: i) para substantivos, as paráfrases são construídas por substantivos com algum atributo (adjetivo) ou por um pronome com oração relativa; ii) para adjetivos, por adjetivos unidos por conjunção, orações relativas ou locuções preposicionais e adjetivas; iii) para advérbios, por advérbios ou locuções adverbiais e; iv) verbos, por verbos unidos por conjunção ou sintagmas verbais.

A estratégia de semagrama é executada a partir de *brainstorming* – mais que uma técnica de dinâmica de grupo, é uma atividade desenvolvida para explorar a potencialidade criativa de um indivíduo ou de um grupo –, com a qual se pode obter diversas lexias e dentre as quais se elegerá, *a posteriori*, a que melhor se adequa na composição da definição. Essa estratégia tem por base cinco elementos: categoria, função, pessoal, localização e tempo.

Como se percebe, as três estratégias podem contribuir de forma significativa para a elaboração de uma definição do tipo lexicográfica, ou seja, aquela em que apenas o *genus proximum* e a *differentia specifica* compõem a codificação da informação de um dado item lexical, prevista para o vocabulário, a fim de evitar a perpetuação da prática da definição sinonímica, por vezes muito utilizada pela lexicografia contemporânea, sobretudo por se tratar de um vocabulário dialetal, ou seja, em que se contemplam, prioritariamente as variantes, os sinônimos em lexicografia.

Além da técnica utilizada como meio de padronização, tornou-se necessário, para a clareza e acuidade definitória, ressaltar o fator sociocultural.

Observe-se, a seguir, os procedimentos usados para a definição da lexia *badogue*, constante da área Jogos e diversões infantis, obtida como resposta à questão 157 do QSL: “Como se chama o brinquedo feito de uma forquilha e duas tiras de borracha, que os meninos usam para brincar?”. No tocante ao item ‘Definição’ da nomenclatura de um verbete.

Elegeu-se, para a definição da maioria dos dados, a estratégia de semagrama, tomando por base os cinco elementos antes citados. Em seguida, buscou-se observar se e como a formulação da pergunta do questionário se enquadra na estratégia, decompondo-a. Na sequência, propõe-se a definição levando em consideração, sobretudo o fator sociocultural.

	Categoria (material)	Função	Fatores sociais			Fator cultural
			pessoal	localização	tempo	
Pergunta do QSL ALiB	<i>o brinquedo feito de uma forquilha e duas tiras de borracha,</i>	<i>usam para matar passarinho.</i>	<i>que os meninos</i>			
Proposta	Instrumento feito de forquilha e elástico	projetar pedras	(+ meninos)	(+ rural)	(+ infância)	usado como brinquedo por algumas comunidades
Definição	'Instrumento para projetar pedras , feito de forquilha e material elástico, utilizado, também, em algumas comunidades como brinquedo'.					

QUADRO 12: Confronto entre a pergunta do questionário ALiB e a proposta de semagrama
Fonte: Elaborado pela autora desta tese (2017)

No caso específico a *badogue*, a pergunta do questionário trata como brinquedo o que pode ser considerado arma. Nesse sentido, tornou-se necessário ponderar a manutenção do aspecto “brinquedo”, especificando, porém, o fator cultural de ser assim considerado e usado por algumas comunidades. A função e a categoria também foram alteradas.

É importante ressaltar que a função da pergunta é contribuir para a compreensão imediata do informante e que, portanto, atende suficientemente, aos critérios e objetivos propostos. No entanto, no que tange no âmbito lexicográfico, sabe-se que a definição precisa ser isenta de quaisquer imprecisões.

Por outra via, considerando os pressupostos da Lexicografia Variacional e da Lexicultura, e admitindo ser impossível recuperar a informação semântica quando se está longe do fato, a inserção da formulação da pergunta à microestrutura foi imprescindível, na sequência à definição, após a área temática do QSL, as quais,

por razões óbvias de metodologia, não poderiam deixar ser inseridas na microestrutura do verbete.

O verbete, pois, contempla além da ‘definição’ a ‘formulação da pergunta’, a fim de que o consulente possa conhecer todas as informações concernentes à lexia e ao seu registro, obtenção do dado.

6.1.1.3.2 Abonação

Outro aspecto lexicográfico-cultural relevante à pesquisa é a abonação que, no caso deste trabalho, se refere ao registro da lexia contextualizada, ou seja, um excerto em que se identifique a lexia na resposta do informante. Observem-se alguns exemplos:

Na tradição lexicográfica, as abonações costumam ser de autores literários consagrados e têm natureza prescritiva, enquanto na lexicografia moderna, de modo geral, são extraídas de um *corpus*, e têm caráter descritivo (PNLD, 2012 – Dicionários, p. 105)

Talvez, a relevância da abonação possa ser melhor evidenciada, se assim interpretada, no excerto do poema *Rios sem discurso*, de João Cabral de Melo Neto (1966¹), utilizada como epígrafe desta tese mas que aqui se decidiu mais uma vez ressaltar:

*Em situação de poço, a água equivale
a uma palavra em situação dicionária:
isolada, estanque no poço dela mesma,
e porque assim estanque, estancada;
e mais: porque assim estancada, muda
e muda porque com nenhuma comunica,
porque cortou-se a sintaxe desse rio,
o fio de água por que ele discorria.*

Assim, esse item se mostra extremamente relevante para a construção do Vocabulário proposto por propiciar ao consulente uma melhor aplicação dos usos contextuais da lexia documentada, além das possíveis caracterizações expressas pelos informantes.

¹ O poema pertence à coletânea Educação pela Pedra, publicada em 1966, pela Editora do Autor, Rio de Janeiro, mas foi encontrada na versão *on line* de Melhores Poemas João Cabral de Melo Neto, disponível em books.google.com.br Disponível em <books.google.com.br>. Acesso em: dia.mês.ano

6.1.1.3.3 Legenda geolinguística

Considerando o caráter sociodialetal do Vocabulário, tornou-se necessária a inserção de itens que reforçassem a descrição da realidade plurilinguística brasileira. Para tanto, buscou-se uma forma de representar, lexicograficamente a identificação da localidade e a descrição social, no que tange ao gênero e à faixa etária. Elaborou-se, pois a legenda geolinguística.

Para o DDB, desde 2010, tem-se adotado a seguinte legenda que leva em consideração as regiões administrativas.

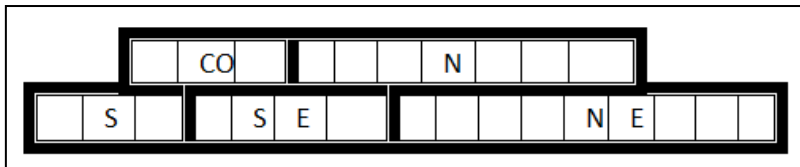


Figura 15 – Legenda geolinguística do DDB
 Fonte: (MACHADO FILHO; NEIVA, 2010)

Para o *Vocabulário Dialetal Baiano*, elaborou-se a legenda contemplando as 22 localidades que constituem o *corpus*, distribuindo-as conforme as mesorregiões baianas. A figura, a seguir, elucida o processo de elaboração da legenda geolinguística construída para o Vocabulário Dialetal Baiano.

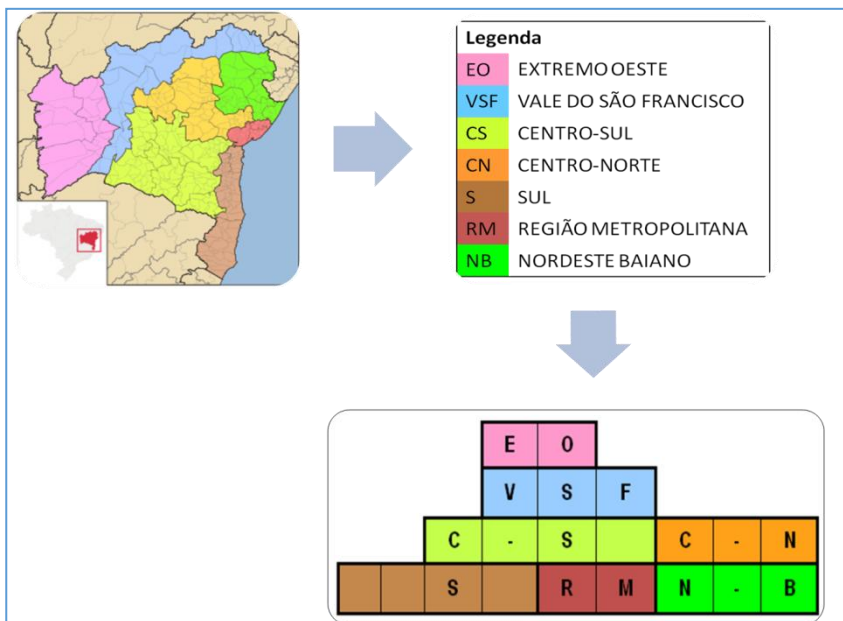


FIGURA 16 – Diagrama do processo de elaboração da legenda geolinguística do Vocabulário Dialetal Baiano.
 Fonte: Elaborada pela autora desta tese (2017).

6.1.1.3.4 Acheegas enciclopédicas

A achega enciclopédica é um item facultativo que compõe o verbete do *Vocabulário*, em que se apresentam a transcrição fonética e observações acerca de fenômeno fônico cuja representação gráfica não pode ser apresentada, ou seja, aqueles que embora façam parte da história da língua portuguesa e continuem em variação no português brasileiro, não se caracterizam sociodialetoalmente como “coisa da Bahia” ou mesmo do Nordeste ou não puderam ser representados graficamente na entrada de verbete. Nela podem ser incluídas, também, marcas de uso.

Considerando haver diferenças não somente na marcação de determinados lexemas e acepções, como também nos conjuntos de marcas adotados nos dicionários, Hausmann (1989) propõe uma divisão de marcas em 11 categorias: i) diacrônicas (*antiquado, envelhecido, neologismo*); ii) diatópicas (aplicadas a acepções restritas a certas regiões ou países); iii) diaintegrativas (usadas para assinalar estrangeirismos); iv) diamediais (diferenciam entre as linguagens oral e escrita); v) diastráticas (por exemplo, *chulo, familiar, coloquial, elevado*); vi) diafásicas (diferenciam entre as linguagens formal e informal); vii) diatextuais (assinalam que o lexema – ou acepção – é restrito a determinado gênero textual; por exemplo, *poético, literário, jornalístico*); viii) diatécnicas (informam que a acepção pertence a uma linguagem técnica, a um tecnoleto); ix) diafreqüentes (em geral: *raro, muito raro*); x) diavaliativas (mostram que o falante, ao usar o lexema, revela certa atitude, por exemplo *pejorativo, eufemismo*) e xi) dianormativas (indicam que o uso de certa acepção – ou lexema – é errado pelas normas da língua padrão)

Considerando o fato de que a diatopia já é marcada pela legenda geolinguística, optou-se por considerar apenas 5 das 11 citadas por Hausmann (1989), em nova perspectiva:

- diacrônicas (*quando a lexia for documentada também no APFB, seguida da sigla APFB*) conservadora
- diaintegrativas (quando a lexia, transcrita em português, for um item lexical estrangeiro – exemplo de *blusch, rouge, corpete* etc. símbolo + estrang.)
- diatécnicas (quando for um jargão – pertence a uma linguagem técnica, a um tecnoleto. símbolo + tec.)
- diafreqüentes (quando o uso da lexia for frequente, for raro.)

- diavaliativas (mostram que o falante, ao usar o lexema, revela certa atitude, por exemplo *pejorativo, eufemismo*)
- tabuísmo – símbolo + tabu. (eufemismos etc)

6.2 GLOSSÁRIO DE TERMOS TÉCNICOS USADOS NO VOCABULÁRIO DIALETAL BAIANO.

Com fins didáticos, para o prosseguimento da leitura, julgou-se ser conveniente apresentar alguns termos que serão utilizados, doravante. As definições subsequentes são adaptações do PNLD 2012- Dicionários e do Brevíssimo glossário do PNLD.

abonação - exemplificação que ilustra o emprego da entrada por meio de fragmentos de textos extraídos do *corpus*.

achega - informação que se acrescenta aos itens obrigatórios e opcionais / optativos de um verbete.

definição - codificação da informação semântica de um item lexical.

entrada - unidade lexical que é objeto das informações seguintes.

entrada falsa – representação lexicográfica que consiste em informar a inexistência da lexia no *corpus*. É, também, usada na canonização dos verbos, a fim de que se apresentem na forma infinitiva.

indicador – marca tipográfica que incidem sobre os itens presentes na microestrutura.

legenda geolinguística - representação lexicográfica das localidades estudadas divididas nas sete mesorregiões da Bahia.

lema principal - a entrada de uma lexia canônica. Variante mais produtiva no *corpus* ou a que se revele mais canônica.

lema secundário - variante lexical cuja alteração é de natureza fônica.

lema múltiplo - variante cuja alteração é de cunho mórfico ou morfossintático.

macroestrutura - design estrutural de um dicionário; projeto lexicográfico. Diferente de nomenclatura.

nomenclatura - conjunto de verbetes. Diferente de macroestrutura.

remissão semântica - qualquer relação semântica não sinonímica entre as lexias documentadas.

variante vocabular – variante cuja base lexical difere do lema principal

PARTE IV

ANÁLISE DOS DADOS

Capítulo 7

ANÁLISES DOS DADOS EM PERSPECTIVA DIALETOLÓGICA E (OU) LEXICOLÓGICA

7.1 “PALAVRA POR PALAVRA”: ANÁLISE DOS DADOS LEXICIS DOCUMENTADAS

A análise dos dados é aqui realizada com base na proposta de léxico em rede ou léxico rizomático, em que se englobam as estratégias de elaboração e estudos dos campos lexicais, semânticos e conceituais.

Considerando que há,

em diversos trabalhos da área, uma tendência a se confundir em linguística a noção de campo semântico com a de campo conceitual, o que deve ser retificado. Enquanto este se refere a um “conjunto informal e heterogêneo de problemas e situações, conceitos, relações estruturas, conteúdos e operações de pensamento, conectados uns aos outros entrelaçados durante o processo de aquisição” (Vergnaud, 1982, *apud* Moreira, 2002, 8), aquele, às possibilidades significativas que uma dada unidade lexical possa assumir em diferentes contextos, ou seja, sua capacidade polissêmica. (MACHADO FILHO, 2015, no prelo)

Com base na citação supracitada, ‘campo conceitual’ pode ser entendido e também denominado ‘área temática’, como faz o ALiB. Esse fato, além de demonstrar como as mudanças sociolinguísticas condicionam o posicionamento teórico das pesquisas, propicia diferentes estratégias de análise.

Considerando o fato de que existem, atualmente, diversos trabalhos, monografias, dissertações e teses com o *corpus* do ALiB, como fora citado na Metodologia, optou-se por analisar a variação lexical documentada a partir das respostas obtidas das 202 questões do QSL ALiB, concernentes a 14 áreas temáticas.

Para melhor apreciação dos resultados, tornou-se relevante estabelecer correlações entre as áreas temáticas do ALiB, a partir da elaboração de

microáreas temáticas. A análise dos dados desenvolve-se, pois, por áreas temáticas¹ (subdivisões das áreas temáticas) e por mesorregiões.

Com vistas a revelar a estrutura e o comportamento léxico-variacional de cada um dos conceitos que fazem parte da microárea temática, e, por conseguinte, mostrar a ação centrípeta e centrífuga na norma em relação ao uso das unidades lexicais, foram elaborados gráficos que propiciassem a representação esquemática da variação.

Para tanto, optou-se pelo modelo de gráfico ‘radar’ com escala invertida. Desse modo, o resultado da frequência de uso da lexia implica na aproximação ao eixo e esse fato pode ser interpretado como a fixação da lexia na língua (na norma). A situação inversa, estar mais distante do eixo, pode pressupor idioleção, mas também pode ser interpretada como resultado da ação centrífuga da norma deixando às margens da estrutura gráfica certas lexias, as quais, não por acaso, também podem coincidir com aquelas que se caracterizam como usos menos prestigiados, na estrutura social.

Pressupõe-se, pois, que as lexias incomuns, as que apresentem modificações metaplásmicas, as estruturas fraseológicas e as resultantes de processos de formação – derivacional e composicional – estejam espalhadas. No entanto, é possível que uma variante canônica possa estar “pendendo” seu lugar de maior fixação na língua em função de outra, independentemente de sua composição estrutural. Esse fato caracterizará, efetivamente, a ação centrípeta / centrífuga da norma.

Pressupõe-se, pois, que as variantes canônicas estejam mais próximas do eixo e que ocupem, portanto, lugar de maior fixação na língua e que as lexias incomuns, as que apresentem modificações metaplásmicas, as estruturas fraseológicas e as resultantes de processos de formação – derivacional e composicional – estejam espalhadas.

Os gráficos podem revelar, ainda, o trânsito possível entre os espaços, de modo que uma variante canônica possa “perder” seu lugar de maior fixação na língua em função de outra, independentemente de sua composição estrutural.

¹ Por questões didáticas, no corpo do texto da tese, apenas se apresentam alguns dos elementos, mais especificamente, aqueles que revelaram traços significantes de análise. Todos os 202 gráficos concernentes a cada uma das questões que compõem o QSL ALiB são apresentados, nos Apêndices do Volume I da tese.

7.1.1 A variação lexical por áreas temáticas: Acidentes geográficos

“Nas águas do rio, homens e mulheres vão desenhando suas histórias de vida”. (OLIVEIRA, 2009, p. 15). Dita sobre o rio São Francisco, essa frase pode ser aplicada, também, a todas as comunidades ribeirinhas, em que se incluem os mais de 30 rios que recortam a Bahia.

Considerando que cinco das seis perguntas que compõem a área temática *Acidentes geográficos*, constante do QSL ALiB, referem-se a rios – com ênfase no tamanho, curso e movimentos –, para uma análise mais criteriosa, tendo em vista a dimensão da proposta para este trabalho, optou-se por, apenas, analisar dados linguísticos e sociolinguísticos.

Como respostas à Questão 01 do QSL, *Como se chama um rio pequeno, de até dois metros de largura?*, foram obtidas como respostas as seguintes variantes:

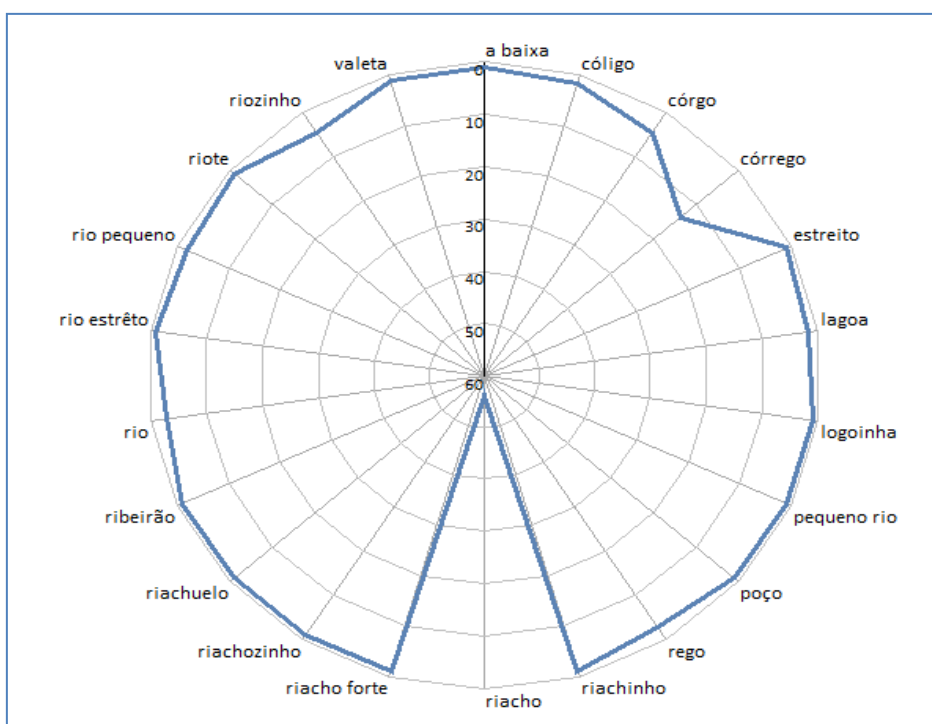


GRÁFICO 1 – Distribuição do comportamento léxico-variacional de 'Rio de pequena extensão'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Compreendendo que quanto mais próximo ao eixo, maior a frequência de uso da lexia, observa-se que **riacho** é a mais predominante (60 ocorrências), seguida de **córrego**, embora para esta apenas se registrem 15 realizações.

Outro dado relevante que se pode depreender dos resultados apresentados no gráfico refere-se ao léxico periférico. Observe-se que o espraiamento da variação se revela na alternância de unidades léxicas as quais se caracterizam i) pela presença de metaplasmos, como ocorre em *cóligo* e *córgo* equivalentes de *córrego*; ii) pelo uso de estruturas fraseológicas como *pequeno rio*, *rio estrêto*, *riacho forte*; iii) pelo uso de lexias constituídas pelo processo de derivação, mais especificamente, na construção de diminutivos, como ocorre em *riote*, *riozinho*, *riachozinho* e *riachinho*.

Com vistas a perceber a relação do homem com o rio, em função de agente construtor, apresentam-se os resultados obtidos como respostas à questão 02 do QSL, cuja formulação “[...] tronco, pedaço de pau ou tábua que serve para passar por cima de um _____ (cf. item 1)?” expressa o caráter improvisador com que é feita a estrutura de madeira.

“E foi assim que”, paralelamente às *pontes* e *pinguelas* – variantes mais produtivas, com respectivas 31 e 16 ocorrências –, foram construídas outras 26 lexias documentadas nas respostas dos informantes, as quais podem ser visualizadas no gráfico, a seguir.

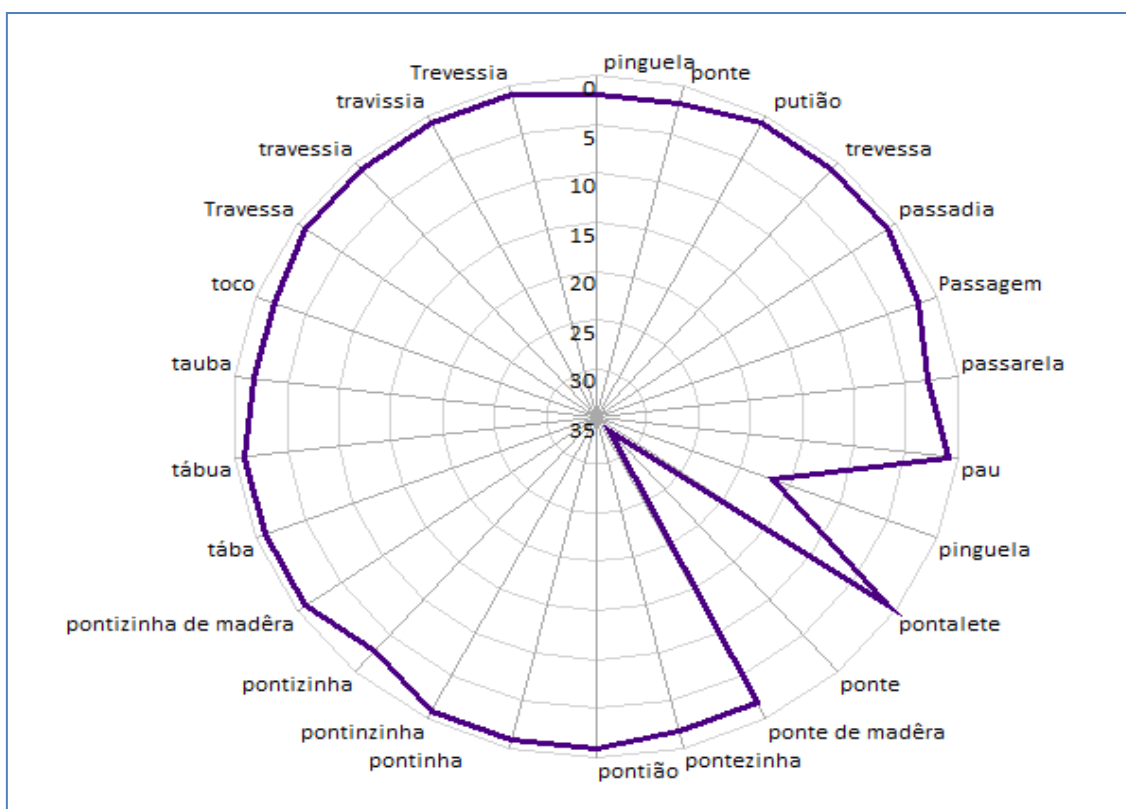


GRÁFICO 2 – Distribuição do comportamento léxico-variacional de 'Qualquer elemento de madeira que sirva de passagem de pessoas por cima de um rio pequeno'.
Fonte: Gráfico elaborado pela autora desta tese (2017)

Dentre as variantes se destacam alterações metaplásmicas como em *travessia*, *treveçia*, *taba* e *tauba*. Observa-se, ainda que da variante mais produtiva *ponte* são produzidas outras variantes resultantes de alterações mórficas como ocorre em *pontinha*, *pontezinha* e morfossintáticas com a inserção de elemento especificador, como ocorre em *ponte de madêra* e *pontezinha de madêra*.

Para o referente 'lugar onde o rio termina ou se encontra com outro rio', a lexia *barra* é a preponderante. *Foz*, a forma canônica, no entanto, parece sofrer ação centrífuga que a leva para a margem, no caso, ao desuso, em função de outras formas que passaram a ocupar o seu lugar, revelando a ação centrípeta da norma, como se observa no gráfico, a seguir:

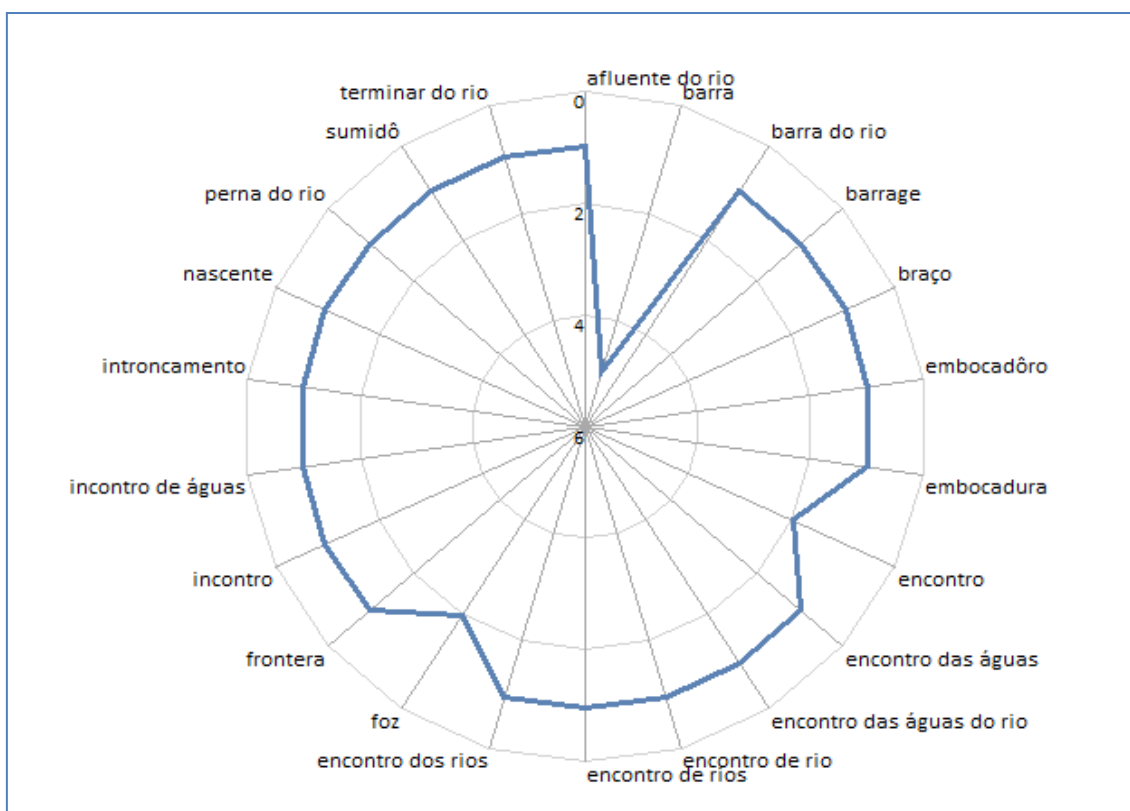


GRÁFICO 3 – Distribuição do comportamento léxico-variacional de 'Ponto de término de um rio ou de desaguamento em outro'.

