

Áreas lexicais no Centro-Sul do Brasil sob uma perspectiva geolinguística

Lexical areas in Central-South of Brazil under a Geolinguistic Perspective

Valter Pereira Romano

Universidade Federal de Lavras, Lavras, Minas Gerais / Brasil

valter.pereira.romano@gmail.com

Resumo: Este trabalho apresenta resultados principais de uma pesquisa que visa discutir, sob o ponto de vista lexical, a divisão dialetal proposta por Antenor Nascentes (1953) no que tange à área geográfica denominada pelo estudioso como subfalar sulista. Foram adotados os procedimentos teóricos e metodológicos da Dialectologia e da Geolinguística e, como corpus de análise, utilizados os dados do Projeto Atlas Linguístico do Brasil, correspondentes a 118 municípios brasileiros, totalizando 472 informantes. Quatro questões do Questionário Semântico-Lexical (QSL) foram selecionadas para validar a hipótese de que a área geográfica investigada é heterogênea linguisticamente. Com base na observação do comportamento diatópico de algumas variantes lexicais, conclui-se que a porção setentrional do território investigado apresenta diferenças em relação à parte meridional, evidenciando a existência de dois possíveis falares: o paulista e o sulista.

Palavras-chave: Projeto ALiB; subfalar sulista; variação lexical.

Abstract: This work presents the main results of a research that aims to discuss, under the lexical perspective, the dialectal division proposed by Antenor Nascentes (1953) with regard to the geographical area called by the scholar as sulista subspeech. The theoretical and methodological assumptions of Dialectology and Geolinguistics were adopted and

the analytical corpus used was the data of the Linguistic Atlas Project of Brazil corresponding to 118 Brazilian municipalities, totaling 472 informants. Four questions from the Lexical-Semantic Questionnaire were selected to validate the hypothesis that the geographical area investigated is linguistically heterogeneous. From the observation of the diatopical behavior of some lexical variants, it is concluded that the northern portion of the territory investigated presents differences in relation to the southern part, showing the existence of two types of speech: the paulista and the sulista.

Keywords: ALiB Project; Sulista subspeech; Lexical variation.

Recebido em 26 de setembro de 2016.

Aprovado em 24 de abril de 2017.

1 Introdução

Os trabalhos de natureza geolinguística buscam, entre outros objetivos, apresentar dados que comprovem ou não a existência de áreas dialetais em determinado território, uma vez que os dados linguísticos coletados empiricamente são registrados em “mapas especiais” (COSERIU, 1987).

Embora Amadeu Amaral, em 1920, já chamasse a atenção para a necessidade de trabalhos empíricos com a finalidade de retratar “com segurança quais os caracteres gerais do dialeto brasileiro, ou dos dialetos brasileiros” (AMARAL, 1981, p. 44), as tentativas de divisão e sistematização de tais ‘dialetos’ vieram a se consolidar na obra *O linguajar Carioca*, de Antenor Nascentes (1953), transformando-se na clássica divisão dialetal do Português Brasileiro (PB), que até hoje é referenciada, embora apresente algumas inconsistências.

Este trabalho visa discutir a divisão dialetal de Nascentes (1953) no que se refere à área geográfica designada pelo dialetólogo como subfalar sulista de forma a contribuir para uma possível redefinição dessa divisão sob uma perspectiva lexical.

Parte-se do pressuposto de que a área geográfica estudada, atualmente, não apresenta homogeneidade lexical; há variantes típicas de determinadas áreas que evidenciam traços dos processos de povoamento e ocupação humana. O estudo justifica-se pela necessidade de (i) descrever

a realidade do PB pautando-se em dados empíricos; (ii) preencher a lacuna existente, na Geolinguística, com um trabalho sistemático de uma área não contemplada por atlas linguísticos regionais e estaduais em sua totalidade e (iii) comparar dados da língua oral de estados pertencentes a regiões administrativas diferentes. Dessa forma, torna-se possível comparar dados da língua falada no Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Minas Gerais, Mato Grosso, Mato Grosso do Sul e Goiás, o que revela semelhanças e diferenças entre as regiões, bem como alguns traços linguísticos que definem e caracterizam cada um dos estados, evidenciando áreas lexicais que extrapolam os limites administrativos das unidades federativas em questão.

2 Léxico e variação diatópica do PB

A variação diatópica do PB é uma realidade comprovada empiricamente pelos inúmeros trabalhos dialetológicos e geolinguísticos e subjaz aos demais tipos de variação. Considerando-se a extensão territorial do Brasil e os diferentes ‘brasis’ que surgiram em decorrência do processo de ocupação e povoamento do território (RIBEIRO, 2006), a diversidade diatópica do PB tem-se revelado como um traço característico dessa língua, principalmente no que se refere ao léxico, e a sua descrição e análise tem sido empreendida sob diferentes perspectivas e aportes teóricos.

No que se refere às abordagens dialetais e geolinguísticas, os dados coletados para a elaboração do Atlas Linguístico do Brasil (ALiB) têm contribuído para a descrição do léxico das diferentes regiões do país. Nessa seara, os trabalhos mais recentes são os de Freitas-Marins (2012), Ribeiro (2012), Razky (2013), Portilho (2013), Romano e Aguilera (2014), Romano e Seabra (2014a, 2014b), Romano (2015), D’Anuniação (2016), Santos (2016), Romano e Seabra (2017), entre outros. Esses trabalhos destacam as diferenças regionais, ora confirmando a proposta de divisão de Nascentes (1953), sob a perspectiva lexical, ora evidenciando particularidades que caracterizam cada região, sobretudo por influências sócio-históricas do processo de ocupação e povoamento.

2.1 Áreas dialetais do PB

Na segunda edição da obra *O linguajar carioca* (1953), Nascentes discute as primeiras propostas de divisão dialetal que remontam o ano de 1881, como a de Julio Ribeiro. Revisitando outras propostas e fazendo

ponderações àquelas que sugeriram na primeira edição de seu livro (1922), o autor reformula seu mapa dialetal do Brasil justificando que

Hoje que já realizei o meu ardente desejo de percorrer todo o Brasil, do Oiapoc ao Xuí, de Recife a Cuiabá, fiz nova divisão que não considero nem posso considerar definitiva, mas, sim, um tanto próxima da verdade. (NASCENTES, 1953, p. 24)

Em 1953, o estudioso apresenta um novo mapa dialetológico. Com base em critérios prosódicos e fonéticos – a cadência e a abertura das vogais médias ([e] / [o]) em posição pretônica, como em p[e]cado/p[ɛ]cado e c[o]ração/c[ɔ]ração – Nascentes divide o PB em seis subfalares, compreendidos em dois grandes grupos, o do Norte e o do Sul (FIG. 1).

FIGURA 1 – Divisão dialetal de Nascentes (1953)



Fonte: Nascentes (1953)

O falar do Norte engloba o subfalar amazônico e nordestino, e o falar do Sul contempla o subfalar baiano, intermediário entre os dois

grupos, o fluminense, o mineiro e o sulista. Soma-se a esse conjunto uma área considerada como Território Incaracterístico que, na época, era praticamente despovoado.

Essa proposta sedimentou-se entre os dialetólogos, tornando-se a clássica divisão do PB, confirmada em trabalhos como os de Cardoso (1986; 1999) sob a perspectiva fonética. Todavia, trabalhos recentes discutem essa divisão, sob o ponto de vista lexical.

Ribeiro (2012), por exemplo, estudou o *subfalar baiano*; Portilho (2013), o *amazônico*; Romano (2015), o *subfalar sulista*; e Santos (2016), o *fluminense*. Essas pesquisas atestam o fato de que esses falares não se limitam à faixa territorial estabelecida por Nascentes, mas adentram a outras áreas.

Ao fazer uma retrospectiva das divisões dialetais, já em 2006, Mota (2006) ressaltara que não fora proposta uma nova divisão, mesmo havendo um substancial avanço dos estudos dialetológicos e geolinguísticos, principalmente, em razão da dificuldade de intercomparação entre os dados dos atlas estaduais e regionais por motivos de ordem metodológica e temporal. Para a autora, faz-se necessário maior conhecimento das áreas dialetais brasileiras, “especialmente daquelas que ainda não dispõem de atlas regionais assim como de uma amostra atualizada, recolhida simultaneamente, com mesma metodologia e sob coordenação geral em todo o País” (MOTA, 2006, p. 351). Nesse sentido, atualmente, os dados coletados para o Projeto ALiB podem contribuir para uma nova proposta de divisão dialetal cientificamente justificável.

Partindo do prognóstico de Mota (2006) sobre a importância do Projeto ALiB para a delimitação de áreas dialetais do PB, encontram-se, atualmente, inúmeros trabalhos com base no corpus do Projeto. Esta pesquisa vem contribuir ao desenvolvimento do Projeto ALiB, pois traz uma análise sistemática de dados de uma extensa área geográfica que compreende o denominado subfalar sulista.

3 Materiais e Métodos

O corpus analisado refere-se aos dados inéditos coletados pela equipe do Projeto em 118 cidades. Em cada uma das localidades, foram entrevistados quatro informantes, dois homens e duas mulheres, pertencentes a duas faixas etárias: Faixa I, de 18 a 30 anos, e Faixa II,

de 50 a 65 anos, naturais da localidade e tendo o ensino fundamental (completo ou incompleto) como grau máximo de escolaridade.¹

A área investigada desta pesquisa contempla localidades de nove estados: Rio Grande do Sul (RS), Santa Catarina (SC), Paraná (PR), São Paulo (SP), Rio de Janeiro (RJ), Minas Gerais (MG), Goiás (GO), Mato Grosso (MT) e Mato Grosso do Sul (MS). Ela se difere da exata área demarcada por Nascentes, pois a dificuldade em trabalhar com o mapa de Nascentes (1953) (Figura 1) foi a imprecisão e a falta de detalhes nos limites geográficos estabelecidos pelo autor para a delimitação de cada subfalar.

Sobre os limites do subfalar sulista, o autor descreve que compreendem os estados de SP, PR, SC, RS, MG, GO e MT, não apresentando mais especificações sobre esses limites. Vale notar que a imagem do mapa do Brasil subdividido em falares não apresenta elementos geográficos que subsidiem a precisão das fronteiras dos subfalares.

Desse modo, depois de criterioso estudo do mapa de Nascentes (1953) e a distribuição da rede de pontos do ALiB por meio do georreferenciamento, chegou-se a um mapa parecido ao do autor, no que tange à área geográfica do subfalar sulista. Além dessa área, foram considerados pontos linguísticos adjacentes aos limites desse subfalar, denominados “pontos de controle”, como já o fizeram Ribeiro (2012) e Portilho (2013).

A rede de pontos do subfalar sulista compõe-se de 108 localidades, às quais foram acrescentados 10 pontos, pois se considera importante verificar a fronteira desse subfalar com os outros subfalares (mineiro e fluminense) e também com o território incaracterístico. Assim, foram incluídas para o estudo duas cidades fluminenses – ponto 205 (Barra Mansa) e 206 – Parati; quatro cidades mineiras: 130 (Unai), 132 (Pirapora), 141 (Formiga) e 145 (São João Del Rei); duas cidades goianas: 118 (Porangatu) e 121 (Formosa) e duas cidades mato-grossenses: 103 (Aripuanã) e 104 (São Félix do Araguaia). Em sua totalidade, o estudo engloba nove localidades mato-grossenses, seis sul-mato-grossenses, oito localidades goianas, duas fluminenses, 11 municípios mineiros, 38

¹ Nas capitais, foram entrevistados, além de informantes de nível fundamental, quatro informantes de nível superior. Considerando-se os objetivos desta pesquisa, os dados dos informantes de nível superior não foram considerados.

idades paulistas, 17 paranaenses, 10 localidades catarinenses e, por fim, 17 cidades gaúchas².

Foram selecionadas questões que evidenciam diversidade de formas lexicais, que podem apresentar variantes diatópicas que possibilitem o traçado de isoglossas e que revelem áreas lexicais no território investigado.

Compreende-se por isoglossa a linha virtual que marca limites também virtuais de formas e expressões linguísticas em determinada área geográfica (FERREIRA; CARDOSO, 1994). Como este trabalho se refere ao nível lexical, optou-se pelo termo isoléxica (mesmo léxico) para denominar as linhas virtuais que evidenciam áreas em que há a ocorrência de duas variantes ou de uma forma lexical em relação à ausência de variante. Por extensão ao conceito de heteroglossa, proposto por Chambers e Trudgill (1994), optou-se pelo termo heteroléxica (léxico diferente), para denominar as linhas que delimitam áreas de coocorrência de duas variantes. Reconhece-se, todavia, que os limites geográficos que definem a abrangência de uma variante em detrimento de outra são fluidos, sobretudo pela natureza dos dados em análise. Desse modo, entende-se que as linhas de isoléxicas e heteroléxicas revelam a “arealidade” de determinada variante, neologismo aqui utilizado para denominar a distribuição espacial ou areal de uma forma linguística.

