

PINGUELA OU PONTE? CÓRREGO OU RIACHO? VARIAÇÃO SEMÂNTICO-LEXICAL NOS DADOS DO ATLAS LINGUÍSTICO DO BRASIL (ALiB)

PINGUELA OR PONTE? CÓRREGO OR RIACHO? SEMANTIC-LEXICAL VARIATION IN THE DATA OF THE LINGUISTIC ATLAS OF BRASIL (ALiB)

Genivaldo da Conceição Oliveira¹
Universidade Federal do Recôncavo da Bahia

Resumo: Este trabalho apresenta os resultados da análise das denominações registradas, nos estados da Bahia e do Paraná, no campo semântico *acidentes geográficos* nas cidades que constituem a rede de pontos do Atlas Linguístico do Brasil (ALiB) nesses dois estados. Baseamos nosso estudo nos princípios teóricos da Dialetoлогия, Sociolinguística, Lexicologia e Lexicografia. O *corpus* é constituído das perguntas 1 a 4 do Questionário Semântico-Lexical (QSL) e se utiliza do método da Geolinguística para a análise espacial dos dados. Enfatizamos o aspecto diatópico, contudo, recorreremos, de maneira periférica, à análise de outras variáveis como a diastrática e a diageracional.

Palavras-Chave: Dialetoлогия; Geolinguística; Português Brasileiro; Variação.

¹ valdooli@yahoo.com.br

Abstract: *This study presents the results of the analysis of names registered in the states of Bahia and Paraná presented in the cities that make up the network of localities of the Linguistic Atlas of Brazil (ALiB – Atlas Linguístico do Brasil) in both states. We base our studies on theoretical principles of Dialectology, Sociolinguistics, Lexicology and Lexicography. The corpus consists of questions 1 to 4 of the Semantic-Lexical Questionnaire (QSL – Questionário Semântico-Lexical). We use the Geolinguistic method for the spatial analysis of the data. We focus on the diatopic aspect. However, we turn, in a peripheral way, to the analysis of other variables such as diastratic and diagenational.*

Keywords: *Dialectology; Geolinguistics; Brazilian Portuguese; Variation.*

INTRODUÇÃO

A identidade e a cultura de uma comunidade estão sintonizadas com a linguagem utilizada por esta comunidade. Esta linguagem, além de ser um objeto desta cultura, é a principal ferramenta para sua transmissão. A língua é um instrumento por meio do qual se dão as relações sociais entre habitantes do mesmo local e de locais diferentes, ou seja, dentro da mesma coletividade ou entre povos distintos. A linguagem, na fala ou na escrita, representa a nossa única fonte, por meio da qual temos acesso à realidade imaterial que é a língua. Biderman (2001, p. 3) afirma que "nosso conhecimento sobre a estrutura e a realidade linguísticas será sempre precário porque a língua é uma realidade mental que, nos seus limites, se confunde com o próprio pensamento". Portanto, a língua desempenha um papel importante na sociedade uma vez que é por meio dela que as pessoas se comunicam, expressam seus pensamentos e sentimentos.

Desse modo, investigamos, nesta pesquisa, a relação entre o léxico referente à área semântica acidentes geográficos documentada no estado da Bahia e no estado do Paraná, com base no *corpus* do Atlas Linguístico do Brasil (ALiB). Utilizamos os dados de 164 informantes, distribuídos: 16 nas duas capitais, 84 nas cidades do interior da Bahia e 64 nas cidades do interior do Paraná. Objetivamos, neste trabalho, identificar e estabelecer semelhanças e diferenças encontradas neste recorte estudado nos dois estados com base teórica

na abordagem apresentada pela Geolinguística e nos valem de pressupostos teóricos Dialetologia, Lexicologia, Sociolinguística e Lexicografia.

1 FUNDAMENTAÇÃO TEÓRICA

Descoberto em 1500, o Brasil só começou propriamente a ser colonizado em 1534 com o regime das capitanias hereditárias. A partir desta data tem-se a efetiva introdução da língua portuguesa.

A língua portuguesa no Brasil, no curso da sua história, vem passando por muitas modificações, realidade para a qual, nos meados do século XX, já chamava a atenção de Nascentes (1953) ao registrar que:

A língua portuguesa no Brasil sofreu grandes alterações porque teve de ser aprendida por homens de duas raças que falavam línguas de estrutura inteiramente diversa do tipo flexional. O índio foi o primeiro que aprendeu o português; é natural, pois foi o povo autóctone. Só mais tarde aparece o outro fator etnográfico, o negro. [...] A escravidão vermelha precedeu à negra e daquela já se fala em 1531, quando Martim Afonso concedeu a Pedro de Góis permissão de levar para Europa dezessete escravos índios; mas desde cedo na capitania de S. Vicente são escravos negros que trabalham na agricultura da cana. (NASCENTES, 1953, p. 9-10).

O falar brasileiro, apesar de sua relativa uniformidade, apresenta diversidade linguística e os estudos dialetológicos, calcados na perspectiva da Geografia Linguística, ajudam-nos a entender esta diversidade.

1.1 Estudos dialetológicos – um breve histórico

A Dialetologia identifica, situa e descreve os usos diferentes em que a língua varia de acordo com sua disposição espacial, histórica e sociocultural, respondendo a um pensamento mais amplo, pois, como afirma Cardoso (2010, p. 27),

O interesse pelo estudo da diversidade de usos da língua e a evidência de certa preocupação universal com as diferenças dialetais perpassam a história

dos povos em todos os momentos, ora como simples constatação, ora como instrumento político, ora como mecanismo de descrição das línguas.