Fonte: Gráfico elaborado pela autora desta tese (2017)

Para o 'fenômeno natural que implica no movimento em rotação espiral, ocasional / eventual, que forma um funil na água', foram documentadas 30 lexias distintas, das quais se destacam as alterações metaplásmicas. Observe-se:

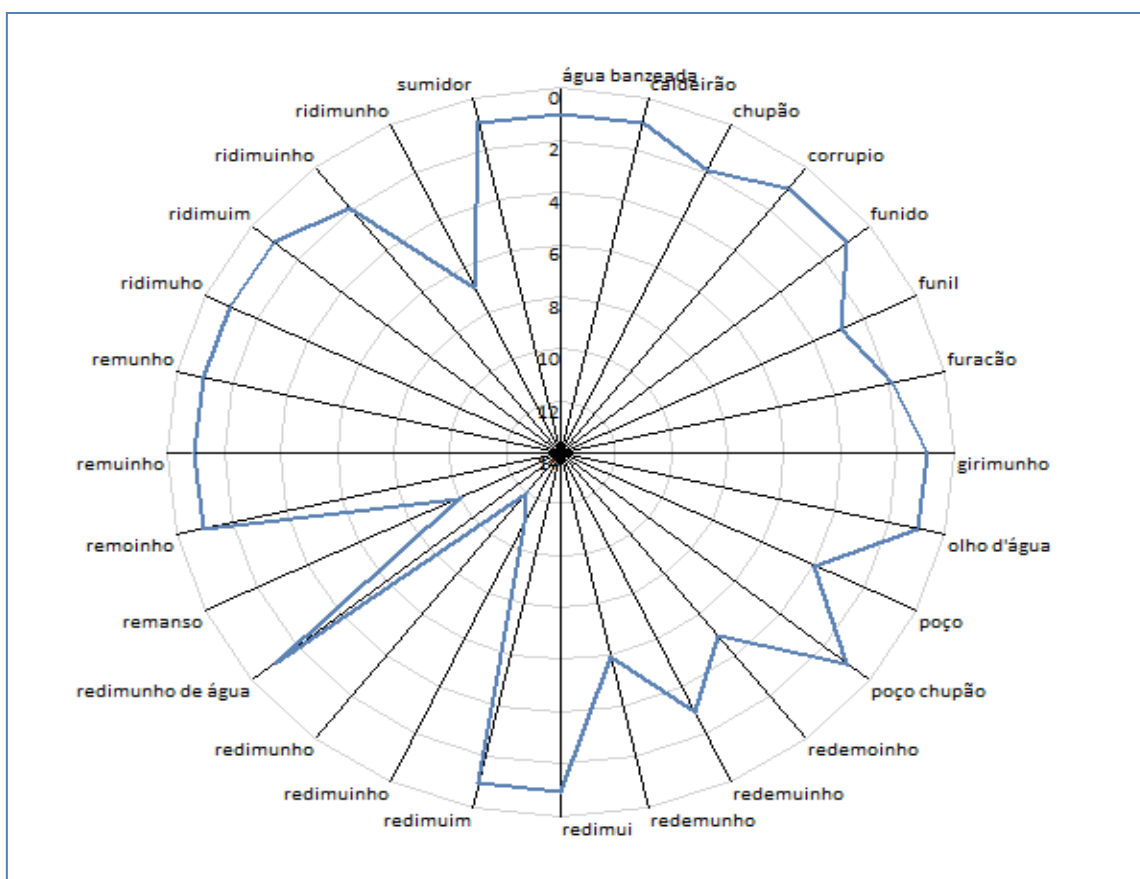


GRÁFICO 4 – Distribuição do comportamento léxico-variacional de 'Fenômeno natural que implica no movimento em rotação espiral, ocasional / eventual, que forma um funil na água.'

Fonte: Gráfico elaborado pela autora desta tese (2017)

7.1.1 A variação lexical por áreas temáticas: Atividades agropastoris

A área temática Atividades agropastoris conta com 25 questões (do QSL 39 ao QSL 63), dentre as quais se apresentam alimentos, frutas, características, instrumentos agropastoris, dentre outros.

Foram documentadas 11 lexias obtidas como respostas ao QSL 39², que tinha como resposta prevista a variante *tangerina*. Dentre tais, *tangerina*, *mixirica* e *pocã* são as variantes mais produtivas.

² Como se chama as frutas menores que a laranja que se descasca com a mão e, normalmente, deixam um cheiro na mão? Como elas são?

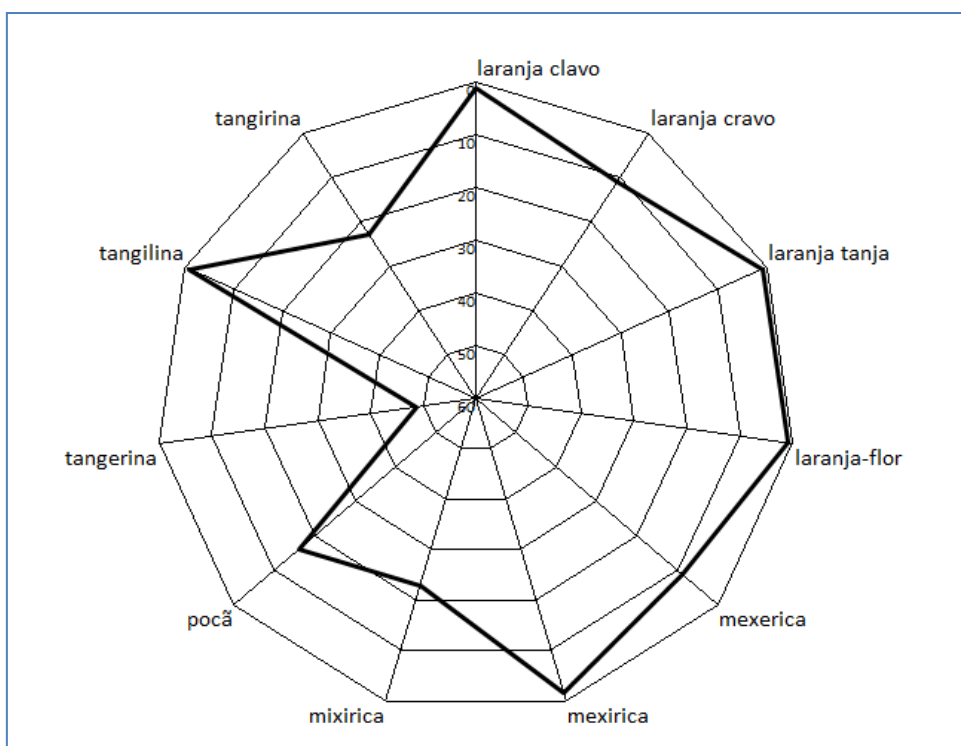


GRÁFICO 5 – Distribuição do comportamento léxico-variacional de 'Fruta cítrica, pouco ácida, de cheiro forte, cuja casca se solta, facilmente, dos gomos, podendo ser descascada com as mãos'

Fonte: Gráfico elaborado pela autora desta tese (2017)

É válido ressaltar, porém, que em algumas localidades, *pocã* se registra como diferente de *tangerina*, pelo tamanho ou pelo sabor.

Lexicograficamente, esse fato é relevante para a atribuição ou não das entradas lexicais. Nesse caso, especificamente, no Vocabulário, contam *pocã*¹ e *pocã*², sendo apenas a primeira, uma variante de *tangerina*.

Para o QSL 40, previa-se a resposta *amendoim*. Como se observa, foram obtidas formas que se caracterizam por alterações metaplásmicas de *amendoim*.

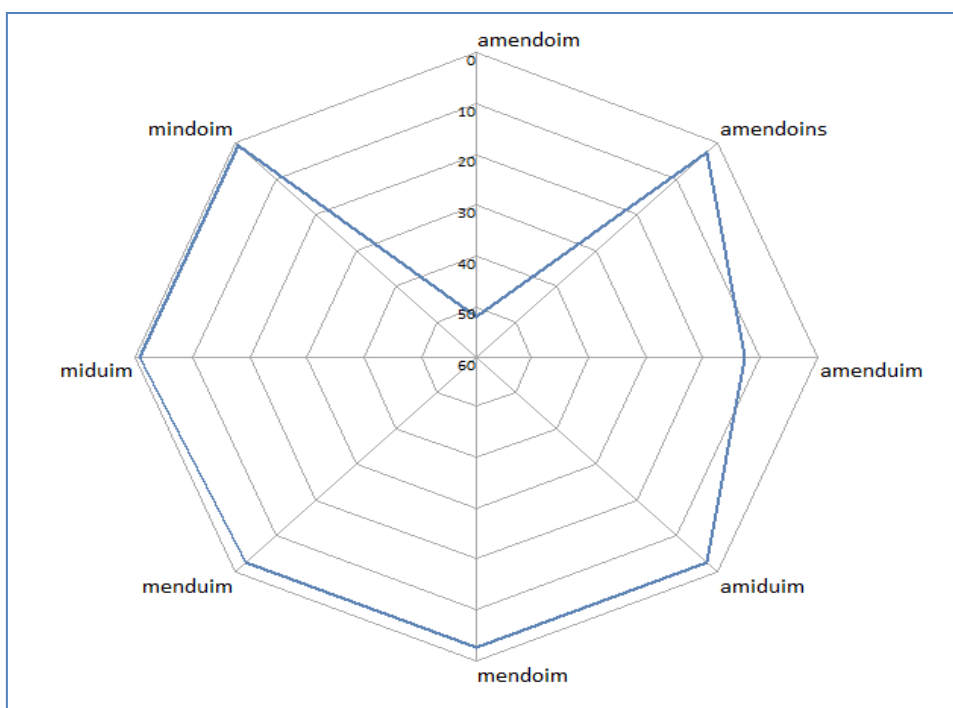


GRÁFICO 6 – Distribuição do comportamento léxico-variacional de ‘Leguminosa de formato longitudinal e aspecto gômico, em cujo interior constam sementes de cor avermelhada usada na alimentação humana em diversas formas de cocção.’

Fonte: Gráfico elaborado pela autora desta tese (2017)

Há, atualmente, uma tendência a considerar essas alterações como variação fônica. Sob a perspectiva do léxico, mais especificamente, sob a perspectiva da Lexicografia Variacional, entretanto, essas formas que podem ser admitidas como variantes lexicais se se considerada a distinção entre léxico de língua e léxico de norma, já comentada.

Fato similar ocorre com as respostas do QSL 41, como se pode observar no gráfico, a seguir:

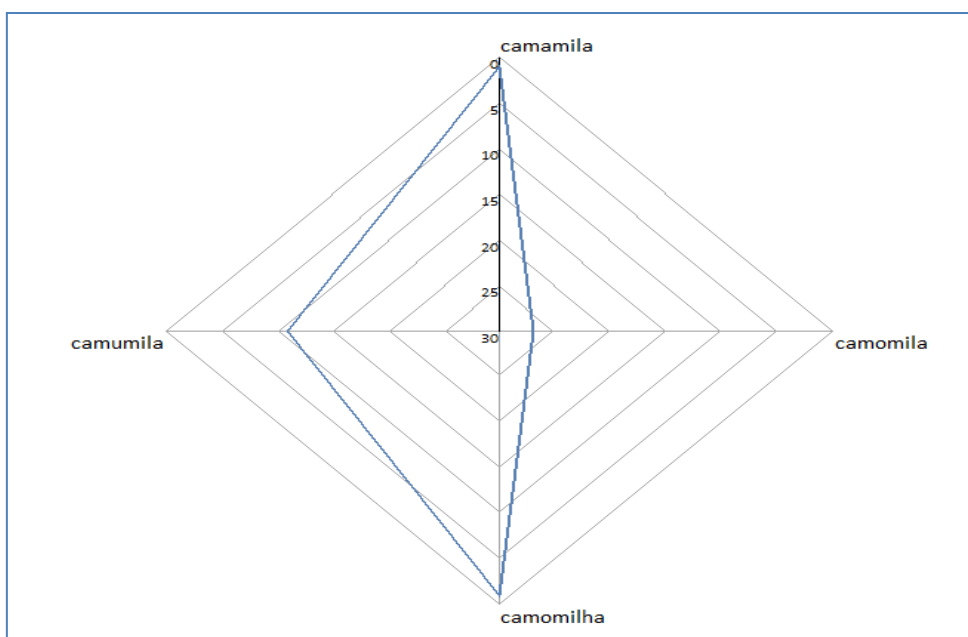


GRÁFICO 7 – Distribuição do comportamento léxico-variacional de 'Tipo de flores brancas e amarelas com as quais se prepara uma infusão empregada como calmante e digestivo.'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Para o referente 'penca de banana', QSL 42, foram obtidas como respostas 5 lexias distintas, das quais a mais produtiva é penca.

As alterações mórficas *cacho da banana* e *cacho de banana* são ressaltadas aqui pela importância que possam ter lexicalmente.

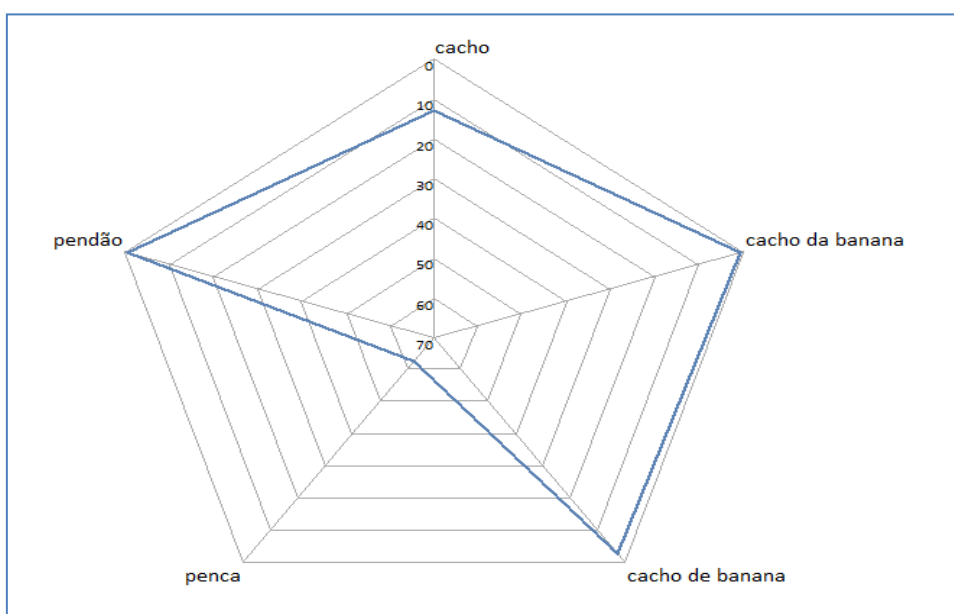


GRÁFICO 8 – Distribuição do comportamento léxico-variacional de 'Cada um dos grupos de cachos de banana'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Tipicamente tropical, a bananeira é uma planta rústica e produtiva, cultivada em quase o Brasil. No panorama nacional destacam-se Bahia e São Paulo como os maiores produtores, com cerca de 23% da produção brasileira. A população brasileira consome mais de 40 kg por habitante de banana.

Dentre as partes dessa planta, destaca-se, aqui como mostra a imagem ao lado –, a inflorescência – espécie de espiga simples, terminal, que emerge do centro das bainhas foliares – sobretudo em função dos diversos nomes a ela atribuída – flor, coração, umbigo, dentre outras.

Observe-se:

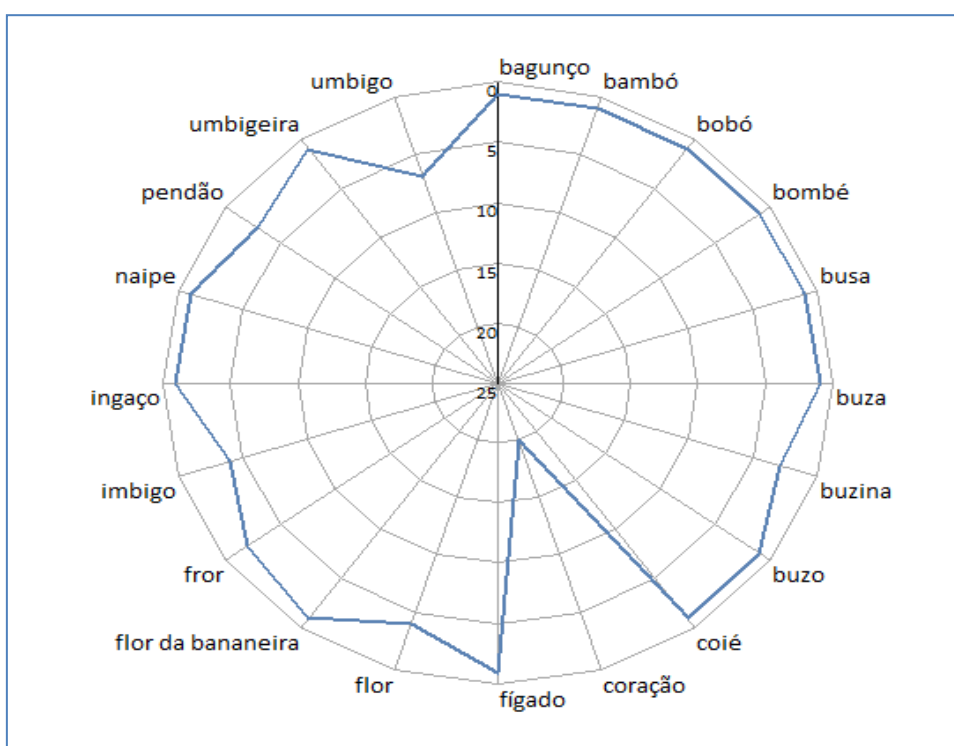


GRÁFICO 9 – Distribuição do comportamento léxico-variacional de 'Inflorescência da bananeira'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Confrontando com os dados constantes nas cartas do APFB, notam-se um total de nove variantes distintas – *coração*, *búzio* ou *búzia*, *umbigo*, *buzina*, *ingaço*, *naipe* e *pendão*, *bagunço* e *bombó*. Diante do que se observa, evidenciam-se:

- a) desuso das formas *bagunço* e *bombó*, por apenas se registrarem no APFB;

- b) presença de lexias *buzina*, *ingaço*, *naipe* e *pendão*, somente no ALiB, caracterizando inovação;
- c) coincidência das formas *búzio*, *coração* e *umbigo* entre os atlas, das quais pôde-se observar:
- i. forte redução de *búzio*, forma que, no APFB, se documenta em quatro localidades – Cabrália, Itaberaba, Jeremoabo e Santana. No ALiB, há o registro de *buza*, que pode ser inferido como uma variante de *búzia* em que houve erosão fonológica;
 - ii. ascendência do uso de *coração* – a forma que, no APFB, só se registrou em Barra, no ALiB passa a ser documentada também em Caetité, Carinhanha, Conquista, Itaberaba e Santana;
 - iii. permanência de *umbigo*, em Jacobina. Essa lexia pode ser considerada inovadora em Cabrália e em Santana, haja vista somente se registrar nos dados do ALiB.

Para o referente ‘banana que nascem grudadas’, foram documentadas 24 lexias. Como se observa, *gêmias* é a forma mais produtiva (mais próxima do eixo), seguida de *felipe*. Observe-se, a seguir, a configuração esquemática da variação lexical.

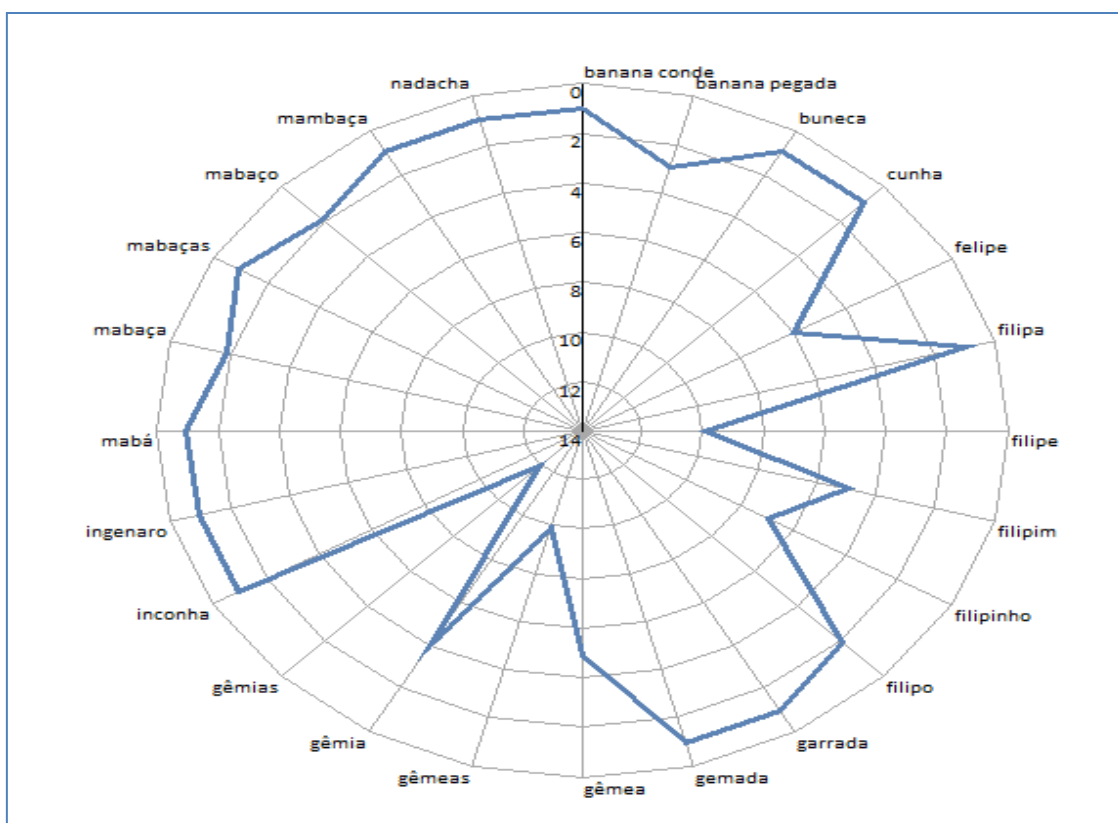


GRÁFICO 10 – Distribuição do comportamento léxico-variacional de ‘Bananas que nascem grudadas’

Fonte: Gráfico elaborado pela autora desta tese (2017)

Se se comparar com as designações dadas às crianças que nascem no mesmo parto – constante do QSL 125, referente à área Ciclos da vida –, nota-se que o traço [+ humano] pode interferir na oscilação de formas, de modo que as variantes *felipe* e suas alterações metaplásmicas apenas são utilizadas para designar ‘banana que nascem pegadas’ mas não são utilizadas para as crianças que nascem no mesmo parto. Observe-se:

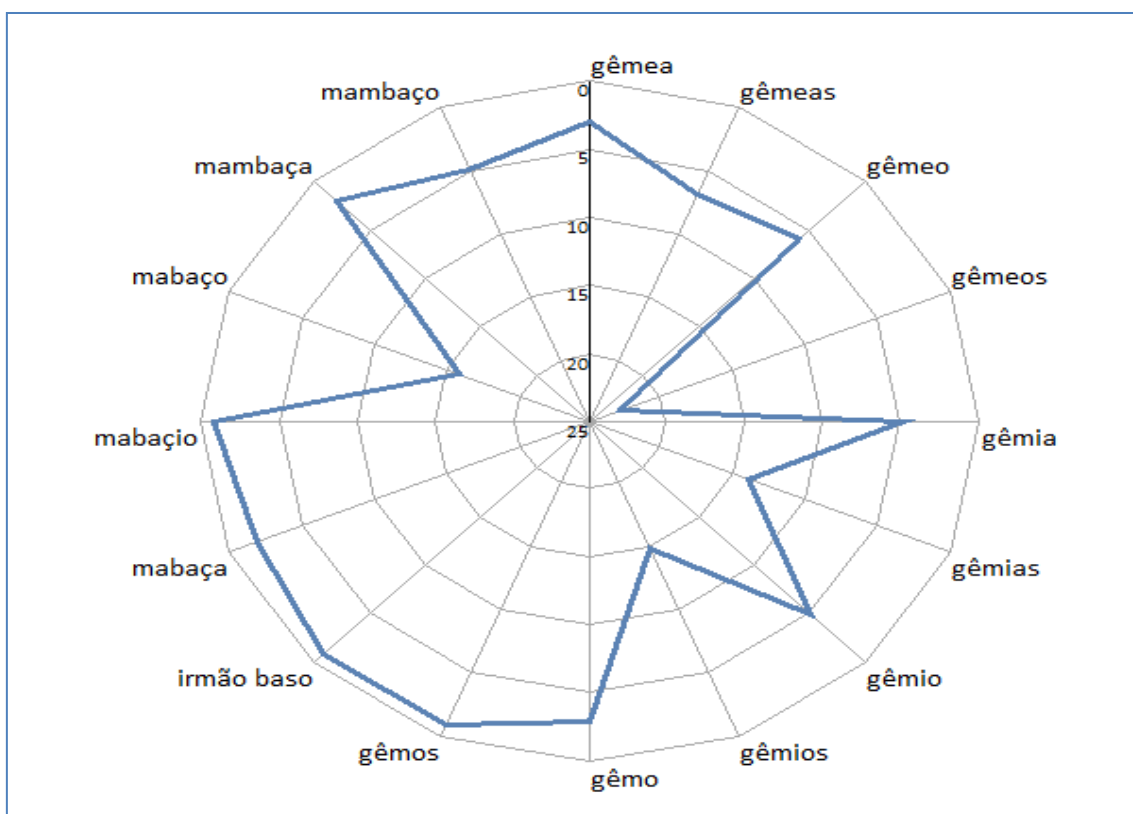


GRÁFICO 11 – Distribuição do comportamento léxico-variacional de 'Gêmeos'
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Passando-se às questões 50 e 51, registram-se, respectivamente as formas **aipim**, **aimpim**, **impim**, **mandioca¹**, **mandioca-doce**, **macaxeira** e **macaxera**, como se apresenta no gráfico , e **mandioca²**, **mandioca azeda** e **mandioca-brava**, presentes no gráfico

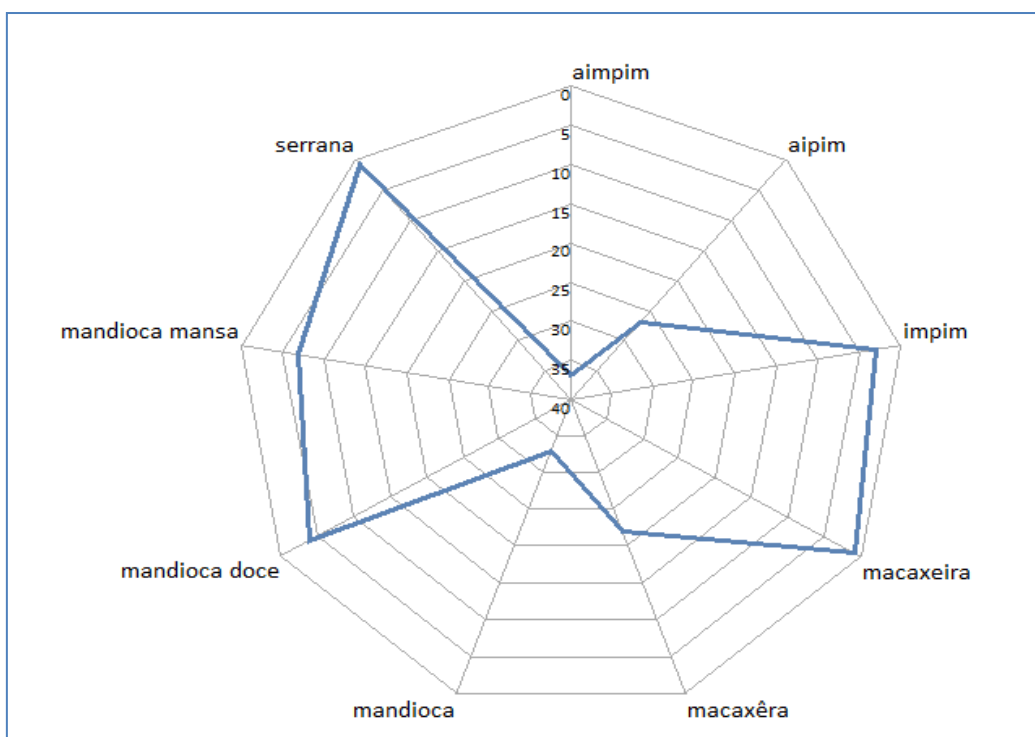


GRÁFICO 12 – Distribuição do comportamento léxico-variacional de 'Raiz comestível, de cor branca por dentro e coberta por uma casca marrom'.