Neste artigo, são analisados os dados das seguintes questões do Questionário Semântico-Lexical (COMITÊ NACIONAL DO PROJETO ALiB, 2001):

- Questão 001 – Córrego – “Como se chama um rio pequeno, de uns dois metros de largura?”, da área semântica Acidentes geográficos;
- Questão 132 – Menino – “Criança pequenininha, a gente diz que é bebê. E quando ela tem de 5 a 10 anos, do sexo masculino?”, da área Ciclos da vida;
- Questão 156 – Bolinha de gude – “Como se chamam as coisinhas redondas de vidro com que os meninos gostam de brincar?”, da área Jogos e diversões infantis;
- Questão 177 – Geleia – “Como se chama a pasta feita de frutas para passar no pão, biscoito?”, da área semântica Alimentação e cozinha.

² A rede de pontos linguísticos do Projeto ALiB é identificada por números. A lista completa está disponível no site: https://alib.ufba.br/sites/alib.ufba.br/files/rede_de_pontos_pdf

Esses dados foram coletados do corpus mediante consulta às transcrições/revisões e às gravações em áudio. O corpus passou pelo processo de revisão e posterior armazenamento no banco de dados do software desenvolvido para agilizar o processo de geração de relatórios e de cartografia linguística, o [SGVCLin] (ROMANO; SEABRA; OLIVEIRA, 2014).

4 Análise dos dados

Nesta seção, apresentam-se a descrição e análise das variantes registradas para as quatro questões supramencionadas. Cada questão apresenta especificidades tratadas individualmente, seja com a criação de rótulos para alguns itens, seja com agrupamento de variantes quando necessário. A discussão atém-se à distribuição diatópica de algumas variantes com vistas a discutir as áreas lexicais.

A validade de uma forma lexical considerada ou não como um designativo para o referente foi pautada nas acepções constantes dos dois principais dicionários da língua portuguesa, Houaiss e Villar (2001) e Ferreira (2004), e na descrição do informante, quando questionado sobre tal designativo. Quanto à etimologia de alguns vocábulos, recorreu-se ao Dicionário Etimológico de Cunha (2010) e ao Dicionário da Real Academia Espanhola (2001).

4.1 Córrego

A questão 001 do QSL busca obter as variantes lexicais para “o rio pequeno de dois metros de largura”. Foram documentados 644 registros, distribuídos em 25 formas. Diante desse polimorfismo, foi necessária a criação de alguns rótulos bem como o agrupamento de variantes. Os rótulos, criados com a finalidade de facilitar a compreensão e descrição dos dados, enquadram-se em três categorias:

- “sugestão na pergunta”: agrupa formas lexicais cujo sema já está expresso na formulação da questão, por exemplo, as variantes *rio*, *rio pequeno*, *rio estreito* e *rio raso*.
- “formas inadequadas”: contempla os itens lexicais que, apesar de terem traços semânticos comuns ao referente, não designam, especificamente, o item buscado, como ocorre em *lagoa*, *lago*, *poço*, *açude*. Essas variantes denominam outro referente, que tem como traço característico a presença de água represada.

- Sob o rótulo “formas pouco produtivas” estão englobadas as *hápx legomena*: *bocaina*, *braço de rio* e *lajeado* que, tanto em Houaiss e Villar (2001) quanto em Ferreira (2004), figuram com pelo menos uma acepção análoga ao referente *córrego*.

Além dos rótulos criados, foram necessários agrupamentos de variantes morfofonêmicas,³ como é o caso de itens em que houve

- redução da proparoxítona (*córrego* > *corgo*);
- formas no diminutivo (*corgo* > *corguinho*; *riacho* > *riachinho/riachozinho*; *sanga* > *sanguinha*, *rego* > *reguinho*, *valeta* > *valetinha*)
- formas que apresentam o mesmo radical (*valo*, *vala*, *valão*)
- formas compostas (*rego d'água* > *rego*).

No caso das variantes de *valeta*, foram feitos os agrupamentos considerando-se os critérios (ii) e (iii); para as variantes de *rego*, consideraram-se os critérios (ii) e (iv). Realizados os devidos agrupamentos, os 644 registros estão distribuídos em 14 itens.

As cinco variantes mais produtivas são: *córrego* e suas variantes morfofonêmicas (261 ocorrências – 40,53%), *riacho* e variantes (122 ocorrências – 18,94%), *riozinho* (82 – 12,73%), *ribeirão* (38 ocorrências – 5,9%) e *sanga/sanguinha*, com 37 registros (5,75%), representando, juntas, 83,85% das respostas.

Seguem-se a essas formas os itens *valeta* e variantes (26 ocorrências – 4,04%). As formas contempladas sob o rótulo “sugestão na pergunta” (*rio*, *rio pequeno*, *rio estreito* e *rio raso*) apresentam 21 registros (3,26% do corpus). A variante *arroio* figura com 19 ocorrências (2,95%), e *rego* obtém 17 registros (2,64%).

Com menor índice, encontram-se oito “respostas inadequadas” (*lago*, *lagoa*, *açude* e *poço*). A variante *corixo* obteve apenas cinco registros, representando 0,78% do corpus. Ainda em menor produtividade,

³ Carreter (2008, p. 281), acerca das variantes morfofonêmicas, afirma que “los fonólogos han propuesto este término para designar ‘la idea compleja de todos los miembros (dos o más) de una alternancia. Así, en la alternancia que se produce en las formas alemanas *geben-gab-gib*, las vocales *e*, *a*, *i* (llamadas *alternantes*) constituyen uno morfofonema.”

encontram-se *canal* (0,47%), *grota d'água* (0,31%) e as *hapax legomena* (*bocaina*, *braço do rio* e *lajeado*) que, sob o rótulo “formas pouco produtivas”, representam juntas 0,47%⁴.

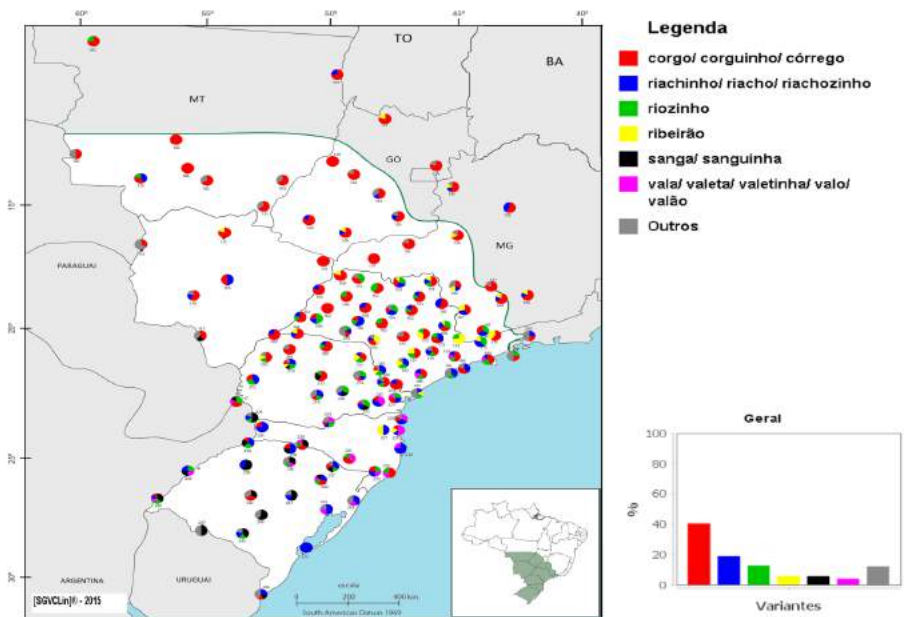
A Figura 2 apresenta a distribuição diatópica de seis das 14 formas. Embora o [SGVCLin] possibilite a cartografia de até 10 variantes, com a finalidade de facilitar a visualização, optou-se por representar cartograficamente as seis formas mais produtivas, para evitar o comprometimento na distinção das cores. As demais variantes estão agrupadas no item *outras*, representadas pela cor cinza. Esse item da legenda contempla o rótulo *sugestão na pergunta, rego, arroio, formas inadequadas, canal, grota d'água* e as *hápx legomena: bocaina, braço de rio e lajeado*.

Há uma concentração da variante *córrego* com suas variantes morfofonêmicas em localidades situadas no estado de SP (não ocorrendo em cinco dos 38 pontos paulistas), e sua presença é registrada em todos os pontos selecionados no RJ, MG, MS, MT e GO (FIG. 2). À medida que se avança no sentido meridional, *córrego* vai perdendo produtividade, dando espaço para variantes regionais como *valeta*, *sanga* e *arroio*.

A carta de a realidade gradual de *córrego* (FIG. 3) ilustra a distribuição do item considerando-se a produtividade por ponto linguístico. Observa-se que essa forma lexical, embora esteja amplamente difundida pelo território, com exceção do RS, apresenta variações quanto à produtividade na rede de pontos, sendo representativa, sobretudo, em grande parte do território do Centro-Oeste e no Triângulo Mineiro, onde a produtividade chega a 100% de incidência. Por outro lado, em SP há uma oscilação entre os índices de 75% a 25%, e à medida que adentra os estados da região Sul, sobretudo a partir do centro-sul do PR, o índice de representatividade não passa de 50%. Há, portanto, um comportamento heterogêneo de *córrego* no território do subfalar sulista.

⁴ Neste artigo, a análise restringe-se ao comportamento diatópico das variantes: *córrego*, *sanga* e *arroio*.

FIGURA 2 – Distribuição diatópica das variantes lexicais para a questão 001 do QSL



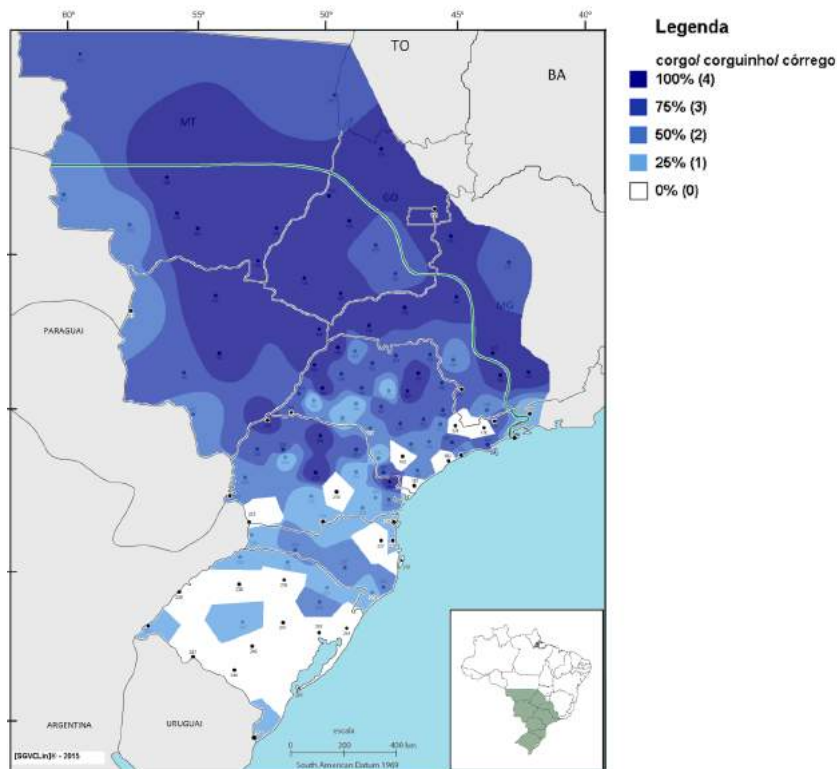
Fonte: Banco de dados do ALiB (2015)

A variante *sanga* é a quinta mais produtiva entre os informantes (37 ocorrências), equivalendo a 5,75% do *corpus*. Apresenta-se em quatro dos nove estados que compõem a pesquisa: MS, PR, SC e RS, com diferentes índices de produtividade. A variante é mais produtiva no RS, onde obtém 27,91% de representatividade (24 ocorrências). No PR, *sanga* apresenta 10 ocorrências (9,17%) e, com menor produtividade, encontra-se em SC, com apenas dois registros (3,82%). No MS, há somente uma ocorrência (3,03%).

A distribuição de *sanga* permite o traçado de uma isoléxica (FIG. 4) que contempla grande parte do território do RS, com exceção da faixa leste do estado (ponto 243 – Porto Alegre, 244 – Osório e 249 – São José do Norte). Essa área passa por uma localidade catarinense – ponto 229 (Concórdia), até chegar ao sudoeste paranaense (pontos 223 – Barracão e 217 – São Miguel do Iguaçu). Embora não prossiga numa faixa contínua, *sanga* ocorre também em uma localidade sul-mato-grossense – ponto 117 – (Ponta Porã). No PR, *sanga* forma uma segunda área lexical

(no sentido transversal do Estado) percorrendo um feixe do noroeste em direção ao sudeste e englobando quatro localidades: 212 – Campo Mourão, 213 – Cândido de Abreu, 218 – Imbituva e 222 – Lapa. Essas duas áreas de ocorrência podem evidenciar traços da história social da região, sobretudo ao se associar esse fato linguístico ao caminho dos tropeiros, nos séculos XVIII e XIX (ROCHE, 1969).

