



UNIVERSIDADE
ESTADUAL de LONDRINA

VALTER PEREIRA ROMANO

**EM BUSCA DE FALARES A PARTIR DE ÁREAS LEXICAIS
NO CENTRO-SUL DO BRASIL**

(Volume 1)

Londrina
2015

VALTER PEREIRA ROMANO

**EM BUSCA DE FALARES A PARTIR DE ÁREAS LEXICAIS
NO CENTRO-SUL DO BRASIL**

(Volume 1)

Tese apresentada ao Programa de Pós-graduação em Estudos da Linguagem, da Universidade Estadual de Londrina, como requisito à obtenção do título de Doutor.

Orientadora: Prof.^a Dr.^a Vanderci de Andrade Aguilera .

Londrina
2015

Dados Internacionais de Catalogação-na-Publicação (CIP)

R759e Romano, Valter Pereira.
Em busca de falares a partir de áreas lexicais no centro-sul do Brasil /
Valter Pereira Romano. - Londrina, 2015.
2 v. : il.

Orientador: Vanderci de Andrade Aguilera
Tese (Doutorado em Estudos da Linguagem) - Universidade Estadual de
Londrina, Centro de Letras e Ciências Humanas, Programa de Pós-Graduação em
Estudos da Linguagem, 2015.
Inclui bibliografia.

1. Língua portuguesa x Variação - Teses. 2. Língua portuguesa x Dialetos z
Brasil, Centro-Sul. - Teses I. Aguilera, Vanderci de Andrade. II. Universidade
Estadual de Londrina. Centro de Letras e Ciências Humanas. Programa de Pós-
Graduação em Estudos da Linguagem. III. Título.

CDU 806.90(81)-087.9

VALTER PEREIRA ROMANO

**EM BUSCA DE FALARES A PARTIR DE ÁREAS LEXICAIS NO
CENTRO-SUL DO BRASIL**

Tese apresentada ao Programa de Pós-graduação em Estudos da Linguagem, da Universidade Estadual de Londrina, como requisito à obtenção do título de Doutor.

BANCA EXAMINADORA

Orientador: Prof.^a Dr.^a Vanderci de Andrade
Aguilera
Universidade Estadual de Londrina - UEL

Prof.^a Dr.^a Vansersí Sant'Ana Castro
Universidade de Campinas - UNICAMP

Prof. Dr. Felício Wessling Margotti
Universidade Federal de Santa Catarina -
UFSC

Prof.^a Dr.^a Jacyra Andrade Mota
Universidade Federal da Bahia - UFBA

Prof.^a Dr.^a Aparecida Negri Isquerdo
Universidade Federal do Mato Grosso do Sul -
UFMS

Londrina, 09 de março de 2015.

AGRADECIMENTOS

Olhar para trás após uma longa caminhada pode fazer perder a noção da distância que percorremos, mas se nos detivermos em nossa imagem, quando a iniciamos e ao término, certamente nos lembraremos o quanto nos custou chegar até o ponto final, e hoje temos a impressão de que tudo começou ontem... (Guimarães Rosa)

Parece que começou ontem, não a tese, mas a caminhada pela Linguística, em especial pela Dialetologia. A feliz escolha em seguir pelo caminho mais difícil, mais apertado, mais desafiador, fez-me chegar até aqui. Transpor as barreiras e contornar os obstáculos foram apenas algumas estratégias para chegar ao resultado que tenho hoje diante de mim. Primeiro veio a graduação, a formação inicial, muito bem amparado, logo de início, por minha primeira orientadora, Dr.^a Aparecida Negri Isquierdo, prof.^a Cida, que me conduziu no primeiro trabalho geolinguístico. Dado o início, não mais voltaria, não tomaria outro caminho senão aquele que levaria à minha própria realização pessoal e profissional.

Depois veio a honra de trabalhar com a estimada prof.^a Vanderci, figura fundamental em minha vida. O exemplo, o carinho, o respeito e a acolhida junto aos seus fez-me persistir cada vez mais no aprimoramento, na superação e na conquista dos meus objetivos. Parafraseando Guimarães Rosa, Vanderci, *se a gente quer atravessar um rio a nado, a gente passa* e muitas vezes isso foi necessário, nadar contra a correnteza, seguir na contra mão dos fatos e das circunstâncias e necessidades materiais que se impunham diante de mim. Lembro-me, agora, dos anos na moradia estudantil, da sala de projetos 169, das infindáveis transcrições, todas agora retomadas na tese, dos relatórios de bolsa de inclusão... dos almoços e jantares em sua casa... Minha história a senhora conhece, sabe das lutas, sabe dos sonhos e sabe da obstinação. Aqui se encerra mais um ciclo... Se necessário fosse, começaria tudo novamente, porque *o que a vida quer da gente é coragem*, não é?! *E todo caminho da gente é resvaloso. Mas também, cair não prejudica demais - a gente levanta, a gente sobe, a gente volta!*... Esses anos de convivência com a senhora me deram a certeza de que, quando se tem confiança, fé e esperança em concretizar nossos objetivos, um manancial de força, coragem e segurança se constrói dentro de nós.

Expressar o significado desse trabalho vai muito além de sua possível relevância acadêmica. Trata-se do resultado de um período de convivência e amadurecimento de um espírito obstinado em busca de sua realização pessoal e de

seu próprio caminho. Convivência esta que se deu em meio a um renomado projeto de pesquisa, o ALiB, do qual tanto me orgulho por fazer parte, e que agora tem mais um dos seus, formado! Formei-me como linguista e como dialetólogo neste projeto.

Nesse sentido, quero agradecer aos competentíssimos pesquisadores que integram esta equipe e ao Comitê Nacional, que, sob a figura da professora Suzana Cardoso, permitiu a utilização dos dados ainda inéditos em mais este trabalho. Espero, assim, pelo menos em parte, contribuir para o desenvolvimento deste nosso atlas que revela o quão colorido e fascinante é o estudo de nossa língua. Tenham certeza que serei um dos continuadores de Geolinguística no Brasil.

Agradeço à minha sempre orientadora, Vanderci, que soube, com todo o entusiasmo que lhe é próprio e pelas suas sábias mãos, mostrar os caminhos desde o início da graduação, orientar-me nas iniciações científicas, no mestrado e agora no doutorado. Mais que orientadora, mentora intelectual e espiritual, que soube transmitir ensinamentos, confiança, vibrações do bem e entusiasmo com a pesquisa geolinguística. Soube ouvir, soube compreender, soube deixar seguir em frente, soube incentivar e jamais duvidou da capacidade de superação, e do espírito crítico deste que agora avança mais um degrau dos muitos que já foram subidos e dos muitos que ainda hei de subir. Vanderci, obrigado por não desamparar e estar sempre de prontidão a me atender, mesmo à distância.

Agradeço à prof.^a Cida, com toda sua sabedoria e indagações intrigantes, sempre me acompanhou de perto, me estimulou à pesquisa e ao amadurecimento teórico. Foi debatedora em ambos os SEDATAs e muito atenciosa na ocasião do exame de qualificação. Seus apontamentos durante todo o percurso foram importantes para o resultado que agora vos apresento. Sua existência em minha vida acadêmica é de excepcional importância! Deixo aqui expresso o imenso carinho e respeito que tenho por esta grande mulher.

Agradeço a outros dois professores de minha banca de qualificação: Prof. Felício Margotti, da UFSC, pelas proveitosas discussões que tivemos durante o III CIDS e aprofundamentos teóricos, e terminológicos, que, aliás, muito me ajudaram. Obrigado pelo estímulo e incentivo na caminhada geolinguística. E à Prof.^a Silvia Brandão, da UFRJ, exemplo de pesquisadora, cujos textos sempre impecáveis foram me ensinados a ler desde o início de minha formação. Suas contribuições no exame de qualificação foram muito importantes para mim!

À prof.^a Jacyra, da UFBA, meus sinceros agradecimentos por ter aceitado participar da banca de defesa e por todas as oportunidades no âmbito da pesquisa e do Projeto, saiba o quanto a admiro!

Agradeço à prof.^a Vandersí Sant'Ana Castro, da Unicamp, pela oportunidade de tê-la em minha banca de defesa. Quero que saiba da minha admiração e que é um referencial para mim, sobremaneira, pela primorosa escrita e clareza de seus textos.

Agradeço a duas professoras suplentes da banca de defesa, Prof.^a Suzana Cardoso, da UFBA, e prof.^a Fabiane Altino, da UEL.

Dois outras pessoas também foram fundamentais para o trabalho que agora se apresenta: Rodrigo Seabra e Nathan Oliveira.

Agradeço ao Rodrigo pelo apoio constante, pelo que é e o que significa. Agradeço por todo auxílio e ensinamentos nos tratamentos estatísticos, revelando-me o quão importante é este ponto de vista para as análises de dados linguísticos. Soube com muita maestria estabelecer o feliz diálogo entre a Linguística e a Computação, orientando, ensinando, aprendendo. Sem sua figura, o caminho teria sido bem mais difícil, a vida seria mais difícil!

Agradeço ao Nathan Oliveira, exímio desenvolvedor de sistemas, aplicado, aluno dedicado. Mais do que aluno, parceiro no projeto do software [SGVCLin]. Nathan, você foi fundamental para este trabalho, sua parceria foi e continua sendo de extrema importância. Minha sincera gratidão!

Agradeço, agora, a duas pessoas pelas quais sempre terei eterna admiração, respeito e gratidão, meus pais. Sou grato a eles pela educação que me deram, pela simplicidade de suas vidas e pela maravilhosa oportunidade de estar neste mundo. Agradeço à minha mãe, Joaquina Romano. Um presente em minha vida, tão doce e tão disposta ao diálogo. Exemplo de caráter, desde cedo me ensinou a importância do trabalho sério, honesto e do estudo. Obrigado por acreditar, por me incentivar, pois, mesmo abdicando da presença física de seu filho, sempre esteve presente em minha vida, deixando-me encontrar o meu próprio caminho, a minha direção. Agradeço ao meu pai, Adenir Romano, homem simples e trabalhador que, com toda a paciência que lhe é própria, acompanha-me nesta caminhada. Pai, mãe, sou grato por tudo que fizeram e fazem, obrigado por não duvidarem em momento algum de meu sucesso pessoal e profissional!

Agradeço aos meus irmãos Wagner Romano, Valber Romano e Viviane Romano que, com suas respectivas famílias, mesmo sem entender de que se tratava o meu estudo, sabiam da sua grandiosidade e a importância que eu lhe dava. Estamos cada um em uma cidade diferente, mas saibam que vocês são de fundamental importância para mim. Conto com vocês e contem comigo, sempre!

Agradeço a todos que acompanharam desde o início ou que de uns anos para cá vem acompanhando esta caminhada. D. Rose Seabra, Seu Mauro Seabra, obrigado pelo incentivo constante.

Agradeço aos amigos, colegas e professores da pós-graduação. Agradeço ao PPGEL e toda equipe deste Programa que cada vez mais caminha para a excelência em pesquisa no Brasil e à CAPES pelo auxílio financeiro que, pela bolsa de doutorado, fomentou minha pesquisa.

Agradeço ao Grande Deus, pela onipresença, pelo estímulo e força em todas as fases, em todos os degraus e pelo maravilhoso dom da vida. A minha gratidão a esse Deus que se fez presente, em forma de pessoas ou de circunstâncias, revelando-me que coincidências não são por acaso!

Olho para trás e vejo a imagem de quando comecei e, para finalizar, retomo mais uma vez Guimarães Rosa, porque agora fica a certeza de que *o mais importante e bonito, do mundo, é isto: que as pessoas não estão sempre iguais, ainda não foram terminadas - mas que elas vão sempre mudando (...) porque a cabeça da gente é uma só, e as coisas que há e que estão para haver são demais de muitas, muito maiores diferentes, e a gente tem de necessitar de aumentar a cabeça, para o total.* E, ao final, não fica a dúvida de que *eu quase que nada não sei, mas desconfio de muita coisa!* Em busca do constante aperfeiçoamento, sigo. Meu muito obrigado a todos!

ROMANO, Valter Pereira. **Em busca de falares a partir de áreas lexicais no Centro-Sul do Brasil**. 2015. 2v. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2015. Volume1: 296 f.. Volume 2: Uma carta introdutória, 71 cartas linguísticas e 10 relatórios.

RESUMO

A necessidade de estabelecer áreas dialetais brasileiras subjaz aos trabalhos de cunho dialetológico implícita ou explicitamente. Antenor Nascentes, em *O Linguajar Carioca* (1922), apresenta um mapa dialetológico do Brasil, reformulado em 1953. O autor divide o português brasileiro em seis subfalares compreendidos em dois grandes grupos – Norte e Sul. O primeiro engloba o subfalar Amazônico e Nordesteño e o segundo compreende os subfalares Baiano, Fluminense, Mineiro e Sulista. Soma-se a esse conjunto de falares uma área denominada como Território Incaracterístico. Discute-se, nesta pesquisa, a divisão dialetal de Antenor Nascentes (1953) no que se refere à área geográfica do subfalar sulista, sob a perspectiva lexical, com vistas a propor uma reformulação. Como objetivos específicos, busca-se revelar as diferenças lexicais dessa área geográfica e delimitar linhas de isoléxicas e/ou heteroléxicas que podem refletir traços da história social do português. A hipótese central do trabalho pauta-se no pressuposto de que a região correspondente ao subfalar sulista não apresenta homogeneidade no que diz respeito ao léxico. Para alcançar os objetivos estabelecidos, adotam-se os pressupostos teóricos e metodológicos da Dialetologia e da Geolinguística e utilizam-se como *corpus* de análise os dados do Projeto ALiB correspondentes a 118 municípios brasileiros, perfazendo o total de 472 informantes. A rede de pontos contempla os Estados do Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Mato Grosso do Sul e parte do território de Minas Gerais, Goiás e Mato Grosso. Foram selecionadas 5 questões do Questionário Semântico-Lexical para validar a hipótese da pesquisa, a saber: 001 – córrego; 039 – tangerina; 132 – menino; 156 – bolinha de gude e 177 – geleia. A partir da observação do comportamento diatópico das variantes documentadas para as questões supramencionadas e do tratamento estatístico para validar hipóteses sobre a homogeneidade ou heterogeneidade lexical na área do subfalar sulista, pode-se chegar à conclusão de que a porção setentrional do território investigado apresenta diferenças em relação à parte meridional, evidenciando a existência de dois grandes falares: o paulista e o sulista. Apesar de os limites e abrangência desses dois falares serem virtuais e fluidos, o falar paulista localiza-se, principalmente, no Estado de São Paulo e adjacências: norte do Paraná, oeste do Mato Grosso do Sul, sul de Goiás, de Minas Gerais e Triângulo Mineiro. Este falar caracteriza-se pelo menor número de variantes lexicais e áreas de coocorrência de itens, predominando o uso da forma considerada padrão. Já o falar sulista, de influência sul-rio-grandense, localiza-se no Estado do Rio Grande do Sul expandindo-se para o oeste de Santa Catarina e sudoeste paranaense. Este falar revela traços do português em contato com outras línguas por conta de movimentos migratórios e apresenta maior número de variantes lexicais delimitadas pelas linhas de heteroléxicas. A esses dois falares somam-se subáreas lexicais que ora se identificam com o falar paulista, ora com o sulista.

Palavras-chave: Falares. Centro-Sul do Brasil. Antenor Nascentes. Variação Lexical. Projeto ALiB.

ROMANO, Valter Pereira. **In search of speech from lexical areas in Central-South of Brazil**. 2015. 2v. Thesis (Doctoral Degree in Language Studies) – Universidade Estadual de Londrina, Londrina, 2015. Volume1: 296 p.. Volume 2: One introductory map, 71 linguistic maps and 10 reports.

ABSTRACT

The need to establish Brazilian dialect areas underlies the dialectological nature works implicitly or explicitly. Antenor Nascentes, in *O Linguajar Carioca* (1922), presents a dialectological map of Brazil, redesigned in 1953. The author divides Brazilian Portuguese into six types of subspeech included into two major groups – North and South. The first encompasses the subspeech Amazônico and Nordestino and the second comprises the subspeech Baiano, Fluminense, Mineiro and Sulista. Added to this set of speech, an area known as Território Incaracterístico. This research discusses the dialectal division of Antenor Nascentes (1953) with regard to the geographical area of the sulista subspeech under the lexical perspective, to propose a reformulation. As specific objectives, the research aims to reveal the lexical differences of this geographical area and to delimit isolexical and/or heterolexical lines that can reflect traces of the social history of the Portuguese language. The central hypothesis of work is guided in the assumption that the region corresponding to the sulista subspeech has no uniformity as regards the lexicon. To achieve the objectives established, the theoretical and methodological assumptions of Dialectology and Geolinguistics were adopted and the analytical corpus used was the data of the ALiB Project corresponding to 118 Brazilian municipalities, totaling 472 informants. The point network covers the states of Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Mato Grosso do Sul and part of the State of Minas Gerais, Goiás and Mato Grosso. Five questions from the Lexical-Semantic Questionnaire were selected to validate the research hypothesis, namely: 001 – córrego; 039 – tangerina; 132 – menino; 156 – bolinha de gude and 177 – geleia. From the observation of the diatopical behavior of variants documented for the above issues and from statistical treatment to validate assumptions regarding lexical homogeneity and/or heterogeneity in the area of sulista subspeech, the research concluded that the northern portion of the territory investigated presents differences in relation to the southern part, showing the existence of two large types of speech: the paulista and the sulista. In spite of the limits and scope of these two types of speech being virtual and fluid, the paulista speech is located mainly in the state of São Paulo and neighboring areas: the north of Paraná, the west of Mato Grosso do Sul, the south of Goiás, Minas Gerais and Triângulo Mineiro. This speech is characterized by a smaller number of lexical variants and areas of co-occurrence of items, and the predominant use of forms considered standard. In turn, the sulista speech of sul-rio-grandense influence is located in the state of Rio Grande do Sul expanding to the west of Santa Catarina and the southwest of Paraná. This speech reveals traces of the Portuguese language in contact with other languages due to immigration and has a higher number of lexical variants bounded by heterolexical lines. These two types of speech added to the lexical sub-areas, are sometimes identified with the paulista speech, sometimes with the sulista

Key words: Speech. Central-South of Brazil. Antenor Nascentes. Lexical Variation. ALiB Project.

LISTA DE FIGURAS

Figura 1 -	Principais rotas de atuação dos bandeirantes e os três principais ciclos das bandeiras	32
Figura 2 -	Mapa da localização das reduções jesuíticas e rota das bandeiras	42
Figura 3 -	Rotas dos tropeiros	52
Figura 4 -	Entrada de imigrantes no Brasil	53
Figura 5 -	Números de imigrantes subsidiados pelas Províncias de São Paulo entre 1882-1886.....	54
Figura 6 -	Origem dos povoadores de Santa Catarina	56
Figura 7 -	Imigração alemã no Brasil entre 1824-1869	60
Figura 8 -	Regiões de origem dos emigrantes italianos entre 1870-1920	60
Figura 9 -	Emigração italiana para o Brasil, segundo as regiões de procedência - período 1876/1920	62
Figura 10 -	Áreas geográficas da Região Sul do Brasil caracterizadas como bilíngues (português/italiano) nos dados do ALERS	65
Figura 11 -	Principais fluxos migratórios no Brasil entre 1950-1980	66
Figura 12 -	Migração interna (1990-2000).....	69
Figura 13 -	Mapa do crescimento da população segundo os municípios no período de 2000-2010.....	72
Figura 14 -	Gráfico comparativo da naturalidade dos brasileiros segundo o Estado federativo	103
Figura 15 -	Divisão dialetal do Brasil, de Julio Ribeiro (1881).....	104
Figura 16 -	Divisão Dialetal de Maximino Maciel (1950)	105
Figura 17 -	Divisão Dialetal de Rodolfo Garcia (1915)	106
Figura 18 -	Divisão Dialetal Antenor Nascentes (1922)	108
Figura 19 -	Divisão Dialetal de Antenor Nascentes (1953)	110
Figura 20 -	Carta linguística para questão 039 do QSL- Capitais	112
Figura 21 -	Carta linguística para questão 085 do QSL- Capitais	113
Figura 22 -	Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Centro-Oeste.....	114

Figura 23 -	Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Centro-Oeste.....	114
Figura 24 -	Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Sudeste	115
Figura 25 -	Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Sudeste.....	115
Figura 26 -	Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Sul	116
Figura 27 -	Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Sul.....	118
Figura 28 -	Isoléxica da variante moleque na rede de pontos do subfalar sulista.....	118
Figura 29 -	Isoléxica da variante guri na rede de pontos do subfalar sulista.....	119
Figura 30 -	Isoléxica da variante piá na rede de pontos do subfalar sulista.....	121
Figura 31 -	Carta linguística para os designativos de bala – QSL 185.....	122
Figura 32 -	Carta linguística para os designativos de geleia – QSL 177	123
Figura 33 -	Carta linguística para os designativos de menstruação – QSL 121	123
Figura 34 -	Carta linguística para os designativos de prostituta – QSL 142.....	125
Figura 35 -	Produtividade dos indigenismos segundo a área semântica	127
Figura 36 -	Distribuição diatópica das designações para pernilongo (QSL 88), segundo a base linguística	128
Figura 37 -	Carta linguística para os designativos de calçada – QSL 196	131
Figura 38 -	Subárea dialetal A no subfalar baiano	133
Figura 39 -	Subárea dialetal B no subfalar baiano	133
Figura 40 -	Subárea dialetal C no subfalar baiano	134
Figura 41 -	Subárea dialetal D no subfalar baiano	135
Figura 42 -	Logotipo do sistema [SGVCLin]	146

Figura 43 -	Tela principal do software	147
Figura 44 -	Exemplo de relatório geral por sexo	149
Figura 45 -	Exemplo de carta diatópica.....	149
Figura 46 -	Exemplo de seleção de gráfico	150
Figura 47 -	Exemplo de carta diatópica gerada para impressão	151
Figura 48 -	Exemplo de carta diatópica/diassexual	152
Figura 49 -	Exemplo de carta isoléxica	153
Figura 50 -	Produtividade da variante córrego por Estado	162
Figura 51 -	Zoom do Rio Grande do Sul - foco na produtividade do ponto 242 (Santa Maria) carta 1	163
Figura 52 -	Produtividade da variante riacho por Estado	164
Figura 53 -	Produtividade da variante riozinho por Estado.....	165
Figura 54 -	Produtividade da variante ribeirão por Estado	167
Figura 55 -	Principal rota dos tropeiros – Caminho de Viamão	169
Figura 56 -	Produtividade da variante sanga/sanguinha por Estado.....	171
Figura 57 -	Carta linguística 163 do ALPR	173
Figura 58 -	Carta linguística 007 do ALERS.....	174
Figura 59 -	Produtividade da variante valeta por Estado.....	175
Figura 60 -	Percentual da abstenção de resposta para a questão 177 do QSL, segundo o perfil do informante	191
Figura 61 -	Produtividade da variante geleia por Estado.....	193
Figura 62 -	Produtividade da variante chimia por Estado.....	194
Figura 63 -	Carta linguística 357 do ALERS.....	196
Figura 64 -	Produtividade das “formas genéricas” relativas à geleira por Estado.....	197
Figura 65 -	Produtividade da variante bolinha de gude por Estado.....	206
Figura 66 -	Produtividade da variante bolita por Estado.....	207
Figura 67 -	Carta linguística bolita do ALMS (com adaptações)	208
Figura 68 -	Carta linguística 302 do ALERS (com adaptações)	210
Figura 69 -	Produtividade da variante búlica por Estado	212
Figura 70 -	Carta linguística 311 do ALPR II	213
Figura 71 -	Notas da carta linguística 311 do ALPR II.....	213
Figura 72 -	Produtividade da variante bolinha de vidro por Estado.....	215
Figura 73 -	Produtividade da variante bilosca por Estado	216

Figura 74 -	Carta 30 do EALMG (com adaptações)	217
Figura 75 -	Produtividade da varitante menino por Estado	229
Figura 76 -	Produtividade da varitante moleque por Estado	230
Figura 77 -	Produtividade da variante guri por Estado	232
Figura 78 -	Produtividade da variante garoto por Estado	233
Figura 79 -	Produtividade da variante piá por Estado	234
Figura 80 -	Carta linguística 270 do ALERS	236
Figura 81 -	Produtividade da variante mexerica por Estado	251
Figura 82 -	Produtividade da variante poncã por Estado	253
Figura 83 -	Produtividade da variante tangerina por Estado	255
Figura 84 -	Produtividade da variante bergamota por Estado	258
Figura 85 -	Carta linguística 062 do ALERS	259
Figura 86 -	Produtividade da variante morcote por Estado	261

LISTA DE TABELAS

Tabela 1 -	Resumo estatístico segundo o Estado para a questão 001 do QSL (verificação da normalidade dos dados)	181
Tabela 2 -	Teste de homogeneidade da distribuição de produtividade da variante córrego com base nos Estados	182
Tabela 3 -	Teste de homogeneidade da variante riacho com base nos Estados	184
Tabela 4 -	Teste de homogeneidade da variante riozinho com base nos Estados de SP, PR, SC e RS.....	185
Tabela 5 -	Teste de homogeneidade da variante ribeirão com base nos Estados de MG, SP, PR e SC.....	186

LISTA DE QUADROS

Quadro 1 – Procedência de alguns grupos alemães vindos ao Brasil no século XIX.....	58
Quadro 2 – População residente, por Unidades da Federação, segundo o lugar de nascimento - Brasil - 2010.....	71
Quadro 3 – Síntese dos testes de hipóteses para a questão 001 do QSL.....	186
Quadro 4 – Síntese das hipóteses para a questão 177 do QSL.....	222
Quadro 5 – Síntese das hipóteses para a questão 156 do QSL.....	224
Quadro 6 - Formas pouco produtivas para a questão 132 do QSL.....	240
Quadro 7 - Síntese das hipóteses para a questão 132 do QSL.....	245
Quadro 8 – Síntese das hipóteses para a questão 039 do QSL.....	268

LISTA DE ABREVIATURAS E SIGLAS

ALERS	Atlas Linguístico Etnográfico da Região Sul do Brasil
ALiB	Atlas Linguístico do Brasil
ALMS	Atlas Linguístico do Mato Grosso do Sul
ALPB	Atlas Linguístico da Paraíba
ALPI	Atlas Linguístico da Península Ibérica
ALPR	Atlas Linguístico do Paraná
ALS	Atlas Linguístico de Sergipe
APFB	Atlas Prévio dos Falares Baianos
EALMG	Esboço de um Atlas linguístico de Minas Gerais
GO	Goiás
IBGE	Instituto Brasileiro de Geografia e Estatística
KS	Kolmogorov-Smirnov
MG	Minas Gerais
MS	Mato Grosso do Sul
MT	Mato Grosso
PB	Português brasileiro
PR	Paraná
QSL	Questionário Semântico-Lexical
RJ	Rio de Janeiro
RS	Rio Grande do Sul
SC	Santa Catarina
SGVCLin	Software para Geração e Visualização de Cartas Linguísticas
SP	São Paulo
SPSS	Statistical Package for the Social Sciences
TN	Tradução nossa

SUMÁRIO

(Volume 1)

INTRODUÇÃO	21
CAPÍTULO 1 - FUNDAMENTOS HISTÓRICOS	27
1.1 OS PRIMEIROS CONTATOS	27
1.2 OCUPAÇÃO E POVOAMENTO DO TERRITÓRIO	28
1.2.1 Pontos de Ocupação e Povoamento: As Sociedades Coloniais	29
1.3 BRASIL MERIDIONAL: O MOVIMENTO DOS BANDEIRANTES	30
1.3.1 A Primeira Fase das Bandeiras: O Despovoamento	33
1.3.2 A Segunda Fase das Bandeiras: O Repovoamento	34
1.3.2.1 A mineração	36
1.4 TROPEIRISMO	39
1.4.1 Os Caminhos das Tropas e as Feiras de Sorocaba	41
1.4.2 O Legado do Tropeirismo	44
1.5 A ECONOMIA CAFEIEIRA	45
1.5.1 Os Meios de Transporte	48
1.6 A IMIGRAÇÃO	50
1.6.1 Imigração nos Estados da Região Sul do Brasil	54
1.7 BRASIL CONTEMPORÂNEO: MIGRAÇÃO INTERNA	62
CAPÍTULO 2 - FUNDAMENTOS LINGÜÍSTICOS	73
2.1 A LÍNGUA PORTUGUESA NO BRASIL	74
2.2 A LÍNGUA PORTUGUESA DO BRASIL	80
2.2.1 Ancianidade do PB	80
2.2.2 Das Origens Geográficas do Português Brasileiro	82
2.2.3 Das Origens do Português Popular Brasileiro	83
2.2.4 A Questão da Língua Brasileira	86
2.2.5 Português Brasileiro vs. Português Europeu	89
2.2.6 Português Brasileiro: Unidade e Diversidade	90
2.2.7 Português Brasileiro: Língua, Dialeto ou Falar?	91
2.2.8 Sobre os Estudos do Português Brasileiro	98

2.2.9	Sobre a Periodização da Geolinguística Brasileira	100
2.2.1	0 Novos Rumos da Geolinguística Brasileira: O Projeto ALiB - (1996)	101
2.3	SOBRE A DIALETAÇÃO DO PORTUGUÊS BRASILEIRO: AS PROPOSTAS DE DIVISÃO DIALETAL	102
2.3.1	Para Uma Divisão Dialetal do Português Brasileiro a partir dos dados do ALiB	109
2.3.1.1	Contribuições do projeto ALiB: perspectiva lexical	109
2.3.1.1.1	<i>Romano e Aguilera (2009)</i>	109
2.3.1.1.2	<i>Aguilera (2010)</i>	111
2.3.1.1.3	<i>Romano e Seabra (2014a)</i>	113
2.3.1.1.4	<i>Yida (2011)</i>	120
2.3.1.1.5	<i>Benke (2012)</i>	123
2.3.1.1.6	<i>Costa (2012)</i>	123
2.3.1.1.7	<i>Freitas-Marins (2012)</i>	126
2.3.1.1.8	<i>Ribeiro (2012)</i>	132
2.3.1.1.9	<i>Portilho (2013)</i>	136
2.4	ALGUMAS CONSIDERAÇÕES SOBRE AS ÁREAS DIALETAIS SOB A PERSPECTIVA LEXICAL	137
CAPITULO 3 – MATERIAIS E MÉTODOS		140
3.1	CONSTITUIÇÃO DO CORPUS	140
3.1.1	A Área Geográfica Estudada	141
3.1.2	Seleção das Questões	142
3.2	O SISTEMA [ΣGVCLIN] – SOFTWARE PARA GERAÇÃO E VISUALIZAÇÃO DE CARTAS LINGUÍSTICAS	143
3.3	ABORDAGEM ESTATÍSTICA	153
3.3.1	Tratamento dos Dados.....	155
CAPÍTULO 4 - DESCRIÇÃO E ANÁLISE DOS DADOS		159
4.1	QUESTÃO 001 (CÓRREGO).....	159
4.1.1	Tratamento Estatístico.....	180
4.2	QUESTÃO 177 (GELEIA)	190
4.2.1	Tratamento Estatístico.....	199

4.3	QUESTÃO 156 (BOLINHA DE GUDE)	204
4.3.1	Tratamento Estatístico	220
4.4	QUESTÃO 132 (MENINO)	228
4.4.1	Tratamento Estatístico	241
4.5	QUESTÃO 039 (TANGERINA)	248
4.5.1	Tratamento Estatístico	264
CONSIDERAÇÕES FINAIS		271
REFERÊNCIAS.....		276
ANEXOS		293
ANEXO A –	Decreto de 1952	294
ANEXO B –	Autorização do Comitê Nacional para uso do corpus do Projeto do Alib	296

SUMÁRIO (Volume 2)

APRESENTAÇÃO.....	5
1. Relação das cartas.....	6
2. Divisão dialetal de Antenor Nascentes (1953).....	9
3. Base Cartográfica.....	10
4. Rede de Pontos Linguísticos.....	11
5. Cartas Linguísticas.....	14
6. Relatórios	86

INTRODUÇÃO

Uma língua não se espalha através de uma região sem alterar-se aqui e ali. (NASCENTES, 1960, p. 253)

O português brasileiro apresenta o aspecto marcante da diversidade linguística que se manifesta não somente no eixo vertical (social), mas também no horizontal (espacial). Desse modo, dada a extensão territorial de um país de dimensões continentais e os diferentes “brasis” que surgiram em decorrência do processo de ocupação e povoamento do território (RIBEIRO, 2006), o português europeu implantado em um novo ambiente e falado pela imensa massa de iletrados dos primeiros séculos da colonização foi se distanciando cada vez mais da língua trazida de além-mar, tomando pouco a pouco identidade própria.

Este país interétnico e multicolorido pelo sangue de diferentes raças adotou como língua de Estado e nação um veículo de comunicação que, ao mesmo tempo em que se apresenta diversificado, mostra-se unitário, conferindo coesão e mútua inteligibilidade a um povo estratificado em classes. De fato, Silva Neto (1979) já afirmara que o português brasileiro apresenta unidade em meio à diversidade, configurando-se, como se sabe hoje, em uma língua plural e polarizada (LUCCHESI, 2006).

A pluralidade de formas linguísticas manifesta-se não somente nos dois polos da sociedade (classe alta e baixa) que se caracterizam por diferenças sociais, econômicas e culturais antagônicas, mas também entre as diferentes regiões geográficas, separadas pelas longas distâncias e aspectos sócio-históricos característicos que conferem identidade a cada região. O português do Brasil (doravante PB) é uma língua variada que evidencia a diversidade cultural de seu povo pelas suas diferentes pronúncias, pelos diferentes vocábulos e estruturas morfossintáticas.

Ao considerar que as diferenças linguísticas no Brasil são menos geográficas que socioculturais, Teyssier (1994) afirma que “a dialetologia brasileira será menos horizontal que vertical” (TEYSSIER, 1994, p. 79). Todavia, este prognóstico do eminente estudioso não se verificou, uma vez que os estudos dialetais e geolinguísticos do PB, como apresenta Romano (2013), têm alcançado

inúmeros avanços, revelando o quanto a diversidade diatópica é um traço característico dessa língua¹.

De fato, será que um paraibano, sem escolaridade ou mesmo escolarizado, saberia dizer o que é um *arroio* ou uma *sanga*? E um *piá* paranaense conseguiria identificar o *marraio* e a *bila*? Obviamente, preferiria brincar de *burquinha* ou soltar *pandorga*, mas jamais chuparia *caramelo* como o fazem os *guris* de Cuiabá, e muito menos chuparia *bombom* como os *meninos* do Pará. Para um paulista, talvez, seja difícil associar *musse* à *geleia*, enquanto que, para o catarinense seria uma tarefa mais fácil, principalmente no litoral leste. Já o gaúcho insiste em dizer que *guris* adoram a *chimia*, vivem comendo *bergamotas* e correndo atrás das *cigarras*, enquanto paulistas observam os *moleques* nos *faróis* da grande São Paulo e os do interior vivem *caçando as libélulas* ou seriam as *pitás/os*? A *mimosa* de Curitiba é *tanja* em São Luiz/MA, a *tangerina* de Manaus, a *laranja-cravo* de João Pessoa. Estes breves exemplos, longe de serem fatos curiosos sobre a variação linguística do PB, são uma realidade comprovada cientificamente por trabalhos empíricos que, de algum modo, discordam de Teyssier (1994, p.79) ao afirmar que as diferenças linguísticas no Brasil são menos geográficas que socioculturais².

A polarização sociolinguística do PB é uma realidade que os trabalhos sociolinguísticos têm revelado pelos inúmeros projetos de que se tem conhecimento, mas o destino da dialetologia brasileira tem sido outro. Os estudos dialetais brasileiros, em especial os de natureza geolinguística, embora tenham se adentrado aos “veios sociolinguísticos” (CARDOSO, 2010), apontam para a diversidade horizontal do PB, uma língua com diferenças diatópicas evidentes que se manifestam nos diferentes níveis, mas é o nível lexical que chamou a atenção para este estudo.

¹ Cabe, no entanto, salientar que sob o ponto de vista fonético ou mesmo morfossintático, as diferenças diastráticas são mais evidentes, podendo-se, portanto, considerar pertinentes as afirmações de Teyssier (1994). Tomem-se como exemplos casos de metátese, rotacismo, concordância dos constituintes nos sintagmas nominais. Todavia, evitando-se generalizações, neste trabalho, exemplificam-se aspectos da variação diatópica a partir do léxico, uma vez que este nível evidencia claramente este tipo de variação.

² Há também aspectos fonéticos e morfossintáticos do PB que revelam marcas diatópicas como, por exemplo, o comportamento dos róticos e das sibilantes em coda silábica, uso dos pronomes de segunda pessoa, presença ou ausência de artigos antes de substantivos próprios, a abertura de vogais mediais em posição pretônica, entre outros.

O léxico, como se sabe, é o conjunto de palavras criadas e assimiladas pelo homem no decorrer da história e resulta do processo de nomeação e de interação daquele com o meio em que vive (BIDERMAN, 1992). É gerado ao longo do tempo e, por meio de seu estudo, é possível detectar influências socioculturais e históricas adquiridas em seu processo de formação. O léxico representa ainda “um aspecto diferenciador no que se refere à variação linguística, pois, além de evidenciar diferenças de uma região para outra, demonstra também a consequente mobilidade dessas diferenças de um espaço para outro” (ISQUERDO, 2007, p. 533). Dessa forma, verificar as especificidades lexicais de cada área geográfica é uma tarefa desafiadora, dada a dinamicidade da língua e a disseminação de variantes.

Esta tese se concentra na variação lexical do PB, buscando respaldo teórico na Dialetoлогия que, por meio da Geolinguística³, dedica-se ao estudo da variação espacial das línguas, embora não exclusivamente. A pesquisa enquadra-se também na linha da Linguística Histórica sob a perspectiva *lato sensu* (MATTOS e SILVA, 2008), buscando correlacionar os fatos linguísticos registrados a aspectos sócio-históricos do português, e contribuindo, assim, para a história da língua portuguesa, uma vez que o léxico de uma língua evidencia traços da cultura e da história de um povo (BIDERMAN, 1992).

A descrição do léxico do PB tem sido amplamente empreendida pelos trabalhos acadêmico-científicos 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 têm contribuído para a descrição do léxico das diferentes regiões do país. Dentre os trabalhos mais recentes, citam-se como exemplos os de Isquerdo (2009), Romano e Aguilera (2009), Silva e Aguilera

³ Neste trabalho, defende-se que a Geolinguística é uma área de interesse atrelada diretamente à Dialetoлогия, que, atualmente, deve ser entendida como uma subdisciplina ou um desdobramento dos estudos dialetais. Reconhece-se, todavia, que na tradição dos estudos dialetológicos, estudiosos como Veny (1986), Coseriu (1987), Montes Giraldo (1987), García Mouton (1996), entre outros afirmam ser a Geolinguística um método do qual se vale a Dialetoлогия para documentar formas e expressões linguísticas no espaço geográfico. Entretanto, considerando-se todo o aparato teórico e metodológico já desenvolvido na bibliografia da área - Margotti (2008), Cardoso (2002), Thun (1998), Radtke e Thun (1996) - bem como o constante interesse dos estudiosos para elaboração de atlas linguísticos, o grande volume de obras produzidas de natureza variada, os distintos espaços geográficos recobertos e descritos pelos trabalhos geolinguísticos e a grande contribuição dos atlas linguísticos para a Dialetoлогия e a Linguística como um todo, compreende-se que a Geolinguística possui seus métodos próprios, como já apontaram Chambers e Trudgill (1994), sendo um deles o método cartográfico. Em Romano (2014, no prelo), encontram-se mais detalhes sobre este posicionamento.

(2010), Paim (2011), Yida (2011), Costa e Isquerdo (2012), Freitas-Marins (2012), Ribeiro (2012), Razky (2013), Portilho (2013), Romano e Aguilera (2014 no prelo), Romano e Seabra (2014a), dentre outros. Esses trabalhos apontam para as diferenças regionais do léxico do PB, ora confirmando a proposta de divisão dialetal de Antenor Nascentes (1953), sob a perspectiva lexical, ora evidenciando particularidades que caracterizam cada uma das regiões, sobretudo, por influências sócio-históricas do processo de ocupação e povoamento.

Dessa forma, com base nos dados do Projeto Atlas Linguístico do Brasil (doravante ALiB), esta pesquisa tem como objetivo geral discutir a divisão dialetal de Antenor Nascentes (1953) no que se refere à área geográfica designada pelo dialetólogo como *subfalar sulista* e contribuir para uma possível redefinição desta divisão. Como objetivos específicos, busca-se (i) revelar as diferenças e semelhanças na norma lexical da área geográfica investigada; (ii) delimitar áreas lexicais que podem refletir traços da história social do português; e (iii) desenvolver uma ferramenta computacional que possibilite a cartografia dos dados geolinguísticos de forma automatizada.

A hipótese central do trabalho pauta-se no pressuposto de que a região correspondente ao *subfalar sulista*, na atualidade, não apresenta homogeneidade no que se refere ao léxico, ou seja, 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 geográfica não contemplada por atlas 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 pode revelar semelhanças e diferenças entre as regiões, bem como traços linguísticos que definem e caracterizam cada um dos Estados, evidenciado áreas lexicais que extrapolam os limites administrativos das unidades da federação em questão.

Vale notar que a maioria dos trabalhos que se propõe a estabelecer áreas dialetais para o português brasileiro pauta-se no nível fonético da língua.

Traçar áreas dialetais com base no léxico não é uma tarefa fácil, considerando-se a dinamicidade da língua e a conseqüente disseminação de variantes de um espaço geográfico para outro, decorrentes de processos migratórios. Dessa forma, os questionamentos que motivam o desenvolvimento desta pesquisa são os seguintes:

- Por meio do léxico é possível delinear ou estabelecer áreas dialetais?
- A divisão de Antenor Nascentes (1953), no que se refere ao subfalar sulista, ainda se sustentaria no século XXI sob o ponto de vista lexical?
- Atualmente, qual seria a definição e quais os limites e abrangência deste subfalar ou falar sulista?
- A área geográfica correspondente ao subfalar sulista, sob ponto de vista do léxico, apresenta, internamente, mais semelhanças ou diferenças?
- Há subáreas lexicais na área geográfica investigada? Como elas estão configuradas?
- A distribuição das variantes lexicais reflete os processos de ocupação e povoamento, bem como os fluxos de migração observados na área em foco?

Para responder às questões de pesquisa supramencionadas e alcançar os objetivos estabelecidos, utiliza-se como *corpus* de análise os dados coletados pela equipe do Projeto ALiB para a elaboração do Atlas Linguístico do Brasil, referente à língua falada de 472 informantes naturais de 118 cidades brasileiras, contribuindo, desse modo para o desenvolvimento do tão almejado atlas nacional.

Esta tese está organizada em dois volumes. O volume 1 apresenta aspectos introdutórios e teórico-metodológicos da pesquisa, além da análise dos dados, e contém quatro capítulos. No Capítulo 1 – Fundamentos históricos – apresenta-se um resgate da história social da região investigada com a finalidade de subsidiar a análise do *corpus*, para uma melhor compreensão da variação lexical nesta região do país. O Capítulo 2 – Fundamentos linguísticos – apresenta os fundamentos da história linguística do PB, com foco na diversidade e nos trabalhos sobre o léxico que a evidenciam. O Capítulo 3 – Materiais e métodos – traz uma descrição dos materiais e métodos utilizados para o desenvolvimento do trabalho. O Capítulo 4 – Análise dos dados – discute os dados analisados, tratando-os

quantitativa e qualitativamente. Seguem-se a esse último capítulo, as considerações finais do trabalho, bem como as referências e anexos.

O volume 2 – Cartas e Relatórios – traz um conjunto de 71 cartas linguísticas numeradas de 1 a 6, subdivididas em letras do alfabeto, por exemplo, 1a, 1b, 1c e, assim, progressivamente. Cada número corresponde a uma questão do Questionário Semântico-Lexical (QSL). Desse modo, as cartas de número 1 referem-se à questão 001 do QSL (*córrego*), as cartas de número 2 dizem respeito às variantes lexicais para a questão 177 (*geleia*), cartas de número 3 documentam variantes lexicais para a questão 156 do QSL (*bolinha de gude*), cartas de número 4 apresentam dados referentes à questão 132 do QSL (*menino*) e por fim, cartas de número 5 dizem respeito à questão 039 (*tangerina*). A carta 6 é um mapa elaborado a partir das análises empreendidas com a finalidade de localizar os dois grandes falares: o *paulista* e o *sulista*.

Ainda nesse volume constam cartas introdutórias e 10 relatórios de produtividade utilizados para a descrição dos dados em termos absolutos e relativos.