Fonte: Gráfico elaborado pela autora desta tese (2017)

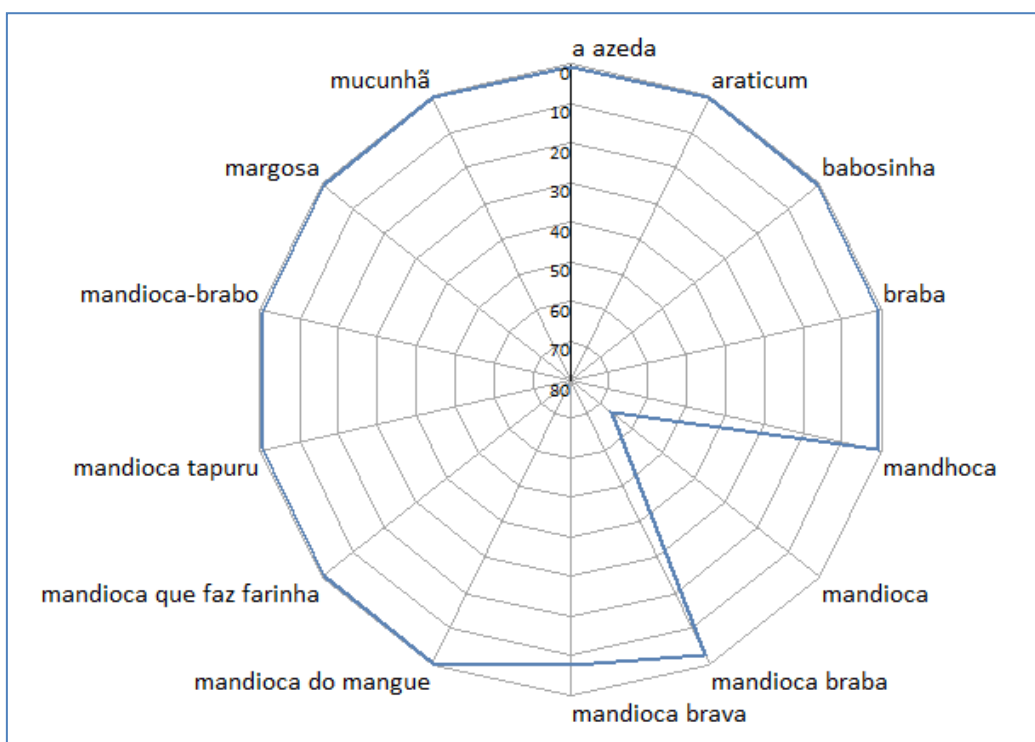


GRÁFICO 13 – Distribuição do comportamento léxico-variacional de 'Raiz branca por dentro, coberta por uma casca marron destinada, exclusivamente, à elaboração de farinha'.

Fonte: Gráfico elaborado pela autora desta tese (2017)

Da observação de ambos os gráficos, podem-se ressaltar:

- a) A correlação entre as lexias **mandioca¹** e **mandioca²**, **mandioca doce** e **mandioca azeda**, **mandioca mansa** e **mandioca brava**, expressa a característica de ser comestível ou não.
- b) A presença de metaplasmos comuns no português brasileiro como aférese (**aipim** ~ **impim**) nasalização do ditongo -ai, (**aipim** ~ **aimpim**).

A área Atividades agropastoris conta com questões que se referem a utensílios comuns ao âmbito rural agropastoril. Dentre tais, carrinho de mão, bruaca e balaio, canga, cangalha e forquilha.

Como variantes de 'carrinho de mão', foram documentadas 16 lexias, das quais 13 são formadas pelo radical '**carr_**', apresentando alterações metaplásmicas como em *carrim*, móficas como em carrinho e morfossintáticas, com a inserção de especificadores, a exemplo de carro de mão, carrinho de ferro.

Destacam-se, também, a presença de variantes *galinhota*, *galiota* e *galinota*.

É válido ressaltar que alguns informantes fazem diferença entre os referentes quanto ao material do qual é feito, de modo que para alguns galinhota é feito de madeira e carrinho de mão, feito de ferro.

Esse fato é relevante para a constituição dos verbetes. Foi necessário, pois, distinguir acepções, a partir da elaboração de entradas distintas para cada referente, levando em consideração o traço material [madeira x ferro].

Observe-se, a seguir, a distribuição dos itens lexicais:

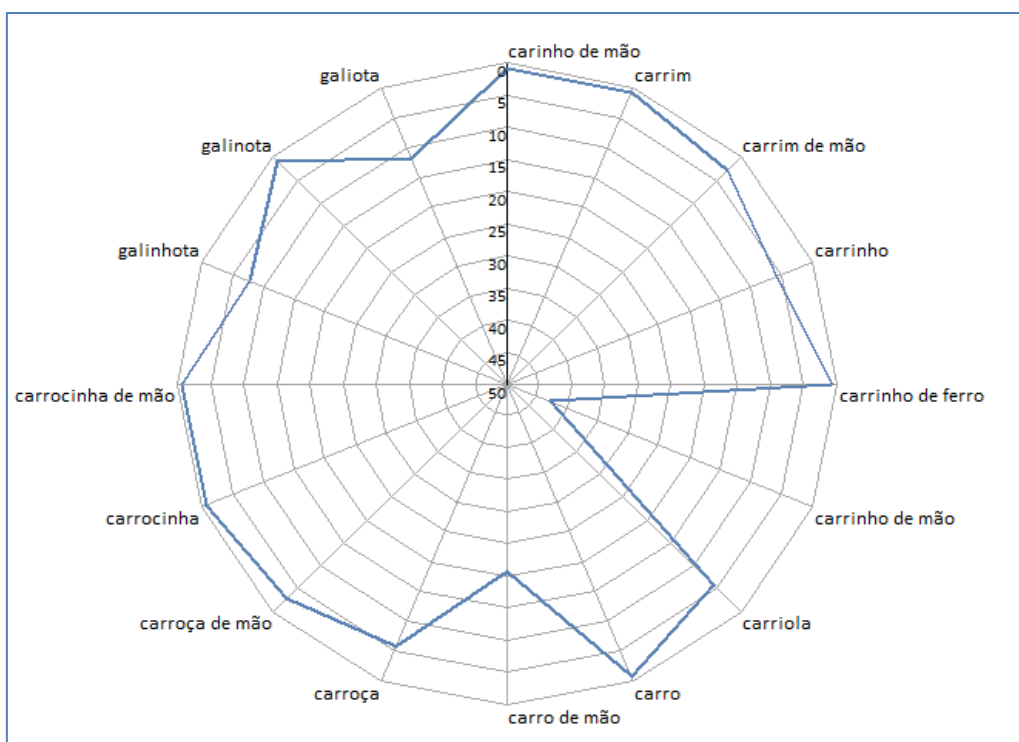


GRÁFICO 14 – Distribuição do comportamento léxico-variacional de 'Veículo de propulsão humana, constituído de uma única roda dianteira e duas hastes na parte oposta usado para transporte de pequenas cargas'.

Fonte: Gráfico elaborado pela autora desta tese (2017)

7.1.2 A variação lexical por áreas temáticas: Fauna

A área temática Fauna constitui-se de 26 questões, das quais se pode obter designações para aves, a mamíferos terrestres e insetos ou invertebrados. Há também, questões que se referem às partes do corpo dos animais. Algumas destas poderiam ser comparadas com partes do corpo humano, de modo a refletir se o léxico utilizado seria diferente se se considerado o traço [+humano]

Por questões didáticas, no corpo do texto desta tese, apenas se apresentam alguns dos elementos, mais especificamente, aqueles que revelaram traços significantes de análise.

Para a presente análise, selecionaram-se as questões: i) QSL 68 cuja resposta prevista é papagaio, em função das relações semânticas possíveis com outros referentes, os quais, inclusive, pertencem a outras áreas temáticas; ii) QSL 67 e 69, por terem sido incluídas no APFB.

Como respostas para o referente ‘ave de penas coloridas que aprende a falar’, foram obtidas 9 lexias distintas, dentre as quais *papagaio* se revela preponderante. Lôro é a segunda variante mais falada.

Destaca-se a alteração mataplásmica em *papagai* e *papagalho*.

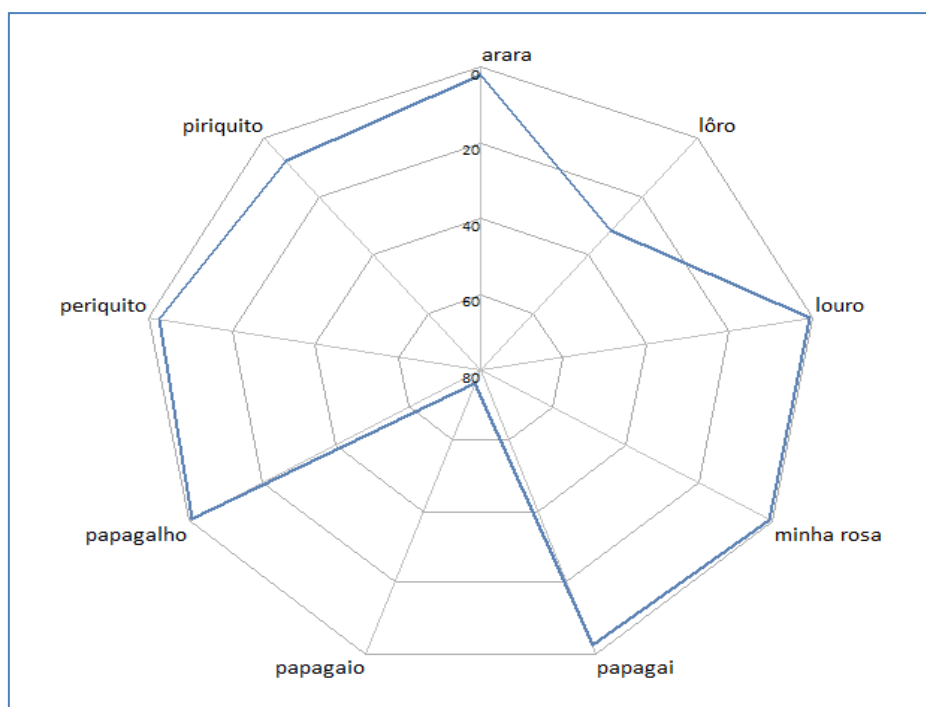


GRÁFICO 15 – Distribuição do comportamento léxico-variacional de ‘Ave de penas coloridas que quando domesticado, pode emitir sons como se na linguagem humana’.

Fonte: Gráfico elaborado pela autora desta tese (2017)

Para além da observação dos itens lexicais que se apresentam nesta rica área temática, considerando a biodiversidade e a diversidade sociolinguística e cultural do e no Brasil, buscou-se investigar como a fauna brasileira tem influenciado e (ou) se refletido no léxico, mais especificamente, através da apropriação de designações de animais no uso de lexias, de frases e de possíveis culturemas.

Dentre as designações fáunicas, *papagaio* se destaca pela relação semântica com outros referentes. Papagaio é uma resposta dada ao brinquedo e também é usado para designar a pessoa que fala demais.

Como brinquedo, *papagaio* é usado em Caetité, Caravelas, Euclides da Cunha, Juazeiro, Santana, Santo Amaro, Seabra e Vitória da Conquista. Dentre

tais localidades, é possível dizer que se caracteriza como um uso da mesorregião Centro-Sul da Bahia.

Como dito anteriormente, na Bahia, conforme os dados, *papagaio* também é utilizado para designar a pessoa que fala demais. Observe-se que há inserção de elementos como “papagaio de marinheiro” e como “que nem papagaio” e “igual a um papagaio”.

Observem-se, a seguir, amostras de verbetes de ‘papagaio’, os quais constam do Volume 2 do *Vocabulário Dialeto Baiano*.

papagaio¹ `Ave, similar a uma galinha, de penas pretas e pintas brancas. //...
QSL 068 - Como se chama a ave de penas coloridas que, quando presas, podem aprender a falar? ~ papagai
 ~ **papagalho** ≡ **arara** ≡ **louro** ~
lôro ≡ **minha rosa** ≡ **periquito** ~
piriquito → **papagaio²** →
papagaio³ {É porque *pã* velho num tem coisa melhó que um **papagaio**, né. Só quem pode tulerá velho é um papagaio} **088.3**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

FIGURA 17 – Verbetes de ‘papagaio’ do *Vocabulário Dialeto Baiano*, relativo à Fauna. Fonte: NEIVA (2017)

papagaio² ↔ **piriquito(inho)** →
papagaio¹ → **papagaio³** →
papagaio⁴

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

papagaio³ ↔ **pipa(s)** → **papagaio¹** →
papagaio³ → **papagaio⁴**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

FIGURA 18 – Verbetes de ‘papagaio’ do *Vocabulário Dialeto Baiano*, relativo à Brinquedos e diversões infantis. Fonte: NEIVA (2017)

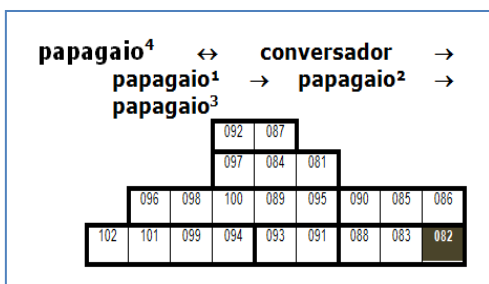


FIGURA 19 – Verbo de ‘papagaio’ do *Vocabulário Dialectal Baiano*, relativo à Convívio e comportamento social.

Fonte: NEIVA (2017)

O destaque para gambá, nesta análise refere-se ao fato de que se trata de um dos referentes com menor índice de variação. Foram documentadas apenas 6 realizações de *saruê*, as quais ocorrem em Caetité, Caravelas, Euclides, Irecê e Santana e 1 de *sariguê*, em Jacobina.

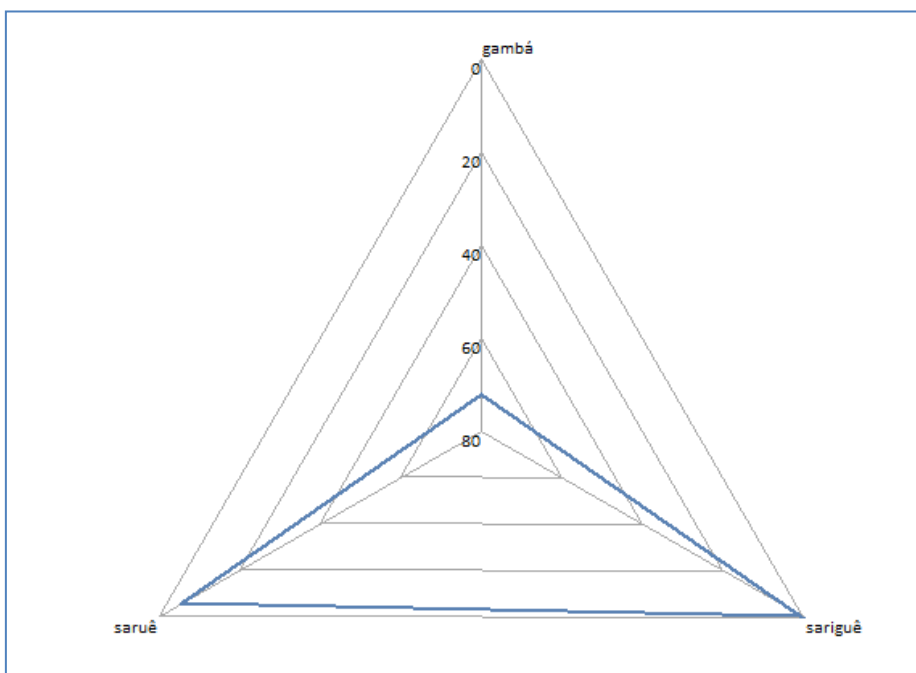


GRÁFICO 16 – Distribuição do comportamento léxico-variacional de ‘Animal mamífero que, quando se sente ameaçado, emite fedor.’

Fonte: Gráfico elaborado pela autora desta tese (2017)

Para a presente análise a questão 80 do QSL – “Em que parte da vaca fica o leite?” – foi selecionada a fim de discutir acerca da apresentação das

variantes léxicas que se caracterizam como registros fônicos de outras e como essas lexias seriam apresentadas no vocabulário como lemas secundários.

Observe-se, a seguir, a distribuição de todas as variantes documentadas.

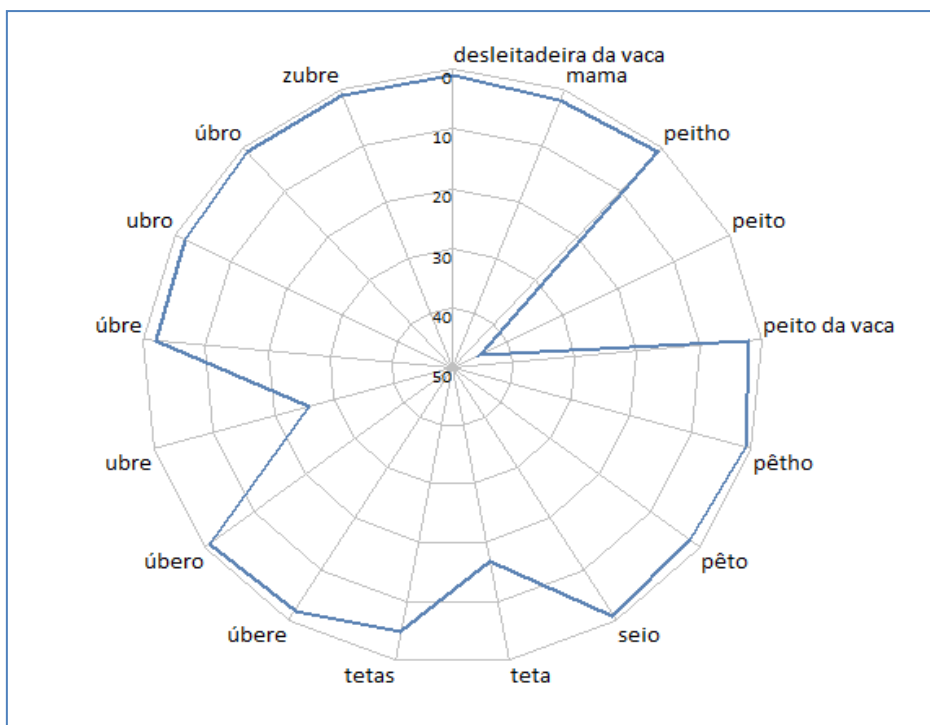


GRÁFICO 17 – Distribuição do comportamento léxico-variacional de 'Glândula mamária da vaca'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Como se percebe, a forma mais produtiva é *peito* (40 ocorrências), seguida das variantes *ubre* (forma erosiva de 'úbere') e *teta* (respectivamente, 24 e 14 realizações).

Como fora visto a forma **peito** é predominante. Mas, para além desse uso, foram documentadas formas que se caracterizam como lemas secundários de *peito* – **peitho** e **pêto** – além de **peito da vaca** que pode ser interpretada como lema múltiplo. Observe-se, a seguir, como tais formas se distribuem:

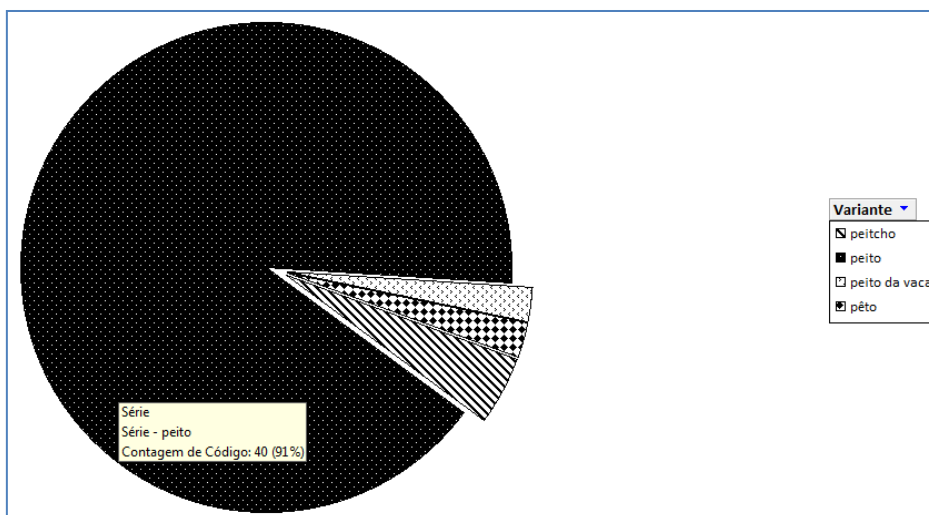


GRÁFICO 18 – Distribuição das variantes secundárias de ‘peito’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Observe-se, ainda, que embora a forma canônica “úbere” não tenha sido documentada no *corpus*, há cinco formas que, por analogia, a sugere: **úbero**, **ubre**, **ubro** e **zubre**. O gráfico 16, mostra a distribuição dessas variantes, em que a forma **ubre** é predominante.

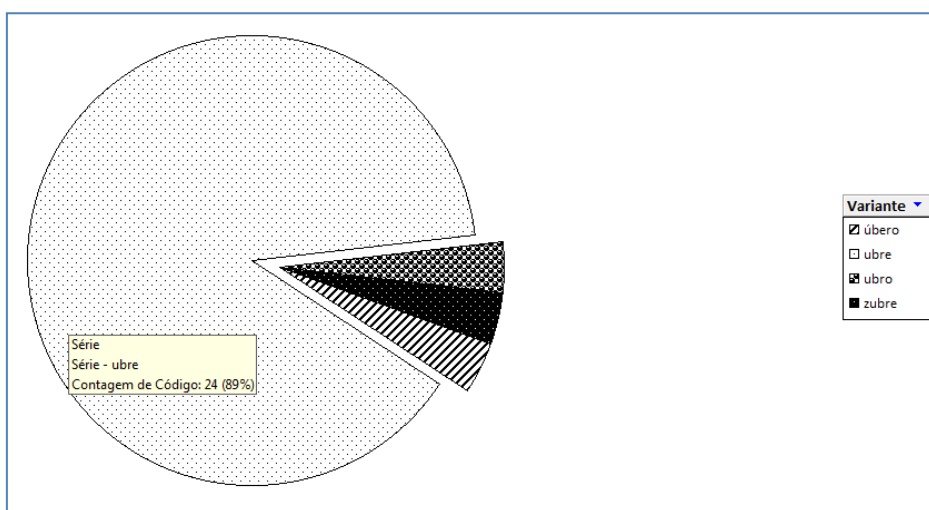


GRÁFICO 19 – Distribuição das variantes secundárias de ‘úbere’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Vale destacar que o registro lexicográfico dessas variantes, além do respeito às variantes menos prestigiadas e, por vezes, consideradas estigmatizadas, pode ser admitida como uma maneira de salvaguardar os usos atuais com vistas à preservação da história da língua, a fim que não se percam com o tempo, já que se tratam de realizações de fala.

O destaque é feito aqui em função das repostas serem muito próximas à formulação da pergunta: “Como se chamam as patas dianteiras do cavalo?”.

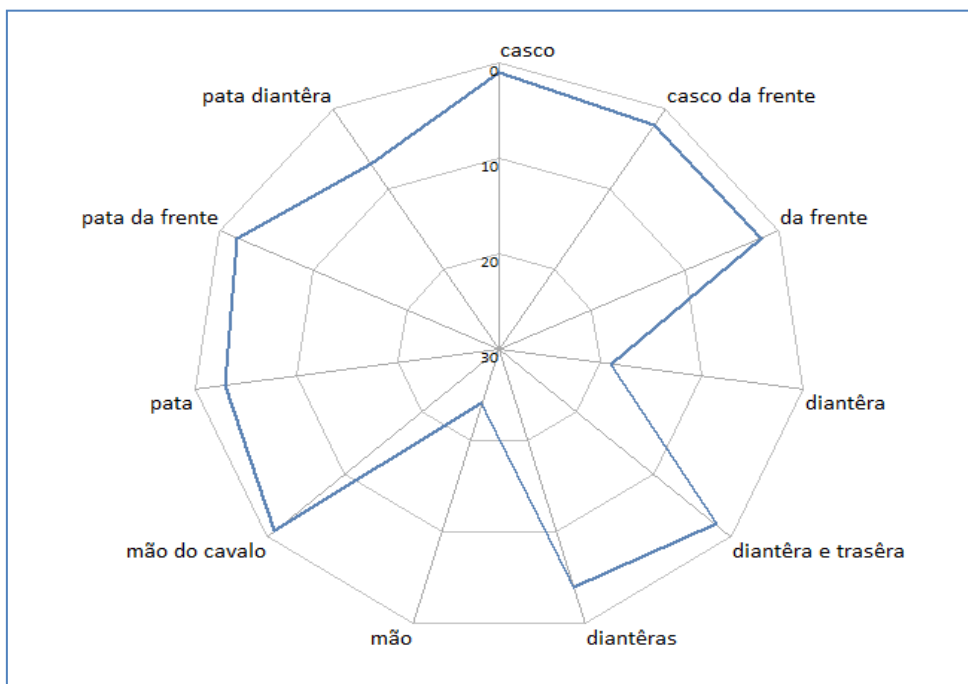


GRÁFICO 20 – Distribuição do comportamento léxico-variacional de 'Patas dianteiras do cavalo'

Fonte: Gráfico elaborado pela autora desta tese (2017)

O destaque aqui concerne ao alto índice de alterações metaplásmicas: 11 lexias distintas. Ainda mais relevante é ver que a lexia canônica *sanguessuga* tem “perdido lugar” para seu lema secundário – em termos lexicográficos – a variante *chamechuga*, predominante entre as demais lexias.

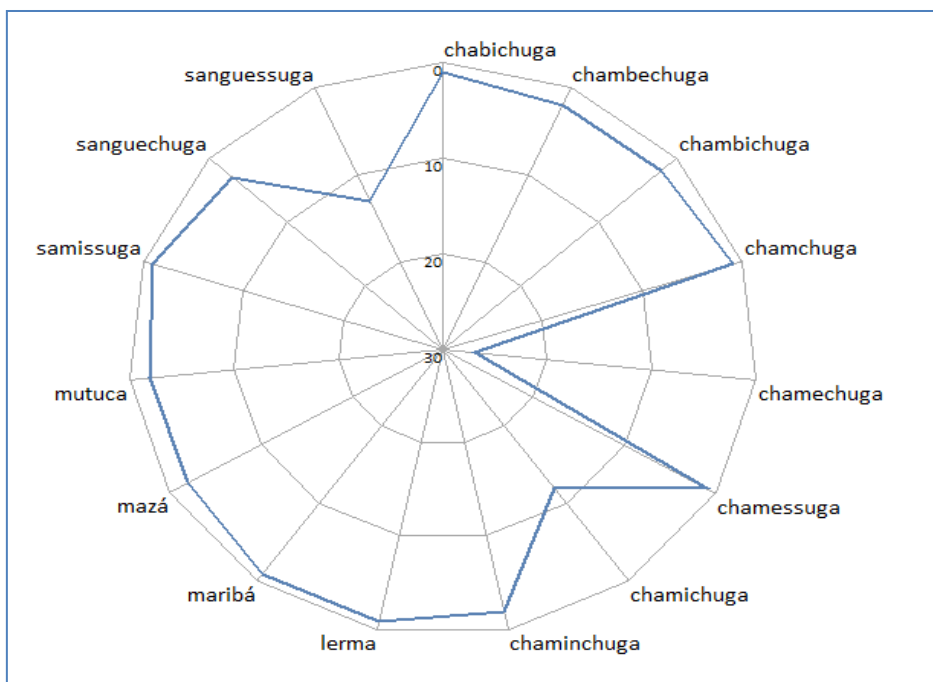


GRÁFICO 21 – Distribuição do comportamento léxico-variacional de 'Sanguessuga'

Fonte: Gráfico elaborado pela autora desta tese (2017)

7.1.2.1 A variação lexical por áreas temáticas: Corpo Humano

Tocando o *Corpo humano*, optou-se pela análise em subgrupos, selecionando uma questão para cada: i) calcanhar, para partes do corpo; ii) cócegas, para sensações; iii) corcunda para características físicas.

Calcanhar é única lexia utilizada para designar. No entanto, se observa a presença do rotacismo, em *carcanhar*. Embora não seja possível presumir se este rotacismo caracterizará alternância léxica, em algum momento da história linguística, o registro lexicográfico dessa variante se faz necessário.