FIGURA 3 – Arealidade gradual da variante *córrego*



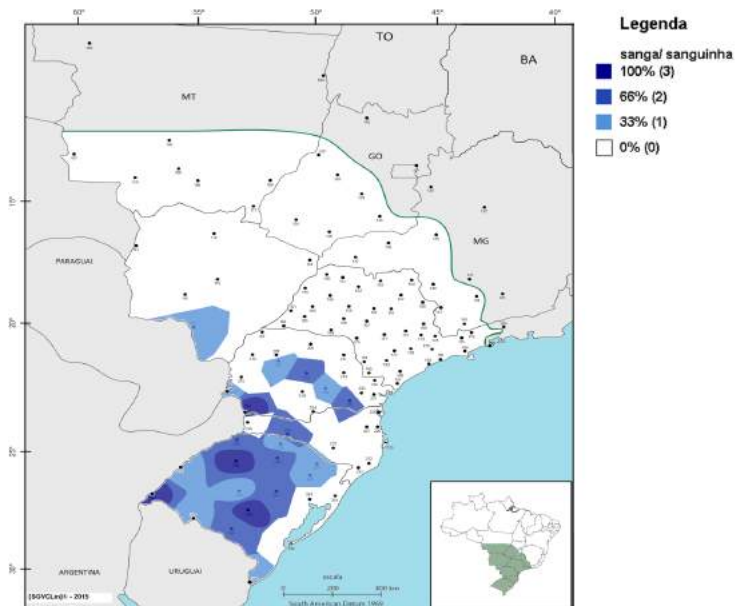
Fonte: Banco de dados do ALiB (2015)

Retomando-se aspectos históricos sobre o povoamento da região Sul, sabe-se que o Tropeirismo consistiu em um importante movimento econômico e de ocupação humana que contribuiu para o surgimento de inúmeras cidades (ALMEIDA, 1981). Sabe-se ainda que as tropas

eram constituídas, em sua maioria, por gaúchos, paulistas, mamelucos e castelhanos, que deixaram traços de sua cultura por onde passaram. A ocorrência de *sanga* nesta região do subfalar sulista (oeste de SC e sudoeste paranaense) coincide com uma das rotas dos tropeiros (o Caminho das Missões), aberto no século XIX, depois de expulsos os jesuítas.

Abandona-se o traçado que cortava o Rio Grande transversalmente (Urugaiana, Alegrete, Viamão, Santo Antônio da Patrulha, São Francisco de Paula, Bom Jesus, Vacaria), adotando-se o traçado São Borja, Santo Ângelo, Cruz Alta, Carazinho, Passo Fundo, Lagoa Vermelha, Vacaria. Com exceção das duas primeiras que já existiam como povoações missioneiras, as demais citadas surgiram como consequência do novo traçado que ficou conhecido como o Caminho das Missões [...]. (BRUM *apud* ROCHA, 2008, p. 61)

FIGURA 4 – Arealidade gradual da variante *sanga*



Fonte: Banco de dados do ALiB (2015)

Ferreira (2004), na primeira entrada do verbete *sanga* para designar “um pequeno regato que seca facilmente”, traz o vocábulo como um brasileirismo do RS e de SC cuja etimologia remete ao espanhol platino *zanja*. Contudo, tanto em Ferreira (2004) quanto em Houaiss e Villar (2001), há a indicação de outra possível origem etimológica para o vocábulo, podendo *sanga* ter vindo do quicongo ou do quimbundo, *dizanga*.

É interessante notar que, apesar de o termo ter duas possibilidades etimológicas e indicar sua presença na história da língua portuguesa desde tempos seculares, a distribuição diatópica do item *sanga* no corpus investigado parece revelar traços do contato linguístico do português com o espanhol e acompanha o processo de movimentação interna dos gaúchos rumo aos estados de SC e PR.

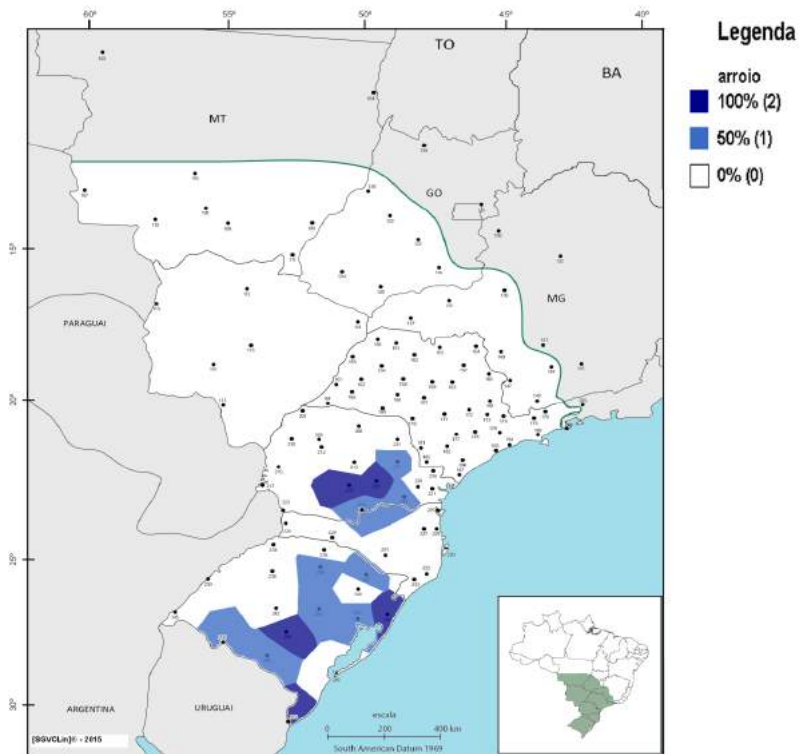
Os dados do Projeto ALiB têm mostrado que *sanga* não se restringe ao RS e à SC, mas adentra o PR, ocorrendo, inclusive, em pontos de passagem de outras rotas dos tropeiros, como em Lapa (ponto 222), situada no antigo Caminho de Viamão. Cabe notar que, apesar de no corpus do Projeto ALiB encontrar-se o registro de uma ocorrência de *sanga* no ponto 117 – Ponta Porã, no Atlas Linguístico do Mato Grosso do Sul (OLIVEIRA, 2007) e no Atlas Linguístico do Município de Ponta Porã (REIS, 2006), não se documenta tal variante, o que não possibilita afirmar, por ora, a influência de gaúchos nessa localidade do MS, pelo menos no que se refere ao uso de *sanga*.

O Atlas Linguístico do Paraná (AGUILERA, 1994), na carta 163, apresenta as isoléxicas das variantes para o “rio pequeno”, e, no que se refere à variante *sanga*, nota-se que essa forma lexical ocorre, principalmente em localidades do sudoeste paranaense, além de outros pontos em que coocorre com outra variante. O Atlas Linguístico Etnográfico da Região Sul do Brasil (ALTENHOFEN; KLASSMANN, 2011) apresenta a uma carta linguística que documenta as variantes lexicais para o “rio pequeno”, com destaque ao item *sanga*, evidenciando que a forma se registra em localidades do sudoeste paranaense, além de algumas ocorrências no centro do estado, embora com menor representatividade da que se encontra no corpus do ALPR e do Projeto ALiB.

Outra variante que também pode evidenciar traços da sócio-história da região Sul é a variante *arroio* (FIG. 5), apresentando-se com 19 registros (2,95% do corpus) distribuídos nos três estados da Região

Sul: RS (12 ocorrências – 13,95%), SC (uma ocorrência – 1,96%) e PR (seis ocorrências – 5,5%).

FIGURA 5 – Arealidade gradual da variante *arroio*

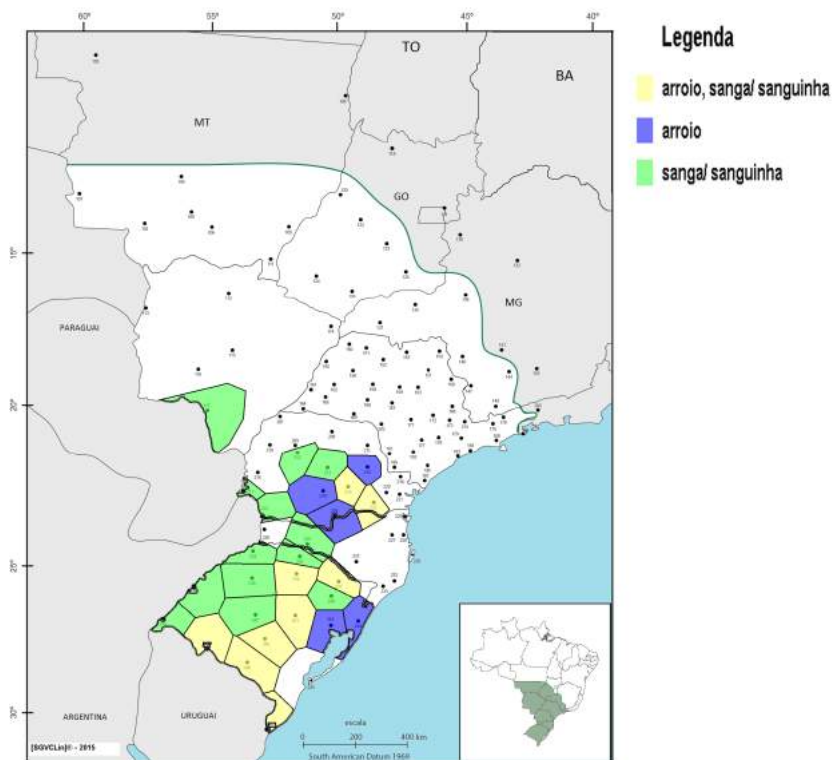


Fonte: Banco de dados do ALiB (2015)

Assim como *sanga*, *arroio* é uma palavra de origem hispânica (HOUAISS; VILLAR, 2001), procedente do latim vulgar *arrugiu*, no sentido de “pequeno curso d’água permanente ou não” (FERREIRA, 2004). Observando a área de ocorrência dessa variante na Região Sul, verificam-se novamente as marcas deixadas no léxico pelo movimento dos gaúchos rumo ao PR, possivelmente durante o ciclo do Tropeirismo e, posteriormente, da reimigração gaúcha de descendentes alemães, eslavos e italianos (ROCHE, 1969).

Arroio é uma forma que ocorre no extremo sul do país – ponto 250 (Chuí) – fronteira com o Uruguai, atravessando o interior do estado até chegar ao norte gaúcho (ponto 237 – Vacaria). No território catarinense, essa variante volta a ocorrer em Porto União – ponto 224 – já na fronteira com o PR, por onde se difunde em três localidades do centro-sul do estado – Lapa (ponto 222), Guarapuava (219), Imbituva (218) e, subindo na direção norte, chega ao ponto 214 (Pirai do Sul). Ao lado da ocorrência de *sanga*, pode-se observar a sobreposição dessas isoléxicas (Figura 6), o que confirma as premissas aventadas sobre a possível influência do movimento do tropeirismo e da situação de contato do português com o espanhol, formando uma área lexical no território investigado.

FIGURA 6 – Arealidade da variante *arroio*

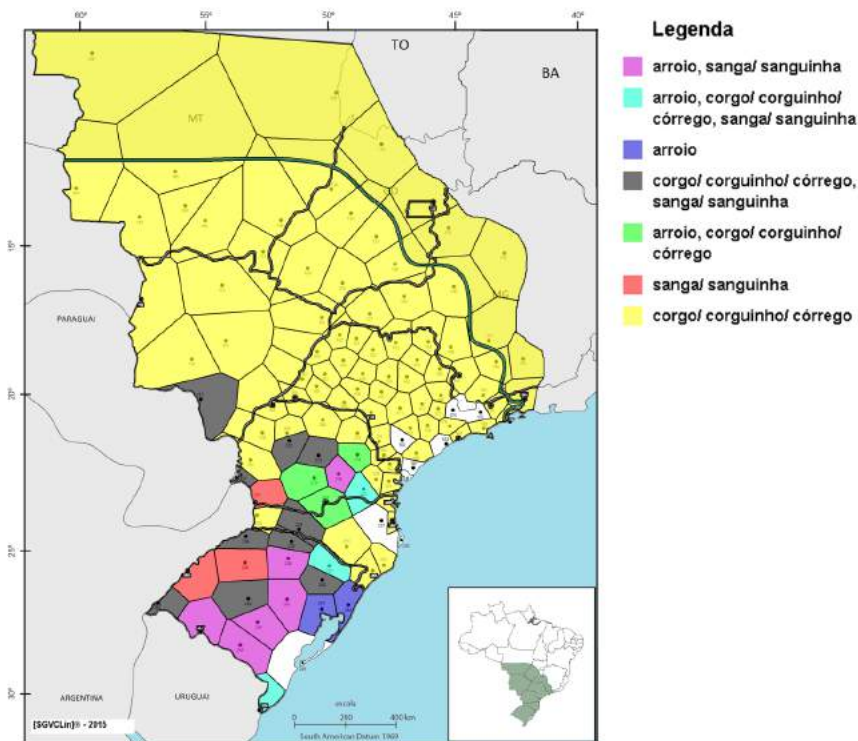


Fonte: Banco de dados do ALiB (2015)

Distinguem-se na Figura 6 a variante *sanga* no oeste rio-grandense e *arroio* ao leste. Na região central, há a sobreposição de formas, revelando áreas lexicais delimitadas por linhas de heteroléxicas. Nessa carta, é possível visualizar também uma área em SC pela qual se liga ao PR, denominada *área lexical do interior catarinense* (ROMANO; AGUILERA, 2014).⁵ No PR, as isoléxicas se sobrepõem em Lapa e Imbituva, pontos 222 e 218, respectivamente. *Sanga* segue rumo ao MS, refletindo-se em Ponta Porã – ponto 117 e *arroio* avança rumo ao Norte Pioneiro paranaense, chegando até o ponto 214 – Piraí do Sul, município que fazia parte do antigo Caminho de Viamão durante o ciclo do tropeirismo. Desse modo, não seriam *sanga* e *arroio* itens lexicais que podem tipificar um possível falar sulista de influência sul-rio-grandense que também evidencia traços do contato do português com outras línguas, no caso o espanhol? Em busca desse possível falar sulista é que se empreendem as análises deste trabalho.