A Dialetoлогия é, como a própria palavra sugere, o estudo dos dialetos. Contudo, precisamos definir o que é dialeto. Na linguagem coloquial, diz-se que um dialeto é uma língua não padrão, que geralmente se associa a grupos que não ostentam prestígio ou uma língua que não tem tradição escrita. Entretanto, Chambers e Trudgill (1994, p. 19) não aprovam estes conceitos e defendem que “todos os falantes são falantes de, pelo menos, um dialeto”² e que a forma padrão de uma língua constitui-se em si um dialeto. Esta definição, contudo, esbarra em um problema: Como distinguir língua de dialeto? Esses autores afirmam que uma língua é um conjunto de dialetos mutuamente inteligíveis, embora salientem que essa definição não seja totalmente satisfatória porque esta tal inteligibilidade nos traz alguns problemas. Eles citam as línguas norueguesa, sueca e dinamarquesa como três línguas distintas, mas mutuamente inteligíveis. Em contrapartida, a língua alemã, considerada como uma única língua, apresenta problemas de comunicação entre os diferentes falantes do alemão, usuários de distintos dialetos. A inteligibilidade entre as línguas escandinavas apresenta graus para mais ou para menos e pode não se apresentar igual nas duas direções. Segundo Chambers e Trudgill (1994), os dinamarqueses, por exemplo, entendem melhor os noruegueses do que os noruegueses conseguem entender os dinamarqueses. Neste sentido, comparamos o português brasileiro e o espanhol falado na América Latina. É muito comum escutar um falante do espanhol reclamar que não entende o brasileiro ao passo que se escuta um brasileiro dizer que entende claramente o espanhol. Dessa maneira, é possível dizer que dialeto é uma variedade linguística baseada em fatores geográficos com formas sintáticas, aspectos fonético-fonológicos e itens lexicais distintos.

² “... todos los hablantes lo son al menos de un dialecto...” Tradução nossa.

A Dialetoologia era vista nos seus primórdios sob uma perspectiva preponderantemente diatópica. Os primeiros estudos dialetológicos eram focalizados dentro de um espaço e tinham uma abordagem monodimensional. Aos poucos, a monodimensionalidade foi perdendo sua hegemonia para um estudo mais pluridimensional. De acordo com Cardoso (2010, p. 15), “a Dialetoologia 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”. O estudo dialetológico obedece a três passos importantes: identificar, descrever e situar a variação linguística. Após a realização do primeiro passo, que é a identificação do fenômeno linguístico, passa-se a descrevê-lo fazendo o levantamento das variantes. Descrever é enumerar as variantes lexicais possíveis e que tenham o mesmo valor de verdade, como em *macaxeira*, *aipim* e *mandioca*. A Dialetoologia inventaria, sistematiza e descreve estas variações.

O espaço físico exhibe variedades linguísticas que ocorrem de uma região para outra. O interesse por este tipo de informação não está apenas em catalogar dados intercomparáveis, mas também em registrar a ausência de tais dados. Podemos, então, dizer que são intercomparáveis tanto os dados presentes em uma região e outros presentes em outras, como também os existentes em uma região comparados à sua ausência em outra.

Com base em Cardoso (2010), observamos que há duas características importantes na origem da Dialetoologia independentemente do princípio metodológico usado. A primeira é o reconhecimento das diferenças ou das semelhanças que a língua transmite. A segunda é o estabelecimento das relações entre as diversas manifestações linguísticas documentadas entre elas e a ausência de dados registrados, circunscritos a espaços e realidades prefixados. A Dialetoologia não pode ignorar fatores extralinguísticos, próprios do falante, da mesma maneira que não pode desprezar as implicações que estes fatores

acarretam nos atos da fala. Portanto, idade, sexo, escolaridade e características socioculturais tornam-se elementos de pesquisa que convivem com a busca de identificação de áreas dialetais. Neste ponto, é possível ver uma confluência de propósitos entre a Dialectologia e a Sociolinguística uma vez que ambas as disciplinas analisam a variação linguística. Portanto, os enfoques diatópico e sociolinguístico estão presentes tanto na Dialectologia quanto na Sociolinguística. Todavia, o que as distingue é a forma de tratar os fenômenos e a perspectiva que cada uma imprime à abordagem dos fatos linguísticos.

1.2 Variação linguística

A variedade linguística depende de variedades geográficas ou diatópicas, bem como de variedades socioculturais ou diastráticas. As variedades diatópicas acontecem em um plano horizontal da língua e se originam dos dialetos ou falares locais, que se mostram por meio de uma linguagem aparentemente comum do ponto de vista geográfico. Estas variedades se distinguem em linguagem urbana e a linguagem rural. A linguagem urbana é influenciada por fatores culturais como a escola, meios de comunicação de massa e literatura e está mais próxima da linguagem comum. A linguagem rural é mais isolada e conservadora e vem desaparecendo com a chegada da civilização. Araújo e Sousa (2018, p. 6), fundamentados em Bortoni-Ricardo (2005), argumentam que "numa linha imaginária há duas extremidades". Em um extremo estão as variedades rurais isoladas, no outro as variedades urbanas padronizadas e no meio os dialetos rurais e urbanos, uma área rurbana.

Entretanto, a Dialectologia e a Geografia Linguística têm se empenhado em catalogar e analisar as particularidades linguísticas de comunidades rurais. As variedades socioculturais ou diastráticas, por seu turno, surgem em um plano vertical dentro de uma comunidade urbana ou rural e podem estar atreladas a

fatores relacionados ao falante – ou ao grupo a que pertence- ou à situação ou a ambos ao mesmo tempo.