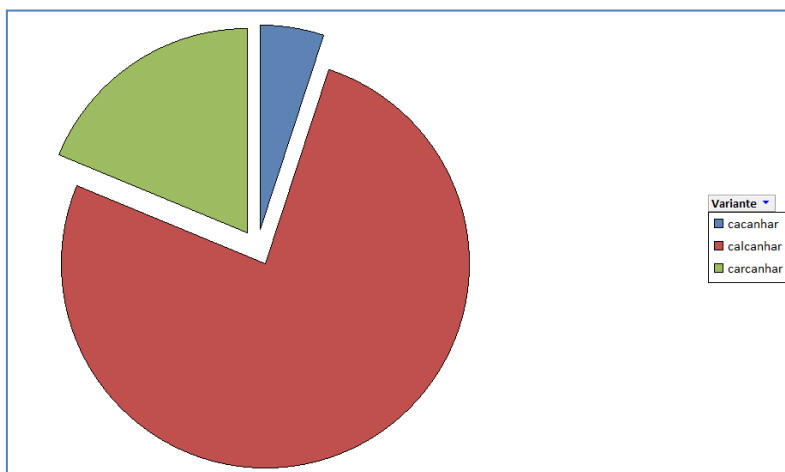


GRÁFICO 22 – Distribuição do comportamento léxico-variacional de 'Calcanhar'

Fonte: Gráfico elaborado pela autora desta tese (2017)

Situação similar ao que fora visto ocorre com as variantes de *cócegas*. Nota-se que a forma canônica “perdeu” seu lugar ao eixo para a variante *cósca* cujas realizações passam de 40, o que pode comprovar sua lexicalização social.

relacionadas a cada item, sobretudo no que tange às alterações mórficas – os lemas múltiplos – à classificação gramatical, mais especificamente no tocante ao gênero e aos processos formativos, os quais requerem a inserção e caracterização clara e específica no verbete.

Obviamente, nem sempre é possível identificar processos formativos, já que muitas lexias são construídas por processos metafóricos e metonímicos, por analogia, dentre outros. No entanto, observando, por exemplo, a relação entre **gimbra** e **gimbrundo**, percebe-se, nitidamente, a influência do processo formativo de derivação por sufixação, fato e fenômeno que merecem registro.

7.1.3 A variação lexical por áreas temáticas: Ciclos da vida

No que tange à área Ciclos da vida, selecionaram-se, para a presente análise, dados obtidos como respostas à questão QSL 131: “[...] o filho que nasceu por último”, com a qual foram registradas as seguintes formas: (.....), das quais se destacam *caçula*, *caçulo* e *mais novo*. Com a investigação foram documentadas 13 lexias, como mostra o gráfico, a seguir:

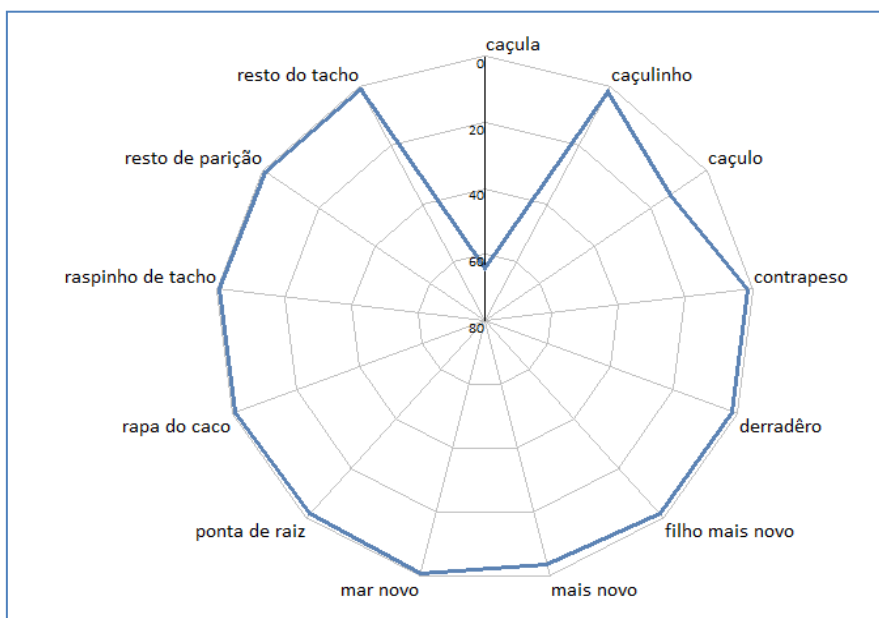


GRÁFICO 24 – Distribuição do comportamento léxico-variacional de ‘Caçula’
Fonte: Gráfico elaborado pela autora desta tese (2017)

Considerando a ordem alfa e o critério de frequência para a escolha do lema principal, elegeu-se a forma **caçula**, cujo verbete identifica, já na entrada,

a possibilidade de lema múltiplo **caçulo** por apresentar particularidades que denunciam alguma modificação de natureza mórfica.

Curioso é ver que o informante considera a vogal temática como morfema flexional de gênero, gerando a partir da forma caçula entendida como feminino, o masculino, fugindo à regra do português.

Na nomenclatura, em sequência, apresenta-se **caçulo** em remissão a **caçula**.

Nota-se, ainda, que lexias compostas ou complexas também podem se apresentar como lema principal. Para tanto, leva-se em consideração critérios como a colocação (frequência / recorrência de uso dos itens), a canonização e a comutação.

Utilizando o programa *WordSmith 4.0*, podem-se ver as colocações. No caso de **mais novo**, percebe-se que, no *corpus* analisado, **mais** precede **novo** por 10 vezes, sendo a segunda realização mais frequente. A primeira é **de** que junto a **novo** constitui uma locução adverbial e é, portanto, mais gramatical do que a expressão **mais novo**. Como se sabe, elementos de gramática são mais usuais na língua.

N	Word	With	elation	Total	tal Left	tal Right	L5	L4	L3	L2	L1	Centre	R1	R2	R3	R4	R5
1	NOVO	NOVO	0,000	62	1	1	0	1	0	0	0	60	0	0	0	1	0
2	DE	NOVO	0,000	39	37	2	0	0	0	0	37	0	0	0	1	1	0
3	MAIS	NOVO	0,000	13	10	3	0	0	0	0	10	0	1	0	0	1	1
4	O	NOVO	0,000	13	10	3	2	1	6	1	0	0	0	2	1	0	0
5	AÍ	NOVO	0,000	12	7	5	0	1	1	5	0	0	4	0	0	1	0
6	QUE	NOVO	0,000	10	6	4	3	2	1	0	0	0	2	1	1	0	0
7	ELE	NOVO	0,000	6	6	0	2	1	2	1	0	0	0	0	0	0	0
8	PRA	NOVO	0,000	5	4	1	0	2	2	0	0	0	1	0	0	0	0
9	É	NOVO	0,000	5	4	1	0	0	0	2	2	0	1	0	0	0	0
10	OU	NOVO	0,000	5	3	2	0	2	0	1	0	0	2	0	0	0	0
11	FALE	NOVO	0,000	5	5	0	0	0	0	5	0	0	0	0	0	0	0
12	BROTA	NOVO	0,000	5	5	0	0	0	0	5	0	0	0	0	0	0	0
13	ANO	NOVO	0,000	5	4	1	0	0	1	0	3	0	1	0	0	0	0

Aqui, veem-se alguns usos de **mais novo**.

N	Concordance	Set	Tag	Word #	t. #	os.	. #	os.	. #	os.	t. #	os.	File	%
1	E quem é que chama, o povo mais novo ou o povo mais velho? INF. – É o			1.240	171	7%	0	1%	0	1%	a	090_04_qls.txt	53%	
2	o povo mais antigo ou o povo mais novo ? INF. – Esse pessoal qui não tem			2.087	412	0%	0	8%	0	8%	a	100_01_qls.txt	50%	
3	– Tem outro jeito. INF. – Filho mais novo . [□□□□□□□□□□□□] (132)			2.939	667	0%	0	4%	0	4%	a	100_04_qls.txt	56%	
4	dá esse nome? INF. – Inquanto ele tá novo , mais depois que já crece é o			1.653	301	0%	0	9%	0	9%	s	102_03_qls.txt	39%	
5	INQ. – O senhor chama assim mais novo ? INF. – Era, até hoje ainda chamo			7.043	998	0%	0	7%	0	7%	a	083_03_qls.txt	76%	
6	(121) INF. – Olha, quando eu era mais novo eu ouvia falar em boi □□□□□□.			5.412	773	7%	0	9%	0	9%	a	083_03_qls.txt	58%	
7	nome, esse criado, apareceu mais novo , era dordoio □□□□□□□□			2.470	336	1%	0	7%	0	7%	a	083_04_qls.txt	66%	
8	é o mais velho e esse o...? INF. – Mais novo □□□□□□□□□□ (132) INQ. -			2.110	428	7%	0	1%	0	1%	a	084_02_qls.txt	61%	
9	peessoa mais velha. Que os mais novo não. INQ. – Já chamam logo de			2.627	543	3%	0	0%	0	0%	é	096_04_qls.txt	79%	
10	□□□□□□ hoje é mais novo . Agora naquele tempo era assim.			149	34	0%	0	4%	0	4%	é	095_03_qls.txt	4%	
11	(131) INF. – Fala filho mais novo , caçulo □□□□□□□□ (132) INF.			1.840	259	0%	0	5%	0	5%	a	083_02_qls.txt	64%	
12	(131) INF. – Filho mais novo □□□□□□□□□□□□□□ INQ.			1.580	260	1%	0	3%	0	3%	s	087_02_qls.txt	63%	

Há casos em que **mais novo** aparece em estruturas sintáticas da língua não representando uma unidade significativa. Em outras, constitui-se em itens lexicais.

Retomando a nomenclatura, na sequência, há outras remissões para **caçula**, constituídas das variantes vocabulares **ponta de raiz**, **rapa do caco**, **resto de parição**, **resto do tacho**.

O estudo das designações para menstruação tem sido bastante recorrente, sobretudo dentre os que tratam acerca de tabuísmo. Notam-se estratégias multivariadas que podem ser consideradas como esquiva. Muitos desses usos se configuram como eufemismos como, por exemplo, ocorre com *amiguinha* e *visitante*. Em contrapartida, outros se revelam estigmatizantes e estigmatizadores.

Outro ponto interessante a ser comentado refere-se à mudança de classe gramatical para a designação, ou seja, de um substantivo menstruação passa-se a um verbo.

Lexicograficamente isso se constitui um problema. No entanto, a Lexicografia Variacional precisa estar atenta à inserção desses usos, se não como entradas de verbete, como achegas enciclopédicas.

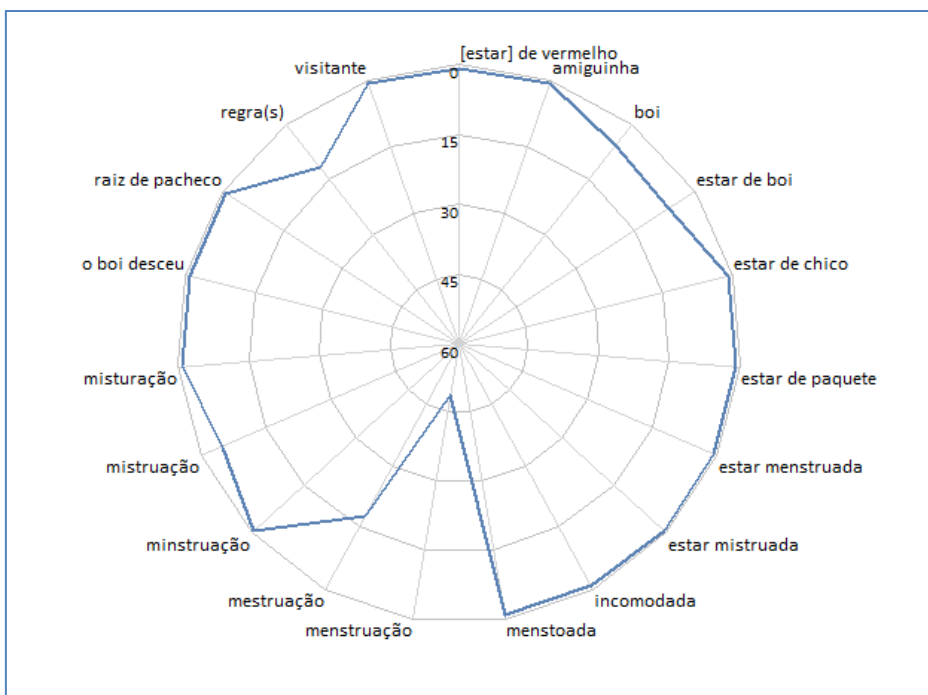


GRÁFICO 25 – Distribuição do comportamento léxico-variacional de 'menstruação'
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Considerando os resultados de ambas as questões, é possível estabelecer relações semânticas entre *estar de boi* para menstruação e *terminar o boi* para menopausa.

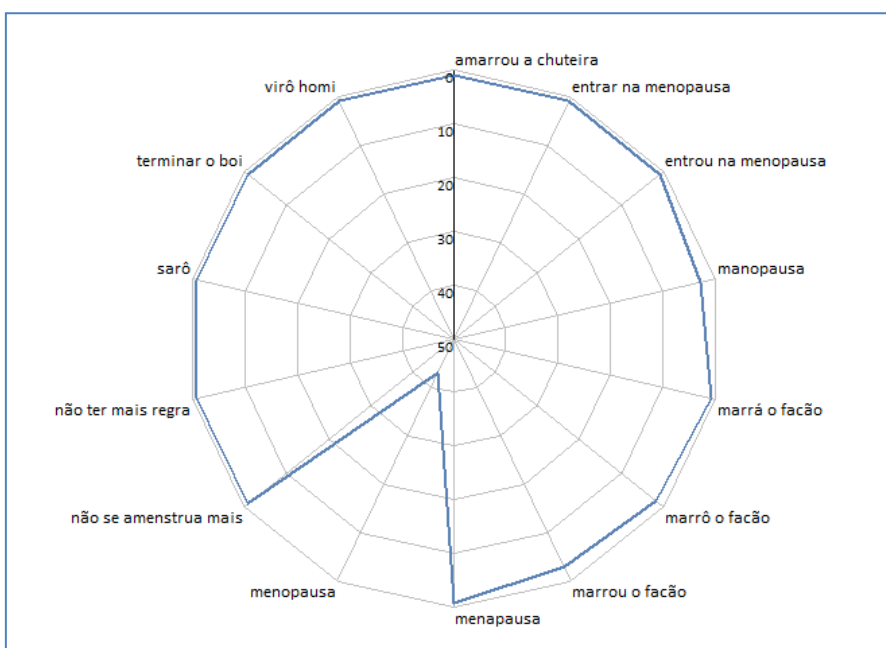


GRÁFICO 26 – Distribuição do comportamento léxico-variacional de 'menopausa'
 Fonte: Gráfico elaborado pela autora desta tese (2017)

7.1.4 A variação lexical por áreas temáticas: Convívio e comportamento social

Para a presente investigação, optou-se por apresentar resultados da análise dos dados obtidos com base nas respostas ao QSL 138 “Como se chama a pessoa que não gosta de gastar seu dinheiro e, às vezes, até passa dificuldades para não gastar?”

Segundo Montesquieu (2000, p. 274), a avareza, bem como a luxúria, “aumenta a sua própria sede com a aquisição de tesouros”. Ora em oposição, ora em paralelismo, ambos os comportamentos se relacionam de maneira a ofuscarem o bem mais precioso: o valor do ser humano, que não tem preço, embora seja possível, lamentavelmente, comprar e vender.

A Revista *Superintessante*, de março de 2012, versão *on line*, publicou que “há duas formas de perceber avareza: aquela necessária à sobrevivência, intrínseca a todos os seres vivos e a que trata do acúmulo desnecessário, humana e mais vinculada ao traço cultural”. A primeira, considerada no texto como “avareza do bem” difere-se, pois, da segunda que se caracteriza como um dos sete pecados capitais no que tange ao traço cultural, já que se trata de um comportamento.

Este comportamento egoísta e de apego aos bens materiais inspirou a criação caricaturada de muitos personagens e também trouxe à tona histórias reais. Caracterizam-se, aqui quatro personagens da dramaturgia nacional e mundial – Harpagon, em “O Avarento”, de Molière; Nonô Correia, da novela “Amor com Amor se paga”; Conde Claus, na novela “Chocolate com Pimenta” e Tio Patinhas, de Barks – dentre os quais se podem destacar traços de avareza. Observe-se o quadro a seguir:

PERSONAGENS	SINOPSE	TRAÇOS DE AVAREZA
“O Avaro”, de Molière. (1622 – 1673),	<i>Harpagon é o personagem que tem ouro enterrado no quintal e vive o terror de ser roubado, além de não deixar sua filha casar com um pobretão. A moral dessa obra é a de que a avareza traz discórdia e desentendimento familiar.</i>	acúmulo e apego ao dinheiro.
Nonô Correia, da novela “Amor com Amor se paga” (1984)	<i>O personagem levava uma vida miserável e mesquinha mesmo contendo uma fortuna escondida. A comicidade estava em trancar sua geladeira para que ninguém roubasse sua comida.</i>	mesquinho(a) é uma variante de avareza; acúmulo de dinheiro e não generosidade
Conde Claus, na novela “Chocolate com Pimenta” (2003-2004)	<i>O personagem era um banqueiro egoísta que costumava aplicar golpes nas pessoas para acumular dinheiro.</i>	acúmulo de dinheiro através da desonestidade.
Tio Patinhas, de Barks (criado em 1947)	<i>O personagem é um bilionário que mora em uma mansão onde toda a sua fortuna está guardada num imenso cofre mas não gasta uma única moeda e deixa seus sobrinhos em má situação financeira.</i>	acúmulo de dinheiro e não generosidade.

Como se pode perceber *acúmulo de dinheiro* é o traço comum a todos os quatro personagens aqui analisados. Aliados a este traço, são citados a não generosidade e até mesmo a desonestidade como meio para obter bens materiais, aos quais se é demasiadamente obcecado.

Ora, sendo parte do convívio sociocultural dos humanos, assim como são muitos os tipos avarentos, muitas são as designações usadas. Afinal, tudo o que faz parte da vida dos seres humanos tem um nome, em alguns casos, diversos nomes, de modo que conforme “a língua, conforme a época ou grupos sociais, conforme o contexto presente tal ou qual objetivo específico, vê-se dominar ora uma forma ora outra, ora uma variante ora outra.” (BAKHTIN, 2006, p. 150)

No que tange à origem e/ou etimologia, Cunha (1982) apenas apresenta informações acerca de *avarento* – “do lat. *avaritia*”. Tornou-se, pois, necessário investigar processo de formação. Embora não seja objetivo deste trabalho observar as motivações extralinguísticas, dentre várias, uma especulação a respeito da origem da variante *pão duro*, expressa em um texto extraído do *blog* de Elaine Paiva (2006), se revelou bastante curiosa:

*Nos primórdios do século passado, vagava pelas ruas cariocas um mendigo muito popular (naquela época os mendigos eram escassos e populares). **A criatura esmolava comida, sempre implorando um pedaço de pão duro.** A reiteração da súplica imprimiu-lhe o apelido: **Pão Duro.** Falece o mendigo e descobre-se que seu refúgio **abrigava uma pequena fortuna em dinheiro, contas bancárias e títulos.** (...)E aí a expressão PÃO-DURO passou a designar o avarento [...]*

A hipótese além de curiosa serve de reflexão acerca dos processos de formação por extensão de sentido e por metáfora para a lexicalização de compostos ou composição de itens lexicais.

7.1.5 A variação lexical por áreas temáticas: Jogos e brincadeiras infantis

Essa área temática foi tese de Ribeiro (2012), intitulada “Brinquedos e brincadeiras infantis na área do *falar baiano*”, obra em perspectiva dialetológica e sociolinguística, cujos dados correspondem à pesquisa realizada em localidades de 10 estados brasileiros (Tocantins, Maranhão, Piauí, Pernambuco, Alagoas, Sergipe, Bahia, Mato Grosso, Minas Gerais e Espírito Santo), “perseguido” a proposta de Nascentes (1953), quanto ao que denomina “falar baiano”.

Dessa maneira, considerando a amplitude e grandiosidade do trabalho supracitado e que envolve todas as localidades baianas com quais se trabalha nesta tese, torna-se impraticável (re)discutir os dados apresentados. No entanto, é relevante admitir que a perspectiva lexicográfica não invalide a presente análise. Inversamente, propõe-se agregar contributos para as sociedades, não se restringindo ao âmbito acadêmico.

Para a presente análise, selecionou-se a questão 157 do QSL: “[...] o brinquedo feito de forquilha e duas tiras de borracha (mímica), que os meninos usam para matar passarinho”.

Uma discussão relevante e que se pode fazer aqui, ainda que sucintamente se refere à definição.

Ribeiro (2012, p. 211) considera que o *brinquedo não tem fronteiras, nem territoriais e nem de função ou utilidade: brinquedo ou arma? Como arma, ele invade os espaços e, muitas vezes, rouba a infância e a inocência das*

crianças. Assim, em perspectiva lexicográfica, mais especificamente a lexicográfica variacional, se em algumas comunidades o brinquedo não é utilizado como brinquedo, mas como arma, ambas as informações devem estar presentes na definição do verbete, admitindo os seus usos possíveis, tornando, pois, imprescindível agregar elementos socioculturais à codificação semântica da informação.

7.1.6 A variação lexical por áreas temáticas: Habitação

Para essa área, selecionou-se a questão 169 do QSL: “Quando uma janela tem duas partes, como se chama a parte de fora que é formada de tirinhas horizontais que permitem a ventilação e a claridade?”. Foram documentadas 10 lexias, como se observa, no gráfico, a seguir:

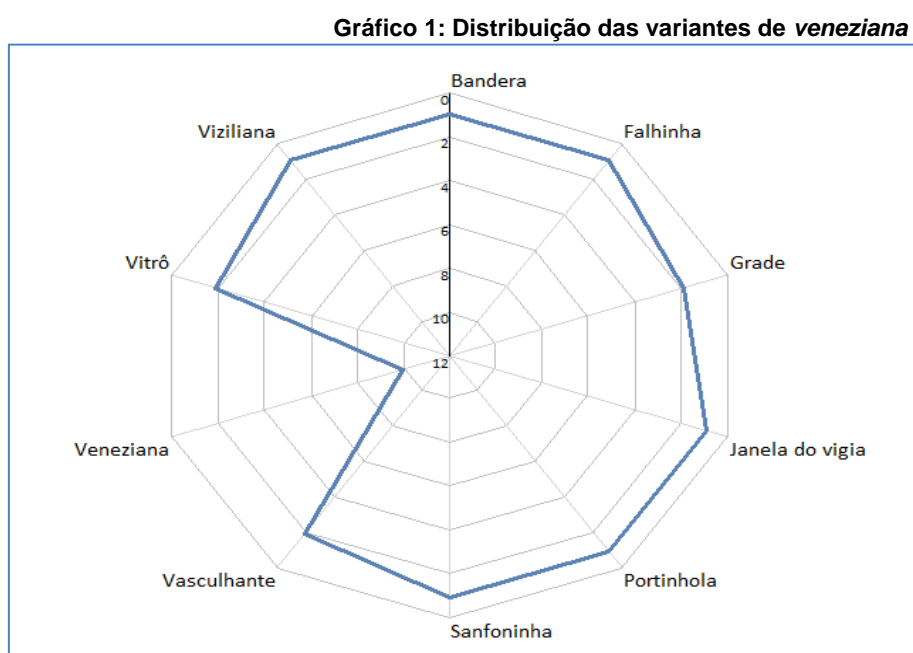


GRÁFICO 27 – Distribuição do comportamento léxico-variacional de ‘veneziana’

Fonte: Gráfico elaborado pela autora desta tese (2017)

Não obstante se perceba a predominância da lexia *veneziana*, diante das demais lexias, essa questão apresentou alto índice de ausência de resposta, ou seja, os informantes não “souberam” ou não costumam denominar a parte de fora da janela (que possui duas partes) formada de tirinhas horizontais que permitem a ventilação e a claridade. Observe-se o gráfico, a seguir:

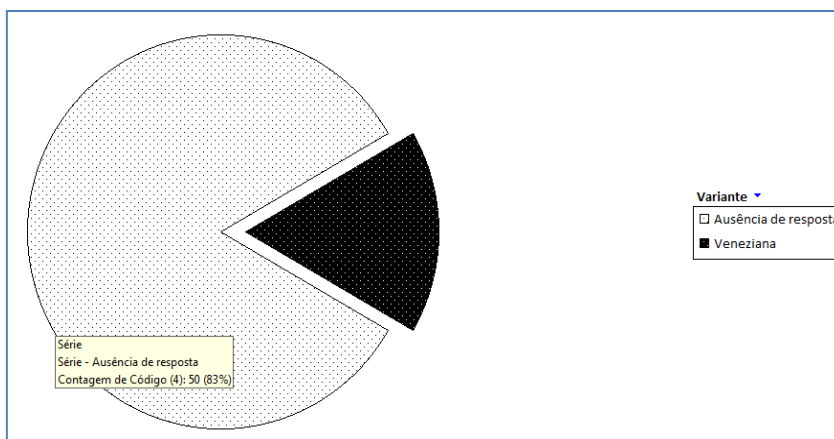


GRÁFICO 28 – Ausência de resposta para o referente ‘veneziana’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Observou-se, ainda, o comportamento sociolinguístico de *veneziana* e percebeu-se que essa variante se caracteriza diageneracionalmente, já que seu uso sendo muito pequeno, apenas 7 de 92 realizações possíveis, equivalente a 7%, ocorre apenas entre falantes da faixa 2 – os mais velhos –, com exceção a Euclides da Cunha, em que a lexia é documentada na fala de um homem jovem. Esse fato pode revelar mudança em curso.

7.1.7 A variação lexical por áreas temáticas: Alimentação e cozinha

Entre café da manhã, pães, geleia, guloseimas, carne moída, curau, canjica e mungunzá, itens que também compõem essa área temática optou-se pela *cachaça*, que, nos dados, se revela como a forma mais frequente em resposta à pergunta “[como se chama a] bebida alcoólica feita de cana-de-açúcar?” – formulação da questão 182 do QSL., como mostra o gráfico:

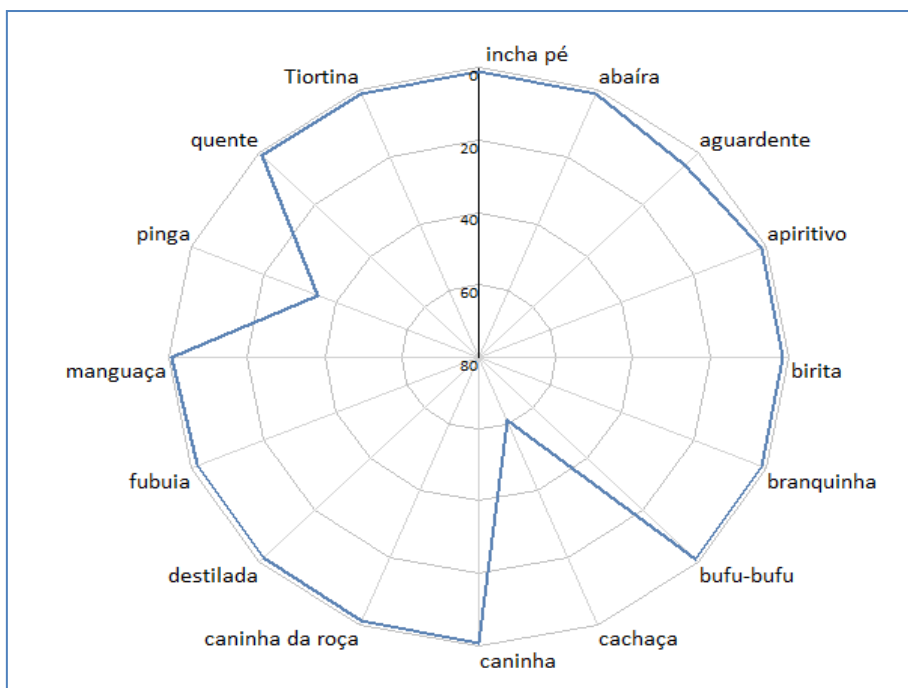


GRÁFICO 29 – Distribuição do comportamento léxico-variacional de ‘cachaça’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

7.1.8 A variação lexical por áreas temáticas: Vestuário e acessórios

Selecionou-se, para essa área, a questão 191 “[...] aquilo que as mulheres passam no rosto para ficarem mais rosadas?”, a partir da qual foram documentadas 17 lexias, cuja distribuição por ocorrência, pode ser percebida no gráfico, a seguir:

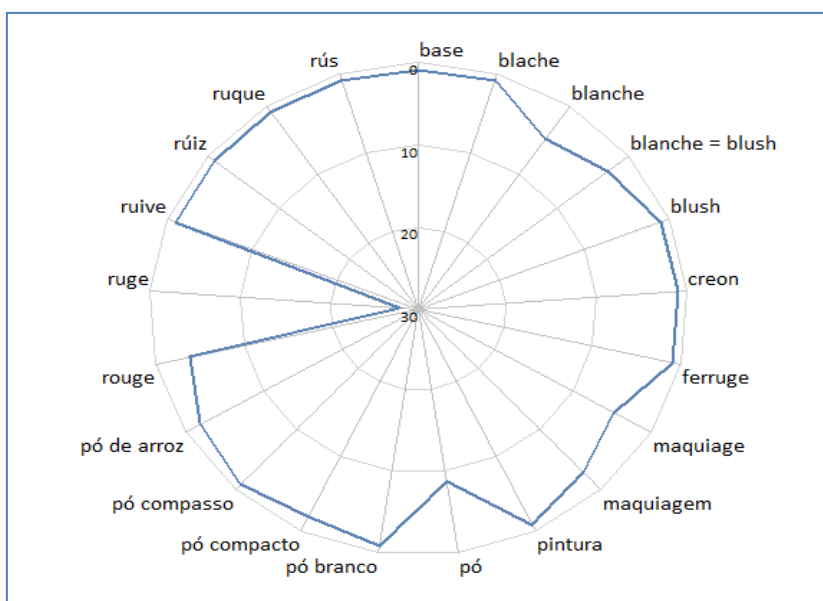


GRÁFICO 30 – Distribuição do comportamento léxico-variacional de ‘rouge’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Como se percebe, **ruge** é a lexia majoritária, possuindo, inclusive, por analogia, cinco formas que se configuram como lemas secundários, em que há alteração fônica, **rús**, **rúiz**, **ruive**, **rouge** e **ferruge**.