Observa-se, diante desse cenário, o polimorfismo acentuado para a questão, do qual se tem uma visão ampla no mapa de sobreposição de isoléxicas, formando áreas lexicais (FIG. 7).

⁵ Com base nos trabalhos de Koch (2000) e Altenhofen (2005), os autores analisam a distribuição diatópica das variantes lexicais para *bolinha de gude*, considerando-se o *corpus* do Projeto ALiB, confirmando três das oito ‘áreas dialetais’ estabelecidas por Altenhofen (2005), e propõem outras três áreas, entre as quais, a *área lexical do interior central de Santa Catarina*, caracterizada pela ocorrência de variantes que revelam o contato entre os gaúchos e paulistas durante o ciclo do Tropeirismo.

FIGURA 7 – Arealidade das variantes lexicais *córrego*, *sanga* e *arroyo*

Fonte: Banco de dados do ALiB (2015)

Considerando apenas as variantes *córrego*, *sanga* e *arroyo*, verifica-se que, na área correspondente ao subfalar sulista, os estados pertencentes ao Sudeste e Centro-Oeste apresentam menor número de sobreposição de isoléxicas em comparação aos da Região Sul, onde se observa o maior número de formas regionais e combinação dessas formas com outras variantes. Ou seja, considerando essas três, a região do subfalar sulista, de Nascentes (1953), apresenta dois comportamentos: (i) uma parte setentrional caracterizada pela difusão da variante padrão (*córrego*) e (ii) a parte meridional em que ocorre maior número de formas lexicais, delimitadas por linhas de heteroléxicas, correspondendo em grande parte aos estados do PR, SC e RS. Neste trabalho, para se referir

a (i), opta-se pela expressão *falar paulista* e, quando se tratar de (ii), *falar sulista*, sempre grafados em itálico. Salienta-se, todavia, que ao se referir a toda a área geográfica discriminada por Nascentes (1953) como subfalar sulista, não se utiliza o itálico e o adjetivo sulista sempre está acompanhado do substantivo subfalar, ou seja, a expressão subfalar sulista se refere à nomenclatura de Nascentes e, no estudo em tela, designa toda a área geográfica investigada, com exceção dos pontos de controle.

4.2 Geleia

A questão 177 do QSL registra os designativos que recobrem o conceito da “pasta feita de frutas para passar no pão, biscoito”. Para essa questão, foram documentados 420 registros, além de 85 abstenções de respostas. As 420 ocorrências distribuem-se em três variantes: *geleia*, *chimia/michia* e *musse*, além de itens lexicais agrupados sob três rótulos: “formas genéricas”: *doce*, *doce de fruta*, *doce de creme* e *doce em pasta*; “formas inadequadas”: *creme*, *creme de fruta*, *melado*, *patê*, *patê de fruta* e *polpa de fruta*; “sugestão na pergunta”: *pasta* e *pasta de fruta*.

A forma lexical *geleia* predomina em 267 das 420 ocorrências (63,57%), seguida de *chimia* – 71 registros (16,90%), formas genéricas (53 – 12,62%) e, em menor produtividade, encontram-se as formas inadequadas (11 ocorrências – 2,62%), itens em que a própria pergunta sugere a resposta – 10 registros (2,38%) e, por fim, a variante *musse*, com oito ocorrências (1,90%). Para essa questão, não é documentado o polimorfismo lexical pelo fato de o índice de abstenção de resposta ser alto. No entanto, as ocorrências das variantes delimitam áreas lexicais no território investigado (FIG. 8).

Na carta (FIG. 8), observa-se a maior concentração de *geleia* no estado de SP, difundindo-se pelos estados da região Centro-Oeste e pelo território mineiro. No PR, há uma concentração de formas genéricas relacionadas à forma lexical *doce*, enquanto *musse* se apresenta significativamente em SC, e *chimia* se difunde pelo RS.

A segunda variante mais produtiva é *chimia* e a sua variante morfofonêmica *michia*. Essas formas representam juntas 16,9% do corpus, apresentando-se em 71 das 420 ocorrências. De acordo com Houaiss e Villar (2001), *chimia* é uma variante fonética de *chimiê* e *chimier*, caracterizadas como um regionalismo do RS para designar a “geleia para passar no pão da merenda, chimiê”, e provém do alemão

Schimiere, que significa lubrificante, graxa, proveniente do verbo *schmieren*, que quer dizer ‘untar, lubrificar, borrar, sujar’.

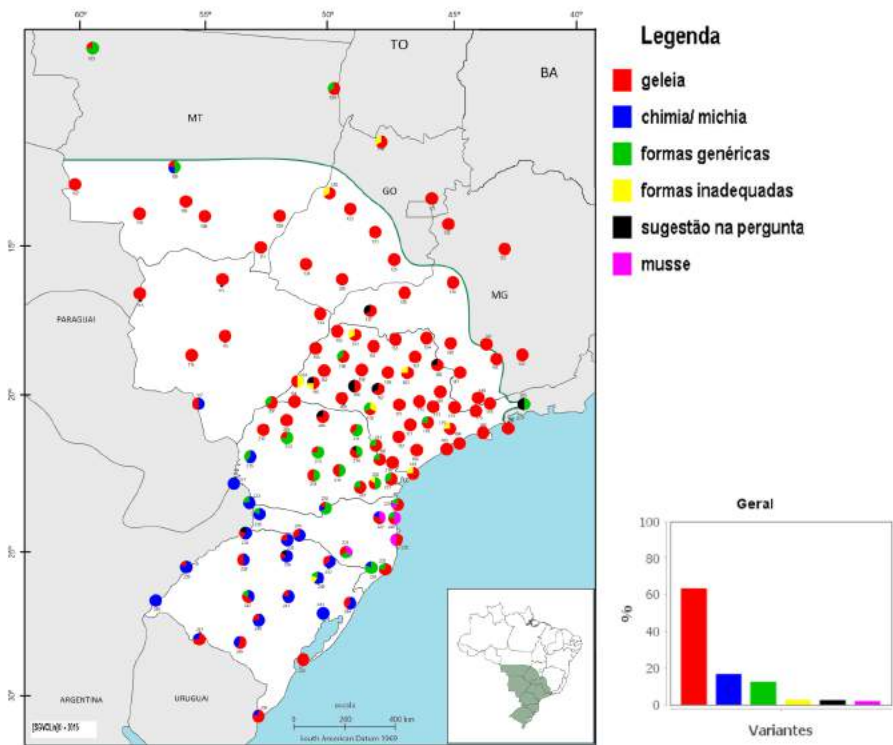
Considerando-se a etimologia do vocábulo, bem como a história social da Região Sul, a presença de *chimia* evidencia traços de bilinguismo e empréstimo do português de línguas de colonização. Dos nove estados contemplados pelo estudo, esse item lexical está presente em cinco, com diferentes índices de representatividade (FIG. 9).

Observa-se que, no MT e no MS, *chimia* apresenta baixa representatividade, figurando como ocorrência única no ponto 117 – Ponta Porã, localizado no sudoeste sul-mato-grossense e no ponto 105 – Diamantino, situado no centro do MT. Assim, pode-se inferir que não se trata de uma forma típica desses estados, considerando, inclusive, relatos dos informantes:

INF.- Ó... a língua do pessoal paranaense que eu já vi de muito tempo até o meu minino invocô, é ele comia nói chamava doce [’dosɪ], né: “ô muié, traiz o doce aí” e a gente já falava do que que é, traiz o doce de banana ou doce de mamão, ou doce de abróba, né, que é o jerimum, a língua certa é o jerimum mesmo, então traiz aí pa nós aqui. Já na língua paranaense chama chimia [ʃi’miɐ], né, tudo quando é feito de doce pa passá no pão é chimia.⁶

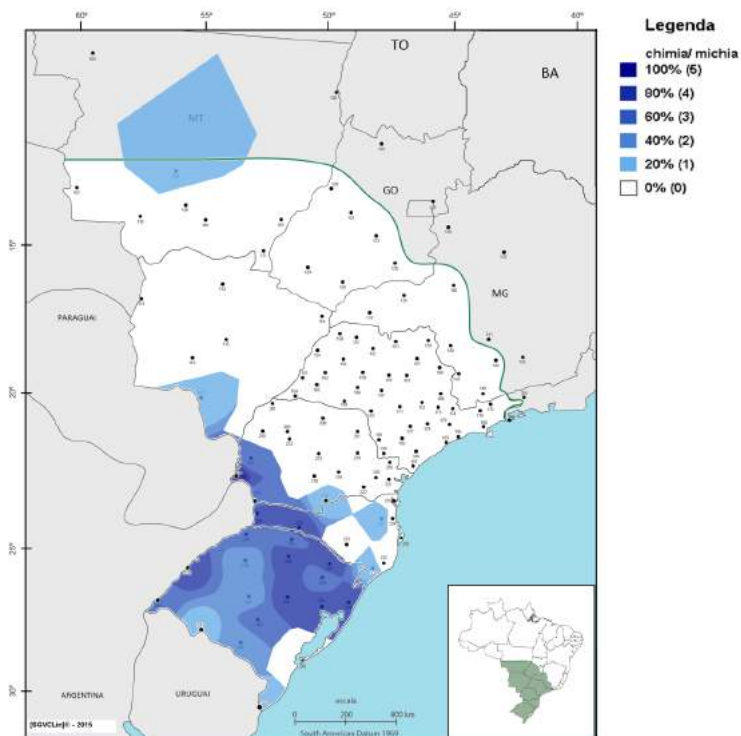
⁶ Informante 3 do ponto 103 (Homem, Faixa etária II, de Aripuanã-MT).

FIGURA 8 – Distribuição diatópica das variantes lexicais para a questão 177 do QSL



Fonte: Banco de dados do ALiB (2015)

O próprio informante mato-grossense reconhece que *chimia* é a denominação que pessoas da Região Sul, genericamente rotulados como “paranaenses”, usam para nomear o que, para o informante, é designado como *doce* (FIG. 9).

FIGURA 9 – Arealidade gradual de *chimia* / *michia*

Fonte: Banco de dados do ALiB (2015)

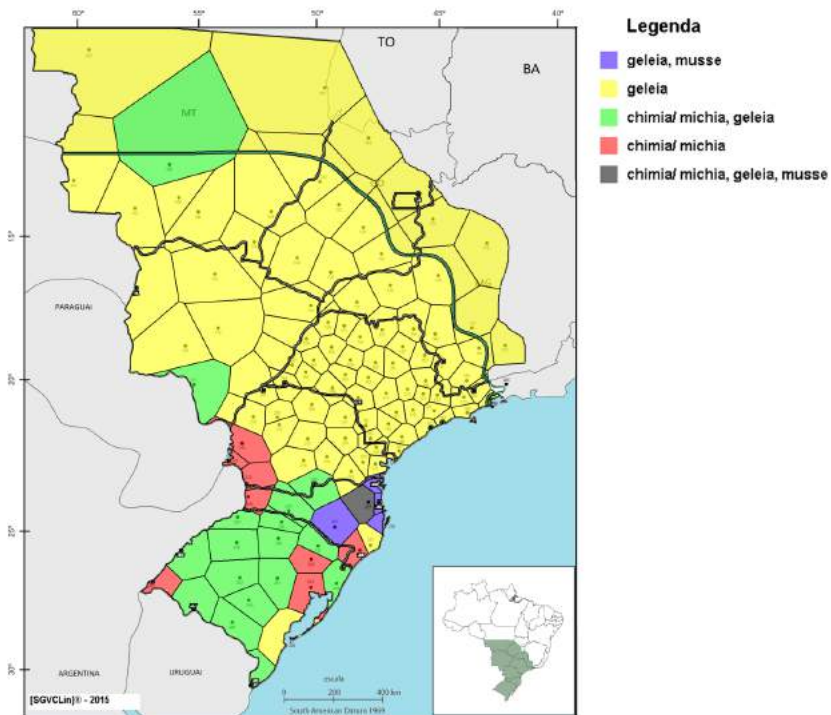
A distribuição diatópica de *chimia* (FIG. 8 e 9) revela uma área que contempla o estado do RS quase em sua totalidade, adentra SC pelo centro e oeste e, em um feixe lateral, perpassa o oeste paranaense até atingir uma localidade sul-mato-grossense – ponto 117 (Ponta Porã).