Formas linguísticas em variação estão presentes em todas as comunidades de fala. Estas formas são chamadas de variantes que são maneiras diferentes de falar a mesma coisa no mesmo contexto e, geralmente, com o mesmo valor de verdade. Estas variantes, por sua vez, estão sempre competindo dentro da comunidade de fala à qual pertencem. Desta maneira, temos as variantes padrão e não padrão, aquelas que são conservadoras contra as que são inovadoras e as variantes que recebem algum tipo de estigma em oposição àquelas de prestígio. Uma variante padrão é considerada conservadora e possui maior importância sociolinguística dentro da comunidade. Em contrapartida, uma variante inovadora tende ser não-padrão e é, portanto, muitas vezes, estigmatizada pelos falantes da comunidade a que pertence. A título de ilustração, trazemos a presença do segmento fônico /s/ como marca de plural no sintagma nominal que é a forma padrão, conservadora e, portanto, de prestígio. Ao passo que a não marcação do plural /s/ no sintagma nominal é estigmatizada. Sobre variação linguística, assim se manifesta Moreno Fernández:

Às vezes, o uso de um elemento em lugar de outro do mesmo nível não supõe nenhum tipo de alteração semântica, de maneira que se usar um ou usar outro está se dizendo a mesma coisa. Isto é o que os sociolinguistas chamam de variação linguística”. (MORENO FERNÁNDEZ, 1998, p. 17)³.

Variável linguística é o elemento, traço ou unidade linguística que pode se apresentar de maneiras diferentes, isto é, de forma variável. Portanto, uma variável linguística é um conjunto de expressões do mesmo elemento e cada uma das manifestações ou expressões de uma variável recebe o nome de variante linguística.

³ “Hay ocasiones em que el uso de un elemento en lugar de otro del mismo nivel no supone ningún tipo de alteración semántica: tanto si se usa uno como si se usa outro, se está diciendo lo mismo. Esto es lo que los sociolinguistas denominan variación linguística”. Tradução nossa.

1.3 Geografia Linguística

As diferenças dialetais estão entre as observações mais comuns que os seres humanos fazem no dia a dia. Estas observações sobre a diversidade dialetal são tão comuns que é surpreendente que o estudo sistemático de dialetos só tenha começado na segunda metade do século XIX. Chambers e Trudgill (1994) descrevem as representações de áreas dialetais até a metade do século XIX como intuitivas e fortuitas. As primeiras tentativas de sistematizar a análise sobre diferenças dialetais surgiram como uma reação aos avanços da Filologia e outros estudos sobre as línguas. Foram os neogramáticos que começaram a busca por princípios gerais da mudança linguística. Para eles, as mudanças fonéticas são governadas por uma regra e o princípio seguido é que todas as mudanças fonéticas não admitem exceções. Como consequência disso, houve o desenvolvimento da geografia linguística, uma metodologia, ou seja, um conjunto de métodos para compilar de maneira sistemática as demonstrações das diferenças dialetais. A Geografia Linguística (Geolinguística) procura criar uma base empírica sobre a qual se possam extrair conclusões a respeito da variedade linguística que ocorre em um lugar determinado. A Geolinguística revelou uma heterogeneidade que não se concebia antes e, portanto, jogava por terra toda e qualquer pressuposição de ausência de exceções.

1.4 O Brasil e a língua portuguesa

As línguas tendem a se alterar por diversas razões e as alterações são mais expressivas quando uma língua é transplantada de uma região para outra. Um exemplo disto são as línguas românicas oriundas do Latim. Estas línguas, por sua vez, quando trazidas pelos colonizadores para o novo mundo, diferiam da variedade falada nas metrópoles de origem, como o francês transplantado para o Canadá e o português trazido para o Brasil. Com relação à língua portuguesa

falada no Brasil, o povo autóctone, o índio cuja escravidão tem registro datado de 1531, a aprendeu primeiro. Logo em seguida, o negro entrou na cena da escravidão brasileira cujo registro pode datar de 1532. O índio e o negro aprenderam o português no Brasil quase que simultaneamente. Entretanto, os colonos portugueses, índios, africanos e seus descendentes, desde o século XVI, começaram a mudar a língua portuguesa que paulatinamente veio a ser o falar brasileiro, que assumiu características próprias. Segundo Lucchesi (2009, p. 71):

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.

1.5 O Projeto ALiB – Novos Caminhos e Perspectivas da Geografia Linguística no Brasil

Em 1996, retomou-se a ideia da construção de um atlas nacional com o surgimento do Projeto Atlas Linguístico do Brasil - ALiB. Durante o *Seminário Nacional Caminhos e Perspectivas para a Geolinguística no Brasil*, realizado em Salvador, na Universidade Federal da Bahia (UFBA), se constituiu um comitê nacional para elaboração do atlas. Este comitê, presidido pela Professora Doutora Suzana Alice Cardoso da UFBA, contou com a participação de autores de atlas publicados e em andamento. O ALiB documenta dados linguísticos no Brasil de Norte ao Sul, de Leste ao Oeste.

2 O LÉXICO E A LEXICOLOGIA

O léxico de uma língua é entendido, de forma geral, como o conjunto de todas as palavras que dela fazem parte. Correia e Almeida (2012), no entanto, consideram que a quantificação do léxico se estabelece com um sem-número de problemas e de questões de difícil resposta. Estas autoras usam palavra e unidade

lexical como expressões sinônimas. Para elas, a palavra típica corresponde, no discurso escrito, a uma sequência de caracteres delimitados por espaços em branco, ou seja, a palavra gráfica. Contudo, as palavras de uma língua assumem outras formas, uma vez que há unidades lexicais de dimensão maior que a palavra gráfica, isto é, os compostos sintáticos, como *casa de saúde* ou *sala de jantar*. O léxico de uma língua também se constitui de unidades de dimensão menor que a palavra gráfica, as chamadas unidades infralexicais. Estas unidades podem ser de significado lexical, as raízes, e podem também ser unidades infralexicais de significado gramatical ou instrucional, os afixos. As primeiras não são unidades autônomas e não apresentam padrão flexional e podem ocorrer apenas como elementos de construção de outras palavras como *psic-* e *metr-*. Os afixos – sufixos e prefixos para o português – têm apenas significado gramatical e podem se associar a unidades de significado lexical para construir novas palavras.