CAPÍTULO 1- FUNDAMENTOS HISTÓRICOS

O estudo científico de uma língua é fundamentalmente o estudo da cultura de que ela é forma e produto. Estudar o português do Brasil é, pois, em grande parte, estudar a história de nossa formação. (CUNHA, 1981, p. 13)

As palavras do insigne prof. Celso Cunha introduzem de forma cabal a importância deste capítulo. O resgate da sócio-história das regiões contempladas neste estudo revela a intrincada e intrigante formação do país, caracterizada, principalmente, pela miscigenação de três tipos humanos: o elemento indígena, o africano e o europeu. De acordo com Ribeiro (2006), formou-se aqui um país autêntico, altamente colorido pelo sangue de diferentes raças. A miscigenação bem como os movimentos de ocupação e povoamento do território deram a identificação ao Brasil.

Neste capítulo, apresenta-se um breve quadro histórico do Brasil, com destaque à área geográfica investigada, correspondente a oito Estados federativos: Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Minas Gerais, Mato Grosso do Sul, Mato Grosso e Goiás.

Não se empreendeu um estudo exaustivo e detalhado dos aspectos historiográficos da formação do país, todavia, buscou-se certo fio cronológico no desenvolvimento dos fatos. Recuperam-se aspectos dos primeiros contatos dos colonizadores com os autóctones, as sociedades coloniais e os pontos de povoamento instalados para a conquista da 'nova terra'. Tecem-se algumas reflexões sobre o movimento dos Bandeirantes, do Tropeirismo e de iniciativas que contribuíram para a imigração e de migração no Brasil contemporâneo. Todo este aparato histórico-social tem por finalidade subsidiar as análises empreendidas no capítulo 4, podendo-se, assim, explicar os aspectos lexicais identificados no *corpus* sob análise.

1.1 OS PRIMEIROS CONTATOS

A capacidade de adaptação do colonizador português que aqui aportou nos primeiros anos da colonização possibilitou que se desenvolvesse uma

típica e vigorosa cultura entre os povos aborígenes e os europeus, propiciando, pelo menos em um primeiro momento, a fusão de culturas e a exploração do ambiente. Desse modo, para a ocupação do território, o português valeu-se, dentre outras estratégias, dos traços culturais dos índios tupi. Segundo Cortesão (1966),

os tupi cativavam, como é sabido, os inimigos para imolá-los em sacrifícios rituais e antropofágicos: os colonos portugueses utilizaram as tendências belicosas do tupi, mas eliminando essa prática e instigando-os a trocar os vencidos por escravos. (CORTESÃO, 1966, p. 79)

Se não fosse essa compreensão de acolhimento da cultura indígena, provavelmente, não teriam conseguido ocupar o território, haja vista que os indígenas encontrados no litoral “somavam, talvez, 1 milhão de índios, divididos em dezenas de grupos tribais, cada um deles compreendendo um conglomerado de várias aldeias de trezentos a 2 mil habitantes” (FERNANDES apud RIBEIRO, 2006, p. 28).

Ainda nos primeiros anos de contato, os indígenas que ainda o puderam fazer fugiram mata adentro, levando, porém, em seus corpos vírus mortais contraídos dos brancos. Foi assim que a civilização europeia se impôs, primeiramente, como uma epidemia de doenças, depois com guerras de extermínio e escravidão. Dessa forma, com o colonizador vieram as três pragas do homem branco: a guerra, a peste e a escravidão, além do seu desejo obsessivo de

multiplicar-se no ventre das índias e pôr suas pernas e braços a seu serviço para plantar e colher suas roças, para caçar e pescar o que comiam. Os homens serviam principalmente para tombar e juntar paus-de-tinta ou para produzir outra mercadoria para o seu lucro e bem-estar. (RIBEIRO, 2006, p. 43)

Muitos indígenas, no entanto, foram aliciados pelos europeus e não fugiram para o interior do país. Preferiram o convívio com os novos senhores devido à irresistível atração das ferramentas, dos adornos. Associavam-se aos colonizadores para conseguir tais bens, tornando-se flecheiros em guerras contra os índios arredios, fugidos, designados pelos de além-mar como *tapuias* ou inimigos.

1.2 OCUPAÇÃO E POVOAMENTO DO TERRITÓRIO

Além da exploração do território, a liberdade de costumes propiciou também a miscigenação, inicialmente, entre os portugueses e as índias, e, posteriormente, com a inclusão do elemento negro. O processo de miscigenação foi imprescindível para a ocupação efetiva do novo território, pois o velho uso indígena de incorporar elementos estranhos à comunidade (denominado *cunhadismo*) passou a ser uma prática fundamental na nova sociedade que surgia.

Nas aldeias, os chefes das tribos davam a cada colono português uma moça indígena como esposa e “assim que ele a assumisse estabelecia automaticamente laços que o aparentavam com todos os membros do grupo” (RIBEIRO, 2006, p. 72), e como cada europeu podia fazer muitos desses casamentos, a instituição do *cunhadismo* funcionava como uma forma eficaz de recrutamento de mão de obra. Figuras máximas do *cunhadismo* no Brasil colonial foram João Ramalho e Antonio Rodrigues que, segundo relatos, eram capazes “de levantar 5 mil índios de guerra, enquanto todo o governo português não conseguiria 2 mil” (RIBEIRO, 2006, p. 74).

Outro aspecto importante para a conquista da nova terra foi o apoio da Igreja Católica, que, no cenário das grandes descobertas, dava total respaldo a Portugal e Espanha na disseminação da fé cristã. Para a América, foram recrutados os jesuítas da Companhia de Jesus, que, nas terras de domínio espanhol, já catequizavam os aborígenes, tornando-os cristãos e reduzindo-os nas missões jesuíticas para protegê-los da escravidão. Há de se notar, no entanto, que os próprios missionários foram os principais dizimadores da população indígena, pois o trabalho de retirar os índios de suas aldeias dispersas para concentrá-los nas reduções deixava-os vulneráveis aos vírus com que os próprios missionários, sem querer, os contaminavam.

1.2.1 Pontos de Ocupação e Povoamento: As Sociedades Coloniais

Sabe-se que a ocupação e o povoamento do Brasil deram-se em face das condições apresentadas pelo meio – o colonizador ocupou, de início, o litoral, sem se preocupar, de imediato, com o interior. O processo de ocupação do interior só foi possível após alguns anos, quando o elemento mestiço encaminhou os povoadores a desbravar as matas antes temidas por eles.

De acordo com Ribeiro (2006), em 1570, a dominação portuguesa estava assentada solidamente no território brasileiro e equivalia a aproximadamente 40 mil colonizadores, a sua maioria constituída por mamelucos⁴.

A sociedade colonial estabelecia-se em quatro principais implantações – Pernambuco, Bahia, Espírito Santo e São Paulo – desenvolvendo-se três tipos de sociedades - duas delas, localizadas na região Nordeste (a agrária e a pastoril) e a terceira, no Centro-Sul do país (a sociedade bandeirante).

Ricardo (1970) afirma que naquela época encontram-se: (i) a sociedade agrária do litoral nordestino baseada na monocultura, no latifúndio e no escravocratismo. Constituía-se de uma minoria de brancos dominadores ‘patriarcais e polígamos’, de um grande número de escravos criados nas senzalas, bem como de agregados, moradores de casas de taipa e palha – vassalos das casas-grandes; (ii) a sociedade pastoril, localizada no interior do nordeste, que buscava a possibilidade de uma vida mais ‘democrática’ fugindo da área açucareira, dedicada ao criatório de gado e deslocada para os sertões e, por fim, (iii) a sociedade bandeirante, ao sul, localizada intrasserra, no Planalto de Piratininga, mestiçada com aborígenes, praticante da policultura em larga escala em pequenas propriedades, constituindo-se como uma sociedade esparsa.

1.3 BRASIL MERIDIONAL: O MOVIMENTO DOS BANDEIRANTES⁵

Enquanto os pontos de povoamento do Brasil setentrional (Pernambuco e Bahia) se desenvolviam com a economia da cana-de-açúcar e o emprego da mão de obra escrava negra, na capitania de São Paulo, encontrava-se uma economia de pobreza, pois “os engenhos de açúcar não prosperavam nem surgiram outras lavouras. Mesmo a produção de pau-brasil foi sempre medíocre comparada com a de outras províncias” (RIBEIRO, 2006, p. 85). Assim, “durante muitas dezenas de anos, os herdeiros dos donatários de São Vicente e Santo Amaro mostraram grande desinteresse por essa faixa sulina da colônia” (SIMONSEN, 1937, p. 313) e para Ribeiro (2006):

⁴ Mamelucos são os filhos originados da relação entre o homem branco e as índias (HOUAISS; VILLAR, 2001).

⁵ Sobre o movimento dos bandeirantes há uma extensa bibliografia. Neste trabalho, apresentam-se as referências das obras a que se teve acesso. Salienta-se ainda que a figura do bandeirante e as qualidades da sociedade paulista do século XVII foram exaltadas, principalmente, por historiadores de São Paulo como Alfredo Ellis Jr.(1926) e Afonso Taunay (1924/1950).

Essa pobreza, que está na base tanto das motivações quanto dos hábitos e do caráter do paulista antigo, é que fazia deles um bando de aventureiros sempre disponível para qualquer tarefa desesperada, mais predispostos ao saqueio que à produção. (RIBEIRO, 2006, p.330)

Assim, da fusão dos portugueses com os aborígenes nasceu o movimento das bandeiras⁶ que, em sua essência, era uma expedição de alguns poucos brancos e mamelucos e muitos índios que objetivou a expansão econômica, geográfica ou política da então colônia portuguesa.

Em dados quantitativos, Fausto (2003) ilustra que “a grande bandeira de Manuel Preto e Raposo Tavares que atacou a região do Guaíra em 1629, por exemplo, era composta de 69 brancos, 900 mamelucos e 2 mil índios” (FAUSTO, 2003, p. 94), revelando que o âmago da bandeira consistiu na colaboração íntima do elemento português com o indígena, pois o primeiro organizou e deu sentido para o objetivo nômade do aborígine e este colaborou com elementos indispensáveis na técnica do conhecimento do território até então desconhecido (RIBEIRO, 2006).

Para Simonsen (1937), o ambiente pobre onde viviam os colonos do Brasil meridional (Capitania de São Paulo) fazia-os lançar mão do braço escravo para lhes assegurar os meios de subsistência. Sem comércio e sem exportação não podiam pagar pelos escravos africanos, tinham que se contentar com o indígena. Com o crescimento do núcleo social primitivo cresciam, também, as necessidades e a ambição de acumular sobras. Assim, “com a crescente procura de braços começaram a progredir as investidas para o preamento de índios” (SIMONSEN, 1937, p. 315), motivadas também pela esperança de encontrar pedras e metais preciosos.

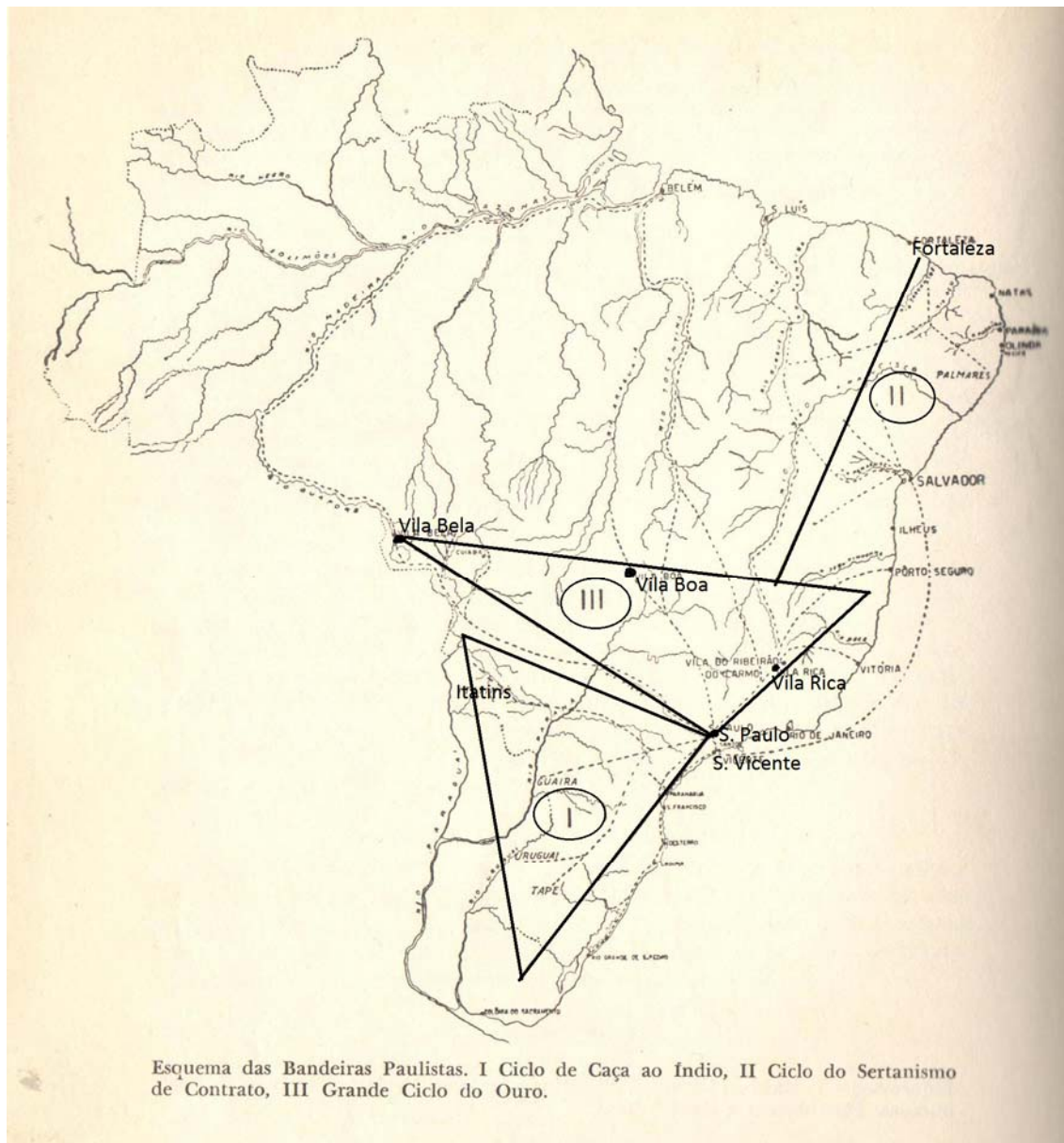
Nas palavras de Diégues Jr. (1960, p. 235), “Saem de São Paulo as bandeiras que vão devassar os sertões; penetração que se inicia no século XVI, se desenvolve no XVII e alcança a sua fase de esplendor no século XVIII.” Foi a partir da capitania de São Vicente que se irradiou a ocupação do interior da colônia. Em direção ao norte de São Paulo, o bandeirismo realizou o desbravamento do território

⁶ Nesta pesquisa, não é feita a distinção entre os nomes *bandeira*, *entrada* e *moções*, adotando-os como equivalentes, sinônimos. Em Vianna (1970), é possível encontrar as características que distinguem cada um desses movimentos de ocupação territorial

atual das Minas Gerais e, em direção oeste, foi alcançada, por um caminho, Cuiabá, e por outro, atravessando o território mineiro – Goiás (DIÉGUES Jr., 1960).

Vianna (1970) apresenta um mapa (Figura 1) localizando as principais rotas dos bandeirantes e sua influência na formação territorial do país, apontando as regiões em que ocorreram os três principais ciclos da bandeira: (i) caça ao índio; (ii) sertanismo de contrato e (iii) grande ciclo do ouro.

Figura 1 - Principais rotas de atuação dos bandeirantes e os três principais ciclos das bandeiras



Fonte: Vianna (1970, p. 198) – com adaptações

O autor completa,

Quanto à localização, ocorreram as bandeiras do *ciclo da caça ao índio* principalmente nas regiões do Sul e Oeste do Brasil, de São Vicente ao futuro Rio Grande do Sul e Mato Grosso, excepcionalmente atingindo, antes do grande ciclo do ouro, as do Centro do país (Minas Gerais e Goiás); o *ciclo de ouro de lavagem* foi essencialmente vicentino, compreendendo zona depois pertencente ao Estado do Paraná; o *ciclo de sertanismo de contrato* consistiu nas expedições que para repressão ao indígenas sublevados dirigiram-se às Capitânicas da Bahia e Ilhéus, posteriormente às do Rio Grande (do Norte) e Ceará, bem como na expedição de Domingos Jorge Velho contra os negros aquilombados nos Palmares, ao sul da Capitania de Pernambuco, principalmente em terras depois alagoanas; o *grande ciclo do ouro* realizou-se nas Minas Gerais, Goiás e Mato Grosso, também atingindo regiões centrais da Bahia e Ilhéus; ao *ciclo das bandeiras de povoamento* coube, afinal, passado o fastígio da mineração, povoar, notadamente com fazendas de criação de gado, os campos dos atuais Paraná, Santa Catarina. (VIANNA, 1970, p. 192-193)

1.3.1 A Primeira Fase das Bandeiras: O Despovoamento

É fato conhecido que os índios, nas suas migrações ou simples relações entre aldeias, traçaram todo o território de caminhos, alguns datando de épocas imemoriais e, muitos deles, chegando a ser aproveitados para estradas até os tempos atuais e, obviamente, pelos bandeirantes paulistas⁷.

O primeiro impulso das bandeiras paulistas consistiu no preamento dos indígenas, chamados por Simonsen (1937) como ciclos despovoadores do sertão, pois,

Além da serventia para suas próprias lavouras, verificaram os paulistas que os índios — mercadorias que se transportavam pelos seus próprios pés, numa época em que escasseavam os meios de transporte — constituíam apreciável elemento para o comércio exportador de suas capitânicas. Incentivaram-se, então, as bandeiras de prêa. (SIMONSEN, 1937, p. 319)

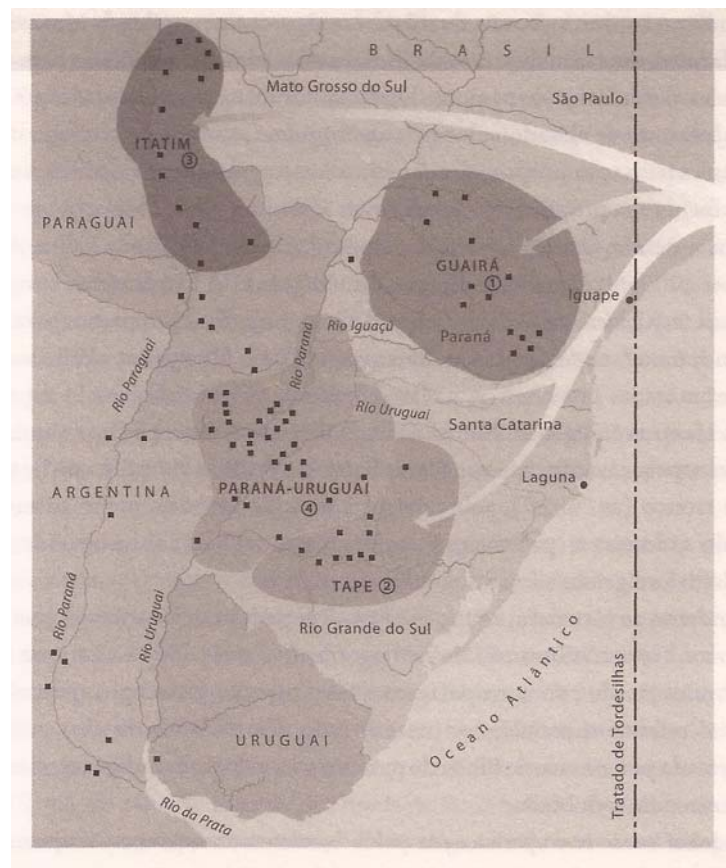
O grande incremento no tráfico de índios parece ter se verificado a partir de 1628, quando os bandeirantes foram assaltar as missões jesuíticas, de onde trouxeram grande número de escravos já um tanto afeitos ao trabalho. “Consta que grande parte foi vendida no mercado do Rio de Janeiro, onde alcançara, então, cerca de 20 mil réis por cabeça. Várias referências se encontram acerca desse comércio” (SIMONSEN, 1937, p. 331).

⁷ Para uma descrição detalhada dos antigos caminhos indígenas, confira Franco (1940, p. 15).

Simonsen (1937) relata ainda que os paulistas encabeçados pelas bandeiras de Manoel Preto e Raposo Tavares desceram pelas bacias dos afluentes do Paraná e investiram contra as missões jesuíticas de Vera e Guayrá, destruindo-as e expulsando os jesuítas dessa área do Brasil em 1628.

Em Bueno (2012), encontra-se um mapa (Figura 2) com a localização dessas reduções jesuíticas e as principais rotas das bandeiras que as destruíram.

Figura 2 - Mapa da localização das reduções jesuíticas e rota das bandeiras



Fonte: Bueno (2012, p. 65)

1.3.2 A Segunda Fase das Bandeiras: O Repovoamento

As bandeiras seguiram, a princípio, o curso dos rios como, por exemplo, o Tietê. Assim, os bandeirantes paulistas

se encontravam algum rio e prestava para a navegação, improvisavam canoas ligeiras, fáceis de varar nos saltos, aliviar nos baixios ou conduzir à sirga. Por terra aproveitavam as trilhas dos índios: em falta delas seguiam córregos e riachos, passando de uma para outra banda, conforme lhes convinha (...) evitavam naturalmente as matas, e de preferência caminhavam pelos espigões. (CASTRO, 1941, p. 112)

Em suas entradas para o sertão, além do preamento de indígenas, a grande esperança dos paulistas era encontrar o ouro, que veio aparecer nos pobres garimpos de Taubaté e, posteriormente, nas prodigiosas morrarias de Minas Gerais (RIBEIRO, 2006) no final do século XVII (1698), depois em Mato Grosso (1719) e, mais tarde, em Goiás (1725).

De acordo com Simonsen (1937, p. 326),

essa mesma gente que despovoara os sertões em busca do "remédio para sua pobreza", levando-lhes os habitantes para as suas fazendas ou para os mercados de escravos, como máquinas de trabalho ou artigo de comércio, promoveria o repovoamento deles, após os descobertos do ouro em Minas, Goiás e Mato Grosso.

Assim, as bandeiras, no século XVIII, tomaram o caráter repovoador, transformadas em expedições mineradoras e colonizadoras, introduzindo nos sertões, em larga escala, os escravos africanos, cessando os paulistas, quase por completo, o tráfico dos nativos da terra. Desse modo,

em poucos anos, aquelas regiões desertas [Minas Gerais] transformaram-se na área mais densamente povoada das Américas, concentraram-se cerca de 300 mil habitantes por volta de 1750. Os ricos vinham com toda a sua escravaria, pleiteando grandes lavras; os remediados, com o que tinham, e os pobres, com uns poucos negros, com apenas um, ou com nenhum, mas também tentando a sorte. (RIBEIRO, 2006, p. 336)

Traça-se, dessa forma, um novo panorama na história do Brasil colonial, o período aurífero, caracterizado pelo luxo de costumes citadinos, e pela ganância de riqueza fácil, resultando em um processo migratório sem precedentes, pois "das regiões do açúcar, que se debatiam em acentuada crise, começou também a debandada em direção às minas (SIMONSEN, 1937, p. 345). Segundo relatos de Antonil (1837),

Cada anno vem nas frotas quantidade de Portuguezes, e de estrangeiros, para passarem às minas. De cidades, voltas, recôncavos, e sertões do Brasil vão brancos, pardos, e pretos, e muitos Indios de que os Paulistas se servem. A mistura he de toda a condição de pessoas: homens, e mulheres; moços e velhos; pobres e ricos: nobres e plebeos, seculares clérigos, e religiosos de diversos institutos, muitos dos quaes não tem no Brasil convento nem casa. (ANTONIL, 1837, p. 149)

Apesar de os bandeirantes terem descoberto as minas, para Diégues Jr. (1960), não foi o bandeirismo em si um fator de povoamento ou de ocupação humana. Para o autor, o movimento foi um fator de devassamento do território, pois, à medida que um grupo bandeirante levantava o acampamento de um ponto à procura de outro, a área deixada era ocupada pelo movimento de retaguarda. “Os paulistas devassaram o território, realizaram a conquista e descobrimento das minas, mas a outros grupos, os baianos, negros, reinóis, é que se deve a fixação” (DIÉGUES, Jr., 1960, p. 239). Constituiu-se, portanto, nessas áreas, uma sociedade miscigenada pelos diferentes tipos humanos que as ocuparam construindo uma reciprocidade de influências no processo de formação humana⁸.

1.3.2.1 A mineração

O bandeirismo contribuiu também para a fundação de vilas e cidades. Diégues Jr. (1960) afirma que, inicialmente, essas vilas eram simples ranchos (arraiais), que se construíam para abrigar os que se dedicavam à mineração. Nas rancharias onde a exploração aurífera alcançava sucesso, começavam as construções mais estáveis, erguia-se a igreja, inicialmente, modesta, depois transformada em templo belo e rico. Erguiam-se também os palácios para os representantes do governo e os particulares construíam suas residências com gastos convenientes e correspondentes ao seu *status* social.

Dessa época, surgem cidades ostentadas pelo luxo e pela riqueza como Sabará, Ribeirão do Carmo (atual Mariana), Ouro Preto e São João Del-Rei, conferindo à região mineira, ainda cedo, “um sentido de urbanidade, antes que este

⁸ Era comum, nesses territórios, ocorrerem conflitos entre os paulistas e povos migrantes, sobretudo, os baianos, cada qual reclamando para si o direito de exploração exclusiva da área. O mais famoso desses conflitos foi a Guerra dos Emboabas na região de Ouro Preto – MG, entre os anos de 1707 e 1709.

se tivesse manifestado em outras regiões ou áreas do Brasil, salvo um século antes em Pernambuco, com os holandeses” (DIÉGUES Jr, 1960, p. 242).

A sociedade tornou-se mais complexa e se diversificou mudando os hábitos, incrementando elementos europeus à nascente cultura urbana, desenvolvendo, além de pobres garimpeiros e escravos forros à margem da sociedade, uma classe de intelectuais, compositores, escultores, artistas e outros artífices.

Conforme se verifica, a mineração resultou em um conjunto de transformações na sociedade brasileira que ecoou durante muito tempo, culminando não só na transladação de contingentes humanos como também na transferência da capital nacional para o Rio de Janeiro – até então um pobre arraial tal qual o antigo São Vicente.

Durante o período da mineração, eclodiram também revoltas contra a exploração abusiva da coroa portuguesa, como, por exemplo, a Inconfidência Mineira (1789). Dessa época, surgem também o contrabando e o garimpo que era a “exploração clandestina (...), mineração furtiva, feita às ocultas, chamando-se *garimpeiro* o que a praticava ou exercia” (DIÉGUES, Jr., 1960, p. 252).

Após algumas décadas de exploração intensiva e desordenada começam a se esgotar os ricos aluviões, primeiramente, de Minas Gerais, depois de Goiás e Mato Grosso. Nos fins do século XVIII, começam a aparecer as expressões de decadência das minas. As movimentadas localidades de antes começam a se transformar em cidades mortas, conforme observa Saint-Hilaire (1938) acerca da cidade de Vila Rica em Minas Gerais:

Contam-se em Vila Rica cerca de duas mil casas. Essa vila floresceu enquanto os terrenos que a rodeiam forneciam ouro em abundância; à medida, porém, que o metal se foi tornando raro ou de extração mais difícil, os habitantes foram pouco a pouco tentar fortuna em outros lugares, e, em algumas ruas, as casas estão quase abandonadas. A população de Vila Rica que chegou a ser de 20 mil almas, está atualmente reduzida a 8 mil, e essa vila estaria mais deserta ainda se não fosse a capital da província, a sede da administração, e a residência de um regimento (SAINT-HILAIRE, 1938, p.130-31).

A primeira fase da mineração no Brasil colonial consistiu na exploração desenfreada do ouro em Minas Gerais e, posteriormente, no Mato Grosso e Goiás. A segunda fase, no entanto, fundou-se na busca dos diamantes.

Essa fase inicia-se em 1729, quando uma portaria da Coroa manda suspender a mineração do ouro nas terras do Tijuco (atual Diamantina-MG), tornando a exploração das minas de diamante sob o monopólio da Coroa, o que vedou a mineração aos particulares (DIÉGUES Jr. 1960, p. 250).

O luxo e a ganância de lucros promoveram a exaustão do solo em Minas Gerais. Essa conjuntura, associada aos diversos conflitos de posse de terra frente à onda de migração de colonos de outras províncias e à imigração portuguesa, fizeram os bandeirantes buscar outras regiões nos sertões do Mato Grosso e Goiás.

Saíram os bandeirantes em direção ao Mato Grosso quase ao mesmo tempo em que outros grupos rumavam às minas de Goiás (DIÉGUES Jr., 1960). As comunicações com Mato Grosso eram feitas por Arararituaba (Porto Feliz), pelos rios Tietê, Paraná, Pardo, Coxim, Taquari, Paraguai, São Lourenço e Cuiabá, e demandavam mais de quatro meses em travessias penosas e perigosas, acossados que eram os paulistas pelos bravios paiaguás, exímios nadadores⁹.

Atribui-se a descoberta das minas goianas ao bandeirante Bartolomeu Bueno da Silva, o primeiro Anhanguera, datada dos anos de 1672. No entanto, a vida das minas goianas se dá a partir da bandeira de 1722, de seu filho Bartolomeu Bueno, quando se fundou o primeiro núcleo de povoamento, Ferreiro (1726), transformado anos mais tarde (1739) em Vila Boa de Goiás. Nessa região, o povoamento processou-se rápido não somente por paulistas, mas também por bandeirantes e aventureiros da Bahia, Pernambuco, Maranhão, Pará e Minas Gerais (DIÉGUES, Jr., 1960), desenvolvendo-se uma sociedade instável.

Para Simonsen (1937, p 349), o acesso a Goiás, a partir da capitania de São Paulo, se dava por Jundiaí, Atibaia e Rio Grande. Embora a estrada Anhanguera fosse o único caminho oficial, muitas estradas eram abertas repentinamente (DIÉGUES Jr. 1960).

Goiás e Mato Grosso apresentavam uma sociedade semelhante, caracterizada pela mestiçagem, pela dinamicidade e instabilidade social, pois, à medida que uma mina se esgotava, logo se buscava outra que lhe garantisse a produção. Assim, tanto no Mato Grosso quanto em Goiás não se desenvolveu o sentido de urbanidade ocorrido em Minas Gerais, haja vista que “o período de

⁹ Em Southey (1862, p. 330-337), encontra-se uma descrição detalhada sobre o trajeto e perigos encontrados pelos bandeirantes rumo às minas mato-grossenses e goianas.

apogeu das minas goianas e cuiabanas foi rápido; não demorou o suficiente para se transformarem os arraiais em núcleos urbanos” (DIÉGUES Jr., 1960, p. 283).

A febre do ouro fez os exploradores se esquecerem da agricultura. O custo de vida nessa região era alto e grande foi a mortalidade dos escravos por falta de alimentação adequada. O acesso às regiões era difícil, o que contribuía para o encarecimento dos produtos básicos, importados do Rio de Janeiro e São Paulo. Job (1984), por exemplo, afirma que “o alqueire de milho, de feijão, que em São Paulo valiam, respectivamente, seiscentos e quarenta e novecentos e sessenta réis, nas minas eram vendidos por noventa e seis mil réis” (JOB, 1984, p. 12).

Com a decadência das minas, a população se dispersou e se sedentarizou desenvolvendo outras atividades econômicas como a extração da erva-mate no Mato Grosso e, principalmente, a pecuária nas regiões mato-grossenses e goianas.

Para Ribeiro (2006), a mineração também estimulou a expansão do pastoreio nordestino pelos campos são-franciscanos e do Centro-Oeste, garantindo um novo mercado consumidor. A atividade mineradora possibilitou a ocupação da região sulina, conquistada pelos paulistas antigos, para a criação de gado vacum disperso pelo campo e a criação de muares que faziam todo o transporte terrestre no Brasil colonial.

1.4 TROPEIRISMO

Sabe-se que, já no começo do século XVII, bandeirantes paulistas assolavam o território que é hoje o Paraná, seguindo para a Ilha de Santa Catarina, atualmente Florianópolis e, nessa marcha “para o sul do país, chegaram à Laguna” (DIÉGUES Jr. 1960, p. 305). Porém, com o passar do tempo, muitos homens bandeirantes, que inicialmente se dedicavam ao apresamento dos indígenas e à exploração de metais preciosos, se fixaram em terras da região Sul e, paulatinamente, foram se integrando ao pequeno comércio praticado ali no lombo de mulas.

No início da mineração e até meados do século XVIII, a falta de animais de tração, carga ou sela era quase absoluta não só em Minas Gerais, como também em São Paulo e Rio de Janeiro. O transporte de mercadorias e viajantes, até então, era realizado nos ombros e nas costas dos indígenas, dos negros e dos

mamelucos assalariados (JOB, 1984). O transporte, dessa forma, constituía-se no principal problema dos habitantes das Minas Gerais.

Com o advento das minas de ouro, houve um aumento da demanda de carne (charque) e também de couro para os alforjes do transporte do ouro à corte portuguesa. A solução foi recorrer ao transporte desses materiais nas mulas, animais resistentes à carga e às longas distâncias, em caminhos muitas vezes impraticáveis, vindos da região Sul para a região mineradora. Dessa forma, surgiu uma nova economia baseada no transporte e comércio desses animais, o Tropeirismo.

O pioneirismo das tropas em terras americanas, para Almeida (1981), deve-se aos castelhanos do Rio da Prata e do Peru, pois foram eles

que inventaram não só a coisa, mas a palavra, porque as comunicações do litoral ou dos litorais com os Andes exigiram, antes do que no Brasil, grandes filas de cargueiros, especialmente muares, por causa de sua maior capacidade do que os cavalos nos terrenos montanhosos. (ALMEIDA, 1981, p. 37-38)

De acordo com Mattos (1984), antigos criatórios de mulas eram mantidos pelos espanhóis nas terras da margem direita do rio Uruguai – em Corrientes, Entre Rios e outras províncias que iam até o Paraguai – e forneciam esses animais para as minas de prata do Potosí¹⁰. Assim, a época da ascensão aurífera do Brasil coincidiu com o declínio das minas de prata espanholas e o destino daqueles muares, logicamente, seria o mercado das Minas Gerais, se não fosse a inimizade entre as duas Coroas.

As tropas de muares eram de dois tipos: (i) tropa xucra (palavra de origem quíchua¹¹) e (ii) tropa arreada. A tropa xucra consistia no transporte dos animais trazidos do sul de forma solta. Esses animais seguiam instintivamente outro animal guia colocado à sua frente, o que facilitava o trabalho dos peões. A tropa arreada ou cargueira era composta por animais já acostumados às tarefas pesadas e levavam suspensas em suas cangalhas as canastras, contendo as mercadorias a serem comercializadas. A abrangência das tropas de muares ateuve-se a uma área delimitada, compreendendo os atuais Estados de Mato Grosso, Mato Grosso do Sul,

¹⁰ As Minas de Potosí localizam-se no cerro de Potosí, atual Departamento de Potosí, na Bolívia.

¹¹ Língua indígena ainda hoje falada nos países da cordilheira dos Andes, do Chile à Colômbia, esp. no Equador e Peru [Foi a língua falada no antigo império inca] (HOUAISS; VILLAR, 2001).

Goiás, Minas Gerais, São Paulo, Rio de Janeiro e Espírito Santo, embora tenha seus reflexos também na Bahia e Pernambuco (MATTOS, 1984).

Os tropeiros partiam das campinas gaúchas tão logo terminasse o inverno, quando os pastos, ressequidos pelas geadas, rebrotavam. Viajavam lentamente, parando mais nos locais em que as pastagens fossem melhores, como no sopé da Serra de Santa Catarina, menos onde houvesse escassez de capim, fazendo chegar os animais ao destino (feira de Sorocaba) em boas condições para o comércio, alcançando bons preços. Normalmente em fins de março as tropas já se encontravam nos campos que iam desde Sorocaba até Itapetininga.

A vida do tropeiro era rudimentar e a sua alimentação era basicamente de origem animal. Frioli (1984, p. 27) relata que “a caça era muito praticada entre os tropeiros, e nela não se perdoavam as aves como papagaio e outros pequenos animais” e se matavam um tamanduá, faziam festa por uns dias. Enfrentavam adversidades de toda ordem, ataques de animais ferozes, de índios, perda e desgaste dos animais, enchentes, chuvas constantes.

1.4.1 Os Caminhos das Tropas e as Feiras de Sorocaba

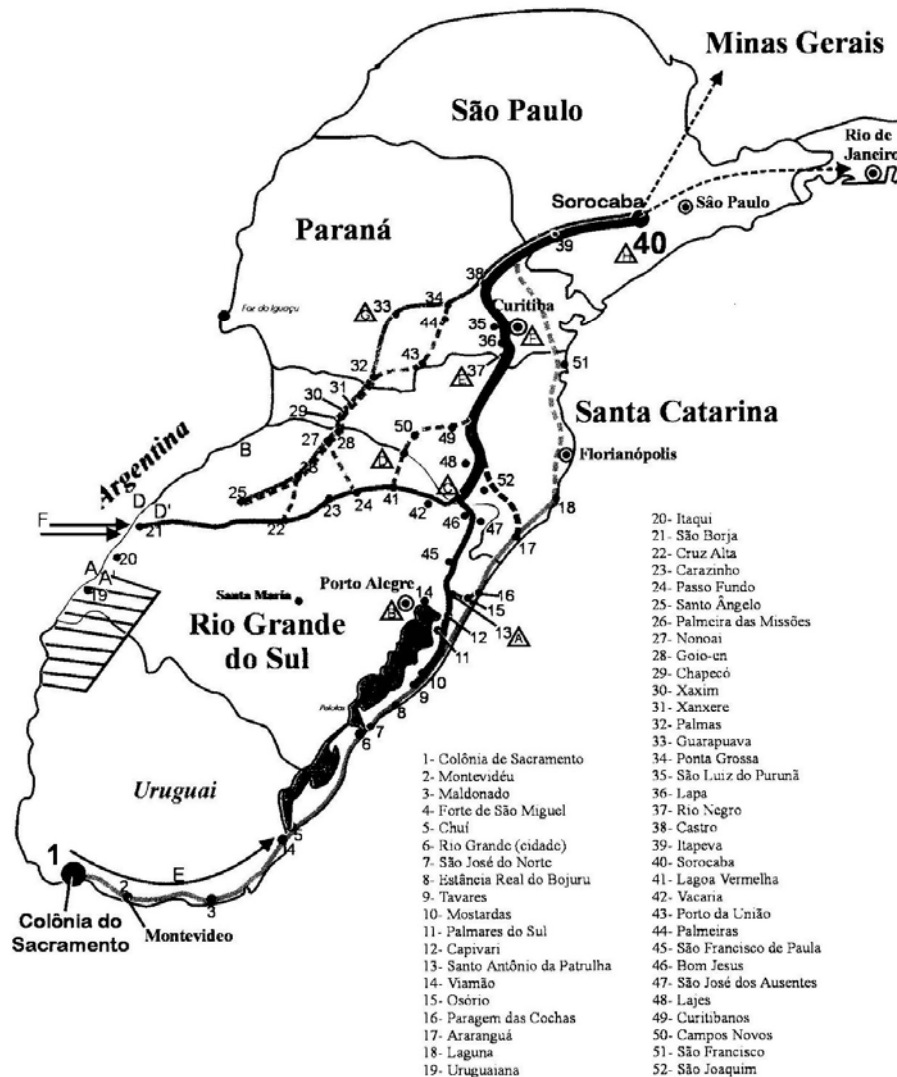
Para assegurar os interesses portugueses no nascente comércio de animais evitando que recaísse nas mãos dos castelhanos contrabandistas, as autoridades coloniais portuguesas incumbiram, em 1727, Francisco Brito Peixoto de abrir caminhos e fundar povoações nas terras rio-grandenses. Respalhado pela tropa de João Magalhães, Peixoto instalou novos pontos de povoamento nos campos de Viamão, na região da Lagoa dos Patos até São José do Norte, construindo a estrada da Serra que inaugurou “verdadeiramente o ciclo do tropeirismo no Brasil” (MATTOS, 1984 p. 19).

Pouco mais tarde, em 1739, o famoso tropeiro Cristóvão Pereira Abreu propôs a ampliação deste caminho ao então governador da Capitania de São Paulo, Rodrigo César de Menezes. Aproveitando a Estrada da Serra, desbravada por Peixoto, avançou em direção ao atual Estado de Santa Catarina passando pelos campos de Lages, Curitiba, Lapa, Ponta Grossa, Castro, Itararé, Itapeva, Itapetininga, e chegando ao destino final, Sorocaba. Estava aberto, portanto, o Caminho de Viamão, a rota mais importante e a mais utilizada pelos tropeiros que integraria o Rio Grande do Sul com o restante do país.

Paixão Côrtes (2000), ao pesquisar os trajes e as danças dos tropeiros, levantou dados sobre a história do tropeirismo, seus caminhos e seu cotidiano nas tropeadas. O autor, em uma das suas obras, representa as principais rotas pelas quais eram conduzidas as tropas de muares, conforme se observa na Figura 3.

Inicialmente, as tropas seguiam para Itu (então cabeça da Comarca) onde deveriam pagar os impostos. A capitania de São Paulo, até então com pouquíssimas rendas que propiciassem arrecadações aos cofres reais, viu, no tropeirismo, uma nova fonte de renda. Para tanto, foi necessário criar uma casa de Registro dos Animais, obrigando os tropeiros a passar por ali e pagar os encargos fiscais (VIEIRA, 1984). O Registro deveria ficar em local que impedisse os descaminhos, evitando o contrabando. Itu, portanto, não serviria devido à localização (no meio dos Campos de Pirapitingui) o que favorecia o extravio dos tropeiros, evitando-se, dessa forma, o pagamento ao governo dos devidos direitos.

Figura 3 - Rotas dos tropeiros



Fonte: Paixão Côrtes (2000, p. 52)

De acordo com Vieira (1984), os oficiais do Governo escolheram a Ponte do Rio Sorocaba para instalação do Registro por conta da topografia da região e pelas qualidades das pastagens que já serviria para o último descanso das tropas, antes de seguirem viagem para os caminhos montanhosos que conduziriam ao Planalto Paulista. Assim, em 1750, fundou-se o Registro de Animais em Sorocaba. Em consequência, começaram a ocorrer as feiras, em diferentes épocas do ano, e com o tempo, estabeleceu-se um período específico para a sua realização, acontecendo de fins de abril até o início de junho¹².

¹² Em Vieira (1984) e Almeida (1981), encontra-se a descrição detalhada sobre a venda e compra desses animais nas Feiras de Sorocaba, bem como aspectos da sociedade sorocabana no período em que ocorriam.

Todavia, pouco a pouco as Feiras de Muares de Sorocaba foram perdendo o prestígio inicial. Os burros e as mulas começaram a ser substituídos por diversos motores. Além disso, um surto de febre amarela, em meados de 1897, assolou a região, afastando os tropeiros, compradores, vendedores, agiotas, artistas e toda a população que vinha para Sorocaba em busca de bons negócios.

Outras feiras foram organizadas, como as Feiras de Itapetininga, sobrevivendo ainda o tropeirismo de forma tímida por mais 30 anos. Porém, a atividade foi minguando, mesmo tentando se adaptar aos novos tempos. Para Mattos (1984, p. 22), “os estudiosos do Tropeirismo têm atribuído o declínio e extinção do ciclo, ocorrido na década de 30, ao deslocamento ou perda de uso do muar no sistema de transportes de carga hoje feito, principalmente, por caminhões”.

1.4.2 O Legado do Tropeirismo

O tropeiro, além de agente comercial, tornou-se indispensável em outras atividades. Era emissário oficial, o correio, o transmissor de notícias, o intermediário de negócios, o portador de bilhetes, recados e encomendas. De acordo com Bonadio (1984, p. 45), os tropeiros “cumpriram um papel da mais alta importância na unificação cultural do país, como veículos difusores de notícias e idéias”.

Apesar da vida rude que levavam, quando se encontravam em Sorocaba, exibiam uma ostentação e esbanjamento demonstrados na rica indumentária de suas montarias e gastos excessivos em jogos, bebedeiras e festas. Eram vistos como homens corajosos que enfrentavam as peripécias de uma vida rigorosa, cheia de riscos nas constantes viagens às regiões inóspitas, “rasgando matas, varando serras, estabelecendo uma teia de caminhos, o tropeiro ligou e manteve vivos os núcleos urbanos isolados, representando efetivamente o consolidador das fronteiras nacionais” (JOB, 1984, p. 14). O ciclo do tropeirismo contribuiu decisivamente para moldar o Brasil tal como o é hoje e “umas das contribuições mais importantes do Ciclo, logo ao seu alvorecer, foi a incorporação do Rio Grande do Sul ao território nacional” (BONADIO, 1984, p. 46).