A observação do comportamento diagenérico e diageracional das variantes tornou-se relevante para a análise haja vista se tratar de um elemento usado por mulheres, conforme a formulação da questão, com a qual os itens lexicais foram obtidos. Buscou-se observar, pois, se e como os homens denominam os elementos que não pertencem à “sua esfera”, de modo a perceber alterações quanto ao gênero, não obstante a oscilação diageracional.

Para tanto, foram feitas duas análises: i) somente com as formas que se configuram lemas de **ruge** e ii) com as variantes vocabulares de **ruge**, excetuando-se essa lexia.

Da análise dos lemas secundários de **ruge**, notou-se que:

- a) **Ruge** é a forma categórica, entre a faixa 1, sendo mais proeminente entre as mulheres, com 4 das 5 realizações.
- b) Entre falantes da faixa 2, ocorrem mais variação, sendo entretanto, a lexia **ruge** a mais predominante.
- c) Coocorrem com **ruge**, entre os homens da faixa 2, as formas **ruive**, **rouge** e **ferruge** e entre as mulheres, as formas **rús**, **rúiz** e **rouge**.

As observações se evidenciam no gráfico, a seguir:

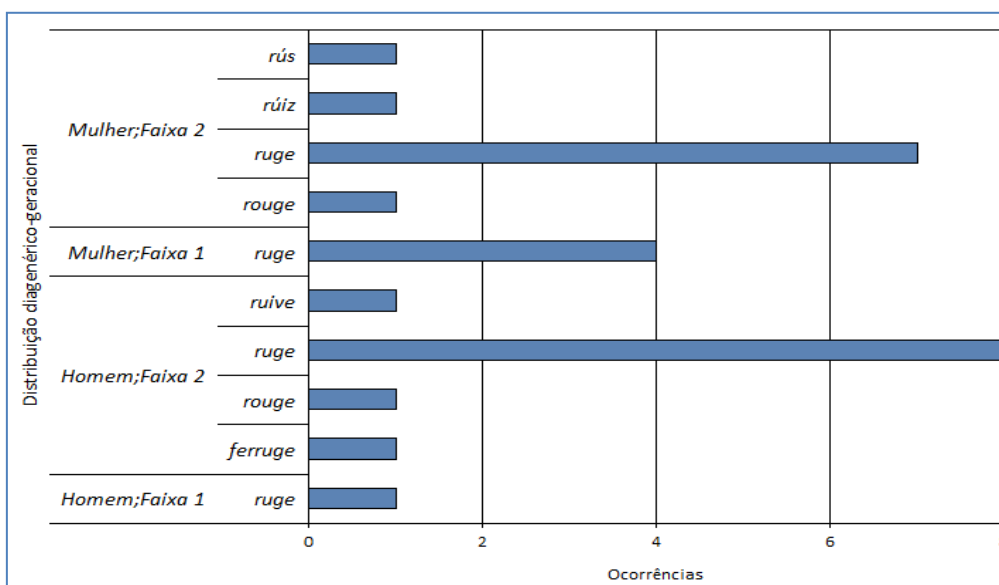


GRÁFICO 31 – Estratificação social das variantes concernentes a ‘rouge’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Da análise das variantes vocabulares de **ruge**, considerando os seus respectivos lemas secundários, notou-se que:

a) Entre os homens da faixa 1, ocorrem a forma mais específica **blanche (= blush)**, bem como as formas mais abrangentes **pó, pó branco, maquiagem, maquiage**, sendo a última a majoritária dentre as demais.

b) Entre os homens da faixa 2, somente ocorrem, de forma equiparada variantes mais abrangentes, **pintura e pó**.

c) As mulheres da faixa 1, usam **blanche (= blush)** e **pó compacto**, equitativamente, mas **pó** se destaca mais usada, embora os índices sejam próximos.

d) As mulheres faixa 2, usam 8 lexias, sendo, dessas, 4 variantes vocabulares. Há realização de: i) **blache (= blush)** e seu lema secundário **blanche (= blush)**; ii) **compasso**, um lema múltiplo – em que há alteração mórfica – do lema secundário **pó compasso** da variante vocabular **pó compacto**, que é o lema principal; iii) além das variantes **maquiagem, pó** e **pó de arroz**.

Paim (2011), na comunicação intitulada *A variação lexical do português falado no Brasil: Reflexões sobre o campo semântico vestuário e acessórios nos dados do Projeto ALiB*, aborda todas as questões constantes da área temática ora em questão, em capitais brasileiras. Em uma das observações gerais, a autora atesta que: “a temática da comparação passado X presente está presente na linguagem dos informantes de faixa etária mais avançada, evidenciado-se na seleção lexical desses informantes como demonstram as estruturas: *porta-seio, ceroula, corpinho, rouge e travessa*.”

No que tange especificamente a **rouge**, confrontando os dados obtidos nesta pesquisa com os dados de Paim (2011), evidenciam-se diferenças que podem caracterizar divergências entre os usos das capitais e das localidades interioranas, já que além de serem obtidas 17 formas distintas, a variante **ruge** – com respectivos lemas secundários – se configura como uma variante altamente produtiva independente do gênero e da faixa etária.

Considerando o fato de que *rouge* e *blush* entraram na língua portuguesa como empréstimos linguísticos, respectivamente do francês e do inglês, outra possível discussão a ser apresentada oportunamente: a influência do francês, do inglês e de línguas africanas e indígenas para o léxico do português brasileiro contemporâneo.

7.1.9 A variação lexical por áreas temáticas: Vida urbana

O conceito de urbano e rural vem sendo discutido ao longo dos anos, sobretudo pela inadequação, no que tange à periferização e favelização de áreas consideradas mais dinâmicas.

No âmbito do senso comum, o “meio urbano” distingue-se do “rural” pelas características que se referem ao modo / estilo de vida, aos hábitos tipicamente “de cidade”, além de características específicas como a densidade populacional e a infraestrutura – escolas, hospitais, vias públicas, transportes, etc.

Partindo desses pressupostos, a questão 194 do QSL: *Na cidade, o que costuma ter em cruzamentos movimentados, com luz vermelha, verde e amarela?*, se revelou interessante para esta análise, em que se buscou observar, previamente, os dados, com vistas ao estabelecimento de uma metodologia adequada para a construção do vocabulário proposto.

Foram documentadas 10 variantes, dentre as quais apenas quatro se configuram como variantes vocabulares, considerando a existência dos lemas secundários e múltiplos para cada uma, exceto para **farol**.

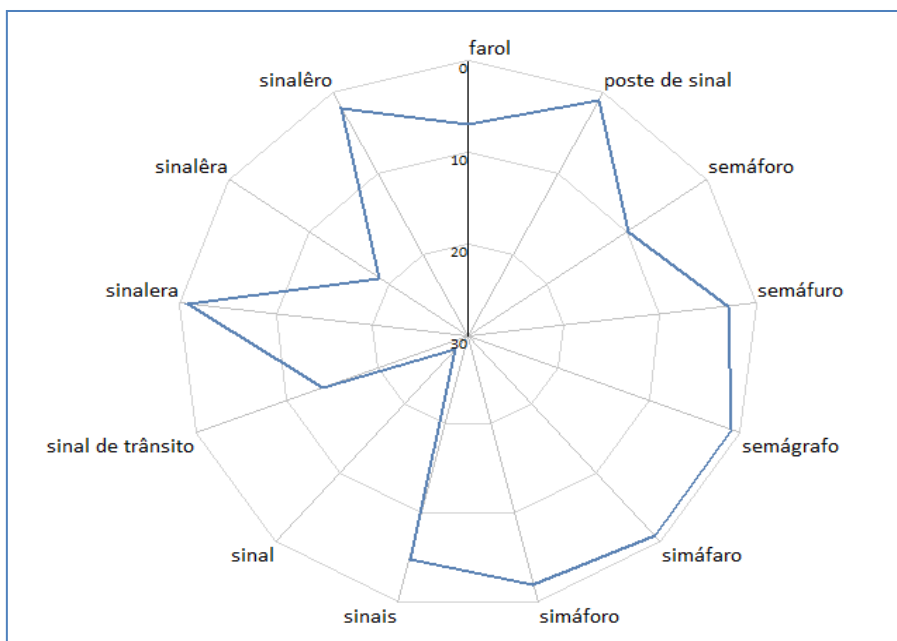


GRÁFICO 32 – Distribuição do comportamento léxico-variacional de ‘rouge’
 Fonte: Gráfico elaborado pela autora desta tese (2017)

Em uma das respostas, ao dizer “o **sinal**, a **sinalêra** da rua, né? Chegô pra nós há pouco tempo”, o informante revela um dado sociocultural relevante à pesquisa, haja vista que a ausência ou a “novidade” do referente pode influenciar os (des)usos ou (não)usos de lexias específicas ou até mesmo denunciar pouco reconhecimento da forma canônica, como se percebe no excerto, a seguir:

- INF. – Tô querendo lembrar o nome *sema...semafo...*
 INQ. – E como é que chama...
 INF. – Eu sei. Eu sei mas eu tô querendo lembrar é o nome.
 INQ. – Ah!
 INF. – *Semágrafo* como é? eu sei. Eu sei o que é mas não tô conseguino falá é o nome.
 INQ. – O nome né.
 INF. – É. *Semáforo* né. *Semáfro*.
 (100/4 – Itapetinga / Mulher, faixa 2, nível fundamental)

7. 2 A VARIAÇÃO LEXICAL NAS MESORREGIÕES DA BAHIA

Apresentam-se, aqui, alguns resultados da análise diatópica dos dados, por mesorregiões.

A presente análise contempla microáreas temáticas de três áreas temáticas. No tocante aos *Acidentes geográficos*, observaram-se a distribuição de lexias relacionadas a ‘rio’. De *Atividades agropastoris*, extraíram-se, para a presente descrição, dados que se referem a ‘alimentos’. Observaram-se, ainda, da área *Ciclos da vida*, especificamente, os “ciclos do período fértil” da mulher.

Para tanto, optou-se pela elaboração de nuvens de palavras¹, tendo por base de ilustração a legenda geolinguística – criada com base nas mesorregiões baianas, como já fora dito – utilizada nos verbetes que constituem o Vocabulário Dialetal Baiano, com a qual se pode verificar as localidades onde a lexia foram documentadas.

Para melhor compreensão, sugere-se o acompanhamento da legenda na folha transparente avulsa, constante no envelope anexo a este Volume 1 da tese.

¹ *Nuvem de palavras* é um gráfico digital que mostra o grau de frequência das palavras em um texto, de modo que quanto mais a palavra é utilizada, mais chamativa é a representação dessa palavra no gráfico. Atualmente há vários programas informáticos que geram nuvens. As que se apresentam, porém, foram criadas, manualmente, pela autora desta tese, haja vista utilizar a legenda geolinguística (também criada pela autora juntamente com o orientador da tese) como base. A frequência das lexias foi computada a partir da base de dados com a qual se constituíram as análises.

NUVEM 1 - Distribuição das lexias relacionadas a 'rio' por mesorregiões

cóligo balancete de água **córrego** pequeno rio

onda(s) riacho forte riachuelo maré

mareta riote marombas **riacho**

mareta(s)

maré

onda

onda(s) amaranta **correnteza**

córgo corredêra

maleta **maré** ondinha

lomba onda

onde d'água **mareta** valeta

bareta **mareta** mareta de onda

maretinha **onda(s)** marola

redimuinho vareta

onda **correnteza**

maleta

correnteza **onda(s)**

correnteza do rio ondinha

remanso remanso de água

NUVEM 2 - Distribuição das lexias relacionadas a 'alimentos' por mesorregiões

aipim gemada mixirica *cacho da banana* aimpim coração tangerina
 imbigo ingenaro *camumila* tangirina *cacho* mexerica umbigueira
amendoim gêmeas *penca* *camomila* umbigo impim macaxera
 mandioca¹ mandioca doce mandioca mansa serrana mandioca
 mandioca brava mandioca que faz farinha

banana pegada *cacho* **amendoim** mixirica *camomila* pocã tangerina *cacho da banana* filipe
 tangirina **amenduim** *camumila* **mendoim** *penca* gêmea(s) bambó
 coração flor pendão aimpim aipim impim macaxera mandioca mandioca¹ doce
 mandioca mansa serrana [mandioca] azeda babosinha braba mandioca² mandioca
 braba mandioca brava margosa

laranja cravo laranja tanja laranja-flor mexerica mixirica pocã tangerina tangirina **amendoim amenduim**
amiduim mendoim menduim mindoim *camomila camomilha camumila* *cacho penca* banana conde
 cunha felipe filipim filipinho filipo gêmeas bagunço bobó bombé buzo coração flor ingaço aimpim aipim macaxera
 mandioca² mandioca mansa araticum mandioca mandioca brava mandioca-brabo mucunhã

laranja clavo **amendoim** mexerica mixirica tangerina
 tangirina **amenduim** camomila cacho cacho da banana
 penca filipe filipim filipinho gêmeas mabá mabaça
 coração flor umbigo aimpim aipim macaxera mandioca¹
mandioca mansa mandioca²

laranja cravo mexerica mexirica mixirica pocã tangerina tangirina amendoim
 amendoins amenduim amiduim miduim camomila camumila cacho cacho de
 banana penca pendão banana pegada buneca felipe filipinho gêmea inconha coié
 fígado flor da bananeira imbigo pendão umbigo aimpim aipim macaxera mandioca¹
 mandioca mansa mandioca² mandioca braba mandioca do mangue mandioca
 tapuru

imbigo laranja cravo mixirica mixiriquêra tangerina tangilina tangirina
 amendoim amendoins amendoins camomila camumila cacho
 penca(s) banana filipinho filipinho gêmeas mabaça(s) mabaço aipim
 aimpim macaxera mandioca²

buza buzina fror naipe pendão umbigo laranja cravo
 mexerica mexirica mixirica tangerina amendoim
 amendoins amenduim mendoim camomila camomila
 cacho penca banana pegada filipa garrada gêmeas
 mabaça nadacha aimpim aipim macaxera mandioca¹
 mandioca²

NUVEM 3 - Distribuição das lexias relacionadas a 'ciclos da vida da mulher', por mesorregiões baianas

de boi *menstruação* **regra**

mestruação menopausa

amarr[ar] o facão

de boi *menstruação* **regra**

mestruação estar de chico sarar

menaspausa **menopausa**

amiguinha boi *estar de boi* **estar de chico** o boi desceu

menstruação raiz de **pacheco** *regras* *estar de vermelho*

não se amenstrua mais *marrou o facão* *menaspausa*

menopausa

boi de boi *menstruação*

mestruação *mestruada*

minstruação **virou homi**

manopausa *marrou o facão*

menopausa

não ter mais regra de boi

visitante *menstruação* **regra**

mestruação *marrou o facão*

menaspausa **menopausa**

boi estar nos dias *menstruação*

regra **mestruação** *incomodada*

climatério **entar na menopausa** *marrou*

o facão *menaspausa* **menopausa**

de boi *estar de pacote*

menstruação **regra**

mestruação

As “nuvens” apresentadas contribuem para o exercício de visualização da legenda geolinguística constante dos verbetes que compõem o *Vocabulário Dialectal Baiano*, com a qual é possível representar, diagramaticamente, a delimitação de possíveis áreas dialetais ou de áreas do léxico sociodialetal.

As amostras de verbetes, a seguir, parecem ser elucidativos.

fanhoso ↔ fonhem

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

cocar ‘Ave, similar a uma galinha, de penas pretas e pintas brancas’. // QSL 067 - *Como se chama a ave de criação parecida com a galinha de penas pretas com pintinhas brancas?* ≅ **caqui** ≅ **cocar** ≅ **cocó** ≅ **conquém** ≅ **coqué** ≅ **galinha d’água** ≅ **galinha d’angola** ≅ **galinha de(a) angola** ≅ **galinha pedrês** ≅ **guiné** ≅ **perdiz** ≅ **pintado** ≅ **saqué** ≅ **sariema** ≅ **tô fraco**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

◇ **APFB-ALIB** n Carinhanha - 097; Jacobina - 086; Santa Cruz Cabralia - 101; Santana - 092

7.3 DO ATLAS PRÉVIO DOS FALARES BAIANOS AO VOCABULÁRIO DIALETAL BAIANO¹: A VARIAÇÃO ESPAÇO-TEMPORAL DAS LEXIAS DOCUMENTADAS

A vida é o dever que nós trouxemos para fazer em casa.
 Quando se vê, já são seis horas!
 Quando de vê, já é sexta-feira!
 Quando se vê, já é natal...
 Quando se vê, já terminou o ano...
 Quando se vê perdemos o amor da nossa vida.
 Quando se vê passaram 50 anos!
 [...]
 (Mário Quintana, O tempo)

Assim como a vida, os estudos linguísticos, mais especificamente os estudos realizados com base na teoria da mudança linguística, pressupõem ser o tempo um fator crucial. Tais pesquisas têm sido desenvolvidas, confrontando usos, não só em tempo real, mas ainda em tempo aparente, cujos conceitos partem do pressuposto central de que as diferenças no comportamento linguístico de gerações diferentes de falantes num determinado momento reflitam diferentes estágios do desenvolvimento histórico da língua. Segundo Chambers e Trudgill (1980), a sua validade

depende crucialmente da hipótese de que a fala das pessoas de 40 anos hoje reflete diretamente a fala das pessoas de 20 anos há 20 anos atrás e pode, portanto, ser comparada com a fala das pessoas de 20 anos de hoje, para uma pesquisa da difusão da mudança linguística. As discrepâncias entre a fala das pessoas de 40 e 20 anos são atribuídas ao progresso da inovação linguística nos vinte anos que separam os dois grupos. (CHAMBERS e TRUDGILL, 1980, p. 165)

Desse modo, ao serem comparadas duas faixas etárias, podem-se pressupor estágios de mudança linguística: i) variação estável – quando duas formas coexistem equiparadamente no que tange às faixas etárias, de modo que não há diferença entre os usos de jovens e idosos – e ii) mudança em curso – quando a fala dos jovens difere-se da dos idosos pela inserção de novos usos, considerados

¹ Parte do texto que integra esta seção é resultante da comunicação intitulada “De ponto a ponto: variação lexical na bahia em cinquenta anos”, apresentada, em outubro de 2015 no V SIMELP, na Itália. O trabalho fora elaborado em coautoria com o então tutorado e bolsista de IC, Ivan Pedro Nascimento. Naquela ocasião, foram analisados sete itens lexicais constantes do APFB e do ALiB: i) onda de rio (C-16/APFB e QSL-06/ALiB); ii) neblina (C-12/APFB e QSL-21/ALiB); iii) inflorescência da bananeira (C-32/APFB e QSL-44/ALiB); iv) sanguessuga (C-128/APFB e QSL-84/ALiB); v) aborto (C-89/APFB e QSL-126/ALiB); vi) avarento (C-103 a 105/APFB e QSL-138/ALiB) e vii) sutiã (C-68/APFB e QSL-188/ALiB).

inovadores. No Brasil, além de estudos da mudança linguística em tempo real de longa duração e em tempo aparente, desenvolvem-se, a partir da implantação do Projeto NURC, estudos da mudança linguística em tempo real de curta duração.

Considerando a distância temporal que se interpõe entre a publicação do primeiro atlas linguístico brasileiro, o *Atlas Prévio dos Falares Baianos (APFB)* – em 1963 – e do *Atlas Linguístico do Brasil (ALiB)* – em 2014 –, considerou-se relevante medir a variação linguística que se pode ter estabelecido nas comunidades investigadas em ambas as pesquisas dialetológicas, confrontando, expressivamente o léxico, com vistas a identificar, categorizar e inserir, *a posteriori*, no DDB, os possíveis processos formativos das lexias documentadas – étimo, origem, possíveis neologismos e desusos – que possam configurar estágios de mudança.

Para a análise dos dados, assumiram-se teoricamente duas perspectivas: i) observação dos estágios de mudança linguística, a partir do cotejo entre as lexias documentadas em ambos os atlas; ii) identificação dos processos de formação de palavras das lexias documentadas, a fim de refletir acerca do registro etimológico em obras lexicográficas.

Os dados do APFB foram documentados com base na utilização de questionário constituído de 182 questões que contemplaram Terra, Vegetais, Homem e Animais Para a construção do ALiB, com base na aplicação do *Questionário 2001*, no que tange ao estudo do léxico variacional, documentaram-se respostas às 202 questões que integram o Questionário Semântico-Lexical (QSL), constituído de 14 áreas temáticas – Acidentes geográficos; Fenômenos atmosféricos; Astros e tempo; Atividades agropastoris; Fauna; Corpo humano; Ciclos da vida; Convívio e comportamento social; Religião e crenças; Jogos e diversões infantis; Habitação; Alimentação e cozinha; Vestuário e acessórios e Vida urbana.

Os dados coletados contribuíram para a inserção de uma marca de uso, como achega de verbete de todas as variantes documentadas no *corpus* do ALiB e que constam também do APFB.

Observem-se, a seguir, os referentes de itens lexicais constantes do APFB e do ALiB.

Area temática (ALIB)	Conceito
Acidentes geográficos	1. onda de rio
Fenômenos atmosféricos	2. neblina
	3. sarolha
	4. arco-íris
	5. trovão / trovuada
Astros e tempo	6. trasanteontem
	7. estrela cadente
Atividades agropastoris	8. trabalhador em roça alheia
	9. sabugo
	10. inflorescência da bananeira
	11. borrego
Fauna	12. sanguessuga
	13. galinha sem rabo
	14. galinha d'angola
	15. cabra sem chifre
Corpo humano	16. nuca
	17. axila
	18. rótula
	19. tornozelo
	20. pessoa que tem uma perna maior ou menos do que a outra
	21. pessoa que tem as pernas arqueadas
	22. dente canino
	23. dente siso
	24. clavícula
	25. útero
	26. calcanhar
	27. cego de um olho
	28. cisco
	29. tersol
	30. conjuntivite
	31. catarata
Ciclos da vida	32. aborto
	33. menstruação
	34. gêmeos
	35. madrasta
Convívio e comportamento social	36. avarento
	37. cigarro de palha
	38. prostituta
Religião e crenças	39. medalha
	40. feiticeiro
Jogos e brincadeiras infantis	41. cambalhota
Habitação	42. borralho
	43. tramela
Vestuário e acessórios	44. sutiã

QUADRO 13 – Questões constantes do APFB e do ALiB.
 Fonte: Elaborado pela autora desta tese (2017)

Da área *Acidentes geográficos*, como designações ao 'movimento da água do rio', foram documentadas 11 variantes vocabulares no *corpus* do ALiB. O gráfico, a seguir, mostra o total de ocorrências de cada uma das lexias documentadas em ambos os atlas.

Com base na observação dos resultados, é possível perceber que, ao longo dos 50 anos, *mareta* – variante catégorica no APFB –, passa a coexistir, no espaço baiano, com outras formas consideradas inovadoras documentadas no ALiB, dentre as quais se destacam: i) *onda*, pela frequência de uso, com 9 ocorrências e; ii) as formas constituídas por processo de composição como *onda de água*, *remanso de água* e *remanso de rio* e de derivação como *ondinha*. No que tange ao uso das lexias constantes do APFB, observe-se:

VARIANTE	Barra		Caetité		Carinhanha		Itaberaba		Jacobina		Jeremoabo		Cabrália		Santana		Conquista	
	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB
mareta				x								x						

Nota-se que o registro de *mareta*, no ALiB:

- ocorre em sete das nove localidades analisadas;
- coexiste com formas inovadoras em Cabrália, Carinhanha, Conquista, Itaberaba, Jacobina e Santana;
- não é usada em Caetité e Jeremoabo, localidades em que seu uso foi substituído por *onda* e formas que se constituem com a base lexical *remanso* – *remanso da água* e *remanso de rio*.

No tocante às *Atividades agropastoris*, analisam-se questões 'trabalhador em roça alheia; sabugo; inflorescência da banana e borrego.

		Barra		Caetité		Carinhanha		Itaberaba		Jacobina		Jeremoabo		Cabrália		Santana		Conquista	
CONCEITO	VARIANTE	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB
trabalhador em roça alheia	pataqueiro	x	x	x	x	x	x	x	x	x	x		x		x	x	x	x	x
	macaqueiro		x			x	x		x		x	x	x		x	x	x	x	x
	diarista	x		x	x	x		x		x		x				x		x	
sabugo	sabugo		x					x				x	x	x					
	capuco		x	x	x	x	x	x	x						x	x	x	x	x
	papuco	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
bananeira	umbigo	x	x		x	x	x	x	x			x	x	x		x		x	x
	búzio	x	x	x	x	x	x		x	x	x		x			x	x	x	x
	bagunço	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x
	coração			x		x		x		x	x	x	x	x	x	x		x	
	bobó	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x
borrego	borrego	x	x	x	x	x	x		x		x				x		x	x	x
	cabrito		x		x							x	x		x				

Sobre 'Trabalhador em roça alheia', podem-se tecer as seguintes considerações:

- a) Há mais de 50 anos, no APFB, a variante *diarista* apenas ocorria em Cabrália. Atualmente, essa que parecia ser a forma mais inovadora, se conserva e se dispersa às demais localidades.
- a) Pataqueiro, que apenas se realizava em Jeremoabo e Cabrália, deixou de ser empregado.
- b) Mataqueiro se conserva em Barra, mas não existem mais em Caetité, Jacobina, Itaberaba e Cabrália.

Destaca-se, aqui, a fala de um informante de Jacobina, em que "Mas antigamente, naquela época [...] às vez tinha um, era dia de macaco, dia de macaquêro." Esse dado se mostra altamente relevante e, por isso, foi incluso ao verbete.

Acerca de 'sabugo', pode-se atestar que:

- a) *Sabugo* prevaleceu aos anos e continua a ser falado em Caetité, Carinhanha, Jacobina, Santana e Conquista. A *lexia* se expandiu também para Itaberaba e Cabrália.
- b) Há conservação de *capuco* Jacobina e em Jeremoabo. Em contrapartida, parece ter sido descartado em Cabrália e em Barra.
- c) O *papuco* de Itaberaba desapareceu.

De 'Borrego e cabrito', diz-se que

- a) *Borrego* e *cabrito* coexistem Itaberaba, Jacobina, Santana e Cabrália.
- b) Atualmente, conforme os dados do ALiB, apenas um borrego foi encontrado em Jeremoabo.
- c) Cabrito ainda é encontrado em Carinhanha, Itaberaba, Jacobina, Santana e em Vitória da Conquista.

Tipicamente tropical, a bananeira é uma planta rústica e produtiva, cultivada em quase o Brasil. No panorama nacional destacam-se Bahia e São Paulo como os maiores produtores, com cerca de 23% da produção brasileira. A população brasileira consome mais de 40 kg por habitante de banana.

Dentre as partes dessa planta, destaca-se, aqui como mostra a imagem ao lado –, a inflorescência – espécie de espiga simples, terminal, que emerge do centro das bainhas foliares – sobretudo em função dos diversos nomes a ela atribuída – flor, coração, umbigo, dentre outras.

Nesta pesquisa, no tocante às designações atribuídas à inflorescência da bananeira, registraram-se um total de nove variantes distintas – *coração*, *búzio* ou *búzia*, *umbigo*, *buzina*, *ingaço*, *naipe* e *pendão*, *bagunço* e *bombó*.

Diante do que se observa, evidenciam-se:

- a) desuso das formas *bagunço* e *bombó*, por apenas se registrarem no APFB;
- b) presença de lexias *buzina*, *ingaço*, *naipe* e *pendão*, somente no ALiB, caracterizando inovação;
- c) coincidência das formas *búzio*, *coração* e *umbigo* entre os atlas, das quais pôde-se observar:
 - i. forte redução de *búzio*, forma que, no APFB, se documenta em quatro localidades – Cabrália, Itaberaba, Jeremoabo e Santana. No ALiB, há o registro de *buza*, que pode ser inferido como uma variante de *búzia* em que houve erosão fonológica;
 - ii. ascendência do uso de *coração* – a forma que, no APFB, só se registrou em Barra, no ALiB passa a ser documentada também em Caetité, Carinhanha, Conquista, Itaberaba e Santana;
 - iii. permanência de *umbigo*, em Jacobina. Essa lexia pode ser considerada inovadora em Cabrália e em Santana, haja vista somente se registrar nos dados do ALiB..