2.1 Distribuição diatópica das variantes lexicais

As primeiras pesquisas de cunho dialetal procuravam mostrar diferenças espaciais. O interesse pelo entendimento da realidade linguística dentro de um espaço físico levou ao desenvolvimento de trabalhos com o objetivo de retratar áreas e de apontar a realidade linguística de um território politicamente definido.

2.2 Distribuição social dos dados – variações diageracional, diassexual, diastrática

Fatores sociais como a escolaridade, profissão, idade e gênero têm se tornado aspectos da variação linguística que têm tomado corpo na pesquisa dialetal. Para Chambers e Trudgill (1994), a variação social de uma língua é tão comum e importante quanto a variação diatópica. Para eles, todos os dialetos são

tão espaciais quanto sociais já que os falantes têm ao seu redor um entorno social e geográfico.

Variação por idade ou diageracional tem sido comentada desde o século XIX, como afirma Cardoso (2010). O conhecimento da idade dos informantes é indispensável para comparar as diferenças que podem existir entre o falar dos jovens e o falar dos mais velhos. Variação por sexo ou diassexual, assim como a variação diageracional, constitui-se em interesse dos estudiosos da Dialectologia. Coulthard (1991) é enfático ao expor que homens e mulheres não falam exatamente da mesma maneira. Segundo ele, a diferença sexual mais patente ocorre provavelmente na altura da voz, pois a voz da mulher é, na maioria dos casos, uma oitava mais alta do que a do homem já que as cordas vocais dos homens são mais longas, vibram mais lentamente produzindo sons mais baixos. Esta característica, por si só, já indica o sexo de um falante ao telefone ou à distância. Entretanto, devemos salientar que esta assertiva de Coulthard nem sempre representa a realidade uma vez que há alguns homens com vozes agudas e algumas mulheres com vozes graves. Ele próprio afirma à página 20 que “a altura da voz pode ser alterada para servir como indicador sexual”. Fenômeno paralinguístico não é o único traço de distinção entre a fala do homem e a fala da mulher. Há outros elementos de diferenciação como entoação, vocabulário, tópico e controle de tópico. Para ele, homens e mulheres têm estilos interativos diferentes e também possuem assuntos preferidos e modos diferentes de usá-los.

Variação por nível de escolaridade ou por profissão se enquadra no que denominamos de variação diastrática. A condição social é um fator importante para o reconhecimento de que podem existir dois usos diferenciados da língua no mesmo lugar.

3 UM OLHAR SOBRE A BAHIA E O PARANÁ – ASPECTOS HISTÓRICOS E GEOGRÁFICOS

Escolhemos realizar o cotejo do léxico entre Bahia e Paraná por não apresentarem contiguidade geográfica e por terem formas de povoamento diferentes. O objetivo é analisar se o léxico apresenta mais homogeneidade ou heterogeneidade entre dois estados localizados em regiões diferentes no Brasil: Nordeste e Sul.

A Bahia se localiza no sul da Região Nordeste do Brasil, limitando-se ao leste pelo oceano Atlântico, ao norte com Sergipe, Alagoas, Pernambuco e Piauí, ao sul com os estados de Minas Gerais e Espírito Santo e oeste com Goiás e Tocantins.⁴

Segundo Reis (2009, p. 19), a província da Bahia era uma das mais prósperas regiões canavieiras das Américas no século XIX. Os engenhos de açúcar, puxados por mão-de-obra escrava, estavam situados especialmente no Recôncavo, região fértil e úmida que envolve a Baía de Todos os Santos. Reis afirma que Salvador, então mais conhecida como Cidade da Bahia, ocupa um dos extremos desse conjunto geográfico que impressionou muitos visitantes estrangeiros por sua beleza. Reis (2009, p. 19) acrescenta que, em 1805, Sir Robert Wilson, comandante da esquadra britânica no Atlântico Sul, conhecedor de muitos mares e terras, escreveu que “a vista desta Baía [...] é talvez a mais magnífica no mundo” (aspas de Reis, 2009).

O Paraná se localiza no Sul do Brasil, limitando-se ao norte com o estado de São Paulo, a leste com o oceano Atlântico, ao sul com o estado de Santa Catarina e a oeste com o estado do Mato Grosso e com as repúblicas do Paraguai e Argentina.⁵

⁴ <<https://www.sogeografia.com.br/Conteudos/Estados/Bahia/>> consultado em 15 set. 20.

⁵ Atlas Linguístico do Paraná. Vol. I. p. 21.

Durante os séculos XVI e XVII, houve uma disputa pela posse do território paranaense pelas missões jesuíticas espanholas e pelas bandeiras paulistas sob ordens do governo português. Durante estes dois séculos, vários núcleos de povoamento começaram a surgir ao longo dos principais rios. Contudo, Aguilera (1990) afirma que estes povoamentos foram abandonados e extintos por causa dos frequentes embates entre os bandeirantes paulistas e os missionários espanhóis. Nesse ínterim, por volta de 1640, o povoamento que começara no litoral prosseguia, com aventureiros a procura de jazidas de ouro. Aguilera acrescenta que chegava Gabriel de Lara a Paranaguá, com a tarefa de proteger a costa paranaense dos ataques estrangeiros. Enquanto isso, Eleodoro Ébano Pereira adentrava os campos de Curitiba, averiguando as minas que lá existiam.

4 METODOLOGIA

Este trabalho se fundamenta na metodologia e no *corpus* do Projeto Atlas Linguístico do Brasil (ALiB) no que tange à área semântica acidentes geográficos do Questionário Semântico-lexical (QSL). Neste estudo priorizamos a variação diatópica, mas, seguindo os passos da Geolinguística Pluridimensional, também consideramos aspectos relativos às variações diastrática, diageracional e diassexual.

