

**UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS HUMANAS E LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
MESTRADO EM LETRAS
ÁREA DE CONCENTRAÇÃO: ESTUDOS DE LINGUAGEM**

RODRIGO ALVES SILVA

**REGIONALISMOS DO CAMPO SEMÂNTICO *JOGOS E DIVERSÕES INFANTIS*
DO *ALIB* NOS DICIONÁRIOS ELETRÔNICOS *HOUAISS* E *AURÉLIO***

TERESINA – PI

2017

RODRIGO ALVES SILVA

**REGIONALISMOS DO CAMPO SEMÂNTICO *JOGOS E DIVERSÕES INFANTIS*
DO *ALIB* NOS DICIONÁRIOS ELETRÔNICOS *HOUAISS E AURÉLIO***

Dissertação apresentada ao Programa de Pós-Graduação em Letras da Universidade Federal do Piauí como requisito parcial para obtenção do título de Mestre em Letras.

Orientador: Prof. Dr. Marcelo Alessandro Limeira dos Anjos

TERESINA – PI

2017

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Setorial do Centro de Ciências Humanas e Letras
Serviço de Processamento Técnico

S586r Silva, Rodrigo Alves.
Regionalismos do campo semântico *jogos e diversões infantis do ALiB* nos dicionários eletrônicos *Houaiss e Aurélio* / Rodrigo Alves Silva. – 2017.
111 f. : il.

Dissertação (Mestrado em Letras) – Universidade Federal do Piauí, 2017.

Orientação: Prof. Dr. Marcelo Alessandro Limeira dos Anjos.

1. Regionalismos. 2. Dicionário Houaiss. 3. Dicionário Aurélio. 4. ALiB. I. Título.

CDD 469.798

RODRIGO ALVES SILVA

**REGIONALISMOS DO CAMPO SEMÂNTICO *JOGOS E DIVERSÕES INFANTIS*
DO *ALIB* NOS DICIONÁRIOS ELETRÔNICOS *HOUAISS E AURÉLIO***

Banca Examinadora

Prof. Dr. Marcelo Alessandro Limeira dos Anjos (UFPI – Presidente)

Profa. Dra. Lucirene da Silva Carvalho (UESPI)

Prof. Dra. Maria Auxiliadora Ferreira Lima (UFPI)

AGRADECIMENTOS

A Deus, autor e mantenedor da vida, provedor de todas as coisas e o principal responsável por essa conquista, pois me deu forças quando mais precisei. Ele é digno de toda honra, glória e gratidão.

Aos meus pais, Ana Paula Alves de Abreu e Raimundo de Moura Silva, por sempre acreditarem em mim. E crendo, forneceram todos os meios necessários, no que puderam e no que não puderam, para que eu chegasse até aqui. Obrigado pela educação, pelos ensinamentos e por todo o amor.

À minha família – irmãos, tios (as), primos (as), avós – que sempre torceram por mim, me apoiaram em minhas decisões e me compreenderam nos momentos de ausência.

Aos meus amigos e irmãos de fé, os quais me deram todo o apoio necessário para que eu chegasse até aqui. Obrigado pelas orações e pelas palavras de estímulo.

Aos meus professores, desde a alfabetização até o ensino médio. Todos eles foram peças fundamentais na construção do que sou hoje, pois de cada um tirei lições e ensinamentos que me serviram para me constituir cidadão e profissional.

Aos meus professores da graduação e da pós-graduação, que me estimularam, em cada disciplina, o interesse pelo mundo das Letras.

Ao prof. Marcelo dos Anjos, meu orientador, que, desde a minha primeira monitoria (2012.1) até aqui, se dispôs a me instruir na formação acadêmica, orientando-me, cuidadosamente, em todos os trabalhos, inclusive na dissertação. Sem essa orientação, o trabalho não se realizaria.

Aos meus colegas pós-graduandos, pelo exemplo na dedicação aos estudos.

Às minhas parceiras de estudo: Gláucia (pela amizade e parceria constante desde a graduação), Meryane (pela amizade e companhia nas atividades do mestrado); Virna (pela amizade e pela ajuda na aquisição no *ALiB*, juntamente com o prof. Marcelo); Layana (pela amizade e parceria nos trabalhos e nos congressos acadêmicos).

Ao Núcleo de Pesquisa e Ensino do Português (NuPEP), pela experiência em pesquisas em grupo.

Às professoras Maria Auxiliadora e Catarina de Sena, pelas preciosas contribuições ao meu trabalho dadas na Qualificação.

À Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES), pela concessão de bolsa durante dezenove meses, que permitiu a minha permanência e produtividade no programa de pós-graduação.

Ao Programa de Pós-Graduação em Letras (UFPI), pela oportunidade de participação no programa e por fornecer os subsídios necessários à minha formação.

Ao Programa de Pós-Graduação em Estudos Linguísticos da UNESP, campus São José do Rio Preto, por me permitir cursar uma disciplina condensada, que, certamente, serviu de direcionamento para a realização deste trabalho.

RESUMO

Esta pesquisa se insere no âmbito da Metalexigrafia e tem, como objetivo geral, analisar o tratamento dado aos regionalismos do campo semântico *Jogos e diversões infantis* do *Atlas Linguístico do Brasil (ALiB)* no *Dicionário eletrônico Houaiss da língua portuguesa (DEH)* e no *Dicionário eletrônico Aurélio da língua portuguesa (DEA)*. Os objetivos específicos são verificar como esses regionalismos e suas variantes são tratados nos dicionários; identificar quais as convergências e as divergências entre os dados do *ALiB* e dos dicionários quanto à localização geográfica dos regionalismos; identificar as variantes que são registradas ou não nos dicionários; e, por fim, propor, quando cabível, algumas possíveis soluções para um tratamento adequado dos regionalismos. Para atingir tais objetivos, foi traçado o seguinte percurso metodológico: 1) levantamento das variantes lexicais do campo semântico *Jogos e diversões infantis* presentes no *ALiB*; 2) consulta no *DEH* e no *DEA* das variantes lexicais selecionadas, a fim de perceber como elas estão registradas e o que se diz sobre elas, sobretudo quanto à localização geográfica; 3) comparação entre as informações dadas no *ALiB* e nos dicionários; 4) elaboração de quadros para a organização dos dados comparados, classificando-os como convergentes ou divergentes; 5) elaboração de quadro para a classificação das variantes em diferentes categorias, segundo o tratamento que elas recebem nos dicionários. A pesquisa partiu das discussões a respeito do problema do tratamento de regionalismos em dicionários, baseadas em Isquierdo (2006, 2007) e em Fajardo (1996-1997). Esses autores afirmam que os dicionários devem se basear em pesquisas dialetológicas e geolinguísticas à hora de precisar a natureza dos regionalismos. Este trabalho também se fundamentou na discussão sobre a formação sócio-histórica do português brasileiro, segundo autores como Lucchesi (2015), Faraco (2016), Rodrigues (2010), entre outros; das discussões sobre a Lexicografia e a Metalexigrafia, baseando-se em autores como Welker (2004), Biderman (1984) e Haensch (1982), entre outros, e dos estudos sobre a Dialetologia, fundamentando-se em Cardoso (2010). Com a análise dos dados, percebeu-se que a maioria das informações dos dicionários sobre determinados regionalismos divergem do *ALiB*. Além disso, foi possível detectar diferentes formas de tratamento de regionalismos nos dicionários analisados, quais sejam: i) não registro de variantes; ii) registro de variantes, mas sem aceção que se refira a jogo, brinquedo ou brincadeira; iii) registro de variantes, mas sem localização geográfica; iv) registro de variantes com localização geográfica. Diante disso, foi possível perceber a necessidade dos dicionários de atualizarem as informações sobre regionalismos, sobretudo quanto à localização geográfica, baseando-se, principalmente, em atlas linguísticos.

Palavras-chave: Regionalismos. Dicionário *Houaiss*. Dicionário *Aurélio*. *ALiB*.

ABSTRACT

This research is inserted in the field of Metalexigraphy and has as general objective to analyze the treatment destined to regionalisms of the semantic field *Games and child enjoyment (ALiB)* on the *Houaiss Electronic Dictionary Portuguese Language (DEH)* and *Aurélio Electronic Dictionary of the Portuguese Language (DEA)*. The specific objectives are to verify how these regionalisms are treated on the dictionaries; to identify what are the convergences and divergences between the data from *ALiB* and the dictionaries regarding the geographic localization of the regionalisms; and, finally, to propose, when possible, some possible solutions to a proper treatment towards regionalisms. In order to achieve such objectives the following methodological path was established: 1) performing a survey on the lexical variants in the semantic field *Games and child enjoyment (ALiB)*; 2) checking on *DEH* and *DEA* for the selected lexical variants in order to understand how they are registered and what is said about them, mainly in regard to the geographic localization; 3) comparison between the data found on *ALiB* and on the dictionaries; 4) elaborating charts to organize the compared data, classifying them as convergent or divergent; 5) elaborating charts for classifying the variants into different categories, according to the treatment they are given on the dictionaries. The research started from discussions over the problem of the treatment of regionalisms on dictionaries, based on Isquerdo (2006, 2007) and Fajardo (1996-1997). These authors assert that dictionaries must be based on dialectical and geolinguistic researches when defining the nature of regionalisms. This study was also based on the discussion about the social-historical formation of Brazilian Portuguese, according to Lucchesi (2015), Faraco (2016), Rodrigues (2010), among others; in addition to the discussions on Lexicography and Metalexigraphy, based on Welker (2004), Biderman (1984) and Haensch (1982), and others, and the studies about Dialectology in accordance to Cardoso (2010). Once the data were analyzed, it was noticed that the majority of information on the dictionaries about certain regionalisms diverge from *ALiB*. Moreover, it was possible to detect ways of treating regionalisms on the analyzed dictionaries, which are: i) non-registering of variants; ii) registering variants, but without meanings referring to games or toys; iii) variant register, but with no geographic location; iv) variant register with geographic location. In face of that, it was possible to perceive the need for the dictionaries to update the information concerning regionalisms, mainly concerning the geographic location, based on linguistic atlas.

Keywords: Regionalism. *Houaiss* Dictionary. *Aurélio* Dictionary. *ALiB*.

LISTA DE FIGURAS

Figura 1: Tipologia lexicográfica segundo Haensch (1982)	55
Figura 2: Tipologia lexicográfica segundo Biderman (1984a).....	58
Figura 3: Tipologia lexicográfica segundo Welker (2004)	59
Figura 4: Microestrutura do DEH no modo tradicional	64
Figura 5: Microestrutura do DEH no modo interativo	65
Figura 6: Microestrutura do DEA	66

LISTA DE QUADROS

Quadro 1: Modelo de quadro para apresentação dos dados	81
Quadro 2: Modelo de quadro para classificação das variantes	82
Quadro 3: Variantes de cambalhota	83
Quadro 4: Variantes de bolinha de gude	88
Quadro 5: Variantes de estilingue	92
Quadro 6: Variantes de <i>brinquedo de empinar (com varetas, com linha)</i>	94
Quadro 7: Variantes de brinquedo de empinar (sem varetas, com linha)	95
Quadro 8: Variantes de cabra-cega (brincadeira).....	98
Quadro 9: Variantes de amarelinha.....	100
Quadro 10: Classificação das variantes.....	102

LISTA DE GRÁFICOS

Gráfico 1: Proporção da classificação das variantes.....	104
---	-----

LISTA DE ABREVIATURAS E SIGLAS

ABL – Academia Brasileira de Letras

ALiB – Atlas Linguístico do Brasil

ANPOLL - Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística

DEA – Dicionário Eletrônico Aurélio da língua portuguesa

DEH – Dicionário Eletrônico Houaiss da língua portuguesa

DPB – Dicionário Português-Brasileiro

DUP – Dicionário de Usos do Português

MEC – Ministério da Educação

PB – Português Brasileiro

PDBLP – Pequeno Dicionário Brasileiro da Língua Portuguesa

PNLD – Programa Nacional do Livro Didático

QFF – Questionário Fonético-Fonológico

QM – Questões Metalinguísticas

QMS – Questionário Morfossintático

QP – Questões de Pragmática

QSL – Questionário Semântico-Lexical

UFRGS – Universidade Federal do Rio Grande do Sul

SUMÁRIO

1 INTRODUÇÃO	14
CAPÍTULO 1 - TEORIA DO LÉXICO E HISTÓRIA DO PB	17
1.1 Léxico: definição e constituição	17
1.2 Formação sócio-histórica do PB.....	24
CAPÍTULO 2 – LEXICOGRAFIA E METALEXICOGRAFIA	36
2.1 Lexicografia e Metalexicografia: definição e função	37
2.2 Dos Primórdios à Contemporaneidade: um percurso histórico da Lexicografia prática e da Metalexicografia.....	38
2.3 Tipologia lexicográfica: diferentes propostas.....	49
2.3 Macro e microestrutura: aspectos estruturais do dicionário	60
2.4 Regionalismos em dicionários.....	66
CAPÍTULO 3 – A DIALETOLOGIA E O ALiB	71
3.1 Aspectos teórico-metodológicos da Dialectologia e os atlas linguísticos	71
3.2 O Projeto <i>ALiB</i>	75
3.1.1 Perfil dos informantes.....	75
3.1.2 O campo semântico <i>Jogos e diversões infantis</i>	76
CAPÍTULO 4 – METODOLOGIA	78
4.1 O que se utilizou do <i>ALiB</i>	79
4.2 Natureza da pesquisa, <i>corpus</i> e procedimentos	80
CAPÍTULO 5 – DESCRIÇÃO E ANÁLISE DOS DADOS	83
CONCLUSÃO	106
REFERÊNCIAS	108

1 INTRODUÇÃO

O léxico de uma língua reflete a diversidade cultural e social de uma comunidade de fala, pois, assim como mudam e variam os costumes, as relações sociais e os valores humanos, a língua também sofre variações e mudanças. Sendo o léxico um campo aberto e dinâmico, ele se torna foco de interesse de várias ciências linguísticas, das quais destacam-se, aqui, a Lexicografia, a Metalexigrafia e a Dialetoлогия, cada uma delas primando por uma perspectiva distinta.

A Lexicografia se utiliza do léxico para registrá-lo em obras, a fim de sistematizar sua descrição e fornecer materiais de consulta (dicionários, vocabulários, glossários etc.). Nessas obras lexicográficas, o léxico é organizado, classificado e definido. A Metalexigrafia lida com o léxico por intermédio das próprias obras lexicográficas, na medida em que analisa e classifica os dicionários, como também oferece o embasamento científico para a elaboração desses materiais.

A Dialetoлогия, por seu turno, foca a variação do léxico, sobretudo a diatópica ou regional. Devido à diversidade cultural e social, o léxico não se mantém uniforme em toda a comunidade de fala. Por isso, existem palavras que podem ser específicas de determinadas regiões geográficas – os regionalismos. Dessa maneira, a Dialetoлогия busca investigar onde e em quais circunstâncias se utiliza determinadas formas do léxico. Essas investigações da Dialetoлогия podem resultar em atlas linguísticos.

Um dos grandes desafios dos estudos linguísticos seria trabalhar com o léxico valendo-se da interface entre duas ou mais ciências. Uma das possibilidades de trabalho seria aliar a Lexicografia e a Metalexigrafia com a Dialetoлогия. Isso é o que se pretende fazer nesta pesquisa: analisar o léxico em dicionários (Lexicografia e Metalexigrafia) confrontando com dados dialetológicos (Dialetoлогия).

De forma mais precisa, o objetivo geral deste trabalho é investigar como são classificados os regionalismos do campo semântico *Jogos e diversões infantis* no Dicionário Eletrônico Houaiss da língua portuguesa (DEH) e no Dicionário Eletrônico Aurélio da língua portuguesa (DEA). Essa classificação consiste em dizer em que espaço geográfico determinado regionalismo é utilizado. Os objetivos específicos consistem em averiguar quais dos regionalismos desse campo semântico do ALiB são ou não registrados nos dicionários; analisar se a localização geográfica apontada nos dicionários são convergentes ou divergentes aos dados do ALiB; e, quando possível, propor meios de um tratamento mais adequando de regionalismos quando se perceber que há divergência de informações entre os dicionários e o ALiB. Essa

divergência pode ser tanto a localização geográfica do regionalismo, como também a ausência de uma acepção no verbete que remeta a/à brinquedo/brincadeira infantil ou, até mesmo, a falta de registro do regionalismo nos dicionários. A hipótese inicial foi que a maioria dos regionalismos registrados nos dicionários *DEH* e no *DEA* divergem do *ALiB* quanto às marcas diatópicas.

A pesquisa realizada se justifica pela necessidade de dicionarização de regionalismos baseada em dados de pesquisas dialetológicas e geolinguísticas, segundo Isquierdo (2007) e Fajardo (1996-1997). Isso porque as pesquisas dialetológicas trazem dados mais seguros em relação variação diatópica do léxico que podem servir de base para os lexicógrafos. O Atlas Linguístico do Brasil (*ALiB*) é a mais recente publicação que descreve a realidade linguística do Brasil, registrando a diversidade fonética, semântico-lexical e morfossintática em cartas linguísticas. Por isso, ele pode ser considerado um dos instrumentos que podem auxiliar o lexicógrafo no registro e na classificação de um regionalismo quanto à sua localização geográfica. Por isso mesmo, utiliza-se o *ALiB* como parâmetro na análise de dados desta pesquisa.

A dissertação está estruturada em cinco capítulos. No capítulo 1, faz-se uma discussão sobre o léxico e o problema da definição de palavra. Além disso, faz-se uma discussão sobre a formação histórica do português do Brasil, a fim de explicar a origem da diversidade cultural e linguística do país.

No capítulo 2, faz-se uma discussão sobre a Lexicografia e Metalexigrafia. Num primeiro momento, é feito um resgate histórico da Lexicografia geral e de língua portuguesa. Posteriormente, apresentam-se os principais conceitos da Lexicografia, a fim de embasar as discussões sobre os dicionários analisados. Segue-se também uma discussão sobre a história da Metalexigrafia e suas contribuições na análise de dicionários.

O capítulo 3 foi dedicado à discussão dos pressupostos teórico-metodológicos da Dialetologia e da Geolinguística, haja vista que foi tomado, como objeto de análise, além dos dicionários, o *Atlas Linguístico do Brasil*. Compreender em que se baseia a Dialetologia, quais as pesquisas mais relevantes realizadas na história dessa ciência e como se caracteriza sua metodologia elucidará os caminhos de pesquisa e elaboração do *ALiB*.

No capítulo 4, apresenta-se a metodologia utilizada nesta pesquisa. Nessa seção, será possível ver quais regionalismos foram pesquisados, em quais dicionários e como os dados foram organizados. No capítulo 5, esses dados são apresentados e discutidos, mediante o

confronto entre os dicionários e o *ALiB*. Por último, fazem-se as considerações finais da pesquisa.

Acredita-se que os resultados desta pesquisa podem contribuir tanto para a Metalexigrafia, pois acrescenta aos estudos dessa ciência mais investigações sobre obras lexicográficas, como também para a Lexicografia, mostrando quais os principais problemas presentes nos dicionários *DEH* e *DEA* em relação aos regionalismos e que caminhos podem ser seguidos para dirimir essas lacunas. A necessidade de aprimoramento dessas obras lexicográficas faz-se mister, uma vez que, sendo o dicionário o principal instrumento de registro do léxico de uma língua, Lara (*apud* BIDERMAN, 2003, p. 53) afirma que ele “representa a memória coletiva da sociedade e é uma de suas mais importantes instituições simbólicas”.

CAPÍTULO 1 - TEORIA DO LÉXICO E HISTÓRIA DO PB

Neste primeiro capítulo, é feita uma discussão a respeito do conceito de léxico e o problema da definição de palavra. Será destacada a questão da variedade e dinamicidade do léxico de uma língua. Além disso, será discutida a formação do PB e as influências recebidas pelo contato com outras línguas faladas aqui no Brasil na época da colonização. Essa discussão possibilita compreender a origem da diversidade cultural e linguística do país, a qual se reflete nos regionalismos.

1.1 Léxico: definição e constituição

Toda língua é composta por dois elementos essenciais: o léxico e a gramática. O léxico, objeto de estudo de várias ciências da linguagem, como a Lexicologia, a Lexicografia, a Terminologia, a Terminografia, a Toponímia, entre outras, consiste no acervo de palavras¹ que uma língua possui e se caracteriza por ser um campo aberto, sujeito a modificações, uma vez que sempre surgem novas palavras, outras desaparecem ou se modificam. Para Villalva e Silvestre:

o léxico de uma língua é, pois, uma entidade abstrata que se obtém por acumulação: às palavras em uso por cada falante, no seio de uma dada comunidade de falantes, juntam-se as palavras em uso por outras comunidades linguísticas falantes da mesma língua; às palavras em uso na contemporaneidade, somam-se as que estiveram em uso em sincronias passadas, do que temos notícia pela documentação escrita e que, por vezes, ressurgem; aos dados da escrita, unem-se os da oralidade, quando é possível apreendê-la, dada a muito maior fluidez da oralidade face à escrita (VILLALVA; SILVESTRE, 2014, p. 23).

Conforme o trecho acima, observa-se que o léxico se constitui não somente das palavras em uso em um determinado período de tempo em uma comunidade de fala, mas agrega palavras que já foram utilizadas em tempos pretéritos e que ficaram em desuso, chamadas de *arcaísmos*, como também as diferentes palavras utilizadas em comunidades linguísticas distintas. O léxico inclui também *neologismos*, que podem ser novas palavras ou novos significados agregados às palavras que já existem (ALVES, 2011), bem como *empréstimos*, que são palavras estrangeiras que passam para um língua de chegada, assimilando ou não as características fonéticas e gráficas dessa língua. Quando essa assimilação não ocorre, chamam-se *estrangeirismos*,

¹ Entende-se que o conceito de *palavra* é um tanto controverso. Essa discussão será feita mais adiante, ainda nesta seção.

conforme Villalva e Silvestre (2014). A dinamicidade do léxico é também citada por Biderman (2001), para quem:

o léxico é um sistema aberto e em expansão. Incessantemente novas criações são incorporadas ao léxico. Só existe uma possibilidade para o sistema lexical se cristalizar: a morte da língua. Foi o que sucedeu ao latim. Se a língua, porém, continuar a existir como meio de comunicação oral (e também escrito), seu léxico se ampliará sempre. Por essa razão, não se poderá censurar em demasia os lexicógrafos se os seus dicionários não registrarem todos os vocábulos e significados que estão em uso na língua, pois tal obra é praticamente inexecutável (BIDERMAN, 2001, p. 203).

O fragmento acima ajuda a entender que, ao ser tomado como um “sistema aberto e em expansão”, o léxico, em sua totalidade, não pode ser confundido com um dicionário, uma lista de palavras etc., uma vez que se configura como um repertório bem mais amplo e em constante mudança. Além disso, como ressalta a autora, não se pode exigir que um dicionário, por mais exaustivo que seja, inclua todas as palavras de uma língua em seu acervo.

É importante ressaltar também que nem o falante de uma língua domina o léxico totalmente. Cada falante se apropria do léxico de diferentes maneiras, a partir dos estímulos a que são expostos e das experiências linguísticas individuais. Villalva e Silvestre (2014) chamam esse recorte do léxico de *léxico mental*. Considerando a aquisição individual do léxico, é possível afirmar que cada falante possui um *léxico mental* peculiar, uma vez que cada um tem experiências linguísticas distintas. É certo que o indivíduo tem a capacidade de cada vez mais ampliar seu repertório lexical, aprendendo novas palavras, como também, ao mesmo tempo, pode esquecer outras, por conta do desuso. O *léxico mental* de um falante pode ser dividido, segundo Villalva e Silvestre (2014), baseados em Aitchinson (1987), em *léxico passivo* e *léxico ativo*. O *léxico passivo* é mais extenso e é usado para o reconhecimento de enunciados linguísticos. Já o *léxico ativo* é mais reduzido e é utilizado na produção de enunciados.

Devido à diversidade de correntes de descrição lexicais, é difícil estabelecer uma uniformidade na terminologia e nos conceitos. Termos como *palavra*, *vocábulo*, *lexema*, *unidade lexical* etc. são assumidos conforme a corrente teórica a que se filia. No entanto, dentre esses termos, o mais usual é *palavra*.

Ao se conceituar léxico como “acervo de palavras”, surgem, inevitavelmente, dois pontos de discussão: a inclusão ou não dos morfemas neste acervo e o conceito de palavra. Welker (2004) afirma que autores como Alain Rey defendem a exclusão dos morfemas neste acervo, enquanto outros, como Wolfgang Schindler, os incluem. Quanto ao conceito de *palavra*, há também diferentes concepções, dependendo do ponto de vista em que se analisa: se fonológico, morfológico, sintático ou semântico.

As palavras carregam em si uma série de informações fonológicas, morfológicas, sintáticas e semânticas e todas elas interessam para o estudo do léxico. Além disso, a palavra pode representar uma série de formas que são utilizadas em diferentes contextos linguísticos. Por isso, Villalva e Silvestre (2014) conceituam *palavra* da seguinte maneira:

enquanto unidades lexicais, as palavras são entidades abstratas de um ponto de vista semântico, aquele que permite sua associação a um conceito, mas também são abstratas de um ponto de vista fonético, morfológico e sintático: enquanto unidades lexicais, elas estão associadas a diversas possibilidades de realização fonética, de variação morfológica, de função sintática e de significação (VILLALVA; SILVESTRE, 2014, p. 78).

Esse conjunto de possibilidades morfossintáticas e semânticas são representadas por um *lema*. Um *lema* é estabelecido por convenção (WERNER, 1982) e cada língua estabelece seus critérios. O processo de estabelecimento de um *lema* é chamado de *lematização*. Segundo o autor, geralmente se escolhem como *lema* palavras sem elementos gramaticais ou com poucos deles.

No âmbito da Lexicografia, ciência que trata da técnica de elaboração de dicionários, o *lema* é a unidade básica de descrição. Sendo assim, cada verbete é encabeçado por um *lema*. Ele representa as formas concretas de uma palavra. Por exemplo: quando se trata de verbos, a forma escolhida para representar as diversas formas verbais conjugadas é o infinitivo. Sendo assim, as formas ‘falo’, ‘falaríamos’, ‘falando’ são representadas, no dicionário, pela forma verbal ‘falar’. Diferentemente, o latim elege como *lema* a primeira pessoa do singular do presente do indicativo: *falo* (inf. *falare*). A eleição de uma forma para representar outras de mesma natureza evita repetições desnecessárias de verbetes em dicionários.

Ainda quanto à terminologia do elemento básico do léxico, nota-se que Camara Junior (1998) reconhece o termo *palavra*, no entanto, prefere utilizar *vocábulo*, por considerar *vocábulo* uma significação geral e *palavra* um tipo especial de *vocábulo*. Assim, para o autor, normalmente, se considera *vocábulo*, na escrita, o conjunto de letras que fica entre dois espaços em branco, já, na fala, o *vocábulo*, segundo o critério fonológico, seria grupos de sons limitados por pausa na emissão oral. Camara Júnior (1998) afirma que o *vocábulo* escrito nem sempre corresponde ao *vocábulo* falado (ou *vocábulo* fonológico, como ele denomina), pois os espaços em brancos, na escrita, não são indicações para a pausa na emissão de voz.

Para exemplificar tal afirmação, Camara Júnior (1998) cita o caso de sintagmas formados por substantivos e adjetivos, entre os quais não há pausa na pronúncia, como em *campo aberto*, *livro excelente* etc. Nesses casos, há uma mesma emissão oral para dois

vocábulo. Por isso, o autor distingue *vocábulo fonológico*, relacionado à modalidade oral, e o *vocábulo formal*, relacionado à escrita.

Quanto aos vocábulos formais, Camara Junior (2009) classifica-os em *formas livres*, quando consistem em formas autônomas, que funcionam isoladamente como comunicação suficiente (ex. *feliz*); *formas presas*, que só funcionam ligadas a outras formas (*prefixos* e *suffixos*, por exemplo, como em *in + feliz*), e as *formas dependentes*, que também se ligam às formas livres, no entanto, com mais alternativas quanto à forma de se unir à forma livre. É o caso dos pronomes clíticos que, ligados a um verbo, podem vir antes ou depois dele (ex. *fala-se* ou *se fala*).

Biderman (2001), por sua vez, utiliza o termo *palavra*. Segundo a autora, todo falante tem uma consciência intuitiva do que seja palavra e são elas mesmas as primeiras a serem aprendidas no processo de aquisição de linguagem. A autora também afirma que, para a Psicolinguística, nas fases iniciais de aquisição, a criança pronuncia palavras isoladas, as quais podem representar uma sentença inteira. Cita-se o caso da palavra ‘papá’, a qual pode significar ‘quero papá’, ‘tenho fome’ etc. Essa fase é chamada pela Psicolinguística de ‘fala holofrástica’. Nessa fase, a fala da criança realiza três funções de linguagem: a conativa, a expressiva e a referencial, ou seja, a criança estabelece uma comunicação com o outro, expressa seu sentimento em relação a alguma coisa e, em sua comunicação, refere-se às coisas ao seu redor (BIDERMAN, 2001).

Esses argumentos da autora, baseados nos dados da Psicolinguística, dão suporte à sua defesa de que “a palavra é uma entidade psicolinguística primordial, a primeira que articula o discurso humano” (BIDERMAN, 2001, p. 102). Ou seja, a palavra está relacionada a fatores mentais e psicológicos. Prova disso, como afirma Biderman (2001), são as afasias (doenças mentais que afetam os mecanismos de linguagem), as quais dificultam ou impedem a produção de palavras.

Ainda que se reconheça a realidade psicolinguística da palavra, há linguistas que consideram o termo *palavra* pré-científico e o excluem de seus estudos. Biderman (2001) menciona, por exemplo, F. Mikus, o qual defende a total exclusão da noção *palavra*, incluindo as noções de *sintagma lexicalizado* e de *sintagma semi-automatizado*. O autor defende que a língua é um *contínuo sintagmático*, portanto, formada por sintagmas.

Martinet também exclui o conceito de palavra e insere a noção de *monema* e *simonema*. Além dele, L. Bloomfield, Ch. Hockett e Z. Harris também foram radicais na exclusão da noção de palavra. Bloomfield prefere então trabalhar com as noções de *formas livres mínimas* e *formas*

presas. As *formas livres mínimas* são cada uma das unidades que compõe uma sentença. Por exemplo: em “o céu está nublado”, ‘céu’ e ‘nublado’ são exemplos de formas livres mínimas. As formas presas se acoplam às formas livres mínimas. No caso de ‘work’, por exemplo, a forma presa ‘-ing’ se une à primeira para formar ‘working’ (BIDERMAN, 2001).

Biderman (2001) explica que, para Hockett, a palavra está entre o *morfema* e a *forma livre mínima*. Duas palavras podem constituir uma *forma livre mínima*. Citando o exemplo da autora, em “twentyeight”, tem-se duas palavras (‘twenty’ e ‘eight’) e uma forma livre mínima (twentyeight). Diferentemente dessa noção, Harris propõe um estudo em que a unidade mínima de análise seja o morfema e não a palavra. Para ele, o discurso nada mais é do que uma sequência de morfemas e a palavra é irrelevante.

Biderman (2001) também defende a relatividade do conceito de palavra. Para ela, a *palavra* é concebida de diferentes formas em diferentes línguas. Baseando-se na teoria do relativismo linguístico, de Sapir-Worf, cuja principal defesa é a de que cada língua determina seu modo peculiar de ver o mundo, segundo suas categorias gramaticais e léxicas, a autora acredita que o conceito de palavra não pode ser o mesmo para todas as línguas, e que:

qualquer definição mais precisa que tenha validade universal parece-nos prematura no atual estágio de desenvolvimento da Linguística. Por ora, endossamos a teoria whorfiana, considerando que só é possível identificar a unidade léxica, delimitá-la e conceituá-la no interior de cada língua. E insistir também na postura teórica de que, na sua análise, o linguista nunca deve superpor a estrutura léxico-semântica e gramatical da sua língua aos dados de uma outra língua (BIDERMAN, 2001, p. 115).

Fica claro, pois, que, para Biderman, o conceito de palavra não deve ser universal, mas sim particular, levando em conta a estrutura léxico-semântica e gramatical de cada língua. Além disso, a autora defende que é necessário estabelecer critérios de delimitação e de definição de palavra. Para tanto, ela menciona três critérios: o fonológico, o gramatical (morfofossintático) e o semântico.

Segundo o critério fonológico, a palavra é caracterizada como uma “sequência fônica que constituísse uma emissão completa, após a qual a pausa é possível” (BIDERMAN, 2001, p. 137). O falante, ao enunciar seu discurso, emite palavras que são delimitadas por pausas. Dessa forma, cada sequência fonológica entre pausas seria uma palavra. Esse critério, no entanto, não é de todo incontestável. Isso porque, muitas vezes, as palavras não são pronunciadas separadamente, mas por grupos rítmicos. Pensando, por exemplo, no PB, pode-se citar como exemplo o sintagma “os amigos”, o qual pode ser transcrito foneticamente como [wzamisus], em que se observa apenas um grupo fônico, ainda que se trate de duas palavras

(‘os’ e ‘amigos’). Por conta disso, o critério fonológico é impreciso e insuficiente para delimitar uma palavra (BIDERMAN, 2001).

Para assegurar a delimitação de uma palavra, é necessário unir ao critério fonológico o critério gramatical (morfofssintático). Este critério leva em conta as características formais e funcionais da palavra: a qual grupo de palavras uma unidade léxica pertence e qual a função que ela exerce numa sentença (BIDERMAN, 2001).