Em síntese, para a área Atividades agropastoris, destacam-se:

- a) o desuso das lexias *pataqueiro* para trabalhador de roça alheia, *papuco* para sabugo e de *bagunço*, *búzio* e *bobó* (na localidade)
- b) a permanência *borrego* em Jeremoabo e de *cabrito* apenas em Carinhanha, Itaberaba, Jacobina, Santana, Vitória da Conquista.
- c) a difusão lexical de *coração* para inflorescência da bananeira que somente ocorria em Barra, e surge, conforme a base nos dados do ALiB, em Caetité, Carinhanha, Itaberaba, Santana e Vitória da Conquista.

De *Fauna*, destacam-se, aqui, as respostas das questões que se referem a 'sanguessuga', 'galinha sem rabo', 'galinha d'angola'

Sanguessuga (C-128/APFB e QSL-84/ALiB)

- a) *Sanguessuga* é a única forma coincidente no ALiB e no APFB.
- b) *Chupão* desapareceu.
- c) No APFB, *mazá* ocorria apenas em Jacobina. Já nos dados do ALiB, apenas em Santo Amaro. Ou seja, a lexia entra em desuso na localidade mas surge em outra.

Sem rabo

- a) *Nambu*, que somente ocorria em Barra, desapareceu.
- b) Há conservação de *suro(a)* em Caetité, Jeremoabo e Cabrália.
- c) *Suruco(a)* se mantém em Caetité, Carinhanha, Itaberaba e Santana. Passa a existir em Jacobina.

Galinha d'angola

- a) *Quenquém* desaparece.

- b) Há manutenção de uso de *saqué* em Itaberaba e em Jacobina; de *guiné*, em Jeremoabo e de *cocar* em Carinhanha, Jacobina, Cabália e Santana.

Observe-se o quadro, a seguir:

VARIANTE	Barra		Caetité		Carinhanha		Itaberaba		Jacobina		Jeremoabo		Cabrália		Santana		Conquista	
	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB	APFB	ALiB
sanguessuga				x	x	x		x		x		x		x		x		x
chupão		x	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x
mazá	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x
nambu		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
suruco(a)	x	x							x		x	x	x					x
suro(a)		x			x			x		x					x		x	
quenquém		x		x	x	x	x	x	x	x	x	x	x	x	x	x		x
saqué	x	x		x	x	x					x	x	x	x	x	x	x	
cocar	x		x				x				x	x					x	x
guiné	x	x	x	x	x	x	x	x	x	x			x	x	x	x	x	x
mocha																		
gambá	x										x						x	
cangambá	x		x		x		x		x				x		x		x	

Para a análise das respostas constantes da área *Corpo Humano*, optou-se pela subdivisão em: i) partes do corpo; ii) patologias e iii)

No tocante às 'Partes do corpo', analisaram-se:

Nuca

- a) *Cangote* se mantém em Caetité e em Cabrália e desaparece em Barra, Itaberaba, Jacobina, Jeremoabo.
- b) *Nuca* permanece nas localidades em que ocorria há 50 anos, conforme consta no APFB – Carinhanha, Cabrália e Santana. Houve, ainda, difusão lexical para as demais localidades, sendo hoje a forma predominante.
- c) *Cabeloro* é, conforme os dados do ALiB, uma forma em desuso na Bahia

Axila

- a) No APFB, *quiquio* era a forma majoritária, ocorrendo em Carinhanha, Itaberaba e Jacobina. Atualmente, conforme os dados do ALiB, apenas ocorre em Barra.

Rótula do joelho

- a) Há permanência de *pataca* em Caetité mas desaparece em Santana.
- b) Das localidades em que se realizava *bolacha* apenas Caetité sofreu desaparecimento.
- c) O comportamento de *rótula* é curioso: i) permanece apenas em Barra; ii) desaparece em Jacobina e em Jeremoabo e iii) é forma inovadora em Itaberaba, Cabrália, Santana e Conquista.

Tornozelo

As variantes *junta* e *peadouro* desapareceram.

Dente canino

Inversamente, *presa* se mantém em todas as localidades em que fora documentada há mais de 50 anos.

Dente siso

Dente quêro ou apenas *quero* permanecem.

Clavícula

- a. Jeremoabo se caracteriza pela manutenção das lexias *osso da fome*.
- b. Atualmente, se percebe a inserção de lexias decorrentes da alteração metaplásmica de clavícula. É possível pressupor que a escolarização ou meios de comunicação possam ter influenciado o uso

Útero

As lexias *cumade*, *sinhora do corpo*, *mãe do corpo* e *dona do corpo* não foram mais documentadas após 50 anos. Em lugar desses itens lexicais realizam-se alterações metaplásmicas de útero.

Calcanhar

O *calcanhar* permanece intacto.

Observe-se, a seguir, um quadro-resumo das variantes documentadas no APFB e sua manutenção ou desaparecimento conforme os dados do ALiB.

No tocante às patologias, destacam-se aqui, a variação para:

Terçol

- a) Das localidades coincidentes, apenas em Cabrália ainda se fala *terçol*
- b) Em contrapartida, *treiçol* que apenas ocorria em Jeremoabo, passa a ser realizada em Itaberaba, Jacobina, Cabrália e em Conquista.
- c) Espinha permanece em Barra, Caetité, Santana e Conquista.
- d) Não há mais *treçol* nas localidades em que ocorria. No entanto, essa variante ocorre em outras localidades que não foram documentadas pelo APFB.

Conjuntivite

- a) *Dordolho* em Jacobina, Jeremoabo e em Conquista.
- b) *Dordói* permanece em Caetité, Carinhanha, Itaberaba, Jacobina e Jeremoabo.

Argueiro / Cisco

Conforme as cartas do APFB, havia argueiro em Barra, Itaberaba, Jacobina, Jeremoabo e em Cabrália. Atualmente, segundo os dados do ALiB, houve apenas uma realização de argueiro em Juazeiro, localidade que não coincide com o APFB.

A questão do aborto na sociedade é, sempre, polêmica. Na língua, não parece ser diferente. A diversidade de usos para a designar a ação.

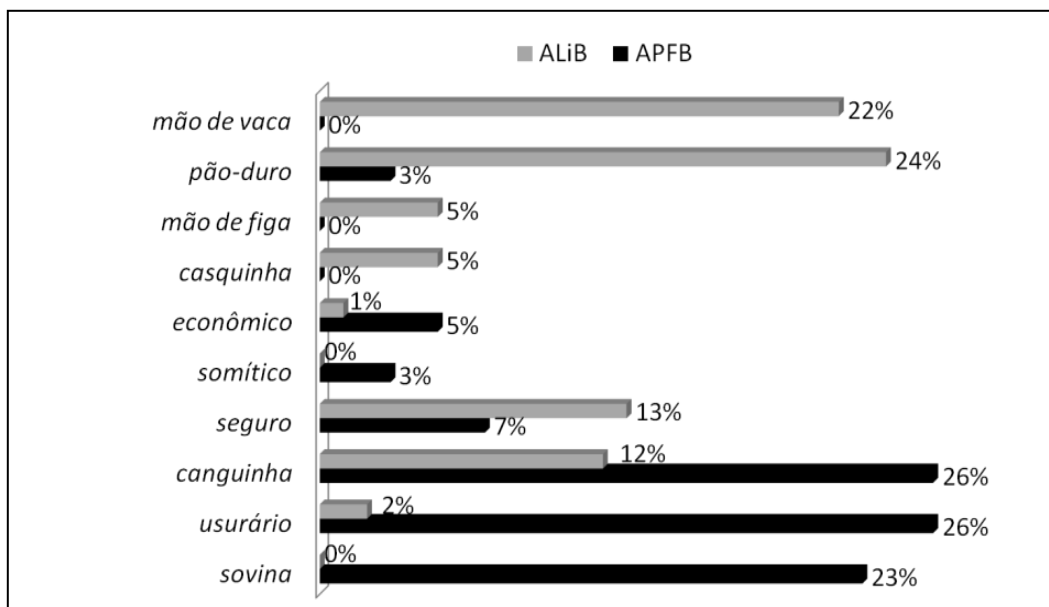
Questão constante da área *Ciclos da vida*, os dados da presente análise atestam:

- a) desuso de *móvito*, haja vista ser documentada apenas no APFB, possivelmente em função da inserção de *aborto*, que se caracteriza como forma majoritária no ALiB;
- b) inovação de seis lexias construídas por processos de composição tendo como base os verbos ‘perder’ – *perdeu, perdeu o nenê, perdeu o menino, perdeu a criança* – e ‘ter’ – *teve um aborto e teve uma perca*;
- c) redução significativa do item lexical *perca* – de 67% no APFB para 19%, no ALiB; iv)

O *Convívio e comportamento social* também apresenta sua diversidade.

Em busca dos “avarentos”, buscou-se comparar as lexias que se registram em ambos os atlas.

Gráfico: Confronto entre as formas mais produtivas do ALiB e do APFB



Diante do que se observa, é possível considerar que no período entre a publicação do APFB e a documentação feita pelo Projeto ALiB, tenha havido:

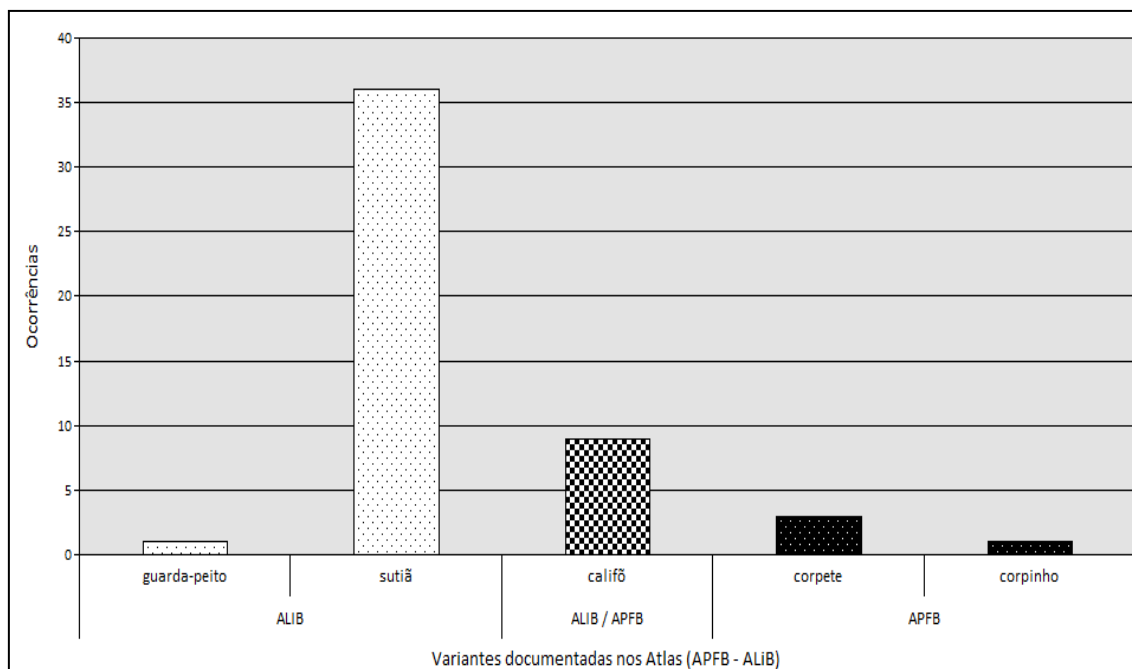
- a) ascendência, já que a forma *pão-duro*, passa de 3 a 24%;
- b) permanência, haja vista persistirem, embora com redução, o uso de *canguinha* (12%), *seguro* (7%), *usurário* e *econômico* (1%);
- c) desuso, com base na não realização das lexias *somítico* e *sovina* no ALiB –as quais no APFB se destacavam pela frequência 23 e 26%;
- d) inovação, tendo em vista que as formas *casquinha* (5%), *mão de figa* (5%) e *mão de vaca* (22%) não foram documentadas no APFB.

O esquema proposto, a seguir, ilustra os processos supracitados.

Esquema 1: Processo de mudança linguística entre as formas documentadas no APFB e pelo Projeto ALiB

ASCENDÊNCIA	PERMANÊNCIA				DESUSO		INOVAÇÃO		
<i>Pão duro</i>	<i>Canguinha</i>	<i>Seguro</i>	<i>Usurário</i>	<i>Econômico</i>	<i>Somítico</i>	<i>Sovina</i>	<i>Casquinha</i>	<i>Mão de figa</i>	<i>Mão de vaca</i>

Do *Vestuário*, mais especificamente, do vestuário feminino, a variação no uso do *sutiã* se destaca. Há, porém, outras opções como *guarda-peito*, *califom*, *corpete* e *corpinho*.



Do que se apresenta no gráfico acima, destacam-se:

- Apenas *califom* é uma forma coincidente entre os atlas.
- As formas constituídas com o radical *corp-*, – *corpete* e *corpinho* – apenas ocorrem no APFB.
- Guarda-peito* e *sutiã* são documentados apenas no ALiB.

No que tange à diatopia, a lexia *califom* ocorre, tanto no APFB quanto no ALiB, em três localidades – Jeremoabo Itaberaba e Jacobina. No ALiB, *califom* se realiza apenas entre indivíduos da faixa 2, o que pressupõe processo de mudança em curso, já que os jovens tendem ao uso de forma mais inovadora ou menos antiga.

Ao final das análises, torna-se relevante destacar fatos acerca do registro etimológico ou a origem de algumas variantes documentadas, a partir de uma análise criteriosa do étimo e ou origem dos itens lexicais, admitindo, sobretudo, os processos de formação.

- a) *Sanguessuga* (var. lat. vulgar, do lat. clássico *sanguisuga*), sobretudo por ressaltar a importância da observação do registro etimológico, já que em um primeiro momento, um pesquisador poderia ser levado a pensar que fosse uma forma constituída de processo de composição.
- b) *Corpete* (do it. *corpetto*) e *sutiã* (do fr. *soutien*), em função da acomodação fônica e adaptação ortográfica, encontram-se amplamente divulgada nos dicionários de língua portuguesa.
- c) *Perca* (var. port. *perda*, do lat. *perdere*) – variante da aborto (do lat. *abortus*), mesmo possuindo certa carga de estigma social quando usada em outros contextos, fato que põe em evidência a discussão do prestígio e agramaticalidade.

No tocante aos processos de formação de palavras, foram documentadas 27 formas que podem ser categorizadas: i) linguisticamente, por derivação e por composição, em que se incluem as colocações – probabilidade de uma unidade lexical coocorrer com outra – e os frasemas; ii) extralinguisticamente, por processos metafóricos e metonímicos.

Interessante é, também, discutir acerca do registro de lexias – compostas e complexas –, em obras lexicográficas, inclusive em dicionários etimológicos.

Tomando como exemplificação, as 10 variantes de *avarento*, notou-se que Cunha (1982) apenas apresenta informações acerca da forma canônica *avarento* (do lat. *avaritia*). Esse fato reforçou a ideia de investigar, também, processos de formação de ordem extralinguística, dentre os quais se destacam os processos metafóricos e metonímicos e a etimologia popular.

Embora não tenha sido objetivo deste trabalho observar as motivações extralinguísticas, dentre várias, uma especulação a respeito da origem da variante *pão duro*, expressa em um texto extraído de um blog, se revelou bastante curiosa:

Nos primórdios do século passado, vagava pelas ruas cariocas um mendigo muito popular (naquela época os mendigos eram escassos e populares). **A criatura esmolava comida, sempre implorando um pedaço de pão duro.** A reiteração da súplica imprimiu-lhe o apelido: **Pão Duro.** Falece o mendigo e descobre-se que seu refúgio **abrigava uma pequena fortuna em dinheiro**, contas bancárias e títulos. (...)E aí a expressão PÃO-DURO passou a designar o avarento [...] (PAIVA, 2006)

A hipótese além de curiosa serve de reflexão acerca dos processos de formação por extensão de sentido e por metáfora para a lexicalização de compostos ou composição de itens lexicais.

Como *pão-duro*, a forma em permanência *seguro* pode ser considerada como proveniente de metáfora. Notou-se, ainda, que as lexias inovadoras e em ascendência, *pão-duro*, *mão de vaca* e *mão de figa*, além de sugerirem ser provenientes de processo metafórico, são constituídas por processos de formação de palavras, mais especificamente, a composição, processo de formação de palavras bastante frequente nos dados desta pesquisa’.

Observou-se estreita relação entre a inovação – com base na documentação de lexias apenas no ALiB – e os processos de formação, sobretudo por composição, além da constituição de fraseologismos:

- a) inovação que se configura como variantes de *aborto* construídas por processos frásticos, tendo como base os verbos ‘perder’ – *perdeu*, *perdeu o nenê*, *perdeu o menino*, *perdeu a criança* – e ‘ter’ – *teve um aborto* e *teve uma perca*;
- b) criação do neologismo *guarda-peito*, que além de ser produto do processo de composição de dois itens lexicais do português;
- c) *remanso de água* (do lat. *remansus* + de, prep. lat. + água (do lat. *aqua*), *água corrente* (água, do lat. *aqua* + corrente, do lat. *currens,entis*), dentre outras formas parecem ser colocações, diferentemente do que ocorre com os compostos, os quais se encontram fixos na língua.

“A verdade prova que o tempo é o senhor dos dois destinos”.

Em contrapartida ao excerto da música Girassol, interpretada pela banda Cidade Negra, no tocante à língua, porém, não podendo evidenciar verdades mas, apenas, fatos e fenômenos presentes em um *corpus*, pode-se pressupor que o tempo, associado a outras tantas variáveis linguística e extralinguísticas, seja, apenas, fator do destino do léxico do português.

CONSIDERAÇÕES FINAIS

CONSIDERAÇÕES FINAIS

Like the legend of the phoenix
All ends with beginnings¹
(Daft Punk)

Acredita-se ser a epígrafe supracitada, uma oportuna forma de iniciar essas considerações que se traduzem em reflexões do que se pôde aprender, apreender e depreender do fazer lexicográfico, partindo da seguinte inquietação: Quais as implicações socioculturais da dicionarização de uma unidade lexical?

Não raro, no Brasil, são noticiados – e, por vezes, vividos e(ou) experienciados – fatos que revelam discriminação social, nas mais diversas possibilidades – raça, cor, credo, nível de escolaridade, gênero, faixa etária, etc.

Alguns ganham repercussão midiática, outros não, de modo a ser possível estabelecer uma escala de valoração: i) ampla e abrangente, quando se trata estritamente de preconceito social, grupos minoritários acolhidos pela mídia; ii) restrita a adeptos e simpatizantes, quando se refere a um grupo social com algum prestígio e visibilidade ou a um grupo político e; iii) ínfima, em que não há uma reação em massa, quando relacionada a um grupo minoritário menor e específico ou se refere à língua.

Muito além disso, se quis tratar, aqui, ainda que indiretamente, do preconceito linguístico, tendo em vista o fato de que a repercussão pública e midiática quanto aos atos discriminatórios parece se restringir ao poder que determinados grupos detenham socialmente e (ou) politicamente, como se língua não fosse, também, identidade.

Ora, num país em que parte da população defende a legalização de drogas, a liberdade de expressão e de afetividade, em que os direitos humanos são pregados, sobretudo no que concerne à defesa de raça e cor, é, no mínimo intrigante, que não se admita a variação linguística, em sua inteireza. Diante disso, pressupõe-se que o discurso de quebra de preconceitos e valorização do indivíduo tem sido hipócrita.

Em suma e mais uma vez, a Educação é a grande vítima do descaso social e político.

¹ Como a lenda da fênix, tudo termina com um novo começo (Daft Punk, tradução nossa)

É válido lembrar que entre maio e junho de 2011, a mídia brasileira protagonizou uma cena polêmica acerca de um livro didático que previa a variação linguística no âmbito do nível morfossintático. Linguistas – sociolinguistas, em geral – tiveram seus “quinze minutos de fama” e foram vistos, por muitos, como “vilões” que assassinaram ou, ao menos, intentaram contra a vida da língua portuguesa.

A fim de solucionar esses equívocos, ressalta-se a importância do registro lexicográfico de usos lexicais, em perspectiva sociodialetoal, que hoje são consideradas “erros” ou até mesmo os que possuem pouca expressividade quanto à veiculação, sobretudo, as alterações de caráter fônico no português brasileiro contemporâneo, identificando-as, sistematicamente, e relacionando-as aos processos históricos de mudança ocorridos na língua portuguesa no seu processo formativo. Acredita-se, pois, que dicionarizar um uso que é estigmatizado possa ser um grande avanço para a diminuição dos estigmas e preconceitos sociolinguísticos.

Esta tese foi construída sob essas bases, sob essa perspectiva, sob a base teórica metodológica da Lexicografia Variacional e da Dialetologia, aberta ao diálogo com outras áreas como a Antropologia Linguística e a Lexicultura, dentre outras. Embora seus pressupostos não estejam manifestos no texto, na parte visível da escrita desta tese, atesta-se que as ideias latentes foram cruciais para a elaboração deste trabalho e servirão, *a posteriori*, para o desenvolvimento de novas pesquisas.

Considerando o fato de que a Lexicografia Contemporânea tem optado por restringir documentar apenas o que é ou se tornou, por influência impositiva da norma, prestigiado, esta tese se propôs a fazer o inverso: dicionarizar, concomitantemente ao “mais normal”, ao padrão e, inclusive, ao erudito, caso ocorresse no *corpus*, o que não é prestigiado, o estigmatizado, a fim de que “flores” e “fiores” sejam reconhecidas como igualmente legítimas e exalem o perfume da última Flor do Lácio que brotou no Brasil, cuja beleza exuberante se revela na diversidade.

A “luta continua”. Para tanto, faz-se necessário que estudiosos de língua continuem a penetrar no reino das palavras como o poeta “à procura da poesia” e dos “poemas que esperam ser escritos”, com atenção e sem omitir os fatos e sem desconsiderar a sua supremacia, despindo-se, sobretudo, de todo e qualquer prejulgamento.

Afinal, os usos linguísticos ecoam por descobertas, se insinuam insistentemente e, muitas vezes, explodem na voz abafada dos excluídos. Aliás,

talvez por essa razão, seus utentes, abandonados, estigmatizados, marginalizados, vítimas da discriminação, se tornem invisíveis, camuflados por tantos preconceitos sociais. A tentativa de torná-los audíveis soa, para alguns puristas, como alguém que oferece seu corpo em troca da própria sobrevivência ou até mesmo como quem deseja morbidamente revelar a todos o prazer do diferente, o prazer daquilo que foge aos padrões de uma sociedade que não se (re)conhece em sua história.

Com vistas, pois, a tornar conhecida a história do léxico português, o Capítulo 2 desta tese foi construído. A partir da observação dos caminhos do léxico português, pôde-se comprovar que o objeto com que se trabalhou – ao qual se quis dar ênfase a sua própria legitimidade – se justifica na história.

Coube ao Capítulo 1, uma discussão acerca das unidades lexicais e de seu caráter nomeador a fim de refletir sobre o papel de significação dos usos linguísticos, sob a perspectiva de que “não existe um uso significativo de língua fora das interrelações pessoais e sociais situadas”. (MARCUSCHI, 2008, p.23)

Segundo Picoche (1992, p.19) “o significado de palavra não é só juntar morfemas, mas constitui uma síntese original, estabelecida por uma explicação histórica”. Leia-se, aqui, motivação pragmática, semântica, discursiva, cultural. A busca por motivações para os usos documentados se revela como possível proposta de trabalho.

Para além do produto final, o Vocabulário Dialectal Baiano, a pesquisa teve como objetivo, dentre os citados na Introdução, a reflexão sobre o papel do lexicógrafo variacionista e das bases teórico-metodológicas necessárias, de modo a constituir, para tanto, um novo ou um reciclado e (ou) reciclável objeto teórico-observacional, com vistas a contribuir, por conseguinte, para a consolidação de uma área ainda pouco explorada – se considerados os parâmetros adotados nos capítulos 3 e 4 – ainda pouco explorada.

Pois, se em dada medida os estudos lexicográficos têm, de certa maneira avançado no método em outra se mostram incipientes e até mesmo preconceituosos discriminando usos ou estigmatizando outros. Já os estudos dialetais, como visto no Capítulo 3 têm, ao longo dos tempos, se voltado à construção de Atlas linguísticos ou trabalhos de cunho lexicológico.

Esse fato justifica a constante inquietação para o estabelecimento de propostas de trabalhos lexicográficos, de cunho dialetológicos, ou dialetológicos, de cunho lexicográficos – citada, incessantemente, em toda tese.

Assim, sob a perspectiva da diversidade e da variação linguística, este trabalho, que visou à elaboração de um vocabulário dialetal, ou seja, um vocabulário que privilegiasse a descrição dos usos sem ignorar a realidade plurilinguística brasileira, mais especificamente a baiana, optou por desenvolver um método aplicável na interseção entre a Lexicografia e a Dialectologia cujas bases foram discutidas nos capítulos 3 e 4.

Como *venezianas* e *portinholas*, a análise permitiu a observação dos dados, vislumbrando aspectos relevantes à construção dos verbetes que constituíram o vocabulário.

A primeira epígrafe desta tese

Em situação de poço, a água equivale a uma palavra em situação dicionária: isolada, estanque no poço dela mesma, e porque assim estanque, estancada; e mais: porque assim estancada, muda e muda porque com nenhuma comunica, porque cortou-se a sintaxe desse rio, o fio de água por que ele discorria.

(MELO NETO, 1999, p. 350-351)

é, aqui, (re)citada no intuito de contrariá-la, parcialmente, pois, em perspectiva variacional, a situação lexicográfica não parece se equivaler a um “poço” mas a uma “lagoa” ou até mesmo um “riacho”, em que a água não se estagna em função da presença do fluxo da variação lexical.

Ao final, atesta-se que os caminhos percorridos, até aqui, embora árduos, não cansaram as mãos, pela certeza da contribuição sociocultural para os estudos linguísticos e de outras áreas.

Acredita-se sob a ótica de Bataille (1968) – segunda epígrafe principal deste trabalho – de que um dicionário tem início a partir do momento em que mostra os usos das palavras, ter encontrado resposta para a questão Qual o papel do lexicógrafo variacionista? Dar vez e voz, na escrita, às vozes, por vezes estigmatizadas, excluídas.

Com a licença do discurso acadêmico, trazem-se à tona as palavras de Jorge Amado em seu discurso de posse na Academia de Letras.

Com o povo aprendi tudo quanto sei, dele me alimentei e, se meus são os defeitos da obra realizada, do povo são as qualidades porventura nela existentes. (AMADO, 1981, p. 12)

Espera-se, enfim, que esta tese cumpra a missão que lhe foi proposta: a de servir de contributo sócio-pedagógico para os estudantes de letras e de áreas afins e, sobretudo, para sociedade brasileira, em específico, à baiana, revelando a língua dos guetos e das mansões desse berço de diversidade linguístico-cultural chamada Bahia tornando, por conseguinte, conhecido e reconhecido um pouco do vocabulário dialetal baiano.

REFERÊNCIAS

- AGUILERA, Vanderci de Andrade. *Atlas Lingüístico do Paraná*. Curitiba: Imprensa Oficial do Estado, 1994.
- ALMEIDA, Edilene Maria de Oliveira. *Atlas Linguístico da Mata Sul de Pernambuco-Almaspe*. 2009. 149f. Dissertação (Mestrado em Linguagens e Cultura) – Universidade Federal da Paraíba, João Pessoa, 2009.
- ALVES, Rubem. **O prazer de aprender**. Conferência proferida no Encontro de Psicopedagogos. São Paulo, julho/1990.
- AMADO, Jorge. **Discurso de Posse na Academia Brasileira de Letras**. In: CURRAN, Marck. J. Jorge Amado e a literatura de cordel. Fundação Cultural do Estado da Bahia - Fundação Casa de Rui Barbosa. 1981.
- AMARAL, Amadeu. **O dialeto caipira**: gramática, vocabulário. 4ª ed., São Paulo: Hucitec / Brasília: INL, 1982 (reprod. facsimil da 2ª ed.; 1ª ed. 1920).
- ANDRADE, Carlos Drummond de. **Poesia e prosa**. Rio de Janeiro: Nova Aguilar, 1988.
- ARAGÃO, Maria do Socorro Silva de.; BEZERRA DE MENEZES, Cleusa P. *Atlas Lingüístico da Paraíba*. Brasília: UFPB/CNPq, Coordenação Editorial, 1984; v. 1, 2.
- BAKHTIN, Mikhail (Volochinov). **Marxismo e filosofia da linguagem**. São Paulo: Hucitec, 2004.
- BAKHTIN, Mikhail (Volochinov). **Marxismo e filosofia da linguagem**. Trad. Michel Lahud e Yara Frateschi Vieira, 12 ed. São Paulo: Hucitec, 2006, p. 150.
- BARCELOS, A. M. F. **Understanding teachers' and students' language learning beliefs in experience: A Deweyan Approach**. Tese (Doutorado) - The University of Alabama, Tuscaloosa, 2000.
- BASÍLIO, Margarida. **Teoria Lexical**. São Paulo: Ática, 1987, p. 12.
- BASSETTO, Bruno Fregni. **Elementos de filologia românica**. EdUSP, 2001, p. 74.
- BATAILLE, Georges. **Documents**. Paris: Mercure de France, 1968.
- BECHARA, Evanildo. **Dicionário da língua portuguesa** Evanildo Bechara. Rio de Janeiro: Nova Fronteira, 2011.