Além de esse movimento efetuar o escoamento da produção econômica, deve-se a ele a difusão de hábitos e costumes, haja vista que os tropeiros levavam consigo a sua cultura por onde passavam, sendo responsáveis

pelo florescimento de cidades que se iniciaram com os ranchos das tropas (STECA; FLORES, 2002).

Considerando as longas distâncias percorridas pelas tropas, os fazendeiros construíam ranchos para que os tropeiros que transportavam os seus produtos pudessem descansar e seguir viagem logo no início do outro dia. No entanto, muitos desses locais foram criados não só para o descanso, mas também para a engorda dos animais ou mesmo para esperar que as águas dos rios baixassem visando à travessia segura da tropa e, em muitos casos, para que as mulas parissem.

De acordo com Frasson e Gomes (2010), o tempo de espera não era pré-determinado, visto que dependia do que se apresentava, em decorrência do Estado nutricional da tropa, das condições físicas do tropeiro e das condições do tempo. Essas circunstâncias poderiam determinar invernações de quatro a seis meses, se não mais. Nesse período de espera surgiram as primitivas cidades separadas entre si por uma distância de cinco a seis léguas, ou seja, aproximadamente trinta e seis quilômetros, o que correspondia a um dia de caminhada da tropa.

Estabelecidos esses ranchos de pouso e descanso, os fazendeiros não tardavam em construir uma capela, e em seguida ocorria a implantação de pequenas vendas para o suprimento das necessidades essenciais dos tropeiros e viajantes que passassem por ali. Com o tempo, algumas famílias fixavam-se ao entorno e estava dado o ponto de partida para o surgimento de vilas no interior do país (ALGATÃO, 2010). Muitas dessas primitivas vilas tornaram-se prósperas cidades como Campinas e Jundiaí em São Paulo, e Pouso Alegre em Minas Gerais (ALMEIDA, 1981). Lapa, Palmeira e Ponta Grossa no Paraná (FRASSON; GOMES, 2010) e Lages, no planalto catarinense.

1.5 A ECONOMIA CAFEEIRA¹³

¹³ Um estudo detalhado sobre a história geral do café pode ser encontrado na obra clássica de Francis B. Thurber (1886): *Coffee from plantation to cup: a brief history of coffee production and consumption*. Nesse livro, o autor apresenta dados quantitativos e historiográficos do café no Brasil e no mundo. Disponível em: <<http://www.brasiliana.usp.br/bbd/handle/1918/01922400#page/8/mode/1up>>

A decadência do Tropeirismo esteve atrelada também ao surgimento de outra atividade econômica, a cafeeira, que impulsionou o Brasil no cenário internacional, exigindo, conseqüentemente, outras formas de transporte, como por exemplo, as ferrovias.

A então sociedade pobre e nômade, que se desenvolveu na Capitania de São Paulo, viu uma nova fase iniciar, a de esplendor e riqueza, caracterizada pelas grandes fazendas de café e pelo poderio exercido pelos barões.

A desilusão das minas de ouro e diamantes, cada vez mais decadentes, obrigou o paulista antigo a se fixar cada vez mais à terra. Assim, nos fins do século XVIII e, principalmente, começo do XIX, pouco a pouco a sociedade paulista começou a se sedentarizar no cultivo da terra com a cultura da cana e algumas de subsistência, dando um novo sentido agrícola à região. Porém, essa atividade era incipiente, ainda rudimentar, e somente começou a tomar vulto quando o café atingiu o território paulista, vindo do Rio de Janeiro e Minas Gerais, pelo vale do Paraíba.

Segundo Milliet (1982), as primeiras propriedades de café foram estabelecidas no lado fluminense do vale do rio Paraíba do Sul, em algumas cidades como Barra Mansa, Barra do Piraí, Valença, Vassouras, com irradiação para Iguazu e outras localidades. Mais tarde foi-se alastrando à área denominada São Gonçalo, Cantagalo e Angra dos Reis, sempre em busca de portos marítimos para o escoamento da produção (DIÉGUES Jr., 1960).

O plantio de café em Vassouras começou em 1816, tornando-se a cidade, na primeira metade do século, um importante foco da produção. O desenvolvimento nessa região atraiu, antes mesmo de 1820, os mineiros que em grande número se deslocavam de sua Província, abandonando as minas em decadência para se dedicarem à lavoura. Esse processo migratório avançou até o Espírito Santo. Assim, de Vassouras, o café prosseguiu e alcançou as ricas terras de Valença, propiciando o desenvolvimento regional culminar na construção de duas estradas de ferro para o escoamento da produção (DIÉGUES Jr., 1960).

A cultura do café expandiu-se à margem esquerda do rio Paraíba do Sul invadindo a zona da Mata, em Minas Gerais, e atingindo as antigas regiões já transitadas pelos primitivos mineradores, onde se localizavam núcleos de populações, vilas e aldeias, fundados ou mantidos pelos seus descendentes

(SIMONSEN, 1973). Rio de Janeiro transformara-se, assim, no escoadouro da produção, enriquecendo-se cada vez mais.

Operou-se, pois, na década 1820-1830, uma transformação profunda na província do Rio de Janeiro. Foram abertas grandes fazendas, que passaram a importar braços de Minas Gerais e mesmo da África. As condições da nova cultura, face aos recursos locais, criaram o tipo da fazenda de café fluminense com sua organização peculiar. (SIMONSEN, 1973, p. 180)

Mas foi no lado paulista do Vale do Paraíba que houve o incremento prodigioso da cultura cafeeira. Em pouco tempo, as lavouras tomaram os espaços agricultáveis do sul fluminense e adentraram o território paulista por meio das cidades limítrofes de Bananal e Ubatuba, além de Ilha Bela (MILLIET, 1982).

De acordo com Amaral (1958, p. 214), “foi a terra roxa que fez São Paulo eleger-se a pátria do café. É de tal modo extraordinária essa terra, que, enquanto todas as lavouras cafeeiras do mundo decaem ou estacionam, a do Brasil vive em superprodução”.

Pestana (1927) salienta que em 1817 a cultura cafeeira já adquirira certa importância e, sobre a sua propagação no Vale do Paraíba, tece as seguintes considerações:

Mais próximo de Rezende, o município de Areias, principal produtor, exportava seu café para o Rio de Janeiro, produzindo anualmente cada escravo 50 arrobas e vendendo-se a arroba á razão de 2\$000. Lorena, Garatinguetá, Pindamonhangaba e Taubaté produziam “muito café” em suas terras. Jacarehy e Mogy das Cruzes também plantavam a promissora rubiácea, mas em escala menor. No entanto, para o lado do fertilíssimo e vagem Oeste, em Itú, cultivava-se “pouco café”, merecendo preferencia a canna. (PESTANA, 1927, p. 9)

As cidades à beira do Rio Paraíba do Sul utilizaram seus espaços agricultáveis quase em sua totalidade, dedicando-se à cultura do café. Os nobres, com o dinheiro obtido, dotaram suas cidades com infraestrutura somente comparável com a do Rio de Janeiro, capital imperial; caso da cidade de Lorena (SOBRINHO, 1967), exibindo a riqueza no modo de vida dos barões. Cidades acanhadas ganharam prédios luxuosos construídos aos moldes das construções da corte, mostrando que a nobreza rural mantinha íntimo contato com a urbanidade, acompanhando-a em todas as suas tendências (ALGATÃO, 2010).

Porém, tanta riqueza e exuberância das cidades do Vale do Paraíba foram pouco a pouco esvaecendo, pois já no decênio de 1870, as terras desgastadas apresentavam seus primeiros indícios de decadência¹⁴. Ademais, a questão da mão de obra majoritariamente escrava empregada na lavoura contribuiu para o declínio da atividade açucareira¹⁵.

Neste panorama, Taunay (1939, p. 141), acerca da situação de Campinas no ano de 1856, salienta que “a necessidade mais instante de que se ressentia a lavoura do município prendia-se ao melhoramento das vias de comunicação”. Acrescenta que não menos urgente era a necessidade de prover o município de mão de obra. O autor acrescenta:

Alargavam-se as lavouras de café e avultavam as colheitas. Em sérios embaraços, achar-se-iam no futuro os fazendeiros com a deficiência de braços que já ia fazendo sentir, principalmente entre aqueles que só tinham escravos ao seu serviço, cujo numero minguava dia a dia. (TAUNAY, 1939, p. 172)

Outro fator que também contribuiu para o enfraquecimento da produção cafeeira no Vale do Paraíba foi o surgimento de outra frente de produção, na região do Oeste Paulista, com terras de qualidade superior à do Vale do Paraíba, e com o emprego de técnicas mais modernas de cultivo e de mão de obra livre, na maior parte composta de imigrantes que vieram ao Brasil fugindo dos conflitos e da fome que os assolavam na Europa.

1.5.1 Os Meios de Transporte

De acordo com Cechin (1978), “até a chegada da ferrovia, quer pelo alto valor quer porque era cultivado, a princípio, próximo do litoral, o café era transportado, basicamente, em lombo de burro. A fronteira, no entanto, se expandiu e este sistema atingira a saturação” (CECHIN, 1978, p. 20). Urgiam, portanto, nesse cenário, alternativas de escoamento da produção, considerando a inexistência das estradas de rodagem e os péssimos caminhos das tropas.

¹⁴ Smith (1941), na obra: “Brazil: the Amazons and the Coast”, originalmente publicada em Nova York em 1879, descreve com detalhes o auge e o declínio do café no Brasil.

¹⁵ Desde que a Lei Eusébio de Queiróz (nome mais comum pelo qual se conhece a Lei nº 581, de 4 de setembro de 1850, promulgada pela Assembleia Geral do Império, pondo fim ao tráfico internacional de escravos no Brasil) foi promulgada, houve acentuado declínio nos braços disponíveis para serem empregados no cultivo e na colheita do café. Sobre a situação do negro no mercado de trabalho pós-abolição, consulte Jacino (2012).

A estrada de ferro seria a solução obrigatória e única, uma vez que os cursos d'água corriam em sentido contrário ao litoral (CECHIN, 1978). Dá-se início, dessa forma, à construção de vias férreas que, entre 1850 e 1890, ocorreu de forma frenética, principalmente, a partir da década de 1870 (LAMOUNIER, 2008)¹⁶, extinguindo, pouco a pouco, o transporte no lombo das mulas.

Diferentemente do que encontrara na Europa e nos Estados Unidos, a ferrovia, ao chegar ao Brasil, não encontrou um parque industrial significativo, ou seja, a rede ferroviária surgiu para auxiliar a produção cafeeira, ligando as áreas produtoras ao porto de Santos. Porém, conforme salienta Carvalho (2007), nos fins do século XIX e início do século XX, a ferrovia assumiu um novo papel na integração regional do Brasil:

Até 1890, o café era quem ditava o traçado das ferrovias, que eram vistas apenas como auxiliadoras da produção cafeeira. Após 1890, há uma inversão nessa ordem e as estradas de ferro agora passam a ser construídas tendo a função de abrir novas áreas para a expansão agrícola, principalmente do café, e conseqüentemente para o povoamento e comércio de terras. (CARVALHO, 2007)

A ferrovia chegou ao Vale do Paraíba ao final da década de 1870, fase em que a produção já se encontrava em declínio. Como as linhas férreas não beneficiaram todas as cidades da região, o tropeirismo teve certa sobrevida. A produção rural das cidades que não contaram com ferrovias passando por seu território continuou a ser transportada em lombo de mulas, com jornadas mais curtas, em alguns casos sendo possível que em apenas um dia de jornada os produtos fossem embarcados para São Paulo ou para o Rio de Janeiro.

Nesse sentido, Smith (1941, p. 17), acrescenta:

Com o tinir de guisos e o bater das ferraduras, os comboios de mulas percorrem a estrada; os animais avançam em fila, cada um com um par de sacas de café dependuradas nas cangalhas grosseiras. Atrás deles vem o tropeiro, montado ou a pé, e vestido pitorescamente, com trajes meio europeus, como os dos camponeses brasileiros. Quasi todo o café é levado para a estação desta maneira. Antigamente os péssimos caminhos não permitiam melhor condução; agora, há várias estradas boas, entre Rio e São Paulo, mas mesmo quando elas existem os agricultores apegam-se ao velho sistema; só alguns utilizam carros. (SMITH, 1941, p. 17)

¹⁶ Dados quantitativos sobre o avanço das ferrovias no Brasil podem ser encontrados em Thurber (1886, p.122-123), mencionado na nota 13 deste trabalho (p.43).

Porém, à medida que se expandiam as fazendas ao interior do Estado, mais necessária se fazia a presença dos trens para o escoamento da produção¹⁷. De acordo com Lacerda (1901), a partir de 1870, novas culturas surgiram e “de 1870 para cá se estabeleceram muitos agricultores vindos de Minas, do Rio e das muitas províncias do Norte, trazendo trabalhadores e formando grandes culturas” (LACERDA, 1901, p.5).

1.6 A IMIGRAÇÃO

O Vale do Paraíba, com seu solo desgastado, mão de obra escrava e produção declinando ano a ano não teve condições de competir em igualdade com a nova terra (oeste paulista), com produção abundante e mão de obra livre. Com o fim da escravidão e o crescimento da produção cafeeira, a necessidade de mão de obra para a lavoura foi um dos grandes problemas enfrentados pelos fazendeiros. A solução encontrada foi a vinda de imigrantes.

Simultaneamente ao aumento da produção cafeeira e à multiplicação das ferrovias, o crescimento populacional no atual Estado de São Paulo se expandiu. A necessidade de um maior contingente de mão de obra para a lavoura levou os cafeicultores a exercerem uma política de incentivo à imigração, sendo os imigrantes “parcela importante da mão-de-obra das fazendas de café” (SAES, 1981, p. 42).

O povoamento, inicialmente, acompanhava as linhas férreas, sendo estas responsáveis também pela implantação dos centros urbanos. Porém, em áreas onde o café já estava em decadência, a população permaneceu estável ou com um pequeno decréscimo. Observa-se, portanto, uma estreita relação entre o aumento demográfico das diferentes zonas cafeeiras com seu respectivo desenvolvimento econômico e a expansão da via férrea.

Sabe-se que o ponto de partida para imigração no Brasil foi o decreto de 25 de novembro de 1808, de autoria de Dom João VI, que visava a atrair europeus à colônia, subsidiando o acesso à terra aos estrangeiros, principalmente para povoar a Bacia do Prata, no extremo sul do país. No entanto, essa primeira

¹⁷ Saliente-se a importância da Companhia de Estrada de Ferro Noroeste do Brasil, construída entre 1904 e 1914, que estabeleceu a ligação entre Bauru, no centro-oeste paulista, a Corumbá-MS. Ghirardello (2002) discute a importância deste empreendimento, sobretudo para o desenvolvimento do Estado do MS e para o oeste do Estado de SP.

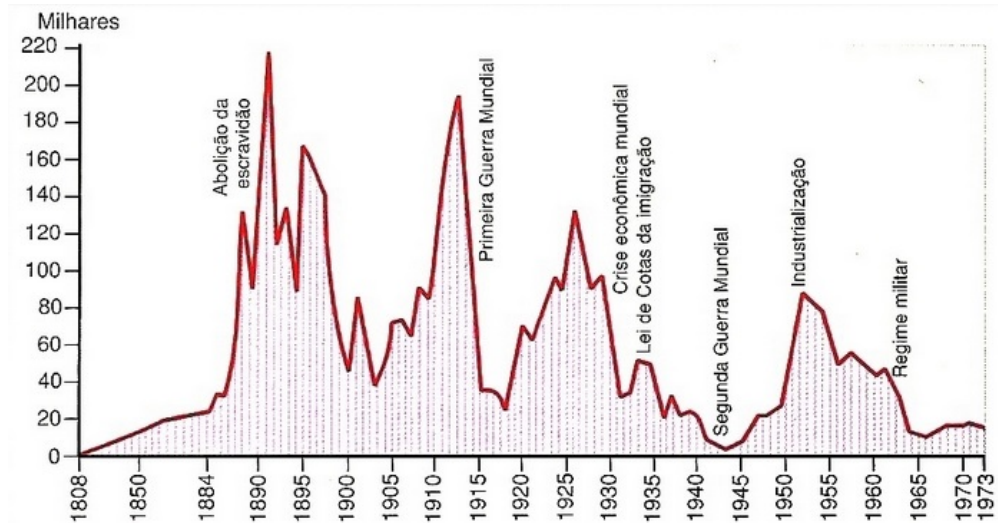
iniciativa de incentivo à imigração teve pequena repercussão, devido a fatores diversos como: a facilidade de mão de obra escrava, a instabilidade política do período regencial, os conflitos da Revolução Farroupilha (1835 a 1845) e o medo que o imigrante tinha de se tornar escravo.

Em 1827, todavia, com vistas a diminuir os problemas de mão de obra nas lavouras cafeeiras, o Ministério do Império destinou um contingente de imigrantes ao Estado de São Paulo sob o regime do colonato, mais especificamente destinando-os à região de Itapeçerica, Embu, Santo Amaro, áreas de difícil acesso, solo pobre e longe do mercado consumidor (PETRI, 2005). Nesse sistema, os colonos eram contratados na Europa e trazidos para o Brasil sob os subsídios do Império, que, como forma de adiantamento, cobria as despesas a serem pagas posteriormente. Num trabalho de parceria, fazendeiros e colonos dividiriam despesas e lucros e o colono via-se preso à terra e obrigado a saldar as dívidas oriundas de sua viagem transatlântica¹⁸.

Por conta disso, segundo Ribeiro (2006, p. 360), a introdução do imigrante europeu nas fazendas de café foi um processo lento e gradual. As primeiras tentativas, subjugando-o a um sistema de parceria, provocaram “reclamações consulares e escândalos na imprensa europeia”. Apesar da penúria que assolava a Europa na época, por causa do capitalismo-industrial, o imigrante não aceitava a coexistência com o escravo. Assim, somente após a abolição estabeleceu-se uma onda regular e ponderável de provimento de mão de obra europeia, que, nos fins do século XIX, atingira 803 mil trabalhadores, sendo 577 mil provenientes da Itália (RIBEIRO, 2006).

O uso dos imigrantes na economia cafeeira deu-se de forma mais rentável do que o emprego do escravo africano, uma vez que aqueles eram mais eficazes aos novos regimes produtivos, pois já estavam ambientados a essa prática nos seus países de origem. Os imigrantes, primeiramente, emigraram das Ilhas Britânicas, depois da França, mais tarde da Alemanha, e da Itália; por fim da Polônia, da Rússia e dos países bálticos (RIBEIRO, 2006). Adas (2004) apresenta em termos quantitativos, considerando-se o período de 1808 a 1973, dados da imigração no Brasil (Figura 4).

¹⁸ Sobre este sistema e conflitos decorrentes, confira Petri (2005).

Figura 4 - Entrada de imigrantes no Brasil

Fonte: Adas (2004, p.292)

O que se nota é a coincidência entre a intensificação do movimento imigratório e a marca da campanha abolicionista. De acordo com Camargo (1952), é justamente entre a promulgação da Lei do Ventre Livre e a da Lei Áurea que esse movimento começa a tomar vulto. Pode-se realmente assinalar o ano de 1888 como um marco na substituição do trabalho escravo pelo trabalho livre, pois, “nesse ano chega a primeira grande leva de 92.086 imigrantes para a Província [de São Paulo]” (CAMARGO, 1952, p. 57).

O processo imigratório aumentou no Brasil quando se passaram as responsabilidades de subsídios para os governos provinciais e também à iniciativa privada. Camargo (1952), baseando-se no Boletim da Diretoria de Terras, Colonização e Imigração da Secretaria da Agricultura, Indústria e Comércio de São Paulo (1937), discorre sobre a imigração na então província paulista.

O período que vai de 1827 a 1886 é aquele em que poucos imigrantes recebeu São Paulo: de 1837 a 1854, 3.227; de 1855 a 1874, 6.965, somando nesses quarenta e sete anos 11.274. Já entre 1875 e 1886 as entradas atingiram 42.343, o que quer dizer que dos 53.517 imigrantes chegados entre 1827 e 1886, os daquele período compreendiam 79,1%. (CAMARGO, 1952, p. 57)

O Governo Provincial passa a investir recursos financeiros na imigração a partir de 1881. Apresentando o número de imigrantes por país de origem que entraram na província pelos favores das leis provinciais entre os anos de 1882 e

1886, encontra-se a seguinte distribuição (Figura 5), o que demonstra a importância desses investimentos.

Figura 5 - Números de imigrantes subsidiados pela Província de São Paulo entre 1882-1886

Nacionalidades	A N O S				
	1882	1883	1884	1885	1886
Italianos	1.857	3.005	2.215	3.270	5.785
Portugueses	567	1.432	2.211	1.641	2.668
Espanhois	220	329	145	39	122
Austriacos	—	2	52	1.469	83
Dinamarqueses	—	—	2	—	448
Alemães	87	11	166	71	189
Suecos	—	—	—	3	30
Belgas	—	—	—	—	117
Franceses	2	20	32	1	49
Suiços	10	4	31	1	13
Turcos	—	—	15	—	—
Norte-americanos	—	—	—	—	12
Russos	—	—	—	2	14
Ingleses	—	9	—	2	6
Holandeses	—	—	10	1	—
Total geral	2.743	4.912	4.879	6.500	9.536

Fonte: Camargo (1952, p. 58)

De acordo com o autor, durante muito tempo os italianos tiveram a primazia nessas correntes migratórias, seguidos pelos portugueses. Dados de Holloway (1984) indicam que os italianos representaram 46% de todos os imigrantes no período de 1887 a 1930 e, durante a transição para o trabalho livre e o surto cafeeiro, representaram 73% de todas as chegadas de 1887 a 1900. Durante o período de 1887 a 1900, a Espanha forneceu 11% dos imigrantes de São Paulo, Portugal 10% e outros países 6% e somente a partir de 1901 a distribuição de nacionalidades foi mais diversificada. A proporção de italianos caiu para 26%, a de espanhóis subiu para 22%, a de portugueses para 23%, e a das outras nacionalidades alcançou 28%.

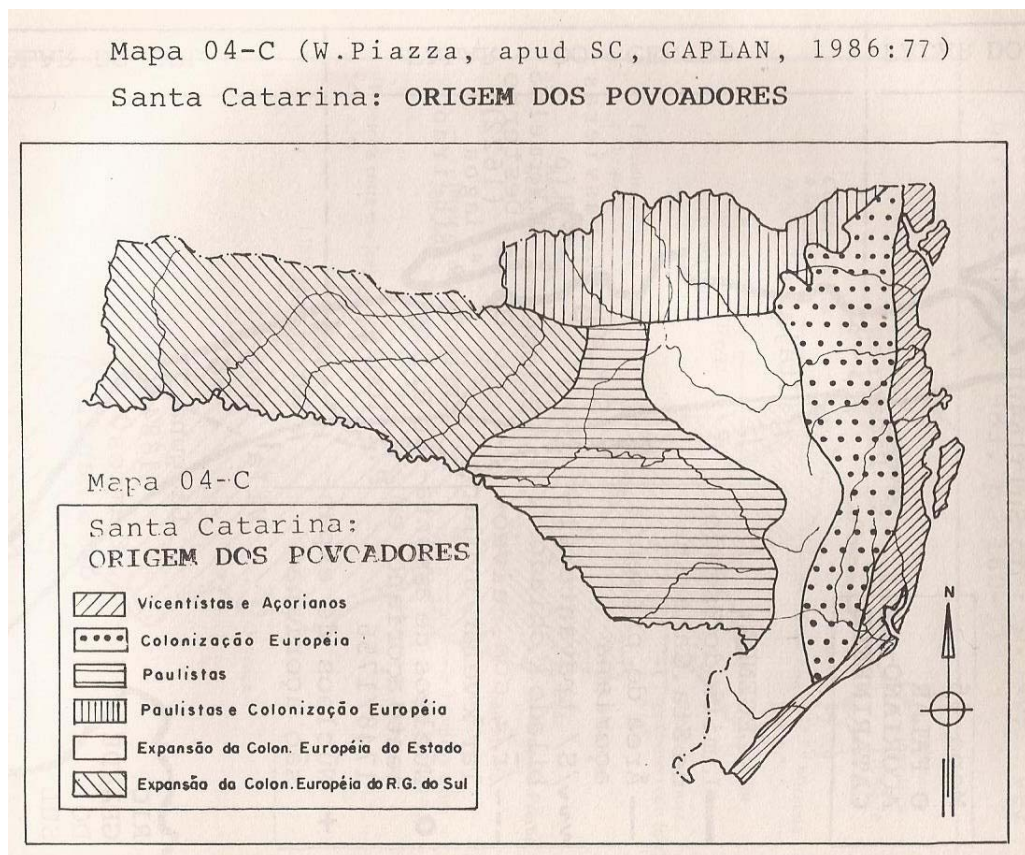
Dentre essas outras nacionalidades, Petri (2005, p.6) destaca que

o mais importante grupo de uma única nacionalidade foi de japoneses, que começaram chegando em pequeno número, em 1908, e se transformaram numa corrente contínua depois de 1917. No período de 1911 a 1930, mais de 96.000 japoneses foram para São Paulo. O Japão, dessa forma, seguiu a Itália, Espanha e Portugal como a mais importante fonte de força de trabalho para as fazendas de café.

1.6.1 Imigração nos Estados da Região Sul do Brasil

Ademais, antes mesmo da maciça imigração europeia nos Estados das regiões Sul e Sudeste do Brasil, ocorrida no século XIX, verifica-se a presença de migrantes vicentinos e imigrantes açorianos na faixa leste do atual Estado de Santa Catarina. Nesse sentido, afirma Furlan (1989), em estudo acerca da influência dos açorianos no português falado em Santa Catarina, que este Estado apresenta uma população “cuja grande maioria descende das três correntes migratórias ocorridas sobretudo entre 1748 e 1880, a saber, a açoriana (1748/1756), a germânica e a italiana, no fim do século 19” (FURLAN, 1989, p. 25). A partir de dados históricos e demográficos, o autor apresenta um mapa com a distribuição dos grupos povoadores considerando a área de sua radicação no território catarinense (Figura 6).

Figura 6 – Origem dos povoadores de Santa Catarina



Fonte: Furlan (1989, p. 233)

Desse modo, aos 4.197 catarinenses recenseados à época de 1747, acrescentaram-se 6.071 açorianos, imigrados, segundo Furlan (1989), sobretudo, das ilhas de São Jorge (36%), Pico (22,7%), Faial (15,4%), Terceira (11,6%), Graciosa (9,9%), São Miguel (4,2%), além de 579 madeirenses. Este processo emigratório decorreu da necessidade de diminuir “a sobrecarga demográfica dos Açores e para tomar posse dos extremos Norte e Sul do Brasil, expostos à invasão espanhola” (FURLAN, 1989, p. 31). Assim, a Coroa portuguesa estimulou e dirigiu de 1748 a 1756 a vinda de açorianos ao Norte (Maranhão e Grão Pará) e ao Sul do país (hoje Rio Grande do Sul e Santa Catarina) (FURLAN, 1989).

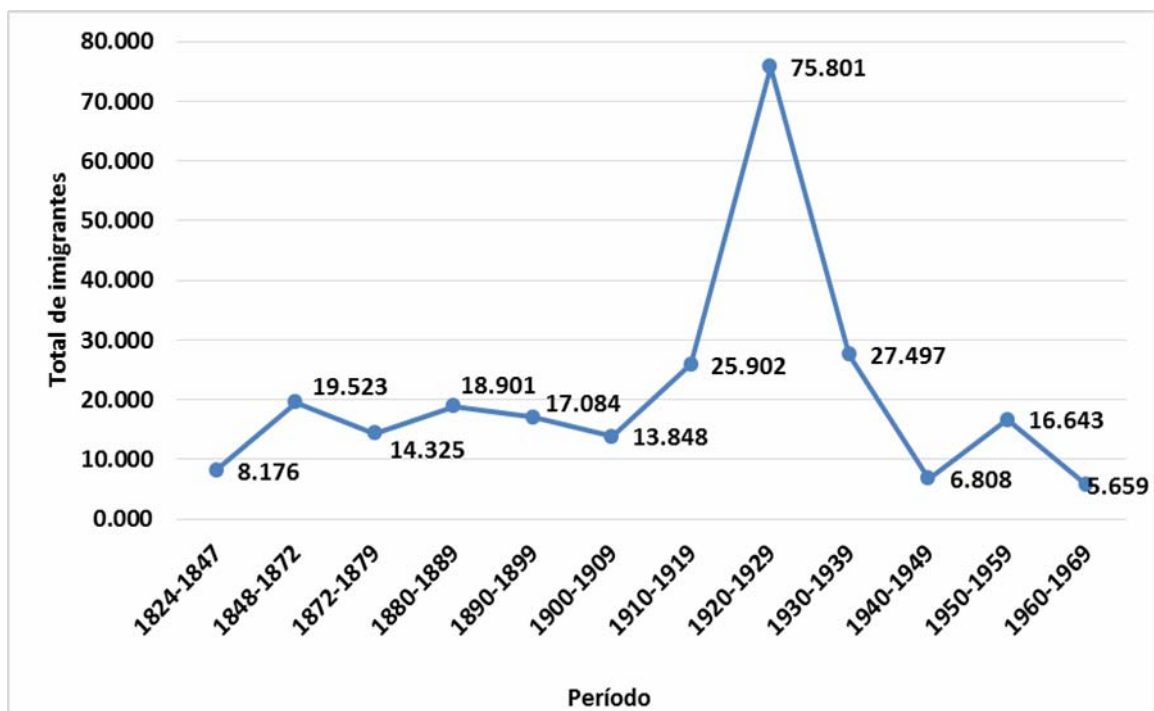
Em Santa Catarina, os açorianos emigrados radicaram-se, principalmente, na faixa litorânea. Walter Piazza, citado por Furlan (1989, p. 35), confirma que “os povoadores açorianos se fixaram em todo o litoral catarinense, desde São Francisco até ao sul de Laguna” (PIAZZA apud FURLAN 1989, p. 35), adentrando, todavia, o atual Estado do RS. Em detalhes, Piazza arrola as cidades fundadas a partir das influências açorianas:

Fundaram-se freguesias, logo de início, como São Miguel, São José, Enseada de Brito e Santa Ana (Vila Nova Mirim), no continente fronteiriço à ilha de Santa Catarina e, nesta, N. Sra. das Necessidades e Santo Antônio e N. Sra. Da Conceição da Lagoa. Posteriormente, no decorrer do século XIX, na ilha de Santa Catarina, como consequência da evolução demográfica das populações luso-açorianas tem-se a criação de N. Sra. da Lapa do Ribeirão, São João Batista do Rio Vermelho, São Francisco de Paula de Canasvieiras e Santíssima Trindade. No continente poder-se-ia, também, dizer de novas freguesias que se criaram no século XIX, mas há necessidade de novos estudos para corroborar a afirmação (PIAZZA apud FURLAN, 1989, p. 35).

Entretanto, a grande corrente imigratória na região Sul do país teve início, de fato, na primeira metade do século XIX. Em 1824, começaram a chegar os primeiros imigrantes alemães ao atual Estado do Rio Grande do Sul e Santa Catarina estabelecendo-se uma economia agrícola na área da Campanha Gaúcha, que se expandia mais tarde para o noroeste daquele Estado e para o território catarinense. Os alemães fundaram vários núcleos muitos dos quais se tornaram as atuais cidades de Novo Hamburgo e São Leopoldo (PETRI, 2005). Em termos quantitativos, Mauch e Vasconcelos (1994) apresentam a seguinte distribuição da imigração germânica no Brasil (Figura 7).

Observa-se que o maior contingente imigratório germânico concentra-se entre os anos de 1910 a 1939. Na década de 1920, este contingente chega a 75.801 alemães emigrados ao Brasil, atingindo o teto máximo, sendo este o período em que o Brasil mais recebeu imigrantes alemães. A emigração alemã, como toda a emigração europeia, se explica pelas grandes transformações sócio-político-econômicas pelas quais passou a Europa no século XIX. Sem dúvida, entretanto, a consolidação do Estado nacional alemão foi de primordial importância para o crescimento do fluxo emigratório. Acrescente-se a isto o fato de que no Brasil do século XIX abriram-se excepcionais condições favorecendo a imigração europeia. Assim, chegaram os imigrantes europeus com a finalidade de prover mão de obra para as lavouras do café e fornecer camponeses para núcleos coloniais que iam sendo criados pelo governo brasileiro.

Figura 7 – Imigração alemã no Brasil entre 1824-1969



Fonte: Mauch e Vasconcelos (1994) - adaptado

A imigração alemã na região Sul do país foi diferente do processo imigratório verificado, por exemplo, na província de São Paulo, cujo objetivo era suprir as carências de mão de obra nas propriedades de café. Nos Estados da região Sul, o objetivo era fazer do povoamento e da colonização mecanismos de conquista e de manutenção do território, povoar áreas de florestas próximas a vales

de rios. A colônia de São Leopoldo (Rio Grande do Sul) foi a primeira experiência de povoamento do Sul do país, tendo se transformado num dos grandes sucessos da política de colonização do governo imperial.

Todavia, com o passar do tempo, os colonos alemães¹⁹ expandiram-se pelo território brasileiro e levaram consigo esse sistema de colonização para além da região Sul: Espírito Santo, Minas Gerais e Bahia, por exemplo. Em um processo de migração interna, saindo de um meio rural em busca de novas terras dada a escassez das que lhes foram designadas e o crescimento populacional que reclamava por mais produção, os alemães conquistaram novos territórios. A esse movimento de deslocamento de imigrantes de um meio rural para outro, motivados por esgotamento da terra e crescimento da população, Roche (1969) denomina “enxameamento”.

Depois de ter feito recuar a floresta, esgotou o solo, obrigando os colonos das gerações seguintes a emigrar para novas zonas a desbravar ou, mais recentemente, para os centros urbanos. (...) Se o esgotamento das terras e o crescimento da população, entre outros fatores, motivaram essas migrações internas, a natalidade e a técnica incompatíveis com a estrutura agrária é que tornaram fatal o êxodo da população excedente (ROCHE, 1969, p. 319).

Por outro lado, cabe acrescentar ainda que nem todos os imigrantes alemães foram ou se tornaram proprietários de terras na ocasião de sua chegada. Muitos deles eram artesãos, industriais, comerciantes e profissionais do meio urbano, bem como religiosos e professores²⁰.

A partir do desenvolvimento de regiões coloniais e do crescimento demográfico, além das migrações para novas colônias e frentes de colonização, ocorreu continuamente o êxodo rural, que se intensificou com a modernização agrícola dos anos sessenta no século XX.

Dados de Willems (1980) (Quadro 1) revelam a diversidade e a heterogeneidade cultural dos grupos de alemães que aportaram no Brasil no século XIX. Eles vieram para povoar, preferencialmente, as colônias das Regiões Sudeste e Sul do país, onde foram estabelecidas, por iniciativa do governo imperial, as colônias de São Leopoldo (RS), São Pedro de Alcântara e Maфра (SC) e Rio Negro (PR).

¹⁹ Os colonos eram os pequenos agricultores que produziam alimentos básicos em lotes de sua propriedade.

²⁰ Para mais informações sobre a colonização alemã no Rio Grande do Sul, o leitor poderá consultar a meticulosa obra de Roche (1969).

Ainda no século XIX, os colonos alemães foram conduzidos também para outras regiões do país, como Espírito Santo, Minas Gerais, Rio de Janeiro e Bahia.

Quadro 1 – Procedência de alguns grupos alemães vindos ao Brasil no século XIX

Localidade	Fundação	Origem dos imigrantes germânicos
São Leopoldo RS	1824	Hunsrück, Saxônia, Württeerg, Saxônia-Coburg
Sta Cruz RS	1849	Renânia, Pomerânia, Silésia
Sto Angelo RS	1857	Renânia, Saxônia, Pomerânia
Nova Petrópolis RS	1859	Pomerânia, Saxônia, Boêmia
Teutônia RS	1868	Westfália
São Lourenço RS	1857	Pomerânia, Renânia
Blumenau SC	1850	Pomerânia, Holstein, Hannover, Braunschweig, Saxônia
Busque SC	1860	Bade, Oldenburgo, Renânia, Pomerânia, Schleswig-Holstein, Braunschweig
Joinville SC	1851	Prússia, Oldenburgo, Schleswig-Holstein, Hannover, Suíça
Curitiba PR	1878	Teutos do Volga
Sta Isabel ES	1847	Hunsrück, Pomerânia, Renânia, Prússia, Saxônia
São Leopoldina ES	1857	Pomerânia, Renânia, Prússia, Saxônia

Fonte: Willems (1980)

Os imigrantes alemães se adaptaram ao Brasil sem abdicar de sua cultura e seu estilo de vida, mantendo uma forte ligação com a cultura e a sociedade de origem, apesar das pressões governamentais que incentivavam o “branqueamento” dos brasileiros, fomentando a criação de colônias mistas. Por outro lado, também foram eles assimilados e aculturados pelas culturas locais. Assim, pouco a pouco, os traços de germanidade foram se enfraquecendo. A língua alemã passou a ser falada menos em público. Diminuíram também as atividades das sociedades e clubes recreativos. A educação passou a ser feita na língua portuguesa, e, em certos meios, ser alemão passou a assumir uma conotação inferior, de negação ou de exclusão (WILLEMS, 1980).

Como traço de resistência cultural formou-se um ideal de germanismo, sobretudo a partir da criação espontânea de escolas, uma vez que os imigrantes estavam distantes e isolados de outros grupos. Dessa forma, em torno de escolas e igrejas buscavam preservar elementos culturais que se estenderam por diversas gerações. A língua, como traço de identidade e cultura, passa, desse modo, a ser preservada²¹.

²¹ Dados quantitativos sobre escolas fundadas sob os auspícios de colonos alemães podem ser encontrados em Mauch e Vasconcelos (1994).

Petri (2005) salienta que, além dos alemães, vieram também suíços para o Brasil. Em 1818, eram fundados os primeiros núcleos desses imigrantes em Nova Friburgo – RJ e, em seguida, núcleos do Espírito Santo, São Paulo, Santa Catarina e Rio Grande do Sul, porém, a segunda onda imigratória de grandes proporções no Brasil foi a italiana.

Historiadores afirmam que a “grande imigração italiana” na região Sul do Brasil remonta o ano de 1875, tanto no Rio Grande do Sul quanto em Santa Catarina. Neste último Estado, de acordo com Pauli (1997), na região do vale do Rio Tijucas (leste catarinense), verifica-se “uma progressiva presença italiana, a partir de diversos núcleos pioneiros” como Tigipió (inicialmente Nova Itália), Nova Trento, Canelinhas, expandindo-se para o Vale do Itajaí até cidades como Rodeio e Rio dos Cedros. Ainda segundo este historiador, as colônias italianas no sul do Estado de Santa Catarina se estabeleceram a partir de Azambuja, localizadas às margens do Rio Pedras Grandes, constituindo núcleos como Grão Pará (1882) e Orleans (1885), em Urussanga (1878), Criciúma (1880), Nova Veneza (1891), “estes todos bem situados e com forte irradiação em torno” (PAULI, 1997).

No Rio Grande do Sul, de acordo com De Boni e Costa (1984), citados por Margotti (2004), os imigrantes italianos chegaram antes de 1875, vindos como ‘austriacos’ para o Brasil por residirem em territórios então sob o domínio austro-húngaro e se estabeleceram na região do norte rio-grandense, imprópria à produção pastoril. Segundo Margotti (2004, p. 33),

Conde D’Eu (Garibaldi), Dona Isabel (Bento Gonçalves) e Caxias foram as três primeiras colônias. Em 1877, foram fundadas Silveira Martins, a quarta colônia, em terras mais baixas, mas também montanhosas e de florestas, nas proximidades de Santa Maria, no Oeste do Rio Grande do Sul. Logo, seguiram-se outras.

Dados do IBGE revelam que os primeiros imigrantes a deixarem a Itália na época da “grande imigração” (1870-1920) ultrapassam um milhão de italianos (1.243.633 emigrados), constituindo-se, em sua maioria, de vênnetos, cerca de 30% do total, seguidos dos habitantes de Campânia, Calábria e Lombardia. (Figuras 8 e 9)²².

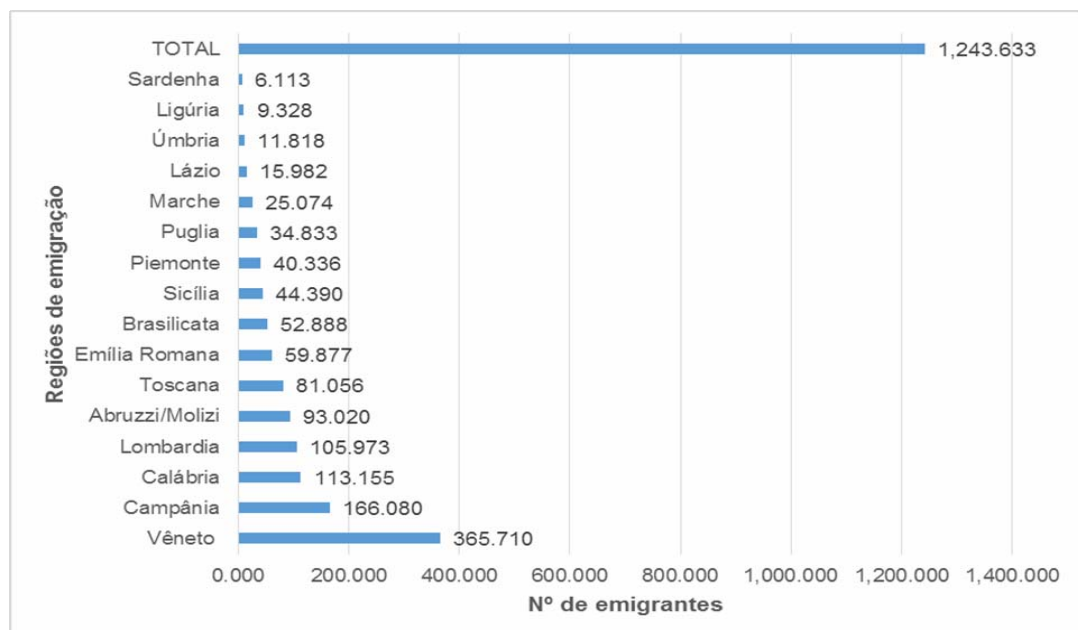
²² Sabe-se que o destino dos imigrantes no período da imigração subvencionada foram as fazendas de café de São Paulo e os núcleos de colonização, principalmente os oficiais, localizados no Rio Grande do Sul, Santa Catarina, Paraná e Espírito Santo. Entretanto, fora esses dois objetivos, uma terceira parte de imigrantes localizou-se nas cidades, como o Rio de Janeiro e São Paulo,

Figura 8 – Regiões de origem dos emigrantes italianos entre 1870-1920



Fonte: IBGE (2000)

Figura 9 - Emigração italiana para o Brasil, segundo as regiões de procedência - período 1876/1920



adensadas por indivíduos que abandonavam o campo, reemigravam de outros países ou mesmo burlavam a vigilância, não seguindo para o interior (PAULI, 1997).