Para mostrar como o critério fonológico e gramatical podem ser aplicados para delimitação de uma palavra, toma-se o exemplo do latim, citado por Biderman (2001): *Arma virumque cano* (Canto as armas e o varão). Segundo a autora, se se aplica o critério fonológico a esta sentença, chega-se à conclusão de que se trata de três palavras: ‘arma’, ‘virumque’ e ‘cano’. No entanto, ao se fazer uma análise gramatical da sentença, observa-se que em *virumque* há duas palavras unidas: o substantivo ‘virum’ no acusativo singular e a conjunção posposta ‘-que’. Dessa maneira, segundo o critério gramatical, a sentença passa a ter quatro palavras. Essa mesma análise gramatical pode ser aplicada ao português quando se encontram, em sentenças, palavras que são resultados de contrações, como, por exemplo, *no* (*em + a*), *nas* (*em + as*), *do* (*de + o*), *das* (*de + as*) etc.

Biderman (2001) cita ainda duas operações formais que podem ser aplicadas a fim de delimitar uma palavra: a inserção e a permutação. A inserção está relacionada ao princípio de coesão interna da palavra, ou seja, toda palavra “tende a ser internamente estável” (LYONS 1968 *apud* BIDERMAN 2001). Para elucidar o processo de inserção, utiliza-se o exemplo citado pela autora. Observando palavras como *bebericar* e *imperecível*, observa-se que a primeira é formada pela base *beber* e o sufixo *-icar*. Já a segunda é formada pela base *perecer* os afixos *im-* e *-ível*. Aplicando o processo de inserção, nota-se que qualquer morfema que seja inserido entre *beber* e *-icar* ou entre *im-*, *perecer* e *-ível* destruiria a palavra. Da mesma forma ocorre em *guarda-roupa* e *beija-flor*, às quais também não se pode inserir nenhum morfema ou palavra, pois as palavras deixarão de existir. Com isso, observa-se que *bebericar*, *imperecível*, *guarda-roupa* e *beija-flor*, segundo o processo de inserção, são palavras. A autora contrasta os exemplos anteriores com o caso do sintagma *os olhos*. Aqui é possível inserir algum elemento entre *os* e *olhos*, como a palavra *seus*: *os seus olhos*. Nesse caso, em *os olhos*, há duas palavras: *os* e *olhos*.

Quanto ao processo de permutação, nota-se que ele está relacionado ao princípio de permuta da posição das palavras em uma sentença. Tomando, mais uma vez, o exemplo de Biderman (2001), observa-se que a frase *Vi a lua no céu* pode ser organizada de diferentes

formas: 1) *Vi no céu a lua*; 2) *No céu vi a lua*; 3) *No céu a lua vi*; 4) *A lua vi no céu*; 5) *A lua no céu vi*. Apesar da tendência à fixação da posição dos termos na oração, viu-se que há um grau de liberdade na colocação das palavras na oração em português. Diante disso, infere-se que cada elemento permutado corresponde a uma palavra.

Além do critério fonológico e gramatical, Biderman (2001) discute o critério semântico, que também ajuda a delimitar a palavra. Segundo ela, os estudos de Semântica foram deixados de lado nas décadas de cinquenta e sessenta do século XX. Os estudos linguísticos prezavam pela forma e pela estrutura, sem levar em conta a significação. Um dos poucos defensores dos estudos semânticos foi o linguista Ullman, o qual define palavra como “a unidade semântica mínima do discurso”. Nessa perspectiva, a palavra é definida como portadora de um significado autônomo. Essa definição é compartilhada por ciências que estudam a palavra numa perspectiva mais semântica, como a Lexicologia, a Lexicografia e Etimologia (BIDERMAN, 2001). Na defesa da inserção do critério semântico, a autora afirma que:

insistimos na postura teórica dos semanticistas para lembrar que a definição e a delimitação da unidade léxica não podem prescindir do critério semântico. Se a fonologia nos ajuda a reconhecer segmentos coesos fonicamente e se a gramática nos leva a identificar as formas linguísticas manifestas nesses segmentos, só a dimensão semântica nos fornece a chave decisiva para identificar a unidade léxica expressa no discurso. (BIDERMAN, 2001, p. 155).

Tendo em vista os três critérios para definição de delimitação da palavra, observa-se que cada um deles tem seus alcances, mas também suas limitações. Por isso, não se pode defender a supremacia de um em detrimento de outro. Biderman (2001) defende, pois, que, para determinar uma palavra, é necessário operar simultaneamente os três critérios analisados.

Diante da imprecisão do que seja *palavra*, opta-se, geralmente, sobretudo em estudos lexicológicos e lexicográficos, por se trabalhar com os termos *lexema* e *lexia*. O lexema seria a unidade léxica abstrata da língua, como já comentado anteriormente, o qual se manifesta no discurso por meio de formas concretas, as lexias. Por exemplo: a forma *cantar* é o lexema, e as formas *cantei*, *cantas*, *cantaremos* são algumas lexias do lexema *cantar*.

Outro ponto que merece atenção no estudo do léxico é sua variação. Devido à sua expansão e constante mudança, o léxico de uma língua também sofre variações. No caso do português, por exemplo, a variação lexical não ocorre apenas entre as variedades nacionais (português brasileiro, português europeu, português moçambicano etc.), mas também dentro de uma mesma variedade. O PB apresenta variações lexicais que atestam a vitalidade da língua e que também caracterizam as diferenças sociais e culturais de determinadas regiões. Essas

variações lexicais, objeto de estudo deste trabalho, podem ser explicadas pela formação sócio-histórica do PB, como se verá a seguir.

1.2 Formação sócio-histórica do PB

No Brasil, devido à grande extensão territorial, à diversidade cultural, como também ao convívio com línguas indígenas, africanas e europeias – principalmente entre os séculos XVI a XIX –, o português tem se constituído como um conjunto de variedades linguísticas (FARACO, 2008). Segundo o autor:

quando, portanto, dizemos português, este nome não designa um objeto empírico uno, homogêneo, claramente delimitável e objetivamente definível por critérios apenas linguísticos (léxico-gramaticais). O nome singular recobre, de fato, uma realidade plural, ou seja, um conjunto de inúmeras variedades reconhecidas histórica, política e culturalmente como manifestações de uma mesma língua por seus falantes (FARACO, 2008, p. 34).

As diferenças lexicais, morfossintáticas e semânticas do PB têm suas razões históricas, que começam com a colonização do Brasil. Lucchesi (2015) faz essa constituição histórica em seu recente livro *Língua e sociedade partidas: a polarização sociolinguística no Brasil*. Segundo autor, apoiado em Mattos e Silva, no período em que se inicia a efetiva colonização do Brasil, em 1532, até meados de 1695, período da descoberta das minas de ouro no estado de Minas Gerais, o Brasil vivia uma situação de *multilinguismo generalizado* (LUCCHESI, 2015, p. 85). Para ele, esse cenário caracterizava uma polarização linguística bem marcada, de modo que:

[...] opondo a língua da minoritária elite colonial às centenas de línguas indígenas e africanas faladas pela população subjugada e escravizada, além das variedades bastante alteradas da língua portuguesa faladas por esse contingente e seus descendentes, entre os quais se podem incluir variedades pidginizadas ou mesmo crioualizadas do português, como as que hoje se falam em Cabo Verde e São Tomé e Príncipe, na África (LUCCHESI, 2015, p. 85).

Dessa maneira, observa-se que, nesse período, o português era apenas uma das línguas faladas no Brasil e, por muito tempo, representou a língua da minoria da população. A convivência da língua portuguesa com as línguas indígenas e africanas possibilitou influências linguísticas. Segundo Rodrigues (2010), em termos cronológicos, o contato inicial dos portugueses, no Brasil, se deu, primeiramente, com os índios, na década de 30 do século XVI, o que implicou, certamente, influências lexicais no falar português, e, possivelmente, influências fonológicas, morfológicas e sintáticas.

Segundo Fausto (2001), quando os portugueses chegaram ao Brasil, em 1500, encontraram uma população indígena bastante homogênea em termos culturais e linguísticos, divididos em dois grandes grupos: os tupis-guaranis e os tapuias. O grupo tupi-guarani pode ser dividido em dois, segundo o autor: os tupis e os guaranis. Eles habitavam quase toda a costa brasileira, sendo que os tupis ocupavam a região da faixa litorânea do norte até o sul de São Paulo, e os guaranis ocupavam a bacia Paraná-Paraguai e do sul de São Paulo até o extremo sul do Brasil (FAUSTO, 2001).

Os tapuias, segundo Fausto (2001), também se localizavam em alguns pontos do litoral brasileiro. Eles se dividiam em alguns grupos, como os goitacazes, na foz do rio Paraíba, os aimorés, no sul da Bahia e no norte do Espírito Santo, e os tremembés, entre o Ceará e o Maranhão. Sendo assim, os portugueses que chegaram ao Brasil em 1500 encontraram um país habitado por grupos indígenas diversos. Sobre as impressões tidas pelos índios sobre os portugueses, Fausto (2001) comenta que:

a chegada dos portugueses representou para os índios uma verdadeira catástrofe. Vindos de muito longe, com enormes embarcações, os portugueses e em especial os padres foram associados na imaginação dos tupis aos grandes xamãs, que andavam pela terra, de aldeia em aldeia, curando, profetizando e falando de uma terra de abundância. Os brancos eram ao mesmo tempo respeitados, temidos e odiados, como homens dotados de poderes especiais (FAUSTO, 2001, p. 16).

A relação entre brancos e indígenas foi se intensificando, à medida que os colonizadores recorriam aos índios para conquistar seus interesses. Fausto (2001) afirma que, em alguns casos, os portugueses conseguiam aliados indígenas para lutar com outros grupos de índios mais resistentes. Além disso, houve a tentativa dos portugueses de escravizar os índios, mas sem êxito, devido à resistência dos nativos ao trabalho compulsório, ao recebimento de proteção das ordens religiosas contra a escravidão e à baixa densidade demográfica de índios, já que muitos deles foram dizimados por causa de doenças transmitidas pelos colonizadores.

Quanto ao aspecto linguístico, não se sabe ao certo o número de línguas indígenas faladas quando da chegada dos portugueses (FARACO, 2016). No entanto, segundo Rodrigues (2010), havia duas variedades linguísticas de base tupi-guarani: o *tupi* e o *tupinambá*. Segundo o autor, “durante o século XVII, o nome tupi ficou mais associado, sobretudo no sudeste e sul do Brasil, aos índios recrutados pelos bandeirantes para invadir e dominar as missões jesuíticas e outras comunidades de indígenas guaranis na bacia do rio Paraná” (RODRIGUES, 2010, p. 29). Os bandeirantes, filhos de mães tupis e pais portugueses, falavam uma língua diferenciada da língua dos tupis. Já a língua *tupinambá* era utilizada pelos tupinambás do Rio de Janeiro, da

Bahia, do Maranhão e Pará, pelas tribos caetés da Bahia, Sergipe, Alagoas e Pernambuco e pelos índios potiguares da Paraíba e do Rio Grande do Norte, conforme Rodrigues (2010).

O contato linguístico estabelecido pelos portugueses se dava principalmente com as duas variedades mencionadas, segundo Rodrigues (2010). Desse contato, os portugueses passaram a utilizar várias palavras indígenas, sobretudo nomes de plantas, animais, artefatos e conceitos da cultura indígena, rios, riachos etc. Por conta disso, muitas palavras do léxico do PB têm origem indígena. Além disso, a relação de colonização entre portugueses e indígenas implicou outras consequências linguísticas, como afirma Faraco (2016):

essa desestruturação das sociedades originárias e o assentamento das bases de uma nova formação socioeconômica – a sociedade colonial – configuraram um novo quadro de relações sociointeracionais que afetou profundamente as línguas nela envolvidas, desde seu completo desaparecimento (com o extermínio de seus falantes ou com sua diluição na nova formação socioeconômica) até direções evolutivas acionadas pelas inusitadas situações de contato, aloglossia e bilinguismo que aqui se criaram, fazendo, de um lado, emergir as chamadas línguas gerais (paulista e amazônica) e, de outro, traçando as primeiras grandes linhas que resultaram no modo polarizado pelo qual se deu a disseminação da língua portuguesa no Brasil (FARACO, 2016, p. 121).

Segundo Noll (2010), o afastamento dos casais formados por brancos e mulheres indígenas do ambiente indígena genuíno e também o uso de uma língua diferenciada pelos filhos destes casais (os mestiços) fez com que, no século XVII, surgissem as *línguas gerais*, as quais se tornaram o principal meio de comunicação do Brasil até o século XVIII. O uso intenso das *línguas gerais* se dava também devido ao maior contato dos mestiços com seus familiares maternos (indígenas) do que com seus familiares paternos (portugueses), conforme afirma Rodrigues (2010):

aos poucos foi crescendo o número de mestiços, filhos de pais portugueses e mães indígenas, falando a língua das mães e dos demais parentes, todos estes indígenas, já que do lado paterno os avós, tios e primos ficavam em Portugal. Embora, quando crescidos, os rapazes fossem atraídos para as atividades dos pais, nem todos aprenderiam bem a língua destes, mas todos falavam normalmente a de suas mães e demais parentes com quem haviam convivido (RODRIGUES, 2010, p. 37).

Com isso, havia duas *línguas gerais* que eram faladas no Brasil: a *língua geral paulista* e a *língua geral amazônica*. A *língua geral paulista*, segundo Rodrigues (2010), é de origem tupi e era falada pelos mamelucos (mestiços) que integravam as bandeiras que se dirigiam para o interior de São Paulo, Minas Gerais, Goiás e Mato Grosso. Os mamelucos também utilizavam o português. Nesses lugares em que os bandeirantes passavam e criavam povoados, o uso da *língua geral* e do português ocorria simultaneamente.

Ainda, segundo Rodrigues (2010), a *língua geral paulista* durou entre a segunda metade do século XVI até a segunda metade do século XIX e foi extinta devido à progressiva extinção dos indígenas e à crescente chegada de africanos, mestiços de outras regiões e novos colonos. Já a *língua geral amazônica* foi desenvolvida a partir do tupinambá do Pará, a qual se expandiu ao longo do rio Amazonas e prevalece até hoje, mas com o nome *nheengatu*, que significa ‘língua boa’. Sobre o desenvolvimento da *língua geral amazônica*, Rodrigues comenta que:

note-se que a colonização portuguesa da Amazônia desenvolveu-se com grande independência em relação à administração da costa leste e nordeste. No Pará, a colonização portuguesa começou só no século XVII, após a expulsão dos franceses do Maranhão e Grão-Pará, independente da administração do Estado do Brasil. Aí multiplicaram-se rapidamente os mestiços falando a língua dos tupinambás, os quais inicialmente foram chamados de mamelucos, como os do Sul, depois também de *tapuias*, pronúncia portuguesa de um termo de sua própria língua, *tapy'ýia*, na qual significava ‘índios’. Assim que foi distinguindo-se da língua falada as aldeias dos tupinambás, a língua desses mamelucos do norte ou tapuias passou a ser chamada também de “língua geral, expressão que na Amazônia se tornou corrente nos séculos XVII, XVIII, XIX e primeiras décadas do século XX (RODRIGUES, 2010, p. 30-39, grifos do autor).

Estabelecendo uma comparação entre *língua geral paulista* e a *língua geral amazônica*, Rodrigues (2010) afirma que a primeira quase não tem registros em documentos, enquanto que a segunda foi registrada em vários documentos ao longo de quatrocentos anos. Isso porque, na Amazônia, a atividade dos missionários jesuítas era muito forte, e eles sempre escreviam textos na *língua geral*. Em São Paulo, por seu turno, a atividade missionária era quase inexistente e o que se destacava eram as expedições das bandeiras que iam em busca de novos territórios e de mão de obra indígena. Além disso, o autor destaca uma semelhança entre as duas *línguas gerais* quanto às relações de contato:

ambas as línguas gerais, convivendo com a língua portuguesa e sendo faladas em contextos socioculturais bastante diferentes daqueles em que antes eram faladas as línguas tupinambá e tupi, afastaram-se em sua evolução mais ou menos sensivelmente das suas respectivas matizes indígenas. Aliás, o contato linguístico de uma e outra foi não só com o português, mas também, em distintas situações e proporções, com várias línguas indígenas. [...] As duas línguas gerais mantiveram, entretanto, as principais características morfossintáticas da família tupi-guarani e quase todo o léxico original. (RODRIGUES, 2010, p. 40).

Conforme Noll (2010), as *línguas gerais* deixaram de ser utilizadas com a expulsão dos jesuítas e a promulgação do *Diretório dos Índios*. Esse Diretório foi um conjunto de medidas administrativas aplicadas ao Estado do Maranhão e Grão-Pará, publicado em 1757, que, entre

outras coisas, determinava que o ensino das crianças, que era feito em *língua geral*, passasse a ocorrer somente em português. Sobre o *Diretório dos Índios*, Faraco (2016) comenta que:

o *Diretório* [...] foi um ambicioso projeto, obra de gabinete inspirada pela crença iluminista de que com régua e compasso (*i. e.*, com o uso da Razão *in abstracto*) era possível constituir uma outra sociedade, desconsiderando as condições reais da sociedade existente. Nele se definiram as linhas básicas de uma ampla política indigenista (uma espécie de estatuto das populações indígenas, sob o pressuposto de sua emancipação), com diretrizes sociais, culturais e, principalmente, econômicas (FARACO, 2016, p. 97).

A questão do *Diretório dos Índios* estava estritamente ligada ao fato político e administrativo, com o intuito de fortalecer a colonização no Grão-Pará (Maranhão e Pará). Quanto ao fator linguístico, que era apenas um dos pontos mencionados no *Diretório*, a proibição do uso de *línguas gerais* e a valorização do português fazia com que o poder dos missionários fosse enfraquecido, uma vez que estes só pregavam e ensinavam nas línguas nativas. Como afirma Faraco (2016):

afrontar o uso corrente da língua geral e das demais línguas próprias era uma das formas de desmorrionar o domínio dos missionários (em especial, dos jesuítas— eleitos pelas autoridades pombalinas os inimigos maiores das novas políticas), atingindo seu coração simbólico, revertendo o caminho seguido por eles, com apoio da própria Coroa, durante duzentos anos (FARACO, 2016, p. 112).

A despeito da proibição do uso das *línguas gerais* pelo *Diretório dos Índios*, Noll (2010) ressalta que não foi o *Diretório* o principal causador da substituição da *língua geral* pelo português, até mesmo porque ele não obteve sucesso e, por isso, foi revogado anos depois. O que de fato contribuiu para o desuso das *línguas gerais*, segundo o autor, foi o fracasso das atividades missionárias dos jesuítas como também o crescimento da população que falava português nos estados que representavam o centro das atividades econômicas e administrativas (Minas Gerais, Bahia, Pernambuco e Rio de Janeiro).

Faraco (2016) comenta, baseado em Mariani (2014), que a promulgação do *Diretório dos Índios* serviu para propagar a ideia, principalmente entre os grupos dominantes, de que apenas o português era falado no Brasil, dada proibição do uso das línguas indígenas. Essa concepção errônea de monolinguismo no Brasil perdura até hoje. Contudo, as *línguas gerais* continuaram sendo usadas ainda por muito tempo, uma vez que um regulamento não consegue impedir, de imediato, o uso de uma língua. Se por um lado, o imaginário de unidade linguística dominava a classe dominante:

essa rede de sentido, porém, esteve claramente distante da maioria da população, em especial a da Amazônia. Esta, como vimos, levou mais de um

século para ver a língua portuguesa se tornar hegemônica na região – menos pelo disposto no *Diretório* ou pela ação do poder público e pelas grandes mudanças econômicas e demográficas (FARACO, 2016, p. 113).

Diante da discussão a respeito da presença das línguas indígenas no Brasil e o uso das *línguas gerais* inclusive pelos portugueses, pergunta-se: qual seria então o papel do tupi na formação do PB? Dietrich e Noll (2010) argumentam que a maior influência do tupi no português é no campo do léxico e que “não há influência tupi nem na fonética, nem na morfologia do PB” (DIETRICH; NOLL, 2010, p. 85). Apesar disso, os autores expõem as principais influências fonéticas e fonológicas apontadas por outros autores, mas refutam cada uma delas, por considerar que não há “nenhuma prova substancial de mudança linguística condicionada preponderantemente por sistemas alheios” (DIETRICH; NOLL, 2010, p. 86). As mudanças do PB seriam uma tendência imanente da própria língua, como defendem Naro e Scherre (2007), em seu conceito de *deriva*.

Conforme Dietrich e Noll (2010), as contribuições das línguas indígenas no léxico do PB ajudam a diferenciá-lo do português europeu, no qual esses empréstimos são quase desconhecidos. Essas contribuições, sobretudo do tupi, se dá, principalmente, no campo dos nomes da fauna, flora, natureza e cozinha, como também no campo dos nomes próprios – toponímia, hidronímia, coronímia e antroponímia. Os autores também asseveram que esses empréstimos se motivaram devido à necessidade dos portugueses de nomear objetos e realidades que lhe eram, até então, desconhecidas e que eram familiares aos indígenas.

Dietrich e Noll (2010) trazem vários exemplos dos empréstimos indígenas, dos quais se pode destacar apenas alguns: na fauna – tatu, tamanduá, capivara, jararaca, sucuri, surumi, parati, tambaqui, pirarucu, cupim, muquirana, saúva etc.; na flora – capim, cipó, taquara, ipê, jacarandá, jenipapo, babaçu, urucum, guaraná, caju etc.; terrenos e paisagens – caatinga, igapó e igarapé; na cozinha – mandioca, tapioca, abati, pirão, manipueira e tucupi; topônimos – Anhangabaú, Araçui, Baurú, Curitiba, Iguatemi, Ipanema, Piracicaba etc.; hidrônimos – Iguazu, Uruguai, Ibicuí, Parnaíba etc.; antropônimos – Jaraguá, Caipora, Curupi, Jussara etc. Os autores ressaltam que os nomes próprios vieram tardiamente, quando a língua geral já não era falada. Quanto ao registro desses empréstimos em fontes lexicográficas, Dietrich e Noll (2010) afirmam que:

em relação aos empréstimos do tupi, o *Aurélio eletrônico* (v. 5.11) enumera, para o português brasileiro, 2.533 verbetes, e o dicionário Michaelis 3. 238, enquanto os *Dicionários PRO da língua portuguesa* de Portugal apontam apenas 504 palavras. O total de tupinismos é avaliado, no português brasileiro, incluindo os topônimos, em 10 mil. [...] No *Dicionário histórico das palavras*

portuguesas de origem tupi fala de 60 mil topônimos, incluindo, contudo, denominações repetidas (DIETRICH; NOLL, 2010, p. 89).

Rodrigues (2010), por seu turno, defende a influência das línguas indígenas para além do léxico. Segundo o autor, no âmbito da fonética, algumas associações podem ser feitas entre as línguas tupi e tupinambá e algumas variedades do PB. O autor cita o caso das consoantes laterais: dada a inexistência de consoantes laterais nas duas línguas indígenas, é possível que isso justifique a substituição da álveo-palatal λ <lh>pela palatal j ([‘pa λ a] > [‘pa j a], ‘palha’). Ainda quanto à consoante lateral, o autor comenta o caso da alteração dessa forma quando colocada no final da sílaba, mas conservada no início dela: *alta* [‘alta] > [‘arta] ou [‘a λ ta] ou [‘awta], mas *lata* [‘lata], *mala* [‘mala].

Rodrigues (2010) também comenta as influências das línguas tupi e tupinambá na gramática do PB, quanto à flexão de número dos nomes. Dada a inexistência dessa flexão nas línguas indígenas, é possível, segundo o autor, que isso tenha contribuído para alteração da regra de concordância de algumas variedades do PB que flexiona apenas os especificadores, como em *As menina brinca*.

Desse modo, fica clara a forte influência das línguas indígenas na formação do PB. Hoje ainda existem cerca de 170 línguas indígenas, mas com um número pouco representativo de falantes. Ao contrário do que aconteceu no período da colonização do Brasil, “na atualidade, estas línguas não têm influência nenhuma no português, mas, pelo contrário, são contaminadas pelo português e, em muitos casos, pouco a pouco abandonadas pelas nações indígenas assimiladas pelo mundo moderno” (DIETRICH; NOLL, 2010, p. 82).

Além das línguas indígenas, o PB se constituiu também a partir do contato com as línguas africanas trazidas para o Brasil pelos negros escravizados. Lucchesi (2009) afirma que os estudiosos que se dedicam à história da língua portuguesa no Brasil convergem ao afirmar que “o elemento africano desempenhou um papel bem mais relevante no processo de constituição da nossa realidade lingüística atual do que o elemento indígena” (LUCCHESI, 2009, p. 57). Isso porque a população indígena foi diminuindo ao longo do período da colonização, à medida que o genocídio indígena se intensificava, por causa, principalmente, de doenças transmitidas pelos portugueses para os índios, o que fez dizimar grande parte da população nativa. Por outro lado, a população de negros no Brasil só aumentava cada vez mais, haja vista que os colonizadores iam em busca de mão-de-obra para trabalhar nas lavouras de cana-de-açúcar no Nordeste, nas minas de ouro e, por fim, nas lavouras de café. Estima-se que foram trazidos de 4 a 5 milhões de africanos para o Brasil (FARACO, 2016).

Com a resistência indígena para o trabalho escravo, os portugueses começaram a transportar populações africanas, já no século XVI, para o Nordeste brasileiro, onde o cultivo da cana-de-açúcar, do algodão e do fumo estava em voga. O transporte de negros africanos também se justifica pelo fato de ser uma prática muito lucrativa para os portugueses. Segundo Lucchesi (2009), com o alvará de 1549, expedido por D. João III, que autorizava a importação de escravos para o Brasil, o tráfico negreiro se intensificou cada vez mais, principalmente em Pernambuco e Bahia. Consequentemente:

é possível que os primeiros escravos africanos tenham tido contato com a língua geral, mas, com a redução da presença indígena na zona açucareira, os escravos, desde cedo, passaram a ter contato com o português. Os escravos se dividiam, assim, entre os *ladinos*, que tinham alguma proficiência em português, e os *boçais*, que eram incapazes de se comunicar nessa língua. E, se a língua de intercurso entre os colonizadores e escravos era o português, talvez tenha sido possível que, em alguns agrupamentos mais homogêneos, em termos étnicos, os escravos pudessem lançar mão de uma língua franca africana para a comunicação entre eles. O uso da língua geral tupinambá tornava-se, assim, residual (LUCCHESI, 2009, p. 47 – grifos do autor).

Lucchesi (2009) ressalta ainda que os escravos domésticos, os que trabalhavam na casa-grande, falavam o português com mais proficiência, sem deixar de influenciar a língua dos seus senhores. Segundo o autor, apesar de serem advindos de diferentes regiões da África e, portanto, falantes de diferentes línguas, a maioria dos africanos trazidos para o Brasil eram falantes de línguas do subgrupo banto, como afirmam Mattos e Silva e Machado Filho (2009):

sabe-se que, embora tenham chegado ao Brasil entre duzentas e trezentas línguas africanas durante todo o período do tráfico, a grande maioria delas, à exceção do hauçá – que só aportou a partir do século XIX e em número reduzido –, pertencia ao tronco Congo-Cordofaniano, a uma única família linguística, isto é, a Níger-Congo [...] o certo é que o subgrupo banto sempre foi hegemônico no cenário de todas as línguas da África que atravessaram o sal do imenso oceano Atlântico em direção ao Brasil (MATTOS e SILVA; MACHADO FILHO, 2009, p. 299-300).

Os autores afirmam ainda que o contato linguístico que se operou entre as línguas africanas e a língua portuguesa resultou em empréstimos lexicais, até mesmo em Portugal, antes da colonização do Brasil, como também em influências morfossintáticas – estas, por sua vez, mais difíceis de serem atestadas (MATTOS e SILVA; MACHADO FILHO, 2009).

Além do aumento do número de africanos trazidos no Brasil, aumentava também a imigração de portugueses, sobretudo com a descoberta do ouro e de pedras preciosas, em Minas Gerais, no fim do século XVII. Com isso, a língua portuguesa começava avançar no território

brasileiro, uma vez que, antes, as *línguas gerais* eram majoritárias. Como afirma Lucchesi (2009):

a grande onda migratória vinda de Portugal com o ciclo do ouro certamente favoreceu a difusão da língua portuguesa no Brasil, aumentando o acesso dos escravos aos modelos da língua-alvo do segmento dominante e penetrando nas regiões do interior de São Paulo, Minas Gerais, e Goiás, onde antes predominava a língua geral. Portanto, esses processos sociodemográficos podem explicar bem mais o retrocesso da língua geral no século XVIII, do que a tão propalada reforma pombalina e a expulsão dos jesuítas do Brasil, na década de 1750 (LUCCHESI, 2009, p. 48).

Com a produção da riqueza das minas, começaram a surgir os centros urbanos, como Vila Rica e São João Del Rey, em Minas Gerais. Além disso, a cidade do Rio de Janeiro teve um grande processo de urbanização, com a chegada da corte portuguesa, e passou a ser o maior porto de importações de escravos do país e se tornou também a capital da colônia.

Ainda no século XIX, a economia brasileira foi deixando de explorar apenas as minas e passou a cultivar também o café, que, com o crescimento das lavouras, passou a ser o principal produto exportado no Brasil. Com isso, muitos escravos foram trazidos para o sul do país, onde estavam os principais centros produtores do café (LUCCHESI, 2009).

Toda essa movimentação de negros no Brasil para o trabalho escravo trouxe implicações linguísticas. Há algumas hipóteses que tentam explicar o formato que o PB apresenta na contemporaneidade. Segundo Mattos e Silva e Machado Filho (2009), destacam-se o processo de criouliização, a deriva secular – mudanças previstas na própria estrutura da língua – e a transmissão linguística irregular. Lucchesi (2009, p. 70) contesta a hipótese do processo de criouliização apresentando as seguintes razões:

- (i) a proporção entre a população de origem africana e branca, que proporcionou uma maior acesso à língua-alvo do que o observado nas situações típicas de criouliização;
- (ii) a ausência de vida social e familiar entre as populações de escravos, provocada pelas condições sub-humanas de sua exploração, pela alta taxa de mortalidade e pelos sucessivos deslocamentos;
- (iii) o uso de línguas francas africanas como instrumento de interação dos escravos segregados e foragidos;
- (iv) o incentivo à proficiência em português;
- (v) a maior integração social dos escravos urbanos, domésticos e das zonas mineradoras;
- (vi) a miscigenação racial.

A primeira razão apontada pelo autor se justifica pelo fato de que, segundo ele, para que houvesse criouliização seria necessário que a língua-alvo fosse de difícil acesso. Isso ocorreria devido à baixa população do grupo dominante, falantes da língua-alvo, em relação à população do grupo dominado. No entanto, no Brasil, ainda que a população dominada fosse sempre maior

que a do grupo dominante, este grupo sempre consistia em mais de 30% da população do país, conforme Lucchesi (2009).

A segunda razão tem a ver com a forma de vida dos africanos no Brasil. Para que houvesse a criouliização, segundo Lucchesi, seria necessário que o uso da língua-alvo fosse socializado entre o grupo dominado. Diante de situações do dia-a-dia, os falantes do grupo dominado iriam criando novas estruturas linguísticas que, até então, não existiam, para preencher certas lacunas, o que ocorre em processos típicos de criouliização. Contudo, no Brasil, essa socialização não ocorreu, devido à “ausência da vida social e familiar”, uma vez que os escravos eram tratados em condições sub-humanas e, por isso, não tinham condições de construir famílias. Os deslocamentos de negros no Nordeste para a região das minas, no século XVIII, e para as lavouras de café, no século XIX, já comentado anteriormente, também contribuíram para a não socialização da língua-alvo.

Outra razão apontada por Lucchesi para a não-criouliização no Brasil é o uso de línguas francas africanas pelos negros. Para que o processo de criouliização ocorresse, deveriam os escravos utilizar a língua dominante até mesmo nas interações entre si. No entanto, eles utilizavam línguas francas, sobretudo aquelas de base banto (quimbundo ou quicongo), tanto nas senzalas quanto nos quilombos. Segundo Lucchesi (2009), ainda que os colonizadores tentassem manter a heterogeneidade linguística entre os escravos, importando negros de diferentes lugares da África, eles conseguiam se comunicar nas línguas francas. O autor afirma ainda que “com o tempo, essas línguas foram caindo em desuso, sendo mantidas apenas em situações especiais e muito restritas, e substituídas por variedades de português reestruturadas pelos afrodescendentes” (LUCCHESI, 2009, p. 66).

Com a integração de muitos escravos à vida urbana, para trabalhar nos serviços domésticos, o acesso à língua-alvo – o português – se tornou mais fácil, por isso, esses tinham maior proficiência na língua do grupo dominante. Isso também impossibilitou o processo de criouliização no Brasil.

Além disso, o último argumento apontado por Lucchesi (2009), como visto na citação, é a miscigenação racial no Brasil. Segundo o autor, esse foi o fator mais decisivo para que não houvesse representativa criouliização. Para que esse processo ocorresse, seria necessário o isolamento da população negra, sem estabelecer contatos entre os brancos. Com isso, a conservação da cultura africana e a manutenção da língua crioula seriam naturalmente resultantes dessa segregação. No entanto, a miscigenação ocorreu, no Brasil, desde o início da colonização entre os homens brancos e as mulheres índias e negras. No século XIX, os mulatos

(filho de branco com negra) passaram a ser quase metade da população (LUCCHESI, 2009). Além disso, com essa profunda miscigenação, a população mestiça tinha mais possibilidade de ascender socialmente, implicando o aprendizado do português mais culto, o que de fato aconteceu com alguns deles. Destarte, essa realidade social tornou inviável, segundo Lucchesi, o processo de crioulização.