Por fim, a variante *musse* delimita uma área lexical (FIG. 8). Apesar de obter apenas oito ocorrências (1,9% do corpus), *musse* apresenta-se em uma área específica do subfalar sulista, em cinco cidades catarinenses: três localizadas no nordeste do estado: pontos 225 (São Francisco do Sul), 227 (Blumenau) e 228 (Itajaí), na capital – ponto 230 (Florianópolis) e 231 – Lages, no sul de SC. No ALERS, *musse* ocorre também no RS e no PR, embora com baixa produtividade.

Para uma visão ampla do comportamento das variantes, a Figura 10 apresenta a realidade dos três itens lexicais detalhados: *geleia*,

chimia/michia e *musse*. Situação análoga à que foi apresentada na Figura 7 repete-se para os designativos da *geleia*. Na área correspondente ao subfalar sulista, os estados pertencentes às regiões Sudeste e Centro-Oeste apresentam menor número de sobreposição de isoléxicas em comparação com os da Região Sul, onde se observa maior número de formas regionais e combinação dessas formas com outras variantes, formando heteroléxicas. Observando a carta, verifica-se a presença de uma grande área nos estados de SP, MS, MT, GO, MG e noroeste paranaense, em que ocorre, exclusivamente, a forma mais difundida do corpus, *geleia*, inclusive nas adjacências do subfalar sulista em MG e GO. A alternância de formas se acentua na Região Sul, que apresenta maior número de combinação de variantes em áreas de heteroléxicas.

FIGURA 10 – Arealidade das variantes lexicais geleia, chimia, formas genéricas e musse



Fonte: Banco de dados do ALiB (2015)

Considerando essas três variantes, a região do subfalar sulista novamente apresenta dois comportamentos: (i) a parte setentrional, caracterizada pela difusão da variante padrão (*geleia*) e (ii) a parte meridional, em que ocorre maior número de formas lexicais e áreas de heteroléxicas, correspondendo, em grande parte, aos Estados do PR, SC e RS.

4.3 Bolinha de gude

A pergunta 156 do QSL – *coisinhas redondas de vidro com que os meninos gostam de brincar* – apresenta 40 formas, incluindo as variantes fonéticas e morfofonêmicas. Foram necessários agrupamentos de variantes para o tratamento dos dados, conforme os seguintes critérios:

- (i) Formas no diminutivo: *bola/bolinha; bulica/buliquinha; burca/burquinha;*
- (ii) As formas compostas: *bola de gude/bolinha de gude/gude; bolinha de clique/clica; bolita de gude/bolita; bolinha de búrico/búrico,* entre outros.
- (iii) Formas que apresentam o alçamento da vogal pretônica: *bolito/bulita; bolica/bulica;*
- (iv) Variantes fonéticas que apresentam a alternância morfofonêmica da vogal átona final [o] / [a]: *bolito/bolita; búlico/búllica;*
- (v) Itens em que há alternância da lateral alveolar para vibrante na sílaba tônica: *bulita/burita; biloca/biroca; bilosca/birosca;*
- (vi) Formas em que há a alternância da lateral para a vibrante na sílaba pós-tônica: *búllica/búrica;*
- (vii) Formas proparoxítonas e paroxítonas: *búllica/bulica;*
- (viii) Variantes fonéticas que apresentam síncope da sílaba pós-tônica: *búrica/burca/buque;*
- (ix) Vocábulos que apresentam o fenômeno de suarabácti: *clica/quilica;*
- (x) Variante em que houve anteriorização da oclusiva velar surda (posterior) para uma oclusiva bilabial surda (anterior): *quilica/pilica.*

Para algumas das formas, utiliza-se mais de um critério. Fez-se necessária a inclusão do rótulo “formas pouco produtivas” para agrupar as *hápax legomena*: *biribinha*, *bili*, *biroquê*, *boleja*, *pelota* e *pedrinha de vidro*.

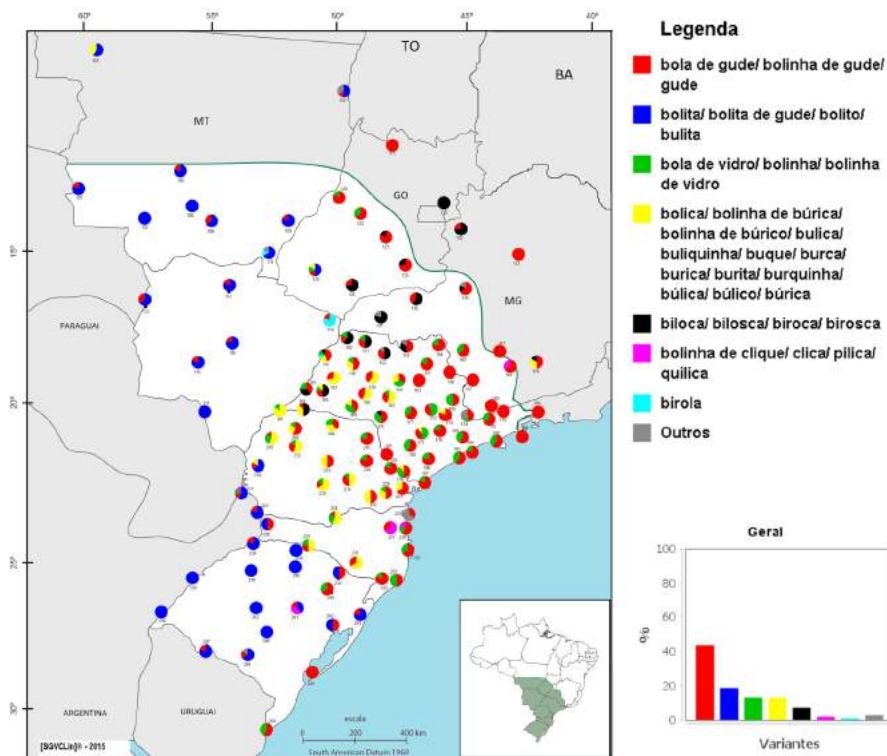
Para essa questão, são documentados 663 registros distribuídos em 11 itens, além de duas abstenções de resposta. A variante mais produtiva é *bolinha de gude* (43,44%), seguida por *bolita* (18,7%), *bolinha de vidro* (12,97%) e *búrca* (12,67%). Com menos de 10% de representatividade, encontra-se o item *biloca* (6,94%) e, em menor índice, encontram-se *clica* (1,81%), *birola* (0,90%), *fubeca* (0,75%), *peca* (0,60%) e *peteca* (0,30%). As *hápax legomena*, agrupadas sob o rótulo “formas pouco produtivas”, representam juntas 0,9% do corpus.

A Figura 11 apresenta a distribuição diatópica de sete das 12 formas lexicais. Os itens que representam menos de 1% de produtividade, na cartografia estão documentadas pelo item *outros*, que compreende as variantes: *birola*, *fubeca*, *peca*, *peteca*. Observa-se que *bolinha de gude* e variantes estão amplamente difundidas pelo território do subfalar sulista e adjacências com diferentes índices de ocorrência. O item é mais produtivo em seis dos nove estados: GO, MG, PR, RJ, SC e SP, atingindo, no entanto, em apenas quatro desses, produtividade acima de 50%. Os estados em que há o menor índice de *bolinha de gude* são o MT (13,04%), o MS (20,69%) e o RS (25%), em detrimento da ocorrência representativa de outros itens, por exemplo, *bolita* e suas variantes morfofonêmicas.

Bolita apresenta-se em seis estados (RS, SC, PR, MS, MT e GO) com diferentes índices de produtividade, não ocorrendo na região Sudeste, sendo representativa no Centro-Oeste e Sul. *Bolita* predomina, majoritariamente, em localidades matogrossenses e sul-mato-grossenses. Em GO, a produtividade de *bolita* atinge 7,69%. Já na região Sul, SC apresenta o menor índice de produtividade (6,67%) e, no PR, o item representa 11,61% das respostas. Nas localidades gaúchas, *bolita* é a forma mais produtiva (63,75%). No MS, *bolita* ocorre em cinco dos seis pontos (com exceção do ponto 114 – Paranaíba – localizado próximo à fronteira política do MS com SP, MG e GO). Em GO, a variante é registrada apenas em um ponto no sul do estado (124 – Jataí) e, em território mato-grossense, esse item lexical está amplamente difundido, ocorrendo, inclusive, nos pontos de controle (103 – Aripuanã e 104 – São Félix do Araguaia), localizados no extremo norte do estado.

A área correspondente ao uso da variante *bolitana*, Região Sul, contempla quase todo o RS, principalmente uma porção central, adentrando SC e PR pela região oeste desses estados, ou seja, pela área fronteira do Brasil com a Argentina e o Paraguai.

FIGURA 11 – Distribuição diatópica das variantes lexicais para a questão 156 do QSL



Fonte: Banco de dados do ALiB (2015)

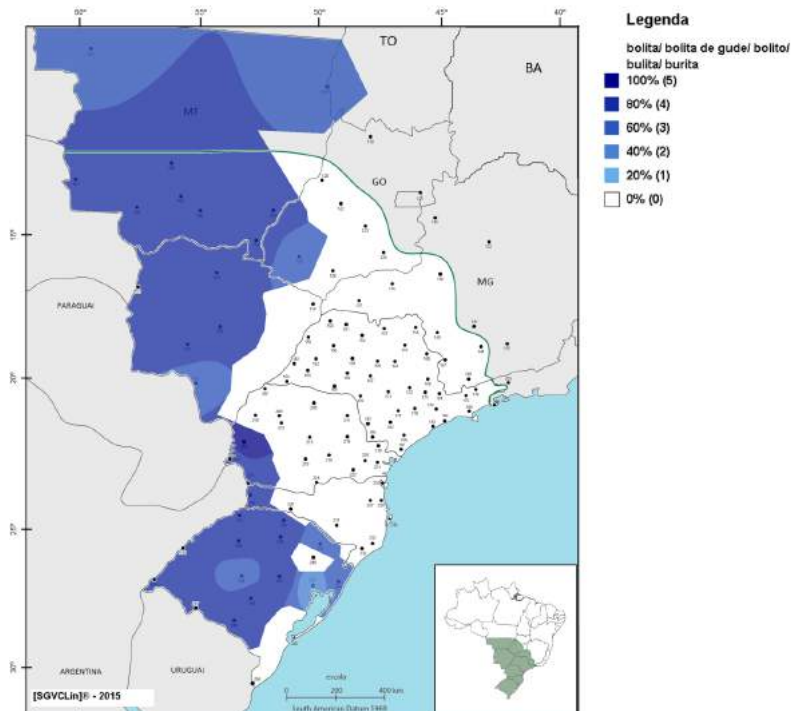
Ferreira (2004) e Houaiss e Villar (2001) registram *bolita* como uma forma vinda do espanhol platino típica do RS na acepção da *bolinha de gude*. Entretanto, os trabalhos geolinguísticos têm mostrado que não se trata de um regionalismo/brasileirismo exclusivo do RS, como

afirmam os lexicógrafos. Na carta linguística 109 do *Atlas Linguístico da Mesorregião Sudeste do Mato Grosso* (CUBA, 2009), encontra-se registrada em 100% das respostas. O *Atlas Linguístico do Mato Grosso do Sul* (OLIVEIRA, 2007) também documenta os designativos para o referente. Observa-se, desse modo, a ampla distribuição desse item pelo estado, não ocorrendo, nesse atlas, também na área correspondente à fronteira do MS com MG e GO, assim como no corpus do ALiB (FIG. 11).

Reis (2006), na carta 218, apresenta a distribuição das variantes para *bolita* em Ponta Porã/MS, cidade localizada na fronteira do Brasil com o Paraguai. Nesse trabalho, encontra-se a variante como forma mais produtiva na maioria dos pontos linguísticos, concorrendo apenas em três localidades com a variante *bola/bolinha de gude*.

No *Atlas Linguístico-Etnográfico da Região Sul do Brasil* (ALTENHOFEN; KLASSMANN, 2011), a carta linguística 302 documenta as variantes lexicais para a *bolinha de gude*. Na referida carta, verifica-se que a área correspondente à *bolita* coincide, em grande parte, com a documentada pelo ALiB, sobretudo no que tange à faixa oeste de SC e PR. *Bolita* ocorre no centro e oeste do RS, caracterizando-se como a variante majoritária. À medida que se avança no território de SC e do PR, a representatividade diminui, figurando na área fronteiriça desses dois últimos estados com a Argentina e o Paraguai (FIG. 12).