Fonte: IBGE (2000) - *adaptado*

Para Gomes (2000), a importância da imigração italiana no Brasil deve-se a fatores de ordem quantitativa, conforme se observa nos dados apresentados, mas também qualitativa. Segundo o autor,

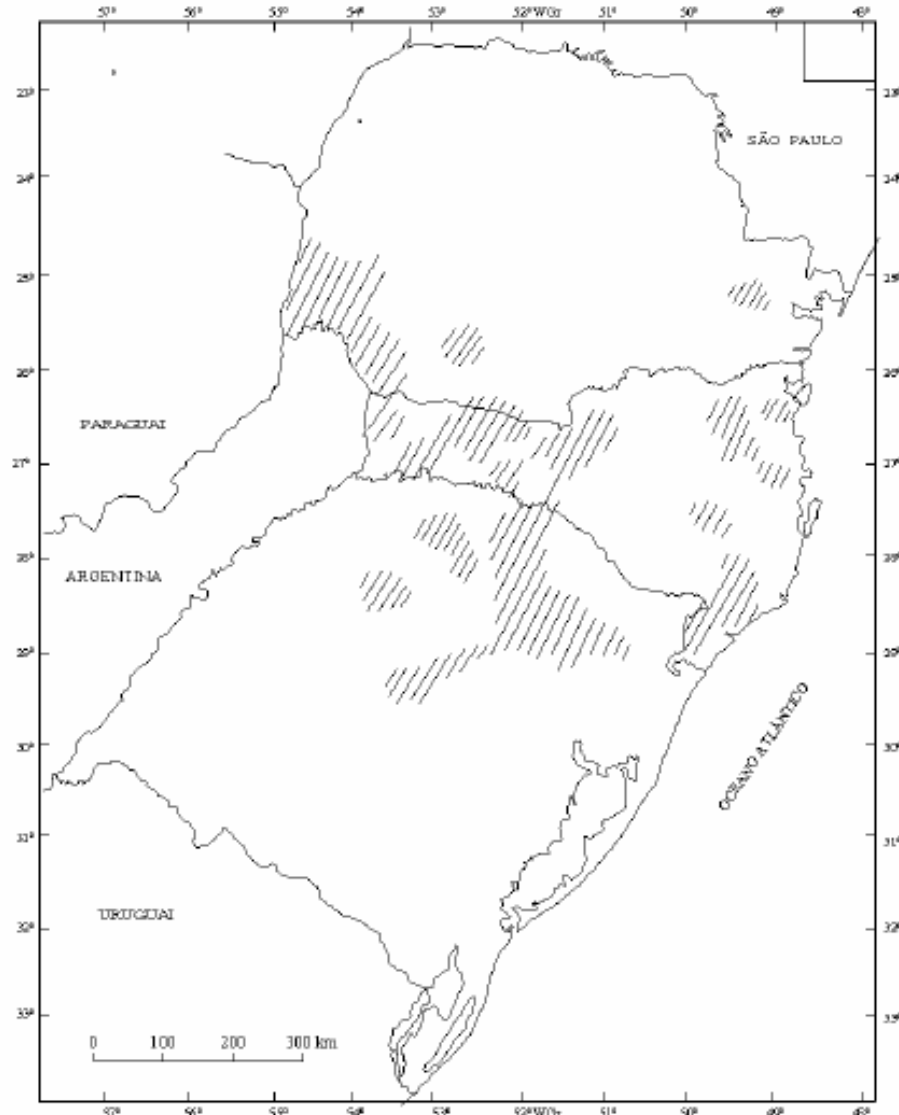
o italiano reuniu as duas condições de imigração mais valorizadas por autoridades públicas, por intelectuais e por empresários privados. A proximidade de língua, religião e costumes, fez o imigrante italiano mais facilmente assimilável por nossa sociedade do que os alemães ou japoneses, por exemplo; além disso, correspondeu aos ideais de branqueamento de nossa população, acreditado como desejável para que nos tornássemos mais "civilizados" diante de nossos próprios olhos e aos olhos do mundo (GOMES, 2000).

Margotti (2004) afirma que as colônias italianas estabelecidas na região Sul do Brasil representavam um espaço geográfico considerável, sobretudo em decorrências de processos migratórios internos, uma vez que, inicialmente, os italianos se assentaram, na região Sul do Brasil, em terras montanhosas e de floresta, ou seja, desfavoráveis à sobrevivência. Com isto, os imigrantes italianos foram impelidos a alcançarem novas terras (colônias novas) propícias, pelo menos, para a subsistência. Para o autor,

Esse espaço é composto, principalmente, por diversos municípios nos vales dos rios Itajaí-Açu, Tubarão e Araranguá, situados em terras entre o mar e a Serra Geral, em Santa Catarina, por diversos municípios nas proximidades de Santa Maria/RS, e por uma centena de outros municípios que ocupavam vasta área, desde a Serra gaúcha até o Sudoeste do Paraná (MARGOTTI, 2004, p. 41)

Nesse sentido, o autor apresenta um mapa representativo dessa distribuição espacial (Figura 10), considerando as áreas do Atlas Linguístico-Etnográfico da Região Sul do Brasil cujos informantes são bilíngues ítalo-brasileiros, ou seja, "áreas em que os ítalo-brasileiros representam mais de 50% dos habitantes" (MARGOTTI, 2004, p. 41). Todavia, reconhece Margotti (2004) que o espaço geográfico ocupado por descendentes de italianos é bem maior do que o mapa representa.

Figura 10 – Áreas geográficas da Região Sul do Brasil caracterizadas como bilingües (português/italiano) nos dados do ALERS



Fonte: Margotti (2004, p.42)

1.7 BRASIL CONTEMPORÂNEO: MIGRAÇÃO INTERNA

Nesse cenário de imigração tanto na região Sul quanto no país como um todo, o Brasil centrava suas atividades econômicas, quase que exclusivamente, na economia monoexportadora do café, pautada, principalmente, na mão de obra estrangeira, sobretudo na então província de São Paulo. Porém, com a crise mundial de 1929, que atingiu o capitalismo em função da superprodução da indústria de bens de consumo, a indústria cafeeira também sofreu consequências. Com a decadência do café, em 1930, devido à queda do preço do produto no mercado internacional, a economia passa novamente por um processo de transformação (VALLONE, 2009).

A grande depressão de 1929 não interferiu somente na economia do país, mas também na estrutura política do Brasil. Entre os anos de 1894 e 1930, o presidente da República era eleito pelos paulistas barões do café num mandato, e no outro, pelos pecuaristas mineiros, constituindo a política do “café com leite” (BUENO, 2012). A partir da crise, no entanto, o foco do poder ficou sob o gaúcho Getúlio Vargas, que se tornou o presidente da República após a Revolução de 1930, dando início a um projeto de industrialização e ao incentivo à migração interna para as regiões até então pouco povoadas.

A expansão da cultura cafeeira no oeste paulista alcançou as férteis terras do norte do Paraná e ao sul do Mato Grosso como uma ‘onda verde’ apoiada por correntes migratórias do Nordeste e de Minas Gerais, constituindo o maior movimento interno de migração, responsável por uma mudança profunda na distribuição populacional do país (MARTINS; VANALLI, 2004).

A partir da crise de 1929, inicia-se a exportação de outros produtos da agricultura e da pecuária, e algumas indústrias (de calçados, roupas e alimentos) apresentam um grande crescimento. O desenvolvimento industrial do país determinou também a existência de fortes movimentos migratórios internos, sobretudo, do Nordeste para o Rio de Janeiro e São Paulo, onde este segmento da população constituiu a base do operariado.

De acordo com Zamberlam (2004, p.68), “a migração interna é fruto da exclusão, da estrutura fundiária e das riquezas concentradas, além da imposição de políticas estatais feitas a partir de 1930, com vistas à universalização do modo de produção capitalista na economia brasileira”.

Assim, um importante movimento de iniciativa governamental contribuiu para o desenvolvimento de áreas até então pouco povoadas, como por exemplo, a região Centro-Oeste. A denominada Marcha para o Oeste teve seu início durante o Estado Novo, no mandato do presidente Getúlio Vargas, tendo como subterfúgio as ameaças internacionais de apropriação das terras pouco habitadas no Brasil.

No cenário da Segunda Guerra Mundial, difundia-se a ideia do “Espaço Vital”, ou seja, o ideal de que os países mais desenvolvidos deveriam ocupar áreas pouco exploradas pertencentes aos países menos desenvolvidos, o que colocava a região Centro-Oeste do Brasil como um alvo potencial para ocupação (MANDEL, 1989). Ameaçado por esse prognóstico, o governo brasileiro

decidiu promover a ocupação e o progresso do Centro-Oeste, subsidiando o acesso às terras com a finalidade de diminuir o desequilíbrio demográfico e incentivar a migração para as áreas que produziam a matéria-prima e os alimentos, o que contribuiria, conseqüentemente, para a industrialização do Sudeste do país. Dessa forma, essa região até então pouco habitada passou a receber migrantes de todos os Estados brasileiros, gerando um intenso progresso e um aumento demográfico significativo (DUARTE, 1988).

Dentre outros fatores que contribuíram para a migração interna, encontra-se também a modernização do campo que originou uma massa de 'desocupados', em busca novas de fronteiras agrícolas em direção às regiões Norte, Centro-Oeste e para as grandes cidades, ocasionando o êxodo rural. No Nordeste, a seca continuou gerando migrações (retirantes), especialmente em direção aos centros urbanos industrializados do Sudeste (ZAMBERLAM, 2004).

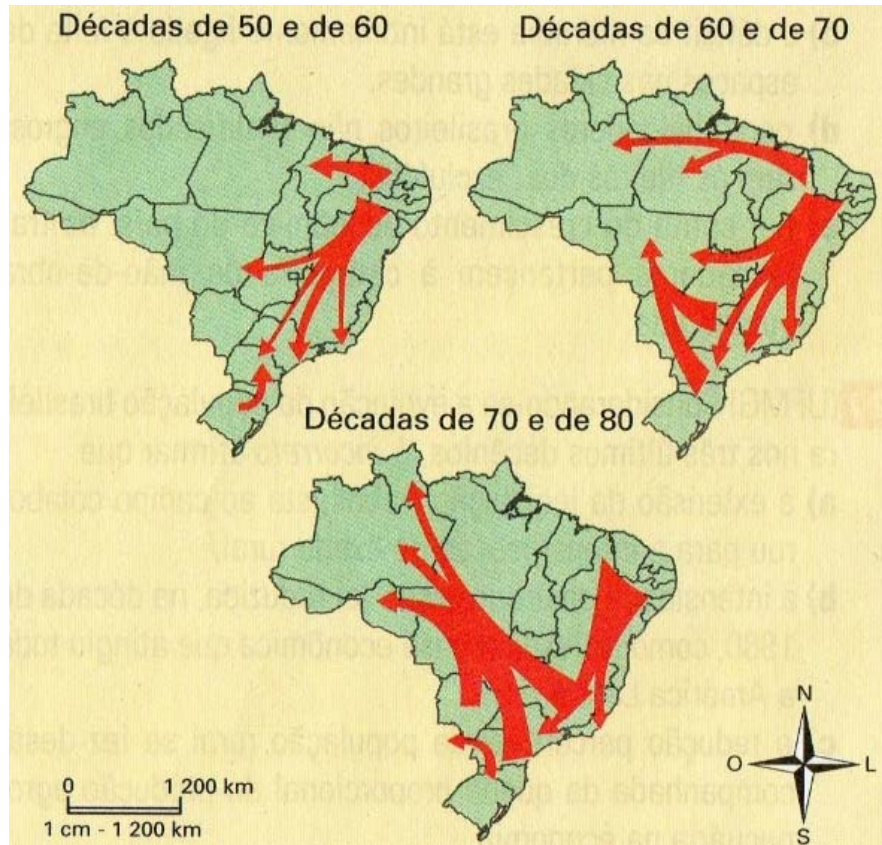
Para representar as correntes migratórias no país, Santos (1994) apresenta três mapas com os principais fluxos considerando o período de 1950 a 1980 (Figura 11).

O primeiro fluxo indica um grande movimento da região Nordeste para a Sudeste decorrente do êxodo rural, verificado no Brasil a partir de 1930. Parte da população rural foi praticamente expulsa do campo, em consequência da instalação de médias e grandes propriedades, da modernização da agricultura, da falta de incentivo ao pequeno produtor rural e da concentração de terras nas mãos de poucos proprietários. Esses fatores aliados às secas em determinadas regiões do país, provocou o desemprego de milhares de trabalhadores (COSTA, 2006). Sem possibilidade de sobreviver dignamente, muitas famílias deslocaram-se para a cidade. Famílias inteiras, na maioria das vezes, fugiam da seca, do desemprego e da falta de investimento dos governos nos lugares onde moravam.

O Brasil dos anos 1950 e 1960 se industrializava rapidamente. A fim de atrair as indústrias estrangeiras, em especial as do setor automobilístico, o Governo Federal, principalmente, garantia os investimentos necessários ao território: construíram-se estradas, siderúrgicas, redes de energia e de comunicação. A região Sudeste foi a que recebeu mais investimentos. Em cidades como São Paulo, Rio de Janeiro e Belo Horizonte, a crescente industrialização atraiu muitos migrantes, o que aumentou expressivamente a população das metrópoles da região Sudeste. Como

principal resultado, a mão de obra superou a procura de trabalhadores pela indústria.

Figura 11 - Principais fluxos migratórios no Brasil entre 1950-1980



Fonte: Santos (1994)

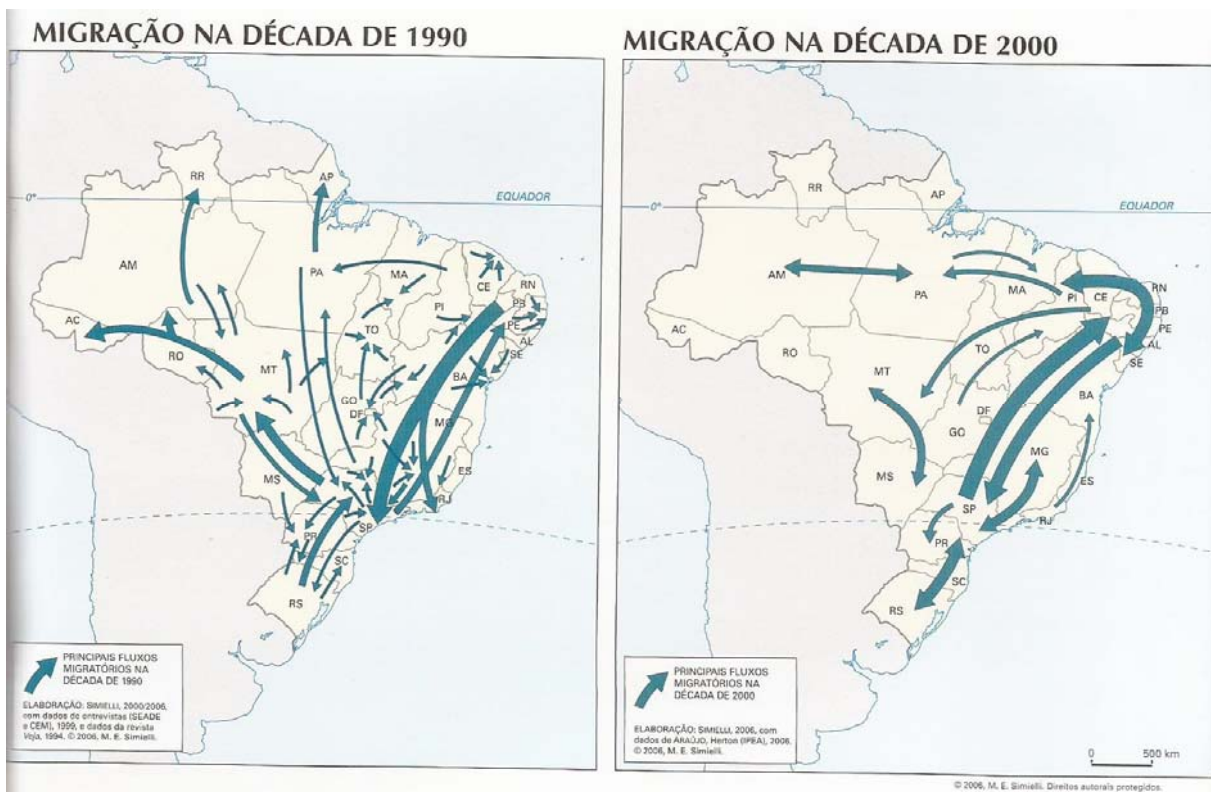
O segundo fluxo migratório correspondente ao período de 1960 a 1970, além de indicar um significativo movimento em direção ao Sudeste, registra dois novos movimentos importantes para a compreensão da migração: (i) o primeiro representado pela seta que se origina na região Sul e vai até a região Centro-Oeste. Diferentemente das migrações ocorridas nas décadas de 1950 e 1960, esse fluxo migratório não se relacionava com o desenvolvimento industrial, mas com a ampliação da fronteira agrícola, isto é, com a formação de uma nova área para a agricultura. As terras do Sul ficaram mais caras, e a concorrência com grandes empreendedores (grandes proprietários) levou os pequenos agricultores a buscar novas propriedades no Centro-Oeste. Nesta época, o governo brasileiro incentivou a migração de famílias de trabalhadores agrícolas, que saíram da região Sul e Sudeste e foram abrir fazendas no Mato Grosso e Goiás; e (ii) o segundo movimento

tem como destino a Amazônia para a qual se destinam, principalmente, os nordestinos que vão trabalhar no garimpo e em projetos agropecuários para atender aos interesses das grandes empresas; além disso, a construção da rodovia conhecida por Transamazônica estimulou a chegada de novos migrantes, conforme se verifica nas setas do mapa.

O terceiro fluxo migratório, ainda de acordo com Santos (1994), compreende o período de 1970 a 1980. Nesse período, o fluxo de migrantes do Nordeste para o Sudeste continua, porém em menor ritmo. O que se destaca, nessas décadas, é a migração rumo à região Norte, em busca dos empregos proporcionados por novas áreas de pecuária e agricultura.

A partir da década de 90, no entanto, observam-se novas tendências para as correntes migratórias. O mapa de Simielli (2006) representa esses novos movimentos (Figura 12).

Figura 12 - Migração interna (1990-2000)



Fonte: Simielli (2006)

Sobre esses novos movimentos, Menezes (2000) tece as seguintes considerações:

No Brasil de hoje o modelo dos grandes fluxos e as grandes distâncias parece acabado como dominância. Especialistas apontam para a situação de "migrações de curta distância", predominantemente intra-regionais; podendo ser sazonais em áreas de modernização agrícola ou inter-municipais em áreas de maior urbanização. A "migração de retorno" estaria associada a uma gama muito variada de contextos, o que torna bastante problemática a própria concepção de retorno: trabalhadores aposentados ou que se integram a uma rede estratégica familiar que retornam ao nordeste brasileiro – origem do grande fluxo que marcou décadas passadas; migrantes retornados da fronteira agrícola - os não retidos, migrantes que retornam dos países vizinhos, migrantes retornados a uma etapa anterior, migrantes – legais ou ilegais – de retorno de outros países. Também coloca-se em pauta a tendência à "migração solitária" e a constituição de famílias unipessoais que comportam subcategorias como migrantes albergados, migrantes de rua, migrantes de comunidades, migrantes de família quebrada, etc. (MENEZES, 2000).

Gonçalves (2001) salienta que a migração para o norte, na década de 1990, não se destinava para a área rural como ocorreu no segundo quartel do século XX, mediante os programas de colonização dirigida. Atualmente, os migrantes buscam a área urbana, como é o caso do crescimento de algumas cidades que expandiram durante os anos 90: Buritis (RO), 29,09%; Campo Novo (RO), 23,20%; Sapezal (MT), 21,54%; Palmas (TO), 21,34%; Confresa (MT), 20,84%, adquirindo muitos vícios de metrópoles sem, contudo, lograr seus serviços e benefícios.

Observa-se também que grande parte dos migrantes ainda continua migrando do nordeste para o sudeste, dirigindo-se aos grandes centros como Belo Horizonte, São Paulo e Rio de Janeiro, mas, em seguida, deslocam-se para o "interior, buscando regiões como Ribeirão Preto, Triângulo Mineiro e outras, para onde muitas indústrias também estão se mudando" (GONÇALVES, 2001), evitando as capitais devido às notícias de desemprego e violência. Vale citar Londrina (PR), Caxias do Sul (RS), Campina Grande (PB), Feira de Santana (BA), Caruaru (PE) como novos polos de atração de migrantes.

Salienta-se ainda o crescimento das áreas litorâneas decorrente do incremento do turismo em regiões como Porto Seguro (BA), litoral fluminense e litoral paulista, como também o litoral dos Estados nordestinos. Como se sabe, o turismo abre um número considerável de postos de trabalho no campo dos serviços em geral, o que se torna fator de grande atração para os migrantes. Além disso, a

presença de turistas amplia o chamado mercado informal, atraindo grande número de "trabalhadores autônomos" ou vendedores ambulantes (GONÇALVES, 2001).

O estudo realizado por Oliveir, Ervatti e O'Neill (2011), com base nos resultados do Censo 2010 do IBGE, aponta que as cidades que mais crescem no país são as que têm menos de 500 mil habitantes, o que demonstra a influência da migração, muito embora as grandes cidades ainda continuem concentrando parcela expressiva da população - aproximadamente 30% - (OLIVEIRA; ERVATTI; O'NEILL, 2011). O referido estudo apresenta também um mapa com a taxa de crescimento da população por município considerando o período de 2000 a 2010, no qual é possível visualizar alguns eixos de crescimento populacional pelas diversas regiões (Figura 13).

No Rio Grande do Sul, ao longo da BR-116, um eixo de crescimento localiza-se nas aglomerações urbanas de Porto Alegre, incorporando o município de Novo Hamburgo, São Leopoldo e de Caxias do Sul. Em Santa Catarina e no Paraná, o crescimento encontra-se entre Curitiba e Florianópolis, com destaque para a o território de Joinville e Vale do Itajaí (SC). No norte do Paraná, Londrina e Maringá compõem outro eixo com municípios com crescimento populacional como Cianorte (PR), Sarandi (PR) e outros, apoiados na rodovia PR-323 (OLIVEIRA; ERVATTI; O'NEILL, 2011).

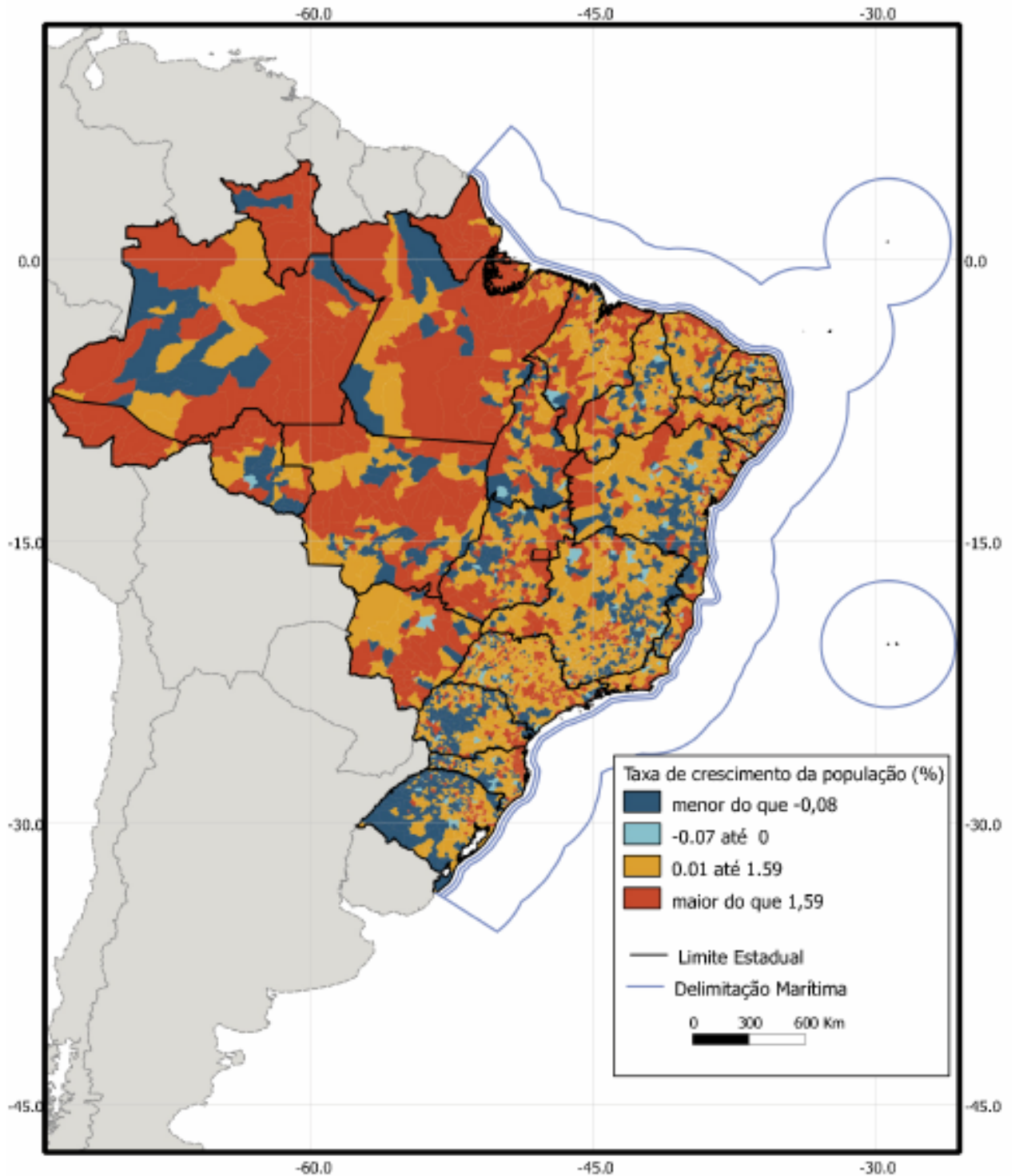
No Estado de São Paulo, observa-se o crescimento na região da capital, com a aglomeração principal, e em cinco outras regiões: Campinas, Jundiaí, Sorocaba, Santos e São José dos Campos (SP). Essas cidades absorvem o extravasamento populacional de São Paulo, além de outros migrantes de diferentes origens, conforme as causas já apontadas por Gonçalves (2011).

As aglomerações de São Paulo e Rio de Janeiro formam, ao longo da BR-116, um eixo de crescimento populacional dos mais importantes; nele localizam-se também as em regiões próximas a São José dos Campos (SP) e Barra Mansa e Volta Redonda (RJ), além de outros centros urbanos (OLIVEIRA; ERVATTI; O'NEILL, 2011).

Para o interior, as áreas de Ribeirão Preto, São José do Rio Preto e Araçatuba (SP) apresentaram dinamismo populacional na última década. Este dinamismo se estende para o Triângulo Mineiro, em direção a Mato Grosso do Sul, em uma extensa região central de Mato Grosso, extremo sul de Goiás e uma área

que incorpora Goiânia (GO), Brasília (DF) e municípios na divisa com Minas Gerais (OLIVEIRA; ERVATTI; O'NEILL, 2011).

Figura 13 - Mapa do crescimento da população segundo os municípios no período de 2000-2010



Fonte: IBGE, Censo Demográfico 2000/2010.

Notas: 1. Para o Censo Demográfico 2010, foram utilizados os primeiros resultados.

2. Para obtenção da taxa média geométrica do período 2000/2010 foram utilizadas as populações residentes em 2000 e 2010, sendo que para este último ano foi incluída a população estimada para os domicílios fechados.

De acordo com os autores, Brasília (DF), com status de metrópole nacional, atrai migrantes de todo o País. É um modelo moderno de cidade e consolida seu papel na hierarquia do sistema urbano brasileiro por força de sua criação e função de capital Federal. Goiânia e Brasília reestruturaram o sistema urbano brasileiro criando um novo arranjo espacial no centro do País. Goiânia é uma aglomeração urbana dinâmica e elemento importante na estruturação do espaço regional, tornando-se polo de relações comerciais no Centro-Oeste.

No Mato Grosso, as áreas de maior crescimento são as de ocupação mais antiga como Cuiabá, ao norte do Estado, nas áreas de ocupação mais recente, tendo como principais eixos de fixação populacional as rodovias BR-158, de Barra do Garças até o Pará, e a BR-163, Cuiabá-Santarém.

Essa rede de crescimento populacional, decorre, dentre outros fatores, das correntes migratórias e, conforme esclarece Oliveira; Ervatti e O'Neill (2011),

deve-se aos municípios em que a economia apoia-se no agronegócio, em áreas de domínio de lavouras temporárias. Em São Paulo, são as culturas da cana-de-açúcar e da laranja, em Minas Gerais o café, Mato Grosso, Mato Grosso do Sul, Goiás, Tocantins, norte do Maranhão, Piauí e oeste Baiano, o aumento populacional reflete a expansão da fronteira agrícola com a presença do cultivo, em larga escala, de produtos como milho, soja e algodão.

Verifica-se, pois, que a intrincada rede de migração interna tem propiciado certo sentido de mobilidade humana, dinamizando a população brasileira e propiciando o crescimento de médias e pequenas cidades. Dados do Censo 2010 do IBGE têm revelado, em termos quantitativos, o número de habitantes em cada um dos Estados da Federação de acordo com o seu Estado de origem.

Considerando os oito Estados que compõem esta pesquisa (PR, SC, RS, SP, MG, MT, MS e GO), antes de finalizar este capítulo, no Quadro 2, é possível verificar o número de habitantes de acordo com o Estado federativo de sua naturalidade, o que revela aspectos da mobilidade migratória/imigratória da população brasileira.

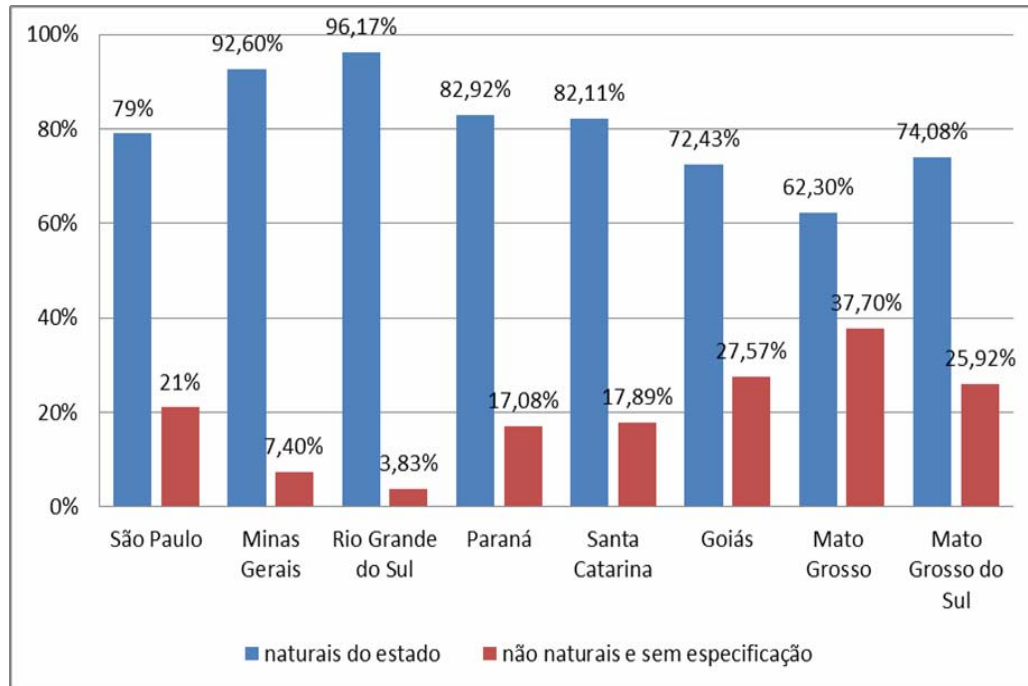
Quadro 2- População residente, por Unidades da Federação, segundo o lugar de nascimento - Brasil - 2010

Lugar de nascimento	População residente							
	Unidade da Federação							
	Paraná	Santa Catarina	Rio Grande do Sul	Mato Grosso do Sul	Mato Grosso	Goiás	São Paulo	Minas Gerais
Total	10 444 526	6 248 436	10 693 929	2 449 024	3 035 122	6 003 788	41 262 199	19 597 330
Rondônia	12 928	2 804	1 354	5 576	38 289	5 327	14 929	4 917
Acre	955	529	348	708	2 195	2 472	2 996	1 117
Amazonas	2 014	1 335	1 646	1 273	2 924	3 421	10 964	3 127
Roraima	1 671	2 082	386	480	1 015	731	1 304	363
Pará	7 368	5 275	3 261	3 196	18 743	50 824	53 534	15 296
Amapá	451	228	157	33	324	975	2 785	772
Tocantins	997	736	705	908	13 933	107 842	8 629	5 988
Maranhão	7 529	4 670	2 656	3 734	55 380	161 712	163 447	27 803
Piauí	5 049	3 804	2 310	4 352	12 864	89 628	278 511	14 712
Ceará	25 675	11 319	7 662	19 398	20 692	69 124	487 620	37 568
Rio Grande do Norte	5 024	2 980	2 072	3 646	5 845	33 197	122 755	22 223
Paraíba	12 176	4 889	2 459	6 180	8 362	43 069	349 791	21 476
Pernambuco	41 669	9 693	4 668	23 217	21 638	45 143	987 689	39 232
Alagoas	22 491	4 227	1 153	15 968	25 935	12 159	376 915	31 533
Sergipe	14 185	2 142	973	7 198	6 031	3 702	159 554	8 705
Bahia	66 276	15 550	6 682	24 863	47 674	218 329	1 702 676	181 416
Minas Gerais	192 857	20 118	10 013	39 191	96 821	302 813	1 616 885	18 164 853
Espírito Santo	14 457	2 802	1 817	2 063	14 505	6 296	50 117	85 417
Rio de Janeiro	35 633	25 535	18 383	11 212	6 978	17 795	217 956	156 256
São Paulo	552 191	99 395	33 068	214 946	136 721	92 336	32 763 387	432 236
Paraná	8 667 283	411 178	72 971	122 668	252 872	23 063	1 003 286	72 258
Santa Catarina	296 354	5 130 746	156 509	14 462	45 649	6 411	58 980	7 042
Rio Grande do Sul	279 758	422 139	10 284 936	37 741	77 640	19 999	79 210	14 920
Mato Grosso do Sul	35 758	9 909	5 805	1 814 461	92 781	13 290	97 493	8 343
Mato Grosso	25 414	7 997	4 856	30 331	1 891 052	46 452	57 096	11 884
Goiás	7 659	3 772	2 743	14 486	111 064	4 349 089	63 195	101 795
Distrito Federal	4 057	3 098	2 381	1 606	4 206	234 718	26 032	32 830
Brasil sem especificação	56 227	21 864	27 713	10 448	17 052	35 590	237 680	68 581
Exterior	50 417	17 622	34 244	14 679	5 935	8 278	266 782	24 666

Fonte: IBGE, Censo demográfico 2010.

Resumidamente, de acordo com os dados constantes do quadro 2, verifica-se que os Estados da região Centro-Oeste são os que mais apresentam moradores não naturais dos Estados, conforme se observa no gráfico comparativo elaborado a partir dos dados do IBGE (Figura 14).

Figura 14 - Gráfico comparativo da naturalidade dos brasileiros segundo o Estado federativo



Fonte: IBGE, Censo demográfico 2010

O Mato Grosso é o que apresenta maior percentual de pessoas não naturais do Estado; em contrapartida, o Rio Grande do Sul apresenta o maior percentual de pessoas naturais das cidades gaúchas (96,17%) contra apenas 3,83% de não naturais. Dentre outras conclusões, os dados revelam que a rede de migração é mais intensa nos Estados da região Centro-Oeste, ao se considerar os oito Estados federativos contemplados por este estudo.

O panorama sócio-histórico traçado neste capítulo revela que os movimentos de ocupação e povoamento do território nesta porção do país deixaram marcas na economia e, principalmente, na cultura dessas regiões. A rede de migração/imigração pode justificar a heterogeneidade cultural e até mesmo linguística nesta região do país (Centro-Sul), revelando os rastros de um mosaico lexical que se compõe e evidencia traços da história social de cada região.

CAPÍTULO 2

FUNDAMENTOS LINGUÍSTICOS

Erro de português

*Quando o português chegou
Debaixo de uma bruta chuva
Vestiu o índio
Que pena!
Fosse uma manhã de sol
O índio tinha despido
O português.*

(ANDRADE, 1972)

Processos sócio-históricos ocorridos no Brasil desde o início da colonização deram ao país uma identidade própria. A língua portuguesa trazida para cá, nos fins do século XV e início do XVI, foi enriquecida por elementos exógenos a ela, recebendo influências diretas de línguas indígenas e africanas.

Antes de impor a sua cultura e, literalmente, vestir o índio, dando-lhe fisionomias de ‘civilidade’, o colonizador precisou aprender não somente a língua dos nativos, como também seus hábitos e costumes. Esse processo ocasionou lutas e conflitos obrigando, de certo modo, o português a se ‘indianizar’ e o índio a se ‘europeizar’. Língua, cultura e raça mesclaram-se ainda mais a partir da inserção de um novo elemento, o africano, e esse constante processo de miscigenação propiciou a gestação e o nascimento do povo brasileiro.

A língua portuguesa encontrou dois momentos na nova terra, ou melhor, um período em que era uma língua transplantada de além-mar para um território novo: *A Língua portuguesa no Brasil*. Depois de algum tempo de ambientação, às antigas raízes europeias foram acrescentados novos matizes e a premente necessidade da então nascente nação exigia uma língua própria, transformando-se na *Língua portuguesa do Brasil*. Não uma língua nova, totalmente desassociada da europeia, mas uma variante brasileira da língua portuguesa na qual se encontram os indícios de sua formação.

Neste capítulo, tecem-se algumas reflexões sobre a história do português brasileiro que se fundou na nova terra de modo tão diversificado, mas também unitário, conforme já asseverara Silva Neto (1979). A evidente diferenciação da língua portuguesa do Brasil face ao português de Portugal, de certo modo,

despertou o interesse de estudiosos, preocupados com o nome que se daria a essa língua. Língua brasileira? Idioma nacional? Português do Brasil? Ou seria um dialeto do português europeu esse que se encontrava na nova terra? Língua, dialeto e falar são conceitos discutidos também neste capítulo.

A diferenciação dessa língua não se encontra somente no ponto de vista geográfico, uma vez que o Brasil possui dimensões continentais, mas também sob a perspectiva social. Desse modo, considerações sobre as origens do português popular falado no Brasil merecem ser abordadas nesta oportunidade.

Mas o foco do capítulo recai na diferenciação e distribuição espacial da língua portuguesa no território brasileiro. Assim, apresenta-se um retrospecto de trabalhos que se preocuparam em apresentar uma divisão dialetal do português neste país, bem como mapas dialetológicos com foco principal na proposta de Nascentes (1953 [1922]) até hoje referida e utilizada pelos dialetólogos brasileiros.

O avanço dos estudos linguísticos no país, porém, seja de natureza dialetológica ou sociolinguística, permite uma revisão da proposta de Nascentes (1953), sobretudo porque no início da segunda década do século XXI, o português falado no Brasil encontra-se documentado em áudio pelos pesquisadores do Projeto ALiB. Assim, cabe também apresentar este projeto e refletir sobre os resultados de trabalhos realizados sob o ponto de vista lexical com o *corpus* do ALiB que têm dado indícios da importância e da necessidade de conhecer melhor o português brasileiro, associando as diferentes áreas lexicais aos processos sócio-históricos que identificam cada região do país.

2.1 A LÍNGUA PORTUGUESA NO BRASIL

A língua portuguesa transportada para o Brasil no início do século XVI deparou-se com um grande número de línguas indígenas em uso na nova terra. Na avaliação de Rodrigues (2000, p. 25), cerca de mil e duzentas línguas eram faladas pelos habitantes da então subjugada “Índias”, pertencentes, basicamente, a três grandes famílias – o tronco Tupi-Guarani, a família Aruák e a família Karib.

Considerando a superioridade numérica dos indígenas, os portugueses tiveram que aprender as línguas dos aborígenes, instaurando durante séculos uma situação de bilinguismo no país (CASTILHO, 2004). Foi assim que uma variante do tupinambá, então chamada de língua brasílica pelos portugueses, falada

pelos povos do litoral, serviu de veículo de comunicação entre os casais de portugueses com mulheres índias e seus filhos mestiços, além dos jesuítas no processo de catequização.

Essa língua não consistia no tupi genuíno, mas era uma forma modificada, em consequência de sua disseminação pelo território brasileiro, passou a ser denominada língua geral que, para Dietrich e Noll (2010), englobava as

variedades que surgiram no Brasil devido à mestiçagem da população a partir do século XVII. Independentemente da missão dominicana, jesuítica e franciscana inicial, tornaram-se uma forma do tupi mais evoluída e simplificada, sobretudo na sua fonética e morfossintaxe. (DIETRICH; NOLL, 2010, p. 81)

Ou seja, a língua geral²³ era uma língua de base tupi que funcionou como língua veicular durante os séculos XVI, XVII e meados do XVIII até a proibição de seu uso, em 1757, pelo Marquês de Pombal.

Foi pela língua geral que o colonizador conseguiu conquistar a nova terra, estabelecendo as relações do escambo para obtenção de força de trabalho, subjugando o indígena à escravidão. Porém, a resistência dos autóctones ao trabalho forçado, aliada às campanhas jesuíticas contra a escravidão, obrigou os colonos a buscar mão de obra escrava na África para suprir as necessidades da indústria açucareira no Nordeste do país.

Embora não se disponha de números exatos sobre o contingente africano chegado ao país, sabe-se que já no século XVI os negros africanos eram trazidos para o Brasil Setentrional (Região de Pernambuco e Bahia) como mão de obra nos engenhos de açúcar. Castilho (2004, p. 239), por exemplo, aponta que de 1538 a 1855 foram traficados cerca de 18 milhões de escravos negros, sujeitos a um intenso contato com a população branca, falando um “português simplificado e deturpado” (SILVA NETO, 1979, p. 603), o que contribuiu, enormemente, para a diversificação do português no Brasil.

Os africanos trazidos para cá, ainda de acordo com Castilho (2004), compreendiam duas culturas: (i) Cultura Banto – um grupo ocidental originário do Congo e de Angola e o grupo oriental, procedente de Moçambique, Tanganika e

²³ Para Rodrigues (1996), no período colonial, existiram três línguas gerais faladas no Brasil: LGP (Língua geral paulista); LGA (Língua geral amazônica) e a GNC (Guarani Criollo). A primeira (LGP) era utilizada no centro-sul do país pelos bandeirantes paulistas; a segunda (LGA) era empregada na região Norte do país (Amazonas Pará) e a terceira no atual Estado do Paraná.

Região dos Lagos e (ii) da Cultura Sudanesa, procedente da costa oeste da África: Sudão, Senegal, Guiné, Costa do Ouro, Daomé e Nigéria.

Conforme se observa, os grupos africanos trasladados para a América apresentavam uma diversidade cultural e uma conseqüente complexidade linguística que geravam empecilhos à comunicação. Esses dois aspectos aliados à prática de misturar os escravos africanos aos indígenas para dificultar as revoltas, com o tempo, acabaram dando origem a uma espécie de koiné²⁴, conforme constata Castro (1980), a que teria sucedido um ‘dialeto português rural’, com o aumento do aportuguesamento dos africanos e a entrada de africanismos no português.

A diminuição do elemento indígena na indústria açucareira fez o escravo desde cedo ter contato com a língua portuguesa. De acordo com Lucchesi (2009, p. 46), esses africanos eram divididos em duas categorias: (i) os ladinos, aqueles que tinham proficiência na língua do colonizador e (ii) os boçais, incapazes de se comunicar nessa língua. Desse modo,

se a língua de intercurso entre colonizadores e escravos era o português, talvez tenha sido possível que, em alguns agrupamentos mais homogêneos, em termos étnicos, os escravos pudessem lançar mão de uma língua franca africana para a comunicação entre eles. O uso da língua geral tupinambá tornava-se, assim, residual. (LUCCHESI, 2009, p.46-47)

O que se verifica é que o centro de cultura açucareira (Brasil Setentrional) estava se aportuguesando cada vez mais com a vinda maciça de escravos africanos para o trabalho nos engenhos. Enquanto isso, na região periférica da colônia (de São Paulo até o Maranhão), predominava o uso da língua geral em função da submissão e aculturação dos indígenas.

Desse modo, o processo de submissão e aculturação dos indígenas contribuiu para a diversificação da língua portuguesa, conforme retratam Sampaio (1901) e Eugênio de Castro (1941), pois, ao mesmo tempo em que núcleos eram formados pela ação dos bandeirantes, ia se “dando a associação da língua portuguesa arcaica, avançando através do povoamento, com a do aborígine, que

²⁴ Derivação por extensão de sentido para qualquer língua ou dialeto comum utilizado como padrão numa vasta área (HOUAISS; VILLAR, 2001).

àquela corrompia e acrescentava novas vozes, novos valores linguísticos” (CASTRO, 1941, p. 87)²⁵.

A sociedade açucareira do nordeste desenvolveu-se sob os moldes europeus, ostentando uma cultura de requinte e luxo. Elia (1979) aponta que o jesuíta Fernão Cardim, em 1584, já asseverara que “em Pernambuco se acha mais vaidade que em Lisboa” (CARDIM apud ELIA, 1979, p. 40). Obviamente, nessa região, cultivaram-se também os padrões linguísticos da Metrópole, principalmente, em ambientes finos dos dois núcleos urbanos mais importantes da época: Salvador e Recife.

Ao contrário, na casa-grande²⁶, os escravos domésticos que possuíam maior proficiência em português não deixavam de influenciar a língua dos senhores, sobretudo as escravas encarregadas de criarem os filhos, imbuindo-lhes o seu português adquirido de forma imperfeita (LUCCHESI, 2009). Relatos dessa situação encontram-se em Silva Neto (1963 [1951]) ao apontar que o bispo Azevedo Coutinho, em 1728, já fazia recomendações às professoras de um colégio recifense, a fim de corrigirem “formas imperfeitas” do português das mocinhas:

principalmente o de pronunciarem as palavras invertendo a ordem das letras, como *breço*, em lugar de *berço* (...); ou suprimindo no meio das palavras algumas letras, como *teado* em vez de *telhado* (...); ou, ainda, engolindo a letra última, principalmente, no número plural, e nos nomes acabados em agudo, como *muitas flore* em lugar de *muitas flores*, *Portugá* em vez de *Portugal*. (SILVA NETO, 1963, p.72)

A descoberta das minas de ouro, inicialmente, em Minas Gerais e, posteriormente, em Goiás e Mato Grosso, contribuiu para o avanço da língua portuguesa no território brasileiro. Conforme já foi assinalado no capítulo 1, esse período da história do Brasil resultou em um conjunto de transformações na sociedade brasileira que ecoou durante muito tempo, culminando não só no traslado de contingentes humanos como também na transferência da capital nacional para o Rio de Janeiro – até então um pobre arraial tal qual o antigo São Vicente (DIÉGUES Jr., 1960).