Para Lucchesi (2015), a hipótese que melhor explica as influências das línguas africanas no PB é a *transmissão linguística irregular*, que consiste no surgimento de uma nova variedade de língua a partir do contato entre a língua dos colonizadores e as línguas dos grupos dominados e sua aquisição defectiva pelos grupos envolvidos. Nesse contato linguístico, os grupos dominados são forçados a aprender a língua dos colonizadores, mas esse aprendizado se dá de forma irregular, pois os membros dos grupos dominados, em sua maioria já adultos, têm dificuldade de aprender uma segunda língua. Segundo ele:

nessa situação, a língua do grupo dominante é imposta aos grupos dominados; mas sua aquisição é defectiva, pois ocorre em condições adversas, por indivíduos em sua maioria já adultos. E, por ser imposta, além das limitações decorrentes do fato desses indivíduos adultos não terem mais acesso aos dispositivos que facilitam a aquisição da língua materna, deve-se acrescentar a potencial resistência cultural e ideológica que faz os indivíduos do grupo dominado não almejarem alcançar uma proficiência plena na língua de seus senhores. Dessa forma, o restrito código de comunicação que se estabelece entre dominantes e dominados é formado por um reduzido elenco de itens lexicais da língua do grupo dominante e é praticamente desprovido de qualquer estrutura gramatical regular; sendo denominado jargão ou pré-pidgin. (LUCCHESI, 2015, p. 96).

A nova variedade de língua, chamada de *jargão* ou *pré-pidgin*, tem seu léxico, predominantemente, oriundo da língua do grupo dominante. A estrutura gramatical, por seu turno, se diferencia das línguas anteriores, porque sofrem alterações pelos membros que a utilizam. O que ocorre, na verdade, é uma reestruturação da gramática, que pode ocorrer de duas formas, segundo Lucchesi (2015): 1) pelos mecanismos gramaticais das línguas nativas dos falantes adultos dos grupos dominados; ou 2) pelos dispositivos inatos que atuam no processo de aquisição da língua materna nas crianças que nascem nesse contexto linguístico. Esse processo explica o que aconteceu no Brasil quando do aprendizado da língua portuguesa pelos africanos, que contribuiu para a formação das variedades populares da língua, como afirma o autor:

a aquisição precária do português pelos escravos trazidos da África e pelos índios integrados na sociedade brasileira e a nativização desse modelo defectivo de português como língua segunda nas gerações seguintes de seus descendentes endógamos e mestiços desencadearam um processo de transmissão linguística irregular que teve importantes consequências para a

formação da atual realidade linguística brasileira, nomeadamente para as suas variedades populares (LUCCHESI, 2009, p. 71).

Lucchesi (2015) explica ainda que, quando os adultos desenvolvem a gramática do *jargão* ou *pré-pidgin*, surge uma língua *pidgin*. Já quando as crianças desenvolvem a gramática dessa variedade, no processo de nativização, surge uma língua crioula. O termo *pidgins*, segundo Bortoni-Ricardo (2014):

denota uma língua de emergência desenvolvida para propiciar o contato entre estrangeiros, mais propriamente entre os colonizadores europeus – ingleses, franceses, portugueses, espanhóis e holandeses – e as populações aborígenes nos territórios por ele colonizados [...] (BORTONI-RICARDO, 2014, p. 29).

Nos últimos anos do século XVIII até o início do século XIX, a língua portuguesa, que era minoria entre as línguas existentes, passou a ser a língua mais falada em todo o território. Segundo Faraco (2016):

A hegemonia da língua portuguesa na sociedade brasileira não se deu, portanto, do dia para a noite, nem estava dada já no século XVI. Foi antes o resultado de um processo de longa duração intrinsecamente interligado com fatores socioeconômicos e não como pretenderam, algo ingenuamente, alguns filólogos que atribuíram a vitória do português à sua (suposta) superioridade [...] (FARACO, 2016, p. 142).

O processo de expansão do português pelo território brasileiro se deu por diversos fatores, em diferentes períodos. Lucchesi (2015) comenta que, entre os anos de 1770 e 1800, houve um expressivo crescimento populacional no Brasil, que chega a ser maior do que o crescimento do século XX. Isso se deu pela chegada de colonos portugueses interessados em explorar as minas de ouro e diamante, como também pela importação de mais africanos para realizarem o trabalho.

Com a população portuguesa aumentando, o uso do português também crescia. Além da imigração, houve um contínuo deslocamento de pessoas do Nordeste para o Sudeste com interesses econômicos. Isso também acentuou o uso da língua portuguesa na região em detrimento do uso da *língua geral*. Esses fatos constituem o que Lucchesi (2015) chama de *primeira onda de lusitanização*.

A *segunda onda de colonização* (LUCCHESI, 2015) ocorreu com a chegada da Família Real ao Brasil em 1808, que trouxe consigo muitos membros da elite portuguesa, que se instalaram na cidade do Rio de Janeiro. A consequência da chegada da Família Real foi a crescente urbanização. Outro fator que desencadeou a expansão da língua portuguesa foi a

intensa dizimação dos índios, o que diminuía o uso das línguas indígenas, e o fim da importação de negros africanos, os quais importavam sua língua nativa.

Assim, o português ia se fortalecendo e se tornando a língua hegemônica no país, mas não a única. Conforme Faraco (2016), nos séculos XIX e XX, mais precisamente entre os anos de 1870 a 1920, cerca de 4.000.000 de europeus não-portugueses vieram ao Brasil em busca de prosperidade econômica. Entre eles estão italianos, espanhóis, alemães, japoneses, sírios e libaneses, poloneses e ucranianos.

Todos esses fatos que compõem a sócio-história do Brasil – chegada dos portugueses, colonização do Brasil, tráfico de negros africanos e chegada de imigrantes europeus etc. – contribuíram para a formação do português eminentemente brasileiro, que cada vez mais tomava uma nova forma, diferenciando-se da norma europeia.

As influências linguísticas sofridas ao longo da história contribuíram para que o PB se tornasse um conjunto de variedades, como é considerado hoje. Além disso, essa formação histórica pode justificar as diferenças linguísticas regionais, sobretudo no campo lexical, que constituem o português do Brasil.

CAPÍTULO 2 – LEXICOGRAFIA E METALEXICOGRAFIA

Neste capítulo, serão abordados os principais aspectos que envolvem as ciências que se voltam para a elaboração e para o estudo das obras lexicográficas em geral, sobretudo os dicionários (foco central das discussões): a Lexicografia (ou Lexicografia prática) e a Metalexigrafia (Lexicografia teórica).

Portanto, este capítulo se constitui, inicialmente, pela distinção entre as duas áreas mencionadas; em seguida, pelo regaste histórico da produção lexicográfica desde a Antiguidade até a contemporaneidade, como também dos estudos metalexigráficos, sobretudo no Brasil. O capítulo contempla ainda uma discussão a respeito das principais propostas de tipologia lexicográfica e, por fim, apresenta os conceitos-chave da Lexicografia prática, relacionados à estrutura do dicionário, que baseiam a produção e o estudo dele. Essa discussão se faz necessária, visto que este trabalho analisa dicionários. Sendo assim, faz-se mister conhecer a teoria, a estrutura, a tipologia entre outros aspectos que envolvem as obras lexicográficas.

2.1 Lexicografia e Metalexigrafia: definição e função

Segundo Welker (2004), a Lexicografia “designa a ‘ciência’, ‘técnica’, ‘prática’ ou mesmo ‘arte’ de elaborar dicionários” (WELKER, 2004, p. 11), enquanto que a Metalexigrafia (Lexicografia teórica) envolve “o estudo de problemas ligados à elaboração de dicionários, a crítica de dicionários, a pesquisa da história da lexicografia, a pesquisa do uso de dicionários [...] e ainda a tipologia” (WELKER, 2004, p. 11). Além disso, o autor afirma que o caráter de ciência não deve ser atribuído à Lexicografia prática, mas sim à teórica, conforme se vê no fragmento abaixo:

[...] a *lexicografia prática* não é uma ciência. Ela é uma técnica – e também uma prática – para a qual se precisa de muita ciência (num outro sentido, a saber, “conhecimento atento e aprofundado de alguma coisa”), pois quem elabora, ou compila, um dicionário tem que conhecer não somente fatos linguísticos, principalmente o léxico, como também as maneiras em que esses fatos podem ser apresentados num dicionário. Já na lexicografia teórica, cada vez mais chamada de *metalexigrafia*, estuda-se tudo o que se diz respeito a dicionários. Essa área, sim, pode ser considerada uma ciência (na primeira das acepções citadas). Seus produtos são os conhecimentos adquiridos e divulgados (SEABRA; WELKER, 2011, p. 30-31 – grifos do autor)².

Seabra (2011), por sua vez, reconhece a distinção entre Lexicografia e Metalexigrafia, no entanto, concebe o termo *Lexicografia* como aquele que contempla as duas áreas. Além disso, a autora não considera a Lexicografia como uma técnica ou como uma ciência, mas sim como uma disciplina, conforme se vê na citação a seguir:

contemporaneamente, acredita-se que essas duas vertentes se complementam, e a *lexicografia* é vista como uma disciplina linguística de caráter científico que contempla os aspectos teóricos e práticos da elaboração de um dicionário.

² Ainda que esta citação seja extraída de um texto de autoria de Seabra e Welker, cada um dos autores escrevem em momentos diferentes, uma vez que o artigo, do qual fora tirada a citação, organiza-se em forma de entrevista. Sendo assim, é possível distinguir as ideias de ambos.

A *lexicografia* se insere, portanto, no domínio da linguística aplicada (SEABRA; WELKER, 2011, p. 30)³.

Outra definição para Lexicografia é dada por Barbosa (1991), em que distingue *lexicografia* e *metalexigrafia*, e afirma que há uma relação de interdependência entre as duas áreas. Nas palavras da autora:

a palavra também é objeto de exame da **Lexicografia**, que a toma, no entanto, de outro ângulo, de vez que se define como uma tecnologia de tratamento daquela, de compilação, classificação, análise e processamento, de que resulta, por exemplo, a produção de dicionários, vocabulários técnico-científicos, vocabulários especializados e congêneres. [...] Os discursos lexicográficos são simultaneamente registros de palavras e objeto de estudo da Lexicografia, enquanto investigação fundamental, esta, a seu turno, é objeto da Metalexigrafia, que se define, por sua vez, como epistemologia da ciência lexicográfica (BARBOSA, 1991, p. 184-185 – grifo no original).

Diante das definições propostas pelos autores, é possível estabelecer os limites entre a Lexicografia e a Metalexigrafia, em que a primeira se propõe a elaborar dicionários, valendo-se de conhecimentos de outras áreas, enquanto que a Metalexigrafia se propõe a estudar, analisar e historiar os dicionários já existentes. Dessa forma, esta pesquisa se insere no âmbito da Metalexigrafia, pois analisa o *Dicionário Eletrônico Houaiss da língua portuguesa (DEH)* e o *Dicionário Eletrônico Aurélio da Língua Portuguesa (DEA)*. Na seção seguinte, apresenta-se o percurso histórico dessas duas áreas, mostrando seus avanços e contribuições no intuito de registrar o léxico.

2.2 Dos Primórdios à Contemporaneidade: um percurso histórico da Lexicografia prática e da Metalexigrafia

Neste percurso histórico, dar-se-á ênfase à Lexicografia das línguas clássicas (latim e grego) e, depois, às línguas neolatinas, sobretudo, o português. Sendo assim, far-se-á, primeiramente, um levantamento histórico da Lexicografia prática, iniciando com os povos antigos da Mesopotâmia, perpassando a Antiguidade Clássica, Idade Média, Idade Moderna, na Europa, culminando com a Lexicografia e a Metalexigrafia do português do Brasil até a atualidade. Esse resgate histórico se justifica pela necessidade de se conhecer a historiografia das obras lexicográficas para, então, compreender a configuração atual dos dicionários.

Tudo começa, pois, com o surgimento da escrita, por volta de 3.000 anos a. C. Nesse período, a prática constante era a listagem de palavras para diferentes fins. Essas listagens não

³ Ver nota 2.

são consideradas obras lexicográficas propriamente ditas, as quais só irão surgir na Idade Moderna, contudo, é possível enquadrá-las na “paleolexicografia” (termo cunhado por BKB⁴, 1981 *apud* WELKER, 2004, p. 61), que são produções que se aproximam de obras lexicográficas propriamente ditas, mas que ainda não são.

Welker (2004) destaca os povos sumérios como os primeiros a exercerem a atividade de listagem de palavras, que estava relacionada a uma atividade pedagógica. O autor cita Kramer (1986) para descrever como a listagem de palavras servia de exercício nas escolas dos escribas:

os professores sumérios inventaram um sistema de instrução que consistia sobretudo no estabelecimento de repertórios; quer dizer que eles classificavam as palavras de sua língua em grupos de vocábulos e de expressões ligados uns aos outros pelo sentido e depois mandavam os alunos decorá-los e copiá-los. (KRAMER, 1986 *apud* WELKER, 2004, p. 62)

Logo mais tarde, surgiram, ainda entre os sumérios, as listas bilíngues sumério-acadianos e sumérios-eblaítas, para facilitar o contato linguístico entre os povos, já que, entre eles, havia forte relação, sobretudo política. Outras listas também foram produzidas na língua egípcia e na língua sânscrita.

Na Antiguidade (do séc. VIII a. C ao séc. V d. C.), surgiram, entre os gregos, os glossários da biblioteca de Alexandria, dentre os quais, se pode destacar a obra de Aristófano de Bizâncio, diretor da biblioteca. Entre os romanos, havia o *Appendix Probi*, que consistia numa lista de palavras escritas em latim clássico acompanhada de suas formas variantes em latim vulgar. A lista era prescritiva e, portanto, tencionava impor a língua latina clássica aos falantes do latim vulgar, na tentativa de “impedir” as mudanças naturais dessa variedade. Esse apêndice é hoje uma das fontes para o estudo do latim vulgar. Além dessa lista, cita-se a obra de Varrão (*De lingua latina*) como importante obra lexicográfica da época, porquanto continha informações semânticas e etimológicas de algumas palavras. Outro dicionário latino de grande importância, conforme Auroux (2009), foi o *Libri de significatu verborum*, de Marcus Verrius Flaccus.

Na Idade Média (séc. V ao séc. XV), inicia-se uma intensa elaboração de glossários, enciclopédias e dicionários, que, segundo Biderman (1984), não se tratava de obra lexicográfica propriamente dita. Os glossários, que surgiram já na Grécia Antiga, foram úteis na Idade Média para auxiliar o entendimento de textos em latim e em grego. Inicialmente, as glosas eram

⁴ Abreviação adotada por Welker (2004) para se referir a Boisson, Kirtchuk e Béjoint, autores do texto *Aux origines de la lexicographie: les premiers dictionnaires monolingues et bilíngues* (1991).

escritas às margens do texto ou entre suas linhas. Com o tempo, elas foram organizadas por diferentes critérios (ordem alfabética, domínios temáticos etc.) e passaram a compor um texto autônomo. Entre os glossários deste período, citam-se o de Reichenau (séc. VIII) e o de Cassel (séc. IX).

As enciclopédias também tiveram muita importância na Idade Média. A obra mais representativa desse período é o conjunto de vinte livros que compõe a *Etimologias*, de Santo Isidoro de Servilha (sec. VII). Segundo Verdelho (1995), esses volumes se tornaram obrigatórios em quase todas as bibliotecas europeias.

Ainda no período da Idade Média, foram produzidos alguns dicionários, os quais, para Verdelho (1992, p. 146), seriam “o embrião dos dicionários modernos e lançam as bases da própria lexicografia [...]”. Os principais dicionários desse período, segundo o autor, são *Elementarium* (1050), de Papias; *Derivationes* (fins do século XII), de Osberno de Gloucester; *Liber derivationum* (fins do século XII), de Hugúcio de Pisa, entre outros. Verdelho (1992) destaca que os dicionários medievais tinham problemas quanto à alfabetação, ou seja, a organização das palavras em ordem alfabética era muito escassa, distanciando-se, pois, da principal característica de um dicionário moderno.

Junto a essas obras medievais está também o primeiro dicionário bilíngue latim-português, que, segundo Welker (2004), continha cerca de 3000 verbetes latinos, transcritos no início do século XIV, aos quais foram acrescentados os equivalentes em português, já no século XV. Esse dicionário representa um marco importante na lexicografia de língua portuguesa, porque se trata de uma das primeiras tentativas de registros dessa língua em uma obra lexicográfica.

Uma curiosidade a respeito dos dicionários medievais era o seu difícil acesso. Segundo Verdelho (1992), apesar de haver, nesse período, uma intensa elaboração deles, os alunos quase não podiam consultá-los nas escolas. Além disso, eram materiais de grande porte físico e, por isso, não eram manuseados facilmente. No fragmento abaixo, vê-se como o autor descreve essa situação:

embora aparentemente estivessem ligados à gramática, e alguns deles aparecessem mesmo acompanhados por um manual de gramática, na realidade, os grandes dicionários, ao contrário das gramáticas, não eram livros de acesso fácil, nem eram verdadeiramente manuais escolares, nem tinham certamente utilização escolar, ao nível dos alunos. Eram então sobretudo uma obra de consulta, exclusivamente destinada aos mestres e aos estudiosos. Sabe-se da sua divulgação generalizada pelas escolas e pelos mosteiros, mas só em casos muito raros haveria mais do que um exemplar em cada casa. E esse, deveria ser devidamente protegido e “concatenado”, com ferros bem seguros para impedir o extravio e garantir o seu uso comum. [...] Eram códices

extremamente volumosos, distribuídos por vezes em vários tomos, difíceis de transportar e de manusear. Deveriam inculcar um espaço de leitura bem marcado, entre os rituais do livro medieval (VERDELHO, 1992, p. 147).

Essa restrição diminuiu com o surgimento da imprensa, que permitiu uma melhor divulgação e circulação desses materiais e que, provavelmente, simplificou sua estrutura física.

Na Idade Moderna (séc. XV ao séc. XVIII), surge a Lexicografia propriamente dita (BIDERMAN, 1984). Esse período da história é marcado, principalmente, pela formação dos estados nacionais europeus e pela expansão marítima dessas nações. Nesse contexto político e econômico, fazia-se necessário também a valorização das línguas vernaculares. Até então, a língua de prestígio era o latim e todos os textos eram escritos nessa língua. Com a formação dos estados, cada nação passou a sistematizar sua língua em gramáticas, seguindo os moldes das gramáticas latinas, conforme se vê no fragmento abaixo:

a constituição das nações europeias corresponde a uma profunda transformação das relações sociais (nascimento do capital mercantil, urbanização, mobilidade social, extensão das relações comerciais etc.), inclusive em seus aspectos religiosos (Reforma (1517) e Contra-Reforma). A expansão das nações acarreta indiscutivelmente uma situação de luta entre elas, o que traduz, ao final, por uma concorrência, reforçada porque institucionalizada, entre as línguas. A velha correspondência *uma língua, uma nação*, tomando valor não mais pelo passado mas pelo futuro, adquire novo sentido: as nações transformadas, quando puderam, em Estados, estes vão fazer da aprendizagem e do uso de *uma língua oficial* uma obrigação para os cidadãos. (AUROUX, 2009, p. 53 – grifos do autor).

Para o aprendizado compulsório dessa língua oficial, começam a elaborar os instrumentos linguísticos que permitiram a sistematização desse ensino: as gramáticas e os dicionários. A produção desses instrumentos foi ainda mais reforçada com o surgimento da imprensa, que possibilitou a multiplicação do texto e o aumento de sua difusão. A imprensa é um dos fatores que contribuíram para a *gramatização* (AUROUX, 2009). Por gramatização, Auroux entende que é o “processo que conduz a *descrever* e a *instrumentar* uma língua na base de duas tecnologias, que são ainda hoje os pilares de nosso saber metalinguístico: a gramática e o dicionário” (AUROUX, 2009, p. 65 – grifos nosso).

A produção massiva dos dicionários das línguas vernaculares ainda estava ligada ao latim. Por conta disso, os primeiros dicionários, como se conhecem hoje, eram bilíngues e serviam para o estudo de textos antigos clássicos.

Citam-se, então, como obras dessa época, o dicionário castelhano-latim *Universal Vocabulário*, de Alonso Palencia (1490); o vocabulário *Latino Español e Español Latino*, de Antonio de Nebrija (1507), autor da primeira gramática do espanhol; e, na França, o

Dictionarium latino-gallicum e o *Dictionnaire françois-latin*, de Robert Estienne, publicados em 1539. Biderman (1984, p. 2) afirma que esses dicionários “eram cheios de lacunas e os dicionaristas da época copiavam-se uns dos outros”.

Os dicionários monolíngues só começaram a surgir tempos depois, no final do Renascimento. Eles têm sua origem nos glossários, que foram bastante produzidos durante a Idade Média para o estudo de textos latinos e gregos. A finalidade dos dicionários monolíngues passa a ser, então, predominantemente prescritiva, como afirma Auroux:

o dicionário monolíngue de uso dos nacionais, cujas entradas, pelo viés das definições, ligam-se entre si, é herdeiro desse trabalho lexical, mas corresponde a uma outra finalidade prática que é a mesma da gramatização das línguas nacionais: a normatização dos idiomas. A elaboração estender-se-á por todo o século XVII. (AUROUX, 2009, p. 75).

São exemplos de dicionários monolíngues, no contexto francês e espanhol, o *Tesoro de La Lengua Castellana*, de Covarrubias (1611); o dicionário da Academia Espanhola - *Diccionario de Autoridades* (1739), com sucessivas edições; os dicionários franceses *Richelet* (1680), *Furetière* (1690) e o *Dicionário da Academia Francesa* (1694).

A Lexicografia portuguesa, por sua vez, foi fundada com a obra *Dictionarium ex lusitanico in latinum sermonem* (1562), de Jerônimo Cardoso, em meados do século XVI. Antes disso, como afirma Nunes (2006), havia um dicionário de verbos latim-português produzido na Idade Média. Essa obra foi completada anos depois com novos dois volumes, publicados em períodos diferentes. No entanto, é a obra de Jerônimo Cardoso que ganhou destaque. Verdelho (1992) reconhece a importância do trabalho de Cardoso, afirmando que: “sendo um dicionarista do latim, não podemos deixar de o considerar como o primeiro lexicógrafo do português, e sendo mestre da língua latina, não deixou de contribuir para formar bons cultores da língua portuguesa” (VERDELHO, 1992, p. 20).

Depois do dicionário de Cardoso, surgiram também o *Dictionarium lusitano-latinum* (1611), de Agostinho Barbosa e o *Thesouro da Língua Portuguesa* (1634-1647), de Bento Pereira.

Outro dicionário de grande importância e que, para Biderman (1885, p. 4), “é o melhor dentre os mais antigos dicionários do português”, é o *Vocabulario Portuguez e Latino*, de Rafael Bluteau. É um dicionário bilíngue e de caráter enciclopédico, pois, como atesta Murakawa (*apud* WELKER, 2004, p. 68): “suas definições são extensas e detalhadas, abrangendo não só a explicação do termo e a relação de seus sinônimos, mas também pormenores descritivos e

históricos, alguns, até certo ponto, curiosos”. A obra é composta por 10 volumes e foi produzida entre os anos de 1712 e 1721.

É importante observar que, na história da Lexicografia de língua portuguesa, os primeiros dicionários são bilíngues, mantendo forte relação com o latim. O primeiro dicionário monolíngue surge anos depois e trata-se do *Diccionario da língua portuguesa*, de Antonio de Moraes Silva, autor brasileiro, mas com militância intelectual em Portugal. Esse dicionário teve duas edições: a primeira foi publicada em 1789 e consistia num resumo da obra de Rafael Bluteau, por isso mesmo, o nome de Moraes não ganhou destaque; a segunda foi publicada em 1813 e contém certas reformulações como a retirada de termos arcaicos e acréscimos de outros mais presentes na literatura portuguesa do século XV e XVI. Assim, ao reconhecer o trabalho de Moraes, Biderman afirma que:

o dicionário de Moraes (2. ed.,1813) constitui um marco na lexicografia de língua portuguesa. É o primeiro dicionário de uso da língua, muito avançado para os padrões lexicográficos da época. Apesar de ter-se baseado na obra do Padre Bluteau, sobretudo na primeira edição, na segunda edição Moraes libertou-se de seu modelo, ampliou consideravelmente a obra com respeito ao número de verbetes, incluídos, e mais que isso, apurou o seu trabalho lexicográfico. Omitiu informações de tipo enciclopédico incluídas no Bluteau, revelando consciência de que um dicionário da língua não é uma enciclopédia (BIDERMAN, 1884, p. 5).

Outro dicionário monolíngue surgido nessa época foi o *Elucidário das palavras, termos e frases que em Portugal antigamente se usaram e que hoje regularmente se ignoram*, de Frei Joaquim de Santa Rosa de Viterbo, publicado em 1798. Esse dicionário, como o próprio título sugere, pretendia facilitar a leitura de textos antigos, trazendo o significado de termos obsoletos da língua.

A história da Lexicografia no Brasil, por seu turno, está relacionada ao período da colonização, no século XVI. As primeiras manifestações de lexicografia neste país foram as listas bilíngues (português-tupi) elaboradas pelos jesuítas, a fim de catequizar os índios.

Segundo Nunes (2006), as línguas indígenas estudadas durante o período da colonização eram o tupi, o kariri e o manau, além das *línguas gerais*. Desses estudos, surgiram as seguintes obras: as gramáticas do tupi, de Anchieta e de Figueira; da língua kariri, de Mamiani, e o *Vocabulário na Língua Basílica*, de autoria anônima. Os estudos sobre as línguas indígenas prevaleceram durante os séculos XVI, XVII e XVIII e são eles que fazem surgir os primeiros dicionários bilíngues no Brasil. Entretanto, Nunes (2006) afirma que, antes disso, viajantes e missionários, como Staden, Thevet e Léry, já faziam descrições de palavras indígenas que podem ser consideradas precursoras da Lexicografia no Brasil. Essas obras tinham fins

catequéticos e inseriam muitas palavras que não eram naturais das línguas indígenas, como “padre” e “batizar”.

Ainda no período colonial, surgiram outras obras brasileiras: o *Dicionário Português-Brasileiro* (DPB), publicado em Lisboa, em 1795, cuja autoria é controversa, mas, geralmente atribuída ao frei Onofre, conforme Edelweiss (1969), e os manuscritos do *Dicionário Brasileiro-Português*, de Frei Veloso e do *Vocabulário na língua geral*, de Frei Prazeres do Maranhão, os quais só foram publicados anos depois.

O primeiro dicionário monolíngue de um autor brasileiro seria o já mencionado *Diccionario da língua portuguesa*, de Antonio de Moraes Silva. Apesar de sua obra ser um marco na lexicografia de língua portuguesa, ela não é considerada um marco na Lexicografia brasileira, porque, segundo Orlandi (2013), “mesmo sendo feito por um autor brasileiro, este se coloca na posição autor de dicionário português, publicado em Portugal. No entanto, ele terá seus ecos na elaboração de dicionários brasileiros” (ORLANDI, 2013, p. 114).

Ainda que Moraes tenha acrescentado em sua obra palavras brasileiras, como afirma Nunes (2010), o dicionário ainda estava muito distante da realidade brasileira e não refletia a língua que era falada nesse período. Por causa disso, surgiram outros dicionários monolíngues que serviram de complementos ao dicionário de Moraes, acrescentando novas palavras de origem brasileira e também dicionário de regionalismos. São eles o *Vocabulário Brasileiro para servir de complemento aos dicionários da língua portuguesa*, de Brás da Costa Rubim (1853) e a *Coleção de Vocábulo e Frases usados na Província de São Pedro do Rio Grande do Sul*, de Antonio Coruja (1853). Já no final do século XIX, foram publicados, no Rio de Janeiro, o *Dicionário Brasileiro da Língua Portuguesa*, de Macedo Soares (1888); e o *Dicionário de Vocábulo Brasileiros*, do Visconde de Beaurepaire-Rohan (1889).

No século XIX, em Portugal, estavam sendo publicados o *Grande Diccionario Português* ou *Thesouro da Língua Portuguesa* (1871-74), de Frei Domingos Vieira; e o *Dicionário Contemporâneo da Língua Portuguesa* (1881), de Caldas Aulete. Sobre este dicionário, Verdelho afirma que foi

o primeiro grande dicionário do séc. XIX que se manteve no mercado até a actualidade. Foi na sua maior parte elaborado por António Lopes dos Santos Valente (1839-1896), dando seguimento a um plano de Caldas Aulete (1823-1878) que faleceu quando a redacção do dicionário que ele dirigia, tinha apenas chegado ao final da letra A. [...] O dicionário de Caldas Aulete pode caracterizar-se em relação à dicionarística do seu tempo, por uma importante actualização da nomenclatura lexical da língua portuguesa, por um esforço de rigor na utilização e nas referências das abonações, pela informação etimológica e gramatical e por uma cuidadosa classificação das variedades

diacrônicas, geográficas e estilísticas” (VERDELHO, 1994 *apud* WELKER, 2004, p. 71).

Ainda no contexto português, é importante citar o *Novo Dicionário de Língua Portuguesa*, de Candido de Figueiredo, que teve sua primeira edição publicada em 1899. Biderman (1984) afirma que o autor pretendia contemplar grande parte do léxico do português e critica a estruturação dos verbetes. Nas palavras da autora:

esse dicionário pretendia ser o repositório mais completo do léxico português de todos os tempos bem como de regionalismos portugueses, brasileirismos, e de territórios onde se falava e fala o português. Cândido de Figueiredo publicou cinco edições do seu famoso dicionário que tem qualidades mas tem também terríveis defeitos. É um dicionário rico sobretudo com respeito ao número de palavras incluídas no seu acervo léxico. Há nele muitas palavras raras. A forma do verbete é muito simples. As definições são curtas, às vezes erradas e tolas, especialmente as relativas a termos técnicos ou de procedência técnico-científica. Muitos significados registrados são obsoletos há muito tempo (BIDERMAN, 1984, p. 7).

Voltando ao contexto brasileiro, Nunes (2010) afirma que os primeiros grandes dicionários monolíngues verdadeiramente brasileiros só surgiram no século XX, uma vez que os anteriores serviam apenas de complemento a outros dicionários. Ele cita, então, o *Pequeno Dicionário Brasileiro da Língua Portuguesa* (PDBLP), do qual Aurélio Buarque de Holanda Ferreira participou como colaborador e redator, e o *Grande e Novíssimo Dicionário da Língua Portuguesa*, de Laudelino Freire (1939-44). Essas obras são consideradas algumas das fundadoras da lexicografia verdadeiramente brasileira.

Segundo Müller *et. al.* (2009), o intuito de elaborar um dicionário verdadeiramente brasileiro, que contemplasse apenas *brasileirismos*, foi cultivado também pela Academia Brasileira de Letras (ABL). A incumbência ficou para Antenor Nascentes, que publicou o *Dicionário da Língua Portuguesa* (1961-1967). Esse dicionário, no entanto, não contemplou todo o léxico esperado.

Somente na segunda metade do século XX, os dicionários brasileiros passam a ser mais utilizados do que os dicionários portugueses (NUNES, 2010). Nesse período, em 1975, Aurélio Buarque de Holanda publica uma nova versão de seu dicionário, o *Novo Dicionário da Língua Portuguesa*, aquele que viria a ser um dos dicionários mais populares até os dias de hoje. Ao comentar sobre o dicionário de Aurélio, Biderman (1984) o compara ao dicionário de Cândido Figueiredo, apontando convergências e divergências:

O Aurélio chega a lembrar a configuração do Cândido de Figueiredo (5.a ed.) com relação ao número de entradas e ao tipo de palavras do léxico que ele abriga. Evidentemente distancia-se do Cândido de Figueiredo, relativamente

à seriedade e probidade no tratamento de muitos verbetes, como aqueles que descrevem a terminologia técnico-científica. O Aurélio acolheu muitas palavras raras, um grande número de regionalismos, de vocábulos desusados ou obsoletos, de termos exclusivamente literários, um vasto acervo de termos técnicos e científicos, o que inchou demais a sua nomenclatura (BIDERMAN, 1984, p. 7-8).

A autora também destaca que o dicionário se aproxima a um *thesaurus*, pois contempla muitas palavras do léxico do português. Contudo essa contemplação não é total, uma vez que “nenhum falante por mais competente que seja em matéria vocabular, jamais conseguirá incluir no seu léxico ativo e passivo grandes parcelas do léxico geral da língua” (BIDERMAN, 1984, p. 8). O dicionário de Aurélio passou por reformulações e novas edições foram publicadas; hoje, ele está entre os grandes dicionários de língua portuguesa do Brasil, segundo Welker (2006).

Outros dicionários de grande importância para a Lexicografia contemporânea, citados por Welker (2006), são: o *Michaelis: moderno dicionário da língua portuguesa*, publicado em 1998, contendo mais de 200.000 verbetes; o *Dicionário Houaiss da língua portuguesa*, publicado em 2001, com 228.500 entradas; o *Dicionário de usos do Português do Brasil* (DUP), de Francisco Soares Borba, publicado em 2002, com mais de 62.000 entradas; e o *Dicionário UNESP da língua portuguesa*, também coordenado pelo mesmo autor e publicado em 2004, com mais de 58.000 entradas.

Passando para a história da Metalexigrafia, percebe-se que ela é bem mais recente. Segundo Guerra (2003), a Metalexigrafia moderna surgiu entre os anos de 1960 e 1970. O autor afirma que Franz Josef Hausmann assinala dois momentos-chaves para a constituição da Metalexigrafia na Europa: o primeiro momento gira em torno dos anos 1967 e 1968, na França, com a publicação de *Le vocabulaires français*, de Robert-Léon Wagner, seguido da tese de Bernard Quemada – *Les dictionnaires Du français moderne, 1539-1863. Étude sur leur histoire, leur srypes et leurs méthodes* (1968) – e da *Histoire des dictionnaires français* (1968), de Georges Matoré. O segundo momento está entre os anos 1970 e 1971, em que se publica também na França a *Introduction a La lexicographie: Le dictionnaire* (1971), de Jean Dubois e Claude Debois e o *Étude linguistique et sémiotique des dictionnaires français contemporains* (1971), de Josette Rey-Debove. Esses estudos se ramificam para outros países da Europa.