FIGURA 12 – Arealidade gradual de bolita e variantes morfofonêmicas



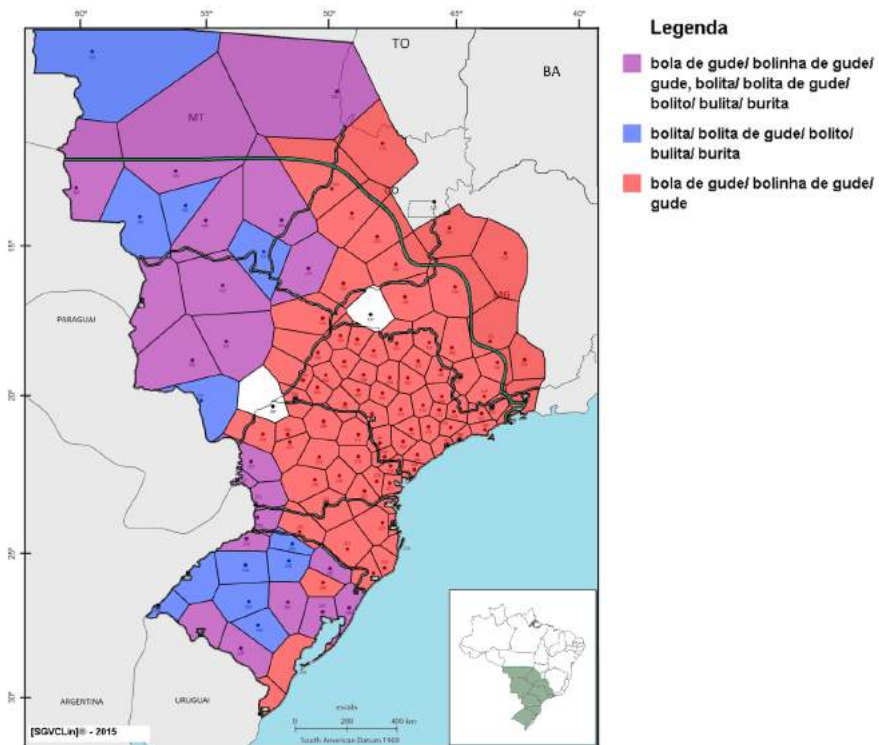
Fonte: Banco de dados do ALiB (2015)

Esse panorama geolinguístico revela traços de empréstimos lexicais do português com línguas de contato fronteiriço, no caso o espanhol. No *Dicionário da Real Academia Española* (2001), *bolita* apresenta uma remissão à forma utilizada na Argentina, *canica*, cujas acepções são “Juego de niños que se hace con bolas pequeñas de barro, vidrio u otra materia dura” ou mesmo, “cada una de estas bolas”. De acordo com o DRAE (2001), o vocábulo *canica* vem do francês *canique*, que, por sua vez, é originado do neerlandês *kinniker*, derivado do verbo *knikkerr*, que significa quebrar, romper.

Ou seja, a forma *bolita* no português evidencia traços das línguas em contato nessa região do país. Nem Houaiss e Villar (2001) nem Ferreira (2004) apresentam a datação desse vocábulo na língua portuguesa, porém constata-se que a variante faz parte da norma lexical da região Sul e dos estados do MS e MT, principalmente na região de

fronteira desses estados com países latino-americanos, embora também possa ocorrer em outros estados. No caso do PR, SC e RS, no contato com a Argentina e Uruguai, e no MS e MT, no contato com a Bolívia e o Paraguai. A Figura 13 apresenta uma carta linguística em que se identificam as áreas de *bolinha de gude* e *bolita*.

FIGURA 13 – Arealidade das variantes *bola de gude* e *bolita*



Fonte: Banco de dados do ALiB (2015)

Observa-se uma ampla área lexical para *bolinha de gude* (e formas agrupadas), compreendendo a faixa leste do território investigado correspondente aos territórios de SP, MG, grande parte de GO, noroeste do MS e grande parte do PR e SC. Já a porção oeste do território apresenta dois panoramas: (i) uma faixa territorial, correspondente à cor roxa, que revela linhas de heterolécia, delimitando a coocorrência das duas formas (*bolinha de gude* e *bolita*). Essas duas variantes são indistintamente

utilizadas em grande parte do território do MT e MS, no oeste do PR e SC e no noroeste, nordeste e extremo sul do RS; (ii) áreas que correspondem às isoléxicas de *bolita*, ou seja, em que ocorre somente esse item, compreendendo o RS, numa faixa que percorre do sudoeste desse estado, adentra o interior e chega ao norte gaúcho. No MS, essa variante restringe-se às redondezas de Ponta Porã – ponto 117. No MT, apresenta-se em três regiões: no ponto 111 – Alto Araguaia – fronteira com MS e GO; nos pontos 108 (Cuiabá) – Centro-sul do estado – e 110 (Cáceres), no sudoeste matogrossense; e no ponto de controle 103 (Aripuanã), localizado fora da área do subfalar sulista, no noroeste do estado.

Nos limites deste artigo, não são abordadas as demais variantes, mas os dados evidenciam heterogeneidade linguística dentro do território investigado e destacam, novamente, a existência de dois falares: um *paulista* (*bolinha de gude*) e o *sulista* (*bolita*). O *falar paulista* distribui-se na parte setentrional da divisão de Nascentes (1953). Compreende o estado de SP, sul de MG, centro e interior de GO, além do noroeste e centro do PR, onde ocorre, predominantemente, a forma padrão: *bola/bolinha de gude*. O *falar sulista*, por sua vez, está no RS, de onde, por um corredor lateral do oeste catarinense e paranaense, pode atingir os estados do MS e MT. Essa distribuição pode ser justificada pelos fluxos migratórios dos gaúchos rumo às novas fronteiras agrícolas do Centro-Oeste. Para esse item específico, esse falar está caracterizado pelo contato do português com outras línguas, no caso, o espanhol. Os limites desses falares têm-se-mostrado fluidos, e cabe notar que cada um deles apresenta outras áreas lexicais.

4.4 Menino/guri

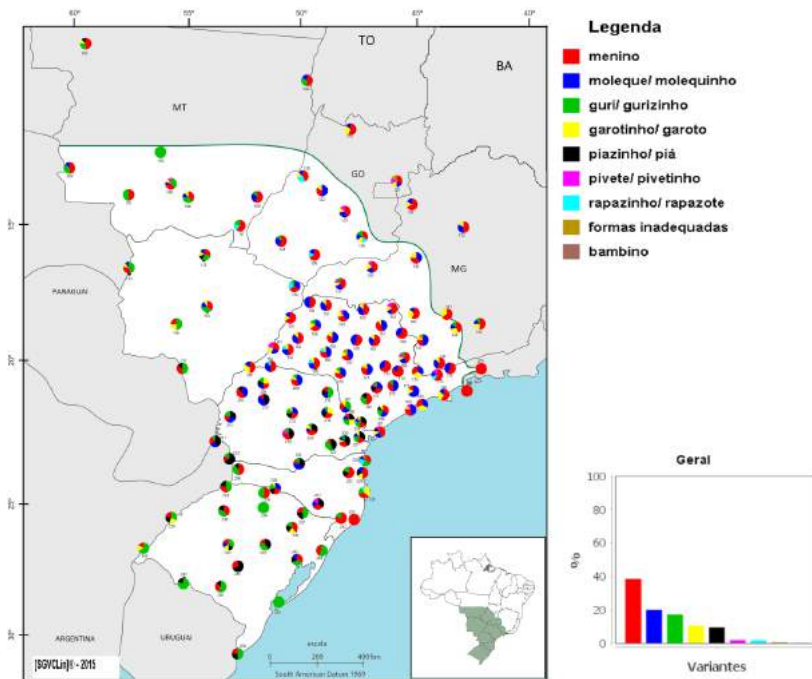
A questão 132 do QSL, que objetiva documentar as variantes para a criança de 5 a 10 anos do sexo masculino, apresenta oito variantes, além de formas inadequadas e abstenção de resposta.

Do conjunto das respostas, são considerados na mesma categoria os itens correspondentes às variantes com a flexão de diminutivo: *moleque/molequinho*; *guri/gurizinho*, *garoto/garotinho*, *piá/piazinho*. As formas *rapazinho/rapazote* também correspondem à mesma variante, e os itens lexicais *pirralho*, *furete* e *fedelho* são considerados como formas inadequadas, uma vez que figuraram como segunda ou terceira resposta na fala de quatro informantes e não são difundidas, salvo em situações de interlocução para acentuar desprezo e características disfêmicas em relação à criança.

A variante mais produtiva, *menino*, representa 38,55% do corpus, seguida por *moleque* (20,14%), *guri* (17,15%), *garoto* (10,47%) e *piá* (9,55%). Em menor índice, encontram-se *pivete* (1,84%), *rapazinho* (1,73%) e, com menos de 1% de produtividade, ocorrem “formas inadequadas” (0,46%) e a *hápaxlegomena*, *bambino*, com 0,12% de representatividade.

A Figura 14 apresenta a distribuição diatópica das nove formas, mostrando ampla distribuição da variante *menino*, presente em quase todas as localidades, com exceção de três cidades gaúchas (ponto 236, Passo Fundo; 247, Santana do Livramento; e 249, São José do Norte) e um município mato-grossense: 105, Diamantino. Nesta carta, observa-se maior concentração de *moleque* no estado de SP, estendendo-se para MG e GO. Por outro lado, a forma *guri* está amplamente distribuída pelo RS, adentrando o MS e MT. No PR e SC, nota-se maior distribuição de *piá*.

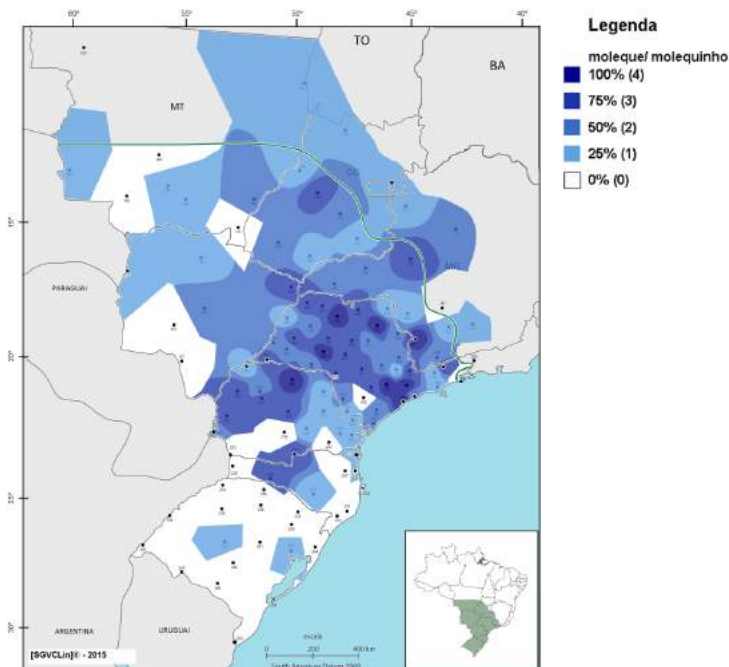
FIGURA 14 – Distribuição diatópica das variantes lexicais para a questão 132 do QSL



Fonte: Banco de dados do ALiB (2015)

Quanto à variante *moleque*, segunda forma mais produtiva, observa-se que esse item apresenta-se com índices percentuais diferenciados (FIG. 15), obtendo baixa produtividade no MT e, principalmente, no RS, aparecendo com maior concentração no território paulista. Observa-se que o índice de 100% de produtividade esteve em um ponto do PR (208 – Nova Londrina), em uma localidade mineira (147 – Poços de Caldas) e em seis localidades paulistas (152 – São José do Rio Preto, 157 – Ribeirão Preto, 166 – Marília, 177 – Itapetininga, 178 – Sorocaba e 183 – Itanhaém). Na maior parte do território paulista, *moleque* figura com 75% atingindo, sobretudo, o noroeste do PR, com o mesmo percentual, e uma faixa central de SC que vai do ponto 225 (Porto União) ao 229 (Concórdia). A variante *moleque* irradia-se também em uma longa faixa territorial do sul de MG, nos estados do MS e MT com 50% e atinge 25% de produtividade nas regiões mais distantes em relação ao estado de SP, por exemplo, localidades gaúchas e pontos situados ao norte do MS, MT e GO.