²⁵ Uma vasta bibliografia de textos clássicos que tratam da língua geral e da formação linguística do Brasil nos primeiros séculos da colonização está disponível para download no site da Biblioteca Digital Curt Nimuendajú: <<http://biblio.etnolinguistica.org>>.

²⁶ Um estudo aprofundado sobre a situação do negro no Brasil colonial pode ser encontrado no livro de Gilberto Freyre (1943) – Casa Grande & Senzala.

Desse modo, a língua portuguesa chegou às regiões onde predominava o uso da língua geral, o que explica o retrocesso dessa última no século XVIII, conforme assinala Teyssier (1994):

Na segunda metade do século XVIII, porém, a língua geral entra em decadência. Várias razões contribuem para isso, entre as quais a chegada de numerosos imigrantes portugueses seduzidos pela descoberta das minas de ouro e diamantes e o Directório criado pelo marquês de Pombal em 3 de Maio de 1757, cujas decisões, aplicadas primeiro ao Pará e ao Maranhão, se estenderam, em 17 de Agosto de 1758, a todo o Brasil. Por elas proibia-se o uso da língua geral e obrigava-se oficialmente o da língua portuguesa (...) Cinquenta anos mais tarde o português eliminaria definitivamente esta última [língua geral] como língua comum, restando dela apenas um certo número de palavras integradas no vocabulário português local e muitos topónimos. (TEYSSIER, 1994, p. 76)

O foco econômico da colônia passa da região Nordeste para o Sudeste, atraindo um contingente humano sem precedentes para aquelas áreas até então desertas (DIÉGUES Jr., 1960). O Rio de Janeiro torna-se capital nacional em 1763, constituindo a principal porta de entrada de africanos e imigrantes de todos os cantos do mundo que vinham ao Brasil em busca de oportunidades e enriquecimento fácil. A mineração aumentou ainda mais o tráfico de escravos, agora para a extração de ouro e diamantes, além de desenvolver um sentimento de urbanidade na região das minas, propagar a valorização da cultura estrangeira e propiciar o surgimento de uma elite literária.

Porém, esse desenvolvimento urbano e valorização de hábitos estrangeiros afetou muito pouco a numerosa população de africanos e descendentes que se esgotavam no exaustivo trabalho da mineração, pois os escravos negros “continuavam a assimilar precariamente o português nas mais aviltantes situações de incivilidade e, em muitos casos, conservavam as suas línguas nativas para a comunicação entre si” (LUCCHESI, 2009, p. 49).

Soma-se a essa conjuntura a vinda da família real para o país, em 1808, culminando numa série de reformas que afetaram de forma decisiva o futuro da então colônia portuguesa. Naquela oportunidade, empreenderam-se investimentos de infraestrutura urbana, bem como a mudança de hábitos e costumes.

Silva Neto (1963, p. 68) relata que, junto à Família Real, vieram quinze mil pessoas trazendo “os hábitos e os esplendores da vida lisboeta”, empreendendo uma progressiva europeização do Brasil em termos demográficos. Pautando-se em relatos de Rugendas, que esteve no Brasil em 1835, o filólogo acrescenta:

Em 1808, o Rio tinha, no máximo, cinquenta mil habitantes e o número de brancos era, sem comparação, muito inferior ao de pretos. Hoje, a população é de cento e dez mil almas e a desproporção entre negros e brancos muito menor, pois, desde então, se estabeleceram na cidade mais de 24 mil portugueses, sem contar grande número de estrangeiros, principalmente ingleses e franceses, atraídos pela abertura dos portos do Brasil (...). (RUGENDAS apud SILVA NETO, 1963, p. 69)

A mudança da família Real para o país fez do Rio de Janeiro a capital do Reino de Portugal e colônias ultramarinas, trazendo vários benefícios além da abertura dos portos. Nessa época, fundou-se o Banco do Brasil; a Imprensa Real possibilitou a impressão de livros e jornais no país e a abertura dos portos possibilitou o intercâmbio com outras nações (LUCCHESI, 2009, p. 50), o que transformou, conseqüentemente, o panorama cultural brasileiro e o ambiente que envolvia a língua.

A aculturação do elemento branco aos indígenas e africanos foi se fundindo pouco a pouco, criando uma sociedade altamente miscigenada que, possivelmente, teria falado um crioulo²⁷ com muitas contribuições linguísticas de línguas indígenas e africanas. No entanto, de acordo com Biderman (2001), a cultura portuguesa era muito mais desenvolvida e os brasileiros de diferentes origens, intuitivamente, reconheciam essa superioridade. “Assim, a língua portuguesa assimilou as demais, ficando profundamente marcada por influências indígenas e africanas. Essa influência se manifestou, sobretudo, na fonética e no vocabulário”

²⁷ Crioulos são “adaptações de uma língua europeia por falantes de outras línguas, em geral africanas e asiáticas, com as quais os europeus entraram em contato por interesse mercantil (...) o crioulo é adquirido na infância, por existirem comunidades de fala crioula (...) é uma língua ‘natural’, no sentido de que uma pessoa adquire essa língua ao nascer, como qualquer outra” (CASTILHO, 2010, p. 187). No cenário nacional e internacional há linguistas que apostam na teoria crioulista do PB como, por exemplo, Coelho (1881; 1967) e Guy (1981) e outros que refutam tal teoria como Tarallo (1996) e Naro e Scherre (2007). Nos limites deste trabalho, não serão abordadas questões mais aprofundadas dessa teoria. Informações mais detalhadas sobre a teoria crioulista do PB podem ser encontradas em Parkvall e Álvares López (2003).

(BIDERMAN, 2001, p. 964) e o português adquiriu um novo sentido, o brasileiro ou do Brasil²⁸.

2.2 A LÍNGUA PORTUGUESA DO BRASIL

2.2.1 Arcaicidade do PB

Eminentes filólogos, dentre os quais se destacam Serafim da Silva Neto (1979), apontam que o português trazido para o Brasil, no início da colonização, era a língua falada por uma população pertencente às classes humildes da sociedade de diferentes regiões de Portugal, o que implica, do ponto de vista linguístico, traços conservadores da linguagem. Para o autor, “não será exagero classificar a língua dos primeiros colonizadores como pertencente ainda ao século anterior, isto é, representando um sistema linguístico muito antigo” (SILVA NETO, 1979, p.587).

Nesse sentido, Melo, em *A língua do Brasil* (1981 [1946])²⁹, documenta uma série de exemplos do caráter arcaizante do português brasileiro chegando mesmo a indicar as causas deste estagnante Estado do português, e atribui ao movimento dos bandeirantes a difusão de hábitos linguísticos rudimentares que se propalaram interior afora. Nas palavras do autor:

Estou que a nossa língua popular, falando-se de um modo geral, é substancialmente o português arcaico, deformado, ou se quiserem, transformado em certo aspecto da morfologia e em alguns da fonética pela atuação dos índios e dos negros. (MELO, 1981, p.91)

As evidências de traços arcaizantes do português do Brasil são encontradas não somente em trabalhos clássicos como os de Silva Neto (1979) e Melo (1981), mas também em pesquisas mais recentes como as de Penha (1970; 1997), Cohen et al. (1997) e Moraes de Castilho (2005).

Em contraponto à postura arcaizante do PB assumida por diversos linguistas, no texto *Conservação e inovação no português do Brasil*, de 1986, Celso Cunha discute não só a questão da unidade em meio à diversidade do PB, outro

²⁸ Optou-se, indistintamente, pela nomenclatura Português Brasileiro ou Português do Brasil (PB), para denominar a variante brasileira da língua portuguesa de acordo com nomenclatura utilizada por Silva Neto (1979).

²⁹ A primeira edição de *A língua do Brasil* foi publicada, em 1946, pela editora Agir - Rio de Janeiro.

postulado defendido por Silva Neto, mas também aponta traços de inovações do português brasileiro observados na fonética e na morfossintaxe. Para o autor:

O mito da unidade da língua popular está sendo progressivamente desmentido pelos atlas lingüísticos que se vão publicando. Ao contrário, o que eles começam a ensinar-nos é que a característica fundamental dos falares brasileiros reside no seu caráter vacilante, no seu acentuado polimorfismo. O mito da arcaicidade parece mais resistente (...). (CUNHA, 1986, p. 200)

O autor reconhece no PB traços de um português arcaico não mais encontrados no português peninsular de hoje. Para Cunha (1986), a língua portuguesa durante séculos no Brasil desenvolveu-se em condições socioculturais mais propícias à conservação do que à renovação de suas formas:

tendo vivido mais de trezentos anos sem contacto com outros povos sem imprensa, sem núcleos culturais de importância, com um número exíguo de escolas, a América Lusitana foi alcançando nesse largo período algumas das etapas que levam os povos aos Estados lingüísticos paralisantes. Esse imobilismo cultural e, conseqüentemente, idiomático é ainda muito sensível nas regiões interioranas, onde continuam a existir populações marginalizadas dos próprios acontecimentos históricos. (CUNHA, 1986, p.202-203)

Em seu estudo, o autor descreve quinze traços fonéticos e três morfossintáticos que evidenciam características de arcaização da língua portuguesa no Brasil, alguns já apontados por Silva Neto (1979). Cunha (1986), porém, apresenta também diversas realizações fonéticas que podem ser consideradas inovações americanas, considerando a falta de sua documentação nos falares portugueses de qualquer época³⁰.

O autor conclui suas reflexões considerando que os fatos descritos “não se clarificam satisfatoriamente senão com minuciosos estudos diatópicos, diastráticos e diacrônicos, que só agora começam a ser empreendidos” (CUNHA, 1986, p. 212).

³⁰ Alguns dos traços considerados inovadores por Cunha (1986) são a palatalização do /s/ pós-vocálico no Rio de Janeiro; casos de suarabácti com vogal epentética para desfazer encontros consonantais; vocalização da lateral alveolar; palatalização da consoante oclusiva dental seguida de vogal alta anterior, dentre outros traços descritos detalhadamente pelo autor nas páginas 208 a 211.

2.2.2 Das Origens Geográficas do Português Brasileiro³¹

Com a finalidade de explicar as particularidades do PB, estudiosos, por vezes, pautaram-se na região de origem dos primeiros colonizadores. Antenor Nascentes (1960), por exemplo, chega a afirmar que

Da comparação que fizemos, principalmente na fonética, entre o português de Portugal e o do Brasil resulta a conclusão que a língua que falamos é mais parecida com Alentejo e com a do Algarve, do que com as das províncias do Norte e do Centro e com as de Lisboa e Coimbra. (NASCENTES, 1960, p.262)

Todavia, Silva Neto (1979), assim como Teyssier (1994), rechaça a hipótese de que, na colonização do Brasil, predominassem imigrantes da parte meridional de Portugal, considerada pelo filólogo como uma área de colonização recente “mais evoluída e mais uniforme do que a do Norte” (SILVA NETO, 1979, p. 588), até mesmo porque faltavam dados sobre a imigração da época.

As semelhanças que aproximam o Português do Brasil ao português falado no Alentejo e Algarve, por exemplo, podem ser observadas pela ausência de alguns fonemas típicos do Norte de Portugal. Por outro lado, segundo Silva Neto (1979), o Português do Brasil também não tem nenhum traço dialetal característico do Sul de Portugal. O que se verifica é que, na Colônia, portugueses de todas as partes se fundiram em contato e interação, eliminando e expurgando os difíceis fonemas que ‘atrapalhavam’ a intercomunicação. Em suas palavras, “a pronúncia do Sul já era, ela própria, o resultado do contacto entre portugueses de todas as partes do País. A fonética do Norte duas vezes se encontrou com a do Sul: durante a Reconquista e durante a obra da colonização ultramarina” (SILVA NETO, 1979, p. 589).

A heterogeneidade na colonização resultou em características próprias que contribuíram para a formação do Português do Brasil. Assim, numa relação de solidariedade dos grupos colonizadores, quando dois indivíduos de origem diversa entravam em convivência, embora do mesmo domínio linguístico, mas de regiões distintas, deu-se um acordo tácito, eliminando-se os localismos e as características mais salientes da pronúncia de cada um, obtendo-se, nas palavras de

³¹ Uma interessante resenha acerca do assunto pode ser encontrada no último capítulo do livro de Noll (2008), no qual o autor apresenta e discute os principais trabalhos acerca da “Arcaicidade, inovação e regionalismo europeu no português brasileiro” (NOLL, 2008, p.277 a 291).

Silva Neto (1979) “um instrumento dúctil e maleável, harmonioso, capaz de servir plenamente aos interesses da intercomunicação (...) em favor do geral” (SILVA NETO, 1979, p. 588).

Naro e Scherre (2007, p. 27), por seu turno, apontam que à “época dos primeiros contatos com os portugueses, já existia em Portugal um tipo de estratégia para a comunicação com estrangeiro”, decorrente de experiências colonizadoras em explorações da África Ocidental, a denominada *língua de preto*³². Ou seja, os colonizadores não utilizaram a língua portuguesa tal qual se ensinava nas escolas e se utilizava na Corte, mas sim um português simplificado que em terras americanas se aculturou com as línguas indígenas e africanas, sem que, contudo, a sua estrutura base fosse alterada drasticamente, apesar de inúmeras influências africanas e indígenas nos níveis lexical e fonético.

2.2.3 Das Origens do Português Popular Brasileiro

Conforme já foi apontado no capítulo 1, a colonização do Brasil deu-se, primeiramente, no litoral e foi nessa região que se formaram os primeiros grandes centros urbanos, estabelecendo-se uma diferenciação entre o homem do litoral e do interior (DIÉGUES Jr. 1960). Essa diferença não se manifestou somente na economia, nos hábitos e na cultura da população, mas também em traços linguísticos que diferenciaram duas sociedades: a urbana e a rural.

Com o surgimento das cidades costeiras, o implemento da escolarização e a influência das classes mais abastadas, sobretudo a partir da vinda da Corte ao país em 1808, a linguagem adquirida de “forma imperfeita” (SILVA NETO, 1979) por escravos africanos e indígenas em contato com o português pouco a pouco foi sendo influenciada pelos padrões linguísticos utilizados pela elite da época. Porém, aqueles que não tinham acesso à escolarização adentravam cada vez mais no interior, sobrevivendo nas pequenas cidades e, principalmente, na zona rural, distanciando-se da sociedade que se formava no litoral. Para Silva Neto (1979), linguisticamente, encontrava-se o seguinte panorama no Brasil:

³² Por considerar conotação negativa a expressão, Naro (1978) sugere a denominação “língua de reconhecimento”.

Durante séculos estabeleceu-se estreito contacto e intrincada interação entre várias linguagens portuguesas: um vivo contacto vertical em que as influências se davam de alto a baixo, para o alto. Estão em choque (como nos mostram os fatos) a língua geral, a linguagem tosca e rude do negro e do índio, o Português da terra e o Português dos reinóis. Na variegada sociedade colonial brasileira o cume da pirâmide social estava decerto exposto, mais do que hoje às influências das classes sociais inferiores (...) a história do Brasil mostra-nos, em intensidade sempre crescente, a ascensão social do mestiço, o que acarreta evidentemente, a aquisição de um equipamento intelectual mais elevado e com ele um ideal linguístico que se aproxima da língua comum. A linguagem padrão – o Português normal com pronúncia nossa (conservadora em relação à atual portuguesa), algumas insignificantes divergências sintáticas e numeroso vocabulário novo – se foi sobrepondo às demais, como progressiva mancha de azeite. (SILVA NETO, 1979, p. 601-602)

Mattos e Silva (2004) salienta que, nesse processo de miscigenação linguística, o português, apesar de ter sido imposto de forma coerciva como uma língua ideal e de prestígio e até mesmo oficializada por decretos da Corte, adquiriu, no Brasil, traços característicos que o particularizam como língua nacional. Pautando-se nas conclusões de Mussa (1991) acerca de processos fonético-fonológicos do português brasileiro confrontados aos mesmos fatos encontrados nas línguas africanas, a autora afirma que as escolhas linguísticas observadas no PB não vieram a favorecer nem o português europeu nem o português africanizado:

Não há, portanto, uma “vitória do português europeu”, por sua superioridade cultural, como muitos já defenderam, nem tampouco há uma “vitória” da influência africana, como também defenderam outros. Não desconhece, nem descarta A. Mussa, como não poderia deixar de ser, que a língua do estrato dominante tem um efeito modelador, mas isso, contudo, não foi suficiente para dar feição ‘europeizada’ ao português brasileiro. (MATTOS e SILVA, 2004, p. 23)

Para a autora, devem-se evitar definições simplificadoras como a “vitória” da língua portuguesa sob as outras línguas faladas no Brasil ou afirmações genéricas como a da “descruiolização” do português popular brasileiro. Todavia, tornam-se necessários estudos aprofundados na reconstrução dos contatos linguísticos entre falantes de variadas línguas tendo como denominador comum a língua de base, a portuguesa. É imprescindível, dessa forma, “investigar fontes históricas de vária natureza para refazer a história dos fatos das sucessivas distribuições demográfico-linguísticas” (MATTOS e SILVA, 2004, p. 38).

Nesse sentido, Lucchesi, no âmbito de seus trabalhos³³, tem se dedicado a investigar as origens do português popular do Brasil partindo do pressuposto de que a caracterização da realidade linguística brasileira não pode ser feita a partir da abstração dos processos sócio-históricos que determinaram a sua formação. O estudioso entende que o português do Brasil não é apenas heterogêneo, mas também plural e polarizado, o que revela características de dois padrões linguísticos, um das classes altas e médias e outro das classes populares.

No pólo das camadas médias e altas da sociedade brasileira, de um lado, pode-se perceber, nos primeiros séculos da história do Brasil, um comportamento lingüístico conservador de uma numericamente bastante reduzida elite colonial, voltada para os padrões lingüísticos e culturais da Metrópole. No pólo das camadas populares, de outro lado, ocorrem, desde os princípios da colonização, drásticas transformações lingüísticas, decorrentes do extenso, massivo e profundo contato do português com as línguas indígenas e africanas. (LUCCHESI, 2001, p. 100)

O autor levanta a hipótese de que a grande mortandade da população indígena, devido a vários fatores, leva a crer que o papel desempenhado pelo elemento nativo na estruturação do português vernáculo ou norma linguística popular, como define o estudioso³⁴, tenha sido muito menos significativo do que o desempenhado pelos segmentos afro-brasileiros (LUCCHESI, 2001).

É fato que o português popular do Brasil surgiu a partir da língua falada nos longínquos e espacejados pontos de povoação no interior do país, ou seja, tem suas origens longe dos centros urbanos, em comunidades rurais que consistiram, durante muito tempo, a base da sociedade brasileira.

Diégues Jr (1960, p. 38) já afirmara que “foi pela organização rural que começou o Brasil” apontando que, durante séculos, o país foi uma vasta área rural que pouca ou quase nenhuma influência recebia das incipientes cidades, uma vez que “a pujança dos domínios rurais, comparada à mesquinhez urbana, representa fenômeno que se instalou aqui com os colonos portugueses, desde que se fixaram à terra” (HOLANDA, 1979, p. 60).

³³ Lucchesi é o coordenador do Projeto Vertentes do Português Popular do Estado da Bahia, vinculado ao Departamento de Letras Vernáculas da Universidade Federal da Bahia – UFBA. O Projeto objetiva traçar um panorama sociolinguístico do português popular do Estado da Bahia. Mais informações podem ser obtidas em: <http://www.vertentes.ufba.br/home>.

³⁴ Para definição destes conceitos e dos parâmetros sociolinguísticos do PB, confira Lucchesi (2006).

Cunha (1981) constata que, nos primeiros séculos da colonização, as cidades que se desenvolveram, quase todas costeiras, de pequena densidade demográfica e desprovidas de centro culturais importantes, pouca ou nenhuma influência exerciam sobre isoladas povoações interioranas. Desse modo, “cerca de quatrocentos anos, portanto, a língua, assim distribuída espacialmente, seguiu uma deriva diversificadora” (CUNHA, 1981, p. 21). Ou seja, a língua portuguesa seguiu o seu próprio curso, lançando-se à própria sorte, na boca de um grande número de analfabetos, espalhados pela imensidão do território brasileiro (CUNHA, 1970).

O português se diversificou entre os africanos, indígenas e brasileiros, formando uma variante da língua portuguesa distanciada, não apenas geograficamente, mas também linguística e socialmente, da língua falada pela elite urbana escolarizada. Para 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, nomeadamente para as suas variedades populares.

O autor considera que, desde o início da colonização até o período republicano, enquanto uma reduzida elite concentrada nas incipientes cidades guardava uma profunda fidelidade aos modelos de uso da língua provenientes de Portugal, nas imensidões do interior do Brasil grandes contingentes de índios aculturados e africanos escravizados adquiriam a língua portuguesa em condições as mais precárias; e foi essa língua “defectiva” que se converteu em modelo para a nativização do português entre os descendentes mestiços, desencadeando profundas alterações na gramática da língua portuguesa assim adquirida, socializada e nativizada.

2.2.4 A Questão da Língua Brasileira

Apesar dos numerosos ensejos de exaltar a influência indígena e africana com a finalidade de reconhecer uma verdadeira língua brasileira, para Silva

Neto e outros estudiosos, o português falado no Brasil é a mesma língua falada em Portugal, embora apresente características próprias e particularizantes.

Silva Neto (1979) considera que uma língua não age sobre a outra, pois ela não existe senão nos homens que a falam: “a interferência linguística é um dos aspectos da aculturação, isto é, dos fatos que decorrem do contacto dos homens que possuem culturas (e portanto línguas) diferentes” (SILVA NETO, 1979, p. 594).

Considerando a língua como um meio de expressão e comunicação, o linguista norte-americano Edward Sapir (1969)³⁵ observa que ela se move “pelo tempo em fora num curso que lhe é próprio. Tem uma deriva” (SAPIR, 1969, p. 151), ou seja, segue uma direção estreitando possibilidades que serão determinadas pelas condições histórico-sociais. Nesse sentido, Silva Neto (1979) afirma que o português no Brasil, desde o século XVI, tem seguido duas derivas: (i) uma conservadora, que se desenvolve lentamente e (ii) uma deriva a que condições sociais próprias imprimem velocidade inesperada (SILVA NETO, 1979, p. 595).

A primeira deriva estaria representada pelos colonizadores das diversas regiões de Portugal que precisaram articular um sistema de comunicação que visasse a um denominador comum, não utilizando inovações da metrópole e por isso era conservadora. A segunda deriva do Português do Brasil está representada pelo grande número de pessoas que tiveram que aprender a língua dos senhores de modo imperfeito e muito rapidamente num processo de transmissão linguística irregular, conforme, posteriormente, acentua Lucchesi (1994; 2001; 2006; 2009).

No entanto, observam-se na história política e social do país, algumas hipóteses que se fundam em ideais nacionalistas, defendidos por partidários que apregoavam a total autonomia do português brasileiro enquanto sistema linguístico face ao português europeu. Esses ideais estavam motivados em dois fatos da história cultural e política do Brasil: a sua Independência e o movimento romântico na sua literatura (LOBO, 2013). Assim, o “nascimento” do Brasil como Estado político independente em 1822 propiciou o desenvolvimento de uma literatura, essencialmente brasileira e profundamente marcada por um nacionalismo exacerbado que levou “a cabo a codificação escrita da norma brasileira, abandonando pouco a pouco a norma europeia” (BIDERMAN, 2001, p. 965). Estava

³⁵ A obra original de Sapir foi publicada em 1921, sob o título: **Language: An Introduction to the Study of Speech**. New York: Harcourt, Brace, 1921.

em voga o debate sobre a identidade linguístico-cultural do Brasil tendo como porta-vozes ilustres escritores, como por exemplo, José de Alencar com o seu romance *Iracema* (1865) entre outros.

Porém, essa abdicação de uma língua nacional pelos românticos, na época, é rechaçada e altamente criticada por puristas, como, por exemplo, o escritor português Pinheiro Chagas, que criticou Alencar acusando-o de escrever numa língua incorreta (TEYSSIER, 1994, p. 89). Instaurou-se, dessa forma, um período de pacificação dos ânimos sobre a questão da língua brasileira.

Afirma Teyssier (1994, p. 89), ainda, que contistas dos fins do século XIX imitavam com frequência a “língua rudimentar” do povo em suas obras repletas de brasileirismos, mas os grandes escritores tanto na poesia quanto na prosa primavam pelo purismo como Machado de Assis (1839-1908), Rui Barbosa (1849-1923), Olavo Bilac (1865-1918). Para o autor, é com o Movimento Modernista da Semana de 1922 que o ideal nacionalista e propagador de uma ‘língua brasileira’ é retomado com novo vigor.

Alcançando todos os aspectos da cultura (música, pintura, escultura, literatura, cinema etc.), o movimento Modernista representou uma revolução para a cultura brasileira, evidenciando uma ruptura com o passado e uma busca pela autenticidade e identidade da cultura brasileira, sobre a qual a questão da língua foi “um ponto de aplicação particularmente rico” (TEYSSIER, 1994, p. 89). Desse modo, autores como Mário de Andrade, Monteiro Lobato, Oswald de Andrade rebelaram-se contra a gramática tradicional e o purismo lusitano³⁶.

No texto *Manifesto da Poesia Pau Brasil*, publicado em 1924, encontra-se um trecho de total desaprovação do uso de uma língua que não seja natural e espontânea dos brasileiros. Os modernistas reivindicavam uma “língua sem arcaísmos, sem erudição. Natural e neológica. A contribuição milionária de todos os erros. Como falamos. Como somos” (ANDRADE, 1924, p.1).

No entanto, cabe salientar que os modernistas não produziram nenhuma obra importante em matéria filológica e linguística, pois foi mais pela prática do que pela teoria que deixaram as suas marcas na língua (TEYSSIER,

³⁶ Mário de Andrade, por exemplo, projetara escrever uma *Gramatiquinha da Fala Brasileira*, sem, contudo, ter colocado o projeto em prática. Pinto (1991), a partir de textos do autor, reconstrói o que seria o projeto da *Gramatiquinha* de Mário de Andrade, livro este que contempla todas as partes de uma gramática tradicional.

1994). Dessa forma, considerando também que cada escritor possuía um estilo particular que imprimia em suas obras, ora alguns se aproximavam do padrão clássico ora outros praticavam um ‘estilo brasileiro moderado’, permaneceu, entre os modernistas da segunda geração, uma linguagem mais clássica, como a de Graciliano Ramos, José Lins do Rego, Érico Veríssimo, Guimarães Rosa e Carlos Drummond de Andrade. Estava, pois, virada a página sobre a denominação que se daria ao português do Brasil, pelo menos em um primeiro momento³⁷.

2.2.5 Português Brasileiro vs. Português Europeu

Alguns estudiosos, não propriamente ligados ao movimento da literatura, mas preocupados com especificidades que diferenciavam o português do Brasil face ao português europeu, produziram obras que valorizavam aspectos linguísticos da língua portuguesa falada aquém-mar. Já no século XIX, encontram-se obras lexicográficas como o *Dicionário da Língua Brasileira*, de Luiz Maria da Silva Pinto (1832), o *Vocabulário brasileiro para servir de complemento aos dicionários da língua portuguesa* (1853), de Brás da Costa Rubim; o *Diccionario brasileiro da língua portuguesa* (1888), de Antonio Joaquim de Macedo Soares, o *Glossário paraense* (1905), de Vicente Chermont de Miranda, o *Dicionário de brasileirismos* (1912), de Rodolfo Garcia, entre outros. Neste período, surgiu também uma importante obra de cunho gramatical, *O idioma do hodierno Portugal comparado com o do Brasil* (1879), de José Jorge Paranhos da Silva³⁸.

A discussão sobre a língua portuguesa, língua brasileira, idioma nacional (obra homônima de Antenor Nascentes (1960)) ou dialeto brasileiro, conforme apregoava Leite de Vasconcelos (1970), vai repercutir durante longo tempo entre os estudiosos da linguagem.

Reacionário à polêmica, Manuel de Paiva Boléo (1943) posiciona-se sobre a temática fazendo uma revisão crítica de muitos aspectos do português do

³⁷ Nesta oportunidade, o assunto não foi aprofundado no campo de outras teorias como as da Enunciação sob a linha de pesquisa do Projeto História das Ideias Linguísticas, coordenado por Eduardo Guimarães e Eni Orlandi, da Universidade Estadual de Campinas. Outras abordagens sobre o tema podem ser encontradas na bibliografia do Projeto disponível em: <http://www.unicamp.br/iel/hil/publica/publica.html>. Confira Orlandi (2005).

³⁸ De acordo com Guimarães (2005, p.25), “em 1827 houve um grande número de discussões sobre o fato de que os professores deveriam ensinar a ler e a escrever utilizando a gramática da língua nacional”. Também remonta a essa época a reivindicação do deputado José Clemente de que os diplomas dos médicos no Brasil fossem redigidos em linguagem brasileira.

Brasil que, por vezes, são considerados por alguns estudiosos, além dos literatos já mencionados, como formas típicas do português brasileiro, os denominados *brasileirismos*. Partindo de aspectos fonéticos, morfológicos, sintáticos e lexicais, o autor afirma que muitos traços que se supõem serem típicos do Brasil também se encontram nos falares portugueses (BOLÉO, 1943) do que extrai duas conclusões:

1ª – que tanto as diferenças fonéticas e lexicais como principalmente as morfológicas e sintáticas entre o português corrente do Brasil e o de Portugal são menores do que se afirma; 2ª – que abundam em livros de filólogos, mesmo dos mais bem informados, muitos falsos *brasileirismos*. (BOLÉO, 1943, p. 47)

O estudioso concorda que a língua falada no Brasil é a portuguesa, embora enriquecida por traços dos indígenas, africanos e brasileiros que aqui conviveram durante larga escala de tempo, contribuindo de forma particular para uma diferenciação em relação à língua falada na metrópole (BOLÉO, 1943).

É certo que o português brasileiro diferencia-se do português europeu e para Castilho (2010), o simples fato da miscigenação em terras brasileiras justifica tal diferenciação linguística. É óbvio que, em terras americanas, a língua se modificaria, mantendo, no entanto, certa unidade frente às diversidades regionais, conforme já considerara Silva Neto (1979).

2.2.6 Português Brasileiro: Unidade e Diversidade

Apesar da extensão territorial, a unidade da língua portuguesa do Brasil justifica-se pelas condições da formação histórica, pois, logo de início, estabeleceram-se dois padrões linguísticos: o da costa e o do interior (ELIA, 1979). O movimento dos bandeirantes contribuiu para a propagação da língua geral ou das línguas gerais, propiciando o intercâmbio linguístico e, conseqüentemente, cultural de diferentes regiões. Nas palavras de Melo (1981),

Uma vez difundidas por todo o *hinterland* brasileiro, a civilização, a população e, por conseguinte, a linguagem bandeirante, essa primeira camada de cultura se fixou e tendeu a se conservar sempre a mesma, dado o isolamento em que ficaram os diversos núcleos, sem contatos com elementos ádvenas. (MELO, 1981, p. 92-93)

Na visão de Melo (1981), além da importância dos bandeirantes como propagadores da "linguagem rude" e mais "arcaica", "figurou na nossa geografia linguística um importantíssimo elemento de unificação, que é o rio São Francisco. Manteve ele (e mantém) em relações constantes o centro com o Nordeste, funcionando, portanto, como nivelador linguístico" (MELO, 1981, p. 93). Ou seja, elementos de ordem natural e também dos processos de ocupação humana contribuíram para a unificação linguística apesar da diferenciação regional decorrente de

vicissitudes históricas por que passaram as diversas regiões do Brasil: Nordeste, Amazônia, São Paulo (com Paraná), e o Rio Grande do Sul (com Santa Catarina). Mas estas tendências divergentes têm sempre esbarrado com a força de inércia das populações rurais, de maneira que não puderam até agora desenvolver dialetos propriamente ditos. (MELO, 1981, p. 93)

2.2.7 Português Brasileiro: Língua, Dialeto ou Falar?

Uma vez constatadas diferenças linguísticas entre o português do Brasil em relação ao português de Portugal, os estudiosos começaram a se preocupar com o processo de dialetação da língua portuguesa falada no Brasil. Alguns estudiosos, já na primeira metade do século XX, reconheciam a necessidade de se empreender trabalhos empíricos, executados com rigor metodológico, com a finalidade de se constatar se realmente há um dialeto brasileiro ou dialetos brasileiros. De acordo com Amaral (1981 [1920]), "só então se saberia com segurança quais os caracteres gerais do dialeto brasileiro, ou dos dialetos brasileiros, quantos e quais os subdialetos, o grau de vitalidade, as ramificações, o domínio geográfico de cada um" (AMARAL, 1981, p. 44).

Boléo (1943, p. 9) reconheceu que este é um ponto nevrálgico para os que se dedicam a estudar o português do Brasil e afirma veementemente que "o português falado no Brasil não é um dialecto e muito menos uma língua; é apenas uma variante do português europeu, ou seja, um falar da língua portuguesa".

Dessa forma, discute-se no decorrer desta seção o conceito de dialeto e falar, pois se observa que não há unanimidade entre os estudiosos na definição dos termos, uma vez que alguns consideram o critério geográfico e outros, os critérios sociais, que dariam origem a dialetos sociais.

Conforme se vem assinalando ao longo deste estudo, de acordo com o referencial teórico adotado - Silva Neto (1979), Boléo (1943), entre outros - não há uma língua brasileira. Para o conceito de língua, adota-se a definição de Alvar (2007 [1961])³⁹, para quem

Lengua es, en la acepción que aquí nos ocupa, el sistema lingüístico del que se vale una comunidad hablante y que se caracteriza por estar fuertemente diferenciado, por poseer un alto grado de nivelación (...). (ALVAR, 2007 [1961])⁴⁰.

Não cabe, portanto, mais detalhamentos desse assunto, uma vez que não se reconhece o PB como uma língua desvinculada do português europeu, nem mesmo como um dialeto da língua portuguesa transportada para a América.

Porém, estudiosos como Leite de Vasconcelos (1970), um dos maiores nomes da Dialectologia portuguesa, com a finalidade de sistematizar o estudo da língua portuguesa, divide-a em dialetos, reconhecendo que estes podem se subdividir em subdialetos e variedades. O dialetólogo português guiou-se pelo critério geográfico para a sua classificação dos dialetos em: continentais, insulares e ultramar.

O dialeto continental é o resultado da evolução do latim vulgar para o português, ou seja, é a língua portuguesa falada em Portugal, enquanto os insulares e os de ultramar representam o comportamento do português europeu transplantado para outro território. Para esse estudioso, a língua portuguesa do Brasil encontra-se entre os dialetos de ultramar, sobre a qual faz as seguintes considerações:

La langue nationale du Brésil est le portugais, qui, transporté dans un milieu si différent de celui de son origine, y a éprouvé beaucoup de modifications (...) Si j'appelle *dialecte*, par exemple, le portugais de Trás-os-Montes, à plus forte raison je dois donner ce nom au portugais du Brésil, ou *brésilien* (...). (LEITE DE VASCONCELOS, 1970, p.132)⁴¹

³⁹ O texto original de Alvar foi publicado na *Nueva Revista de Filología Hispánica*, Año 15 (1961), pp. 51-60. A versão digital do texto está na Biblioteca Virtual Miguel de Cervantes. Disponível em: < <http://www.cervantesvirtual.com/obra/hacia-los-conceptos-de-lengua-dialecto-y-hablas-0/>> Acesso: 07 set. 2013.

⁴⁰ “Língua é, no sentido que nos interessa aqui, o sistema linguístico que uma comunidade de falantes utiliza e que é fortemente caracterizado por ter um elevado grau de nivelamento e diferenciação,” (TN = tradução nossa).

⁴¹ “A língua nacional do Brasil é o português, que, transportado em um meio tão diferente do de sua origem, sofreu muitas modificações (...) Se chamo dialeto, por exemplo, o português de Trás-os-Montes, com mais razão devo dar esse nome ao português do Brasil ou brasileiro” (TN).

Leite de Vasconcelos considera, pois, o português do Brasil um dialeto do português europeu que, em solo americano, adquiriu uma nova configuração, mesmo reconhecendo que tal posicionamento feriria os “melindres nacionais”, conforme já afirmara Nascentes (1953).

Ferreira e Cardoso (1994), para definir dialeto, mobilizam o conceito de isoglossa⁴², pois, para as pesquisadoras, o dialeto define-se como um feixe de isoglossas, ou seja, “um conjunto de isoglossas que se somam e que, portanto, mostram uma relativa homogeneidade dentro de uma comunidade linguística em confronto com outras” (FERREIRA; CARDOSO, 1994, p. 16).

A definição das autoras vai ao encontro da postura de Montes Giraldo (1987), que, por sua vez, também associa o termo dialeto ao conceito de isoglossa. Para o pesquisador colombiano,

una isoglossa es la línea que señala en la representación gráfica el área o dominio de vigencia de una determinada norma. Y, por lo tanto, un conjunto de isoglossas más o menos coincidentes nos determina un dialecto, esto es, una forma idiomática distinguida por un conjunto de normas. (MONTES GIRALDO, 1987, p. 55-56)⁴³

Neste trabalho, adota-se o conceito de norma na perspectiva de Coseriu (2004 [1980]), ou seja, norma é aquilo que se diz e não o que se deve dizer. A norma é aquela que o indivíduo segue por fazer parte de um grupo. Em suma, norma é “tudo o que, no falar correspondente a uma língua funcional, é fato tradicional, comum e constante, ainda que não necessariamente funcional: todo fato que diz e se entende dessa maneira e não de outro modo” (COSERIU, 2004, p.122). Vale notar que o conceito de norma é muito discutido entre os linguistas. Nesta oportunidade, não se aprofunda esta discussão, uma vez que em Romano (2012) encontram-se as definições. Ademais, Faraco (2002), Castilho (2002), Lucchesi (2002) discutem com propriedade conceitos de norma culta, norma padrão, norma objetiva e norma subjetiva.

⁴² “uma linha virtual que marca o limite, também virtual, de formas e expressões linguísticas. As isoglossas podem delinear contrastes e conseqüentemente apontar semelhanças em espaços geográficos (isoglossas diatópicas), podem mostrar contrastes e mostrar semelhanças linguísticas sócio-culturais (isoglossas diastráticas) ou ainda podem configurar diferenças de estilo (isoglossas diafásicas)” (FERREIRA; CARDOSO, 1994, p.13).

⁴³ “Uma isoglossa é uma linha que assinala na representação gráfica da área ou domínio de vigência de determinada norma. E, portanto, um conjunto de isoglossas mais ou menos coincidentes nos determina um dialeto, isto é, uma forma idiomática que distingue por um conjunto de normas”. (TN).

Na perspectiva de Montes Giraldo (1987), os dialetos são subdivisões de uma língua estabelecendo uma relação de subordinação a “una entidad mayor, su inclusión en un conjunto más vasto, por lo que he definido dialecto en la forma más general possible como D ç L (D incluído en L, es decir, dialecto como parte de una lengua – lengua histórica o lengua-idioma)” (MONTES GIRALDO, 1987, p. 50)⁴⁴.

Segundo Câmara Jr. (2007, p. 115)⁴⁵, “dialetos, do ponto de vista puramente linguístico, são falares regionais que apresentam entre si coincidência de traços linguísticos fundamentais”, ou seja, o dialeto para este linguista está relacionado à distribuição da língua no espaço geográfico, portanto, conceito ao conceito de Leite de Vasconcelos (1970). “Entretanto, ao conceito linguístico se acrescenta 9...) um conceito extralinguístico de ordem psíquica, social ou política: a) a existência de um sentimento político comum (...); b) a existência de uma língua culta, superposta aos dialetos, que assim ficam limitados ao uso cotidiano, sem maior expressão cultural ou literária; c) a subordinação política das respectivas regiões como partes de um Estado político nacional” (CÂMARA Jr. 2007, p. 116).

Para o linguista, levando-se em conta a existência de “traços linguísticos essenciais” (i), pode-se “dividir a língua portuguesa em dois grandes dialetos, correspondentes a nações distintas: o lusitano (ou português europeu), em Portugal; e o brasileiro, ou português americano, no Brasil. Para cada uma dessas nações, há, por sua vez, uma divisão em dialetos menores e subdialetos, que leva em conta as condições extralinguísticas (b) e (c).

Manuel Alvar (2007), no entanto, buscando certo rigor científico para a definição de dialeto, faz uma crítica às definições registradas em três dicionários especializados⁴⁶.

⁴⁴ “uma entidade maior, sua inclusão em um conjunto mais vasto, pode ser definido dialeto em forma mais geral possível como DçL (D incluído em L, quer dizer, dialeto como parte de uma língua – língua histórica ou língua-idioma)” (TN).

⁴⁵ A primeira edição desse dicionário data o ano de 1956, lançado pelo Centro de Pesquisa da Casa Rui Barbosa sob o título *Dicionário de Fatos Gramaticais*.

⁴⁶ Os dicionários de termos linguísticos criticados por Alvar (1961, reproduzido em 2007) são: o de J. Marouzeau, *Lexique de la terminologie linguistique. Français, allemand, anglais*, Paris, 1943. O *Diccionario de términos filológicos*, Madrid, 1953, de Fernando Lázaro Carreter e o *Dicionário de fatos gramaticais*, [Rio de Janeiro], 1956 de Joaquim Mattoso Camara Jr. O autor considera que os dois primeiros apresentam perspectivas diacrônicas em suas definições, enquanto este considera aspectos sincrônicos. Assim, ambas as definições têm pontos insatisfatórios conforme descreve detalhadamente no artigo. Vale notar que as definições trazidas por Camara Jr., na obra de 1956, têm correspondência unívoca com as encontradas na 26ª edição de seu *Dicionário de Linguística e Gramática referente à língua portuguesa* (Ed. Vozes, 2007).

Segundo Alvar (2007) “la definición del tratadista brasileño involucra una serie de conceptos que no pueden mezclarse. Bien es verdad que su postura está lejos de la objetividad científica y limitada por respetables compromisos sentimentales”⁴⁷. E mais à frente acrescenta:

El propósito de Mattoso Câmara es, simplemente, llegar a una conclusión: la paridad lingüística de Portugal y el Brasil; la no dependencia lingüística de América con respecto a la antigua metrópoli. Este problema queda al margen de mi interés actual. Voy a atender, sólo, a su terminología lingüística. De ella no se puede aceptar la expresión *lengua regional*, porque, en sentido especializado, *lengua* se opone a *dialecto* por ser «un idioma nettement différencié de tout autre» (MAROUZEAU, p. 128, s. v. *langue*). El uso de *lengua* en la definición de dialecto es abusivo (aunque el tratadista vaya guiado por un determinado sentimiento nacional), sobre todo si se tiene en cuenta el término *habla regional* (*falar regional* en portugués), que explícitamente aclara la cuestión. (ALVAR, 2007)⁴⁸

Dessa forma, Alvar (2007) considera dialeto como um sistema linguístico decorrente de uma língua viva ou desaparecida, normalmente com uma concreta limitação geográfica, porém sem uma forte diferenciação frente a outros de origem comum. De modo secundário, o autor declara que “pueden llamarse dialectos “las estructuras lingüísticas, simultáneas a otra, que no alcanzan la categoría de lengua” (ALVAR, 2007)⁴⁹.

O autor inclui ainda na sua classificação duas outras categorias, os falares regionais e os falares locais. Os primeiros são as peculiaridades expressivas próprias de uma região que carecem da coerência que têm os dialetos. De acordo com o autor “las hablas regionales se caracterizan por ser dialectos empobrecidos, que han abandonado el uso escrito para limitarse a ser manifestaciones orales”⁵⁰

⁴⁷ “a definição do tratadista brasileiro envolve uma série de conceitos que não podem ser misturados. É verdade que sua posição está longe de objetividade científica e limitada por respeitáveis compromissos sentimentais” (TN).

⁴⁸ “O propósito de Mattoso Câmara é, simplesmente, chegar a uma conclusão: a paridade linguística de Portugal e Brasil; a não dependência linguística da América em relação à antiga metrópole. Este problema fica à margem de meu interesse atual. Vou me ater, somente, à sua terminologia linguística. Dela não se pode aceitar a expressão língua regional, porque, em sentido especializado, língua se opõe a dialeto por ser “um idioma claramente diferenciado do outro” (MAROUZEAU, p. 128, s. v. *langue*). O uso de língua na definição de dialeto é abusivo (embora o tratadista se guie por um determinado sentimento nacional), sobretudo quando se considera o termo fala regional (*falar regional* em português), que explicitamente esclarece a questão” (TN).