Na Alemanha, os estudos lexicográficos também tiveram bastante representatividade. Segundo Guerra (2003), um dos nomes de maior destaque é Henert-Ernst Wiegand, o qual tem sido o autêntico impulsor da Metalexigrafia dentro e fora do país. Wiegand se dedicou aos trabalhos metalexigráficos a fim de demarcar o terreno próprio da Metalexigrafia dentro das disciplinas linguísticas. Foi fundador da primeira revista alemã de Metalexigrafia

(*Lexicographica*). O autor ainda destaca o trabalho de um grupo de hispanistas alemães da Universidade de Augsburg, entre cujos membros se encontram figuras de destaque para a teoria e a prática lexicográfica modernas, como Güther Haensch y Reinhold Werner.

Guerra (2003) comenta sobre o surgimento da Metalexigrafia nos Estados Unidos, que se desenvolveu através de uma corrente de estudos históricos e descritivos em torno do dicionário. Diversas correntes de estudos metalexigráficos foram surgindo, ao final dos anos sessenta, e contribuíram para a formação da *Dictionary Society of North America*.

Na Espanha, o interesse teórico pela Lexicografia, em seu início, foi influenciado pela Lexicografia teórica francesa. As investigações de Georges Matoré, Robert-Léon Wagner, Paul Imbs, Bernard Quemada, Jean Dubois e Claude Dubois, Josette Rey-Debove têm sido um constante ponto de referência.

Segundo Guerra (2003), embora a Metalexigrafia na Espanha conte com algumas figuras pioneiras, como é o caso de Julio Casares, com sua *Introducción a la lexicografía moderna* (Madrid, 1950), o verdadeiro precursor da Lexicografia teórica na Espanha foi Julio Fernández Servilha. Em sua obra clássica *Problemas de lexicografía atual* (1974), surgem, pela primeira vez, aspectos considerados hoje centrais para lexicografia teórica: a especificidade da lexicografia ou sua relação com outras matérias afins, a definição do conceito de dicionário, as relações que se estabelecem entre Lexicografia e geografia linguística, o vocabulário científico e técnico nos dicionários e os problemas que suscita a elaboração de uma história da Lexicografia espanhola. Além dele, Manuel Alvar Ezquerro, em seu *Projeto de lexicografía espanhola* (1976), discute todo um programa de renovação para a Lexicografia do espanhol, concentrando-se também, como Fernández-Servilha, na demarcação dos conceitos fundamentais. Conforme Guerra (2003), a inovação está, por um lado, na apresentação das dificuldades que exigiria a elaboração de um *Tesuro de La lengua española*, análogo ao *Trésor de la langue française*; e, por outro, nas propostas concretas que se fazem para poder realizar uma obra de semelhante envergadura.

Os estudos metalexigráficos têm como base o manual de Lexicografia de Günther Haensch, Lothar Wolf, Stefan Ettinger e Reinhold Werner: *La lexicografía: de La lingüística teórica a la lexicografía práctica* (1982), que tem servido para instruir sobre os princípios da teoria da lexicografia.

No Brasil, segundo Welker (2006), os estudos científicos sobre os dicionários (Metalexigrafia) podem ter começado com a fundação da disciplina “Lexicologia e

Lexicografia” na USP, em 1971. Somente em 1980, foram defendidas as primeiras dissertações de mestrado na UFRGS, por Leci B. Barbisan e por Maria da Graça Krieger.

O autor também cita a criação do *Grupo de Trabalho de Lexicologia, Lexicografia e Terminologia* da ANPOLL, em 1988, como impulsionador do crescimento da Metalexigrafia brasileira. Grupos mais recentes, como o *Termilex*, dirigido pela professora Maria da Graça Krieger e vinculado à Unisinos, também contribuem para reflexões sobre a Lexicografia no Brasil.

Recentemente, tem surgido, no Brasil, a Lexicografia e a Metalexigrafia escolares. Autores como Krieger (2006) e Gonçalves (2013) defendem a importância da produção de dicionários para fins didáticos. A produção de dicionários escolares teve seu fortalecimento na última década. Segundo Gonçalves (2013), a partir do ano 2000, o “Ministério da Educação e da Cultura (MEC), por meio do Programa Nacional do Livro Didático (PNLD), deu início à inclusão dos dicionários escolares de Língua Portuguesa às políticas oficiais para materiais didáticos” (GONÇALVES, 2013, p. 178). Em 2001, avaliaram-se os dicionários de 1ª a 4ª série do Ensino Fundamental, e o projeto era que cada aluno recebesse um dicionário. Em 2003, o PNLD/MEC avaliou também dicionários de 5ª e 6ª séries, com a mesma sistemática de distribuição.

Em 2005, como afirma Gonçalves (2013), o modo de avaliação e distribuição dos dicionários foi reformulado: os parâmetros de avaliação foram ampliados e os dicionários eram distribuídos em acervos. O alvo era os dicionários voltados para o Ensino Fundamental completo. Já o PNLD/MEC de 2012 passou a contemplar dicionários para o Ensino Médio.

Gonçalves (2013) cita Gomes (2007) para esclarecer o que seria a Metalexigrafia pedagógica (ou escolar). Para elas, a Metalexigrafia pedagógica consiste em analisar obras lexicográficas escolares existentes com o fim de fazer uma reflexão linguística e metodológica a respeito do dicionário escolar, analisando diferentes aspectos como público-alvo, configuração gráfica, discurso lexicográfico e finalidade pedagógica. O principal intuito desse novo campo do saber seria “refletir sobre o fazer lexicográfico de dicionários escolares, tendo em mente as suas contribuições para o ensino” (GONÇALVES, 2013, p. 177). A autora afirma, no entanto, que essa é uma área relativamente nova e que ainda há muito para se refletir.

Tendo em vista o breve percurso histórico da Lexicografia, é possível notar a importância que essas obras tiveram durante todos os períodos da história e em muitas civilizações. A técnica de elaborar dicionários tem se aperfeiçoado com o tempo, mediante critérios mais rigorosos de seleção e organização de verbetes, dentre outros aspectos. A

Metalexigrafia, com uma história mais recente, tem também contribuído no aperfeiçoamento da produção dos dicionários gerais. E agora, mais recente ainda, a Metalexigrafia escolar pode apontar novas contribuições para o estudo e aprimoramento dos dicionários escolares.

2.3 Tipologia lexicográfica: diferentes propostas

A discussão sobre a tipologia lexicográfica é bastante complexa e, por isso, as propostas nunca são unânimes. Cada autor baseia-se em diferentes critérios para propor uma tipologia. Nesta seção, serão apresentadas as propostas de Haensch (1982), de Biderman (1984a) e de Welker (2004). Conforme Haensch (1982), a dificuldade de se estabelecer uma tipologia de obras lexicográficas⁵ se dá por diferentes razões:

[...] em primeiro lugar, porque tem sido não somente os critérios linguísticos, mas também os fatores históricos e culturais que têm influenciado no nascimento e no desenvolvimento dos diferentes tipos de obras lexicográficas; em segundo lugar, porque as obras lexicográficas existentes apresentam, geralmente, uma combinação de características pertencentes a categorias de classificação totalmente diferentes. Portanto, uma caracterização dos diferentes tipos de obras lexicográficas que se podem distinguir deve partir, primeiro, da história da lexicografia, em segundo lugar, dos trabalhos lexicográficos existentes e, em grau menor, de critérios teórico-linguísticos (HAENSCH, 1982, p. 96).⁶

Dada essa complexidade de classificação de obras lexicográficas, Haensch (1982) propõe uma tipologia, a partir da lexicografia espanhola, segundo dois pontos de vista: da linguística teórica e de critérios histórico-culturais e práticos. Partindo do critério teórico-linguístico, o autor se vale da distinção entre *discurso individual* e *discurso coletivo*, e classifica as obras lexicográficas entre as que representam o discurso individual e aquelas que representam o discurso coletivo.

No primeiro grupo, estão os *glossários*, *dicionários* ou *vocabulários de obras literárias* ou de outros textos. Essas obras lexicográficas têm a lista de palavras restrita aos textos aos quais estão ligados, por isso, representam um discurso individual. Já, no segundo grupo, estão os dicionários do tipo *tesouros*. O dicionário *tesouro* registra o maior número de palavras que

⁵ Utiliza-se o termo ‘obras lexicográficas’ ou ‘instrumentos lexicográficos’ para se referir a todos os tipos de dicionários, vocabulários e glossários (HAENSCH, 1982, p. 105).

⁶ Tradução livre para: [...] en primer lugar, porque han sido no sólo criterios lingüísticos, sino también factores históricos y culturales que han influido en el nacimiento y desarrollo de los distintos tipos de obras lexicográficas; en segundo lugar, porque las obras lexicográficas existentes presentan, por lo general, una combinación de rasgos pertenecientes a categorías de clasificación totalmente diferentes. Por lo tanto, una caracterización de los diferentes tipos de obras lexicográficas que se pueden distinguir de hecho ha de partir, primero, de la historia de la lexicografía, en segundo lugar, de los trabajos lexicográficos existentes y, en grado menor, de criterios teórico-lingüísticos (HAENSCH, 1982, p. 96).

uma língua possui, baseando-se em textos escritos e representativos de uma coletividade linguística. O autor afirma também que uma obra lexicográfica pode registrar um subsistema de uma língua (características específicas de determinadas comunidades de fala), como, por exemplo, os *dicionários de regionalismos* e os *dicionários de jargões*.

Do ponto de vista do receptor linguístico, Haensch (1982) inclui nessa categoria os *dicionários semasiológicos* – aqueles que partem do significante para o significado. Esses significantes, em geral, são organizados na ordem alfabética e, além do significado, são indicadas informações linguísticas como ortografia, pronúncia, separação silábica, classe gramatical etc. Nem sempre os dicionários semasiológicos são formados por unidades lexicais simples, podem também ser compostos de fraseologismos, como os *dicionários de fraseologismos* e *modismos*, que são unidades lexicais complexas. Entre os dicionários semasiológicos estão os dicionários monolíngues e plurilíngues, cuja distinção será feita mais adiante, os dicionários históricos, os dicionários etimológicos, os dicionários de uso etc.

Segundo Haensch (1982), a ordem alfabética é o princípio de organização de grande parte dos dicionários semasiológicos, mas não é a única. Conforme o autor, em línguas que não são representadas com escrita alfabética, a organização da lista de palavras pode ser feita por meio de transcrição fonética, organizadas por grupos de sons.

Além do critério linguístico-teórico, Haensch (1982) estabelece uma classificação a partir de critérios histórico-culturais e práticos, como dito anteriormente. Partindo do critério histórico-cultural, o autor faz um panorama histórico da produção de obras lexicográficas, o que também se fez no tópico 2.1 deste capítulo. Com isso, o autor reforça que as produções lexicográficas se iniciaram com os glossários, em seguida começaram a serem produzidas as enciclopédias e os dicionários bilíngues e, posteriormente, os dicionários monolíngues.

Por fim, o autor estabelece uma tipologia segundo critérios práticos. Esses critérios têm a ver com as características de ordem prática das obras lexicográficas. Como o autor estabelece vários critérios, os quais são vistos a seguir, essa tipologia é a mais extensa e que, talvez, contemple o maior número de tipos de obras lexicográficas existentes. São eles:

- 1) formato e extensão da obra lexicográfica;
- 2) caráter linguístico, enciclopédico;
- 3) sistema linguístico em que se baseia;
- 4) número de línguas;
- 5) seleção do léxico:

Vocabulário geral ou parcial;

- Codificação exhaustiva ou seletiva;
- Cr terios cronol gicos;
- Car ter prescritivo ou descritivo;
- 6) Ordena o de materiais;
- 7) Finalidades espec ficas de dicion rios;
- 8) Dicion rio tradicional (impresso) ou eletr nico.

O crit rio *formato e extens o* est  relacionado ao tamanho do dicion rio, bem como ao n mero de entradas (cada palavra listada). Alguns gloss rios cont m um n mero bem restrito de palavras por estarem ligados a um texto espec fico, j  dicion rios do tipo *tesouro* s o mais volumosos porque possuem um n mero expressivo de entradas. O autor n o apresenta uma classifica o baseada em n meros espec ficos de entradas. Biderman (1984b, p. 27), por sua vez, o faz, estabelecendo que os dicion rios infantis possuem cerca de 5.000 verbetes; os dicion rios escolares podem ter at  30.000 verbetes; o dicion rio padr o tem, em m dia, 50.000 verbetes; e os do tipo *tesouro* possuem entre 100.000 a 500.000 verbetes.

Quanto ao *car ter lingu stico ou enciclop dico*, Haensch (1982) classifica os dicion rios em dois tipos: dicion rio lingu stico e o dicion rio enciclop dico (ou misto). O *dicion rio lingu stico* traz informa es lingu sticas sobre as palavras que a comp e (ortografia, pron ncia, classe gramatical, significado, rela es paradigm ticas e sintagm ticas etc.) mediante a metalinguagem. J  o *dicion rio enciclop dico*, que tamb m pode trazer informa es lingu sticas, sobretudo o significado, cont m tamb m informa es extralingu sticas sobre as palavras, com ilustra es, diagramas, mapas etc. Por essa dupla fun o de conter informa es lingu sticas e extralingu sticas, o dicion rio enciclop dico tamb m pode ser chamado de *dicion rio misto*.

Um outro crit rio pr tico apontado por Haensch (1982) como importante para a classifica o de obras lexicogr ficas   o sistema lingu stico em que se baseia a obra. Esse sistema lingu stico pode ser individual, baseado no autor ou no grupo de autores da obra lexicogr fica, ou coletivo, baseado em *corpus* escrito. As obras lexicogr ficas baseadas em um sistema lingu stico mais restrito s o os vocabul rios e os dicion rios especializados, que trazem termos t cnicos representativos de uma  rea do conhecimento.

Das obras lexicogr ficas baseadas no sistema lingu stico coletivo, o autor cita como exemplo os do tipo *tesouro*. No contexto da lexicografia brasileira atual, podem-se citar os dicion rios *Aur lio* e *Houaiss* como obras dessa categoria, j  que tamb m se baseiam na l ngua

comum aos falantes do português, a partir de *corpus* escrito. Segundo Zavaglia (2012), a escolha das palavras que irão compor a nomenclatura⁷ de um dicionário é uma das principais dificuldades dos lexicógrafos. Por isso, segundo a autora, o critério mais utilizado na lexicografia moderna para a seleção de palavras é a frequência: as palavras mais usuais da língua e, portanto, mais recorrentes em textos escritos são as que podem compor o dicionário. Conforme Haensch (1982), seria conveniente que cada obra lexicográfica informasse em que sistema linguístico está baseado e de que fontes tem se valido para constituir sua nomenclatura.

Quanto ao *número de línguas*, os dicionários podem ser, segundo Haensch (1982), monolíngues e plurilíngues. Estes se subdividem em bilíngues (duas línguas) e multilíngues (mais de duas línguas).

Os dicionários também podem ser classificados a partir da seleção que fazem do léxico. Se for um vocabulário geral, comum a todos os falantes de uma língua, tem-se os *dicionários gerais*, como o Aurélio e o Houaiss, na língua portuguesa, já comentados anteriormente. Em grande parte, esse vocabulário advém da língua culta e, em certos casos, não inclui palavras consideradas inadequadas pela norma. Importante ressaltar que esse vocabulário geral não consiste no léxico total de uma língua, mas uma tentativa de aproximação dessa totalidade.

Quanto ao vocabulário parcial, utilizado apenas em uma dada comunidade linguística – uma comunidade geográfica, um grupo social, uma profissão, um autor ou grupo de autores ou uma área do conhecimento –, é possível registrá-lo em obras com *marcação diatópica*, que indicam os regionalismos, de *marcação diastrática e/ou diafásica*, que são marcações que indicam variações decorrentes do nível social, da idade, do sexo, da profissão, da situação de comunicação etc., além de obras que registram subconjuntos léxicos com *marcação diatécnica*, que apresentam variações decorrentes das diferentes áreas de conhecimento e inclui o vocabulário técnico específico de determinadas profissões ou esferas da vida humana (dicionário técnico ou especializado). Essas obras lexicográficas podem ser tanto monolíngues quanto bilíngues (HAENSCH, 1982).

As obras lexicográficas cuja seleção de vocabulário se dá pela *marcação diaintegrativa* (que indica termos estrangeiros integrados à língua), segundo Haensch (1982), consistem em dicionários de estrangeirismos. O autor comenta que muitos deles têm um caráter purista e, nessas obras, registram-se os estrangeirismos, explicando seu uso (ou “mal uso”) e suas variantes gráfica ou fonéticas. Ele cita o *Dicionário de anglicismos*, de Alfaro R. J. (1964). Por

⁷ Conjunto das entradas de um dicionário, comumente organizadas em ordem alfabética, verticalmente (ZAVAGLIA, 2012).

fim, o último grupo apontado pelo autor dentro do critério de seleção do léxico são as obras lexicográficas que registram vocabulário com *marcação dianormativa* (apresenta variações decorrentes das diferentes normas linguísticas). Esses dicionários se comprometem com o “bom uso da língua”, apresentando formas “erradas” e apontando as “correções”. Em língua portuguesa, há, dentre outros, o *Dicionário de erros e problemas de linguagem*, de Rodrigo de Sá Nogueira (1974).

É importante ressaltar que, além dos dicionários especificamente normativos, como os mencionados acima, os dicionários “comuns”, sem o rótulo de *normativos*, também estão comprometidos com a variedade de prestígio de uma língua. Vale lembrar que os dicionários, ao lado das gramáticas normativas, são instrumentos linguísticos que visam à normatização da língua (AUROUX, 2009). Biderman (1997) comenta a respeito da relação entre dicionário e norma, afirmando que:

[...] o dicionário padrão da língua tenta descrever o léxico do idioma na sua totalidade, de acordo com este modelo ideal da língua culta e escrita, só circunstancialmente referindo-se aos padrões sub-cultos, ou desviantes da norma padrão, tais como os usos dialetais, populares, gíriáticos. A primazia dada à língua escrita pelos dicionaristas explica-se facilmente. A escrita é conservadora e evolui lentamente, sendo socialmente valorizada, o que não ocorre com a oralidade, considerada fala livre, sem compromisso, pouco policiada. [...] Dentro dessa perspectiva, o dicionarista é o porta-voz de sua sociedade, registrando, no dicionário, a linguagem aceita e valorizada em sua comunidade (BIDERMAN, 1997, p. 165).

A partir do fragmento, observa-se a justificativa que a autora dá a respeito da valorização da norma de prestígio em dicionários: é a norma presente em textos escritos e, portanto, a mais valorizada socialmente. As demais normas, chamadas pela autora de “sub-cultas”, acabam ficando em nível secundário nos dicionários, quando há alguma menção a usos populares da língua. Por conta disso, Biderman (1997) segue defendendo a inclusão das normas não prestigiadas em dicionários, como se vê a seguir:

para a sociedade brasileira, qualquer dicionário padrão da língua deve descrever a norma do português brasileiro, mas, no país, coexistem diversas variantes igualmente boas, catalogadas como *normais*, sendo alguns usos ou signos preferidos numa ou noutra região para designar o mesmo referente. Num país continental como o nosso, as diversas comunidades brasileiras vivem a sua identidade em português e no português do Brasil. Além disso, numa sociedade muito diversificada socialmente como a nossa, estratificada em classes sociais, coexistem variedades diastráticas diversas. Embora o dicionário privilegie a língua escrita como foi dito, ele deve descrever também os diferentes níveis da linguagem, os registros sociais e, assim, não só

identificar o vocabulário e os usos marcados como típicos da linguagem culta e formal, mas também o da linguagem coloquial, apontando, os itens lexicais característicos de um uso popular, vulgar, chulo, as gírias e palavras e expressões obscenas (BIDERMAN, 1997, p. 166-167 – grifo da autora).

Ainda sobre a relação entre dicionário e norma, Welker (2004) também fomenta a discussão, baseado em autores como Malkiel (1989), Glatigny (1989) e Ripfel (1989). Em síntese, o autor afirma que existem os dicionários que são claramente normativos/prescritivos, porque explicitam em algum lugar da obra – introdução, prefácio, título etc. –, e outros que têm efeito normativo, em que não há uma menção direta a essa função, mas que acabam, de alguma forma, exercendo-a. É o caso dos dicionários descritivos: mesmo cumprindo o papel de descrever a língua, acabam privilegiando certos usos em detrimento de outros, como também pelo fato de o próprio consulente atribuir essa função aos dicionários. Dessa forma, o autor conclui que:

[...] o dicionário é inerentemente normativo. Porém, têm que ser diferenciado vários tipos de normatividade: (1) a intenção de ser normativo é explicitada; (2) tal intenção é velada; (3) o dicionário quer ser descritivo, mas, sem revelá-lo, privilegia certos usos; (4) o dicionário descritivo escolhe, declaradamente, o registro neutro de uma norma culta (WELKER, 2004, p. 188 – grifo meu).

Retomando os critérios práticos de Haensch (1982) para uma tipologia lexicográfica, passa-se à *codificação exaustiva ou seletiva*. Como já mencionado, o dicionário não consegue registrar a totalidade do léxico de uma língua, mas pode ser exaustivo em sua seleção de palavras quando se baseia em *corpus* escrito. Os *tesouros* da língua são os que cumprem esse papel. Segundo Krieger (2006, p. 144), “este modelo chega alcançar algo em torno de 400.000 verbetes em países de grande tradição lexicográfica como a França, a Espanha e a Inglaterra”. Já os dicionários que fazem uma codificação seletiva, que se limitam a registrar determinado campo do léxico, podem ser os *dicionários de uso*, que selecionam as palavras mais correntes de uma língua, os *dicionários de aprendizagem*, que são utilizados para o aprendizado de uma língua estrangeira, apresentando as palavras mais correntes dessa língua (HAENSCH, 1982).

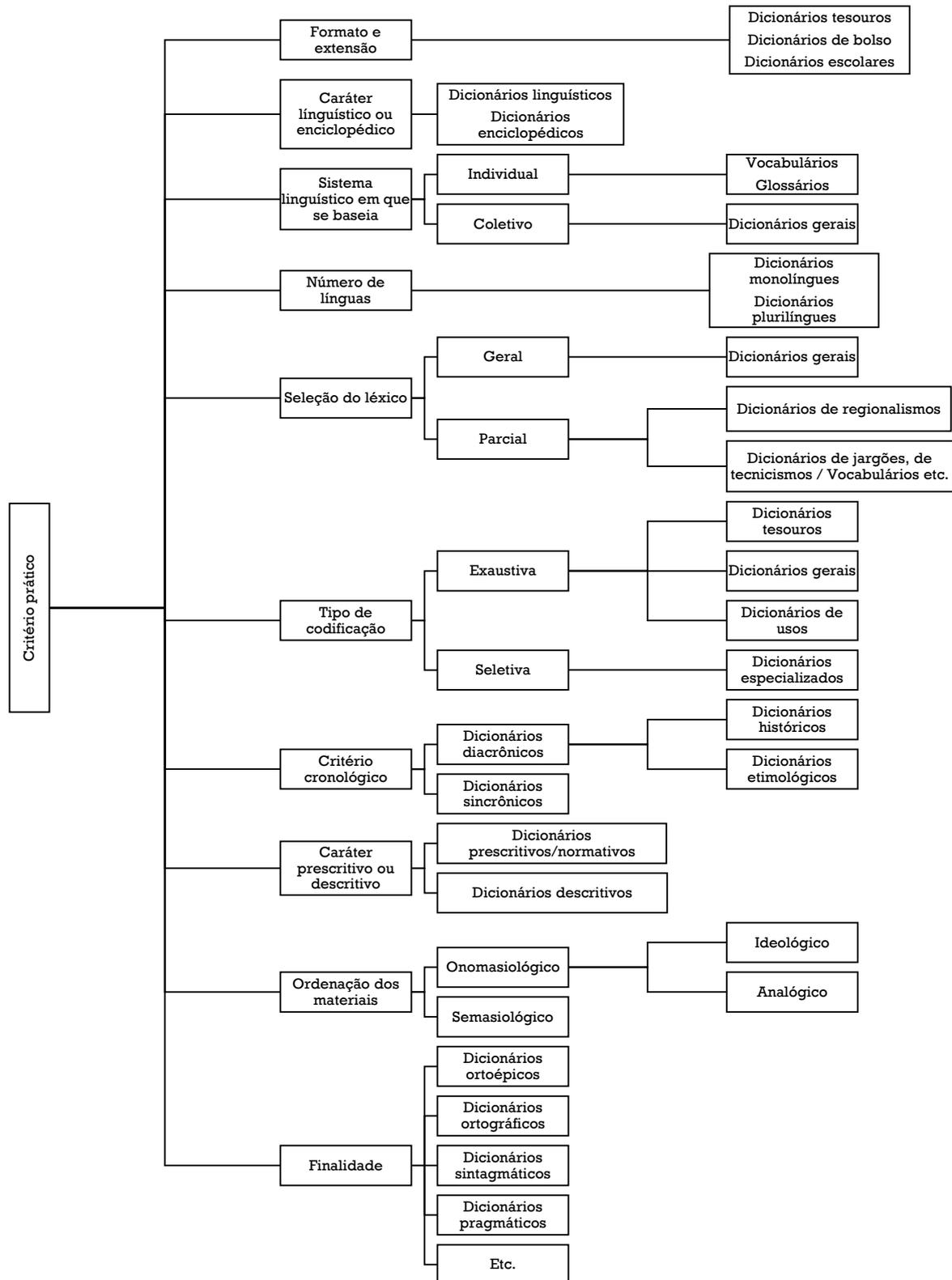
Quanto ao *critério cronológico*, Haensch (1982) distingue os dicionários em *dicionários diacrônicos* e *dicionários sincrônicos*. Os primeiros podem ser divididos em *dicionário histórico*, que estuda a trajetória de uma palavra com as mudanças de forma e significado através do tempo, e o *dicionário etimológico*, que se preocupa em apresentar as origens das palavras. Os segundos registram o vocabulário de uma língua em uma época determinada.

Quanto à *ordenação de materiais*, os dicionários podem ser onomasiológicos e semasiológicos. Os dicionários onomasiológicos ordenam por conceitos, os semasiológicos por significantes. Os primeiros podem ser *analógicos, ideológicos e pictóricos*.

Quanto à *finalidade dos dicionários*, a classificação é bastante exaustiva: dicionários definitórios; dicionários plurilíngues; dicionários de uso; dicionários terminológicos; dicionários dialetais; dicionários de jargões; dicionário de neologismos; dicionário de arcaísmos; dicionários de frequência; dicionários históricos; dicionários etimológicos; dicionário de estrangeirismos; dicionários de abreviaturas; dicionários onomásticos – que explicam a origem dos nomes próprios; dicionários paradigmáticos (de sinônimo, antônimos e parônimos); dicionários ortoépicos – que informam ao consulente a pronúncia correta da palavra; dicionários ortográficos; dicionários sintagmáticos (de regência, colocação, construção, fraseologismos etc.); dicionários gramaticais; dicionários de dúvidas e dificuldades; dicionário convencional/tradicional (impresso) e dicionário eletrônico.

Diante do exposto, observa-se que o critério prático na tipologia lexicográfica de Haensch (1982) é o que mais abrange as obras lexicográficas, sobretudo no contexto espanhol, em que se baseou o autor. Além disso, nota-se a repetição de algumas obras em diferentes categorias, como o próprio autor já advertiu. Diante disso, e, para fim de sintetização e sistematização, elaborou-se um mapa conceitual que ilustra a tipologia proposta por Haensch (1982).

Figura 1: Tipologia lexicográfica segundo Haensch (1982)



Fonte: elaboração própria

A proposta tipológica de Biderman (1984) não apresenta, explicitamente, os critérios que orientam a classificação, como no caso anterior. No entanto, pode-se perceber que a autora

classifica tais obras, sobretudo, mediante a função que elas exercem. O primeiro tipo apontado pela autora é o *dicionário padrão*, definido como:

[...] um instrumento para orientar os seus consulentes sobre os significados e os usos das palavras e para que eles possam expressar suas ideias e sentimentos com maior precisão e propriedade possíveis, utilizado o tesouro léxico que a língua põe à disposição dos falantes do idioma (BIDERMAN, 1984b, p. 27).

O segundo tipo de dicionário apontado pela autora é o *dicionário ideológico* ou *analógico* (tratados como sinônimos). Essas obras são do tipo onomasiológico (partem do conceito para a forma linguística) e organizam os conceitos em campos semânticos. Segundo Biderman (1984a), uma das mais representativas obras dessa categoria é o *Diccionario Ideológico de La Lengua Española*, de Julio Casares (1942). Em português, existe a obra de Francisco Ferreira dos Santos Azevedo, o *Dicionário Analógico da Língua Portuguesa* (1974).

Outro tipo de dicionário apontado por Biderman (1984) é o *dicionário histórico*, que, segundo ela, pode se basear em uma determinada época histórica de uma língua, como também remontar à origem das palavras (dicionários etimológico ou pancrônico). A autora afirma que existem poucas obras desse tipo em língua portuguesa. Ela cita o *Dicionário Etimológico da Língua Portuguesa*, de Antenor Nascentes (1932), o *Dicionário Etimológico da Língua Portuguesa*, de José Pedro Machado (1951) e o *Dicionário Etimológico Nova Fronteira da Língua Portuguesa*, de Antônio Geraldo da Cunha (1982).

Biderman (1984a) também classifica os dicionários de tipo especial, dos quais ela cita vários exemplos: o *Dicionário de Sinônimos e Locuções da Língua Portuguesa*, de Agenor Costa (1954), o *Dicionário de Verbos e Regimes*, de Francisco Fernandes (1968); os dicionários especializados da área de linguística: *Dicionário de Linguística e Gramática*, de Mattoso Camara (1978); dicionários de linguagem popular, entre eles: *Dicionário do palavrão*, Mário Souto Maior (1980) e o *Novo dicionário da gíria brasileira*, de Manuel Viotti (s/d), além de outras obras especializadas de outras áreas do conhecimento. Ao citar essas e outras obras especializadas, Biderman (1984a) chama atenção à necessidade de atualização das obras existentes, como também da criação de outras no contexto português. Segundo ela:

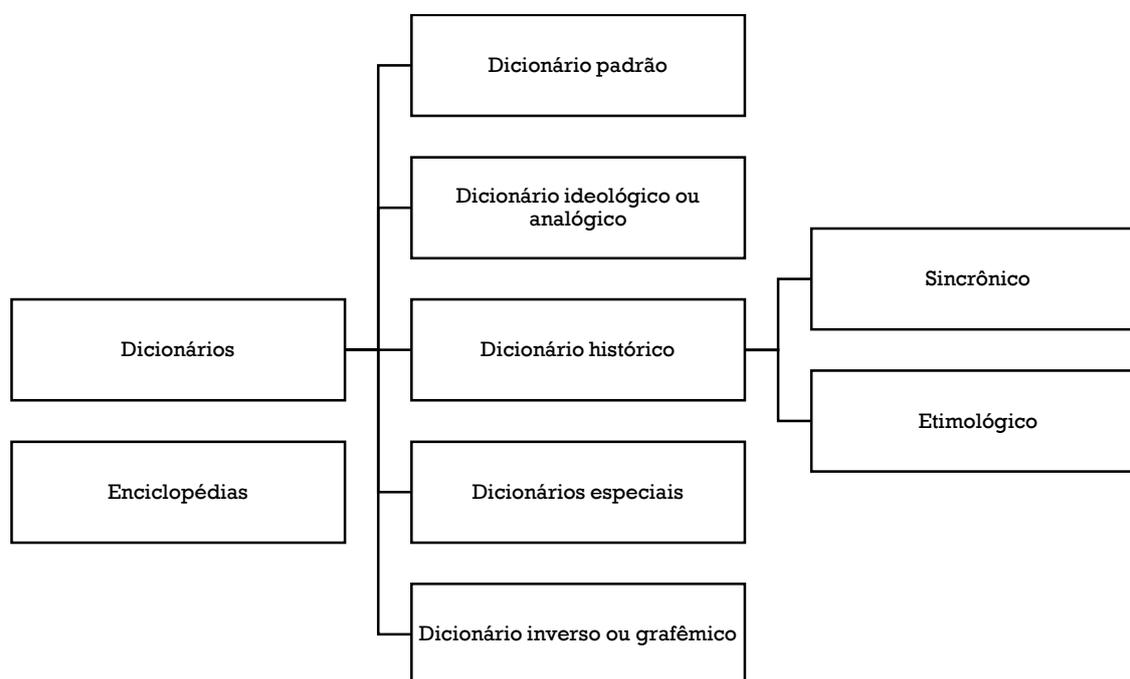
a cultura luso-brasileira precisa refazer muitas dessas obras para atualizá-las e aprimorá-las lexicograficamente; deve também elaborar outros dicionários relativos a áreas do conhecimento que não possuem nenhum dicionário especializado. O poder público deveria criar um órgão para ocupar-se de tão magna tarefa (BIDERMAN, 1984a, p. 16).