FIGURA 15 – Arealidade gradual da variante *moleque/molequinho*



Fonte: Banco de dados do ALiB (2015)

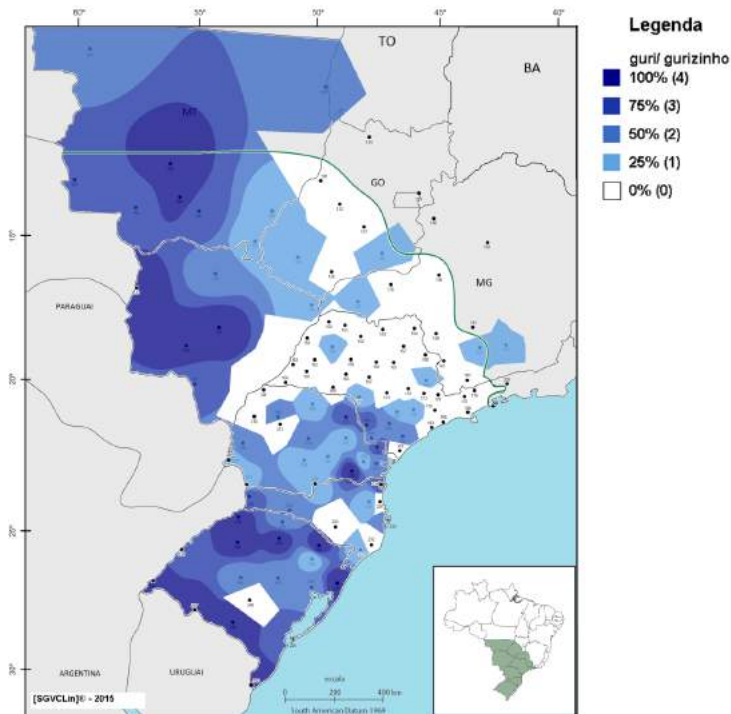
Nas cartas 270 e 271 do ALERS (ALTENHOFEN; KLASSMAN, 2011), encontra-se *moleque* como o quarto item da legenda, não se constituindo como uma forma tão produtiva, o que é compensado pela ocorrência de variantes regionais nos estados do Sul.

Romano e Seabra (2014a) confirmam que o item lexical evidencia a influência das línguas africanas no léxico do português. Cunha (1986, p. 528) registra que esse vocábulo veio do quimbundo mu'leke, para designar o “menino”, “rapazote”, com datação na língua portuguesa de 1731. Ferreira (2004) apresenta dez acepções para o vocábulo, entre as quais figura “Menino de pouca idade” (FERREIRA, 2004). Houaiss e Villar (2001), por sua vez, apresentam 17 acepções para “moleque”, das quais 13 são para emprego como substantivo masculino e quatro como adjetivos. Dessa última obra, merecem destaque as acepções: “1. menino novo, de raça negra ou mista. 2. garoto de pouca idade. 3. menino criado à solta; menino de rua. 4. garoto travesso” (HOUAISS; VILLAR, 2001).

A terceira variante mais produtiva é *guri*, forma majoritária no RS (48,57%), o que compensa a baixa produtividade de *moleque* nas cidades gaúchas. *Guri* concentra-se, além do RS, nos MS (36%) e MT 37,1%. Nos demais estados, há decréscimo na produtividade da variante, representando 17,57% das respostas dos catarinenses e 16,45% dos paranaenses. Em SP, MG e GO, a ocorrência do item é pequena (abaixo de 5%).

A Figura 16 apresenta a distribuição do item no território, revelando que essa variante se irradia a partir do território gaúcho, e, por um corredor oeste catarinense, atinge o MS e o MT e, a partir deste, chega ao sudoeste goiano refletindo-se em uma localidade do Triângulo Mineiro – ponto 137 (Campina Verde). No PR, *guri* está em quase todo o estado, exceto localidades do noroeste paranaense, região de colonização mais recente sob influência de mineiros e paulistas na primeira metade do século XX.

FIGURA 16 – Arealidade gradual de guri/gurizinho



Fonte: Banco de dados do ALiB (2015)

Por outro lado, essa variante reflete-se no estado de SP na região limítrofe com o PR, além de se difundir para as localidades situadas no Vale do Ribeira, conforme relatam Romano e Seabra (2014a, p. 485):

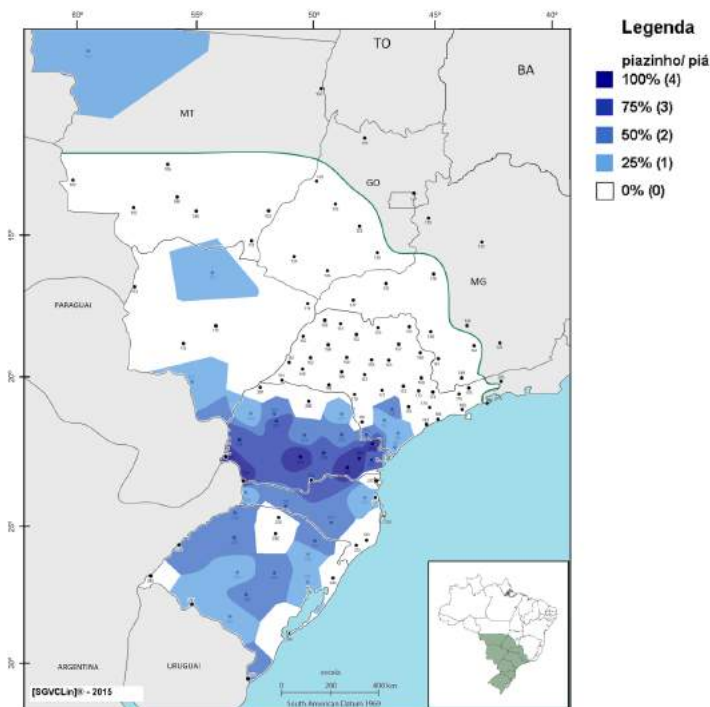
a presença da variante “guri” em determinadas localidades paulistas, tais como Itararé, Itapetininga e Sorocaba, pode ser explicada pela influência vinda do sul do país, haja vista que algumas dessas localidades eram pontos pertencentes à antiga rota proveniente do Caminho das Tropas, pela consequência do Tropeirismo.

A presença de *guri* no MS e no MT, ainda de acordo com esses autores, revela traços dos movimentos recentes de migração de gaúchos que deixaram marcas no léxico da região, bem como influências na formação da cultura local. Esse movimento decorre das novas fronteiras

agrícolas na região Centro-Oeste, sobretudo a partir da década de 1970, com o incentivo governamental.

A variante *piá* também é mais frequente na Região Sul, principalmente, no PR. Em SP, MT e MS, a frequência desse item é pequena, ao passo que, no PR, *piá* ocorre como variante majoritária (28,95%), seguindo-se o RS (18,1%) e SC (13,51%). Na Figura 17, observa-se a distribuição desse item na constituição de uma ampla área lexical na Região Sul. Entretanto, essa área não contempla algumas cidades gaúchas: Erechim (235) e Passo Fundo (236), no norte gaúcho, além dos pontos 245 – Uruguaiana (sudoeste) e 249 – São José do Norte (litoral sul). Observa-se a ausência de *piá* no nordeste do estado (ponto 244 – Osório), no litoral catarinense, pontos – 230 (Florianópolis), 232 (Tubarão), 233 (Criciúma), além do ponto 225 (São Francisco do Sul), no nordeste de SC.

FIGURA 17 – Arealidade gradual da variante *piá*



Fonte: Banco de dados do ALiB (2015)

A distribuição de *piá* não contempla integralmente o noroeste paranaense, de colonização mineira, expandindo-se, no entanto, para uma localidade sul-mato-grossense – 117 (Ponta Porã) e cidades paulistas localizadas no Vale do Ribeira. Assim como ocorre com a variante *guri*, a presença de *piá* nesta área de SP evidencia traços da história social das localidades, uma vez que essa forma típica da Região Sul, possivelmente, teria sido trazida para essa região paulista por meio dos tropeiros, durante o ciclo do Tropeirismo, persistindo até hoje no vocabulário ativo dos naturais dessa região paulista.

O ALPR (AGUILERA, 1994) e o ALPR II (ALTINO, 2007) não apresentam cartas linguísticas que documentem os designativos para o referente. O ALERS (ALTENHOFEN; KLASSMANN, 2011), entretanto, registra, nas cartas 270 e 271, as principais variantes e outras denominações, coincidindo com os dados do ALiB.

Conforme Romano e Seabra (2014a), tanto a variante *guri* quanto *piá* reiteram a influência das línguas indígenas no português. De acordo com Cunha (2010), a forma lexical “guri” veio do tupi *ũ’ri*, que designa o “bagre novo” (tipo de peixe), por extensão de sentido, a criança. O etimologista remete o usuário ao verbete “guiri - ‘sm bagre’ / curi 1587, guori datada, aproximadamente, de 1631”. Houaiss e Villar (2001), em contrapartida, registram a datação de 1890. Para os lexicógrafos, trata-se de um regionalismo brasileiro para designar o “menino”. Ferreira (2004), apesar de apresentar a mesma acepção, traz outra etimologia. Segundo o dicionarista, o vocábulo “guri” também vem do tupi, porém com o sentido de pequeno, não fazendo alusão ao peixe.

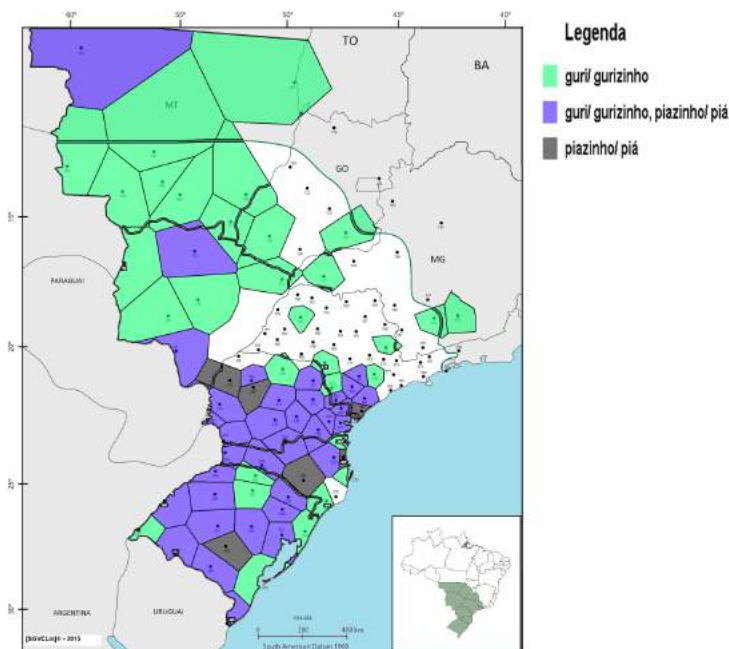
Piá, de acordo com Houaiss e Villar (2001), refere-se ao menino indígena e, analogamente, é empregado para denominar o menino mestiço de indígena com branco ou mesmo qualquer criança do sexo masculino. Segundo esses lexicógrafos, trata-se de um regionalismo de SC e RS para designar o peão menor de idade que não é de raça branca. Ferreira (2004) traz para *piá* a informação de que se trata de um brasileirismo para designar o índio jovem ou mestiço jovem de branco com índio, remetendo ao verbete “menino”, o que revela tratar-se, pois, de um sinônimo. Em SC e RS é uma variante usada para designar “qualquer menor que não é branco e trabalha como peão de estância” (FERREIRA, 2004), ou seja, a mesma acepção dada por Houaiss e Villar (2001).

Voltando-se à distribuição diatópica, observa-se uma ampla área na Região Sul em que ocorrem, indistintamente, as variantes *guri*

e *piá* (FIG. 18), o que ratifica o caráter regional atribuído a esses itens. *Guri* apresenta-se como forma exclusiva em áreas isoladas em SP, MG e GO. Em grande parte do território do MS e MT, observa-se a ampla distribuição desse item. *Piá*, por sua vez, não apresenta distribuição exclusiva no território, não formando grandes áreas lexicais, uma vez que ocorre isoladamente em pontos do RS, SC, PR e SP.

Observando-se o comportamento e a distribuição de três variantes lexicais: *moleque*, *guri* e *piá*, podem-se fazer algumas considerações acerca das áreas em que ocorrem, caracterizando dois padrões de variação lexical: um de influência paulista, denominado *falar paulista*, em que se verifica, entre outras variantes, a maciça presença de *moleque*, que se expande para outros estados como MG e GO; outro, de influência sul-rio-grandense e paranaense, denominado *falar sulista*, em que se observa a difusão das formas *guri* e *piá*, pela Região Sul, com reflexos no MS e MT. Cumpre afirmar que os limites desses dois grandes falares são fluidos, ora apresentando maior número de heteroléxicas, ora menor número.