4.1 *Corpus*

O *corpus* está constituído pelas respostas às quatro questões iniciais do Questionário Semântico Lexical do Atlas Linguístico do Brasil (COMITÊ NACIONAL DO ALiB, 2001) aplicado em 22 cidades na Bahia e 17 no Paraná. Na Bahia, há um total de 92 informantes dos quais oito estão na capital e 84 no interior. Do mesmo modo, o Paraná tem um total de 72 informantes dos quais 64 estão no interior e oito na capital.

Para a constituição do *corpus* partimos da leitura e exame das transcrições de todas as questões de 01 a 04 do QSL⁶, complementadas pela audição dos registros dos inquiridos linguísticos considerados.

Como se trata de um estudo cujo enfoque é o aspecto semântico-lexical, adotamos a transcrição grafemática representativa da forma padrão, considerando, no entanto, para a análise, variantes fonéticas das lexias ou expressões fraseológicas que podem revelar variações de cunho diastrático, diageracional ou diassexual. As lexias apresentadas no diminutivo são computadas como sua forma padrão, exceto a variante *riozinho*. *Riozinho* se apresenta como uma variante independente, de acordo com o que pudemos inferir com os relatos dos informantes, em vez de ser apenas o uso do diminutivo para *rio*.

4.2 Localidades da Bahia

As localidades da Bahia, que perfazem o total de 22 pontos, estão assim distribuídas em suas mesorregiões⁷:

- Mesorregião Centro Norte Baiano – Itaberaba (090)⁸, Jacobina (086), Irecê (085);
- Mesorregião Centro Sul Baiano – Caetité (096), Itapetinga (100), Jequié (095), Seabra (089), Vitória da Conquista (098);
- Mesorregião Extremo Oeste Baiano – Barreiras (087), Santana (092);

⁶ Servimo-nos das transcrições realizadas pelas equipes do Projeto ALiB na Bahia e no Paraná, no que concerne a dados das capitais e do interior.

⁷ Mesorregião. Unidade territorial homogênea, em nível maior que a microrregião, porém menor que o estado ou território, e resultado do grupamento de microrregiões. (Ferreira, 1986).

⁸ O número entre parênteses se refere ao atribuído à localidade na rede de pontos do Atlas Linguístico do Brasil (ALiB).

-
- Mesorregião Metropolitana de Salvador – Capital – Salvador (093), Santo Amaro (091);
 - Mesorregião Nordeste Baiano – Alagoinhas (088), Euclides da Cunha (083), Jeremoabo (082);
 - Mesorregião Sul Baiano – Caravelas (102), Ilhéus (099), Santa Cruz de Cabrália (101), Valença (094);
 - Mesorregião Vale São-Franciscano da Bahia – Barra (084), Carinhanha (097), Juazeiro (081).

4.3 Localidades do Paraná

O conjunto de localidades do Paraná, que perfaz o total de 17 pontos, está assim distribuído em suas mesorregiões:

- Mesorregião Centro Ocidental Paranaense – Campo Mourão (212), Terra Boa (209);
- Mesorregião Centro Oriental Paranaense – Piraí do Sul (214);
- Mesorregião Centro-sul – Guarapuava (219);
- Mesorregião Metropolitana de Curitiba – Capital – Curitiba (220), Adrianópolis (216), Lapa (222), Morretes (221);
- Mesorregião Noroeste Paranaense – Nova Londrina (207), Umuarama (210);
- Mesorregião Norte Central Paranaense – Cândido de Abreu (213), Londrina (208);
- Mesorregião Norte Pioneiro Paranaense – Tomazina (211);

-
- Mesorregião Oeste Paranaense – São Miguel do Iguaçu (217), Toledo (215);
 - Mesorregião Sudeste Paranaense – Imbituva (218);
 - Mesorregião Sudoeste Paranaense – Barracão (223).

4.4 Informantes

Em cada ponto de inquérito no interior dos dois estados foram entrevistados quatro informantes, dois homens e duas mulheres em duas faixas-etárias (18-30 anos e 50-65 anos). Nas capitais dos estados foram entrevistados oito informantes, quatro dos quais têm nível universitário.

Os informantes deste trabalho, que integram o corpo de informantes do ALiB, têm as seguintes características: (i) são naturais da localidade perscrutada, têm pais desta mesma localidade e não se ausentaram dela por períodos longos durante sua vida, evitando-se, assim, informantes cujas profissões requerem deslocamentos frequentes, como comissários de bordo ou caminhoneiros; (ii) os mais jovens têm idade compreendida entre 18 e 30 anos e os mais velhos entre 50 e 65 anos; (iii) são, em representação equitativa, do sexo masculino e feminino; (iv) são alfabetizados com ensino fundamental completo ou incompleto nas cidades do interior, e, nas capitais, quatro dos oito informantes possuem nível universitário.

4.5 Questionário

O *corpus* desta pesquisa se fundamenta nos dados originados da aplicação do Questionário Semântico-Lexical (QSL) integrante dos *Questionários 2001* (COMITÊ NACIONAL DO ALiB, 2001) e constituído de 207 questões divididas em quinze áreas semânticas das quais selecionamos a área semântica acidentes geográficos.

No quadro 1, disposto em quatro colunas, apresentamos o conjunto de questões de 1 a 4 utilizadas para este estudo, com a seguinte distribuição: a primeira coluna mostra o número da questão; a segunda exibe os itens semântico-lexicais mais frequentes no português brasileiro; a terceira coluna indica a maneira como foi formulada a pergunta; e, finalmente, na quarta vem indicada a área semântica a que se refere cada pergunta.