- BECHARA, Evanildo. **Dicionário da língua portuguesa Evanildo Bechara**. Rio de Janeiro: Editora Nova Fronteira, 2011.
- BENEDUZI, Renata. **Valência verbal**: sua apresentação no comentário semântico de um dicionário de falsos amigos. Disponível em: <<http://hdl.handle.net/10183/60573>>. Acesso em: 06 jul. 2017.
- BESSA, José Rogério Fontenele (coordenador). *Atlas Linguístico do Ceará*. Vol.I – Introdução, Vol.II – Cartogramas. Universidade Federal do Ceará. Fortaleza: Edições UFC, 2010.
- BIDERMAN, M. T. C. **Dicionário didático de português**. São Paulo: Ática, 1998, p. 11.
- BIDERMAN, M. T. C. **Léxico e vocabulário fundamental**. São Paulo: Alfa, 1996.
- BIDERMAN, M. T. C. **Teoria Linguística**: Teoria linguística, Linguística Computacional. São Paulo, Martins Fontes, 2001.
- BIDERMAN, M.T.C. **Teoria Linguística**: teoria lexical e linguística computacional. São Paulo: Martins Fontes, 2001.
- BIDERMAN, Maria Tereza Camargo. **Aurélio**: sinônimo de dicionário?. In: ALFA: Revista de Linguística, São Paulo, vol. 44, p. 27-55, 2000. Disponível em: <http://hdl.handle.net/11449/107776>. Acesso em: 06 de jul. de 2017.
- BORBA, F. S. et al. **Dicionário de usos do português do Brasil**. São Paulo: Ática, 2002.
- BORBA, Francisco S. **Dicionário Unesp do português contemporâneo**. Curitiba:Piá, 2011.
- BORBA, Roberta de Carvalho e ABREU, Maria Helena Elpídio. **Questão Social**: Um debate sobre as perspectivas teóricas de análise. In: 4º Encontro Nacional de Política Social, UFES, Vitória/ES, 2009.
- BRASIL. Ministério da Educação Secretaria de Educação Básica. **Com direito à palavra**: dicionários em sala de aula. Brasília: Ministério da Educação Secretaria de Educação Básica, 2012.
- BRASIL. Projeto de Lei 1.676 de 1999. Dispõe sobre a promoção, a proteção, a defesa e o uso da língua portuguesa e dá outras providências. Diário da **Câmara dos Deputados**, Brasília, 4 nov. 1999. p. 106 a 108. Disponível em: <<http://imagem.camara.gov.br/Imagem/d/pdf/DCD04NOV1999.pdf#page=106>>. Acesso em: 06 jul. 2017.
- CÂMARA JUNIOR, Joaquim Matoso. **História e Estrutura da Língua Portuguesa**. 2ª Ed. Rio de Janeiro: Padrão, 1979.

CARDOSO, S. A. M.; FERREIRA, C. da S. **O léxico rural: glossário, comentários.** Salvador: Universidade Federal da Bahia, 2000. 138 p.

Cardoso, Suzana Alice et al. 2014a . **Atlas Linguístico do Brasil:** volume 1. Londrina: Eduel.

Cardoso, Suzana Alice et al. 2014b. **Atlas Linguístico do Brasil:** volume 2. Londrina: Eduel.

CARDOSO, Suzana Alice Marcelino da Silva. *Atlas Lingüístico de Sergipe II.* Rio de Janeiro: S. A. M. da S. Cardoso, 2002. 2v.

CARDOSO, Suzana Alice Marcelino et al. **Atlas Linguístico de Sergipe.** Salvador: Universidade Federal da Bahia; Fundação Estadual de Cultura de Sergipe, 1987.

CARDOSO, Wilton; CUNHA, Celso. **Estilística e Gramática Histórica:** Português através de textos. Rio de Janeiro: Tempo Brasileiro, 1978, p. 198.

CASTRO, I. (ed.) **Sete Ensaios Sobre a Obra de J. M. Piel.** Lisboa, Publicações do Instituto de Linguística da Faculdade de Letras de Lisboa, 1988.

CHAMBERS, J. K.; TRUDGILL, Peter (1980). **Dialectology.** Cambridge: Cambridge University Press.

CIDADE NEGRA. FARIAS, Bino; LAZÃO; GAMA, Da; GARRIDO, Toni. **Estrada.** *In:* Quanto Mais Curtido Melhor. Sony Music Entertainment (Brasil) I.C.L.: 1998. 1 CD. Faixa 5 (4 min 42).

CINTRA, Luís F. Lindley. **Estudos de Dialectologia Portuguesa.** Lisboa: Sá da Costa, 1983.

CINTRA, Luís F. Lindley. **Nova Proposta de Classificação dos Dialectos Galego-Portugueses.** *In:* Boletim de Filologia, Lisboa, Centro de Estudos Filológicos, 22, 1971, pp. 81-116.

COMITÊ NACIONAL DO PROJETO ALiB. **Atlas lingüístico do Brasil:** questionários. Londrina: Editora da UEL, 2001.

COROMINAS, J. **Breve diccionario etimologico de la lengua castellana.** Madrid, Gredos, 1997.

CORREIA, L. M. **Dificuldades de Aprendizagem Específicas:** Contributos para uma definição portuguesa. Coleção Impacto Educacional. Porto: Porto Editora, 2008.

COSTA, M.J.D.; GIRACCA, M.; OYARZABAL, M. **Culturema:** um breve histórico e as possibilidades de resignificação na tradução. I SILLIC – I

Simpósio Internacional de Lexicografia e Linguística Contrastiva - (no prelo), 2013.

CRISTIANINI, Adriana Cristina. *Atlas Semântico-Lexical da Região do Grande ABC*. 2007. 772f. Tese (Doutorado – Programa de Pós-Graduação em Linguística. Área de concentração: Semiótica e Linguística Geral) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2007.

CUBA, M. A. *Atlas Linguístico da Mesorregião Sudeste de Mato Grosso*. Dissertação (Mestrado em Estudo de Linguagens) – Universidade Federal de Mato Grosso do Sul, Campo Grande-MS, 2009.

CUNHA, A. G. **Dicionário etimológico nova fronteira da Língua Portuguesa**. Rio de Janeiro: Nova Fronteira, 1982.

DELICADO, António. **Adágios portugueses reduzidos a lugares communs /** pello lecionado Antonio Delicado, Prior da Parrochial Igreja de Nossa Senhora da charidade, termo da cidade de Euora. Lisboa, officina de Domingos Lopes Rosa, 1651. [12], 190 [i. é 187], [1] p. Disponível em: <http://clp.dlc.ua.pt/Corpus/AntonioDelicado.aspx> Acesso em 19 de dez. 2016.

Disponível em:

Disponível em: <<https://educacao.uol.com.br/disciplinas/filosofia/filosofia-da-linguagem-1-da-torre-de-babel-a-chomsky.htm>>.

ENCARNAÇÃO, Márcia Regina Teixeira da. *Atlas semântico-lexical de Caraguatatuba, Ilhabela, São Sebastião e Ubatuba* - municípios do Litoral Norte de São Paulo. 2010. 741f. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2010.

FARENZENA, Deize & DALPIAN, Laurindo. **A Preposição de:** do Latim ao Português. Disc. Scientia. Série: Artes, Letras e Comunicação, S. Maria, v. 9, n. 1, pp. 193-203, 2008. Disponível em: <<http://sites.unifra.br/Portals/36/ALC/2008/a%20preposi%E2%80%A1%C3%86o.pdf>> Acesso em 7 de jul. 2017.

FERNANDES, Célia Regina Delácio. **Letramento literário no contexto escolar**. In: GONÇALVES, Adair Vieira; PINHEIRO, Alexandra Santos. (Orgs.). Nas trilhas do letramento: entre teoria, prática e formação docente. São Paulo: Mercado das Letras, 2011.

FERNÁNDEZ, F Moreno. **Principios de Sociolingüística y Sociología del Lenguaje**. Barcelona: Ariel, 1998, p. 28-29.

FERREIRA, Carlota et al. *Atlas Linguístico de Sergipe*. Salvador: UFBA - Instituto de Letras/Fundação Estadual de Cultura de Sergipe, 1987.

FERREIRA, Emilia. **Alfabetização em Processo**. São Paulo: Cortez, 1996.

FURLAN, Oswaldo Antônio. **Língua e Literatura e sua Derivação Portuguesa**. Petrópolis, Rio de Janeiro: Vozes, 2006.

GALISSON, R. **Accéder à la culture partagée par l'entremise des mots à CCP. Études de Linguistique Appliquée**, 1987.

GALISSON, R. **Didactologie: de l'éducation aux langues-cultures à l'éducation par les langues-cultures**, 2002.

GALVES, C.M.C.; PAIXÃO DE SOUSA, M. C.; NAMIUTI, C. **Novas respostas para antigas perguntas: a periodização do português revisada**. VI Deutscher Lusitanistentag – VI Jornada Alemã de Lusitanistas, Leipzig, 2005.

GALVES, Charlotte M. C. **A língua das caravelas: periodização do português europeu e origem do português brasileiro**. In: A. Castilho; Maria Aparecida T. Morais; R. E. V. Lopes & S. M. L. Cyrino. (Org.). Descrição, história e aquisição do português brasileiro. Campinas: Pontes, 2007, p. 513-528.

GECKELER, H. **Semántica estrutural y teoría del campo léxico**. Madrid: Gredos, 1971, p 95.

GEIGER, Paulo (org.). **Novíssimo Aulete dicionário contemporâneo da língua portuguesa**. Rio de Janeiro: Lexikon, 2011.

GENETTE, G. **Discurso da narrativa**. Tradução de Fernando Cabral Martins. Lisboa: Vega, 1984, p. 33.

Giracca, Mirella Nunes. **Os Culturemas Presentes nos Folhetos Turísticos da Região Sul do Brasil: as técnicas utilizadas pelos tradutores**. Mirella Nunes Giracca; orientadora, Maria José Roslindo Damiani Costa. Florianópolis, SC, 2013. 141 p. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Comunicação e Expressão. Programa de Pós-Graduação em Estudos da Tradução.

GREIMAS, A.J. **Semântica Estrutural**. São Paulo, Cultrix, 1966, p. 5.

GUILBERT, Louis. **La créativité lexicale**. Paris: Larousse, 1975.

HOUAISS, A.; VILLAR, M. S. **Dicionário Houaiss de Língua Portuguesa**. Elaborado pelo Instituto Antônio Houaiss de Lexicografia e Banco de Dados da Língua Portuguesa S/C Ltda. Rio de Janeiro: Objetiva, 2009.

HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro, Ed. Objetiva, 2001.

HOUSE, R. J. A. **Theory of Charismatic Leadership**. In: J. G. Hunt and L. L. Larson. *In Leadership: The cutting edge*. Shouthern Illinois University. 1977.

HUERTA, Pedro Mogorron; MEJRI, Salah. **Las construcciones verbo-nominales libres y fijas**: Aproximación contrastiva y traductológica. Universidad de Alicante, 2008, p.191-202.

IBFC. Instituto Brasileiro de Formação e Capacitação. **Caderno de Questões: Especialista em Regulação**: Agência Estadual de Regulação de Serviços Públicos de Energia, Transportes e Comunicações da Bahia (AGERBA), 2016. (IBFC). Cargo Especialista em Regulação Escolaridade Superior. O texto de que se extraiu a citação está disponível, apenas em versão digital, *on line*, em: <<http://educacao.uol.com.br/filosofia/ult3323u52.jhtm>>.

ISQUERDO, A. N. Brasileirismos, regionalismos e americanismos: desafios e implicações para a lexicografia brasileira. In: BERLINCK, R. de A.; GUEDES, M. e MURAKAWA, C. de A. A. (Org.) **Teoria e análise linguísticas: novas trilhas**. Araraquara: Laboratório Editorial FCL/UNESP; São Paulo: Cultura Acadêmica, 2006. pp. 11-29. (Trilhas Linguísticas, 8)

JUNG, Milton. **Rebobinar, Cair a Ficha e Virar o Disco**: expressões que ficaram no tempo, 2014. Disponível em: <<https://miltonjung.com.br/2014/11/18/rebobinar-cair-a-ficha-e-vira-o-disco-expressoes-que-ficaram-no-tempo/>>. Acesso em: 06 jul. 2017.

KOCH, Walter; Klassmann, Mário Silfredo; ALTENHOFEN, Cléo. *Atlas Lingüístico-etnográfico da Região Sul do Brasil*. Porto Alegre/Florianópolis/Curitiba: Ed. UFRGS/Ed. UFSC/ Ed. UFPR, 2002. v. 1, v. 2.

KOIKE, Kazumi. **Colocaciones Léxicas en el Español Actual**: estudio formal y léxicosemántico. Madrid: Universidad de Alcalá y Universidad de Takushoku, 2001.

LANGACKER, Ronald W. **A Linguagem e Sua Estrutura**. Petrópolis: Vozes, 1975, p. 100.

LOBATO, Monteiro. **Idéias de Jeca Tatu**. São Paulo: Globo, 2008, p. 67-68.

LOBO, Tânia (2003). **A questão da periodização da história linguística do Brasil**. In: CASTRO, Ivo; DUARTE, Inês. (Orgs.). Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mira Mateus, v. I. Lisboa: IN-CM.

LOBO, Tânia Conceição Freire. **A questão da periodização da história linguística do Brasil**. In: CASTRO, Ivo; DUARTE, Inês (orgs.). Razões e emoção: miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional/Casa da Moeda, 2003.

LUCCHESI, Dante. **A periodização da história sociolinguística do Brasil**. DELTA, São Paulo , v. 33, n. 2, p. 347-382, ago. 2017 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-

44502017000200347&lng=pt&nrm=iso>. acessos em 13 set. 2017.
<http://dx.doi.org/10.1590/0102-445067529349614964>.

MACHADO FILHO, Américo V. L. **Dicionário Etimológico do Português Arcaico**. Salvador: Edufba, 2013.

MACHADO FILHO, Américo V. L. **Um Ponto de Interseção para a Dialectologia e a Lexicografia**: a proposição de um dicionário dialetal brasileiro com base nos dados do ALiB, *Estudos Linguísticos e Literários*, 41, p. 51-52, 2010.

MACHADO, Luiz Carlos Pinheiro; MACHADO FILHO, Luiz Carlos Pinheiro. **Dialética da Agroecologia**. São Paulo: Expressão Popular, 2014.

MARTINET, André. **A linguística sincrônica**. Rio de Janeiro: Edições Tempo Brasileiro, 1974.

MATTOS E SILVA, Rosa Virgínia. **Para a história do português culto e popular brasileiro**: sugestão para uma pauta de pesquisa. *Cadernos de Letras da UFF, Niterói*, v. 34, p. 11-30, 2008.

MATTOS E SILVA, Rosa Virgínia. **Sobre a terminologia metalinguística na primeira metade do século XVI: alguns aspectos**. *Estudos Linguísticos e Literários*, Salvador, n. 37-38, p. 135-164, 2008.

MATTOS E SILVA, Rosa Virgínia. **Teorias da mudança linguística e a sua relação com a(s) história(s) da língua(s)**. *Linguística - Revista de Estudos Linguísticos da Universidade do Porto*, Porto, v. 3, p. 39-53, 2008.

MEJRI, Salah. **Constructions à verbes supports**: collocations et locutions verbales. Disponível em: <<https://hal.archives-ouvertes.fr/halshs-00410950/document>>. Acesso em: 06 jul. 2017.

MEL'CHUK, I. A., A. CLAS; A. POLGUÈRE. **Introduction à la Lexicologie Explicative et Combinatoire**. Louvain-laNeuve: Duculot, 1995.

MELO NETO, João Cabral. **Rios sem discurso**. In: *Obra completa*. Rio de Janeiro: Nova Aguilar, 1999.

MOLINA, Lucía Martínez. **Análisis descriptivo de la traducción de los culturemas árabe-español**. (Tese de doutorado) Bellaterra: Departament de Traducció i d'Intepretació, Universitat Autònoma de Barcelona, 2001.

MOLLARD-DESFOUR, Annie, Le vert: **Dictionnaire de la couleur**, mots et expressions d'aujourd'hui Xxe-Xxle, Paris, Editions du C.N.R.S., 2012.

NASCENTES, Antenor. **O linguajar carioca**. Rio de Janeiro: Organização Simões, 1953.

NEIVA, Isamar. A expressão de futuridade no português brasileiro contemporâneo: a variação em dados do Projeto ALiB. 2012. 193 folhas. Dissertação de Mestrado – Universidade Federal da Bahia.

NEWMARK, Peter. **A Textbook of Translation**. New York: Prentice Hall, 1995.

NEWMARK, Peter. **A textbook of translation**. New York: Prentice Hall, 1988.

NIDA, Eugene. **Language structure and translation**. California: Stanford University Press, 1975.

NORD, C. **Text Analysis in Translation** - Amsterdam: Rodopi. Transl. from the German by C. Nord and P.Sparrow. German original: Textanalyse und Übersetzen. Heidelberg: Julius Groos, 1991.

NORD, C. **Textanalyse und Übersetzen**. Theorie, Methode und didaktische Anwendung einer übersetzungsrelevanten Textanalyse. Tübingen: Stauffenburg 1988.

OLIVEIRA, Dercir. Pedro de (Org.). *ALMS - Atlas Lingüístico de Mato Grosso do Sul*. 1. ed. Campo Grande: Editora UFMS, 2007. 271 p.

OLIVEIRA, Fernão de (2000 [1536]). **Gramática da Linguagem Portuguesa**. Edição crítica, semidiplomática e anastática por Amadeu Torres e Carlos Assunção, com um estudo introdutório do Prof. Eugenio Coseriu. Lisboa, Academia das Ciências de Lisboa.

PAIM, Marcela Moura Torres. **A variação lexical do português falado no Brasil**: Reflexões sobre o campo semântico vestuário e acessórios nos dados do Projeto ALiB.

<[http://www.cchla.ufrn.br/visiget/pgs/pt/anais/Artigos/Marcela%20Moura%20Torres%20Paim%20\(UFBA\).pdf](http://www.cchla.ufrn.br/visiget/pgs/pt/anais/Artigos/Marcela%20Moura%20Torres%20Paim%20(UFBA).pdf) >. Acesso em: 21 mar. 2017.

PAIVA, E. V. de (Org.). **Pesquisando a formação de professores**. Rio de Janeiro: DP&A, 2003.

PEREIRA, Maria das Neves. *Atlas geolinguístico do litoral potiguar*. Rio de Janeiro: UFRJ, Faculdade de Letras, 2007. 2v. Vol I: 123p. mimeo. Vol II 189p. mimeo. Tese de Doutorado em Letras Vernáculas.

PICOCHÉ, J. **Precis de Lexicologie Française**: l'étude et l'enseignement du vocabulaire. Paris: Nathan, 1992.

PIEL, Joseph Maria. **Origens e estruturação histórica do léxico do português**. In: CASTRO, Ivo (org.) Estudos de lingüística galego-portuguesa. Lisboa: Imprensa Nacional – Casa da Moeda, 1989[1976].

POTTIER, B. **Linguistique générale**. Théorie et description. Paris: Klincksieck, 1974.

POTTIER, Bernard; AUDUBERT, Albert; PAIS, Cidmar Teodoro. **Estruturas Linguísticas do Português**, São Paulo, Difusão Européia do Livro, 1973.

QUINTANA, Mario. **O Último Viandante**. In: A cor do invisível, 1989. Disponível em: <http://www.avozdapoesia.com.br/obras_ler.php?obra_id=16604&poeta_id=302>. Acesso em: 06 jul. 2017.

RAZKY, Abdelhak. (Org.) *Atlas lingüístico sonoro do Pará*. Belém: PA/CAPES/UTM, 2004. CDRoom.

RIBEIRO, Silvana Soares Costa. **Brinquedos e brincadeiras infantis na área do Falar Baiano**. Tese. Salvador: 2012.

ROMANO, Valter Pereira. *Atlas Geossolinguístico de Londrina: um estudo em tempo real e tempo aparente*. 2012. 366f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2012.

ROSA, Guimarães. **Primeiras Estórias**. Editora Nova Fronteira: Rio de Janeiro, 1988, p. 13.

ROSSI, Nelson; ISENSEE, Dinah Maria; FERREIRA, Carlota. *Atlas Prévio dos Falares Baianos*. Rio de Janeiro: INL, 1963.

SÁ, Edmilson José de. *Atlas Linguístico de Pernambuco*. Tese (Doutorado em Letras) – Universidade Federal da Paraíba. João Pessoa. 2013.

SACCONI, Luiz Antônio. **Grande Dicionário Sacconi (2009)** São Paulo: Nova Geração, 2009.

SACCONI, Luiz Antonio. *Grande Dicionário Sacconi da Língua Portuguesa*. São Paulo: Nova Geração, 2009.

SANROMÁN, Iriarte Álvaro. **A Unidade Lexicográfica. Palavras, Colocações, Frasemas, Pragmatemas**. Braga: Centro de Estudos Humanísticos: Universidade do Minho, 2001.

SANTOS, Antonia Vieira dos. **Compostos sintagmáticos nominais VN, NN, NA, AN e NprepN no português arcaico (sécs. XIII-XVI)**. Tese. Salvador, 2009.

SCHMITZ, Egídio. **Fundamentos da Didática**. 7ª Ed. São Leopoldo, RS: Editora Unisinos, 2000.

SEABRA, Maria Cândida Trindade Costa de. (Org.) *O Léxico em Estudo*. Belo Horizonte: Faculdade de Letras da UFMG, 2006, p. 7.

SILVA E GUIMARÃES, Maria de Nazaré Serra. **Por uma análise sintático-semântica das preposições**. *Revista de Estudos de Língua Portuguesa*, [S.l.], v. 4, n. 4, p. 63-89, dez. 1987. ISSN 2447-0554. Disponível em:

<http://www.periodicos.letras.ufmg.br/index.php/lingua_portuguesa/article/view/7073/6080>. Acesso em: 06 jul. 2017.

SILVA NETO, Serafim da. **História da Língua Portuguesa**. 4ª ed. Rio de Janeiro: Presença, 1986. [1ª ed. 1950]

SILVA, Josué Cândido da. **Filosofia da linguagem (1): Da Torre de Babel a Chomsky**.

STERKENBURG, Van. **Lexicography: Principles and Practice**. Londres: Academic Press. 1983. p. 80.

TARALLO, Fernando. **A Pesquisa Sociolinguística**. São Paulo: Ática, 1986, p. 8.

TARALLO, Fernando. **Tempos linguísticos**. São Paulo: Ática, 1990.

TEIXEIRA, J. **Palavra**. *In*: Logos. Enciclopédia Luso-Brasileira de Filosofia, vol. 3. Lisboa: Verbo, 1991, p. 1316-1317.

TESNIÈRE, L. **Esquisse d'une syntaxe structurale**. Paris, 1953, p. 53-56.

TEYSSIER, Paul. **História da língua portuguesa**. São Paulo: Martins Fontes, 1997.

TINHORÃO, José Ramos. **Os sons dos negros no Brasil: cantos, danças folguedos: origens**. São Paulo: Art.

TSITSA. **Le français de l'art culinaire un projet dictionnaire**, 2004.

VASCONCELOS, S. D. Re-estruturação do curso de Licenciatura em Ciências Biológicas (UFPE): identificação do perfil do aluno. *In*: ENCONTRO "PERSPECTIVAS DO ENSINO DE BIOLOGIA", 7., 2000, São Paulo. Anais... São Paulo: USP.

VERDELHO, Telmo. **Dicionários: testemunhos da memória linguística**. *In*: Linguística Histórica e História da Língua Portuguesa: actas do Encontro de Homenagem a Maria Helena Paiva. Porto, Universidade do Porto. Faculdade de Letras, 2004, p. 417.

VERMEER, H. J. **Voraussetzungen fur eine Translationstheorie**. Einige Kapitel Kulturund Sprachtheorie: Heidelberg 1986.

VIEIRA, R. **Identidades pessoais: interações, campos de possibilidade e metamorfoses culturais**. Lisboa: Colibri, 2009.

XATARA, C. **Dicionário de expressões idiomáticas francês do Brasil e de Portugal-francês da França, da Bélgica e do Canadá**. São José do Rio Preto, SP, 2013. UNESP/FAPESP (www.deipf.ibilce.unesp.br).

XATARA, C. **Dictionnaire d'expressions idiomatiques portugais du Brésil et du Portugal - français de la France, de la Belgique et du Canada.** 2013. Disponível em: <www.deipf.ibilce.unesp.br>. Acesso em: 06 jul. 2017.

XATARA, C. M.; SECO, M. **Culturemas em contraste: idiomatismos do português brasileiro e europeu.** *In*: Domínios de Linguagem, v. 8, n. 1 (jan./jun. 2014), p. 502-519. ISSN 1980-5799. Disponível em: <<http://www.ileel2.ufu.br/anaisdosilel>>. Acesso em 24 de junho 2014.

ZÁGARI, Roberto L. et al. *Esboço de um atlas lingüístico de Minas Gerais.* Rio de Janeiro: Fundação Casa de Rui Barbosa, 1977.

APÊNDICES

**ACIDENTES
 GEOGRÁFICOS**

onda(s) - Movimento de elevação das águas do mar' // *Como se chama o movimento da água do mar?* ~ **onde** ~ **zonda** ≅ **maré**
 ↔ **marola** ↔ **trombas de água**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

riacho(inho) - Rio de pequena extensão' // *Como se chama um rio pequeno, de uns dois metros de largura?* ~ **riachinho** ~ **riachozinho** ≅ **baixa** ≅ **córrego** ≅ **estreito** ≅ **lagoa(inha)** ≅ **pequeno rio** ≅ **poço** ≅ **rego** ≅ **riacho forte** ≅ **ribeirão** **rio estreito** ≅ **rio pequeno** ≅ **riote** ≅ **riozinho** ≅ **valeta**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

**ATIVIDADES
 AGROPASTORIS**

amendoim Leguminosa de formato longitudinal e aspecto gômico, em cujo interior constam sementes de cor avermelhada usada na alimentação humana em diversas formas de cocção.' // *QSL 40 Como se chama o grão coberto por uma casquinha dura que come assado, cozido, torrado ou moído?* ~ **amendoins** ~ **amenduim** ~ **amiduim** ~ **mendoim** ~ **menduim** ~ **miduim** ~ **minduim**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

girassol 'Flor herbácea, de grandes pétalas amarelas e uma parte central, que segue a trajetória do Sol' // *Como se chama uma flor grande, amarela, redonda, com uma rodela de sementes no meio?* ~ **Mirassol**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

penca(s) 'Cada um dos grupos de cachos de banana' // *Como se chama cada parte que se corta do cacho da bananeira para pôr para madurar / amadurecer?* **cacho** ≅ **cacho de (da) banana** ≅ **pendão**¹ {*Da o nome de penca, as penca. Aí diz vamos tirar as penca*} **087.3**

ALIMENTAÇÃO E COZINHA

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

bala 'Pequena guloseima feita de calda de açúcar solidificado, vendida embrulhada em papel' // QSL 185 - *Como se chama aquilo embrulhado em papel colorido que se chupa? (MOSTRAR) (PEDIR PARA DESCREVER)* ≅ **bala** ≅ **bala doce** ≅ **balinha** ≅ **bombom(ns)** ≅ **caramelo** ≅ **quemado**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

canjica¹ 'Iguaria que consiste em uma papa cremosa, em ponto de corte, de milho verde ralado e cozido com leite, açúcar e com côco' // QSL 179 *Como se chama uma papa cremosa, feita com côco e milho verde ralado, polvilhada com canela?* **curau** ≅ **lele** ≅ **mingau**² ≅ **mingau de milho** ~ **mingau de mí** {*A diferença é que um é líquido e outro é pedaço. Mingau, e a canjica é o que corta. Não, num pode fazer ela sem coco não. Eh, tem que ser coco e o milho junto.*} **093.3** → **canjica**² → **canjica**³

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

café ↔ **café da manhã**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

café da manhã - Primeira refeição do dia, feita pela manhã, com ou sem o uso do café como bebida' // QSL 176 - *Como se chama a primeira refeição do dia, feita pela manhã?* ≅ **café** ≅ **café da manhã** ≅ **cafezin primêro da manhã** ≅ **dejejum** ≅ **quebra de jejum** ≅ **tirar o jejum** {*A gente fala o café da manhã*} 085.3

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

CICLOS DA VIDA

caçula(o) - Forma de tratamento ao mais novo dos filhos ou dos irmãos' // QSL 131: "[...] o filho que nasceu por último?" ≈ **caçulo** ~ **contra-peso** ~ **derradêro** ~ **mais novo** ~ **ponta de raiz** ~ **rapa do caco** ~ **resto de partição** ~ **resto do tacho**.