FIGURA 18 – Arealidade das variantes guri/gurizinho e piá/piazinho

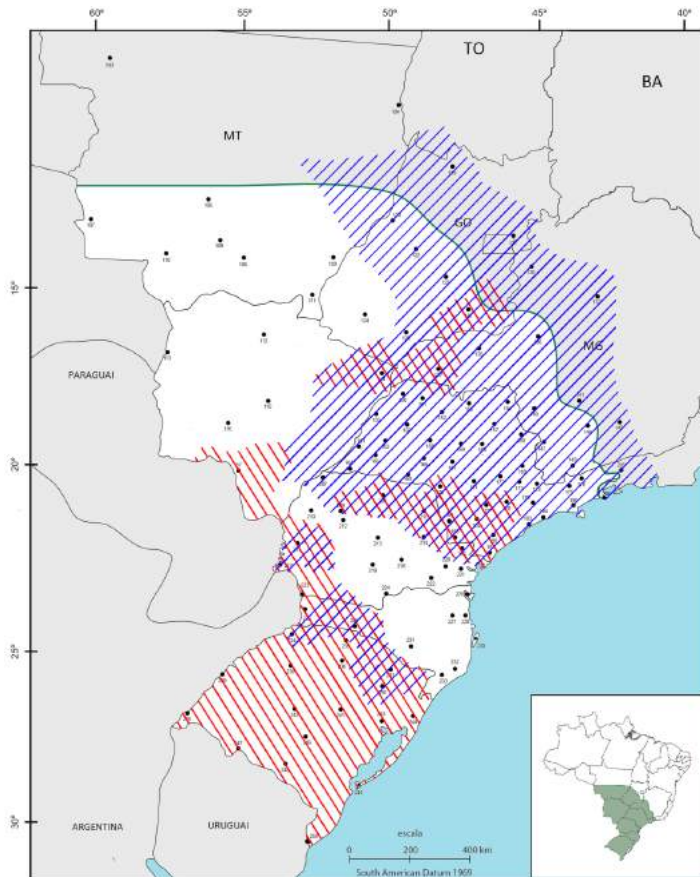


Fonte: Banco de dados do ALiB (2015)

5 O falar paulista e o falar sulista

A análise do *corpus* dá indícios da existência de dois grandes falares na área geográfica do subfalar sulista: o *falar paulista* e o *falar sulista*, conforme se sintetiza na Figura 19.

FIGURA19 – Localização geográfica dos falares *sulista* e *paulista*



-  Falar Sulista
-  Falar Paulista

Fonte: Elaborado pelo autor

O *falar paulista* difunde-se a partir do estado de SP, com influência na região norte do PR, oeste do MS, sudoeste e interior de GO, sul de MG e Triângulo Mineiro. Esse falar alcança os pontos de controle, não se limitando à área geográfica do subfalar sulista definida por Nascentes (1953) e diferencia-se da região meridional do território. Portanto, sob o ponto de vista do léxico, considerar o estado de SP e adjacências no mesmo grupo que o RS, possivelmente seria um equívoco, tendo em vista aspectos sócio-históricos envolvidos no processo de ocupação e povoamento do território, o que se reflete na variação lexical das questões abordadas.

Na amostra selecionada, o *falar paulista* caracteriza-se por uma maior homogeneidade lexical, revelando um menor número de coocorrência de variantes. Na área do *falar paulista*, há o predomínio das formas mais produtivas: *córrego*, *geleia*, *bolinha de gude*, *menino*, *moleque* que, à medida que avançam para a porção sul do país, apresentam uma diminuição em sua produtividade. Esse *falar paulista*, provavelmente, tem suas origens no elemento bandeirante, homem em geral de origem lusa que, em processo de miscigenação constante com o índio e, posteriormente, com o negro, difundiu a língua portuguesa para o interior do país.

Os limites que definem o *falar paulista* são virtuais e fluidos, ora alcançando toda a região Centro-Oeste, ora adentrando o PR, e, por um corredor central em SC (região dos Campos de Lajes), atingindo o norte do RS e, em alguns casos, expandindo-se ao sudoeste paranaense e oeste catarinense.

O *falar sulista*, por sua vez, localiza-se principalmente na porção meridional do Brasil, contemplando o RS e, por um corredor do oeste catarinense e sudoeste paranaense, atinge o MS. Esse corredor lateral em SC e PR já fora identificado sob o ponto de vista fonético e morfossintático, inicialmente, por Koch (2000), posteriormente, por Altenhofen (2005) e, mais recentemente, sob a perspectiva lexical, por Romano e Aguilera (2014). Trata-se, portanto, de um falar de influência sul-rio-grandense, que revela o contato do português com o espanhol em áreas de fronteira, e também de contato com línguas de imigração, o que se evidencia por variantes presentes na norma lexical, como *sanga*, *arroyo*, *chimia*, *bolita*.

Esse falar também apresenta variantes que revelam o contato do português com o tupi, exemplificado por *guri* e *piá*. No tocante a esses itens, o *falar sulista* alcança o MT e MS, em consequência de correntes migratórias mais recentes. Os limites desse falar também são virtuais e

fluidos e, além de se expandir para o Centro-Oeste, adentra SC e, por um corredor central do PR, passando em localidades como Lapa e Piraí do Sul, atinge cidades paulistas localizadas no Vale do Ribeira até chegar a Itapetininga e Sorocaba. Essa configuração diatópica pode evidenciar traços da história social dessas localidades, uma vez que revela rastros do tropeirismo, ou seja, os tropeiros gaúchos, além de heranças deixadas na cultura material de determinadas cidades, deixaram também marcas na norma lexical que até hoje ecoam no vocabulário ativo de informantes.

Além desses dois falares, há indícios de subáreas lexicais no território do PR e SC, com a presença de itens como *valeta*, formas genéricas para a *geleia*, *musse*, *clica*, *búrica*, *piá*, que merecem atenção mais detalhada. São denominadas subáreas pelo fato de ora se identificarem com o *falar paulista*, ora com o *falar sulista*, figurando em uma área geográfica comum aos dois grandes falares. Nos limites desta análise, sob o ponto de vista lexical, este estudo vem ratificar que, atualmente, o mapa dialetal do Brasil estabelecido por Nascentes (1953) merece um olhar mais atento, haja vista a heterogeneidade lexical atestada nessa região do país.

Referências

AGUILERA, V. de A. *Atlas Linguístico do Paraná*. Curitiba: Imprensa Oficial, 1994.

ALMEIDA, A. de. *Vida e morte do tropeiro*. São Paulo: Martins; EDUSP, 1981.

ALTINO, F. C. *Atlas Linguístico do Paraná II*. 2007. v. 1: 182f; v. 2: 347f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2007.

AMARAL, A. *O dialeto capira: gramática, vocabulário*. 4. ed. São Paulo: HUCITEC; Brasília: INL, 1981 [1920].

ALTENHOFEN, C. Áreas linguísticas do português falado no sul do Brasil: um balanço das fotografias geolinguísticas do ALERS. In: AGUILERA, V. de A. (Org.). *A geolinguística no Brasil: trilhas seguidas, caminhos a percorrer*. Londrina: Eduel, 2005. p. 177-208.

ALTENHOFEN, C. V.; KLASSMANN, M. S. (Org.). *Atlas Linguístico-Etnográfico da Região Sul do Brasil – ALERS: Cartas Semântico-Lexicais*. Porto Alegre: Editora da UFRGS; Florianópolis: Ed. UFSC, 2011.

CARDOSO, S. A. M. *Tinha Nascentes razão? Considerações sobre a divisão dialetal do Brasil. Estudos Linguísticos e Literários*, Salvador, UFBA, n. 5, p. 47-59, 1986.

CARDOSO, S. A. M. Vogais médias pretônicas no Brasil: uma visão diatópica. In: AGUILERA, V. de A. (Org.). *Português no Brasil: estudos fonéticos e fonológicos*. Londrina: EDUEL, 1999. p. 95-108.

CARRETER, F. L. *Diccionario de términos filológicos*. Madrid: Gredos, 2008.

CHAMBERS, J. K; TRUDGILL, P. *La Dialectología*. Trad. Carmén Morán González. Madrid: Visor Libros, 1994.

COMITÊ NACIONAL DO PROJETO ALiB. *Atlas Linguístico do Brasil: Questionários 2001*. Londrina: EDUEL, 2001.

COSERIU, E. *O homem e sua linguagem*. Trad. Carlos A. Fonseca e Mário Ferreira. Rio de Janeiro: Presença, 1987.

CUBA, M. A. *Atlas Linguístico da Mesorregião Sudeste de Mato Grosso*. 2009. v. 1: 163f.; v. 2: 521f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Três Lagoas, 2009.

CUNHA, A. G. da. *Dicionário etimológico da língua portuguesa*. 4. ed. Rio de Janeiro: Lexikon, 2010.

D'ANUNCIÇÃO, E. S. *Registrando o léxico dos brinquedos infantis em Minas Gerais*. 2016. 86 f. (Monografia) – Instituto de Letras, Universidade Federal da Bahia, Salvador, 2016.

FERREIRA, A. B. de H. *Novo Dicionário da Língua Portuguesa*. Versão eletrônica 5.0. 3. ed. Curitiba: Positivo, 2004.

FERREIRA, C.; CARDOSO, S. A. M. *A dialetologia no Brasil*. São Paulo: Contexto, 1994.

FREITAS-MARINS, L. G. *O rural e o urbano: novos e velhos falares na região Centro-Oeste do Brasil*. 2012. 310f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2012.

HOUAISS, A.; VILLAR, M. de S. *Dicionário Eletrônico Houaiss da Língua Portuguesa*. Versão 1.0. Rio de Janeiro: Objetiva, 2001.

KOCH, W. O povoamento do território e a formação de áreas linguísticas. In: GÄRTNER, E.; HUNDT, C. SCHÖNBERGER, A. (Ed.). *Estudos de geolinguística do português Americano*. Franckfurt am Maim: TFM, 2000. p. 55-69.

MOTA, J. Áreas dialetais brasileiras. In: CARDOSO, S. A. M; MOTA, J.; MATTOS E SILVA, R. V. *Quinhentos anos de história linguística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p. 321-357.

NASCENTES, A. *O linguajar carioca*. 2. ed. Rio de Janeiro: Organizações Simões, 1953 [1922].

OLIVEIRA, D. G. de (Org.). *ALMS – Atlas Linguístico do Mato Grosso do Sul*. Campo Grande: Editora UFMS, 2007.

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

RAZKY, A. A dimensão sociodialetal do léxico no Projeto Atlas Linguístico do Brasil. *Signum: Estudos da Linguagem*, Londrina, UEL, v. 13, n. 2, p. 247-270, 2013. Doi: <http://dx.doi.org/10.5433/2237-4876.2013v16n2p247>.

REALACADEMIA ESPAÑOLA. Dicionário de la lengua española. 22. ed. Madrid: Real Academia Española, 2001.

REIS, R. C. P. *Atlas Linguístico do município de Ponta-Porã-MS: um registro das línguas em contato na fronteira do Brasil com o Paraguai*. 2006. v. 1: 214f.; v. 2: 253f. Dissertação (Mestrado em Estudos da Linguagem – Universidade Federal do Mato Grosso do Sul, Três Lagoas, 2006.

RIBEIRO, D. *O povo brasileiro: formação e sentido do Brasil*. São Paulo: Companhia das Letras, 2006.

RIBEIRO, S. S. C. *Brinquedos e brincadeiras infantis na área do falar baiano*. 2012. 466f. Tese (Doutorado em Linguística) – Universidade Federal de Bahia, Salvador, 2012.

ROCHA, P. G. da. *O português de contato com o espanhol no sul do Brasil: empréstimos lexicais*. 2008. 148f. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2008.

ROCHE, J. *A colonização alemã e o Rio Grande do Sul*. Trad. Emery Ruas. Porto Alegre: Globo, 1969.

ROMANO, V. P. *Em busca de falares a partir de áreas lexicais no centro-sul do Brasil*. 2015. v.1: 286f.; v. 2: 117f. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2015.

ROMANO, V. P.; AGUILERA, V. de A. Padrões de variação lexical no sul do Brasil a partir dos dados do Projeto ALiB. *Estudos Linguísticos*, São Paulo, v. 43, n. 1, p. 575-587, 2014.

ROMANO, V. P.; SEABRA, R. D. Menino, guri ou piá? Um estudo diatópico nas regiões Centro-Oeste, Sudeste e Sul a partir dos dados do Projeto Atlas Linguístico do Brasil. *Alfa: Revista de Linguística*, São José do Rio Preto, UNESP, v. 58, n. 2, p. 463-497, 2014a. Doi: <https://doi.org/10.1590/1981-5794-1405-9>.

ROMANO, V. P.; SEABRA, R. D. Dados geolinguísticos sob uma perspectiva estatística: a variação lexical no Centro-Oeste, Sudeste e Sul do Brasil. *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG, v. 22, n. 2, p. 59-92, 2014b. Doi: <http://dx.doi.org/10.17851/2237-2083.22.2.59-92>.

ROMANO, V. P.; SEABRA, R. D. Do presente para o passado: a variação lexical em Minas Gerais a partir de corpora geolinguísticos sobre brinquedos infantis. *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG, v. 25, n. 1, p. 111-150, 2017. Doi: <http://dx.doi.org/10.17851/2237-2083.25.1.111-150>.

ROMANO, V. P.; SEABRA, R. D.; OLIVEIRA, N. [SGVCLin] – Software para geração e visualização de cartas linguísticas. *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG, v. 22, n. 1, p. 119-151, 2014. Doi: <http://dx.doi.org/10.17851/2237-2083.22.1.119-151>.

SANTOS, L. A. dos. *Brincando pelos caminhos do falar fluminense*. 2016. 197f. Dissertação (Mestrado em Língua e Cultura) – Universidade Federal da Bahia, Salvador, 2016.