Quadro 1: Extrato do QSL utilizado

QSL Nº	Item Semântico-Lexical	Formulação da Pergunta	Área Semântica
01	CÓRREGO / RIACHO...	... um rio pequeno, de uns dois metros de largura?	ACIDENTES GEOGRÁFICOS
02	PINGUELA...	... tronco, pedaço de pau ou tábua que serve para passar por cima de um_____ (cf. item 1)?	
03	FOZ...	... o lugar onde o rio termina ou encontra com outro rio?	
04	REDEMOINHO (DE ÁGUA) ...	Muitas vezes, num rio, a água começa a girar, formando um buraco, na água, que puxa para baixo. Como se chama isto?	

5 ANÁLISE DOS DADOS

Neste capítulo apresentamos a descrição e a análise dos dados, estruturadas em itens que priorizam a perspectiva diatópica, mas contemplam, também, aspectos de natureza sociolinguística.

5.1 Acidentes geográficos em Salvador e Curitiba

A totalidade dos informantes de ambas as capitais forneceu três lexias e duas expressões fraseológicas⁹ distintas para o conceito *rio pequeno* perfazendo 16 ocorrências. As lexias *riozinho*, *córrego* e *riacho* estão presentes em Salvador e em Curitiba. As ocorrências mais produtivas foram *riacho* e *córrego*. Contudo, há uma preponderância de *riacho* sobre *córrego*. *Riacho* foi fornecido por quatro informantes no ponto 093 e por três informantes no ponto 220. Temos então, um total de sete ocorrências, perfazendo 43,8% das ocorrências nas duas capitais. *Córrego* foi fornecido por dois informantes no ponto 093 e quatro no ponto 220, perfazendo um total de seis ocorrências. Isto significa 37,5% das ocorrências nas duas cidades. *Riozinho* foi fornecido por três dos informantes – dois no ponto 093 e um no ponto 220. *Riozinho* representa 18,7%. Três informantes no ponto 093 e um informante no ponto 220 não souberam responder.

Para a questão 2 do QSL, tivemos 16 ocorrências para o conceito *pedaço de pau ou tábua que serve para passar por cima de um riacho*, das quais a lexia *ponte* apresenta maior índice de ocorrência nas duas capitais. Para doze dos informantes esta lexia afigura-se como primeira resposta – oito informantes do ponto 093 e quatro informantes do ponto 220. Ao mencionarmos primeira resposta, apontamos que o informante forneceu mais de uma resposta para a pergunta. *Ponte*, com as suas variantes como *pontezinha de madeira* e *pontezinha*, representa 85,7% dentro das 14 ocorrências. A variante que apresenta traço de cunho espacial para a questão 2 é *pinguela*, registrada no ponto 220. A forma diminutiva de *ponte* foi utilizada por três informantes do ponto 093/3,7,8 e um destes informantes acrescentou o material com o qual ele acredita que tal ponte deveria ser feita, oferecendo a expressão fraseológica *pontezinha de madeira*.

⁹ Xatara e Parreira (2011) dizem que há "um subconjunto vocabular, que reúne as unidades lexicais complexas, cujo entendimento e descrição compete a uma subárea da Lexicologia, a Fraseologia. Assim, essas unidades complexas são chamadas unidades fraseológicas".

Contudo, *pontezinha* foi registrada como *ponte*, em virtude dos critérios estabelecidos no que concerne ao tratamento dos diminutivos.

No ponto 220, *pinguela* ocorre na fala de dois informantes em Curitiba, contudo um deles aponta diferenças de cunho social entre *ponte* e *pinguela*, como se depreende do diálogo a seguir:

- INQ.- *E se quer passar por cima desse córrego e joga uma madeira, uma árvore...*

INF.- *Tem que ter uma ponte, uma pinguela que a gente fala muito nos mato quando vamo em pescaria, principalmente, atravessá um rio. A árvore... não, tem que atravessá uma pinguela. Eu conheço por pinguela, né.*

INQ.- *Tá certo, é isso mesmo. (220/3)*

Pinguela representa 14,3% do total dos registros e ocorre apenas em Curitiba onde tem maior produtividade. Constatamos, então, que *ponte* é a opção mais frequente nas duas capitais. Quanto às respostas para o conceito *o lugar onde o rio termina ou encontra com outro rio*, os informantes de ambas as capitais forneceram quatro lexias e expressões fraseológicas distintas apresentadas em dez respostas. A lexia *foz* teve três ocorrências em Salvador e quatro em Curitiba. *Foz*, portanto representa 77,8% das ocorrências considerando o número de respostas fornecidas. *Encontro de rio* e *embocadouro* representam 11,1% cada. Esta questão foi pouco produtiva dada a quantidade de informantes que não souberam ou não lembraram a resposta: três no ponto 093 e três no ponto 220. Para a questão 4 – *muitas vezes, num rio, a água começa a girar, formando um buraco, na água, que puxa para baixo* – ocorreu uma única resposta, *redemoinho*. A totalidade dos informantes de Salvador e de Curitiba forneceu esta lexia como única resposta para esta questão.

5.2 A realidade do interior dos dois estados

O que revelam linguisticamente o interior da Bahia e o interior do Paraná? Buscamos resposta ou respostas para esta pergunta no exame do léxico de 21 cidades na Bahia e de 16 cidades do Paraná agrupadas em mesorregiões de acordo com a classificação do IBGE¹⁰.

5.3 Cidades baianas em suas mesorregiões – distribuição diatópica do léxico

Como resposta à pergunta 1, registramos 82 ocorrências que se referem à presença de oito variantes. Destas, destacam-se, pelo número de registros, *riacho* detentor da maior produtividade – 53 ocorrências, representando 64,7% do total – que está presente em todas as regiões; e *córrego* com 24 registros que representam 29,3% do total. As outras variantes apresentam baixa produtividade. *Riozinho* obteve um registro no centro sul e atinge apenas o índice de 1,2%. *Ribeirão* representa 2,4% com duas ocorrências no centro sul baiano. *Rego* foi elicitado por dois informantes na mesorregião Nordeste e, também, representa 2,4% do total das ocorrências.