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

gêmeo(a)(s) Forma de tratamento aos indivíduos que nasceram de um mesmo parto' // QSL 125 - *Como se chama duas crianças que nasceram no mesmo parto?* **gêma** ~ **gêmo(a)** ≈ **gêmea(s)**
irmão baso ≈ **mabaça** ≈ **mabação** ≈ **mabaço** ≈ **mabaça** ≈ **mabaço**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

discansar ↔ **parir** → **parteira** → **abortar**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

irmão de leite Forma de tratamento ao indivíduo que foi amamentado pela mãe de outro, em razão da ausência de leite de sua mãe. // QSL 129 - *O próprio filho da "ama-de-leite" e a criança que ela amamenta são o que um do outro?* **irmã de leite** ≈ **irmã emprestada** ≈ **irmão** ≈ **irmão de lei** ≈ **irmão de mama** ≈ **irmão por parte de leite** ≈ **irmãos** ≈ **irmãos adotivos** ≈ **irmãos de consideração** ≈ **irmãos de leite** ≈ **irmãos por parte do leite** ≈ **irmãs de leite** ≈ **meio irmão** ↔ **irmão de leite** → **mãe de leite**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

finado Forma de tratamento ao indivíduo após sua morte, em algumas comunidades, quando não se deseja falar, apenas, o seu nome de registro, fazendo menção de que está morto' // QSL 135 - *Numa conversa, para falar de uma pessoa que já morreu, geralmente as pessoas não a tratam pelo nome que tinha em vida. Como é que se refere a ela?* ≈ **finada(o)** ≈ **finado(a)** ≈ **afinado** ≈ **defunto** ≈ **difunto** ≈ **falecida** ≈ **falecido** ≈ **falicida** ≈ **falicido**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

menina Forma de tratamento à criança do sexo feminino' // QSL 133 - *E se for do sexo feminino, como se chama?* ≈ **garota(inha)** ≈ **guria** ≈ **menina**
mulé ≈ **minina** ≈ **piquena** ≈ **nigrinha**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

parir Ação de concluir a gravidez, culminando no nascimento de um feto desenvolvido' // QSL 124 - *Chama-se a <<parteira>> quando a mulher está para ___? ≅ dar (a) luz ≅ dar neném ≅ descansar ≅ estar incomodada ≅ ganhar criança ≅ ganhar menino ≅ ganhar neném ≅ ganhar o bebê ≅ ganhar o menino ≅ ganhar o neném ≅ incomodada ≅ parir ≅ pegar neném ≅ ter (a) criança ≅ ter filho ≅ ter menino ≅ ter neném ≅ ter o filho ≅ ter o menino ≅ ter o neném ≅ ter um filho ≅ ter um parto ≅ ter uma parição → parteira → abortar {Pra ganhar neném, pra ter minino, pra ter neném, não falava tá pra **parir** não, **parir** falava não, era vergonhoso, tá pra **parir** não falava não, pra ter minino, pra ganhá neném, pra dar luz mas não falava parir não, de jeito nenhum} **082.3***

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

parteira `Profissional que assiste aos partos e ministra primeiros socorros às parturientes' *que não é médica // QSL 123 - *Como se chama a mulher que ajuda a criança a nascer? aparadeira ≅ enfermeira*

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

CONVÍVIO E COMPORTAMENTO SOCIAL

burro(a)(inho) `Forma de tratamento pejorativo ao indivíduo não inteligente ou que tem dificuldade de aprender' // QSL 137 - *Como se chama a pessoa que tem dificuldade de aprender as coisas? ≅ burrinho ≅ burro(a)(inho) ≅ cabeça-dura ≅ cabiçudo ≅ cavalo ≅ desatento ≅ desinteligente ≅ idiota ≅ ingnorante ≅ jumento ≅ menos inteligente ≅ mente fraca ≅ muco ≅ retardado ≅ rude ≅ rudo(a) ~ rudia ≅ tolo {Eu num... num chamo meus alunos de **burro**, entendeu? As pessoas chamam as outras que têm dificuldade de aprender de **burro**. Agora, obviamente, qu'eu, enquanto pedagoga, não vou chamar meu aluno de **burro**.}* **093.6**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

cornio `Forma de tratamento pejorativa ao indivíduo que foi traído pelo cônjuge' // QSL 141 - *Como se chama a pessoa que é paga para matar alguém? ≅ besta ≅ boi ≅ boinho ≅ chapéu de bode ≅ chibungo ≅ chifrudo ≅ conformado ≅ convencido ≅ cornélio ≅ enganado ≅ galhudo ~ gaiudo ≅ guidom ≅ homi traído ≅ levou ponta ≅ maniacca ~ manicaca ≅ mói de galha ≅ tôro*

			092	087					
			097	084	081				

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialectal Baiano*, por áreas temáticas:
Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma
localidade. Ou em uma mesorregião

	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

pistoleiro 'Profissional contratado para assassinar pessoas' // *QSL 139 - Como se chama a pessoa que é paga para matar alguém?*
assassino ≅ **assassino de aluguel** ≅ **assassino particular** ≅ **assassino por encomenda** ≅ **bandido** ≅ **cangacêro** ≅ **capanga** ≅ **comparsa** ≅ **criminoso** ≅ **jagunço** ≅ **malfazejo** ≅ **malfeitor** ≅ **matador** ≅ **matador de aluguel** ≅ **matador por encomenda**

	092	087						
	097	084	081					
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

xará 'Forma de tratamento aos indivíduos que possuem o mesmo nome' // *QSL 143 - Como se chama a pessoa que tem mesmo nome da gente?*

	092	087						
	097	084	081					
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

CORPO HUMANO

banguelo Forma de tratamento ao indivíduo que não tem alguns ou todos os dentes' *Corpo Humano. QSL 100 - Como se chama a pessoa que não tem dentes? ≈ banguela ≅ disdentado ≅ sem dente*

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

catarata¹ 'Patologia ocular que consiste na opacidade do cristalino (a lente do olho) ou da membrana envoltória' /Corpo Humano/ *QSL 096 - Como se chama aquela pele no olho que dá em pessoas mais idosas? ~ cascarata ~ cataraca ~ cataraquita ~ catarara ≅ carne no olho ≅ lua branca ≅ pili ≅ raspagem ≅ treilé {Catarata. Aqui, o que eu vejo falar é ela.} 088.3 → catarata²(inha)*

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

cisco 'Partícula de poeira, que acidentalmente, entra no olho' // *QSL 090 - Como se chama alguma coisinha que cai no olho e fica incomodando? ≅ arguêro ≅ garce ≅ puera*

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialeto Baiano*, por áreas temáticas:
Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma
localidade. Ou em uma mesorregião

	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

dente quêro Cada um dos terceiros dentes molares que surgem ger. entre os 17 e os 21 anos de idade' // QSL 097 - *Como se chama esses dois dentes pontudos?* (APONTAR) **den'de coelho** \cong **dente da sorte** \cong **dente de quêro** \cong **dente do siso** \cong **dente quêro** \cong **quêro** \cong **siso** {*Eu sei qual é, que eles dói pra daná. É Dente quêro*} 088.4

	092	087						
	097	084	081					
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

gogó 'Saliência na parte anterior do pescoço do homem, formada pela cartilagem tireoide' // QSL 105 - *Como se chama esta parte alto do pescoço do homem?* **caroço** \cong **engato** \cong **garganta** \cong **garganta do homem** \cong **nó** \cong **nó da garganta** \cong **nó da guela** \cong **nó na garganta** \cong **pato** \cong **pescoço** \cong **pomo de adão**

	092	087						
	097	084	081					
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

meleca(inha) 'Mucosidade nasal' // QSL 102 - *Como se chama a sujeirinha dura que se tira do nariz com o dedo?* \approx **melecazinha** **caraca** \sim **caca** \sim **carquereca** \sim **catara** \sim **cataraca** \sim **catarata(inha)** \approx **cataratazinha** \sim **cataratinha** \sim **catareca** \cong **capa de sujera** \cong **capelinha** \cong **catarro** \cong **catela** \cong **catota** \cong **cera** \cong **cera do catarro** \cong **meleca(inha)** \approx **melecazinha** \cong **péla** \cong **perereca** \cong **remela** \cong **sujêra** \cong **tacotaraca**

	092	087						
	097	084	081					

nuca Parte posterior e superior do pescoço, sobre a vértebra chamada atlas' Corpo Humano. QSL 104 - *Como se chama isto?* (APONTAR) \sim **nunca** \cong **cachaça** \cong **cangote** \cong **fonte** \cong **osso da nunca**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

peito(s) Parte do corpo feminino onde se situam as glândulas mamárias' Corpo Humano. QSL 111 - *Como se chama a parte do corpo da mulher com que ela amamenta os filhos?* \sim **peitho** \sim **pêtho** \sim **pêto** **busto** \cong **mama(s)** \cong **seio(s)** \cong **teta**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

presa(s) Dente pontudo e perfurante, situado entre o incisivo lateral e o pré-molar, em número de dois em cada maxilar, que permite rasgar os alimentos' Corpo Humano. QSL 096 - *Como se chama aquela pele no olho que dá em pessoas mais idosas?* **canino(s)** \cong **den'de vampiro** \cong **dente canino** \cong **dente do vampiro** \cong **dente quêro** \cong **incisivos** \cong **quebrado de osso** \cong **quêro**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

lombo² ↔ **anca(s)** → **lombo¹**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

FAUNA

chifre(s) 'Apêndices ósseos pontiagudos presentes na parte superior da cabeça de mamíferos, como bois, cabras etc.'
// *O que o boi tem na cabeça?* ~ **chife**
armação ≅ **gaia** ≅ **gaio** ≅ **ponta**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

muriçoca 'Inseto de pequeno porte que, geralmente à noite, emitem sons desagradáveis' // *Como se chama aquele inseto pequeno de perninhas compridas que canta no ouvido das pessoas, de noite?* (IMITAR ZUMBIDO) ~ **miriçoca** ≅ **borrachudo** ≅ **cadongo** ≅ **chincudo** ≅ **jatinho** ≅ **mosquito** ≅ **muruum** ≅ **musquito** ≅ **mutuca** ≅ **perereca** ≅ **pernelongo** ~ **pernalonga** ~ **pernilonga** ~ **pernilongo** ≅ **tiririca**

gambá 'Animal mamífero que, quando se sente ameaçado, emite fedor' Fauna.
QSL 071 - *Como se chama o bicho que solta um cheiro ruim quando se sente ameaçado?* ≅ **sariguê** ≅ **saruê**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

joão(inho) de barro 'Ave conhecida por seu característico ninho de barro em forma de forno' // QSL 066 - *Como se chama a ave que faz a casa com terra nos postes, nas árvores e até nos cantos da casa?* ≈ **joãozinho de barro jandainha** ≅ **janica de barro** ≅ **janique de barro** ≅ **joana de barro** ≅ **joanica de barro** ≅ **joaninha** ≅ **joaninha de barro** ≅ **joanita de barro** ≅ **joão barrão** ≅ **joão de barro** ≅ **joãozinho de barro** ~ **joãozim de barro** ≅ **maria de barro** ≅ **maria pobre**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

papagaio¹ 'Ave, similar a uma galinha, de penas pretas e pintas brancas. // _QSL 068 - *Como se chama a ave de penas coloridas que, quando presas, podem aprender a falar?* ~ **papagai** ~ **papagalho** **arara** ≅ **louro** ~ **lôro** ≅ **minha rosa** ≅ **periquito** ~ **piriquito** {É porque pa velho num tem coisa melho que um **papagaio**, né. Só quem pode tulerá velho é um papagaio} **088.3** → **papagaio²** → **papagaio³** → **papagaio⁴**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

rabo¹ 'Cauda de animais quadrúpedes' // *Como se chama a parte com o que boi espanta as moscas?* ≅ **cabo** ≅ **cauda** ≅ **coro** ≅ **rabada** → **rabo²**

		092	087						
		097	084	081					

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialeto Baiano*, por áreas temáticas:
Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma
localidade. Ou em uma mesorregião

	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

rabo² 'Pelo flexível e resistente da cauda do cavalo e de alguns outros animais // Como se chama o cabelo comprido na traseira do cavalo? **cabo** ≅ **cauda** ≅ **rabicho** ≅ **rabo** ≅ **rabo do cavalo** → **rabo¹**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

urubu 'Ave de penas preta que se alimenta de carnes putrefadas' // QSL 064 - Como se chama a ave preta que come animal morto podre? ~ **arubu** ~ **orubu** ~ **unrubu** **abutre** ≅ **carnicêro** ≅ **covo** ≅ **uruvango** ↔ **urubu**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

HABITAÇÃO

isqueiro 'Utensílio que substitui o fósforo, geralmente, utilizado para acender cigarro, ' Habitação. Para acender um cigarro, se usa fôforo ou ____? ~ ~ **isqueire** ≅ **artifício**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

lanterna 'Utensílio portátil de iluminação com lâmpada elétrica e pilhas eletroquímicas' Habitação. Como se chama aquele objeto que se usa para clarear no escuro e se leva na mão, assim (MÍMICA)? ≅ **cachulate** ≅ **fechiaile** ≅ **lâmpada** ~ **lâmpida** ~ **lâmpira** ≅ **pilha** {Lanterna, chamava pilha também por conta das pilhas de dento} **093.7**

		092	087					
		097	084	081				
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

tramela 'Trava retangular, geralmente de madeira, que gira em torno de um eixo preso em janela, porta ou portão para trancá-los.' Habitação. Como se chama aquela pecinha de madeira que gira ao redor de um prego para fechar a porta, janela? **bisalha** ≅ **escora** ≅ **feche** ≅ **ferrôio** ≅ **passador** ≅ **táuba** ≅ **tranca** ≅ **trava** ≅ **travanca** ≅ **travessa**

		092	087					
--	--	-----	-----	--	--	--	--	--

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialectal Baiano*, por áreas temáticas:
Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma localidade. Ou em uma mesorregião

			097	084	081				
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

esconder ≅ **brincar de (se)**
esconder ≅ **brincar de bicho**
≅ **brincar de tonga** ≅
esconderijo ≅ **jangolô** ≅ **nego**
fugido ≅ **picula de se**
esconder ≅ **pique-esconde** ≅
(se) esconder

				092	087				
				097	084	081			
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

JOGOS E DIVERSÕES INFANTIS

balanço¹ 'Brinquedo que consiste em um assento suspenso por cordas fixadas num suporte, em que uma criança, sentada e com o impulso do próprio corpo, faz movimentos oscilatórios.' *Em algumas comunidades, apenas uma corda.' // *Como se chama uma tábua pendurada por meio de cordas, onde uma criança se senta e se move para frente e para trás?* ≅ **balango** ≅ **balança**² ≅ **gangorra**² ≅ **zanza**¹ → **balanço**²

				092	087				
				097	084	081			
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

gangorra¹ 'Brinquedo constituído de uma prancha de madeira firmada numa base central de ajustamento oscilante, de modo que duas pessoas, geralmente crianças, possam se sentar em cada uma das extremidades e lhe imprimir movimentos alternados de baixo para cima e vice-versa.' // *Como se chama uma tábua apoiada no meio em cujas pontas sentam duas crianças e quando uma sobe a outra desce?* ≅ **balaço de pau**
balança ≅ **balança mas não cai** ≅ **balanceio** ≅ **balanço** ≅ **zanza** → **gangorra**²

				092	087				
				097	084	081			
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

balanço² ↔ **gangorra**¹ → **balanço**¹

				092	087				
				097	084	081			
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

pipa^{1(s)} 'Brinquedo infantil que consiste em uma armação leve de varetas, recoberta de papel fino, e que se empina no ar por meio de uma linha'. *Difere-se de "**piriquito**^{2(inho)}" pela presença de varetas.' // *Como se chama o brinquedo feito de varetas, cobertas de papel, que se empina no vento por meio de uma linha?* **arara** ≅ **arraia**
~ **raia** ≅
carambola ≅ **caxote** ≅
papagaio ≅ **peixe(inho)** ≅
pipa(s) ≅ **piriquito** ≅ **suru** ≅ **suruco**

				092	087				
				097	084	081			
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

esconde-esconde 'Brincadeira de crianças em que uma destas fica de olhos fechados pelo período de tempo em que as demais se escondem e, depois, sai à procura delas.' // *Como se chama a brincadeira em que uma criança fecha os olhos enquanto as outras correm para um lugar onde não vistas e depois essa que fechou os olhos vai procurar as outras?* ≅ **bacondê**¹ ≅ **brincadeira de**

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialeto Baiano*, por áreas temáticas:
 Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma
 localidade. Ou em uma mesorregião

			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

VESTUÁRIO

calcinha 'Peça íntima do vestuário feminino, que cobre a área pélvica' Vestuário e acessórios. *Como se chama a roupa que a mulher usa debaixo da saia?*
 ≅ **calça de baixo** ≅ **calçola** ~ **caçola** ~ **carçola** ≅ **tanga**

			092	087				
			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

cueca 'Peça íntima do vestuário masculino que cobre a área pélvica' Vestuário e acessórios. *Como se chama a roupa que o homem usa debaixo da calça?*
calção ≅ **ceroula** ~ **cilora** ~ **cirôla** ≅ **samba-canção** ≅ **shortinho**

			092	087				
			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

sutiã 'Peça íntima do vestuário feminino usada para sustentar, modelar e cobrir os seios' Vestuário e acessórios. *Como se chama a peça do vestuário que serve para segurar os seios?* ~ **sutião** ~ **sutien**
califom ~ **calofone** ~ **galifom**
 ≅ **corpete** ≅ **golêro** ≅ **guarda-peito** ≅ **sigura goleiro**

	092	087						
--	-----	-----	--	--	--	--	--	--

VIDA URBANA

ônibus¹ 'Veículo de grande porte, usado para o transporte de passageiros, cuja rota preestabelecida' // *Como se chama a condução que leva mais ou menos 40 passageiros e faz o percurso dentro da cidade?* ~ **ôndibus¹** ~ **ônibo¹**
besta ≅ **bus** ≅ **buzu** ≅ **circular**
 ≅ **coletivo** ≅ **condução** ≅ **lotação** ≅ **micro ônibus** ≅ **ônibu circular** ≅ **ônibus coletivo** ≅ **transporte** ≅ **transporte coletivo** ≅ **transporte urbano** → **ônibus²**

			092	087				
			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

ônibus² 'Veículo de grande porte, usado para o transporte de passageiros, cuja rota preestabelecida' Vida urbana. *Como se chama a condução que leva mais ou menos 40 passageiros de uma cidade para outra?* ~ **ôndibus²** ~ **ônibo²**
buzu ≅ **coletivo** ≅ **condução** ≅ **interurbano** ≅ **ôndibus de viagem** ≅ **ônibu rodoviário** ≅ **ônibus de linha** ≅ **ônibus interurbano** ≅ **ônibus rodoviário** ≅ **ônibus urbano** ≅ **transporte rodoviário** → **ônibus¹**

	092	087						
--	-----	-----	--	--	--	--	--	--

APÊNDICE I – Amostras de verbetes do *Vocabulário Dialeto Baiano*, por áreas temáticas:
 Lexias que ocorrem em todo o espaço baiano analisado ou exceto em uma
 localidade. Ou em uma mesorregião

			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

quebra-mola 'Obstáculo de pouca altura construído transversalmente no leito de ruas e estradas para forçar a redução da velocidade dos veículos' Vida urbana. *Como se chama aquele morrinho atravessado no asfalto para os carros diminuírem a velocidade?* \cong **banguela²** \cong **gelo baiano** \cong **lombada** \cong **tartarura**

			092	087				
			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

terreno(inho) 'Área não construída de uma propriedade, em ambientes mais urbanos' Vida urbana. *Como se chama a área que é preciso ter ou comprar para se fazer uma casa na cidade?* \approx **terreninho** \cong **área de terra** \cong **chão** \cong **lote** \cong **loteamento** \cong **pedaço de terra** \cong **posse** \cong **terreno vadio**

			092	087				
			097	084	081			
	096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082

**ACIDENTES
GEOGRÁFICOS**

mareta(s)² ~ maleta ≈ maretinha ↔
onda(s)(inha) → **mareta(s)¹** {*Não, aqui a gente é... hoje tá maretano muito, né? Aqui no rio a gente chama **mareta**, né?*} 084.3

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

♦ APFB-ALIB ∩ Santana - 092; Caetité - 097; Barra - 084; Juazeiro - 081; Caravelas - 102; Santa Cruz Cabralia - 101

**ATIVIDADES
AGROPASTORIS**

borrego ~ **burrego(inho)(a)** ↔
cabrito(a)(inho)(a)

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

♦ APFB-ALIB ∩ Jeremoabo - 082;

cabrito(a)(inho)(a) Filhote da ovelha com até um ano de idade! // *Como se chama a cria da ovelha logo que nasce? E até que idade se dá esse nome?* ≈
cabritinha ≈ **cabritinho** ~
cabritim ≈ **cabrita bezerro** ≅
bizerrinho ≅ **bizerro(inho)** ≅
bodinho ≅ **borrego** ≅
burrego(inho)(a) ≅ **burreguim** ≅
burreguinha ≅ **burreguinho** ≅
cabra ≅ **carneirinho** ≅
carneiro(inho) ≅ **filhote** ≅
filhote de ovelha ≅ **marrã** ≅
novilho ≅ **ovelha(inha)** ≅
ovelhinha {*Enquanto ele tá mamano é **cabrito**, num sabe, é um cabritinho. Aí, ele passa a ser um carneiro, aí muda o nome*} **095.1**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

♦ APFB-ALIB ∩ Carinhanha - 097; Itaberaba - 090; Jacobina - 086; Santana - 092; Vitória da Conquista - - -

capuco ~ **capu** ↔ **sabugo**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

♦ APFB-ALIB ∩ Jacobina - 086; Jeremoabo - 082

FAUNA

cocar 'Ave, similar a uma galinha, de penas pretas e pintas brancas'. // QSL 067 - *Como se chama a ave de criação parecida com a galinha de penas pretas com pintinhas brancas?* ≅ **caqui** ≅ **cocar** ≅ **cocó** ≅ **conquém** ≅ **coqué** ≅ **galinha d'água** ≅ **galinha d'angola** ≅ **galinha de(a) angola** ≅ **galinha pedrês** ≅ **guiné** ≅ **perdiz** ≅ **pintado** ≅ **saqué** ≅ **sariema** ≅ **tô fraco**

		092	087						
		097	084	081					
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

◇ APFB-ALiB n Carinhanha - 097;
Jacobina - 086; Santa Cruz Cabrália -
101; Santana - 092

sanguessuga // *Como se chama o bicho que se gruda nas pernas das pessoas quando elas entram num <<córrego>>?* ~ **samissuga** ~ **sanguexuga** ~ **xambexuga** ~ **xamessuga** ~ **xamexuga** **lerma** ≅ **maribá** ≅ **mazá** ≅ **mutuca**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

◇ APFB-ALiB n Barra - 084

guiné ↔ **cocar**

		092	087						
		097	084	081					
		098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

◇ APFB-ALiB n Jeremoabo - 082

mocha 'Cabra que não possui chifres' *Fauna. Como se chama a cabra que não tem <<chifre>>?* ≅ **alejada**¹ ≅ **cabra lisa** ≅ **cabra mocha** ≅ **cabra sem chifre** ≅ **cabra sem ponta** ≅ **mocho** ≅ **moja** ≅ **muchu** ≅ **mucho** ≅ **muje** ≅ **sem chifre** ≅ **suro**⁴

		092	087						
		4	3						
		097	084	081					
		1,3,4							
	098	100	089	095	090	085	086		
	3		3	1,4	2	3	1,2,4		
	101	099	094	093	091	088	083	082	
	3		3	5	3		3		

◇ APFB-ALiB n Carinhanha - 097;
Santana - 092; Caetitê - 097; Barra -
084; Juazeiro - 081; Jeremoabo - 082;
Caravelas - 102; Santa Cruz Cabrália -
101; Vitória da Conquista - 098

ACIDENTES GEOGRÁFICOS

(variantes secundárias de ‘redemoinho’)

redimuinho² ↔ onda(s) (inha) → redimuinho¹

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

redimunho ↔ redemoinho

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

remoinho ↔ redemoinho

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

remuinho ↔ redemoinho

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

remunho ↔ redemoinho

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

ATIVIDADES AGROPASTORIS

(variantes secundárias de ‘aipim’)

aïmpim ↔ aipim {Aqui chama de *aïmpim*. Lá pelo sul, Rio e São Paulo o povo chama ela de *mandioca*.} **101.3** APFB

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
		3	2	1, 2	2	2,3,4	2			
102	101	099	094	093	091	088	083	082		
2, 3		4	1,2,3,4	2,5,6,7,8	3, 4	1, 4	2,3,4			

(variantes secundárias de ‘amendoim’)

amendoins ↔ amendoim

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
102	101	099	094	093	091	088	083	082		

amenduim ↔ amendoim

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
		3, 4	2	1	2		1, 3			
102	101	099	094	093	091	088	083	082		
1, 3	4	2, 4		1,4,5,6,7,8				4		

amiduim ↔ amendoim

		092	087							
		097	084	081						
	096	098	100	089	095	090	085	086		
			4							
102	101	099	094	093	091	088	083	082		
4	3									

APÊNDICE III – Amostras de verbetes do *Vocabulário Dialectal Baiano*, por áreas temáticas:
Lexias metaplásmicas (variantes secundárias)

(variantes secundárias de ‘camomila’)

camomilha ↔ **camomila**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

camumila ↔ **camomila**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

**ALIMENTAÇÃO
E COZINHA**

(variantes secundárias de ‘muncunzá’)

mucunzá ↔ **muncunzá**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

mucuzá ↔ **muncunzá**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

mugunzá ↔ **muncunzá**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

muguzá ↔ **muncunzá**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

APÊNDICE III – Amostras de verbetes do *Vocabulário Dialeto Baiano*, por áreas temáticas:
Lexias metaplásmicas (variantes secundárias)

mungunzá ↔ muncunzá

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

munguzá ↔ muncunzá

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

CICLOS DA VIDA

(variantes secundárias de ‘menopausa’)

manopausa ↔ menopausa → menstruação

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

menapausa ↔ menopausa → menstruação

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

menaspausa ↔ menopausa → menstruação

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

CORPO HUMANO

(variantes secundárias de ‘clavícula’)

cavico ↔ clavícula

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

cavícua ↔ clavícula

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

cavícula ↔ clavícula

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

cavingula ↔ clavícula

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

pêtho ↔ peito(s)

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

pêto ↔ peito(s)

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

FAUNA

(variantes secundárias de ‘papagaio’)

papagaio¹ ‘Ave, similar a uma galinha, de penas pretas e pintas brancas. // *_QSL 068 - Como se chama a ave de penas coloridas que, quando presas, podem aprender a falar? ~ papagai ~ papagalho arara ≅ louro ~ lôro ≅ minha rosa ≅ periquito ~ piriquito* {É porque pa velho num tem coisa melhó que um *papagaio*, né. Só quem pode tulerá velho é um *papagaio*}
088.3 → **papagaio²** → **papagaio³** → **papagaio⁴**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

papagai ↔ papagaio

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

papagalho ↔ papagaio

			092	087					
			097	084	081				
		098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

(variantes secundárias de ‘varejeira’)

barijêra ↔ varejeira

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

vaguijêra ↔ varejeira

		092	087						
		097	084	081					
		098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

vaigêra ↔ morotó¹

		092	087						
		097	084	081					
		098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

varêja ↔ varejeira

		092	087						
		097	084	081					
		098	100	089	095	090	085	086	
	101	099	094	093	091	088	083	082	

HABITAÇÃO

lâmpada ~ lâmpida ~ lâmpra ↔ lanterna

			092	087					
			097	084	081				
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

lâmpida ↔ lâmpada ↔ lanterna

			092	087					
			097	084	081				
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

lâmpra ↔ lâmpada ↔ lanterna

			092	087					
			097	084	081				
		096	098	100	089	095	090	085	086
102	101	099	094	093	091	088	083	082	

VESTUÁRIO

caçola ↔ calçola ↔ calcinha

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

**calçola ~ caçola ~ carçola ↔
 calcinha**

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	

carçola ↔ calçola ↔ calcinha

			092	087					
			097	084	081				
	096	098	100	089	095	090	085	086	
102	101	099	094	093	091	088	083	082	