⁴⁹ “Podem se chamar de dialetos as estruturas linguísticas simultâneas a outras que não alcançam categoria de língua” (TN).

⁵⁰ “As falas regionais se caracterizam por ser dialetos empobrecidos, que abandonaram o uso escrito para limitar-se a ser manifestações orais” (TN)

(ALVAR, 2007). Já os falares locais são as estruturas linguísticas que apresentam pequenos traços diferenciados, mas com nuances características dentro da estrutura regional a que pertencem e cujos usos estão limitados a pequenas circunscrições geográficas, geralmente, municípios, distritos, vilas.

O que se observa em Alvar (2007) é uma relação hierárquica entre os termos, partindo da língua que apresenta diferenciações geográficas, portanto, dialetos, que por sua vez compreendem em si falares regionais, e estes comportam os falares locais.

Pautando-se na distinção proposta por Alvar, Brandão (1991), em seu livro *Geografia Linguística no Brasil*, discute brevemente os conceitos de língua, dialeto e falar. Porém a autora não chega a uma nítida distinção entre os termos

Embora Alvar tenha procurado definir, com rigor, tais variantes diatópicas, na prática se torna difícil estabelecer uma precisa diferença entre dialeto e falar (...) Nota-se, no entanto, forte tendência a empregar-se o termo “dialeto” em sentido mais amplo, ou seja, o de considerar-se dialeto como sendo qualquer variedade linguística – quer de natureza geográfica, quer de cunho social – que constitua um subsistema singular, unitário. (BRANDÃO, 1991, p. 13)⁵¹

Uma interpretação diferente ao posicionamento de Alvar (2007) e Brandão (1991) é a de Boléo (1943), que apresenta uma distinção entre os termos “dialecto e falar (ou linguajar, como dizem os Brasileiros)” (BOLÉO, 1943, p. 10).

Revisitando critérios para a distinção entre língua, dialeto e falar, Boléo (1943) afirma que “o falar é que é mais amplo. Partindo do geral para o particular, nós teremos na terminologia que adopto: língua, falar, dialecto, podendo ainda admitir o subdialecto” (BOLÉO, 1943, p. 16) e salienta que apenas uma particularidade linguística para determinar se se trata de dialeto ou falar não é suficiente, uma vez que se faz necessário um conjunto de traços para a delimitação de tais termos.

Para se referir ao Português do Brasil, o autor prefere o termo falar ao termo dialeto justificando que

⁵¹ Em nota, Brandão (1991) completa que para as variações de cunho social também é corrente a designação “socioleto”, conceito este que não é discutido nesta oportunidade.

o que me leva a proceder assim não é um motivo de ordem psicológica: o desejo de evitar uma palavra que fere a sensibilidade dos Brasileiros; seria uma consideração extralinguística, que de modo algum podia influir no meu espírito. É o próprio exame do português de terras de Santa-Cruz, tal como o tenho ouvido a brasileiros que vem à Europa e como nos é apresentado em livros de filólogos desse país, que me leva àquela conclusão. (BOLÉO, 1943, p. 17)

Antenor Nascentes (1953), em *O linguajar carioca* (1953), reconhecendo a polêmica entre portugueses e brasileiros sobre a noção do termo dialeto, e considerando que ela ocorre muito mais por questões políticas do que propriamente linguísticas, opta por adotar o termo *falar* ao se referir ao português brasileiro, assim como o fez Boléo (1943).

Nascentes afirma que, embora seja difícil estabelecer distinções entre *falar* e *dialeto*, é menos inconveniente utilizar o termo *falar* ao se referir ao português do Brasil. Assim, para esse estudioso, dentro do Brasil, existem dois grandes grupos, o do norte e o do sul, e estes se subdividem em subfalares. O grupo do norte contém o subfalar amazônico e o subfalar nordestino; e o grupo do sul abrange os subfalares baiano, o sulista, o fluminense e o mineiro. Com base em Marouzeau (1943), defende Nascentes que o falar é “um conjunto de meios de expressão empregados por um grupo no interior de um domínio linguístico” (MAROUZEAU apud NASCENTES, 1953, p.17).

Após a revisão da literatura e considerando-se a polêmica envolvida nas definições dos termos (*dialeto*, *subdialeto*, *falar*, *subfalar*), adota-se, neste trabalho, o posicionamento de Boléo (1943) e de Nascentes (1953), com ressalvas, pois não se compreende que o falar compõe-se de dialetos. Antes, procura-se não confundir os conceitos. Para o autor desta tese, pautando-se em Montes Giraldo (1987) e em Ferreira e Cardoso (1994), o uso do termo dialeto refere-se a um conjunto de isoglossas mais ou menos coincidentes em determinado território. Desse modo, delimitar áreas dialetais ou a existência de dialetos somente é possível a partir de dados de natureza léxico-semântica, fonético-fonológica, prosódica ou mesmo morfossintática que se coincidem em determinado território. Somente um estudo minucioso das cartas de um atlas linguístico (lexicais, fonéticas, morfossintáticas) pode indicar a presença ou não de dialetos do português usado em determinada região.

Como este trabalho não se destina à discussão sobre a existência de áreas dialetais no Centro-Sul do Brasil, mas sim se propõe a discutir a divisão de Antenor Nascentes e propor uma nova classificação a partir de dados lexicais, acerca da área denominada pelo estudioso como subfalar sulista, procurou-se utilizar a mesma nomenclatura desse autor, embora se reconheça a imprecisão dos conceitos.

Considera-se, portanto, que o subfalar sulista está dentro do grupo do sul denominado por Nascentes (1953), evitando-se o uso do termo dialeto. Todavia, a área geográfica correspondente ao subfalar sulista estabelecida por Nascentes (1953) não pode ser considerada em sua totalidade, uma vez que ao conjunto dos subfalares que formam o grupo do sul, acrescentam-se outro subfalares, localizados na área geográfica do subfalar sulista. Em outras palavras, aos subfalares fluminense, mineiro, baiano e sulista acrescentam-se outros subfalares, sobretudo, o paulista, evidenciados a partir de dados lexicais devidamente mapeados.

2.2.8 Sobre os Estudos do Português Brasileiro

Já nos fins do século XIX e início do XX começam a surgir os primeiros trabalhos de cunho dialetal que visavam a descrever a realidade linguística do português do Brasil considerando-se as diferentes regiões⁵².

Mota e Cardoso (2006, p. 17) registram que a primeira manifestação que se pode caracterizar, numa visão ampla, de natureza dialetal sobre o português do Brasil, ocorre no ano de 1826, e deve-se a Domingos Borges de Barros, o Visconde de Pedra Branca, que, a serviço de Adrien Balbi, anotou algumas peculiaridades do português brasileiro para integrar ao *Atlas Ethnographique du Globe* (1924-1925).

Mas, tradicionalmente, o primeiro trabalho sistemático acerca da diferenciação dialetal observada no PB data de 1920, ano em que se publicou o livro de Amadeu Amaral - *O Dialeto Caipira* (1920), inaugurando, basicamente, os primeiros passos da Dialectologia brasileira. Seguiram-se ao *Dialeto Caipira* outras obras de cunho monográfico como *O linguajar Carioca*, de Antenor Nascentes

⁵² Sobre a periodização dos estudos dialetais no Brasil, confira Ferreira e Cardoso (1994) e Mota e Cardoso (2006).

(1922/1953), *A Linguagem dos Cantadores*, de Clóvis Monteiro (1933); *A Língua do Nordeste*, de Mário Marroquim (1934) e os trabalhos de Teixeira - o *Falar Mineiro* (1938) e *Estudos de Dialectologia Portuguesa: A linguagem de Goiás* (1944).

Somam-se a esses trabalhos iniciais também glossários regionais como o *Vocabulário Gaúcho*, de Roque Callage (1926); o *Vocabulário do Nordeste do Rio Grande do Sul: linguagem dos praieiros*, de Dante de Laytano (1933) e o *Vocabulário Pernambucano*, de Francisco Pereira da Costa (1937), além de obras de caráter geral que tratam de questões mais abrangentes da língua portuguesa no Brasil como: *O português do Brasil*, de Renato Mendonça (1937), *O Problema da Língua Brasileira*, de Sílvio Elia (1940), *Estudos de Língua Nacional*, de Arthur Neiva (1940) e o livro *A Língua do Brasil*, de Gladstone Chaves de Melo (1946). Surgem também nesse período, estudos acerca da contribuição do elemento africano à língua portuguesa, como a obra *O Elemento Afro-negro na Língua Portuguesa* (1933), de Jacques Raimundo e *A Influência Africana no Português do Brasil* (1935), de Renato Mendonça.

Mas foi a partir da segunda metade do século XX que os estudos sobre a diversidade geográfica do PB tomaram incremento prodigioso. Remonta ao ano de 1952 a data em que o governo brasileiro publicou o Decreto 30.643 (anexo A) que definia as finalidades da Comissão de Filologia da Casa de Rui Barbosa, dentre as quais a primordial era a elaboração do Atlas linguístico do Brasil. Logicamente que, dentre as muitas dificuldades de colocar em prática tal desiderato, encontrava-se também a falta da mentalidade dialetológica entre os linguistas da época, já preconizada por Silva Neto (1957).

Reconhecida também a dimensão continental do país e os recursos rodoviários da época, em 1957, Celso Cunha e Serafim da Silva Neto, por ocasião do III Colóquio Internacional de Estudos Luso-Brasileiros, realizado em Lisboa, defenderam a necessidade de elaboração de atlas regionais como forma de elaborar o tão almejado atlas nacional.

Junta-se a esses dois linguistas o professor Antenor Nascentes, que, em 1958, publica o primeiro volume das *Bases para Elaboração do Atlas linguístico do Brasil*, no qual o autor define diretrizes gerais que orientariam a execução de um Atlas linguístico do Brasil, tais como o perfil dos informantes, escolha das localidades, estrutura dos instrumentos de coleta de dados e acentua, claramente, a importância da elaboração de atlas regionais.

Esses estudiosos passam a dar uma nova visão ao tratamento dos fenômenos da variação no Brasil, implantando um novo momento para a Dialectologia no país: o início da Geolinguística brasileira, e, como resultado inicial, é publicado, em 1963, o primeiro atlas linguístico de um Estado brasileiro, o *Atlas Prévio dos Falares Baianos (APFB)*, de Nelson Rossi e suas colaboradoras.

Segundo Cardoso (2010), esse momento da história da Dialectologia brasileira tem como marca identificadora o começo dos estudos sistemáticos no campo da Geolinguística, embora não fiquem ausentes desse período a produção de léxicos regionais e glossários, bem como a elaboração de monografias sobre regiões diversas.

2.2.9 Sobre a Periodização da Geolinguística Brasileira

Assim como os estudos dialetais são sistematizados em fases, conforme se observa em Ferreira e Cardoso (1994) e Mota e Cardoso (2006), os estudos geolinguísticos também podem ser classificados em dois momentos de acordo com a metodologia empregada e as características de cada trabalho⁵³.

Segundo Romano (2012, p.69), um primeiro momento da Geolinguística brasileira tem como marco inicial a publicação do primeiro atlas linguístico estadual, o APFB, em 1963, e se estende até 1996 - ano em que o projeto Atlas Linguístico do Brasil deu início às suas atividades com o esforço de dialetólogos de todo o país. Os atlas linguísticos desenvolvidos ou projetados neste período de 33 anos seguem pressupostos metodológicos anteriores aos do ALiB, portanto, divergem quanto à metodologia, ora priorizando a diatopia: Atlas Prévio dos Falares Baianos (ROSSI et al., 1963), Esboço de um Atlas linguístico de Minas Gerais (RIBEIRO et al., 1977), Atlas Linguístico da Paraíba (ARAGÃO; BEZERRA DE MENEZES, 1984), Atlas Linguístico-Etnográfico da Região Sul do Brasil (ALTENHOFEN; KLASSMAN, 2011) e (KOCH; ALTENHOFEN; KLASSMAN, 2011), ora incluindo outras dimensões da variação, como a diagenérica - Atlas Linguístico de Sergipe (FERREIRA et al., 1987) e o Atlas Linguístico do Paraná (AGUILERA,

⁵³ Romano (2012) preconiza esta divisão e apresenta, detalhadamente, as características metodológicas dos atlas linguísticos estaduais, regionais e de pequeno domínio, publicados, em andamento ou projetados, traçando o panorama atual da Geolinguística brasileira. Informações sobre estes atlas linguísticos também podem ser encontradas em Isquierdo (2006) e, em Romano (2013), há uma atualização dos dados, inclusive com um mapa ilustrativo sobre o Estado da arte dos atlas linguísticos estaduais/regionais.

1994). São trabalhos pioneiros no campo da Geolinguística brasileira que revelam o amadurecimento da mentalidade dialetológica, defendida por Silva Neto (1957). Alguns desses atlas foram divulgados e publicados após 1996, porém pertencem ao primeiro momento por suas particularidades metodológicas.

O segundo momento da Geolinguística brasileira vai de 1996 até os dias atuais. Os trabalhos deste período já apresentam influência dos pressupostos metodológicos do ALiB, caminhando para uma visão pluridimensional da variação. Nesse segundo momento, registra-se um incremento prodigioso da elaboração de atlas linguísticos de pequeno domínio, a partir de projetos de Mestrado e Doutorado na área, com o incentivo dos Diretores Científicos do ALiB.

2.2.1 O Novos Rumos da Geolinguística Brasileira: O Projeto ALiB - (1996)

O marco divisório que separa os dois momentos da Geolinguística brasileira é o Projeto ALiB. Assim, concretizada a mentalidade dialetológica entre os estudiosos brasileiros dos primeiros atlas linguísticos, observa-se que, em muitos casos, não há uniformidade metodológica naqueles trabalhos⁵⁴ que possibilite estudos comparativos e contrastivos com vistas a revelar uma visão geral da língua falada no Brasil.

Dessa forma, em novembro de 1996, por ocasião do Seminário *Caminhos e Perspectivas para a Geolinguística no Brasil*, ocorrido em Salvador/BA, dialetólogos de todo o país, assessorados pelo pesquisador francês Prof. Michel Contini, da Universidade de Grenoble, reuniram-se e decidiram assumir o grande desafio de elaborar um atlas nacional, retomando a ideia dos eminentes linguistas brasileiros do início da segunda metade do século XX. Este atlas, adotaria uma metodologia única e revelaria a realidade linguística brasileira. Estavam, pois, iniciadas as atividades do Projeto Atlas Linguístico do Brasil.

Este projeto interinstitucional, sediado na Universidade Federal da Bahia, sob a presidência da Dr^a. Suzana Alice Marcelino Cardoso, e coordenado por um Comitê Nacional, envolve quinze instituições universitárias brasileiras. Tem como

⁵⁴ Por exemplo, o APFB, o EALMG e o ALPB podem ser considerados como atlas monodimensionais, isto é, privilegiam a dimensão diatópica, pois não apresentam na cartografia variáveis sociais. No ALS e o ALPR, porém, a variável gênero recebeu uma atenção especial, não somente na coleta de dados, como também na cartografia, sendo possível identificar as respostas dos homens e das mulheres nas cartas. Trata-se, pois, de atlas bidimensionais.

objetivo principal descrever a variante brasileira da língua portuguesa nos níveis fonético-fonológico, semântico-lexical, morfossintático e pragmático-discursivo sob a perspectiva da Dialetoologia pluridimensional (THUN, 1998).

Para tanto, a equipe do Projeto, no período que compreende os anos de 2001 a 2013⁵⁵, aplicou questionários a brasileiros de duas faixas etárias (18-30 e 50-65 anos), de ambos os sexos, em todo o território nacional (interior e capitais). Em cada localidade do interior entrevistaram-se quatro informantes de nível fundamental e, nas capitais, somam-se a estes quatro informantes de nível superior.

No total, o *corpus* do projeto consta de um volume de gravações que revela características da língua falada no século XXI por 1100 brasileiros. Atualmente, a coleta de dados nos 250 pontos linguísticos selecionados pelo projeto está concluída e o projeto entra em uma nova fase de sua elaboração, a transcrição e revisão desses dados para posterior armazenamento em um banco de dados geral.⁵⁶

2.3 SOBRE A DIALETAÇÃO DO PORTUGUÊS BRASILEIRO: AS PROPOSTAS DE DIVISÃO DIALETAL

Os trabalhos de natureza geolinguística têm por finalidade apresentar um conjunto de dados que comprove ou não a existência de áreas dialetais (léxicas, fônicas, morfossintáticas) em determinado espaço geográfico, uma vez que os dados linguísticos coletados empiricamente são registrados em “mapas especiais” (COSERIU, 1987).

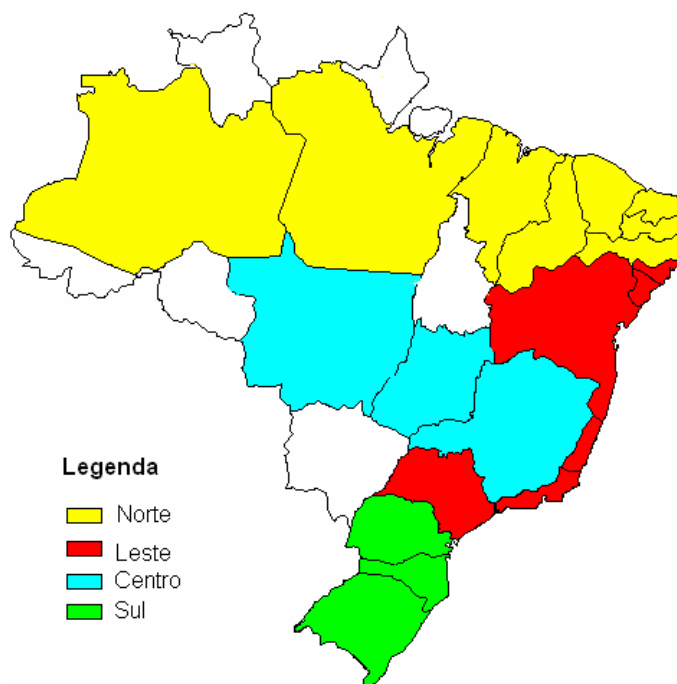
Embora Amaral (1981, p. 44 [1920]) já apontasse a necessidade de trabalhos empíricos com a finalidade de retratar “com segurança quais os caracteres gerais do dialeto brasileiro, ou dos dialetos brasileiros”, as tentativas de divisão e

⁵⁵ O primeiro inquérito para a constituição do *corpus* do ALiB foi realizado no dia 01 jun. de 2001 pela Equipe da Regional Paraná, em Quirinópolis-GO (ponto 126) e o último inquérito foi realizado no dia 18 de set. 2013 pela Equipe da Bahia em Limoeiro-PE e Olinda-PE (pontos 64 e 65).

⁵⁶ Durante o III Congresso de Dialetoologia e Sociolinguística (III CIDS), realizado em Londrina, em homenagem às Profas Suzana Cardoso e Jacyra Mota, em outubro de 2014, ocorreu o lançamento dos dois primeiros volumes do ATLAS LINGUÍSTICO DO BRASIL, publicados pela EDUEL. O volume I é o de Introdução e o Volume II apresenta 159 cartas linguísticas, com dados de 25 capitais de Estados. O próximo volume (vol. III – em elaboração) constará de estudos sobre material cartografado no vol. II. Para mais informações do Projeto ALiB, acesse: <http://twiki.ufba.br/twiki/bin/view/Alib/WebHome>.

sistematização de tais ‘dialetos’ vieram a se consolidar na obra *O linguajar Carioca*, de Antenor Nascentes (1922/1953). Nessa obra, o autor apresenta as primeiras propostas como, por exemplo, a de Júlio Ribeiro (1881), considerada por Nascentes (1953) como “toda ela imperfeita”, pois Ribeiro pautou-se, exclusivamente, no critério geográfico. Na ocasião, Ribeiro dividiu o Brasil em quatro áreas: 1) Norte; 2) Leste; 3) Centro; 4) Sul, conforme se observa na Figura 15.

Figura 15 - Divisão dialetal de Júlio Ribeiro (1881)



Montagem: Valter Romano (2009)
Fonte: Nascentes (1953)

Fonte: Adaptado de Nascentes (1953)

Nascentes (1953) destaca os seguintes defeitos nessa divisão:

Junta o Norte com Nordeste, que é muito diferente dele.
Separa Alagoas dos demais Estados do Nordeste.
Coloca o Espírito Santo (sem discriminar norte e sul) e Rio de Janeiro junto da Baía, tão diferente esta.
Coloca São Paulo, tão caracteristicamente sulino, junto com Alagoas (!) e junto com Sergipe e Baía e junto com Espírito Santo e Rio de Janeiro.
Coloca Minas (sem discriminar *sic*) junto com Goiás e Mato Grosso.
No sul só ha que objetar a falta de S. Paulo (NASCENTES, 1953, p. 21)

Dando continuidade, Nascentes (1953) descreve a proposta de Maximino Maciel (1950) (Figura 16).

Figura 16 - Divisão dialetal de Maximino Maciel (1950)



Fonte: Adaptado de Nascentes (1953); Barbadinho Neto apud Ribeiro (2012)

Segundo Nascentes (1953), essa divisão, além de se pautar, exclusivamente, no critério geográfico, apresenta outros defeitos como:

A língua chamada Guiana Brasileira se estende à região da margem direita do Amazonas; que serão idioletos? A influencia do castelhano platino na língua da fronteira com o Uruguai e com a Argentina não vai ao ponto de dominar um subfalar. (NASCENTES, 1953, p. 21)

Conclui que, sem uma base histórica, não se pode fazer nada sobre o assunto. Nesse sentido, a proposta de divisão do historiador João Ribeiro, constante da obra *História do Brasil*, vem contribuir, pois agrupa o “imenso organismo nacional” (NASCENTES, 1953, p. 23) em cinco grupos locais:

1) o Extremo Norte (a Amazônia, o Maranhão, Piauí, e Ceará); 2) o Norte (Alagoas, Pernambuco, Paraíba e Rio Grande do Norte; 3) o Centro (Sergipe, Baía, Ilhéus e Porto Seguro); 4) o Interior (São Paulo, Minas Gerais, Goiás e Mato Grosso); 5) o Sul (Espírito Santo, Rio de Janeiro, Santa Catarina e Rio Grande do Sul). (NASCENTES, 1953, p. 23)

Nascentes (1953), no entanto, pondera apresentando outra proposta de divisão dialetal, a de Rodolfo Garcia, publicada no *Dicionário de brasileirismos* (1915). Garcia combinou o critério geográfico com o histórico, levando “em conta a continuidade territorial, a facilidade de comunicações terrestres, marítimas, fluviais (naquele tempo, 1915, não havia ainda as aéreas), a homogeneidade ou heterogeneidade de culturas” (NASCENTES, 1953, p. 22). Assim, considerando a distribuição geográfica dos localismos compendiados nos diferentes glossários, Garcia delimitou cinco zonas, conforme se observa na Figura 17⁵⁷.

Figura 17 - Divisão Dialetal de Rodolfo Garcia (1915)



Montagem: Valter Romano(2013)
Fonte: Nascentes (1953)

Fonte: Adaptado de Nascentes (1953)

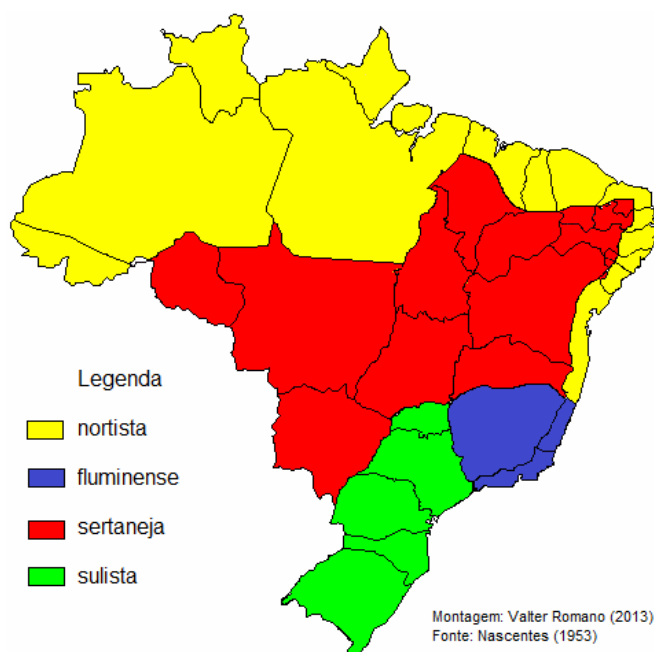
Nascentes (1953) considera positivos os pontos referentes à combinação de critérios históricos e geográficos, porém não deixa de tecer algumas críticas a essa divisão.

⁵⁷ Em 1915, os atuais Estados do Amapá, Roraima, Acre e Rondônia eram Territórios Federais. O Estado do Mato Grosso do Sul foi desmembrado do Mato Grosso em 1977 e o Estado do Tocantins foi criado em 1988. Fonte: <http://www.sogeografia.com.br/Conteudos/Estados/>. Acesso: 06 mai. 2013.

Ha os seguinte (*sic*) defeitos nesta divisão: colocar o Maranhão na zona Norte, quando ele é uma espécie de intermediário entre ela e o Nordeste; colocar o Rio de Janeiro e o sul do Espírito Santo na zona central-marítima; colocar Minas Gerais (sem discriminar) e Goiás junto com Mato Grosso. (NASCENTES, 1953, p. 22)

Considerando a divisão de Rodolfo Garcia, Antenor Nascentes, na primeira edição de *O Linguajar Carioca* (1922), apresenta a sua divisão dialetal do Brasil agrupando os falares brasileiros em quatro áreas: 1) nortista; 2) fluminense; 3) sertaneja; 4) sulista, conforme se observa na Figura 18.

Figura 18 - Divisão dialetal de Antenor Nascentes (1922)



Fonte: Adaptado de Nascentes (1953)

Essa proposta de divisão dialetal recebeu críticas de Lindolfo Gomes e o próprio Nascentes reconheceu a falha, pois, na época, havia percorrido apenas uma pequena parte do nosso território: “Conhecia Minas, Espírito Santo, Estado do Rio, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Recife e Baía” (NASCENTES, 1953, p. 23).

Atendendo às ponderações de Gomes, publicou, em 1933, o mapa dialetológico do Brasil com algumas alterações. Segundo Nascentes (1953), “Criei uma zona Norte para Amazonas e Pará, constituí uma zona Nordeste com o litoral, do Maranhão ao norte do Espírito Santo, passei o Sul de Minas para o subdialeto sulista” (NASCENTES, 1953, p. 23-24).

Porém, a reformulação do mapa, em 1933, também foi criticada, dessa vez por um ex-aluno de Nascentes, Renato Mendonça. Na obra *Influência africana no português do Brasil*, Mendonça (1935) reforça que o equívoco de tal divisão consistia em fazer entrar no subdialeto sulista regiões dialetologicamente diversas como São Paulo e Rio Grande do Sul. Para Mendonça, a linguagem da Amazônia distingue-se nitidamente de Pernambuco ou da Baía, todos enfeixados no subdialeto nortista (NASCENTES, 1953, p.24).

Nascentes, por sua vez, refuta as críticas de Mendonça, afirmando:

O Rio Grande do Sul, com que Rodolfo Garcia quase constituiu uma zona à parte, é muito típico, por vários motivos, mas não a ponto de destacar-se das linhas fundamentais do subfalar sulista.

Por conseguinte, está muito bem, junto com São Paulo.

Na ânsia de censurar, Mendonça esqueceu-se de que o erro de colocar a Amazonia junta com o litoral do Maranhão à Baía, cometido na divisão de 1922, já tinha sido sanado em 1933, na pág. 244 da quarta série do Idioma Nacional por ele citada na pág. 202 de *A influência africana*.

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)

Dessa forma, na segunda edição de *O linguajar carioca* (1953), Nascentes apresenta um novo mapa dialetológico do Brasil. Com base em critérios prosódicos e fonéticos, a cadência e a abertura das vogais médias ([e] / [o]) em posição pretônica, como em *pecado/p[E]cado* e *coração/c[]ração*, o dialetólogo divide o português brasileiro em seis subfalares compreendidos em dois grandes grupos, o do Norte e o do Sul, conforme se observa na Figura 19.

Figura 19 - Divisão dialetal de Antenor Nascentes (1953)



Fonte: Nascentes (1953)

Ribeiro (2012) estudou o falar baiano no nível lexical com base nos dados coletados para o ALiB. Assim, verificou a ocorrência de áreas lexicais nessa faixa territorial e, dentre outras conclusões, constatou que o falar baiano não se limita à faixa territorial estabelecida por Nascentes, estendendo-se também a outras áreas denominadas pela autora como área de controle.

Mota (2006), ao fazer uma retrospectiva das divisões dialetais, ressalta que até hoje não foi proposta uma nova divisão, mesmo havendo um substancial avanço dos estudos dialetológicos e geolinguísticos. Isso se deve, sobretudo, à dificuldade de intercomparação entre os dados publicados nos atlas estaduais e regionais, seja pela adoção de metodologias distintas, seja pela distância temporal que se instituiu entre as publicações (APFB – 1963 ao ALERS – 2011).

Para Mota (2006), atualmente, faz-se necessário um 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, os dados coletados para o Atlas

Linguístico do Brasil podem contribuir para uma nova proposta de divisão dialetal cientificamente justificável.

2.3.1 Para Uma Divisão Dialetal do Português Brasileiro a partir dos dados do ALiB

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 desenvolvidos com base no *corpus* do Projeto, sob diferentes enfoques, contemplando os níveis fonético-fonológico, semântico-lexical e morfossintático. Estes trabalhos contribuem para uma descrição mais adequada de traços linguísticos que caracterizam e identificam as regiões brasileiras. Nesse sentido, apresenta-se, nas subseções seguintes, um levantamento descritivo de alguns dos trabalhos desenvolvidos sob a perspectiva do léxico, com foco especial às conclusões obtidas acerca da definição de áreas lexicais do PB.

2.3.1.1 Contribuições do projeto ALiB: perspectiva lexical

Dentre outros trabalhos que focalizam o léxico do PB com base no *corpus* do ALiB, foram selecionados três artigos: o de Romano e Aguilera (2009), Aguilera (2010), Romano e Seabra (2014a); cinco dissertações de mestrado: Yida (2011), Freitas-Marins (2012), Benke (2012), Costa (2012), Portilho (2013); e uma tese de doutorado, Ribeiro (2012). Esta seleção baseou-se na convergência de aspectos analisados nesses trabalhos que contribuem para a discussão dos dados analisados na presente tese.

2.3.1.1.1 – Romano e Aguilera (2009)

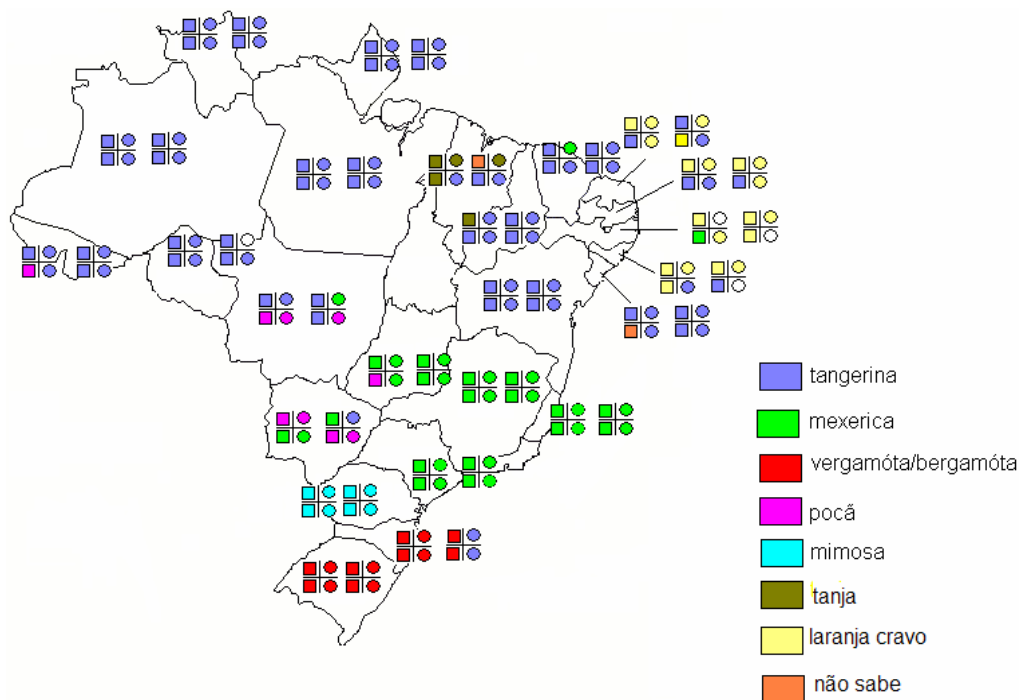
Um dos primeiros trabalhos desenvolvidos a partir dos dados coletados para o ALiB foi o projeto de Iniciação Científica de Romano⁵⁸, no qual foram descritas e analisadas as variantes lexicais para a questão 039 do QSL (*frutas*

⁵⁸ Trata-se do projeto desenvolvido com bolsa da Fundação Araucária (2007-2008) intitulado *Interferência do rural sobre a fala urbana*, sob orientação da diretora científica do ALiB Dr^a. Vanderci de Andrade Aguilera.

menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro na mão?). Dessa forma, Romano; Aguilera (2009) analisaram os dados de 24 capitais brasileiras, sumarizando, na época, a fala de 188 informantes⁵⁹. Sob a perspectiva da Geolinguística, os autores concluem que as variantes lexicais para *tangerina* podem ser estudadas sob duplo aspecto: (a) da distribuição diatópica na formação de zonas de isoléxicas; (b) da produtividade de variantes na fala do mesmo informante, que tanto se refere à forma predominante na localidade, como é capaz de distinguir as variantes de outras regiões (ROMANO; AGUILERA, 2009).

Os dados apontaram que as variantes *tangerina* e *mexerica* são as formas mais difundidas no *corpus* investigado. Em linhas gerais, a variante *tangerina* predomina em capitais da região Norte do país, ao passo que *mexerica* encontra-se mais frequentemente nas capitais do Sudeste e Centro-Oeste. Os autores observam também que foram documentadas outras variantes dependendo da região administrativa. Assim, a variante *laranja-cravo* é uma forma lexical comum nas capitais nordestinas, *bergamota/vergamota* nas capitais do Sul e a variante *mimosa* ocorre, exclusivamente, em Curitiba. Representa-se, na Figura 20, a carta linguística experimental elaborada exclusivamente para aquele trabalho.

Figura 20 - Carta linguística para questão 039 do QSL- Capitais



Fonte: Banco de dados do Projeto ALiB (2008) In: Romano e Aguilera (2009)

⁵⁹ Na ocasião, os autores não tiveram acesso aos dados de todos os informantes que compõem a rede de pontos das capitais.

Apesar das limitações do estudo, bem como das lacunas observadas na representação cartográfica e em parte das análises, encontra-se em Romano e Aguilera (2009) uma primeira tentativa de cartografia dos dados das capitais. Sem entrar em méritos da representação e de “leitura da carta”, observa-se a existência de diferentes formas lexicais dependendo da região que, de certo modo, ratifica, em parte, a divisão dialetal de Nascentes (1953) sobre o grupo do Norte e do Sul, sob a perspectiva do léxico. O Norte, no *corpus* investigado, está representado pelo predomínio da variante *tangerina*, o Sul, pelo uso de *mexerica* e outras formas.

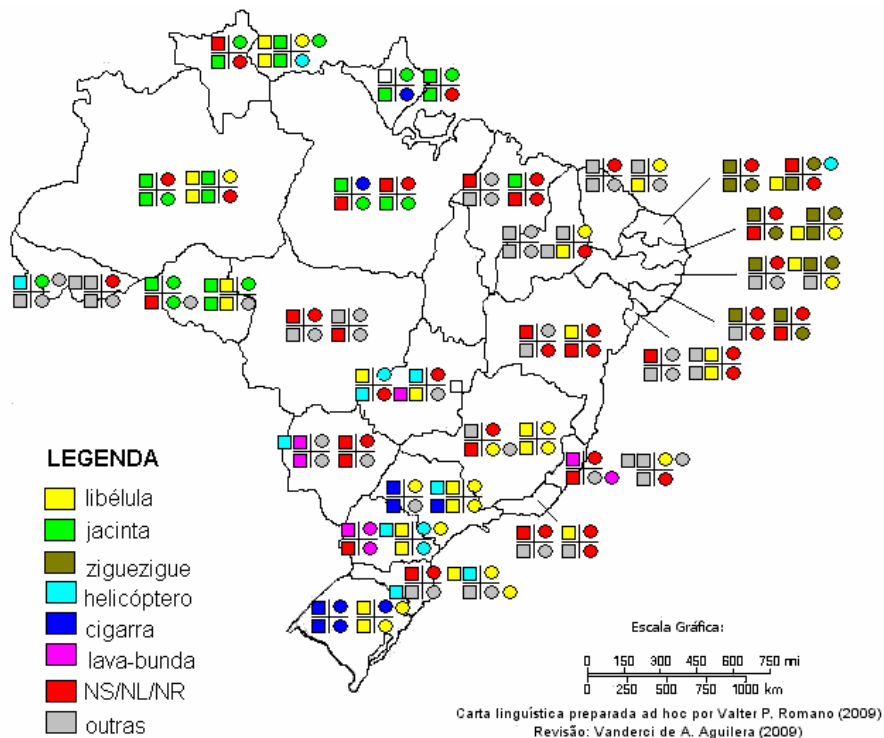
No grupo do Norte, a partir dos dados das capitais, encontra-se também o subfalar Amazônico, caracterizado pelo uso de *tangerina* que se distingue do subfalar nordestino, no qual ocorre a forma *laranja-cravo*. Em meio a esses dois subfalares encontram-se os Estados do Maranhão e Piauí onde ocorreu a variante *tanja*. No falar do Sul, verificam-se, pelo menos, quatro variantes: *mexerica*, *pocã (sic.)*, *mimosa* e *vergamota/bergamota*. Há de se observar também a generalização dos autores ao afirmar o traçado de isoléxicas a partir dos dados das capitais, uma vez que elas não representam um território contínuo. Somente com os dados do interior associados aos das capitais será possível afirmar com mais propriedade a existência dessas áreas de isoléxicas apontadas pelos autores.

2.3.1.1.2 Aguilera (2010)

Em um artigo publicado em 2010, Aguilera também analisou dados do Projeto ALiB referentes a 25 capitais brasileiras. Com base nas respostas de 200 informantes, a autora discutiu aspectos da distribuição diatópica e lexicográficos das variantes obtidas para a questão 085 do QSL (*Como se chama o inseto de corpo comprido e fino, com quatro asas bem transparentes, que voa e bate a parte traseira na água?*). Para esta questão, observa Aguilera (2010, p.293) o alto índice de abstenção de resposta (27%), bem como o acentuado polimorfismo entre os informantes que atribuíram um designativo para o referente em pauta. Foram documentadas 30 variantes lexicais, em sua maioria, não dicionarizadas nas obras lexicográficas consultadas. Quanto à distribuição diatópica, a autora apresenta uma carta linguística experimental, reproduzida na Figura 21, o que possibilita algumas inferências.

Verificam-se, novamente, as diferenças linguísticas entre as capitais da região Norte em relação às capitais do Nordeste. Nestas, dentre outras formas, observa-se a ocorrência da variante onomatopaica *ziguezigue*, ao passo que nas capitais nortistas predomina o uso da variante de étimo indígena, *jacinta*. Na porção centro-sul do país, além da preferência pela variante padrão, *libélula*, principalmente entre os informantes de nível superior, ocorrem outras formas como, por exemplo, *helicóptero*, *lava-bunda* e *cigarra*. Há de se observar, portanto, diferenças linguísticas dentro do território denominado por Nascentes (1953) como subfalar sulista. Dessa forma, um estudo sobre os designativos para *libélula* na rede de pontos do interior localizada na região do subfalar sulista poderá evidenciar áreas convergentes e divergentes quanto à realização dessas e de outras variantes lexicais. Observa-se, pelo estudo de Aguilera (2010), que a variante *cigarra*, por exemplo, foi produtiva em Porto Alegre e em São Paulo, obtendo ocorrência única em Belém e Macapá. Seria *cigarra* uma forma regional que, por conta dos processos migratórios contemporâneos, se encontra em processo de difusão pelo Brasil afora? A este e a outros possíveis questionamentos é que esta tese pretende responder.

Figura 21 - Carta linguística para questão 085 do QSL- Capitais

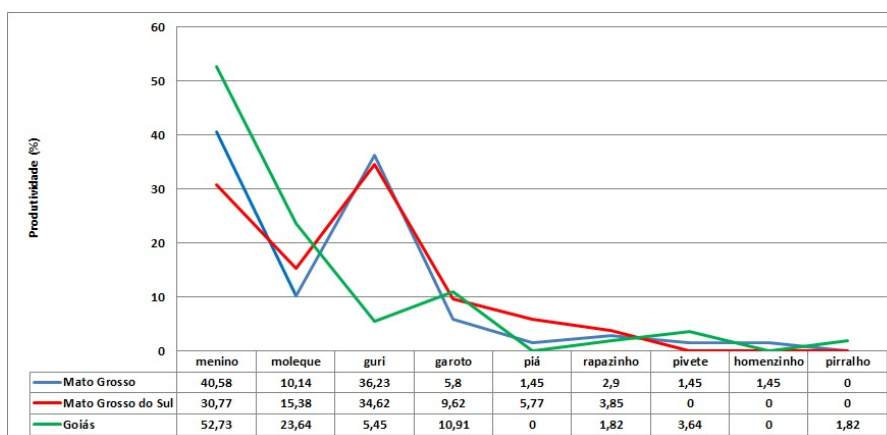


Fonte: Banco de dados do Projeto ALiB (2009). In: Aguilera (2010, p. 301)

2.3.1.1.3 Romano e Seabra (2014a)

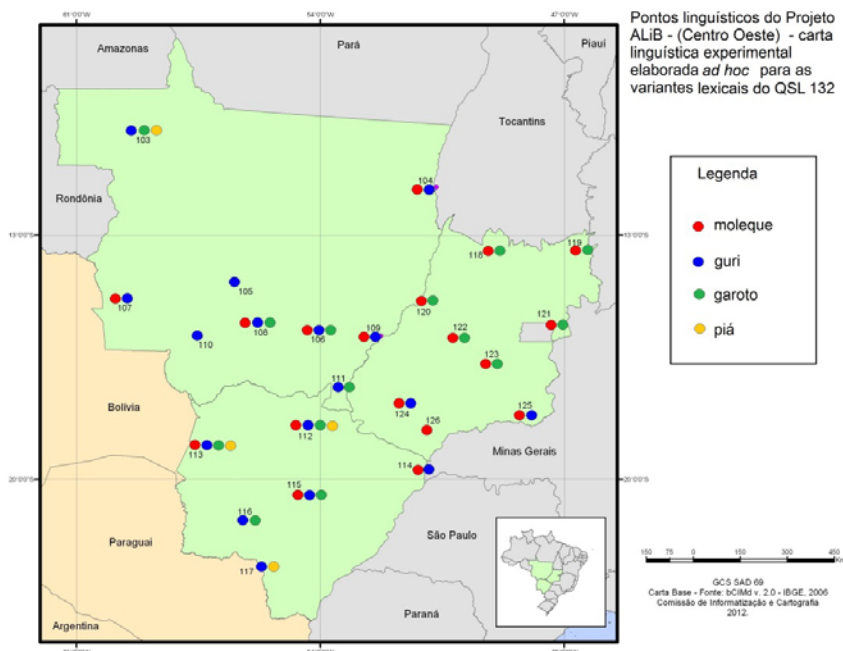
Os dados do recente estudo de Romano e Seabra (2014a) evidenciam diferenças linguísticas nas regiões Centro-Oeste, Sudeste e Sul do Brasil. A partir dos dados coletados pelas equipes do Projeto ALiB em 10 Estados brasileiros, sob uma perspectiva estatística, os autores discutem a distribuição diatópica das variantes lexicais para a questão 132 - *Criança pequenininha, a gente diz que é bebê. E quando ela tem de 5 a 10 anos, do sexo masculino?* – do Questionário Semântico-Lexical (COMITÊ NACIONAL DO PROJETO ALiB, 2001). Buscando respaldo teórico em obras lexicográficas e na história social das regiões estudadas, o estudo comprovou diferenças linguísticas entre os Estados e as regiões investigadas. A variante mais difundida na norma lexical do PB para a questão analisada, considerando o *corpus* investigado, é *menino*, registrada em quase todos os pontos linguísticos. Os autores apresentaram os resultados obtidos por meio de gráficos e cartas linguísticas, considerando-se as regiões administrativas (Figuras 22, 23, 24, 25, 26 e 27).