Houve 79 ocorrências para a questão 2, das quais *ponte* recebeu 48 registros e representa 60,8% das ocorrências. *Pinguela* obteve 26 registros correspondendo a 32,9% das ocorrências. A variante *travessia* conta com três registros e representa 3,7%. As outras variantes *pontaleta* e *passarela* obtiveram um registro cada e, portanto, cada uma destas variantes representa 1,3%.

A questão 3 apresentou um número alto de abstenções: 70 informantes não deram respostas. Houve 15 ocorrências das quais *encontro de rio* e *barra* obtiveram maior índice de resposta. *Encontro de rio* obteve 10 registros representando 66,7% das 15 ocorrências e *barra* com cinco registros representa 33,3%. A questão 4 consta de 62 ocorrências das quais *redemoinho* figura como a variante mais

¹⁰ IBGE <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=22269>> Consulta em 03 set. de 2020.

produtiva com 49 registros e corresponde a 79,1% das ocorrências. Outras variantes que foram registradas e obtiveram três ocorrências cada uma delas foram *funil*, *corrupio*, *chupão* e *sumidouro* e representam, cada uma delas, 4,8% do total dos registros para a questão 4. A variante que se mostrou menos produtiva para esta questão foi a expressão fraseológica¹¹ *olho d'água* com apenas um registro, o que representa apenas 1,7%.

5.4 Cidades paranaenses em suas mesorregiões - distribuição diatópica do léxico

Para a questão 1, registramos sete variantes. Houve 51 ocorrências. Diferentemente das cidades do interior da Bahia onde a lexia *riacho* é majoritária aqui temos a lexia *córrego* com maior produtividade, com 22 registros, 43,1% das ocorrências. *Riacho* obteve 16 registros, 31,4% no total das mesorregiões do Paraná. As demais variantes são *arroio*, *riozinho*, *ribeirão* e *sanga*. *Arroio* corresponde a 9,8% das ocorrências com cinco registros enquanto *sanga* e *riozinho* com três registros cada correspondem a 5,9%. *Ribeirão* aparece com duas ocorrências e representa 3,9%. Para a questão 2, *pinguela* aparece com mais produtividade no interior do Paraná do que no interior da Bahia onde a predominância é da lexia *ponte* para a questão 2. Houve 54 ocorrências. No interior do Paraná, *pinguela* obteve 36 registros e *ponte* 18 representando 66,7% e 33,3% das ocorrências, respectivamente. A questão 3 apresenta um alto nível de abstenções: 48. Houve 16 ocorrências com um considerável número de variantes: cinco. Estas variantes apresentam baixa produtividade, sendo a expressão *encontro de rios* a mais representativa com apenas sete ocorrências, 43,8% das respostas registradas. A segunda variante mais produtiva é *barra* ou *barra do rio*

¹¹ Para um melhor entendimento sobre expressão fraseológica, sugerimos: TRISTÁ, Antonia María. *Fraseologia y Cotexto*. Habana: Editorial de Ciencias Sociales, 1988. ALVAREZ, Maria Luisa; UNTERBAUMEN, Enrique Huelva. *Um (Re)visão da Teoria e da Pesquisa Fraseológicas*. Campinas: Pontes Editores, 2011.

com quatro registros, 25%. *Foz* e *final de rio* aparecem com dois registros cada – 12,5% cada. *Desemboque*, com um registro, representa 6,2% do total registrado. A questão 4 traz 58 ocorrências. Os 55 registros da lexia *redemoinho* representam 94,9%. *Sumidouro*, *rebojo* e *funil* aparecem cada uma com apenas um registro configurando cada uma delas 1,7% do conjunto de variantes registradas.

5.5 Dados etimológicos - uma breve análise

Das 26 variantes documentadas neste trabalho, 21 têm origem no latim e cinco não têm registro etimológico nos dicionários consultados ou o registro etimológico aponta em outra direção que não Latim, como em *rego* que segundo Cunha (2010) tem origem pré-romana. *Sanga*, lexia registrada no oeste e sudoeste paranaense, é citada por Machado (1967) como sendo de origem obscura e por Cunha (2010) como sendo de origem controversa. Outro exemplo é *chupão* documentado no centro sul baiano. Machado não cita *chupão*, mas traz o verbo *chupar* como tendo uma etimologia obscura. Cunha afirma que *chupar* tem origem onomatopaica. As variantes que têm origem no latim representam 80,8% do total ao passo que aquelas que não encontramos nos dicionários estudados (Cunha, 2010; Machado, 1967) ou estes dicionários não exibem a etimologia representam 19,2%, ficando evidente o não registro de africanismos, indigenismos e estrangeirismos, pelo menos nesse extrato do *corpus* do ALiB.

5.6 Um olhar sociolinguístico sobre o tema

Os informantes arrolados para este estudo estão distribuídos de maneira a abranger equitativamente todas estas variáveis consideradas e se encontram, assim, identificados: informantes de 1 a 4 possuem nível fundamental e de 5 a 8 nível universitário. Os números ímpares se referem aos homens e os números

pares se referem às mulheres; os números 1-2 e 5-6 são atribuídos aos informantes agrupados na primeira faixa etária e 3-4 e 7-8 à segunda faixa etária.

Os dados das capitais possibilitam uma análise considerando os níveis de escolaridade, a variação linguística por sexo e a diferença linguística por faixa etária. Contudo, nas cidades do interior, por ter sido considerado apenas um nível de escolaridade — o fundamental — os informantes são considerados segundo a faixa etária e o sexo. Abordamos, neste item, apenas as variantes mais produtivas e aquelas com observações fornecidas pelos informantes durante o inquérito que consideramos relevantes para nosso estudo.