Ainda sobre a tipologia proposta por Biderman (1984a), cita-se o *dicionário inverso* (ou grafêmico), que apresenta um repertório de sufixos e “é muito útil para o estudo dos processos

de sufixação e da produtividade léxica de determinados sufixos” (BIDERMAN, 1984a, p. 16). Por fim, a autora também cita as enciclopédias como obras lexicográficas de referência que buscam reunir o máximo de informações linguísticas e extralinguísticas de palavras de várias áreas científicas, as quais podem estar organizadas em ordem alfabética ou por área do conhecimento (BIDERMAN, 1984a).

Em síntese, pode-se apresentar a seguinte figura que ilustra a classificação de obras lexicográficas apontada por Biderman (1984a):

Figura 2: Tipologia lexicográfica segundo Biderman (1984a)



Fonte: elaboração própria

Por fim, passa-se à discussão feita por Welker (2004) sobre tipologia lexicográfica. O autor faz, em sua obra, um apanhado sobre várias propostas tipológicas de outros autores, tais como: Al-Kasimi (1977), o qual também retoma outros autores (Sčerba, Sebeok, Malkiel e Rey) para discutir a temática; Haensch (1982), cuja proposta foi apresentada anteriormente; Hausmann (1985); Martínez de Sousa (1995); Béjoint (2000); Hartmann e James (1998).

Ao detalhar a proposta de cada autor, Welker (2004) também apresenta seu olhar à classificação de obras lexicográficas e defende que ela deve partir, primeiramente, da divisão entre obras de consulta em formato de livro e as computadorizadas, haja vista que, atualmente, os dicionários eletrônicos têm ganhado espaço no conjunto de produções lexicográficas. A segunda distinção seria quanto ao número de línguas, que são os monolíngues e os multilíngues,

dos quais, certamente, os monolíngues são os mais produzidos. E, por fim, a terceira distinção seria entre dicionários gerais e dicionários especiais.

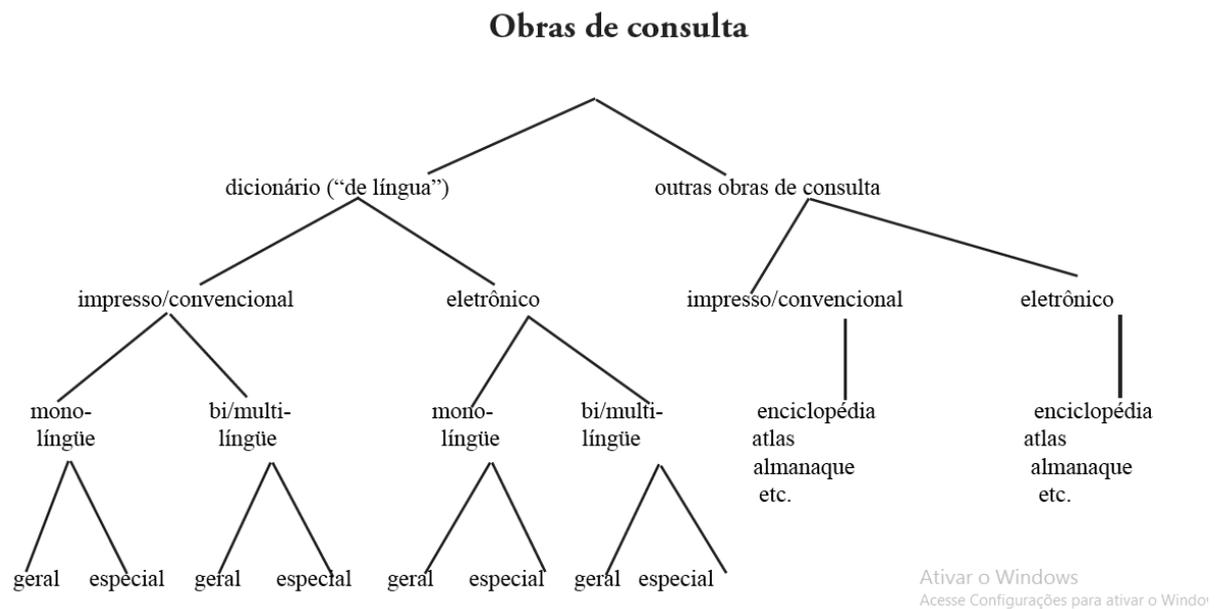
Para Welker (2004), apenas um tipo deve ser classificado como geral – aquele que se caracteriza por ser alfabético, sincrônico, da língua contemporânea, arrolando, sobretudo, os lexemas da língua comum – e os outros como especiais – os históricos, os diacrônicos, os onomasiológicos etc. O autor ainda propõe uma divisão entre os dicionários gerais:

nos gerais, devemos distinguir entre os seletivos, isto é, entre aqueles que registram os lexemas realmente em uso (como o DUP ou Borba 2004) e aqueles muito extensos, às vezes chamados de *tesouros*, que incluem numerosos lexemas e termos não empregados na língua comum, como *Aurélio*, *Michaelis*, e *Houaiss*, que, além de *tesouros*, podemos denominar *gerais extensos*. Embora a definição de *geral* se aplique aos dicionários para aprendizes, estes se destacam por dirigirem-se a um determinado público e por apresentarem certas características que os diferenciam dos “comuns” (WELKER, 2004, p. 43 - grifos do autor).

A distinção estabelecida pelo autor entre dicionário geral e especializado e entre dicionário geral seletivo e extenso, é tomada, neste trabalho, como parâmetro de classificação. Por isso, e conforme o fragmento acima, consideramos os dicionários objeto de estudo desta pesquisa, o *Dicionário eletrônico Houaiss da língua portuguesa* e o *Dicionário eletrônico Aurélio de língua portuguesa*, como dicionários gerais, mais especificamente gerais extensos, por representarem hoje, juntamente com o dicionário *Michaelis*, os grandes *tesouros* da língua portuguesa.

Welker (2004) também organiza sua proposta tipológica em forma de mapa conceitual, o qual é apresentando a seguir:

Figura 3: Tipologia lexicográfica segundo Welker (2004)



Fonte: Welker (2004, p. 44)

Diante do exposto sobre tipologia lexicográfica, viu-se que as propostas de cada autor são distintas, devido aos diferentes critérios que baseiam a classificação. Dessa forma, o ideal seria conciliar tais propostas e observar em que aspectos elas se aproximam ou se distanciam. Ademais, reitera-se que, neste trabalho, assume-se a classificação de dicionário geral proposta por Welker (2004) para designar os dicionários *DEH* e *DEA*, objetos de estudo desta pesquisa.

A seguir, apresenta-se uma exposição sobre a estrutura do dicionário e os principais conceitos da Metalexicografia, a fim de conhecer os termos e os elementos do dicionário que serão utilizados no momento das análises.

2.3 Macro e microestrutura: aspectos estruturais do dicionário

Os dicionários possuem uma estrutura que pode ser dividida em duas: a *macroestrutura* e a *microestrutura*. O termo *macroestrutura* é um tanto polissêmico, porque é utilizado para designar diferentes coisas. Segundo Zavaglia (2012), o termo *macroestrutura* pode ser entendido de três formas:

1. Sinônimo de *nomenclatura*, isto é, o conjunto das entradas de um dicionário, geral e mais comumente organizado em ordem alfabética, submetidas a uma leitura vertical;
2. Todas as partes escritas do dicionário, quer dizer, todo o seu conteúdo, isto é, a nomenclatura acrescida dos textos externos (prefácio, introdução, lista

de abreviaturas e siglas, resumo da gramática, quadro de verbos irregulares, listagem de nomes antropônimos etc.).

3. *nominata*, ou seja, lista completa de palavras definidas no dicionário. (ZAVAGLIA, 2012, p. 243).

Com isso, nota-se que há divergência entre os autores em utilizar o termo. No entanto, a autora opta pelo sentido 1, como sinônimo de *nomenclatura*. Welker (2004) também utiliza o termo de igual modo. O segundo sentido, aquele que define *macroestrutura* como o conjunto de todas as partes do dicionário (elementos pré-textuais e a própria nomenclatura), poderia ser denominada, conforme Zavaglia (2012), como *megaestrutura*. Diante disso, adota-se, neste trabalho, essa distinção entre *macroestrutura* e *megaestrutura*.

Os textos introdutórios dos dicionários, que compõe a *megaestrutura* juntamente com a *nomenclatura*, são muito importantes para o conhecimento da obra lexicográfica, porque, neles, espera-se que o lexicógrafo informe sobre a constituição e caracterização da obra. Conforme Haensch (1982):

a parte introdutória tem que ser muito completa e clara e dar o máximo de informações; como, por exemplo, sobre o tratamento de palavras homônimas e polissêmicas, a lematização de unidades léxicas etc. Para tirar o máximo proveito de um dicionário, é indispensável que o usuário estude a fundo a parte introdutória (HAENSCH, 1982, p. 458)⁸.

A *macroestrutura* consiste na lista de entradas do dicionário. As entradas seriam cada palavra listada em um dicionário, que também pode ser chamada de *entrada do verbete*, *palavra-entrada* ou *lema*. Sendo assim, *lematizar* seria incluir uma palavra como entrada do verbete, o que implica colocá-la na forma de lema (WELKER, 2004).

Compor a *macroestrutura* de um dicionário é organizar a lista de palavras, que, geralmente, se faz em ordem alfabética. Segundo Haensch (1982), a ordem alfabética é a preferida dos dicionários gerais, porque ela permite localizar as palavras com maior rapidez. Mas não é a única forma de organização: as entradas podem também ser organizadas por famílias de palavras. Conforme o autor, essa forma de organização é ideal para o dicionário etimológico, já que permite conhecer, ao mesmo tempo, todas as palavras originadas de um mesmo étimo.

⁸ Tradução livre para: “La parte introductoria tiene que ser muy completa y clara y dar el máximo de informaciones; por ejemplo, también sobre el tratamiento de palabras homónimas y polissêmicas, la lematización de unidades léxicas pluriverbales, etc. Para sacar el máximo provecho de um dicionário, es indispensable que el usuario estudie a fondo la parte introductoria”. (HAENSCH, 1982, p. 458).

Vale ressaltar também que a *macroestrutura* pode ser organizada onomasiologicamente (partindo do significado para o significante) ou semasiologicamente (partindo do significante para o significado).

Outra questão que envolve a *macroestrutura* é a delimitação de sua extensão. Segundo Zavaglia (2012), geralmente, essa é uma decisão que é tomada pelo lexicógrafo ou pela equipe, dependendo dos objetivos da obra e para quem ela vai ser dirigida. Caso seja um dicionário infantil, dicionário especializado ou um dicionário geral, por exemplo, o número de entradas que compõem a *macroestrutura* é diferente. Retoma-se, pois, a classificação feita por Biderman (1984b, p. 27) pelo número de entradas: 1) dicionário infantil e/ou básico – 5.000 verbetes aproximadamente; 2) dicionário escolar ou médio – entre 10.000 a 12.000 verbetes; 3) dicionário padrão – em média 50.000 verbetes; 4) *tesouros* – 100.000, 200.000, 500.000 verbetes.

Além de definir o número de verbetes que irá compor a nomenclatura dos dicionários, é necessário definir o *corpus* que irá fornecer os lemas. Biderman (1994) afirma que, na maioria das vezes, alguns lexicógrafos acabam copiando parte da nomenclatura de dicionários anteriores, sem apontar critérios de escolha. No entanto, os lexicógrafos Borba e Villar (2011) afirmam que a escolha mais segura das entradas de um dicionário deve partir de *corpora* representativos constituídos de textos escritos. Como afirma Borba (2011)⁹, “com um *corpus* controlado por computador, tem-se mais segurança e objetividade na organização tanto da *macroestrutura* quanto da *microestrutura* do dicionário” (BORBA; VILLAR, 2011, p. 21).

Conforme Biderman (1994), para o processo de constituição da nomenclatura de um dicionário, é necessário levar em consideração dois aspectos. O primeiro deles é a necessidade de se basear em uma teoria lexical. Assim afirma: “para selecionar as palavras-entrada ou lemas do dicionário, o dicionarista tem que se fundamentar em um conceito lexicológico de palavra [...]” (BIDERMAN, 1994, p. 26).

O segundo é o critério de seleção das entradas a partir do *corpus*. A autora defende que a metodologia lexicográfica mais confiável para a seleção de palavras da nomenclatura é a frequência de usos na língua, uma vez que “o critério de autoridades literárias não é mais aceitável” (BIDERMAN, 1994, p. 31) e as ocorrências eventuais de determinadas palavras na fala ou na escrita de qualquer falante não deve “merecer registro dicionarístico” (BIDERMAN, 1994, p. 31). O critério *frequência* deve estar aliado a outros critérios de classificação de

⁹ Ainda que esta citação seja referenciada com Borba e Villar (2011), a afirmação é apenas de Borba, uma vez que o texto citado, organizado em forma de entrevista, separa a fala de um e outro autor.

palavras baseados na teoria lexical para que se defina melhor quais palavras devem ou não ser incluídas na macroestrutura, sobretudo, quanto aos casos mais controversos, como os neologismos, derivados morfológicos, terminologismos, regionalismos¹⁰, estrangeirismos, vulgarismos, palavras obscenas e chulas, onomásticos etc., e devem ser explicitados na obra, para que o consulente tenha ciência das concepções do autor da obra.

Quanto à *microestrutura*, é possível afirmar que ela está relacionada ao verbete do dicionário (em espanhol, *artículo lexicográfico*). Segundo Guerra (2003), o verbete é a unidade mínima autônoma em que se organiza o dicionário. Portanto, a *microestrutura* é todo o conjunto de elementos que constitui o verbete. Segundo o autor, os elementos que compõem a microestrutura podem variar, a depender do tipo de dicionário. No entanto, em geral, as principais informações trazidas nos verbetes são: a etimologia, a pronúncia, a ortografia, a categoria gramatical e o número, as restrições de uso, os sinônimos e antônimos, as combinações léxicas em que aparecem os aspectos sintáticos relevantes e a definição com suas diversas acepções (GUERRA, 2003, p. 105).

Zavaglia (2012) concebe a *microestrutura* em três partes: lema, cabeça do verbete e definição. O lema é a *palavra-entrada* do verbete e pode ser chamado também de *entrada*. A *cabeça de verbete*, segundo Zavaglia (2012), são todas as informações que vêm depois do lema e antes da definição. Welker (2004), por seu turno, concebe *cabeça do verbete* como o lema e as informações que a ele se referem e que antecedem a definição. A *cabeça do verbete* pode ser constituída de: *variantes ortográficas*, *pronúncia*, que, segundo Guerra (2003), geralmente só aparece em dicionários bilíngues, mas que também é necessário em dicionários monolíngues, indicação da classe gramatical, indicação etimológica e a indicação das marcas de uso. Nem todos os dicionários seguem a mesma ordem para compor a microestrutura.

A definição, entendida por Biderman (1984b, p. 32) como uma paráfrase da palavra, é também parte fundamental da *microestrutura*. Segundo a autora, ela deve ser redigida em linguagem simples e utilizando palavras correntes na língua. Discutindo sobre a definição de substantivos, especificamente, Biderman (1993, p. 3) ressalta que há seis tipos de definição: *def. sinonímica*, que, segundo ela, deve ser evitada dada a sua pouca precisão, *def. metonímica*, *def. hiperonímica*, que, para ela, seria a definição ideal, *def. enumerativa*, *def. por aproximação* e *def. antonímica*.

A definição pode ser constituída ainda de diferentes acepções. *Acepção* é cada um dos sentidos que uma palavra possui, em diferentes contextos. Quando há mais de uma acepção,

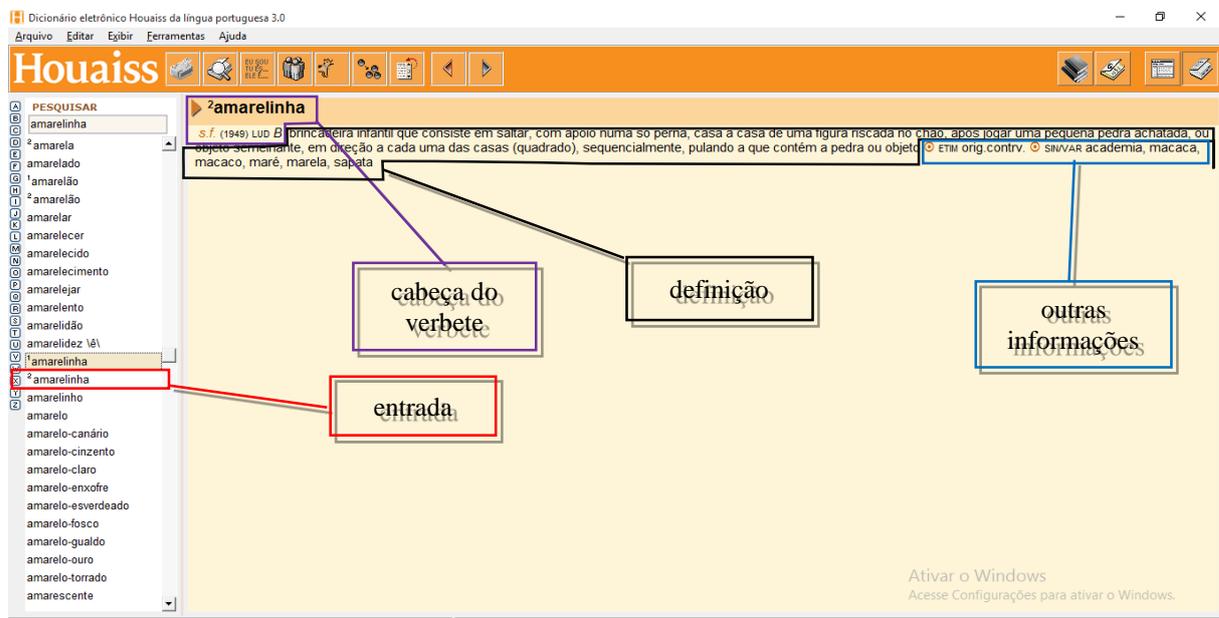
¹⁰ A questão de regionalismos em dicionários será tratada no tópico 2.4.

elas geralmente vêm enumeradas e possuem autonomia, podendo ter suas próprias marcas de uso, exemplos, seus sinônimos, antônimos etc. Segundo Escribano (2003), a ordenação das acepções se dá mediante dois critérios: o sincrônico e o diacrônico. Quanto ao critério sincrônico, as acepções são ordenadas de acordo com a frequência de uso. Já o critério diacrônico leva em conta a etimologia da palavra: começa-se com a acepção mais próxima do étimo.

Em se tratando de dicionários eletrônicos, Escribano (2003) afirma que eles apresentam uma “nova maneira de acesso às informações, já que o lema não é mais a única via de acesso: se pode acessar por abreviaturas, por marcas, por palavras que aparecem na definição, etc.” (ESCRIBANO, 2003, p. 106). Assim, segundo o autor, a microestrutura de dicionários eletrônicos possibilita um melhor acesso às informações.

Tendo em vista que os objetos de estudo desta pesquisa são dois dicionários eletrônicos, apresenta-se, a seguir, a *microestrutura* de cada um deles, para fim de ilustração. Fez-se, nas figuras seguintes, marcações coloridas, a fim de identificar os elementos da microestrutura. Inicia-se com o *DEH*, o qual apresenta a microestrutura de dois modos: no modo tradicional e no modo interativo.

Figura 4: Microestrutura do *DEH* no modo tradicional



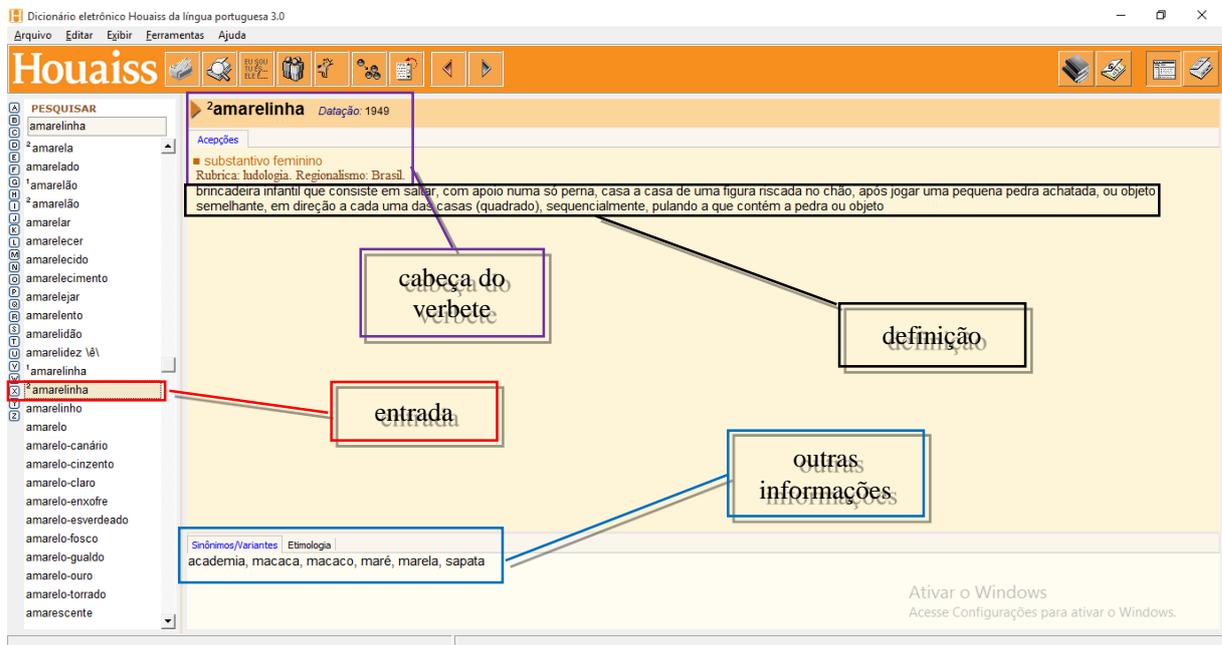
Fonte: dicionário eletrônico Houaiss.

Como se observa, a microestrutura do *DEH* no *modo tradicional*, contém os elementos apresentados anteriormente. Na lateral esquerda, é possível observar parte da *nomenclatura* do

dicionário, referente aos verbetes que iniciam com a letra *a*. Há, na parte superior da nomenclatura, uma ferramenta de busca que facilita o acesso aos verbetes. O verbe selecionado para ilustrar a *microestrutura* foi *amarelinha*. A *cabeça do verbe* é a entrada *amarelinha*, mais a indicação da classe gramatical e do gênero (*sf.*), somada à datação da palavra (1949). Logo em seguida, tem-se a definição, com apenas uma acepção, seguida de outras informações (etimologia, sinônimos e variantes).

A seguir, tem-se esse mesmo verbe no *modo interativo* do *DEH*, o qual apresenta uma configuração um pouco diferenciada.

Figura 5: Microestrutura do *DEH* no modo interativo



Fonte: dicionário eletrônico Houaiss.

Aqui, observa-se a presença da nomenclatura e da ferramenta de busca de verbe na parte superior. A *cabeça do verbe* está presente com indicação da classe gramatical e do gênero da palavra, bem como com a datação. Além disso, outras informações foram acrescentadas à *cabeça do verbe*, como a *rubrica*¹¹ e a indicação de regionalismo. Assim, o dicionário classifica a *rubrica ludologia*¹², indicando que se trata de uma palavra utilizada nessa área de conhecimento, e *Regionalismo: Brasil*, indicando que tal palavra é utilizada no Brasil,

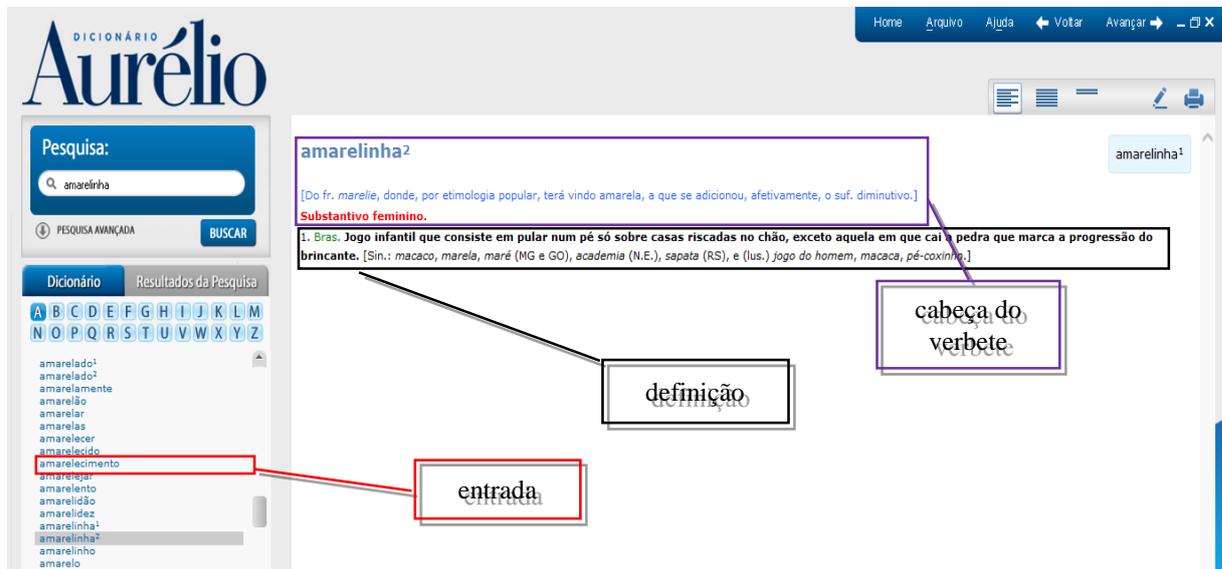
¹¹ *Rubrica* se refere à abreviação que aparece na definição e que indica a área, seja geográfica ou de conhecimento, em que determinada palavra é utilizada.

¹² *Ludologia*: área do conhecimento que se atém a jogos, brinquedos e brincadeiras infantis.

contrapondo o uso europeu. Além disso, na parte inferior, apresenta-se mais informações como *sinônimos/variantes e etimologia*.

Quanto ao *DEA*, a microestrutura é semelhante à do *DEH*, como se vê na figura abaixo:

Figura 6: Microestrutura do *DEA*



Fonte: dicionário eletrônico Aurélio.

Como se nota, o *DEA* contém a nomenclatura na lateral esquerda e a ferramenta de busca na parte superior. O verbete selecionado para análise foi também *amarelinha*. Na *cabeça do verbete* encontram-se a entrada *amarelinha*, a etimologia, a classe de palavra e o gênero. Em seguida, tem-se a definição, que já começa com a indicação de que se trata de um brasileirismo (*Bras.*). E, por fim, vem a indicação de sinônimos e variantes, algumas delas com marcação diatópica.

Com isso, é possível constatar que tanto o *DEH* quanto o *DEA* possuem uma *microestrutura* que facilita a consulta dos verbetes. Além disso, a versão eletrônica possibilita inserir mais informações nos verbetes, inclusive sobre os regionalismos. No tópico seguinte, discute-se a respeito da inserção de regionalismos em dicionários e quais os principais problemas que envolvem essa temática.

2.4 Regionalismos em dicionários

Uma das questões mais controversas que envolvem dicionários é o tratamento dado a regionalismos. Eles podem compor uma obra específica, em um dicionário de regionalismos,

ou um dicionário geral. Seja qual for o tipo de obra, o lexicógrafo precisa levar em conta algumas questões quando da classificação e da inserção de regionalismos em dicionários, como assevera Biderman (1994):

critério ambíguo e difuso é o de regionalismos. Os dicionários são lacônicos e até contraditórios no tratamento dessa matéria. Alguns pontos para serem considerados: qual o ponto de referência para definir um termo como regional? Sendo regional relativamente a um dialeto padrão, qual é esse dialeto padrão? (BIDERMAN, 1994, p. 33).

A discussão sobre o conceito de regionalismos também é bastante complexa. Geralmente, o termo *regionalismo* pode ser entendido como os fatos linguísticos específicos de uma determinada região geográfica. Contudo, alguns autores definem o termo, partindo de algum ponto de referência, como é o caso do linguista Camara Júnior (2011):

Regionalismos – Em sentido lato, traços linguísticos privativos de cada uma das regiões em que se fala uma dada língua assim dividida em dialetos (v.). Em sentido estrito, os regionalismos léxicos, especialmente quando recebem guarida na língua escrita e literária, são de forma (vocabulares) ou de significação (semânticos). O *brasileirismo* (v.), como o *lusitanismo* (v.), corresponde a regionalismos em face do léxico comum português; mas para a língua escrita e literária do Brasil há, por sua vez, regionalismos amazônicos, nordestinos, baianos, fluminenses, mineiros, sulistas, bem como em Portugal para a sua língua escrita e literária há regionalismos interamnenses, transmontanos, beirões, estremenhos, alentejanos, algarvios (CAMARA JÚNIOR, 2011, p. 258).

Observa-se que o autor estabelece o conceito em dois sentidos: sentido lato e sentido estrito. No primeiro, reconhece que são os traços linguísticos das várias regiões que falam determinada língua. No segundo, pontua os regionalismos lexicais e toma como parâmetro a língua escrita e literária.

Isquerdo (2007) também propõe um conceito para regionalismo, partindo da distinção entre vocabulário comum de caráter geral e de cunho regional. Segundo ela:

o primeiro entendido aqui como o vocabulário atualizado em praticamente todo o território onde a língua é usada como instrumento de comunicação e interação, por falantes de diferentes classes sociais e de atividades profissionais diversificadas, enfim, o vocabulário fundamental usado na comunicação cotidiana. Já o vocabulário regional remete à questão dos regionalismos, ou seja, a unidades lexicais, cujo uso é restrito a determinadas regiões, muitas delas relacionadas a atividades específicas do meio rural. (ISQUERDO, 2007, p. 198).

A autora também concebe regionalismos em dois níveis (gerais e específicos), estabelecendo para cada um deles um ponto de referência. Nas palavras da autora:

[...] os gerais, entendidos como aqueles fatos linguísticos no nível lexical, atualizados em todo o território nacional que se opõe à variante europeia do português, e os específicos, ou seja, aqueles que são peculiares a cada estado/região do Brasil. No primeiro caso, o ponto de referência é a norma europeia e, no segundo, a variante brasileira tomada como padrão. Neste caso, classificam-se os regionalismos com difusão em um ou mais estados da Federação (regionalismos do Paraná, do Amazonas, da Bahia, de Goiás...) ou nas diferentes regiões brasileiras (regionalismos do Norte, do Sul, do Centro-Oeste, do Sudeste ou do Nordeste). Esse tem sido o critério adotado pelos maiores dicionários gerais da variante brasileira da língua portuguesa: *Dicionário Aurélio* e *Dicionário Houaiss* (ISQUERDO, 2007, p. 200).

Os regionalismos gerais correspondem aos chamados *brasileirismos*. O conceito de *brasileirismo* é também complexo e tem trazido à baila várias discussões. Isquerdo (2006), ao discutir *brasileirismo*, cita o trabalho de Oliveira (1999), chamando atenção ao fato de que as unidades lexicais classificadas como *brasileirismos* no dicionário Aurélio não deveriam ser classificadas dessa forma, uma vez são utilizadas em estados/regiões específicas, e não em todo o território brasileiro, o que leva a entender que *brasileirismo* pode se caracterizar também por sua difusão em todas as regiões do país. Neste trabalho, utiliza-se este critério para classificar uma variante como *brasileirismo*. Quanto aos regionalismos específicos, toma-se como base o conceito de Isquerdo (2007), a qual os considera como traços linguísticos específicos dos diferentes estados/regiões do Brasil.

Tendo em vista o papel do dicionário de registrar o léxico de uma língua, incluindo aí os regionalismos, é necessário observar como os dois dicionários em análise – *DEH* e *DEA* – concebem regionalismo. No primeiro dicionário, destaca-se no prefácio o seguinte excerto:

a indicação de **regionalismo** recai sobre palavra ou locução (dialetismo vocabular) ou acepção (dialetismo semântico) privativa de determinada região dentro do território onde se fala a língua e desconhecida das demais. É a informação sobre os limites geográficos da utilização de determinada unidade léxica ou acepção sua (HOUAISS; VILLAR, 2009, p. 5-6 – grifo no original).

Como se vê, o *DEH* toma regionalismo como as palavras ou acepções que são exclusivas de certas regiões. Já no *DEA*, não se encontra no prefácio o que se entende por regionalismo e como a obra irá tratá-lo. A menção a regionalismo é feita na seção *Verbetes – entenda sua estrutura*, quando se explica o que é rubrica. Nesse caso, tem-se:

a rubrica, em geral abreviada e em verde, situada antes das definições quando se refere a todas as definições do *verbe*, ou dentro de uma definição quando

se refere apenas a esta, delimita uma área em que a palavra é usada dentro do(s) significado(s) considerado(s), seja área geográfica, de assunto ou disciplina, de uso, etc. As definições assim classificadas formam, quando associadas em uma determinada rubrica, um subdicionário especializado na área delimitada pela rubrica (regionalismos, assunto ou disciplina, etc.). A mesma definição pode estar sob mais de uma rubrica. As abreviaturas das rubricas encontram-se na lista de abreviaturas, siglas e sinais convencionais (FERREIRA, 2010).