Figura 22 - Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Centro-Oeste



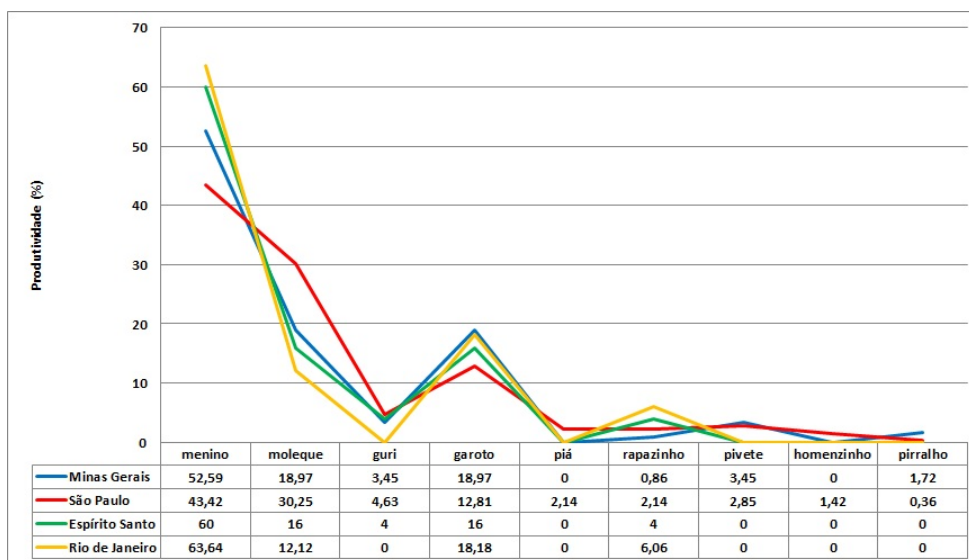
Fonte: Banco de dados do Projeto ALiB. In: Romano e Seabra (2014a)

Figura 23 - Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Centro-Oeste



Fonte: Banco de dados do Projeto ALiB (2013). In: Romano e Seabra (2014a)

Figura 24 - Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Sudeste



Fonte: Banco de dados do Projeto ALiB (2013). In: Romano e Seabra (2014a)

Figura 25 - Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Sudeste

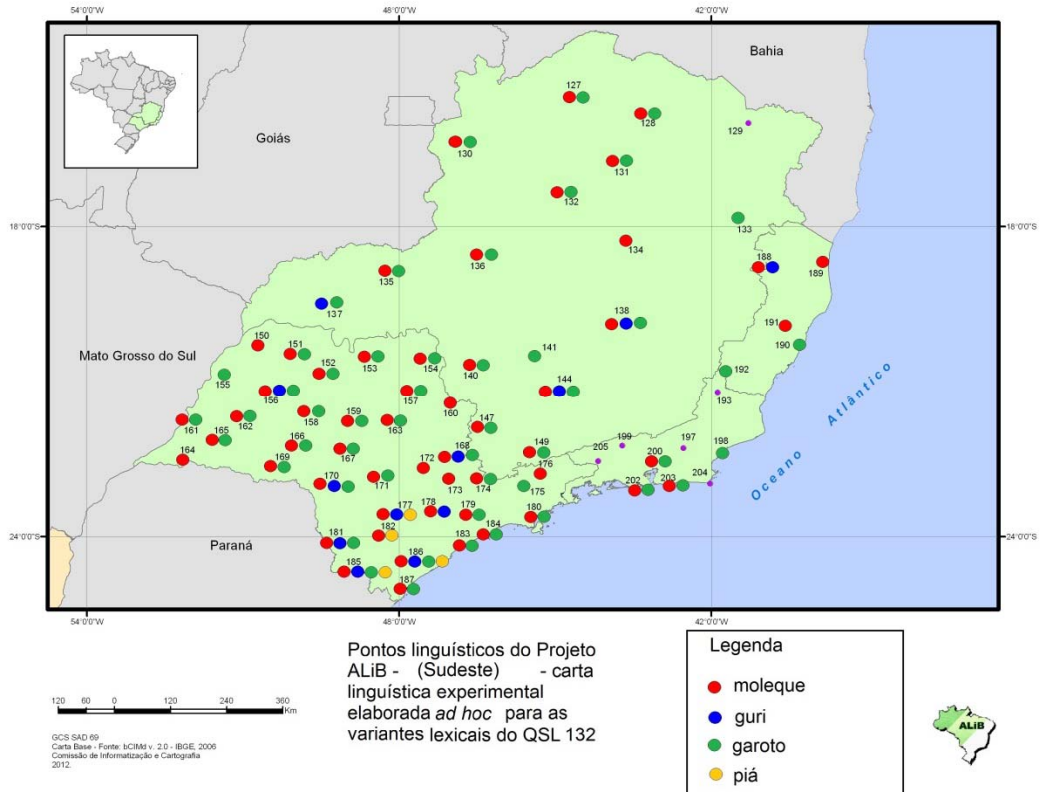


Figura 26 - Produtividade das variantes lexicais para a questão 132 do QSL nos Estados da região Sul

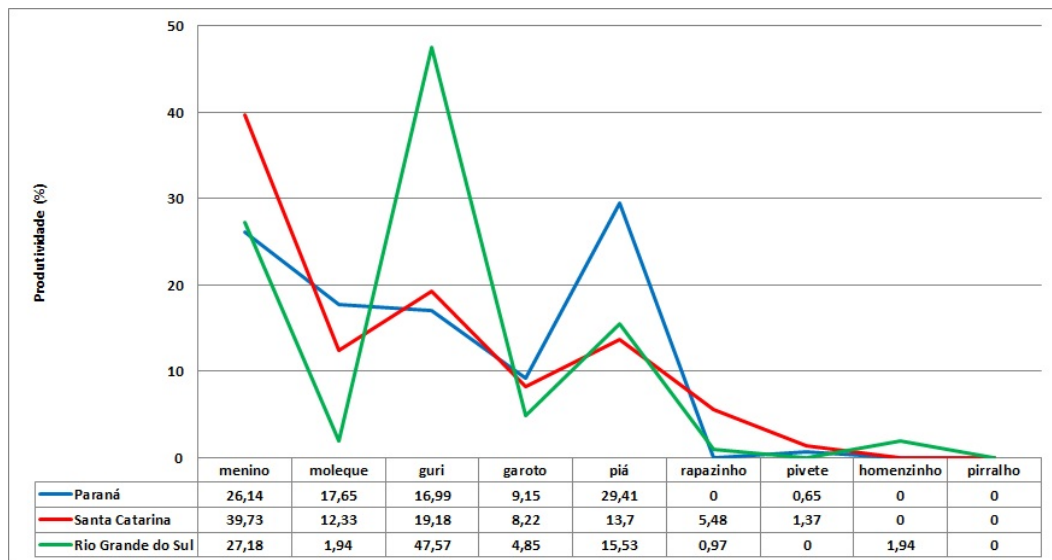
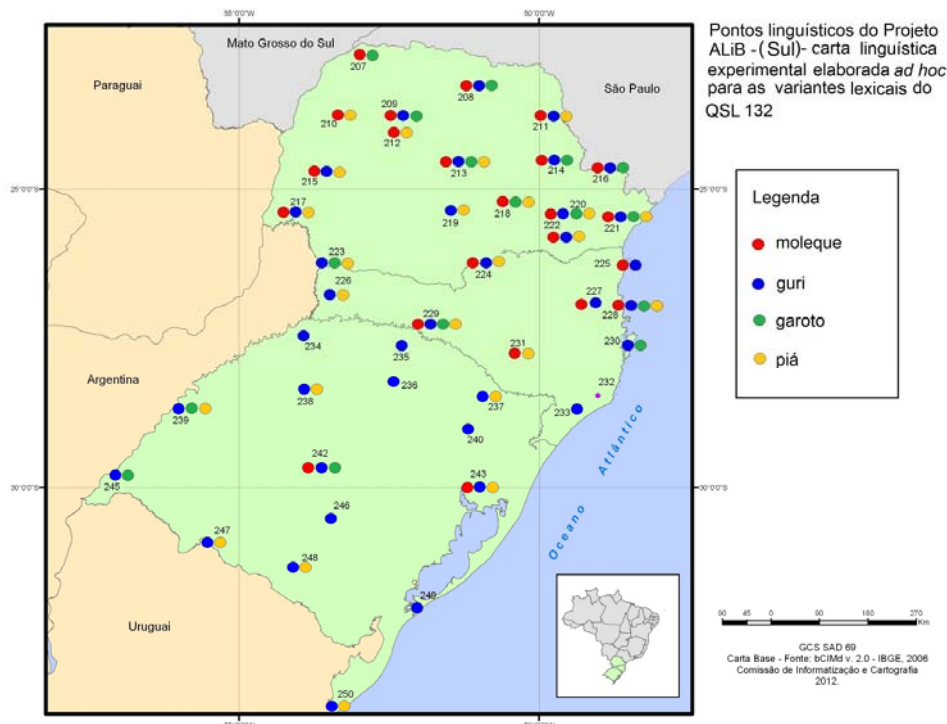


Figura 27 - Distribuição diatópica das quatro variantes lexicais mais produtivas para a questão 132 do QSL na rede de pontos da região Sul



Fonte: Banco de dados do Projeto ALiB (2013).In: Romano e Seabra (2014a)

A pesquisa evidenciou que, no Centro-Oeste, há diferenças linguísticas entre Mato Grosso do Sul e Mato Grosso em relação ao Estado de Goiás. Nos dois primeiros, além da predominância da variante *menino*, há um índice considerável de registros da variante *guri*, típica da região Sul, o que ressalta a semelhança desses Estados em alguns aspectos com o Rio Grande do Sul, provavelmente devido a fluxos migratórios ocorridos nos últimos 50 anos. Em Goiás, por sua vez, o índice de ocorrência da variante *guri* é menor, haja vista que esse Estado recebeu influências também de outros grupos, além dos bandeirantes paulistas durante o seu processo de ocupação.

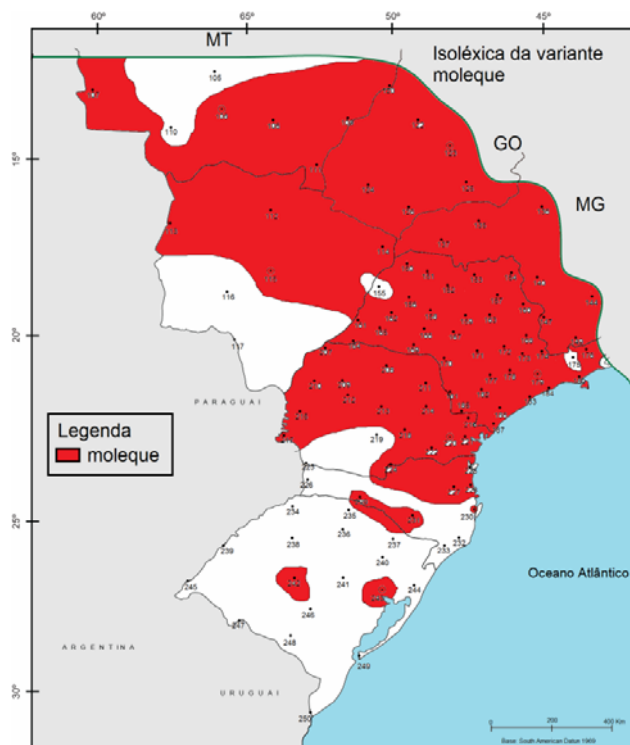
A variante *moleque* difunde-se pelas três regiões, como a segunda variante mais produtiva. Essa forma lexical está presente, sobretudo, nos Estados em que houve a intensa atuação dos bandeirantes durante o período de escravização do índio e, principalmente, do negro nas áreas mineradoras. Essa variante de etimologia africana não obteve índice de ocorrência considerável no Rio Grande do Sul.

Na região Sul, há três variantes que se destacam: (i) *guri* é a forma mais difundida e mais produtiva no Rio Grande do Sul; (ii) em Santa Catarina, há maior índice de ocorrência da variante *menino*; e (iii) *piá* é a forma lexical mais produtiva no Estado do Paraná. Neste Estado, a ocorrência dessa variante concentra-se em municípios do centro e sul, não ocorrendo em cidades do norte e leste paranaense. A não ocorrência de *piá* nesta área justifica-se por sua proximidade linguística com o Estado de São Paulo.

Por fim, no Sudeste, além de outras formas, distribui-se de forma descontínua a variante *guri* em cidades mineiras e capixabas, não possibilitando o traçado de uma linha de isoléxica. No entanto, a ocorrência dessa variante juntamente com a forma lexical *piá* revela uma área de isoléxica em localidades paulistas situadas no Vale do Ribeira. A presença dessas variantes nessa região do Estado, principalmente, em Itapetininga, pode evidenciar traços da sócio-história, pois sabe-se que o Caminho das Tropas que vinha do Rio Grande do Sul tinha como ponto de passagem algumas cidades paulistas antes de alcançar o destino final, Sorocaba. Portanto, o estudo ratifica a importância de associar aspectos sócio-históricos a dados linguísticos, contribuindo, desse modo, para a melhor compreensão e descrição do PB.

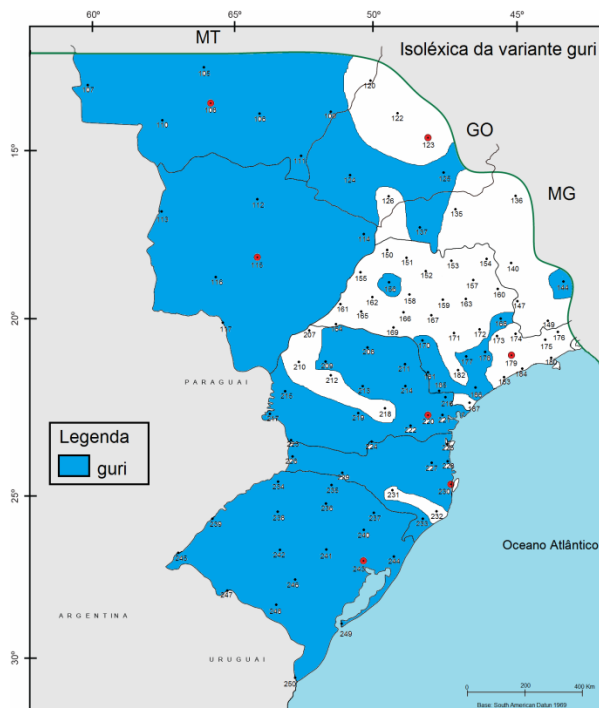
Com base nos resultados encontrados no estudo de Romano e Seabra (2014a), foram elaboradas, especificamente para este trabalho, três cartas de isoléxicas (experimentais) referentes à área geográfica do subfalar sulista (Figuras 28, 29 e 30).

Figura 28 - Isoléxica da variante moleque na rede de pontos do subfalar sulista



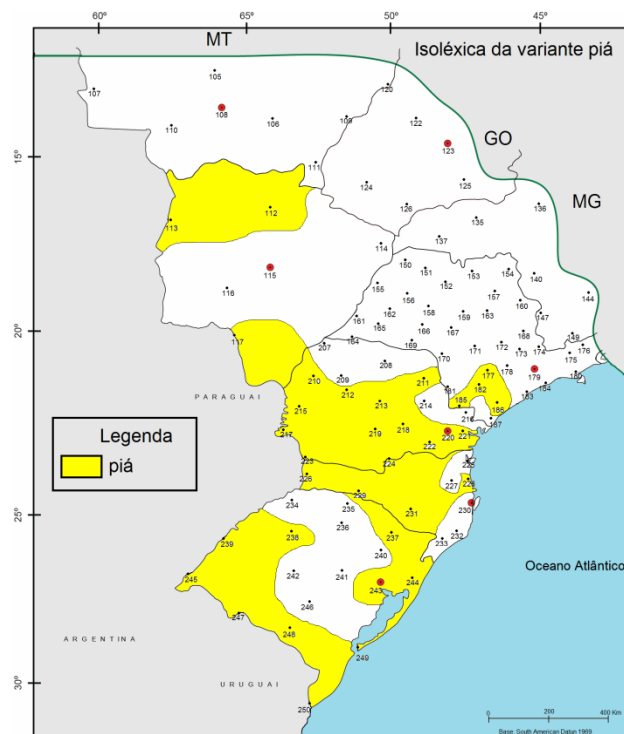
Fonte: Banco de dados do ALiB (2013) – adaptado de Romano e Seabra (2014a)

Figura 29 - Isoléxica da variante guri na rede de pontos do subfalar sulista



Fonte: Banco de dados do ALiB (2013) - adaptado de Romano e Seabra (2014a)

Figura 30 - Isoléxica da variante piá na rede de pontos do subfalar sulista



Fonte: Banco de dados do ALiB (2013) – adaptado de Romano e Seabra (2014a)

Um estudo sistemático que documente a variação lexical na área do subfalar sulista permite verificar as semelhanças lexicais de Estados como Rio Grande do Sul, Mato Grosso e Mato Grosso do Sul, conforme atestam as conclusões obtidas a partir das Figuras 28, 29 e 30.

Em contrapartida, o estudo permite revelar também as diferenças do Estado de São Paulo, por exemplo, em relação ao Rio Grande do Sul, conforme os dados das Figuras 28 e 29. Na Figura 28, a variante de etimologia africana (*moleque*) não é significativa no Rio Grande do Sul, onde se verifica a presença da variante de etimologia indígena *guri*.

A ocorrência de *guri*, no Estado de São Paulo, restringe-se a uma área geográfica que contempla algumas cidades que pertenciam ao antigo *Caminho das tropas* (pontos 177 – Itapetininga, 178 - Sorocaba e 185 – Ribeira), ou seja, permite o traçado de uma área de isoléxica, naquele Estado, que revela traços dos processos migratórios. A pequena ocorrência de *guri* no Estado de São Paulo, em contrapartida, é compensada pela maciça distribuição da forma lexical *moleque*, que adentra o sul de Minas, sul de Goiás, norte do Mato Grosso do Sul e sul do Mato Grosso, região em que escravos africanos foram amplamente utilizados como mão de obra durante o período da mineração, conforme explicitado ao longo desta tese.

A variante *piá*, por sua vez, apresenta ampla distribuição no Paraná Tradicional, bem como em Santa Catarina (Figura 30). A não ocorrência dessa variante, de etimologia indígena, na área limítrofe entre o Paraná e São Paulo revela a proximidade linguística daquela área com este Estado. A presença de *piá*, no Estado de São Paulo, por sua vez, se concentra na faixa territorial do Vale do Ribeira, ocorrendo em cidades como Ribeira, Capão Bonito, Registro e Itapetininga.

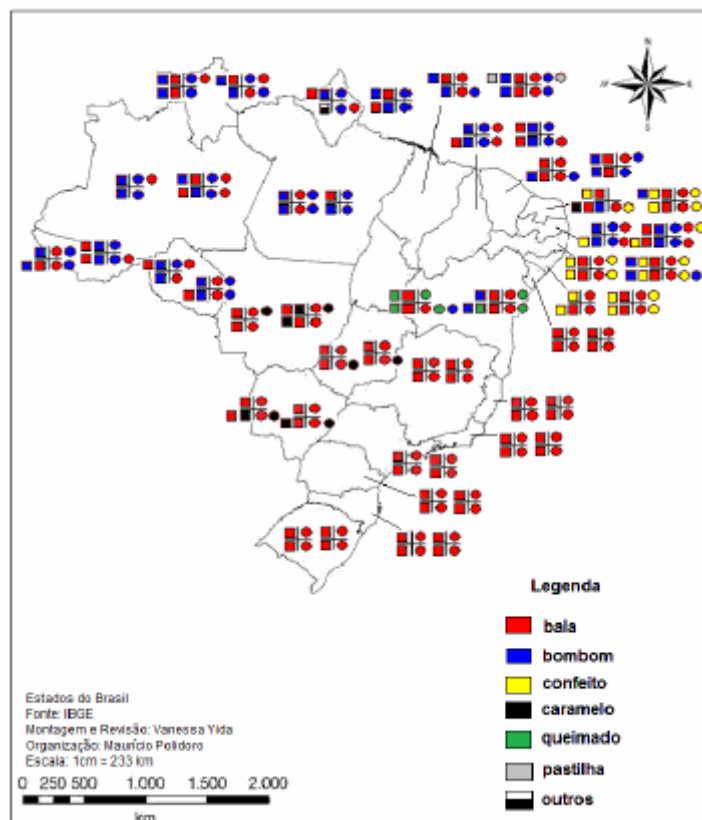
Com base em aspectos sócio-históricos, é possível entender a motivação que propicia a formação dessa área lexical, ou seja, os municípios em que foram registradas ocorrências dessa variante faziam parte das rotas dos Tropeiros. Assim, esse importante movimento de ocupação do território brasileiro deixou marcas que podem ser observadas não somente na formação das cidades originadas a partir dos antigos ranchos, mas também em aspectos culturais e, sobretudo, linguísticos.

2.3.1.1.4 Yida (2011)

Yida (2011), em sua dissertação de mestrado defendida na Universidade Estadual de Londrina, trabalhou com o *corpus* do Projeto ALiB referente a 25 capitais, apresentando uma análise da área semântica *Alimentação e Cozinha* do QSL. Dessa forma, analisou os dados de 200 informantes pautando-se em variáveis extralinguísticas, tais como sexo, faixa etária, escolaridade e ponto linguístico. O material analisado versou sobre os designativos de 12 questões, a saber: 176 - *café da manhã*; 177 - *geleia*; 178 - *carne moída*; 179 - *curau/canjica*; 180 - *curau*; 181 - *mungunzá/canjica*; 182 - *aguardente*; 183 - *empanturrado*; 184 - *glutão*; 185 - *bala/confeito/bombom*; 186 - *pão francês*; 187 - *pão bengala*. Buscando respaldo teórico na Etnolinguística e na Dialetologia pluridimensional (THUN, 1998), a autora interpretou os dados que evidenciaram características próprias de cada localidade, bem como aspectos típicos do perfil dos informantes.

Algumas das análises convergem para as propostas de Nascentes (1953) sobre a divisão dialetal do PB, no que se refere ao grupo do Norte e do Sul. A Figura 31 apresenta uma carta linguística experimental elaborada pela autora, por meio da qual são representados os designativos de *bala* – questão 185.

Figura 31 - Carta linguística para os designativos de *bala* – QSL 185

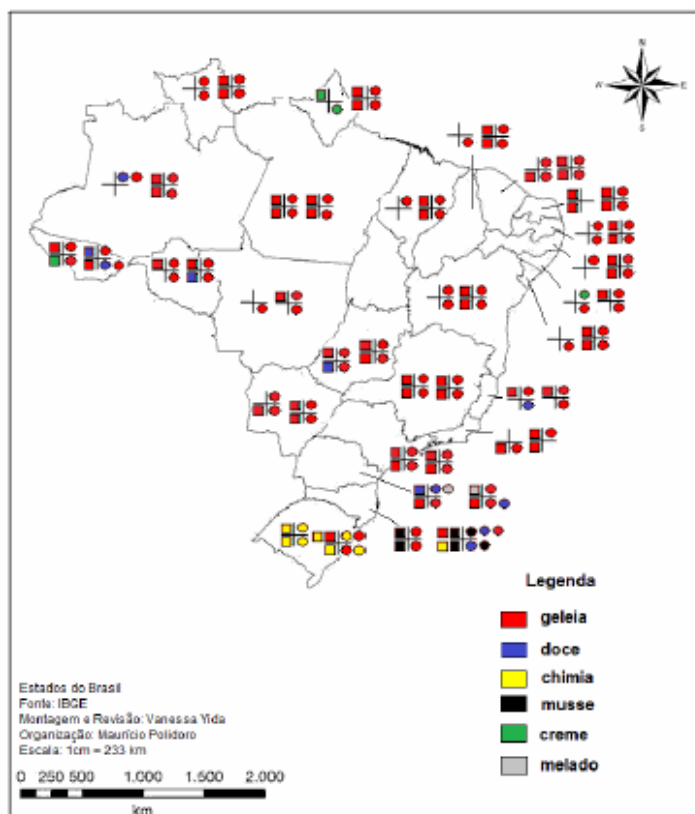


Fonte: Banco de dados do ALiB (2011). In: Yida (2011, p. 158)

Ficam nítidas, nesta carta, as diferenças entre as capitais nortistas, com o predomínio da variante *bombom* diferentemente quatro capitais nordestinas, onde ocorre a variante *confeito*. Salvador foi a única capital em que se encontrou a forma lexical *queimado*, ao passo que no Centro-Oeste documentou-se a variante *caramelo*. Desse modo, nota-se a diferença entre o grupo do Norte e Sul, e naquele a presença dos dois subfalares, o Amazônico e o Nordeste.

Outra questão interessante analisada por Yida (2011) é a de nº 177 do QSL (*Como se chamam aqui a pasta feita de frutas feita para passar no pão, biscoito*). Na Figura 32, visualizam-se as variantes documentadas para esta questão junto às 25 capitais.

Figura 32 - Carta linguística para os designativos de *geleia* – QSL 177



Fonte: Banco de dados do ALiB (2011). In: Yida (2011, p. 102)

O destaque aqui é para a ocorrência da variante *chimia*, registrada em Florianópolis e em Porto Alegre. De acordo com as obras lexicográficas consultadas pela autora, o vocábulo vem do alemão *schimier* que, na variante usada pelos informantes do Projeto ALiB, passou por um processo de adequação fonética resultando no vocábulo *chimia*. Este caso, em particular, revela traços que caracterizam a influência de imigrantes germânicos no vocabulário de usuários desta região do país, podendo, inclusive, ser classificado como um regionalismo.

Dentre as outras questões trabalhadas pela autora, surgem algumas indagações no que se refere à área correspondente ao subfalar sulista. Qual seria a abrangência da variante *caramelo*, como um designativo de *bala*, na rede de pontos do interior? Qual seria o grau de alcance da variante *chimia* na região Sul? Estas duas variantes, *chimia* e *caramelo*, podem ou não evidenciar áreas lexicais no território do subfalar sulista? A estas e a outras possíveis perguntas é que esta tese se propõe a responder.

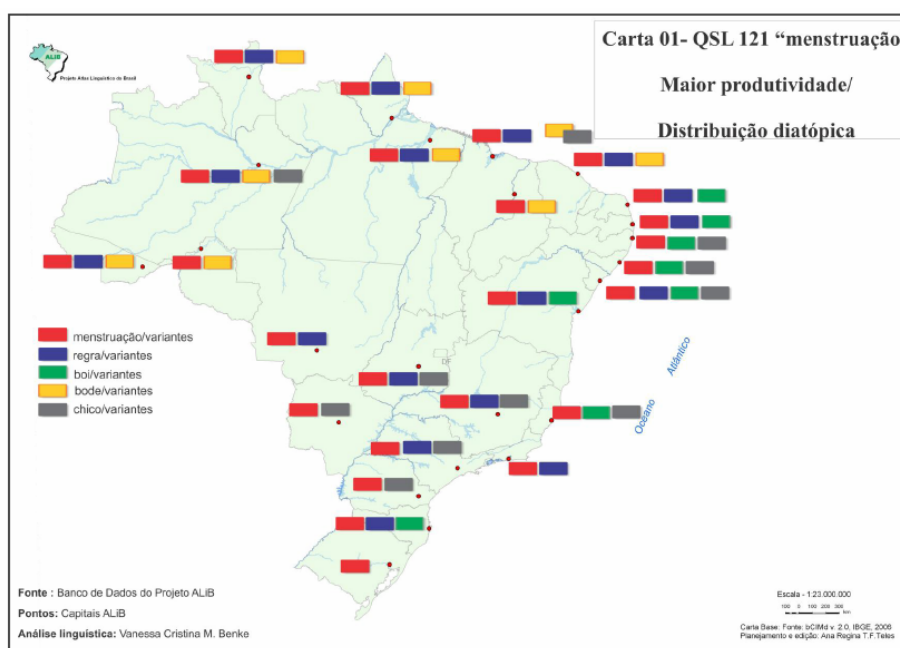
2.3.1.1.5 Benke (2012)

O trabalho de Vanessa Cristina Martins Benke, apresentado a título de mestrado na Universidade Federal do Mato Grosso do Sul, discute questões do QSL que, frequentemente, geram variantes lexicais tabuizadas. Associando o perfil dos informantes às variantes documentadas, Benke (2012), com os dados do ALiB referentes às 25 capitais brasileiras, apresentou uma análise léxico-semântica dos designativos para cinco questões do QSL, a saber: 121 – *menstruação*; 137 – *pessoa pouco inteligente*; 141 – *marido traído*; 142 – *prostituta* e 147 – *diabo*.

Estas questões apresentaram polimorfismo acentuado e, dentre outras conclusões, a autora afirma que “do ponto de vista diatópico, o *corpus* examinado mostrou-se bastante relevante pela disseminação característica de uma determinada região para outras, afetada, por sua vez, por condicionantes histórico-geográficos” (BENKE, 2012, p.234). A autora ilustra esta afirmação com as variantes da questão 121 – *menstruação* (Figura 33). Assim, o uso das variantes *de boi* e *de bode vai*

além dos limites geográficos impostos por sua marca dialetal do Nordeste já dicionarizada, tendo sido evidenciada também nas regiões Norte, Sul e Sudeste em virtude da grande leva de nordestinos que se deslocaram para essas regiões do Brasil. (BENKE, 2012, 234)

Figura 33 - Carta linguística para os designativos de *menstruação* – QSL 121



Fonte: Banco de dados do ALiB (2012). In: Benke (2012, p. 103)

Esta carta apresentada por Benke (2012, p. 103) mais uma vez ratifica o prognóstico de Nascentes (1953) acerca da divisão dialetal do grupo do Norte em dois subfalares, o Nordeste e o Amazônico. Assim, *bode* e suas variantes ocorrem, principalmente, nas capitais nortistas, diferentemente da região Nordeste em que, na maioria das capitais, foi registrada a forma *boi*. O equívoco da autora, no tocante ao emprego destas variantes, está em afirmar que a “concentração da unidade lexical *bode* nas capitais do Norte e Nordeste, sobretudo nas do Norte” evidencia “uma isoglossa no âmbito do vocabulário em exame” (BENKE, 2012, p. 108). Ora, tal afirmação poderia ser confirmada ou não se a análise englobasse também os dados do interior de cada Estado, uma vez que as capitais apresentam relativa distância geográfica que as separam, podendo, pois, a variante *bode* não apresentar uma distribuição diatópica contínua na rede de pontos do interior. Não obstante, traçar isoglossas (isoléxica no caso) com base apenas em dados das capitais seria um posicionamento precipitado, como afirmam Romano e Aguilera (2009).

Na análise de outra questão (Figura 34), a autora apresenta mais dados que evidenciam a dinâmica da população e, conseqüentemente, linguística. De acordo com a autora,

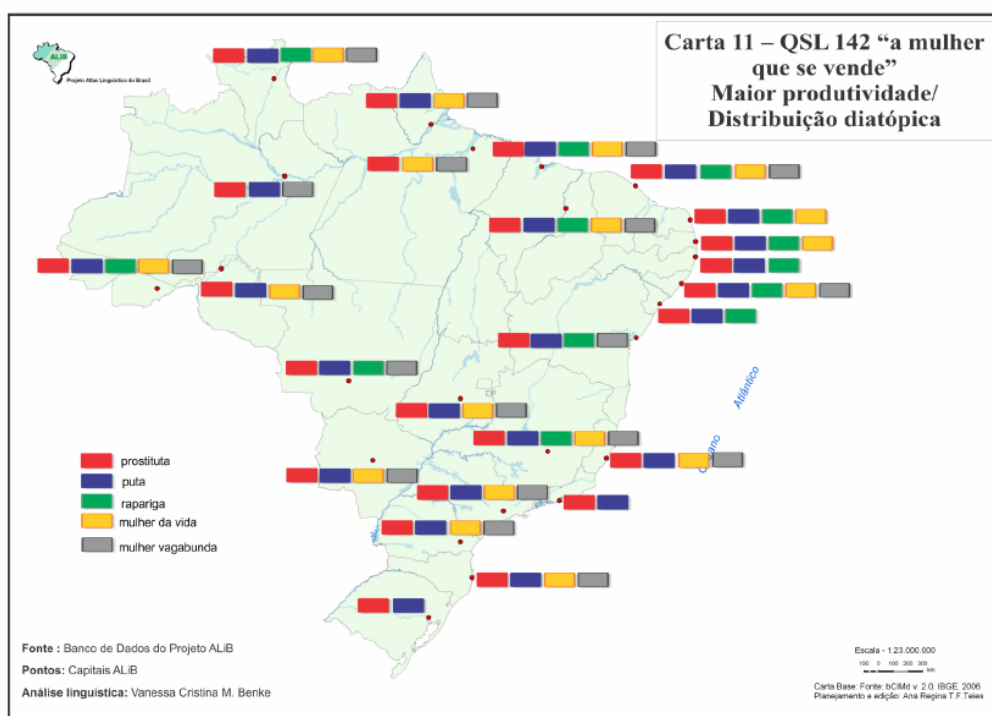
outro exemplo de propagação linguística foi identificado com a unidade lexical *biscate*, designativo para “prostituta”, a que os lexicógrafos consultados atribuem a marca dialetal de São Paulo, ao passo que os dados do Projeto ALiB evidenciam esse item lexical em Campo Grande, em Cuiabá e em Curitiba, o que a descaracterizou como um dialetismo de São Paulo, que possivelmente, em razão de ser um Estado limítrofe com os três Estados, facilitou essa disseminação. (BENKE, 2012, p. 234-235)

Por meio da Figura 34 e com base nas observações de Benke (2012), podem-se extrair algumas conclusões ao relacionar os dados cartografados à divisão dialetal do Brasil proposta por Nascentes (1953) no que diz respeito às duas grandes divisões, o grupo do Norte e o grupo do Sul.

Verifica-se, assim, que a variante *rpariga* predominou em capitais da região Nordeste, ao passo que, com exceção de Belo Horizonte e Cuiabá, não há registro dessa variante nas outras capitais do Centro-Sul do país.

(...) *rapariga*, a terceira variante mais produtiva no Brasil, alcançou maior produtividade no Nordeste, tendo sido registrada em todas as capitais dessa região, ocorrendo, ainda, em duas regiões (*sic* – capitais) nortistas: Boa Vista e Rio Branco, em uma capital do Centro-Oeste, Cuiabá, e em uma capital do Sudeste: Belo Horizonte. Retomando as marcas dialetais atribuídas à unidade léxica *rapariga* por Houaiss (2001) que a classifica como um regionalismo do Nordeste, de Minas Gerais e de Goiás e por Ferreira (2004) que acrescenta, além dessas, também como um brasileiro do Norte, os dados aqui examinados não confirmaram esses dialetismos regionais na sua totalidade pela não documentação de *rapariga* em Goiás. Em contrapartida, os dados do Projeto ALiB documentaram *rapariga* em Cuiabá, capital não mencionada pelos lexicógrafos consultados. (BENKE, 2012, p. 195)

Figura 34 - Carta linguística para os designativos de *prostituta* – QSL 142



Fonte: Banco de dados do ALiB (2012). In: Benke (2012, p. 194)

Cabe ressaltar que a variante em pauta não ocorreu em Goiânia, uma vez que a pesquisa de Benke (2012) não incluiu dados do interior do Estado. Estaria *rapariga* presente na fala dos habitantes de todo o Estado? Somente com os dados do interior se poderia ratificar ou não a afirmação de Benke (2012).

Em nota, na página 190, a autora informa que a variante *china* apresentou-se como ocorrência única na fala de uma informante de Florianópolis (230/4). De acordo com Benke (2012, p. 179), esta forma lexical encontra-se dicionarizada em Ferreira (2004) e Houaiss (2001) como um regionalismo do sul do

Brasil para designar a *prostituta*. Dentre outros questionamentos, estaria esta variante em desuso na fala dos brasileiros sulistas ou os dados do interior do Brasil revelariam traços regionais que a particularizariam?

Vale notar que, já na década de 1980, a forma *china* era pouco frequente como um designativo para a *prostituta*, uma vez que a carta 283 do Atlas Linguístico-Etnográfico da Região Sul do Brasil (ALTENHOFEN; KLASSMAN, 2011) apresenta baixa produtividade para esta variante, constituindo-se como a última variante cartografada, seguida apenas pelas *hápax legomena*. Na carta 283 do ALERS, a variante *china* apresenta-se apenas com sete ocorrências, registradas em pontos linguísticos distribuídos geograficamente por todo o Estado do Rio Grande do Sul: ponto 679 (São Luís Gonzaga), localizado no noroeste do Estado; ponto 685 (Itaqui), no sudoeste do Rio Grande do Sul; ponto 711 (Lagoa Vermelha), no nordeste rio-grandense; 743 (Santa Maria), centro-sul do Estado; ponto 783 (Quaraí) e ponto 789 (Dom Pedrito), localizados no extremo sul do Brasil e no ponto 833 (Canguçu), no sudeste do Rio Grande do Sul. Uma pesquisa sobre esta variante, com os dados do interior, revelaria se ela ainda persiste no vocabulário ativo dos gaúchos evidenciando ou não a existência de uma área lexical no subfalar sulista, no que se refere ao emprego das formas para designar *prostituta*.

2.3.1.1.6 Costa (2012)

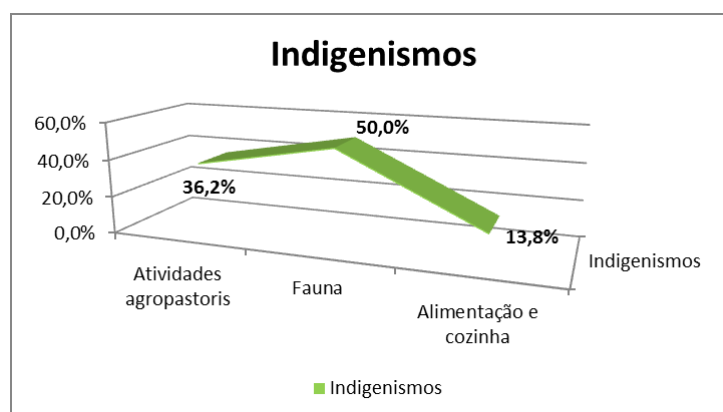
A dissertação de mestrado de Daniela de Souza Silva Costa, defendida na UFMS, analisou a manutenção de indigenismos no léxico do PB com base nos dados das 25 capitais do Projeto ALiB. Dentre as abordagens do trabalho está a disposição para discutir em que medida motivações extralinguísticas, como fatores sócio-histórico-culturais, podem auxiliar na manutenção ou no desaparecimento da presença indígena na identidade brasileira. Para tanto, Costa (2012) pautou-se nos dados de 62 questões do QSL, divididas em três áreas semânticas: fauna (25), alimentação e cozinha (12) e atividades agropastoris (25).

Dos 800 itens documentados para as questões, foram catalogados 59 indigenismos, o que equivale a 7,2% das respostas documentadas. De acordo com a autora, “[...] a maioria das respostas foi de origem portuguesa, que resultaram em 500 unidades léxicas, havendo também a ocorrência de itens lexicais de outras

bases linguísticas” (COSTA, 2012, p.103). Quanto à área semântica (Figura 35), a autora afirma que a área da *Alimentação e cozinha*

revela um menor índice de indigenismos, haja vista que, mesmo tendo quase a metade da quantidade de perguntas das demais áreas semânticas, o registro de ocorrências aproximou-se da terça parte das *Atividades agropastoris* e da quarta parte de *fauna*. (COSTA, 2012, p. 105)

Figura 35 - Produtividade dos indigenismos segundo a área semântica



Fonte: Banco de dados do ALiB (2012). In: Costa (2012, p. 105)

Do ponto de vista diatópico, Costa (2012) afirma que a região Norte apresentou o maior número de indigenismos, 39 dos 58 estudados na pesquisa (COSTA, 2012, p. 194). Apresentando os resultados sob a forma de cartas linguísticas experimentais, a autora assevera que

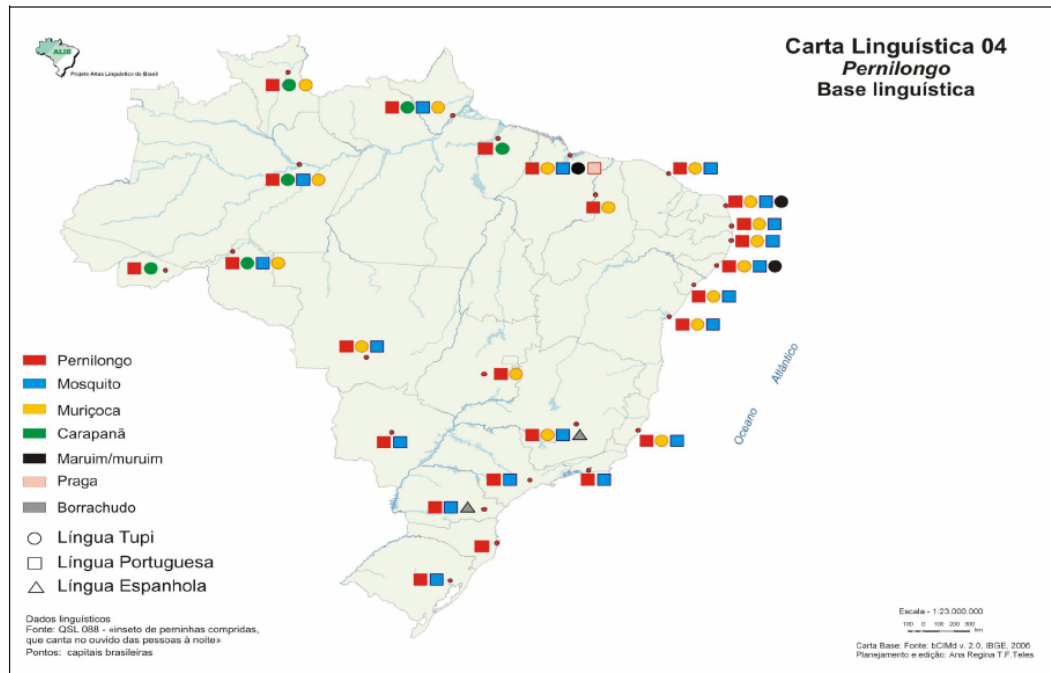
A disseminação das unidades lexicais demonstrou também a presença de isoglossas léxicas em território nacional, mostrando o caminhar das unidades léxicas pelo Brasil e sua maior ou menor concentração, de acordo com as localidades pesquisadas. (COSTA, 2012, p. 197)

Para ratificar essa informação, a autora elaborou a carta experimental acerca dos designativos da questão 88 do QSL – *pernilongo* (Figura 36).

Dois pontos principais podem ser destacados nesta carta linguística. O primeiro deles refere-se à afirmação da autora sobre a disseminação de variantes pelo território nacional, e o segundo, quanto à definição de áreas lexicais a partir de dados das capitais. De acordo com Costa (2012),

Nas obras lexicográficas consultadas, *muriçoca* aparece classificada como marca regional do Nordeste e do Estado de Minas Gerais. Todavia, os dados do ALiB demonstram que a disseminação desse item lexical está para além do dicionarizado, o que pode ser atribuído às migrações de nordestinos para todas as regiões do Brasil. (COSTA, 2012, p. 158)

Figura 36 - Distribuição diatópica das designações para pernilongo (QSL 88), segundo a base linguística



Fonte: Banco de dados do ALiB (2012). In: Costa (2012, p. 159)

Observa-se ainda que, para esta questão, não foram registradas variantes de base tupi em Porto Alegre, Florianópolis, Curitiba, São Paulo, Rio de Janeiro e Campo Grande e que a variante *carapanã* esteve presente nas capitais nortistas o que poderia evidenciar traços do subfalar Amazônico preconizado por Nascentes (1953). A autora afirma que

A disseminação das unidades lexicais demonstrou também a presença de isoglossas léxicas em território nacional, mostrando o caminhar das unidades léxicas pelo Brasil e sua maior ou menor concentração, de acordo com as localidades pesquisadas. Podem-se citar, além de *muriçoca* (área semântica da fauna) e *curau* (área semântica da alimentação e cozinha), também *mangará*, *mandioca*, *macaxeira* e *aipim* (atividades agropastoris) para comprovar a existência dessas linhas imaginárias que delineiam o falar do homem cidadão entrevistado pelo Projeto Atlas Linguístico do Brasil. (COSTA, 2012, 197-198)

Cabe aqui a mesma ressalva apontada acerca das conclusões de Benke (2012) e Romano; Aguilera (2008). Seria possível traçar isoléxicas a partir de dados das capitais? Reitera-se o que já se afirmou a propósito desta questão: somente um trabalho com dados do interior poderia confirmar ou não a existência dessas isoléxicas, bem como definir se existem áreas lexicais ou outros subfalares dentro de territórios, como por exemplo, o sulista.