1) Variação social referente a acidentes geográficos em Salvador e Curitiba:

No primeiro grupo, registros referentes às perguntas 1 a 4, observamos que as lexias *riacho*, *córrego*, *ponte*, *redemoinho* não se caracterizaram como de uso específico de grupo social, apresentando-se, dessa forma, como gerais nas duas capitais.

2) Variação social referente a acidentes geográficos no interior da Bahia e Paraná

No primeiro grupo, que concerne às questões relativas a acidentes geográficos no interior dos dois estados, verificamos que:

- *Riacho*, *córrego*, *ponte*, *pinguela*, *redemoinho* não apresentam variação social.

CONSIDERAÇÕES FINAIS

A variação diatópica registrada na Bahia e no Paraná, no que diz respeito às diferenças registradas no léxico analisado neste recorte e referente a esses estados, não sugere resultar de implicações históricas ou relacionadas ao tipo de povoamento. Embora, a história registre a grande presença de línguas africanas na Bahia por séculos durante o período da escravidão negra no Brasil, a presença de línguas indígenas tanto na Bahia quanto no Paraná e grupos de imigrantes

estrangeiros no Paraná, constatamos que a maioria do léxico considerada nesta pesquisa tem origem no Latim. No que se refere à variação sociolinguística, observamos que as variantes mais produtivas são de uso geral, não apresentando variação social.

Verificamos que no plano da análise diatópica, podemos destacar como resultado mais relevante o fato de a distribuição das variantes ter um caráter mais homogêneo do que heterogêneo considerando os dados de Salvador e Curitiba bem como de todas as cidades do interior dos dois estados.

REFERÊNCIAS

AGUILERA, Vanderci de Andrade. *Atlas Linguístico do Paraná*. 1990. Vol. I. Tese apresentada à Faculdade de Ciências e Letras da Universidade Estadual Paulista – UNESP, Assis.

ARAÚJO, Silvana Silva de Farias; SOUSA, Nilton Carlos Carmo. A História social do português do Brasil e o preconceito linguístico. *Revista Tabuleiro de Letras*. Salvador PPGEL, nº 03, p. 82-96, Dez. 2018. Disponível em: [<https://www.revistas.uneb.br/index.php/tabuleirodeletras/article/view/5568>] Acesso em 07/03/2021.

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

CHAMBERS, J.K.; TRUDGILL, Peter. *La Dialectologia*. Tradução Carmen Morán González. Madrid: Visor Libros, 1994.

COMITÊ NACIONAL DO PROJETO ALiB – Atlas Linguístico do Brasil. Questionários 2001. Londrina: UEL, 2001.

CORREIA, Margarita; ALMEIDA, Gladis M^a de Barcellos. *Neologia em Português*. São Paulo: Parábola, 2012.

COULTHARD, Malcolm. *Linguagem e Sexo*. Tradução Carmen Rosa Caldas-Coulthard. São Paulo: Ática, 1991.

CUNHA, Antônio Geraldo da. *Dicionário Etimológico da Língua Portuguesa*: 4^a ed. revista e atualizada. Rio de Janeiro: Lexikon, 2010.

BIDERMAN, Maria Tereza C. *Teoria Linguística: leitura e crítica*. São Paulo: Martins Fontes, 2001.

BORTONI-RICARDO, Stella Maris. *Nós chegemu na escola, e agora?: sociolinguística e educação*. 2. ed. São Paulo: Parábola Editorial, 2005.

FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro: Nova Fronteira, 1986.

LUCCHESI, Dante. História do contato entre Línguas no Brasil. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (orgs.). *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009. p. 71.

MACHADO, José Pedro. *Dicionário Etimológico da Língua Portuguesa: com a mais antiga documentação escrita e conhecida de muitos dos vocábulos estudados*. 2ª edição. V. I, II, III. Lisboa: Confluência, 1967.

MORENO FERNÁNDEZ, Francisco. *Principios de Sociolingüística y sociología del language*. Barcelona: Editorial Ariel, 1998.

NASCENTES, Antenor. *O Linguajar Carioca*. 2ª Ed. Rio de Janeiro: Organizações Simões, 1953.

REIS, João José. *Rebelião Escrava no Brasil: a história do levante dos Malês em 1835*. São Paulo: Companhia das Letras, 2009.

XATARA, Cláudia; PARREIRA, Maria Cristina. A Elaboração de um Dicionário Fraseológico. In: ALVAREZ, Maria Luisa Ortiz; UNTERNBAUMEN, Enrique Huelva (orgs.). *Uma (Re)visão da Teoria e da Pesquisa Fraseológicas*. Campinas: Pontes, 2011. p. 70.

O AUTOR E O PPGLinC

Genivaldo da Conceição Oliveira

Após algumas idas e vindas dos Estados Unidos da América, voltei para o Brasil em 2008 com um diploma de Mestre em Linguística, com especialização em Fonética e Fonologia, obtido na Universidade do Texas em Austin. Em 2009, passei em dois concursos públicos federais. Optei pela Universidade Federal. Em 2010, participei da seleção para o Doutorado em Linguística no Programa de Pós-Graduação em Língua e Cultura (PPGLinC) da Universidade Federal da Bahia. Fui aprovado e comecei as disciplinas em 2011. Concluí meu Doutorado em 2014, sob a orientação da brilhante, simpática e generosa Professora Doutora Suzana Alice Cardoso, que me ensinou, com sabedoria e profundidade, os caminhos da Dialetoлогия e a grandeza do Atlas Linguístico do Brasil (ALiB). O PPGLinC contribuiu muito para minha formação como linguista, como pesquisador, como cientista da palavra.

Nota do editor:

Artigo submetido para avaliação em: 16 de setembro de 2020.

Aprovado em sistema duplo cego em: 03 de março de 2021.