Com essas menções a regionalismos nos dois dicionários, é previsível a presença deles na nomenclatura. No entanto, como afirma Isquierdo (2011), o dicionário não deve tratar os regionalismos de forma aleatória, mas, primeiramente, deve definir o que ele entende por regionalismo e ter o cuidado de não restringir os regionalismos à determinada localização geográfica, quando, na verdade, seu uso pode ser mais amplo. Além disso, a autora cita as principais fontes que podem servir de base para os dicionários:

em se tratando, por exemplo, dos regionalismos, além dos *corpora* de língua escrita, dados colhidos de atos reais de fala também são muito úteis para atestar o uso de uma unidade lexical. Assim, *glossários, vocabulários regionais* resultantes de pesquisas acadêmicas (dissertações, teses), produzidos nos programas de pós-graduação das diferentes regiões do Brasil, na(s) sincronia(s) contempladas pela obra, configuram-se como fontes fidedignas de regionalismos. De forma similar, os atlas linguísticos também não podem ser desconsiderados como fonte de dados. À medida que documentam a língua em uso num determinado tempo e espaço, e, em virtude da metodologia que orienta as pesquisas geolinguísticas, as cartas lexicais oferecem dados concretos acerca da distribuição diatópica de uma variante lexical, representando fonte segura para a identificação e a classificação de regionalismos (CARDOSO; CASTRO; ISQUERDO, 2011, p. 62)¹³.

Nota-se que a autora aponta várias fontes que podem servir para atestar a existência de regionalismos, e um deles são os atlas linguísticos, os quais, resultantes de pesquisas geolinguísticas, podem contribuir para o adequado tratamento de regionalismos em dicionários, pois são elas que podem fornecer dados mais assertivos sobre a natureza dos regionalismos, como se vê em:

nesse sentido, os atlas linguísticos são de grande valia, pois fornecem dados confiáveis acerca do grau de disseminação de uma variante lexical, o que é garantido pela metodologia que orienta pesquisas geolinguísticas, em especial o uso de um questionário de natureza onomasiológica utilizado na coleta dos dados e uma rede de pontos distribuída na área investigada – um município, um estado, uma região ou um país, dependendo da amplitude do atlas, ou seja, se de pequeno ou de grande domínio (ISQUERDO, 2007, p. 199).

¹³Ainda que esta citação seja extraída de um texto de autoria de Cardoso, Castro e Isquierdo, cada uma das autoras escrevem em momentos diferentes, uma vez que o artigo, do qual fora tirada a citação, organiza-se em forma de entrevista. Sendo assim, este fragmento trata de uma ideia apenas de Isquierdo.

Fajardo (1996-1997) também reconhece as falhas no tratamento de regionalismos em dicionários e também defende o uso dos resultados das pesquisas dialetológicas e geolinguísticas para fundamentar as informações trazidas nos dicionários:

as marcações geográficas, portanto, têm fundamentos teóricos fracos e uma imprecisão característica. Em Lexicografia, despreza-se a descrição impossível da extensão geográfica de cada voz e se dão unicamente “localizações”, indicações ou pistas de lugares nos quais será geralmente compreensível. No entanto, deve existir uma constante preocupação do lexicógrafo em aproveitar as contribuições da dialetologia e da geografia linguística à hora de precisar, cada vez mais adequadamente, as marcas diatópicas no dicionário (FAJARDO, 1996-1997, p. 42)¹⁴.

Reconhecendo a importância das pesquisas dialetais e geolinguísticas na descrição do português do Brasil, como também na possibilidade de orientação de lexicográficos na classificação de regionalismos, este trabalho elege o *Atlas Lingüístico do Brasil (ALiB)* como parâmetro para investigar as marcas diatópicas dos regionalismos nos dicionários. Essa escolha não despreza nem diminui a importância que as outras fontes citadas por Isquierdo (2007), como atlas regionais, vocabulários, glossários e pesquisas acadêmicas, representam para auxiliar os lexicógrafos no tratamento de regionalismos.

Uma vez que se trabalha com dados de pesquisas dialetológicas neste trabalho, faz-se, no capítulo seguinte, uma discussão a respeito da Dialetologia e da Geografia Linguística no Brasil e suas principais contribuições.

¹⁴ Tradução livre para: Las marcación geográfica, por tanto, tiene débiles fundamentos teóricos y una característica imprecisión. En lexicografía se renuncia a la imposible descripción de la extensión geográfica de cada voz y se dan únicamente “localizaciones”, indicaciones o pistas de lugares em los que será generalmente comprensible. Sin embargo, debe existir una constante preocupación del lexicógrafo por aprovechar las aportaciones de la dialectología y la geografía linguística a la hora de precisar, cada vez más adecuadamente, las marcas diatópicas en el diccionario. (FAJARDO, 1996-1997, p. 42).

CAPÍTULO 3 – A DIALETOLOGIA E O ALiB

Neste capítulo, faz-se uma discussão sobre os estudos da Dialectologia, mostrando, primeiramente, como essa área se define e quais os principais pressupostos teórico-metodológicos que a fundamentam, além de apresentar os principais trabalhos dialetológicos realizados ao longo dos anos com o fim de elaborar um atlas, tanto em âmbito mundial quanto nacional. Posteriormente, serão apresentados os princípios metodológicos do Projeto ALiB, que resultou na publicação do primeiro *Atlas Linguístico do Brasil*. O conhecimento da natureza da pesquisa dialetológica, em especial do Projeto ALiB, poderá ajudar a entender como os dados que serão apresentados na seção de análise, no capítulo 5, foram obtidos.

3.1 Aspectos teórico-metodológicos da Dialectologia e os atlas linguísticos

A Dialectologia é uma das ciências que objetiva estudar as variedades da língua, sobretudo diatópicas, numa perspectiva sincrônica. Mais especificamente:

a Dialectologia é um ramo dos 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 (CARDOSO, 2010, p. 15).

Por dar atenção à língua em uso, a Dialectologia tem se assemelhado à Sociolinguística, cujo objetivo também é, *grosso modo*, descrever as variedades linguísticas. No entanto, os limites entre as duas ciências se estabelecem na medida em que, segundo Cardoso (2010), a Dialectologia prioriza a localização geográfica de determinada variedade linguística, enquanto que a Sociolinguística, ainda que se preocupe também com o fator diatópico, se interessa pela relação entre os fatos linguísticos e os fatos sociais, priorizando as relações sociolinguísticas. Nessa tarefa de aliar os usos da língua à geografia, a Dialectologia está intimamente relacionada à Geolinguística, porque esta, também chamada de Geografia linguística, é considerada o método da Dialectologia (CARDOSO, 2010).

Os estudos dialetais, ainda que tenham como foco a variação diatópica, têm se preocupado também com outros fatores sociais – idade, gênero, escolaridade, profissão –, principalmente à hora da seleção de informantes. Esses “veios sociolinguísticos”, como denomina Cardoso (2010), aproxima a Dialectologia da Sociolinguística, pois, como afirma a autora:

a dialectologia e especificamente os estudos geolinguísticos deixam de apresentar-se numa visão predominantemente diatópica e passam a exibir,

também cartograficamente, dados de natureza social. Isso vem mostrar, ainda, que a uniformidade diatópica pode sofrer desdobramentos e, tal como acontece com a célula humana, é passível de dividir-se, sem, contudo, perder a sua inteireza e a sua integridade (CARDOSO, 2010, p. 61-62).

Segundo Cardoso (2010), uma pesquisa de cunho dialetológico se realiza a partir de três aspectos: a rede de pontos, os informantes e os questionários. A rede de pontos é a área geográfica a ser investigada. Essa área pode ser uma única localidade, um estado, uma região, um país ou um continente. A determinação da localidade depende dos objetivos da pesquisa.

Além disso, como afirma a autora, a escolha de uma rede de pontos não deve se pautar apenas no princípio do isolamento, antiguidade e pouco desenvolvimento, como se fazia no começo dos estudos dialetológicos, mas que inclua também localidades com altos índices de desenvolvimento, urbanização e densidade demográfica, para que seja possível refletir as novas configurações do mundo moderno (CARDOSO, 2010).

Mesmo com a seleção de localidades mais urbanas e com uma pluralidade de falantes, como são as capitais do Brasil, as pesquisas levam em conta critérios de seleção de informantes, para que os dados sejam mais precisos. Para essa seleção, é necessário definir o número de informantes, a identificação deles – naturalidade, vinculações familiares, inserção social; e suas características sociais – idade, gênero e escolaridade.

Segundo Cardoso (2010), para a Dialetologia, um único informante de uma determinada localidade já é suficiente para que a informação seja validada, uma vez que, com essa informação, já se pode afirmar qual uso se faz na área investigada. Devido ao pequeno número de informantes selecionados, várias críticas são feitas ao método geolinguístico. Todavia, como afirma Cardoso (2010), o objetivo da Dialetologia não é afirmar se tal uso linguístico se caracteriza como variação estável ou mudança em curso, tampouco quantificar seu uso, estabelecendo percentuais de ocorrência, mas sim afirmar que “num dado lugar, registra-se tal fato, recolhido em tais circunstâncias” (CARDOSO, 2010, p. 92).

Quanto ao perfil dos informantes, Cardoso (2010) afirma que sua seleção deve atender aos objetivos da pesquisa, que pode ser voltada especificamente para a diversidade diatópica ou conjugada a outros fatores sociais. Sendo assim, fatores como idade, sexo, profissão e nível de escolaridade são levados em conta nessa seleção.

O questionário é o principal instrumento de coleta de dados. Ele se organiza a partir do tipo de dado a ser recolhido – fonético-fonológico, semântico-lexical ou morfossintático etc. Para tanto, Cardoso (2010, p. 98) adverte que “[...] é de suma importância a aplicação prévia,

um teste preliminar, antecedendo a realização da pesquisa, para verificar-se a eficácia do questionário e a pertinência da formulação das perguntas que o integram”.

Os trabalhos dialetológicos já têm certa tradição. Um dos primeiros atlas linguísticos produzidos, que, conforme Elia (1978), obteve considerável importância e ao qual se deve o surto moderno da Geografia Linguística, é o *Atlas Linguístico da França*, de Jules Gilliéron (*Atlas Linguístico da França*), publicado entre 1902 e 1910, em Paris. O colaborador para a realização deste atlas foi Edmont, que fez os inquéritos nas localidades pesquisadas.

Posterior a este trabalho, vários atlas foram produzidos em outros países, conforme Elia (1978), quais sejam: o *Novo Atlas Linguístico da França por Regiões*, organizado por A. Dauzat, publicado em 1939; o *Atlas Linguístico da Córsega*, de Gilliéron e E. Edmont; o *Atlas Linguístico-Etnográfico Italiano da Córsega (ALEIC)*, de Gino Bottigliani; o *Atlas Linguístico da Catalunha*, de monsenhor A. Grieria; o *Atlas Linguístico e Etnográfico da Itália e da Suíça Meridional (AIS)*; o *Atlante Linguístico Italiano (ALI)*; e o *Atlas Linguístico Romeno (ALR)*.

Os primeiros estudos de caráter dialetológico e geolinguístico, no Brasil, têm início ainda no século XIX. Aragão (2014) divide a história da Dialetologia e da Geografia Linguística em quatro momentos. O primeiro momento tem como marco inicial o trabalho de Domingos Borges de Barros, publicado em 1826 e intitulado *Les différences que le dialecte brésilien pourrait présenter, compare à la langue du Portugal*. Neste trabalho, o autor descreve as características do português do Brasil à época, destacando as influências das línguas indígenas.

Essa primeira fase também se caracteriza pela produção de vários glossários, vocabulários e dicionários regionais que tentaram registrar, no campo do léxico, os dialetos do Brasil. Entre eles, podem-se citar o *Dicionário da língua portuguesa*, de Luís Maria da Silva Pinto (1832), o *Vocabulário brasileiro para servir de complemento aos dicionários da língua portuguesa*, de Brás da Costa Rubim (1853), o *Dicionário brasileiro da língua portuguesa*, de Antônio Joaquim de Macedo Soares (1888), o *Glossário paraense*, de Vicente Chermont de Miranda (1906), entre outros.

A segunda fase tem como marco inicial a publicação de *O dialeto caipira*, de Amadeu Amaral, em 1920. Dois anos depois, Antenor Nascentes publica *O dialeto carioca*. Além deles, Mário Marroquim publicou, em 1934, *A língua do Nordeste*. Segundo Aragão (2014), nesta fase da Dialetologia, foram publicadas várias pesquisas que investigam o caráter fonético-fonológico e morfossintático de algumas regiões do Brasil, como também pesquisas que tratavam da influência das línguas africanas no português. Entre elas estão *O elemento afro-*

negro na língua portuguesa, de Jacques Raimundo (1933) e *A influência africana no português do Brasil*, de Renato Mendonça (1935).

A terceira fase se inicia com a promulgação do Decreto nº 30.643, de 20 de março de 1952, que determinou a elaboração de um atlas linguístico nacional pela Comissão de Filologia da Casa Rui Barbosa. A partir desse momento, começaram os estudos geolinguísticos propriamente ditos, que tiveram como principais representantes Antenor Nascentes, Serafim da Silva Neto, Celso Cunha e Nelson Rossi. Os dois primeiros publicaram obras basilares para as pesquisas dialetológicas: Antenor Nascentes publicou, em dois volumes, as *Bases para a elaboração do atlas linguístico do Brasil* (1958 e 1961) e Serafim da Silva Neto publicou o *Guia para estudos dialetológicos* (1955). Nessa fase ainda surgiram vários atlas linguísticos regionais. O próprio Serafim da Silva Neto, juntamente com Celso Cunha, defendia a elaboração de atlas regionais, para, só depois, elaborar um atlas nacional. No III Colóquio Internacional de Estudos Luso-Brasileiros, realizado em Lisboa, em 1957, os dois pesquisadores defenderam essa proposta, mas sem perder de vista o propósito maior.

O primeiro atlas linguístico regional do Brasil, publicado em 1963, foi o *Atlas prévio dos falares baiano*, coordenado por Nelson Rossi. A partir desse Atlas, surgiram outros atlas regionais. Hoje, o Brasil conta com dez atlas regionais. São eles: o já mencionado *Atlas prévio dos falares baianos* (1963), o *Esboço de um atlas lingüístico de Minas Gerais* (1977), o *Atlas lingüístico da Paraíba* (1984), o *Atlas lingüístico de Sergipe* (1987), o *Atlas lingüístico do Paraná* (1994), o *Atlas lingüístico sonoro do Pará* (2004), o *Atlas lingüístico de Sergipe II* (2005), o *Atlas lingüístico-etnográfico da Região Sul* (2002; 2013), o *Atlas lingüístico de Mato Grosso do Sul* (2007) e o *Atlas lingüístico de Estado do Ceará* (2010). Além dos primeiros atlas regionais, a terceira fase se caracteriza pela defesa de teses e dissertações que tratam da Dialetologia em diversos programas de pós-graduação.

Aragão (2014) afirma que a quarta fase tem como marco o surgimento do Projeto do Atlas Linguístico do Brasil, em 1996, a partir do qual as pesquisas geolinguísticas tomaram maior impulso. Esse projeto contou com a participação de pesquisadores de várias universidades do país, os quais, reunidos no Seminário *Caminhos e Perspectivas para a Dialetologia no Brasil*, realizado no Instituto de Letras da Universidade Federal da Bahia, em novembro de 1996, retomaram a ideia da construção de um atlas nacional e lançaram bases para, finalmente, pôr em prática o projeto.

Com a realização do projeto, publicou-se o *ALiB* em 2014, em dois volumes, com dados obtidos em 25 capitais do Brasil, exceto Palmas e Brasília. O volume 1 apresenta a história do

atlas e as questões metodológicas. O volume 2 contém as 159 cartas linguísticas que contemplam o nível fonético-fonológico, semântico-lexical e morfossintático. Além disso, há informações introdutórias como a caracterização das capitais e o perfil dos informantes. A seção seguinte traz mais informações sobre o Projeto *ALiB* e como ele foi desenvolvido.

3.2 O Projeto *ALiB*

O Projeto *Atlas Linguístico do Brasil* (Projeto *ALiB*) se insere no quadro metodológico da Geolinguística. O objetivo foi investigar as realizações linguísticas em redes de pontos espalhadas nos 26 estados do Brasil por meio de questionários. Para tanto, foram aplicados diferentes tipos de questionários: questionário fonético-fonológico (QFF), questionário semântico-lexical (QSL), questionário morfossintático (QMS), questões de pragmática (QP), questões metalinguísticas (QM) e aplicação de textos para leitura. Tendo em vista que o foco deste trabalho são regionalismos no nível lexical, destaca-se, entre os tipos de questionários citados, o questionário semântico-lexical (QSL).

Ele é constituído de 202 questões divididas em 14 áreas semânticas: 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; Vida urbana. As questões são de caráter onomasiológico, ou seja, partem do significado para a unidade lexical. Assim, a pergunta podia solicitar uma denominação, por meio de descrições, ou levava o informante a completar uma sentença feita pelo inquiridor.

Dos 14 campos semânticos que constituem o questionário semântico-lexical, apenas 8 deles foram contemplados nas cartas linguísticas do Volume 2. Desses oito, optou-se, neste trabalho, por analisar as variantes lexicais no campo semântico *Jogos e diversões infantis*, presentes nas cartas linguísticas de números L20 a L24. A escolha desse campo semântico se deu pelo fato de ele ser o campo de maior representatividade no *ALiB* e pela grande produtividade quanto às variantes lexicais.

3.1.1 Perfil dos informantes

Quanto ao perfil dos informantes, foi estabelecido, pelo Projeto *ALiB*, não só critério diatópico, mas também as variáveis sexo, faixa etária e o grau de escolaridade. Ao total, foram

1.100 informantes, sendo 550 homens e 550 mulheres. Desse número de informantes, foram selecionados dois do sexo masculino e dois do sexo feminino em cada localidade no interior dos estados. Já nas capitais, cujos resultados estão presentes no volume 2, foram oito informantes, sendo quatro mulheres e quatro homens.

Quanto à variável idade, o Projeto *ALiB* estabeleceu duas faixas: jovens, de 18 a 30 anos, e mais velhos, de 50 a 65. Em algumas situações, afirma Mota (2014), devido à dificuldade de se encontrar informantes dentro desses limites de idade, foi necessário admitir informantes fora dessa faixa etária. O grupo entre 30 a 50 foi excluído por razões operacionais: “para não aumentar demasiadamente o número de informantes, dificultando a tarefa que já se apresentava bastante onerosa” (MOTA, 2014, p. 91). O Projeto *ALiB* resolveu também não eleger informantes que fosse da mesma família, ou que fosse de bairros muito distintos economicamente, para que não houvesse interpretação de casos de variação social.

Em se tratando do grau de escolaridade, foram selecionados dois grupos: os não universitários e os universitários. No primeiro grupo, foram incluídos os que tinham até o ensino fundamental completo, mas dando preferência aos que não o tinham completado e, em alguns casos, admitiu-se os que tinham o segundo grau incompleto. No segundo grupo, foram incluídos os que tinham concluído a graduação ou aqueles que já tinham feito uma pós-graduação. Eventualmente, foram admitidos estudantes que estavam em fase de conclusão do curso de graduação. No entanto, foram evitados profissionais de Letras, Comunicação e outros cursos que trabalham com a linguagem (MOTA, 2014, p. 92).

3.1.2 O campo semântico *Jogos e diversões infantis*

A denominação *Jogos e diversões infantis*, proposta pelo *ALiB*, abrange, ao mesmo tempo, jogos, brinquedos e brincadeiras. Assim, o *ALiB* não faz distinção entre os três, considerando todos numa mesma classificação. Por isso, Ribeiro (2012), que pesquisou a respeito do mesmo campo semântico, optou por distinguir os dados em dois grupos: 1) *brincadeiras e jogos tradicionais*, incluindo *cambalhota*, *cabra-cega* e *amarelinha*; e 2) *brinquedos*, incluindo *bolinha de gude*, *estilingue* e *brinquedo de empinar (com e sem varetas)*.

Em estudos sobre ludologia, a distinção entre jogos, brinquedos e brincadeira também é um pouco conflituosa. No entanto, Kishimoto (2008) estabelece a seguinte distinção: o jogo se caracteriza pela existência de regras explícitas ou implícitas que ordenam e conduzem a brincadeira; o brinquedo se caracteriza pela ausência de um sistema de regras que organizam sua utilização de objetos; e a brincadeira seria “a ação que a criança desempenha ao concretizar

as regras do jogo, ao mergulhar na ação lúdica” (KISHIMOTO, 2008, p. 21). Ainda segundo a autora, as brincadeiras fazem parte da cultura de um povo e, geralmente, são repassadas às gerações pela transmissão oral, e assumem novos valores com o decorrer do tempo.

Neste trabalho, reconhece-se a distinção entre jogo, brincadeira e brinquedo e que, portanto, os regionalismos selecionados para análise do campo semântico *Jogos e diversões infantis*, se subdividem nessas três categorias. Essa distinção, no entanto, não incide, diretamente, na análise dos dados, pois nela não se leva em conta esse critério de classificação.

Alguns trabalhos anteriores a esse também investigaram o léxico desse campo semântico. Entre essas pesquisas, destaca-se a tese de Ribeiro (2012), intitulada *Brinquedos e brincadeiras infantis no âmbito do falar baiano*. A autora realiza inquéritos em várias redes de pontos a fim de pesquisar as variantes lexicais do campo semântico *Jogos e diversões infantis* proposto pelo ALiB. Todo procedimento teórico-metodológico se baseia nesse projeto maior.

No referido trabalho, a autora faz uma pesquisa de cunho dialetológico, investigando 57 localidades pertencentes a 11 estados brasileiros. Para a obtenção dos dados, foram utilizados os mesmos questionários no Projeto ALiB para o campo semântico selecionado. Além disso, a autora consultou as variantes obtidas em três dicionários de língua portuguesa: Houaiss, Aurélio e Aulete; e um de folclore: Cascudo. Os dados foram organizados em cartas linguísticas, seguindo as bases metodológicas da Geolinguística.

A consulta aos dicionários objetivou conhecer as formas lexicais que estavam registradas para cada brinquedo ou brincadeira. Porém, como se trata de uma pesquisa eminentemente dialetológica e Geolinguística, a reflexão sobre os dicionários não foi o objetivo principal.

Outro trabalho que também se preocupa com os estudos dos dialetos é a dissertação de Portilho (2013), intitulada *O falar amazônico: uma análise da proposta de Nascentes (1953) a partir de dados do Projeto ALiB (2013)*. A autora objetivou confirmar ou refutar a área dialetal amazônica proposta por Nascentes, mediante a análise de um acervo lexical restrito ao campo semântico *Jogos e diversões infantis*. A pesquisa é de natureza dialetológica e Geolinguística. Sendo assim, Portilho (2013) fez investigações em 26 localidades consideradas como constituintes da área do falar amazônico. Para a obtenção dos dados lexicais, a autora utilizou o questionário do Projeto ALiB e contou com 128 informantes, dos quais foram fornecidos 196 itens lexicais. A autora catalogou, descreveu e analisou as unidades lexicais do campo semântico selecionado e, posteriormente, fez o mapeamento dos dados em cartas linguísticas.

Portilho (2013) concluiu que há, sim, um falar amazônico, haja vista algumas peculiaridades de unidades lexicais investigadas, mas que não é bem delimitado, devido às influências das áreas limítrofes e dos processos imigratórios. O estudo também traz contribuições para a Lexicografia, identificando regionalismos que são característicos da região Norte do Brasil. Todavia, não se trata especificamente de um estudo metalexigráfico, pois a análise de dicionários também não é o foco.

Ainda que neste trabalho tenha-se analisado o léxico do mesmo campo semântico das pesquisas prévias mencionadas, esta pesquisa se diferencia por ter o caráter metalexigráfico, uma vez que analisa diretamente dicionários, tomando os dados dialetológicos apenas como âncora para as análises.

Diante do que se expôs neste capítulo, é possível afirmar que os estudos dialetológicos têm realizado grandes feitos quanto à descrição da língua portuguesa no Brasil, sobretudo com o Projeto *ALiB*. Com este Projeto de grande envergadura, somado aos atlas regionais anteriormente publicados, foi possível mapear as variedades diatópicas do país. No entanto, ainda que se tome o *ALiB* como parâmetro para a análise dos dicionários neste trabalho, é necessário ressaltar que o atlas não é tido como verdade absoluta e como aquele que deve ser a única base para os lexicógrafos quanto aos regionalismos, mas se reconhece, sim, a contribuição que o *ALiB* pode dar para a atualização dos dicionários.

CAPÍTULO 4 – METODOLOGIA

Neste capítulo, apresentam-se os aspectos metodológicos que nortearam a pesquisa. Para tanto, cumpre, num primeiro momento, apresentar o que se utilizou do *ALiB* para este trabalho e, posteriormente, apresentam-se o tipo de pesquisa, o *corpus* e os procedimentos adotados para a realização deste estudo.

4.1 O que se utilizou do *ALiB*

Considerando que o *ALiB* contém dados de diferentes níveis da língua – fonético-fonológico, semântico-lexical e morfossintático –, optou-se por utilizar apenas os dados do nível semântico-lexical. Esses dados foram coletados a partir de questionários de diferentes campos semântico, como citado no capítulo anterior.

Entre os campos semânticos que constituem o questionário semântico-lexical, elegeu-se o campo semântico *Jogos e diversões infantis*, presentes nas cartas linguísticas de números L20 a L24. A escolha desse campo semântico se deu pelo fato de ele ser o campo de maior representatividade no atlas e pela grande produtividade quanto às variantes lexicais. Além disso, o nome de jogos, brinquedos e brincadeiras reflete a diversidade cultural das regiões investigadas.

As variantes lexicais analisadas neste trabalho referem-se às questões 155, 156, 157, 158, 159 161 e 167 do questionário semântico-lexical do *ALiB*, quais sejam:

155 – Como se chama a brincadeira em que se gira o corpo sobre a cabeça e acaba sentado (*mímica*)?

156 – Como se chama as coisinhas redondas de vidro com que os meninos gostam de brincar?

157 – Como se chama o brinquedo feito de uma forquilha e duas tiras de borracha (*mímica*), que os meninos usam para matar passarinhos?

158 – Como se chama o brinquedo feito de varetas cobertas de papel que se empina ao vento por meio de uma linha?

159 – Como se chama o brinquedo parecido com o(a)_____ (*cf. item 158*), também feito de papel, mas sem varetas, que se empina ao vento por meio de uma linha?

161 – Como se chama a brincadeira em que uma criança, com os olhos vendados, tenta pegar as outras?

167 – Como se chama a brincadeira em que as crianças riscam uma figura no chão, formada por quadrados numerados, jogam uma pedrinha (*mímica*) e vão pulando com uma perna só? *SOLICITAR DESCRIÇÃO DETALHADA*

Para efeito de recorte, optou-se por investigar apenas as variantes que estão presentes nas cartas linguísticas da região Nordeste, das quais muitas estão presentes em outras regiões também. Os regionalismos selecionados neste trabalho e que serão investigados nos dicionários são os seguintes:

- 155 – CAMBALHOTA (carta 17)
- 156 – BOLINHA DE GUDE (carta L18)
- 157 – ESTILINGUE (carta L19)
- 158 – BRINQUEDO DE EMPINAR (COM VARETAS) (carta L20)
- 159 – BRINQUEDO DE EMPINAR (SEM VARETAS) (carta (L21)
- 161 – CABRA-CEGA (carta L22)
- 167 – AMARELINHA (carta L23)

Cada regionalismo desses contém um grupo de variantes, as quais também foram investigadas nos dicionários. A seguir, apresenta-se como a pesquisa foi realizada.

4.2 Natureza da pesquisa, *corpus* e procedimentos

Quanto à sua natureza, esta pesquisa se caracteriza por ser de cunho bibliográfico, uma vez que trabalha com dicionários e com um atlas linguístico, e de caráter descritivo, pois busca descrever o tratamento de regionalismos em obras lexicográficas. Além disso, a pesquisa é de natureza qualiquantitativa, haja vista que, além da descrição do tratamento de regionalismos, fez-se, neste trabalho, a quantificação deles, conforme em diferentes grupos.

A pesquisa utiliza como *corpus* o *Dicionário Eletrônico Houaiss da língua portuguesa (DEH)*, o *Dicionário Eletrônico Aurélio da língua portuguesa (DEA)* e o *Atlas Linguístico do Brasil (ALiB)*. A escolha desses dicionários deve-se à importância que eles têm em meio às obras lexicográficas, pois estão entre os maiores dicionários gerais de língua portuguesa, haja vista que ambos possuem mais de 100.000 verbetes (WELKER, 2004). A preferência pela versão eletrônica justifica-se pela facilidade e celeridade na consulta.

O *ALiB* também constitui o *corpus* por ser considerado hoje, dentre as obras dialetológicas, a que melhor descreve, em nível nacional, a variação linguística do Brasil, por meio de cartas linguísticas. Portanto, pode ser considerado uma fonte segura de descrição da variação diatópica no Brasil. No entanto, ressalta-se que o *ALiB* não é o único material que pode fornecer esses tipos de dados para os dicionários, já que os atlas regionais, os *corpora* de língua oral e escrita, bem como glossários e vocabulários regionais e pesquisas acadêmicas podem conter esses dados. Dentre essa diversidade de materiais, elegeu-se o *ALiB* para servir de parâmetro para as análises devido à sua importância entre as pesquisas dialetológicas, mas reconhecendo a existência e a importância de outras fontes.

No que concerne aos procedimentos de pesquisa, foram obedecidas as seguintes etapas metodológicas: 1) levantamento das variantes lexicais do campo semântico *Jogos e diversões infantis* presentes nas cartas linguísticas da região Nordeste; 2) consulta das variantes lexicais selecionadas no *DEH* e no *DEA*, a fim de perceber como estão registradas as variantes e o que se diz sobre elas, sobretudo quanto à localização geográfica; 3) comparação entre as informações dadas no *ALiB* e nos dicionários; 4) elaboração de quadros para a organização dos dados comparados, classificando-os como convergentes ou divergentes; 5) quantificação dos regionalismos dicionarizados ou não nas duas obras lexicográficas em análise, no intuito de desvelar se determinadas variantes que são de uso corrente na língua estão registradas ou não nos principais dicionários gerais. Isso pode oferecer subsídios para uma futura atualização dos dicionários.

Os quadros elaborados para a apresentação dos dados de cada grupo de variantes seguem o seguinte modelo:

Quadro 1: Modelo de quadro para apresentação dos dados

REGIONALISMOS	Dados do ALiB	Marcação no DEH	Marcação no DEA	Confronto entre os materiais
Variante 1	Apresentam-se as regiões e as capitais em que a variante aparece.	Apresentam-se as informações geográficas trazidas no dicionário	Idem	Classifica-se como Convergente quando as informações do <i>ALiB</i> e dos dicionários são semelhantes
Variante 2		Indica-se com Não há quando o dicionário registra a variante, porém não há, no verbete, informação geográfica	Idem	Classifica-se como Divergente quando as informações do <i>ALiB</i> e dos dicionários são diferentes
Variante 3		Indica-se com Variante não registrada quando o dicionário não registra a variante em análise.	Idem	Utiliza-se apenas (-) quando não é possível confrontar as informações, nos casos em que os dicionários não trazem localização geográfica

Fonte: Elaboração própria.

Além desse, elaborou-se um último quadro que sintetiza as informações obtidas na análise, classificando os regionalismos em categorias relacionadas à menção ou não de localização geográfica nos dicionários, registro ou não das variantes e presença ou não de acepções que remetam ao campo semântico *Jogos e diversos infantis*. Ei-lo:

Quadro 2: Modelo de quadro para classificação das variantes

	<i>DEH</i>	<i>DEA</i>
Variantes não registradas	Citam-se as variantes que não estão registradas no <i>DEH</i>	Citam-se as variantes que não estão registradas no <i>DEA</i>
Variantes registradas sem acepção que se refira a jogos ou diversão infantil	Citam-se as variantes que não possuem a acepção em análise no <i>DEH</i> .	Citam-se as variantes que não possuem a acepção em análise no <i>DEA</i> .
Variantes registradas sem localização geográfica	Citam-se as variantes que estão registradas no <i>DEH</i> , mas que não indicam localização geográfica	Citam-se as variantes que estão registradas no <i>DEA</i> , mas que não indicam localização geográfica
Variantes registradas com localização geográfica divergente ao <i>ALiB</i>	Citam-se as variantes do <i>DEH</i> cujas informações são divergentes do <i>ALiB</i>	Citam-se as variantes no <i>DEA</i> cujas informações são divergentes do <i>ALiB</i> .
Variantes registradas com localização geográfica convergente ao <i>ALiB</i>	Citam-se as variantes do <i>DEH</i> cujas informações são convergentes ao <i>ALiB</i>	Citam-se as variantes do <i>DEA</i> cujas informações são convergentes ao <i>ALiB</i>

Fonte: Elaboração própria.

De posse dessas informações teórico-metodológicas, passa-se à análise dos dados no capítulo que se segue.

CAPÍTULO 5 – DESCRIÇÃO E ANÁLISE DOS DADOS

Neste capítulo, apresenta-se a análise dos dados obtidos, discutindo-se cada regionalismo e suas variantes. O primeiro regionalismo a ser analisado é *cambalhota*, cujo quadro apresenta-se a seguir. Ressalta-se que, nos verbetes em que há mais de uma acepção e/ou rubrica, foram consideradas apenas as marcações geográficas da acepção que se refere à brincadeira infantil ou, no caso do *DEH*, os que têm *rubrica de ludologia*¹⁵.

Quadro 3: Variantes de *cambalhota*

¹⁵ Marcação que remeta a jogos, brinquedos ou brincadeiras infantis (v. notas 11 e 12, p. 66).

REGIONALISMOS	Dados do ALiB	Marcação no DEH	Marcação no DEA	Confronto entre os materiais
Cambalhota	Todas as capitais do Brasil	Não há	Não há	-
Bunda-canastra	Nordeste, exceto em São Luís, Aracaju e Salvador	<i>Variante não registrada</i>	<i>Variante não registrada</i>	-
Carambela	Norte (Macapá e Belém); Nordeste (São Luís e Teresina)	<i>Variante não registrada</i>	<i>Variante não registrada</i>	-
Cabriola	Nordeste (Salvador)	Não há	Não há	-
Cambona	Nordeste (Maceió)	Não há	Não há	-
Cambota	Todas as capitais do Centro-Oeste; no Sul (Curitiba e Porto Alegre); Sudeste (São Paulo); Nordeste (Fortaleza).	Brasil	Brasileirismo	Divergente
Cangapé	Norte (Rio Branco); Nordeste (Fortaleza).	1. Brasil 2. Maranhão/Alagoas	Maranhão a Alagoas	Divergente
Carambola	Todas as capitais do Norte; Nordeste (Teresina)	Não há	Não há	-
Maria-escambona	Nordeste (Aracaju)	<i>Variante não registrada</i>	<i>Variante não registrada</i>	-
Mortal	Norte (Boa Vista); Nordeste (São Luís); Sudeste (Belo Horizonte)	Não há	Não há	-
Piruetta	Norte (Macapá); Nordeste (São Luís); Sul (Curitiba); Centro-Oeste	Não há	Não há	-

	(Campo Grande)			
Aú	Nordeste (Salvador)	Brasil	Brasileirismo	Divergente

Fonte: elaboração própria com base no *DEH*, no *DEA* e no *ALiB*

Observa-se, a partir do quadro, que o *ALiB* apresenta 12 variantes lexicais para *cambalhota* (incluindo este) que aparecem na região Nordeste. Há outras variantes nas demais regiões, no entanto, escolheu-se apenas os que estão presentes na região Nordeste. A variante *cambalhota* é a que mais ocorre entre todas. Ela aparece em todas as capitais do Brasil, com maior expressividade em Vitória, Rio de Janeiro e Florianópolis, onde chega a 100% das ocorrências. Entre as capitais do Nordeste, as que mais registram uso dessa variante é Fortaleza e Salvador. Ainda que os dicionários *DEH* e *DEA* registrem essa variante, não há menção à localização geográfica. Caso houvesse, o ideal seria que fosse registrado como *brasileirismo*, uma vez que se trata de uma variante que ocorre em todo o país.

A variante *bunda-canastra*, segundo o *ALiB*, ocorre apenas no Nordeste, com exceção das capitais São Luís, Aracaju e Salvador. Trata-se, pois, de um regionalismo específico das capitais Maceió, Fortaleza, João Pessoa, Recife, Teresina e Natal. Quanto à marcação nos dicionários, observa-se que nenhum deles registra tal variante. Nota-se que essa variante ocorre apenas na região Nordeste e na maioria das capitais. Nesse caso, pode-se classificá-lo como um regionalismo do Nordeste, levando em conta sua ocorrência exclusiva nessa região e em 6 das 9 capitais. A variante *carambela* também não está registrada nos dicionários. Esta ocorre em duas capitais do Norte (Macapá e Belém) e em duas do Nordeste (São Luís e Teresina). Nesse caso, não se trata de um regionalismo específico do Nordeste, devido à ocorrência em outra região, como também a pouca ocorrência entre as capitais do Nordeste. Os dicionários, além de registrarem a variante *carambela*, poderiam informar, segundo os dados que o *ALiB* fornecem, como também de outras fontes, que se trata de um regionalismo ocorrente em capitais do Norte (Macapá e Belém) e do Nordeste (São Luís e Teresina).

A variante *cabriola* é uma das menos recorrentes: tem registro apenas na capital Salvador, no Nordeste. Mesmo com baixa recorrência, essa variante está registrada nos dois dicionários, no entanto, não possui indicação de localização geográfica. O mesmo acontece com a variante *cambona*, a qual ocorre em apenas uma capital do Nordeste (Maceió). No *DEH*, há quatro entradas diferentes para *cambona* (homônimos), mas nenhuma delas se refere a uma brincadeira infantil, como também não há nenhuma indicação de localização geográfica. Nesse caso, o dicionário poderia incluir a acepção, com a rubrica de *ludologia*. No *DEA*, há três

entradas para *cambona* e a primeira delas se refere à brincadeira infantil. Contudo, não há marcação geográfica. No caso dessas duas variantes – *cabriola* e *cambona* –, em que não há marcação geográfica, seria possível classificá-las, a partir do que os dados do *ALiB* indicam, como regionalismos de Salvador e de Maceió, respectivamente. Por ser esse uso restrito a essas capitais, é inviável classificá-los como do Nordeste.

Quanto à variante *cambota*, observa-se, como mostra o quadro, que ela ocorre em todas as capitais da região Centro-Oeste, em Curitiba e em Porto Alegre, no Sul; em São Paulo, no Sudeste, e em Fortaleza, no Nordeste. Sendo assim, essa variante ocorre em apenas quatro regiões do Brasil. O *DEH* classifica tal variante como *brasileirismo*, utilizando a rubrica *Brasil*, assim como o *DEA*, que utiliza a rubrica *brasileirismo*. No *DEH*, há duas entradas para *cambota*, a segunda é a que se refere à brincadeira infantil, e classifica-a como ‘uso informal’. Já no *DEA*, há três verbetes para *cambota*. O segundo verbe é o que se refere à brincadeira e é indicado com a rubrica Fam. (familiar). Analisando as indicações geográficas dos três materiais – *ALiB*, *DEH* e *DEA* –, é possível afirmar que eles são divergentes, uma vez que o *ALiB* mostra que a variante é utilizada em poucas capitais. Dessa forma, os dicionários poderiam indicar as capitais e/ou as regiões nas quais se utilizam a variante *cambota*, segundo os dados fornecidos pelo *ALiB*. Esses dados também podem ser confrontados com atlas regionais ou outra fonte de pesquisa disponível.

No caso da variante *cangapé*, nota-se, segundo o *ALiB*, que ela é pouco recorrente, porque há registros apenas em uma capital do Norte (Rio Branco) e em uma do Nordeste (Fortaleza). O *DEH* traz duas acepções para o verbe *cangapé*. As duas têm a ver com uma brincadeira, mas, na acepção 1, refere-se a um “pontapé na panturrilha para fazer o adversário cair durante a luta”, e, na acepção 2, refere-se a um “pontapé aplicado dentro da água, em uma espécie de jogo de capoeira”. Para a primeira acepção, indica-se a localização ‘Brasil’ e, para a segunda, ‘Maranhão e Alagoas’. O *DEA* também apresenta duas acepções, referentes à brincadeira infantil, sendo a primeira marcada como *Brasileirismo* e a segunda como variante ocorrente do Maranhão a Alagoas. Percebe-se, comparando os dois dicionários, que as informações que eles trazem são diferentes, uma vez que o *DEH* apresenta como sendo ocorrente no Maranhão e em Alagoas, já o *DEA* parece abranger outras localidades, porque compreende a região que vai do Maranhão até Alagoas. Os dois dicionários também divergem do *ALiB*. Por isso, um tratamento mais adequado da variante seria informar que ela é utilizada em Rio Branco e em Fortaleza, segundo os dados fornecidos pelo *ALiB*. É certo que os dados

do *ALiB* não são definitivos para determinar o tratamento que os dicionários devem dar. É necessário uma consulta a outras fontes que também forneçam dados sobre essa variante.

A variante *carambola* está presente em todas as capitais do Norte e em apenas uma capital do Nordeste – Teresina. O *DEH* não dá indicação de localização geográfica. No *DEA*, a variante é registrada, mas não há acepção que se refira à brincadeira infantil, quanto menos localização geográfica. Nesses casos, para um tratamento mais adequado da variante, o *DEH* poderia se valer dos dados do *ALiB*, informando que se trata de regionalismo recorrente em todas as capitais do Norte e em uma do Nordeste (Teresina); e o *DEA* poderia incluir a acepção que remeta à brincadeira infantil como também os dados geográficos trazidos pelo *ALiB* para a variante *carambola*.

A variante *maria-escambona* ocorre somente no Nordeste, na capital Aracaju, e não está registrada em nenhum dos dicionários analisados. Talvez a baixa recorrência da variante entre as capitais justifique o fato de ela não estar registrada nos dicionários. Como se discutiu no capítulo 2, Biderman (1994) afirma que a frequência é um dos principais critérios que deve orientar o registro de uma palavra no dicionário. Baseando-se nisso, os dicionários podem se eximir de inserir a variante *maria-escabona*, devido a sua baixa recorrência.

Quanto à variante *mortal*, nota-se que ela ocorre em três regiões do Brasil: Norte (Boa Vista), Nordeste (São Luís) e Sudeste (Belo Horizonte). Nos dicionários *DEH* e *DEA*, há registro da variante, mas não há acepção que remeta à brincadeira infantil. Seria, então, o caso de os dicionários incluírem a acepção e informar os dados geográficos que o *ALiB* fornece. Nesse caso, também é possível que se utilize atlas, vocabulários e glossários regionais que também atestem essa acepção.

A variante *pirueta* também ocorre em várias regiões: no Norte (Macapá), no Nordeste (São Luís), no Sul (Curitiba) e no Centro-Oeste (Campo Grande). Porém, não há marcação de localização geográfica nos dois dicionários. Nesse caso, como o *DEH* e o *DEA* já registram a variante com a acepção referente à brincadeira infantil, poderiam acrescentar ao verbete os dados geográficos.

A última variante é *aú*. Segundo o *ALiB*, essa variante ocorre apenas na capital Salvador, na região Nordeste. Trata-se, então, de um regionalismo específico de Salvador, uma vez que não ocorre em outras capitais. No entanto, os dois dicionários analisados registram essa variante com a rubrica de *brasileirismo*, estando divergente dos dados dialetológicos. Um tratamento mais adequado dessa variante seria registrá-la como um regionalismos de Salvador, como atesta o *ALiB*.

Observa-se, com essa análise, que das 12 variantes, 3 não estão dicionarizadas nos dois dos maiores dicionários de língua portuguesa. Além disso, 6 das 9 variantes registradas não têm qualquer marcação de localização geográfica. Outras nem mesmo possuem acepções referentes à brincadeira infantil. Das 3 variantes que possuem localização geográfica nos dois dicionários, apenas 1 contém informações convergentes aos dados do *ALiB*. Isso mostra a necessidade de atualização desses materiais, tanto no registro de variantes, quanto na inserção de acepção que remeta ao campo semântico *Jogos e diversões infantis* e de marcação de localização geográfica.

A seguir, apresenta-se o segundo quadro que contém as variantes de *bolinha de gude* (incluindo esta), como também os dados dos três materiais em análise:

Quadro 4: Variantes de *bolinha de gude*

REGIONALIMOS	Dados do ALiB	Marcação no DEH	Marcação no DEA	Confronto entre os materiais
Bola de gude	Norte (Manaus); Nordeste (todas as capitais); Sudeste (Vitória, Rio de Janeiro e São Paulo); Sul (todas as capitais); Centro-Oeste (todas as capitais)	Brasil	Brasileirismo	Convergente
Bolinha de gude	Norte (Manaus); Nordeste (São Luís, Teresina, Recife e Maceió); Sudeste (todas as capitais); Sul (Florianópolis e Porto Alegre); Centro-Oeste (Goiânia e Campo Grande)	Brasil	<i>Variante não registrada</i>	Convergente (para o DEH)
Bila	Nordeste (Fortaleza e João Pessoa)	<i>Variante não registrada</i>	Ceará	Divergente
Peteca	Norte (todas as capitais e	Pará	Pará	Divergente

	100% em todas elas, exceto em Manaus); Nordeste (São Luís e Teresina)			
Biloca	Nordeste (Natal); Centro-Oeste (Goiânia)	Goiás	Goiás	Divergente
Bola de fona	Nordeste (João Pessoa)	Não há	Brasileirismo	Divergente
Bolinha de vidro	Nordeste (Natal); Sudeste (Vitória, Belo Horizonte e São Paulo); Sul (Florianópolis); Centro-Oeste (Goiânia)	<i>Variante não registrada</i>	Brasileirismo	Divergente (para o DEA)
Cabeçulinha	Nordeste (Fortaleza)	<i>Variante não registrada</i>	Brasileirismo	Divergente (para o DEA)
Marraio	Nordeste (Aracaju)	Não há	Brasileirismo	Divergência (para o DEA)
Ximbra	Nordeste (Maceió)	Alagoas	Alagoas	Convergente

Fonte: elaboração própria com base nos dados do ALiB, do DEH e do DEA

A partir dos dados acima, observa-se que há 10 variantes para *bolinha de gude*. A primeira delas é *bola de gude*, a qual, segundo o ALiB, é a mais recorrente e aparece em todas as regiões do Brasil, na maioria das capitais. No DEH, a variante ‘bola de gude’ não aparece como cabeça de verbete. Ela é encontrada dentro do verbete *bola*, entre as locuções formadas por esta palavra e é registrada com a marca de *Brasil*. No DEA, não há uma entrada para *bola de gude*. Contudo, encontra-se essa variante na acepção 2 do verbete *bolinha*: “jogo com bola de gude”, em que há rubrica de *brasileirismo*. Nos verbetes *bola* e *gude* não há indicação para a variante em questão, mas *gude* é definido como “bolinhas de vidro, etc., usadas nesse jogo”. Comparando as informações dos três materiais, observa-se convergência entre eles, pois a variante em questão ocorre em todo o território brasileiro, como aponta o ALiB, o DEH e o DEA.

Quanto à variante *bolinha de gude*, o ALiB informa que tal variante ocorre também em todas as regiões do Brasil, em algumas capitais. Sendo assim, é uma variante que tem expressividade em todo o território brasileiro, portanto, um *brasileirismo*. Em consonância com essas informações, o DEH classifica a variante como sendo um *brasileirismo*. Importante

ressaltar que, neste dicionário, não há entrada para a variante “bolinha de gude”, mas sim para “bolinha”. Na acepção 5 do verbete *bolinha*, há uma remissão para *bolinha de gude* e também a indicação de regionalismo Brasil. No *DEA*, não há entrada para *bolinha de gude*, mas sim “bolinha”. No entanto, não há uma acepção que apresente a forma *bolinha de gude*, logo essa variante não está registrada no dicionário. Um tratamento adequado no *DEA* seria incluir a forma *bolinha de gude* e marcá-la como *brasileirismo*.

A variante *bila* ocorre apenas no Nordeste, segundo o *ALiB*, mais especificamente nas capitais Fortaleza e João Pessoa. Nesse caso, tem-se mais um regionalismo que pode ser considerado exclusivo de duas capitais do Nordeste. O *DEH* não registra tal variante, mas o *DEA* a registra, indicando que se trata de um regionalismo do Ceará. Como se vê, o *DEA* se aproxima dos dados dialetológicos, no entanto, precisa incluir também o estado da Paraíba (ou mesmo a capital João Pessoa), com base nos dados do *ALiB*. Quanto ao *DEH*, seria necessário registrar a variante *bila* e informar que se trata de um regionalismo de Fortaleza e João Pessoa.

A variante *peteca* ocorre em todas as capitais do Norte, com ocorrência de 100% em quase todas elas, com exceção de Manaus. Ocorre também em São Luís e em Teresina, no Nordeste. Essa variante não aparece nas outras regiões e predomina na região Norte. Segundo as informações do *DEH* e do *DEA*, a variante *peteca* ocorre apenas no Pará. No entanto, como se percebe pelos dados do *ALiB*, o uso de tal variante é recorrente em outros estados. Dessa forma, os dados dos três materiais são divergentes e os dicionários precisam atualizar a informação geográfica, incluindo as outras capitais do Norte, bem como as do Nordeste – São Luís e Teresina.

Outra divergência de informações ocorre com a variante *biloca*. Segundo os dados do *ALiB*, seu uso é recorrente em apenas duas capitais: Natal, no Nordeste, e Goiânia, no Centro-Oeste. Os dicionários *DEH* e *DEA* trazem a mesma localização geográfica para a variante – Goiás. As informações dos dicionários parecem se aproximar dos dados dialetológicos, embora seja necessário acrescentar ainda a capital Natal, que também registra usos da variante *biloca*. Por isso, é possível observar a divergência existente entre as informações dos três materiais.

Quanto à variante *bola de fona*, o *ALiB* registra uso apenas na capital João Pessoa, no Nordeste. Nesse caso, tem-se, então, um regionalismo específico de João Pessoa, pois só ocorre nessa capital. O *DEH* não contém a palavra-entrada *bola de fona*, mas sim *fona*. Há três entradas para essa variante. É a primeira entrada que tem sentido próximo ao de *bolinha de gude*. Porém, não há indicação de localização geográfica. No *DEA* também não há a palavra entrada *bola de fona*, e sim *fona*. Há quatro entradas para essa variante, sendo que a quarta refere-se a uma

brincadeira, na qual há a marcação de *brasileirismo*. Nota-se que essa informação diverge dos dados dialetológicos, uma vez que a variante *bola de fona* ocorre, segundo o *ALiB*, apenas em uma capital do Nordeste. Devida a baixa recorrência dessa variante e retomando o critério de Biderman (1994) sobre frequência, podem os dicionários escusarem-se de registrar tal variante.

A variante *bolinha de vidro* ocorre em quatro regiões, a saber: Nordeste (Natal); Sudeste (Vitória, Belo Horizonte e São Paulo); Sul (Florianópolis); Centro-Oeste (Goiânia). Dessa forma, há pouca representatividade entre as capitais, mas está presente em quase todas as regiões. O *DEA* a registra como *brasileirismo*, no entanto, devido à ausência dessa variante na região Norte, é comprometedor classificá-la como *brasileirismo*, uma vez que, como os dados apontam, seu uso está ainda restrito a determinadas capitais. Por isso, uma possível saída seria indicar as capitais que registram o uso dessa variante. Ainda no *DEA*, a variante *bolinha de vidro* não aparece como cabeça do verbete, mas como uma das acepções da definição de *gude*: “bolinhas de vidro, etc., usadas nesse jogo”. Já no *DEH* não há registro dessa variante de nenhuma forma.

Outro regionalismo analisado foi *cabeçulinha*, o qual é utilizado, segundo o *ALiB*, apenas em Fortaleza, no Nordeste. Tem-se, assim, outro caso de regionalismo que pode ser considerado específico de Fortaleza. O *DEH* não registra essa variante. Portanto, sugere-se que o dicionário inclua na nomenclatura tal variante e indique que se trata de regionalismo de Fortaleza. O *DEA*, por seu turno, inclui a variante *cabeçulinha* como palavra-entrada, no entanto, com uma grafia diferenciada: *cabiçulinha*. Observa-se, então, a existência de variantes ortográficas, mas que não são indicadas nem no *ALiB* nem no *DEA*. Apesar de registrar a variante, o *DEA* diverge dos dados dialetológicos, porque classifica *cabiçulinha* como um *brasileirismo*, sendo que se trata de regionalismo específico de uma capital do Nordeste: Fortaleza. O tratamento adequado quando à marcação diatópica seria especificar a localidade em que a variante é utilizada, segundo o *ALiB*.

Outras duas variantes que também podem ser consideradas específicas de capitais do Nordeste, segundo o *ALiB*, são *marraio* e *ximbra*. A primeira ocorre apenas em Aracaju e não está registrada no *DEH*. Este dicionário poderia incluir a variante e informar que se trata de regionalismo de Aracaju. O *DEA* registra esse regionalismo, mas o classifica como *brasileirismo*, divergindo, assim, dos dados dialetológicos. A segunda, que é utilizada apenas em Maceió, segundo o *ALiB*, está registrada nos dois dicionários e ambos informam que tal variante ocorre em Alagoas, estando, então, convergentes com os dados dialetológicos.

Com a análise das variantes de *bolinha de gude*, nota-se, mais uma vez, o distanciamento entre os dados do *ALiB* e as informações nas obras lexicográficas. Das 10 variantes analisadas, 3 não são registradas no *DEH* e 1 não aparece *DEA*. Todas as variantes de *bolinha de gude* registradas nos dicionários e listadas no **Quadro 4** têm acepção referente à brincadeira infantil, diferentemente das variantes de *cambalhota* listadas no **Quadro 3**, entre as quais houve casos de falta dessa acepção. Na comparação entre os três materiais, apenas 4 variantes apresentam dados convergentes. Isso evidencia, mais uma vez, a necessidade de atualização dos dicionários no registro das variantes que não aparecem nas obras e na atualização dos dados geográficos.

Outro regionalismo investigado pelo *ALiB* é *estilingue*. A seguir, apresenta-se o quadro com suas variantes, os dados do *ALiB* e dos dois dicionários.

Quadro 5: Variantes de *estilingue*

REGIONALIMOS	Dados do <i>ALiB</i>	Marcação no <i>DEH</i>	Marcação no <i>DEA</i>	Confronto entre os materiais
Estilingue	Todas as capitais (100% em Goiânia)	Brasil	Minas Gerais; Centro-Oeste; Sul	Convergente (para o <i>DEH</i>) Divergente (para o <i>DEA</i>)
Baladeira	Norte (todas as capitais); Nordeste (exceto Maceió e Salvador)	Acre a Pernambuco	Acre a Pernambuco	Divergente
Atiradeira	Norte (Macapá); Nordeste (São Luís, Fortaleza, Natal, Recife, Aracaju e Salvador)	Brasil	Brasileirismo	Divergente
Badogue	Norte (Macapá); Nordeste (Recife, Aracaju e Salvador); Sudeste (Belo Horizonte e São Paulo);	Bahia	Bahia	Divergente
Estilete	Nordeste (São Luís)	Não há	Não há	-
Peteca	Nordeste (Maceió)	Não há	Nordeste	Divergente (para o <i>DEA</i>)

Fonte: elaboração própria com base nos dados do *ALiB*, do *DEH* e do *DEA*

Entre as variantes apresentadas acima, *estilingue* é a mais recorrente, segundo os dados do *ALiB*. Ela é utilizada em todas as capitais que foram investigadas pelo *ALiB* e tem 100% de ocorrência em Goiânia. O *DEH* converge com o *ALiB*, classificando essa variante como *brasileirismo*. O *DEA*, por seu turno, diverge do *ALiB*, porque ele restringe o uso de *estilingue* a Minas Gerais, Centro-Oeste e Sul. O dicionário começa citando um estado da região Sudeste (Minas Gerais), mas logo em seguida generaliza o uso citando duas regiões. Seria mais viável citar os estados em que o uso dessa variante é, de fato, comprovada. Com os dados do *ALiB*, é mais seguro afirmar que *estilingue* se trata de um regionalismo geral, como nomeia Isquierdo (2007), ou um *brasileirismo*, uma vez que ocorre em quase todas as capitais do Brasil.

A variante *baladeira* é utilizada, segundo o *ALiB*, em todas as capitais do Norte e em capitais do Nordeste, exceto Maceió e Salvador. Os dicionários *DEH* e *DEA* trazem a mesma localização geográfica para esse regionalismo: Acre a Pernambuco. Com essa informação, não fica muito claro que capitais/estados abrangem essa delimitação geográfica. Além disso, o uso desse regionalismo não tem como extremo Pernambuco, como dizem os dicionários, porquanto ele se estende até o estado de Sergipe, conforme o *ALiB*. Dessa forma, um tratamento adequado seria informar, nos verbetes, que *baladeira* é um regionalismo de todas as capitais do Norte e da maioria do Nordeste, com exceção de Maceió e Salvador.

Outra variante presente no *ALiB* é *atiradeira*. Segundo o *ALiB*, essa variante é restrita às regiões Norte e Nordeste, mais especificamente às capitais Macapá, São Luís, Fortaleza, Natal, Recife, Aracaju e Salvador. Os dicionários *DEH* e *DEA* tratam essa variante como *brasileirismo*, o que diverge dos dados do *ALiB*. Por isso, os dicionários poderiam indicar as capitais em que o uso da variante já foi comprovado pelo *ALiB*.

Quanto à variante *badogue*, tem-se outro caso de divergência de dados. Enquanto que os dicionários dizem que essa variante é utilizada apenas na Bahia, o *ALiB* informa que seu uso é mais abrangente: Norte (Macapá); Nordeste (Recife, Aracaju e Salvador) e Sudeste (Belo Horizonte e São Paulo). Dessa maneira, fica clara a necessidade de incluir, nos verbetes dos dicionários, as localizações mostradas no *ALiB*.

A variante *estilete* é utilizada apenas em São Luís, segundo o *ALiB*. No *DEH*, não há informação quanto à localização geográfica, e o *DEA*, ainda que registre tal variante, não possui nenhuma acepção que se refira a jogos ou diversão infantil. Um tratamento adequado dessa variante seria, no caso do *DEA*, inserir a acepção referente à brincadeira infantil e, em ambos, indicar que tal regionalismo é específico de São Luís.

Por fim, tem-se a variante *peteca*, que também é utilizada em apenas uma capital do Nordeste (Maceió). O *DEH* registra tal variante, porém, não traz uma marcação geográfica. O *DEA*, por sua vez, informa que se trata de um regionalismo do Nordeste. Nesse caso, o dicionário não especifica em que estado/capital do Nordeste essa variante é utilizada, levando a entender que ela é utilizada em toda a região, o que não procede, tendo em vista os dados do *ALiB*. Dessa forma, o *DEA* diverge do *ALiB*. Nesse caso, seria mais adequado tanto o *DEH* quanto o *DEA* informar que se trata de um regionalismo de Maceió.

Os quadros seguintes se referem a dois objetos muito semelhantes. São brinquedos de empinar, mas um é com varetas e com linha, e o outro é sem varetas e com linha. Pela proximidade física, as variantes de ambos são também praticamente as mesmas, com poucas exceções. Os dicionários não fazem distinção dos dois objetos (se um é com vareta ou não). Dessa maneira, os verbetes consultados no **Quadro 6** são os mesmos no **Quadro 7**. Todavia, o *ALiB* trata essas variantes como diferentes, uma vez que nomeiam objetos diferentes. Por conta disso, preferiu-se manter dois quadros, ainda que as informações dos dicionários se repitam. Ei-los:

Quadro 6: Variantes de *brinquedo de empinar (com varetas, com linha)*

REGIONALIMOS	Dados do <i>ALiB</i>	Marcação no <i>DEH</i>	Marcação no <i>DEA</i>	Confronto entre os materiais
Papagaio	Todas as capitais	Não há	Não há	-
Pipa	Todas as capitais	Não há	Brasileirismo	Convergente (para o <i>DEA</i>)
Raia	Nordeste (Fortaleza, Maceió, Aracaju, Salvador); Sudeste (Vitória e São Paulo); Sul (Curitiba); Centro-Oeste (Goiânia)	Brasil	Brasileirismo	Divergente
Coruja	Nordeste (Natal e João Pessoa)	Não há	Não há	-

Fonte: elaboração própria com base nos dados do *ALiB*, do *DEH* e do *DEA*

A primeira variante que designa *brinquedo de empinar (com varetas, com linha)* é *papagaio*. Segundo o *ALiB*, essa variante é recorrente em todas as capitais investigadas. Portanto, dada o seu uso representativo em todo o país, *papagaio* pode ser considerado um

brasileirismo. Os dicionários *DEH* e *DEA* não trazem indicação de localização geográfica. Sendo assim, um adequado tratamento para essas variantes nesses dicionários seria classificá-los como *brasileirismo*.

O mesmo ocorre com a variante *pipa*, que também é recorrente em todas as capitais. Enquanto o *DEH* não traz informação sobre localização geográfica, o *DEA* classifica-a como *brasileirismo*, o que converge com os dados dialetológicos. O *DEH*, baseando-se no *ALiB*, poderia também indicar que se trata de um *brasileirismo*.

A variante *raia* ocorre em algumas capitais específicas. Na região Nordeste, é usada em Fortaleza, Maceió, Aracaju e Salvador; no Sudeste, em Vitória e em São Paulo; no Sul, Curitiba; e no Centro-Oeste, em Goiânia. Não há ocorrência em capitais da região Norte. Quanto ao tratamento dos dicionários, o *DEH* traz duas entradas para *raia*. A primeira contém duas acepções, uma com rubrica de ictiologia (ramo da zoologia que estuda os peixes) e a outra com rubrica de *ludologia* (a que nos interessa). A marcação de regionalismo no Brasil é feita para as duas acepções. O *DEA* também traz duas entradas para *raia*. Na segunda, a acepção 2 remete ao verbete *arraia*², cuja segunda acepção refere-se a *papagaio* (brinquedo de empinar) e em que também é indicada a marca de *Brasileirismo*. Comparando as informações dos três materiais, observa-se que tanto o *DEH* quanto o *DEA* concebem *raia* como *brasileirismo*. No entanto, o *ALiB* não possibilita tal classificação, uma vez que esta variante ocorre em capitais específicas e não em todo o território nacional. Dessa maneira, os dicionários poderiam, aproveitando os dados do *ALiB*, indicar as regiões e as capitais onde *raia* é usada.

A variante *coruja* tem também uso muito restrito, como mostra o *ALiB*. Ela ocorre apenas na região Nordeste, em duas capitais: Natal e João Pessoa. Os dicionários, por sua vez, mesmo registrando a palavra *coruja*, não trazem qualquer acepção que remeta a brinquedo de empinar ou rubrica de *ludologia*, no caso do *DEH*. Por isso, seria ideal que as duas obras incluam, nos verbetes, a acepção que se refira a jogos ou diversão infantil, como também indiquem que o uso da palavra *coruja* nesse sentido de “brinquedo de empinar” ocorre apenas em Natal e João Pessoa.

Agora, passa-se à análise das variantes que nomeiam um *brinquedo de empinar* (*sem varetas, com linha*). Como dito anteriormente, algumas das variantes presentes no **Quadro 6** se repetirão no **Quadro 7**, mas os dados do *ALiB* se diferem. Segue o quadro:

Quadro 7: Variantes de *brinquedo de empinar* (*sem varetas, com linha*)

REGIONALIMOS	Dados do <i>ALiB</i>	Marcação no <i>DEH</i>	Marcação no <i>DEA</i>	Confronto entre os materiais
--------------	----------------------	------------------------	------------------------	------------------------------

Papagaio	Norte (Porto Velho); Nordeste (Fortaleza, Piauí, Recife, Aracaju e Salvador); Sudeste (Belo Horizonte, Vitória e Rio de Janeiro); Sul (Florianópolis); Centro-Oeste (Cuiabá e Campo Grande)	Não há	Não há	-
Pipa	Norte (Boa Vista, Macapá, Manaus, Porto Velho); Nordeste (Teresina, João Pessoa – 100%, Recife, Maceió e Salvador); Sudeste (Vitória); Sul (Curitiba)	Não há	Brasileirismo	Divergente (para o <i>DEA</i>)
Avião	Nordeste (Maceió e Salvador); Sudeste (Vitória); Sul (Florianópolis); Centro-Oeste (Cuiabá)	Não há	Não há	-
Curica	Norte (todas as capitais); Nordeste (São Luís e Teresina)	Não há	Amazonas; Maranhão e Pará	Divergente (para o <i>DEA</i>)
Periquito	Nordeste (Aracaju e Salvador); Sul (Florianópolis)	Não há	Não há	-
Ratinho	Nordeste (Natal 100% e Maceió)	Não há	Não há	-
Raia	Nordeste (Maceió); Sudeste (Vitória, Rio de Janeiro e	Brasil	Brasileirismo	Divergente

	São Paulo); Sul (Curitiba e Florianópolis); Centro-Oeste (Goiânia 100%)			
Bolachinha	Nordeste (Fortaleza)	<i>Variante não registrada</i>	<i>Variante não registrada</i>	-

Fonte: elaboração própria com base nos dados do ALiB, do DEH e do DEA.

A variante *papagaio*, que no **Quadro 6** é recorrente em todas capitais, ocorre nas cinco regiões do país, em capitais específicas: Norte (Porto Velho); Nordeste (Fortaleza, Teresina, Recife, Aracaju e Salvador); Sudeste (Belo Horizonte, Vitória e Rio de Janeiro); Sul (Florianópolis); Centro-Oeste (Cuiabá e Campo Grande). Os dicionários não trazem a localização geográfica. Com isso, tanto o *DEH* quanto o *DEA* poderiam se valer desses dados dialetológicos para incluí-los nos verbetes.

A variante *pipa*, que no **Quadro 6** também é recorrente em todas as capitais, é usada em apenas três regiões, com predominância em capitais do Norte e do Nordeste: Norte (Boa Vista, Macapá, Manaus, Porto Velho); Nordeste (Teresina, João Pessoa, Recife, Maceió e Salvador); Sudeste (Vitória); Sul (Curitiba). Como já dito anteriormente, o *DEH* não indica localização geográfica para a variante *pipa*, e o *DEA* classifica-o como brasileirismo. No caso de *pipa* como *brinquedo de empinar (sem varetas, com linha)*, há divergência entre o *ALiB* e o *DEA*, porque não se trata de um brasileirismo, já que a variante ocorre somente em capitais de três regiões.

Quanto à variante *avião*, percebe-se, segundo o *ALiB*, que ela é pouco recorrente. Foi encontrada em apenas três regiões, em algumas capitais: Nordeste (Maceió e Salvador); Sudeste (Vitória); Sul (Florianópolis) e Centro-Oeste (Cuiabá). Seria necessário, pois, acrescentar essa aceção, indicando as capitais em que tal uso é atestado pelo *ALiB*.

Outra variante que designa *brinquedo de empinar (sem varetas, com linha)* é *curica*. Ela é utilizada em todas as capitais do Norte e em São Luís e Teresina, no Nordeste. O verbe *curica*, no *DEH*, não traz aceção referente à *ludologia*. Já o *DEA* traz essa aceção, como também informações geográficas para o verbe *curica*: diz que se trata de um regionalismo dos estados de Amazonas, Maranhão e Pará. Essa informação se aproxima dos dados do *ALiB*, no entanto, ela restringe o uso da variante a três localidades, sem mencionar os outros estados que também utilizam *curica*. Pode-se, portanto, considerar divergentes os dados do *ALiB* e do *DEA*. Talvez seja possível classificar *curica* como um regionalismo do Norte e do Nordeste, tendo em vista as informações que o *ALiB* fornece.

Além de *curica*, o quadro acima apresenta outra variante que é utilizada para se referir a *brinquedo de empinar (sem varetas, com linha): periquito*. O *ALiB* registrou o uso dela em apenas três capitais: Aracaju e Salvador, no Nordeste, e Florianópolis, no Sul. O *DEH* e *DEA*, não obstante registrem tais variantes, não incluem, no verbete, acepção referente à brincadeira infantil. Tem-se, então, mais um caso em que os dicionários precisam incluir uma nova acepção no verbete.

A variante *ratinho* é utilizada apenas na região Nordeste, mais especificamente na capital Natal, onde tem ocorrência de 100%, e em Maceió. Os dicionários registram a variante, mas não há acepção que remeta ao campo semântico em análise, também não há localização geográfica. O ideal seria incluir a acepção e informar que se trata de um regionalismo de Natal e Maceió.

A variante *raia*, que também se encontra no **Quadro 6**, está presente em capitais do Nordeste (Maceió), do Sudeste (Vitória, Rio de Janeiro e São Paulo), do Sul (Curitiba e Florianópolis) e do Centro-Oeste (Goiânia, com ocorrência de 100%). Nota-se que tal variante, ao referir-se a *brinquedo de empinar (sem varetas, com linha)*, tem certa representatividade no país. No entanto, não há nenhuma ocorrência no Norte, portanto, trata-se de uma variante restrita a determinadas regiões. Por isso, os dados do *ALiB* divergem dos dicionário *DEH* e *DEA*, que classificam *raia* como brasileirismo. Com os dados que o *ALiB* fornecem, seria possível informar que esse regionalismo ocorre em capitais específicas: Maceió, Vitória, Rio de Janeiro, São Paulo, Curitiba, Florianópolis e Goiânia.

Por fim, a última variante desse grupo é *bolachinha*, que tem ocorrência apenas em Fortaleza (Nordeste). O tratamento dado a essa variante no *DEH* e no *DEA* é o mesmo: não há entrada para *bolachinha*, e, no verbete *bolacha*, não há menção à variante *bolachinha* nem acepção referente ao campo semântico *jogos e diversões infantis*. Por isso, os dicionários deveriam incluir a variante e indicar que se trata de um regionalismo específico de Fortaleza.

Diante desses dados, observa-se que os dicionários, ainda que registrem a maioria das variantes, não trazem a acepção referente à brincadeira infantil. Em poucos casos em que há essa menção, os dados geográficos são divergentes.

A seguir, passe à análise de variantes de *cabra-cega*.

Quadro 8: Variantes de *cabra-cega* (brincadeira)

REGIONALIMOS	Dados no <i>ALiB</i>	Marcação no <i>DEH</i>	Marcação no <i>DEA</i>	Confronto entre os materiais
Cabra-cega	Todas as capitais– 100%	Não há	Não há	-

	em Recife e Maceió			
Cobra-cega	Norte (Rio Branco, Macapá e Belém); Nordeste (Teresina, Fortaleza, Natal, João Pessoa, Aracaju e Salvador); Sudeste (todas as capitais); Sul (Curitiba e Porto Alegre); Centro-Oeste (todas as capitais)	Não há	Não há	-
Pata-cega	Norte (exceto Rio Branco); Nordeste (São Luís); Sul (Florianópolis)	<i>Variante não registrada</i>	<i>Variante não registrada</i>	-
Pega-pega	Norte (Macapá); Nordeste (Teresina); Sul (Florianópolis)	Não há	Brasileirismo	Divergente (para o <i>DEA</i>)

Fonte: elaboração própria com base nos dados do *ALiB*, do *DEH* e do *DEA*

Há quatro variantes para *cabra-cega*. Esta é utilizada em todas as capitais do Brasil investigadas pelo *ALiB* e tem ocorrência de 100% em Recife e em Maceió (Nordeste). Os dicionários em análise registram a variante, mas não trazem indicação de localização geográfica. Com o *ALiB*, o *DEH* e o *DEA* poderiam informar que se trata de um *brasileirismo*.

Outra variante é *cobra-cega*, que também é bastante ocorrente. O *ALiB* registra usos no Norte (Rio Branco, Macapá e Belém); Nordeste (Teresina, Fortaleza, Natal, João Pessoa, Aracaju e Salvador); Sudeste (todas as capitais); Sul (Curitiba e Porto Alegre); Centro-Oeste (todas as capitais). Sendo registrada no *DEH* e no *DEA*, os verbetes dessa variante não contêm acepção referente ao campo semântico *jogos e diversões infantis*. Eis mais um caso em que os dicionários devem incluir a acepção que remeta à brincadeira infantil, como também informar as capitais em que a variante é utilizada.

A variante *pata-cega*, que ocorre em capitais do Norte, exceto Rio Branco, São Luís, no Nordeste, e Florianópolis, no Sul, não é registrada em nenhum dos dicionários analisados, o

que poderia ser feito pelas duas obras. Para a última variante, *pega-pega*, que, segundo o *ALiB*, se apresenta apenas em Macapá (Norte), Teresina (Nordeste) e Florianópolis (Sul), não há indicação de localização geográfica no *DEH*. Já no *DEA*, a acepção 5 do verbete *pega-pega*, na qual se encontra a indicação de *brasileirismo*, há remissão para o verbebo *pique*². Neste verbebo, a terceira acepção refere-se a jogo infantil. Ao tratá-lo como *brasileirismo*, o *DEA* diverge do *ALiB*.

Passa-se, então, para a análise das variantes de *amarelinha*.

Quadro 9: Variantes de *amarelinha*

REGIONALIMOS	Dados do <i>ALiB</i>	Marcação no <i>DEH</i>	Marcação no <i>DEA</i>	Confronto entre os materiais
Amarelinha	Todas as capitais	Brasil	Brasileirismo	Convergente
Academia	Nordeste (Natal, João Pessoa e Recife)	Rio Grande do Norte	Nordeste	Divergente
Cancão	Nordeste (Maranhão e Teresina)	Não há	Não há	-
Avião	Nordeste (Maceió)	Não há	Não há	-
Macaca	Norte (todas as capitais); Nordeste (Fortaleza)	Não há	Lusitanismo	Divergente (para o <i>DEA</i>)
Macacão	Nordeste (Aracaju)	Não há	Não há	-
Macaco	Nordeste (Salvador)	Brasil	Brasileirismo	Divergente

Fonte: elaboração própria com base nos dados do *ALiB*, do *DEH* e do *DEA*

A variante *amarelinha* está presente em todas as capitais e tem 100% de ocorrência em Cuiabá, Campo Grande, São Paulo, Florianópolis, Rio de Janeiro e Vitória. Logo, trata-se de um *brasileirismo*. Quanto aos dicionários, tanto o *DEH* como o *DEA* classificam tal variante como *brasileirismo*, o que converge com os dados dialetológicos. Vale ressaltar que as duas obras lexicográficas apresentam duas entradas para *amarelinha*, mas a segunda é a que se refere à brincadeira infantil.

A variante *academia* é tratada pelo *DEH* como sendo regionalismo do Rio Grande do Norte, enquanto que o *DEA* indica a região Nordeste. Nota-se que os dois tratamentos são problemáticos: tendo em vista que *academia* ocorre em apenas três capitais do Nordeste, classificá-lo apenas como do Rio Grande do Norte é desconsiderar a ocorrência em outras

capitais; classificá-lo como do Nordeste é generalizar um uso. Por isso, os dicionários podem indicar, precisamente, os locais em que a variante ocorre: Natal, João Pessoa e Recife.

Cancão é uma variante que só ocorre em São Luís e em Teresina. Trata-se, pois, de um regionalismo maranhense e piauiense. Nenhum dos dois dicionários traz qualquer acepção que remeta à brincadeira infantil. Assim, seria adequado inserir a acepção em estudo, como também informar que o uso é restrito às capitais São Luís e Teresina.

Avião é outra variante que só ocorre em Maceió. Sendo assim, os dicionários poderiam registrá-lo como um regionalismo maceioense. Contudo, os dicionários, apesar de registraram a forma *avião*, não inserem uma acepção que se remete à brincadeira infantil.

O caso de *macaca* é interessante. O *ALiB* atesta que tal variante está presente em todas as capitais do Norte e em Fortaleza, no Nordeste. Há a possibilidade de classificar o regionalismo como do Norte e de Fortaleza, especificamente. O *DEH* contém duas entradas para *macaca*, sendo a primeira indicada com várias acepções, mas nenhuma é sinônimo de *amarelinha*, logo não há marcação de regionalismo para essa acepção. O *DEA*, por seu turno, traz acepção referente à *amarelinha* e a classifica como *lusitanismo*. O termo *lusitanismos* se refere a variedades de português de Portugal em oposição ao português do Brasil. Contudo, é inadequado classificar como *lusitanismo* uma variante que tem ocorrência em várias capitais do Brasil. Por isso, o *DEA* diverge do *ALiB*.

A variante *macacão* só ocorre em Aracaju, no Nordeste. Os dois dicionários em análise registram essa forma, mas não contêm acepção referente à brincadeira infantil, como também não trazem localização geográfica. Um tratamento adequado no *DEH* e no *DEA* seria, primeiramente, inserir a acepção referente à brincadeira infantil e, em seguida, indicar que se trata de um regionalismo aracajuense.

Por fim, a variante *macaco* é registrada apenas em Salvador, no Nordeste, segundo o *ALiB*. O *DEH* e o *DEA* classificam a variante como *brasileirismo*, divergindo mais uma vez com os dados dialetológicos. Seria mais adequado classificá-lo como um regionalismo salvadoreense.

Diante da análise feita, foi possível notar que certos regionalismos recebem tratamento diferenciado, pois alguns deles são registrados nos dicionários, mas não possuem acepção referente a campo semântico *Jogos e diversões infantis*, outros não trazem localização geográfica e outros nem chegam a ser incluídos na nomenclatura.

Quanto aos casos em que não houve a acepção desejada, é possível perceber o problema no tratamento não só geográfico, mas também semântico dos regionalismos. Haensch (1982), falando sobre as diferenças diatópicas em dicionários, adverte que:

para registrar lexicograficamente estas diferenças, não basta apenas limitar-se a lexemas que existem em um território A e que não existem no território B, mas deveriam enfatizar, atualmente, as características semânticas que têm de mais ou de menos os diferentes lexemas. (HAENSCH, 1982, p. 382).

Quando os dicionários privilegiam determinadas acepções em detrimento de outras, eles também cumprem seu papel *normativo*, ao estabelecer que as variantes só significam aquilo que está posto na definição, deixando de incluir os usos que remetem à brincadeira infantil.

Levando em conta as análises feitas e a maneira como estão registrados (ou não) os regionalismos nos dicionários, foi possível verificar diferentes formas de tratamento das variantes, quais sejam:

- a) Não registro de variantes;
- b) Registro de variantes, mas sem acepção que se refira a jogo, brinquedo ou brincadeira;
- c) Registro de variantes, mas sem localização geográfica;
- d) Registro de variantes com localização geográfica divergente do *ALiB*;
- e) Registro de variantes com localização geográfica convergente ao *ALiB*.

Partindo da classificação acima, foi possível dividir as variantes analisadas em grupos que caracterizam a forma de tratamento dos dicionários. Com isso, é possível verificar qual forma de tratamento é mais recorrente. A seguir, tem-se o quadro que organiza as variantes nessa classificação:

Quadro 10: Classificação das variantes

	<i>DEH</i>	<i>DEA</i>
Variantes não registradas	Bila Bolachinha Bolinha de vidro Bunda-canastra Cabeçulinha Carambela Maria-escambona Pata-cega	Bolachinha Bolinha de gude Bunda-canastra Carambela Maria-escambona Pata-cega

<p>Variantes registradas sem aceção que se refira a jogos ou diversão infantil</p>	<p>Avião Cambona Cancão Cobra-cega Coruja Curica Macaca Macacão Mortal Periquito Ratinho</p>	<p>Avião Cancão Carambola Cobra-cega Coruja Estilete Macacão Mortal Periquito Ratinho</p>
<p>Variantes registradas sem localização geográfica</p>	<p>Bola de fona Cabra-cega Cabriola Cambalhota Cambona Carambola Estilete Marraio Mortal Papagaio Pega-pega Pipa Pirueta</p>	<p>Cabra-cega Cabriola Cambalhota Cambona Carambola Mortal Papagaio Pirueta</p>
<p>Variantes registradas com localização geográfica divergente ao ALiB</p>	<p>Academia Atiradeira Aú Badogue Baladeira Bila Biloca Bola de fona Cambota Cangapé Macaco Peteca Raia 1 Raia 2</p>	<p>Academia Atiradeira Aú Badogue Baladeira Bila Biloca Bola de fona Bolinha de vidro Cabeçulinha Cambota Cangapé Curica Estilingue Macaca Macaco Marraio Peteca Peteca Raia 2 Raia 1</p>

Variantes registradas com localização geográfica convergente ao ALiB	Amarelinha Bola de gude Bolinha de gude Estilingue Ximbra	Amarelinha Bola de gude Pipa Ximbra
---	---	--

Fonte: elaboração própria

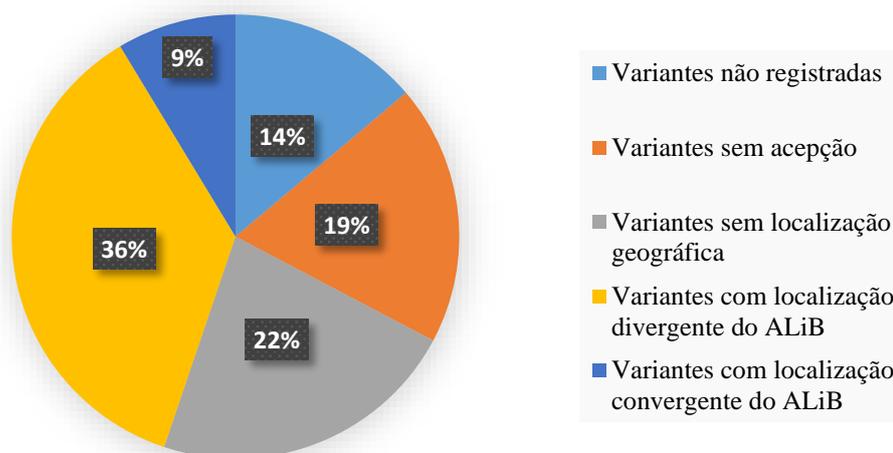
Diante do quadro acima, torna-se possível identificar o tratamento de cada variante. Só para citar alguns casos, a variante *bila*, por exemplo, a primeira que aparece na primeira linha do quadro, na coluna do *DEH*, está no grupo das variantes que não são registradas no *DEH*, mas que é registrada no *DEA*. Variantes como *cambona*, *avião*, *cancão* estão entre as que são registradas, mas sem aceção de remeta a jogo, brinquedo ou brincadeira. Variantes como *bola de fona*, *cabra-cega* e *marraio* estão entre as que não possuem localização geográfica no verbete. *Academia*, *atiradeira* e *aú* têm localização geográfica divergente do *ALiB*, e *amarelinha*, *bola de gude* e *ximbra* têm localização geográfica convergente do *ALiB*, só para citar alguns casos.

Além disso, o que chama atenção, no quadro acima, é o grupo com o maior número de variantes, que corresponde às **variantes registradas com localização geográfica divergente ao ALiB**. Comparando os dois dicionários, é possível verificar que o maior número dessas variantes está no *DEA*. Sendo assim, constata-se que, entre as variantes analisadas, a maioria tem localização geográfica divergente do *ALiB* e essa divergência predomina no *DEA*, o que confirma a hipótese inicial desta pesquisa. Dessa maneira, a forma de tratamento das variantes mais recorrente entre as listadas anteriormente é a que faz o registro das variantes com localização geográfica divergente.

Outro dado que chama atenção é o último grupo de variantes que compõe o quadro, que consiste nas **variantes com localização geográfica convergente ao ALiB**. Esse grupo contém a minoria das variantes analisadas, tanto no *DEH* quanto no *DEA*. Esse dado evidencia que poucos são os casos em que os dicionários trazem informações diatópicas semelhantes aos dados dialetológicos.

Representando essa classificação em gráfico, torna-se mais visível a distribuição das variantes, em proporções mais ou menos definidas:

Gráfico 1: Proporção da classificação das variantes



Fonte: Elaboração própria

O gráfico permite evidenciar a proporção da classificação das variantes proposta acima. Segundo os dados, a maioria das variantes investigadas (36%) tem localização geográfica divergente do *ALiB*. É o caso, por exemplo, de *cambota*, *cangapé*, *aú* etc. (cf. Quadro 10). A segunda maior ocorrência (22%) é das variantes que não possuem localização geográfica no dicionário. Os verbetes dessas variantes não trazem qualquer marca diatópica. Nesse grupo estão *carambola*, *estilete*, *mortal*, *papagaio*, para citar alguns.

Além disso, o gráfico aponta que 19% das variantes investigadas e que estão registradas nos dicionários não tem aceção que remeta à ludologia, como é o caso de *avião*, *cambona*, *cancão*, *mortal* etc. Ademais, 14% das variantes investigadas não estão registradas no dicionário (*bolachinha*, *bunda-canastra*, *carambela* etc.). E, por fim, a menor porcentagem (9%) representa as variantes com marca diatópica convergente ao *ALiB*. Nessa última porcentagem estão *amarelinha*, *bola de gude*, *pipa*, *estilingue* e *ximbra*.

Diante desses dados, é possível concluir que os regionalismos não são tratados da mesma forma nos dicionários e, por isso mesmo, há a necessidade de uniformização desse tratamento.

CONCLUSÃO

A pesquisa, aqui realizada, buscou estabelecer uma interface entre a Lexicografia, a Metalexigrafia e a Dialetoлогия. Isso porque foram analisados os dicionários *DEH* e *DEA*, tomando como base dados dialetológicos, em especial os dados do *ALiB*. Retomando os objetivos do trabalho, o principal interesse da pesquisa foi investigar o tratamento de regionalismos do campo semântico *Jogos e diversões infantis* do *ALiB* nos dicionários *DEH* e *DEA*.

Para atingir os objetivos do trabalho, foi necessário partir de discussões a respeito da variação do léxico e como essa variação se constituiu no PB, o que se fez no capítulo 1. Essa base teórica permitiu entender a língua como um conjunto de variedades que se constitui a partir de culturas diversas.

No capítulo 2, discutiu-se sobre as ciências que se ocupam em elaborar e analisar dicionários – Lexicografia e Metalexigrafia. Fez-se uma reflexão sobre a produção

lexicográfica ao longo dos séculos, bem como do desenvolvimento da ciência lexicográfica. Além disso, discutiu-se sobre o papel do dicionário no registro de regionalismos. Com isso, concluiu-se que os dicionários gerais podem incluir os regionalismos, mas, em alguns casos, há incoerências no tratamento deles. Foi a partir dessa discussão que se pôde justificar a realização deste trabalho.

O capítulo 3 também forneceu embasamento teórico para se entender em que consiste o trabalho dialetológico e como ele pode contribuir para a Lexicografia. Observou-se, pois, que os atlas linguísticos são ferramentas indispensáveis para classificar um regionalismo, e, por isso, podem os lexicógrafos se valer desses atlas para tratar mais adequadamente regionalismos em dicionários.

Com as análises, percebeu-se que há diferentes formas de tratamento de regionalismos nos dicionários e suas variantes, que são i) não registro de variantes; ii) registro de variantes, mas sem acepção que se refira a jogo, brinquedo ou brincadeira; iii) registro de variantes, mas sem localização geográfica; iv) registro de variantes com localização geográfica. Diante dessas diferentes formas de tratamento, observa-se que os dicionários não são uniformes quando ao registro de regionalismos. Por isso, é necessário que os dicionários estabeleçam critérios de seleção e registro de regionalismos. No caso das variantes presentes no *ALiB*, mas não registradas no dicionário, os lexicógrafos podem incluí-las na nomenclatura.

Nas variantes que estão registradas, mas que não possuem acepção que remeta ao campo semântico analisado, é necessário que se insira o campo semântico que se refira a jogo, brinquedo ou brincadeira, a depender do regionalismo. Quanto às variantes registradas nos dicionários, mas sem marca diatópica, pode-se inserir essa informação, baseando-se, principalmente, no *ALiB* ou em atlas regionais, conforme propõe Isquierdo (2007) e Fajardo (1996-1997).

Em relação ao registro de regionalismos com localização geográfica, objetivou-se estabelecer um confronto com as informações dos dicionários e do *ALiB* sobre o local em que tais regionalismos são utilizados. A hipótese inicial era a de que haveria muitas divergências entre os dicionários e o *ALiB*. Tal hipótese foi confirmada, haja vista que a maioria das variantes investigadas nos dicionários tem localização geográfica diferente do que informa o *ALiB*, principalmente no *DEA*. Ainda que se tome o *ALiB* como padrão para informar a localização geográfica do regionalismos, reconhece-se que ele não é o único instrumento que pode fornecer dados sobre a natureza de regionalismos, mas também *corpora* de língua escrita, dados colhidos

de atos reais de fala e glossários e vocabulários regionais resultantes de pesquisas acadêmicas, conforme Castro, Cardoso e Isquardo (2011).

Mesmo com esses problemas de tratamento de regionalismos identificados nesses dicionários, não se pode desprezar a importância que tais obras possuem. Tanto o *DEH* quanto o *DEA* estão entre os maiores dicionários de língua portuguesa e foram elaborados por uma equipe composta por pesquisadores de várias áreas e cada uma delas se valeu de um *corpus* e uma bibliografia específica. Sendo assim, é possível que os lexicógrafos não tenham tido acesso a fontes que registrem as variantes aqui analisadas, o que justificaria o não registro de algumas delas. Ademais, como já se discutiu no início desse trabalho, o léxico é um campo aberto e em constante expansão, por isso, é justificável que haja lacunas nos dicionários quanto ao uso dos regionalismos e quanto às marcas diatópicas.

Esse trabalho não esgota as possibilidades de análise de regionalismos em dicionários, mas suscita novas discussões a respeito do tema. Acredita-se que, com essas reflexões, é possível pensar em um tratamento mais adequado de regionalismos, sejam em dicionários gerais ou em dicionários especializados em regionalismos, trazendo informações geográficas que estejam atestadas em atlas linguísticos.

REFERÊNCIAS

ALVES, Ieda Maria. Neologia e dicionários escolares. In: CARVALHO, Orlene Lúcia de Sabóia; BAGNO, Marcos. **Dicionário escolares: políticas, formas e usos.** São Paulo: Parábola Editorial, 2011.

ARAGÃO, Maria do Socorro Silva de. O *Atlas Linguístico do Brasil* no quadro da geolinguística brasileira. In: CARDOSO, Suzana Alice Marcelino da Silva *et alii*. **Atlas Linguístico do Brasil: introdução.** v. 1. Londrina: EDUEL, 2014, p. 31-36.

AUROUX, Sylvain. **A revolução tecnológica da gramatização.** 2ª ed. Campinas, SP: Editora da Unicamp, 2009.

BARBOSA, Maria Aparecida. Lexicologia, lexicografia, terminologia e terminografia: objeto, métodos, campos de atuação e de cooperação. In: **Estudos Linguísticos.** Franca: Unifran/GEL, p. 182-189, 1991.

BIDERMAN, Maria Teresa Camargo. A ciência da lexicografia. **Alfa**, São Paulo, p. 1-26, 1984a.

_____. O dicionário padrão da língua. **Alfa**, São Paulo, p. 27-43, 1984b.

_____. A definição lexicográfica. In: Cadernos do IL, Porto Alegre, Instituto de Letras da UFRGS, n. 10, p. 23-43, jul. 1993.

_____. A nomenclatura de um dicionário de língua. **Anais de Seminários do GEL**, São Paulo: v.1, n.23, 26-42, 1994.

_____. O dicionário como norma na sociedade. IN: 1º encontro Nacional do GT de Lexicologia, Lexicografia e Terminologia, 1998, Rio de Janeiro. **Anais do 1º Encontro Nacional do GT de Lexicologia, Lexicografia e Terminologia**. Recife: Editora Universitária UFPE, v. 1, p. 161-180, 1997.

_____. **Teoria linguística: teoria lexical e linguística computacional**. 2ª ed. São Paulo: Martins Fontes, 2001.

_____. Dicionários do português: da tradição à contemporaneidade. **Alfa**, 47(1), p. 53-69, 2003.

BORBA, Francisco da Silva; VILLAR, Mauro de Salles. O trabalho do dicionarista. In: XATARA, Claudia et al.(orgs.). **Dicionários na teoria e na prática: como e para que são feitos**. São Paulo: Parábola Editorial, 2011.

CAMARA JÚNIOR, Joaquim Mattoso. **Problemas de linguística descritiva**. 17. ed. Petrópolis, RJ: Vozes, 1998.

_____. **Estrutura da língua portuguesa**. 42. ed. Petrópolis, RJ: Vozes, 2009.

_____. **Dicionário de linguística e gramática: referente à língua portuguesa**. 28. ed. Petrópolis, RJ: Vozes, 2011.

CARDOSO, Suzana Alice. **Geolinguística: tradição e modernidade**. São Paulo: Parábola Editorial, 2010.

_____; CASTRO, Yêda Pessoa de.; ISQUERDO, Aparecida Negri. Quais critérios que deveriam orientar os lexicógrafos na inserção de lusitanismos, africanismos e regionalismos em dicionários gerais? In: XATARA, Claudia et al.(orgs.). **Dicionários na teoria e na prática: como e para que são feitos**. São Paulo: Parábola Editorial, 2011.

DIETRICH, Wolf; NOLL, Volker. O papel do tupi na formação do português brasileiro. In: _____. **O Português e o Tupi no Brasil**. São Paulo, Contexto, 2010.

EDELWEISS, Frederico G. **Estudos tupis e tupi-guaranis: confrontos e revisões**. Livraria Brasileira Editora: Rio de Janeiro, 1969. Disponível em: http://biblio.wdfiles.com/local--files/edelweiss-1969-123-133/edelweiss_1969_123-133.pdf. Acesso: ago. 2016.

ELIA, Sílvio. **Orientações da linguística moderna**. 2ª ed. Rio de Janeiro: Ao Livro Técnico, 1978.

ESCRIBANO, Cecilio Garriba. La microestructura del diccionario: las informaciones lexicográficas. In: GUERRA, Antonia María Medina. **Lexicografía española**. Barcelona: Editorial Ariel S.A., 2003.

FAJARDO, Alejandro. Las marcas lexicográficas: concepto y aplicación práctica en la lexicografía española. **Revista de Lexicografía**. Volumen III, p. 31-57, 1996-1997.

FARACO, Carlos Alberto. **Norma culta brasileira**: desatando alguns nós. São Paulo: Parábola Editorial, 2008.

_____. **História sociopolítica da língua portuguesa**. São Paulo: Parábola Editorial, 2016.

FAUSTO, Boris. **História concisa do Brasil**. São Paulo: Edusp/Imprensa Oficial do Estado, 2001.

GONÇALVES, Sheila de Carvalho Pereira. Metalexigrafia escolar no Brasil: história e contribuições. In: **Entreletras**, Araguaína/TO, v. 4, n. 2, p. 168-186, ago/dez. 2013.

GUERRA, A. M. M. (coord.). **Lexicografía española**. Barcelona: Editorial Ariel S. A., 2003.

HAENSCH, Günther. Aspectos prácticos de la elaboración de diccionario. In: _____; WOLF, Lothar; ETTINGER, Stefan; WERNER, Reinhold. **La lexicografía**: de la lingüística teórica a la lexicografía práctica. Madrid: Editorial Gredos, 1982.

_____. Tipología de las obras lexicográficas. In: _____; WOLF, Lothar; ETTINGER, Stefan; WERNER, Reinhold. **La lexicografía**: de la lingüística teórica a la lexicografía práctica. Madrid: Editorial Gredos, 1982.

HOUAISS, Antônio; VILLAR, Mauro de Salles; FRANCO, Francisco Manoel de Melo. **Dicionário Eletrônico Houaiss da língua portuguesa**. Rio de Janeiro: Editora Objetiva, 2009.

ISQUERDO, Aparecida Negri. Achegas para a discussão do conceito de regionalismos no português do Brasil. In: **Alfa**. São Paulo, 2006, p. 9-24.

_____. A propósito de dicionários de regionalismos do português do Brasil. In: _____; ALVES, Ieda Maria. **As ciências do léxico**: lexicologia, lexicografia e terminologia. v. 3. Campo Grande, MS: Ed. UFMS; São Paulo: Humanitas, 2007.

KISHIMOTO, Tizuko Morchida. O jogo e a educação infantil. In: _____(org.). **Jogo, brinquedo, brincadeira e a educação**. 11. ed. São Paulo: Cortez, 2008, p. 13-43.

KRIEGER, Maria da Graça. Políticas públicas e dicionários para escola: o programa nacional do livro didático e seu impacto sobre a lexicografia didática. In: **Cadernos de Tradução**: UFSC, v. 2, n. 18, 2006.

_____. Tipologias de dicionários: registros de léxico, princípios e tecnologias. In: **Colidoscópio**. Vol. 4, n. 3, p. 141-147, set/dez, 2006.

LUCCHESI, Dante. História do contato entre as línguas do Brasil. In: _____; BAXTER, Alan; RIBEIRO, Ilza. **O português afro-brasileiro**. Salvador: EDUFBA, 2009.

_____. **Língua e sociedade partidas**: a polarização sociolinguística do Brasil. São Paulo: Contexto, 2015.

MATTOS e SILVA, Rosa Virgínia; MACHADO FILHO, Américo Venâncio Lopes. Entre duas diásporas: o português e as línguas africanas no Brasil. In: OLIVEIRA, Klebson; CUNHA E SOUZA, Hirão F.; SOLEDADE, Juliana (orgs.). **Do português arcaico ao português brasileiro**: outras histórias. Salvador: EDUFBA, 2009. p. 297-304.

MOTA, Jacyra Andrade. Percursos metodológicos: questionários e informantes. In: CARDOSO, Suzana Alice Marcelino da Silva *et alii*. **Atlas Linguístico do Brasil**: introdução. v. 1. Londrina: EDUEL, 2014.

MÜLLER, Alexandra Feldekircher *et. al*. A lexicografia brasileira do século XX: dicionários inaugurais e temáticas. In: **Cadernos do CNLF**, Vol. XIII, nº 04, 2009.

NARO, Anthony Julius; SCHERRE, Maria Marta Pereira. **Origens do português brasileiro**. São Paulo: Parábola Editorial, 2007.

NOLL, Volker. O Brasil Colônia entre a língua geral e o português. In: _____; DIETRICH, Wolf (orgs.). **O Português e o Tupi no Brasil**. São Paulo, Contexto, 2010.

NUNES, José Horta. **Dicionários no Brasil**: análise e história do século XVI ao XX. Campinas, SP: Pontes Editores – São Paulo, SP: Fapesp – São José do Rio Preto, SP: Faperp, 2006.

_____. Dicionários: história, leitura e produção. In: **Revista de Letras**. Brasília, UCB, v. 3, n. ½, 2010.

OLIVEIRA, A. M. P. P. de. **O português do Brasil**: brasileirismos e regionalismos. 1999. Tese (Doutorado em Lingüística e Língua Portuguesa) – Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara, 1999.

ORLANDI, Eni P. **Língua e conhecimento linguístico**: para uma história das ideias no Brasil. 2. ed. São Paulo: Cortez, 2013.

PORTILHO, Danyelle Almeida Saraiva. **O falar amazônico**: uma análise da proposta de Nascentes (1953) a partir de dados do Projeto ALiB. 2013. 155 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2013.

RIBEIRO, Silvana Soares Costa. **Brinquedos e brincadeiras infantis na área do falar baiano**. 2012. 752f. Tese (Doutorado do Programa de Pós-graduação em Letras e Linguística) - Instituto de Letras, Universidade Federal da Bahia, Salvador, 2012.

RODRIGUES, Aryon Dall’Igna. Tupi, tupinambá, línguas gerais e português do Brasil. In: NOLL, V.; DIETRICH, W. (Orgs.). **O Português e o Tupi no Brasil**. São Paulo, Contexto, 2010.

SEABRA, Maria Cândida Trindade Costa de; WELKER, Hebert Andreas. Questões teóricas genéricas. In: XATARA, Claudia et al.(orgs.). **Dicionários na teoria e na prática**: como e para que são feitos. São Paulo: Parábola Editorial, 2011.

VERDELHO, Telmo. **As origens da gramaticografia e da lexicografia latino-portuguesas**. Coimbra: Instituto Nacional de Investigação Científica, 1^a. ed. 1995.

VILLALVA, Alina; SILVESTRE, João Paulo. **Introdução ao estudo do léxico**: descrição e análise do português. Petrópolis, RJ: Vozes, 2014.

WELKER, Hebert Andreas. **Dicionários**: uma pequena introdução à lexicografia. 2. ed. revista e ampliada. Brasília: Thesaurus, 2004.

_____. Breve histórico da metalexiconografia no Brasil e dos dicionários gerais brasileiros. In: **MATRAGA**. Rio de Janeiro, ano 13, n.19, p.69-84, 2006.

WERNER, Reinhold. La unidade léxica y el lema. In: HAENSCH, Günther; WOLF, Lothar; ETTINGER, Stefan; WERNER, Reinhold. **La lexicografía**: de La lingüística teórica a la lexicografía práctica. Madri: Editorial Gredos, 1982.

ZAVAGLIA, Claudia. Metodologia em ciências da linguagem: lexicografia. In: GONÇALVES, Adair Vieira; GÓIS, Marcos Lúcio de Sousa (orgs.). **Ciências da linguagem**: o fazer científico. 1. ed. Campinas: Mercado de Letras, 2012, v. 1, p. 231-264.