2.3.1.1.7 Freitas-Marins (2012)

Também como dissertação de mestrado defendida na UFMS encontra-se o trabalho de Luciene Gomes de Freitas Marins. Diferentemente das outras três dissertações, a autora não trabalhou exclusivamente com dados das capitais brasileiras, mas focalizou uma região administrativa, o Centro-Oeste.

A autora analisou algumas questões referentes a três áreas semânticas do QSL: *Atividades agropastoris, Habitação e Vida urbana* com vistas a discutir os traços de ruralidade e urbanidade na fala de 108 informantes radicados na rede de pontos do Projeto ALiB da região Centro-Oeste, totalizando 24 localidades.

Para a constituição do *corpus* de sua pesquisa, Freitas-Marins selecionou as seguintes questões: 052 - *carrinho de mão*; 053 - *haste do carrinho*; 054 - *forquilha*; 055 - *cangalha*; 056 - *canga*; 058 - *bruaca*; 171 - *fuligem*; 172 - *borralho*; 174 - *lanterna*; 194 - *sinaleiro*; 195 - *lombada*; 196 - *calçada*; 198 - *rotatória*; 200 - *ônibus urbano*; 202 - *boteco*, apresentando os resultados de seu trabalho em forma de tabelas, gráficos e cartas linguísticas.

Dentre os resultados, destacam-se as conclusões que vêm ao encontro das asseveradas por Romano e Seabra (2014a) no que se refere à diferenciação linguística do Estado de Goiás em comparação com os demais Estados da região Centro-Oeste.

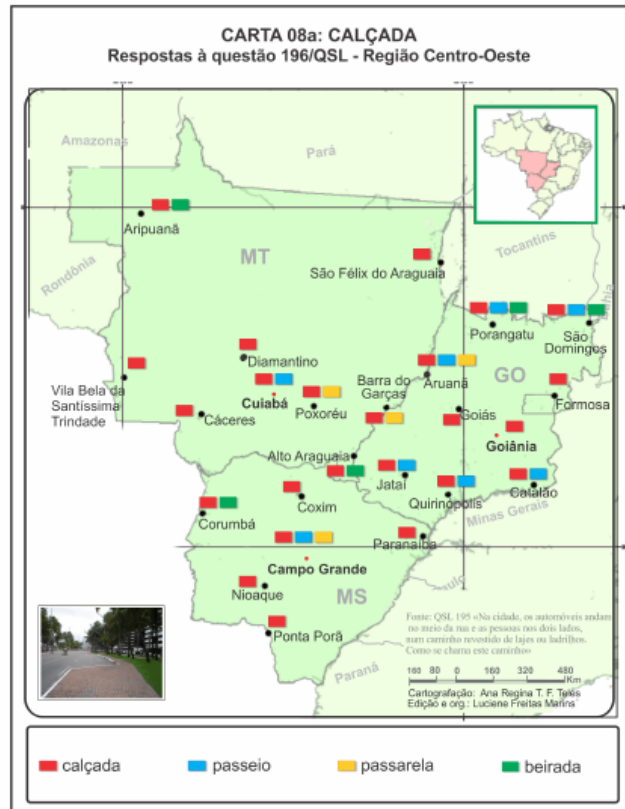
No Estado de Goiás, observou-se (*sic*) tendências diferenciadas dos demais Estados pesquisados: (i) o uso do item lexical *semáforo* obtido como resposta para a pergunta 194/QSL, mais produtivo apenas no interior de Goiás; (ii) a ocorrência da unidade lexical *passeio* como resposta para a pergunta 196/QSL, no interior de Goiás, ao passo que nos demais Estados esse item lexical foi registrado apenas nas capitais. O uso dessa unidade lexical também evidencia traços de conservadorismo linguístico, já que foi mencionada apenas entre os informantes da segunda faixa etária; (iii) Goiás foi o único Estado onde não foi documentada a unidade lexical *bolicho*, de alta produtividade em todas as localidades dos Estados de Mato Grosso e de Mato Grosso do Sul e (iv) a forma *buraca* (variante fonética de bruaca) na região Centro-Oeste foi documentada apenas no interior de Goiás independente do perfil do entrevistado. (FREITAS-MARINS, 2012, 297)

Reproduz-se na Figura 37 a carta elaborada pela autora para representar a distribuição diatópica das variantes lexicais para *Calçada* – QSL 196 – na rede de pontos da região Centro-Oeste. Ainda de acordo com a autora,

Dentre os Estados investigadas, Mato Grosso foi o que apresentou maior índice de unidades léxicas que deixam transparecer marcas do período da escravidão, como o uso de *quitanda* e de *juvu* (da língua africana) e o registro de *venda/vendinha*, unidade léxica que nos séculos passados dava nome ao estabelecimento humilde aberto por negros livres da escravidão. Esses dados confirmam a importância do léxico como um índice revelador da cultura de um povo. (FREITAS MARINS, 2012, 297)

Tomando por base os resultados já encontrados por Freitas-Marins (2012), estes dois Estados (Mato Grosso e Goiás) apresentariam traços lexicais coincidentes, por exemplo, com os Estados da região Sul ou mesmo com o Estado de São Paulo e Sul de Minas Gerais? Seria possível encontrar áreas lexicais nesta região que se diferenciariam das outras constantes do subfalar sulista? Eis mais um motivo para não classificar o subfalar sulista como uma área linguística homogênea, cabendo, portanto, outras áreas dialetais ou outros subfalares dentro do falar do Sul.

Figura 37 - Carta linguística para os designativos de *calçada* – QSL 196



Fonte: Banco de dados do ALiB (2011). In: Freitas-Marins (2011, p. 102)

2.3.1.1.8 Ribeiro (2012)

A primeira tese de doutorado desenvolvida com base nos dados do Projeto ALiB é a de Silvana Soares Costa Ribeiro – *Brinquedos e brincadeiras infantis na área do falar baiano*, defendida na Universidade Federal da Bahia. Em seu trabalho, Ribeiro discutiu a validade da proposta de divisão dialetal de Nascentes (1953) no que se refere ao território classificado como subfalar baiano. Dessa forma, a autora investigou, em 57 localidades junto a 244 informantes, os designativos para 13 questões do Questionário Semântico-Lexical do Projeto ALiB, concernentes à área semântica *Jogos e diversões infantis*.

A área correspondente ao subfalar baiano pertence a quatro das cinco regiões administrativas do país e compreende 11 Estados da Federação. Além da rede de pontos localizada na área geográfica selecionada, a autora considerou pontos adjacentes aos limites do subfalar. Estes pontos adjacentes a autora denominou como *pontos de controle*, pelos quais foi possível constatar que o falar

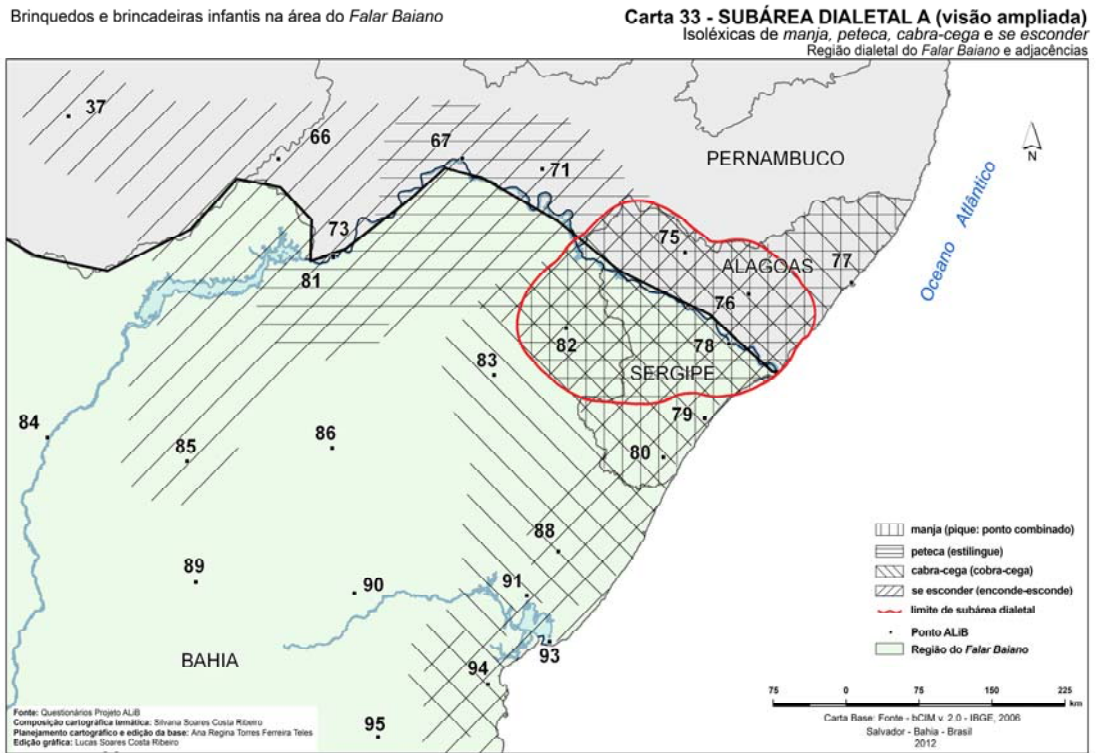
baiano não se limita à faixa territorial estabelecida por Nascentes (1953). Nas palavras da autora:

O Falar Baiano é uma área geográfica que ora se expande para os (...) pontos de controle dos Estados de Alagoas e Pernambuco, ou que ora se elastece para todo o Estado de Goiás (hoje subdividido em GO e TO), ou avança por pontos de controle de Minas Gerais, (...) e, excepcionalmente, ora se vê englobando os pontos de controle dos Estado do Piauí, Maranhão ou Espírito Santo. (RIBEIRO, 2012, p. 439)

Desse modo, Ribeiro conseguiu identificar quatro subáreas dialetais na região do subfalar baiano:

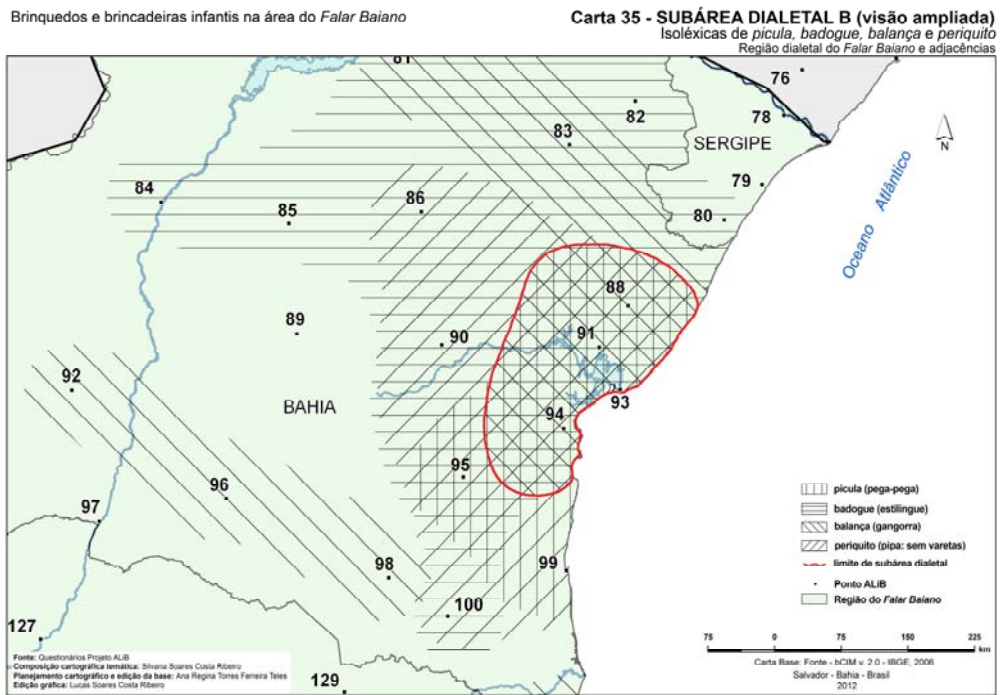
- (1) Subárea dialetal A: localizada no nordeste da área geográfica pesquisada (Estado da Bahia, Sergipe e Alagoas), atestada pela presença simultânea de variantes como: *manja* (pique: local combinado), *peteca* (estilingue), *cabra-cega* (cobra cega) e *se esconder* (esconde-esconde), conforme atesta na carta 33 (Figura 38);
- (2) Subárea dialetal B: localiza-se no Estado da Bahia onde se atesta a presença simultânea de *picula* (pega-pega), *badogue* (estilingue), *balança* (gangorra) e *periquito* (pipa sem varetas), conforme se visualiza na carta 35 (Figura 39);

Figura 38 - Subárea dialetal A no subfalar baiano



Fonte: Banco de dados do ALiB (2011). In: Ribeiro (2012, vol. 2)

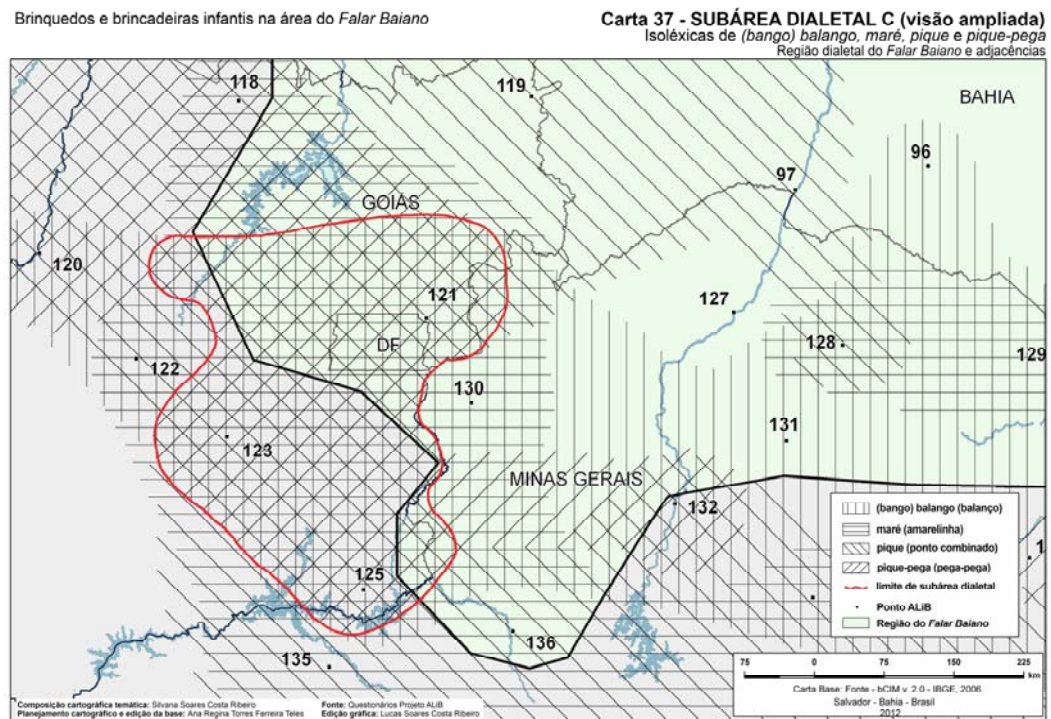
Figura 39 - Subárea dialetal B no subfalar baiano



Fonte: Banco de dados do ALiB (2011). In: Ribeiro (2012, vol. 2)

- (3) Subárea dialetal C: localizada no Estado de Goiás, tanto em área do Falar Baiano quanto em Área de Controle e caracteriza-se pela presença simultânea (*bango*)-*balango* (balanço), *maré* (amarelinha), *pique* (local combinado), e *pique-pega* (pega-pega). Uma visão ampliada dessa subárea dialetal é apresentada na carta 37 (Figura 40);

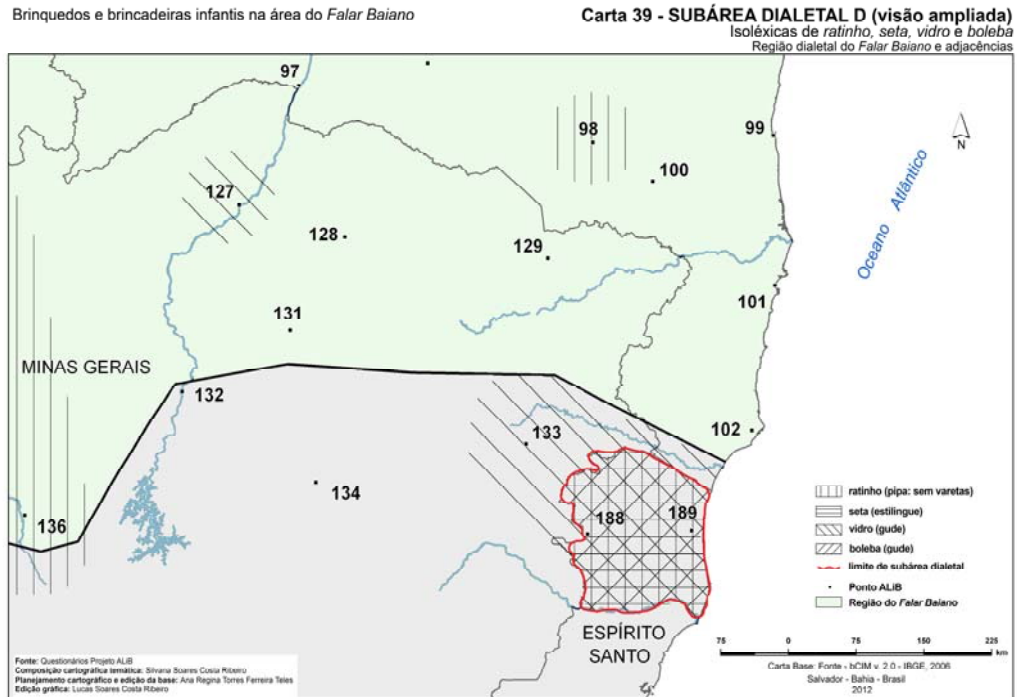
Figura 40 - Subárea dialetal C no subfalar baiano



Fonte: Banco de dados do ALiB (2011). In: Ribeiro (2012, vol. 2)

- (4) Subárea dialetal D: situada no Estado do Espírito Santo e em Área de Controle, configurada pela presença simultânea de *ratinho* (pipa sem varetas), *seta* (estilingue), *vidro* e *boleba* (gude), conforme se visualiza na Figura 41 (Carta 39).

Figura 41 - Subárea dialetal D no subfalar baiano



Fonte: Banco de dados do ALiB (2011). In: Ribeiro (2012, vol. 2)

Embora a pesquisa de Ribeiro (2012) não discuta os resultados associando-os aos aspectos sócio-históricos da região investigada, o que enriqueceria a interpretação dos dados, o trabalho mostra-se relevante para a Geolinguística brasileira, uma vez que aponta a possibilidade de discutir divisões dialetais a partir de dados lexicais. Segundo a autora:

Uma das constatações a que se chegou foi a de que o léxico se mostrou apropriado para traçar áreas dialetais. Para tanto, as isoléxicas estabelecidas demonstraram que os caminhos das palavras são muitos, o que possibilitou atestar áreas dialetais, a partir do léxico, foi o estabelecimento de agrupamentos de lexias distintas usadas para nomear conceitos distintos e que, quando reunidas, por sobreposição de isoléxicas (um feixe de isoglossas), puderam demonstrar a variação diatópica. (RIBEIRO, 2012, p. 449)

O trabalho apresenta, pois, valiosas contribuições no que se refere ao desenvolvimento do Projeto ALiB, evidenciando traços linguísticos que caracterizam as regiões, indicando possibilidades para a cartografia de itens lexicais.

Ribeiro (2012), ao afirmar que a proposta de Nascentes (1953) tem vitalidade na realidade presente, faz uma ressalva de suma importância: “o limite

traçado pelo autor ainda não pode ser alargado ou reduzido sem que antes seja conhecida, com maior profundidade, a área circunvizinha ao Falar Baiano” (RIBEIRO, p. 449). Ou seja, são necessários trabalhos complementares e uma discussão mais acurada sobre as áreas dialetais em outras regiões brasileiras. Neste particular, a presente tese busca alargar os horizontes e verificar a vitalidade da proposta de Nascentes no que se refere ao subfalar sulista, sob o ponto de vista lexical.

2.3.1.1.9 Portilho (2013)

A dissertação de mestrado *O falar amazônico: uma análise da proposta de Nascentes (1953) a partir de dados do Projeto ALiB*, de autoria de Portilho (2013), defendida na UFMS, discute a vitalidade da proposta de Antenor Nascentes na atualidade, pautando-se em dados de 26 localidades e de 128 informantes. Para o estudo, Portilho (2013) selecionou 13 questões do QSL concernentes à área semântica *Jogos e Diversões Infantis* o que possibilita análises comparativas com o trabalho de Ribeiro (2012). Das 26 localidades selecionadas, 20 pertencem à área geográfica do *falar amazônico* (pontos linguísticos localizados nos Estados do AP, RR, AM, PA, AC, RO) e seis localidades de controle que se distribuem pelos Estados de RO, TO, MA e MT.

A análise da autora contempla duas abordagens: a diatópica e léxico-semântica. Acerca das análises diatópicas, subsidiadas por cartas linguísticas elaboradas *ad hoc*, a autora constatou que a vitalidade do falar amazônico pode ser parcialmente atestada, ou seja, as variantes lexicais não permitem a definição de limites rígidos para o *falar amazônico*. Sobre as variantes para a questão 156 (bolinha de gude), Portilho (2013) tece as seguintes considerações:

Peteca foi o item lexical mais produtivo na área em foco, e *bolinha de gude* alcançou maior produtividade nas duas localidades do Estado de Mato Grosso, evidenciando uma “divisão” de falares – o do Norte diferenciando-se do restante do Brasil. Parece que o item lexical *peteca* está se disseminando para a região Nordeste, considerando a sua elevada produtividade nas localidades da área de controle, pertencentes à região Nordeste. (PORTILHO, 2013, p. 183)

Desse modo, em alguns casos, a divisão dialetal do falar amazônico não se confirma, uma vez que a ocorrência de alguns itens lexicais é comum a outras regiões do país, por motivos de migração interna. Não obstante, constataram-se as interinfluências entre os falares, especialmente entre as localidades fronteiriças e a área dialetal investigada. Assim, “foi atestada uma relativa vitalidade do *falar amazônico* no nível lexical, considerando que o léxico dessa área mostrou-se peculiar em relação ao de outras regiões do Brasil” (PORTILHO, 2013, p.138).

2.4 ALGUMAS CONSIDERAÇÕES SOBRE ÁREAS DIALETAIS SOB A PERSPECTIVA LEXICAL

Como se sabe, as propostas de divisão dialetal de Nascentes (1953) foram pautadas em aspectos fonéticos e prosódicos do português e, de certa forma, são confirmadas com trabalhos desta natureza como os de Cardoso (1986; 1999), acerca da abertura das vogais médias em posição pré-tônica.

Discutir as áreas dialetais a partir do léxico, apesar de os estudos apontarem diferenças interessantes, é uma tarefa ainda pouco explorada; os trabalhos pioneiros nesta seara, cujos objetivos explícitos são definir áreas dialetais de amplos espaços geográficos, são os de Aguilera (2009), Ribeiro (2012), Portilho (2013) e o de Romano e Seabra (2014a).

Aguilera (2009) revisita dados lexicais do Projeto ALiB para discutir a proposta de Nascentes (1953) quanto à validade e atualidade daquela divisão, apontando que à época do dialetólogo “faltavam - e ainda faltam - no entanto, estudos que pudessem/possam determinar as isoglossas” (AGUILERA, 2009, p. 4220).

Conforme bem assinala a autora não é só no Brasil que os dialetólogos têm buscado definir áreas dialetais sob a perspectiva lexical. Cintra (1995), por exemplo, apresenta divisões dialetais em Portugal a partir de variantes lexicais registradas nos inquéritos do ALPI (Atlas Linguístico da Península Ibérica), documentando com uma série de mapas, como o das variantes lexicais para *cabrito* (ao norte e oeste) e *chibo* (no sul e leste) e os designativos para a *espiga de milho* em duas áreas delimitadas: *espiga* correspondendo ao centro nordeste de Portugal e *maçaroca* do noroeste ao sul do país.

A opinião aqui defendida é que somente estudos que considerem o critério da continuidade territorial na rede de pontos, podem, com mais propriedade,

afirmar a existência de áreas lexicais. Os estudos lexicais desenvolvidos a partir de dados das capitais brasileiras do ALiB não podem, de fato, definir as áreas lexicais do português, pois não apresentam uma continuidade territorial dada a distância geográfica que separa cada uma das capitais, inviabilizando o traçado de linhas de isoléxicas. Todavia, esses trabalhos têm apontado a importância do estudo do léxico, sobretudo sob o ponto de vista diatópico, uma vez que o léxico “representa um traço diferenciador marcante, no que diz respeito à variação linguística no eixo horizontal” (ISQUERDO, 2003, p. 166).

É inegável o estatuto do léxico para os estudos linguísticos, pois “o léxico da língua é que mais nitidamente reflete o ambiente físico e social dos falantes” (SAPIR, 1969, p. 45), portanto, considerando-o como “o tesouro vocabular de uma língua” (BIDERMAN, 1992, p. 399), pode-se afirmar que ele constitui o patrimônio de uma sociedade juntamente com outros símbolos da herança cultural.

Dessa forma, o estudo do léxico distribuído diatopicamente em uma rede de pontos contínua torna-se interessante para a delimitação de linhas de isoléxicas que, juntamente com dados de outros níveis da língua e com cartas linguísticas desta natureza, poderão indicar a existência de áreas dialetais.

Entendendo que “os limites linguísticos não coincidem necessariamente com os limites naturais” (COSERIU, 1987, p. 95), e, por extensão, com os limites políticos, compreende-se que as especificidades lexicais de dada região estão mais condicionadas a aspectos histórico-culturais do que propriamente geográficos e políticos. Nesse sentido, a presente tese busca, no capítulo 4, associar os resultados encontrados em dados lexicais de natureza geolinguística aos aspectos históricos e sociais de cada uma das regiões contempladas neste estudo. Desse modo, a área correspondente ao subfalar sulista não pode ser considerada homogênea, haja vista a existência de subáreas dialetais e a heterogeneidade da formação étnica nesta porção do país.

Neste sentido, Aguilera (2009) assinala em suas conclusões que a área correspondente aos subfalares apontados por Nascentes (1953) devem ser revistas e estudadas de forma mais pormenorizada.

O falar do Sul, que Nascentes dividiu em subfalares, porém, além de demonstrar menor uniformidade ainda não traz, pelo menos nas variantes aqui analisadas, a divisão nos subfalares *fluminense*, *mineiro* e *sulista*. A denominação deste último eu substituiria por subfalar *paulista* que considero mais apropriada para designar a variante de grande parte dos territórios dos Estados de São Paulo, Paraná, Mato Grosso, Mato Grosso do Sul, e partes menores de Goiás e Santa Catarina. Nesta minha proposta, reservaria a denominação *sulista* para a variante do Rio Grande do Sul, que, pelo menos no nível lexical que aqui analisamos, se constitui diferentemente dos demais Estados. Justifico a denominação de subfalar *paulista* considerando a influência dos paulistas que adentraram o Brasil a partir do século XVII, intensificada nos séculos posteriores em direção ao sul de Minas Gerais, Goiás, Mato Grosso (inclusive o atual Mato Grosso do Sul), Paraná e parte de Santa Catarina. (AGUILERA, 2009, p.4232)

Não seria, então, agora que já se possui a documentação empírica do português falado no Brasil, a oportunidade de discutir a proposta de divisão dialetal de Nascentes (1953) e propor um novo mapa dialetológico para o Brasil, pelo menos no que se refere ao subfalar sulista? Acredita-se, em consonância com as ideias de Aguilera (2009), que ao conjunto dos subfalares fluminense, mineiro e sulista, podem ser incluídos outros subfalares, como, por exemplo, o paulista, principalmente, pelo foco de irradiação e a importância do movimento dos bandeirantes no território brasileiro. Não estaria o subfalar sulista de Nascentes contemplando outros subfalares não discriminados pelo dialetólogo? Pretende-se responder a estas e outras perguntas após a análise dos dados que se empreende nesta tese.

CAPÍTULO 3

MATERIAIS E MÉTODOS

Este capítulo destina-se a apresentar os materiais e métodos utilizados na pesquisa que deram origem a esta tese.

3.1 CONSTITUIÇÃO DO *CORPUS*

O *corpus* analisado nesta pesquisa refere-se aos dados ainda inéditos coletados pela equipe do Projeto ALiB em 118 cidades brasileiras. Considerando a participação do autor deste trabalho junto a esse grupo de pesquisa desde o ano de 2006 e o interesse em continuar trabalhando com os dados do ALiB, agora no doutorado, foi solicitada, previamente, à presidente do Comitê Nacional, Dr^a. Suzana Alice Marcelino Cardoso, a autorização para utilizar os dados do Projeto, no sentido de também contribuir para o desenvolvimento do atlas nacional a partir das análises aqui empreendidas.

Uma vez concedida a permissão para utilização do material (ANEXO B), passou-se a estudar a melhor forma de organizar e selecionar os dados orais dos 472 informantes que compõem a pesquisa desta tese.

Em cada uma das localidades do interior do Brasil, foram entrevistados quatro informantes, dois homens e duas mulheres, pertencentes a duas faixas etárias – Faixa I: 18 a 30 anos e Faixa II: de 50 a 65 anos, naturais da região linguística e com ensino fundamental de escolaridade (completo ou incompleto). Vale lembrar que nas capitais de cada Estado foram entrevistados, além de informantes de nível fundamental, também quatro informantes de nível superior, perfazendo oito entrevistados nestas localidades. Todavia, os dados dos informantes de nível superior não foram considerados para este estudo, uma vez que o número de informantes com este perfil seria bem menor em comparação com os dados de informantes de nível fundamental, o que, metodologicamente, inviabilizaria as análises estatísticas aqui empreendidas. Portanto, os dados analisados nesta tese referem-se apenas a informantes de nível fundamental, tanto das localidades do interior quanto das capitais.

3.1.1 A Área Geográfica Estudada

A área investigada contempla localidades situadas em nove Estados federativos, a saber: Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Rio de Janeiro, Minas Gerais, Goiás, Mato Grosso e Mato Grosso do Sul.

A dificuldade que se teve ao trabalhar com o mapa de Nascentes (1953), reproduzido na Figura 19 (p. 103), 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, Nascentes (1953, p.26) descreve que este compreende os Estados de “São Paulo, Paraná, Santa Catarina, Rio Grande do Sul, Minas (Sul e Triângulo), Goiás (Sul) e Mato Grosso”, não apresentando mais especificações sobre tais limites. Vale notar também que a imagem do mapa do Brasil subdividido em falares, constante da página 19 do livro *O linguajar Carioca* (1953), não apresenta elementos geográficos que subsidiem a precisão das fronteiras de um e outro subfalares.

Desse modo, depois de criterioso estudo do mapa dialetal de Nascentes (1953) e a distribuição da rede de pontos do ALiB, juntamente com um geógrafo contratado para a elaboração da base cartográfica⁶⁰, chegou-se a um mapa parecido ao de Nascentes (1953) (vol. 2) no que se refere à área geográfica do falar sulista. Além desta área do subfalar sulista, foram também considerados pontos linguísticos adjacentes aos limites desse subfalar, aqui também denominados “pontos de controle”, como já o fizera Ribeiro (2012).

Portanto, a rede de pontos do subfalar sulista compõe-se de 108 localidades, às quais foram acrescentados 10 pontos para este estudo, uma vez que se considera importante verificar a fronteira deste falar com os outros falares (mineiro e fluminense) e também com o território incaracterístico, o que ajudará na definição dos limites dos “falares sulistas”, haja vista que se parte do pressuposto de que a área investigada não apresenta comportamento homogêneo sob o ponto de vista lexical⁶¹.

⁶⁰ Sobre definição de base cartográfica e mapa base, confira Teles e Ribeiro (2006). O geógrafo contratado foi Alan Alievi que fez o georreferenciamento da rede de pontos linguísticos e elaborou a base cartográfica utilizada neste estudo.

⁶¹ Ver rede de pontos linguísticos constante do vol. 2 deste trabalho.

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: pontos 118 (Porangatu) e 121 (Formosa) e, por fim, duas cidades mato-grossenses: 103 (Aripuanã) e 104 (São Félix do Araguaia).

No total, portanto, este estudo engloba nove localidades mato-grossenses, seis sul-mato-grossenses, oito localidades goianas, duas fluminenses, 11 municípios mineiros, 38 cidades paulistas, 17 paranaenses, 10 localidades catarinenses e, por fim, 17 cidades gaúchas, conforme se observa com mais detalhes no volume 2 (Rede de pontos linguísticos).

3.1.2 Seleção das Questões

Sobre as questões selecionadas para este estudo e tendo sempre em vista o objetivo principal do trabalho, não se optou por uma área semântica específica, como fizeram Yida (2011) e Ribeiro (2012). Desse modo, após uma pesquisa em estudos elaborados a partir do *corpus* do ALiB e de cartas linguísticas de atlas estaduais já publicados, bem como cartas elaboradas *ad hoc* para apresentação em eventos e seminários, foram selecionadas questões que evidenciam diversidade de formas lexicais, considerando-se a variável espaço geográfico, ou seja, questões que, possivelmente, apresentam 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 também 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 (lexical). 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.

Nesta tese, analisam-se, portanto, os dados de cinco questões do QSL:

- Questão 001 – Córrego – que traz no *caput* a seguinte formulação “Como se chama um rio pequeno, de uns dois metros de largura?”, da área semântica Acidentes Geográficos;
- Questão 039 – Tangerina – “Como se chamam as frutas menores que a laranja, que se descascam com a mão, e, normalmente, deixam um cheiro na mão?”, do campo semântico Atividades Agropastoris;
- 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.

Estes dados foram levantados a partir do *corpus* do ALiB, mediante a consulta às transcrições/revisões e às gravações em áudio. O *corpus* passou pelo processo de revisão para, assim, ser armazenado no banco de dados do software que foi desenvolvido com a finalidade de agilizar o processo de geração de relatórios e de cartografia linguística.

3.2 O Sistema [ΣGVCLin] - Software para Geração e Visualização de Cartas Linguísticas⁶²

⁶² O Software encontra-se em fase de registro de patente junto ao INPI – Instituto Nacional da Propriedade Industrial. Processo nº. BR5120140000486.

Considerando o volume de dados a ser tratado e analisado, foi necessário elaborar um programa computacional que fornecesse não apenas relatórios diversos por meio de consulta a um banco de dados informatizado, mas também a geração de cartas linguísticas de forma automatizada, uma vez que o modo tradicional de elaborar cartas, usando programas de edição de imagem, exigiria demasiado investimento de tempo, além de não permitir a dinamicidade e a produção de grande número de cartas sem a intervenção de terceiros no processo de cartografia.

Com a finalidade de auxiliar esta importante etapa do trabalho geolinguístico – a cartografiação – foi que o Projeto do [ΣGVCLin], já idealizado durante o curso de mestrado do autor da tese, concretizou-se por meio do diálogo de duas áreas distintas: a Linguística e a Computação.

Desenvolvido no *campus* da Universidade Federal de Itajubá (UNIFEI), no Instituto de Matemática e Computação, sob a orientação do prof. Dr. Rodrigo Duarte Seabra, o software contou com a atuação direta de um bolsista de iniciação científica do CNPq, Nathan Oliveira, graduando em Ciência da Computação. O [ΣGVCLin] foi desenvolvido no período de março a dezembro de 2013 e, para seu desenvolvimento, realizaram-se reuniões semanais de planejamento e execução das atividades.

Obviamente, o programa não atende apenas às necessidades do desenvolvimento deste trabalho. Desse modo, o software foi concebido pensando no uso comum dos pesquisadores da área, com a finalidade de divulgá-lo no cenário nacional e internacional, uma vez que foram elaboradas duas versões do programa: em português e em inglês.

Este programa tem o objetivo de facilitar o processo de consulta ao material linguístico transcrito e armazenado em um banco de dados geral que servirá para quaisquer tipos de projetos, baseados nos moldes do ALiB. Os resultados obtidos a partir da ferramenta poderão ser consultados por meio de relatórios diversos, bem como pela geração de cartas linguísticas de produtividade e de cartas de arealidade.

A modelagem do banco de dados seguiu informações básicas pautadas em critérios metodológicos do Projeto ALiB já sedimentados entre os estudiosos da área. Vale lembrar que os pesquisadores que se propõem elaborar atlas linguísticos não podem ignorar os avanços alcançados pelo ALiB,

considerando a sua abrangência e importância para os estudos geolinguísticos brasileiros. Desse modo, com esta ferramenta totalmente gratuita e de fácil interação, dado o caráter indutivo atribuído às funcionalidades do programa, depois de armazenadas as informações necessárias para o estudo, o próprio linguista consegue gerar os relatórios e as cartas linguísticas sem grandes dificuldades⁶³.

A motivação principal que direcionou o desenvolvimento do *software* foi a criação de um ambiente no qual o linguista possa ter independência no manuseio da ferramenta. Entretanto, pelo fato de a Geolinguística estar inserida na interface da Linguística e da Geografia, cabe ao estudioso, antes de iniciar o uso do *software*, contatar um geógrafo para que este prepare sua base cartográfica ou carta base, que esteja devidamente georreferenciada, contendo os elementos indispensáveis, tais como a escala gráfica, os limites políticos do território e a localização dos pontos linguísticos. Uma vez elaborado o mapa em formato de imagem (extensões 'png', 'jpg' ou 'bmp'), o linguista poderá carregá-lo na ferramenta e, assim, representar inúmeros fatos linguísticos (fonéticos, lexicais, morfossintáticos).

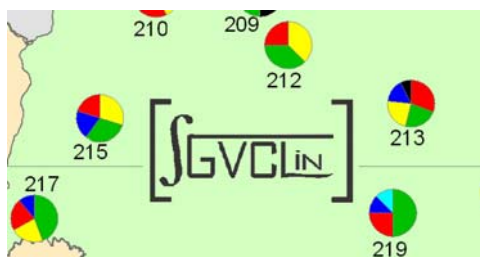
A diretriz norteadora para o desenvolvimento e a implantação do sistema consistiu no projeto de uma interface gráfica que priorizasse a intuição e a simplicidade. Essas qualidades conduziram o processo de desenvolvimento do sistema com a participação efetiva do perfil de usuário desde a coleta dos requisitos, visando a facilitar sua interação e, deste modo, atender a padrões satisfatórios de uso. Para a definição dos requisitos da ferramenta, foram utilizadas variadas cartas linguísticas com dados representativos de amostras reais, com distintas representações, além dos conhecimentos discutidos em reuniões subsidiadas pelo Projeto ALiB – regional Paraná.

Após o processo de análise de cartas linguísticas de atlas já publicados e de cartas experimentais elaboradas *ad hoc* para artigos científicos com os dados do ALiB, os requisitos foram identificados e deu-se início ao desenvolvimento da ferramenta proposta (Figura 42). De modo geral, segundo Romano; Seabra e Oliveira (2014, p. 128) “o sistema consiste em uma ferramenta computacional inovadora projetada para fornecer uma interface simples que permita

⁶³ Informações sobre o *software*, bem como instruções para instalação e uso podem ser encontradas no website dos desenvolvedores: <http://sgvclin.altervista.org/>. Detalhes sobre a ferramenta também podem ser encontrados em Romano; Seabra e Oliveira (2014) e Seabra; Romano e Oliveira (2014).

o armazenamento de dados geolinguísticos e posterior geração de cartas linguísticas por meio de consultas ao banco de dados”. Além disso, o software permite a criação de grupos de questões por usuário e a geração de relatórios pautados por diferentes variáveis, bem como a criação de cartas bidimensionais: diatópica/diassexual e diatópica/diageracional (THUN, 1998), além de cartas de arealidade.

Figura 42 - Logotipo do sistema



Fonte: [GVCLin].

Por se tratar de uma ferramenta para uso no âmbito da pesquisa e, possivelmente, como recurso complementar no ensino da cartografia de dados linguísticos, o sistema utiliza basicamente uma interface convencional composta por janelas, ícones, menus e dispositivo de apontamento (via *mouse*). Toda a interação do usuário com o sistema ocorre via interface, sem que nenhuma manipulação direta ao banco de dados seja necessária.

O software, em sua versão *desktop*, foi desenvolvido utilizando a linguagem Java⁶⁴ e o banco de dados MySQL⁶⁵, o qual foi modelado com base não só nas fichas dos informantes, mas também na metodologia do Projeto ALiB.

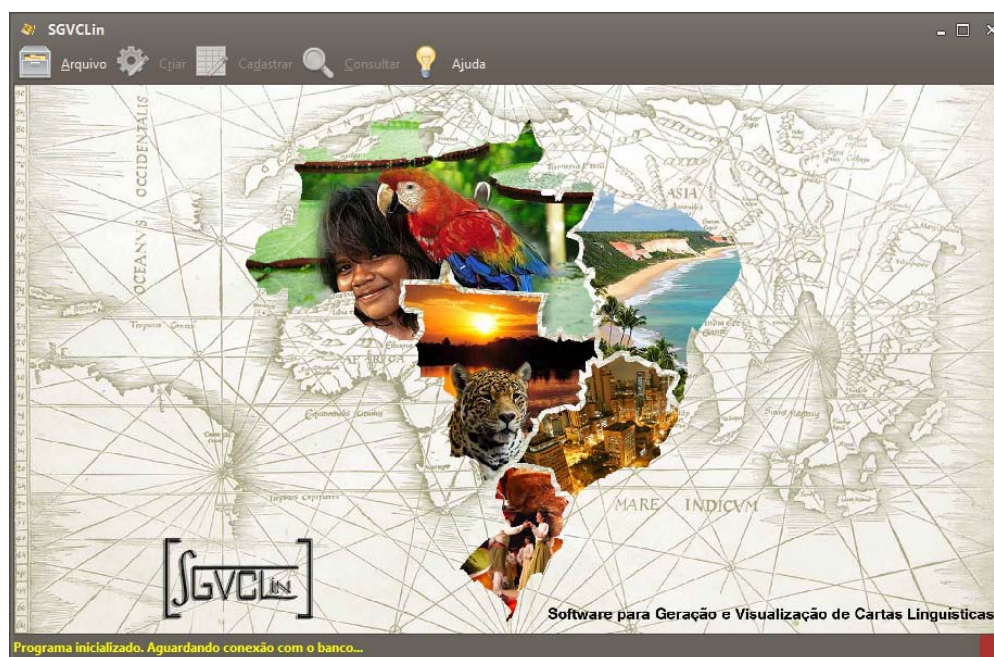
A interface da ferramenta apresenta menus e botões, componentes tradicionalmente disponíveis na maioria das aplicações computacionais convencionais. Inicialmente, são visualizados na tela (Figura 43) uma imagem ilustrativa, a barra de *menu* principal do software na parte superior da janela (com cinco opções) e uma barra de *status*, na parte inferior, que orienta o usuário durante o processo de interação. Pode-se observar ainda no canto direito inferior da Figura 43 um ícone vermelho que indica que o usuário ainda não conectou a ferramenta ao banco de dados responsável pelo armazenamento das informações. Ao estabelecer

⁶⁴ <http://www.java.com/>

⁶⁵ <http://www.mysql.com/>

a conexão, o ícone imediatamente tem sua cor alterada para verde, indicando que ela foi estabelecida com sucesso. Uma descrição completa e detalhada dos componentes e funcionalidades do software é fornecida no artigo de Romano; Seabra e Oliveira (2014).

Figura 43 - Tela principal do software



Fonte: [SGVCLin]

Uma das funcionalidades do software envolve a geração de relatórios, agrupados em três conjuntos e distribuídos em treze formatos diferentes. O primeiro conjunto, o mais amplo, permite a visão abrangente do *corpus* cadastrado no banco de dados e apresenta as seguintes opções: (i) Geral; (ii) Geral por sexo; (iii) Geral por faixa etária; e (iv) Geral por ponto. O segundo conjunto de relatórios – Por Estado – abrange quatro opções: (i) Por Estado; (ii) Por Estado, segundo o sexo; (iii) Por Estado, segundo a faixa etária; e (iv) Por Estado, segundo o sexo e a faixa etária. O terceiro e último conjunto de relatórios diz respeito a algumas informações inerentes aos informantes: (i) Geral por Estado civil; (ii) Geral por escolaridade; (iii) Geral por profissão; (iv) Geral por religião; e (v) Geral por categoria social. Nessa seara, podem-se realizar estudos investigativos correlacionando possíveis respostas às variáveis extralinguísticas⁶⁶.

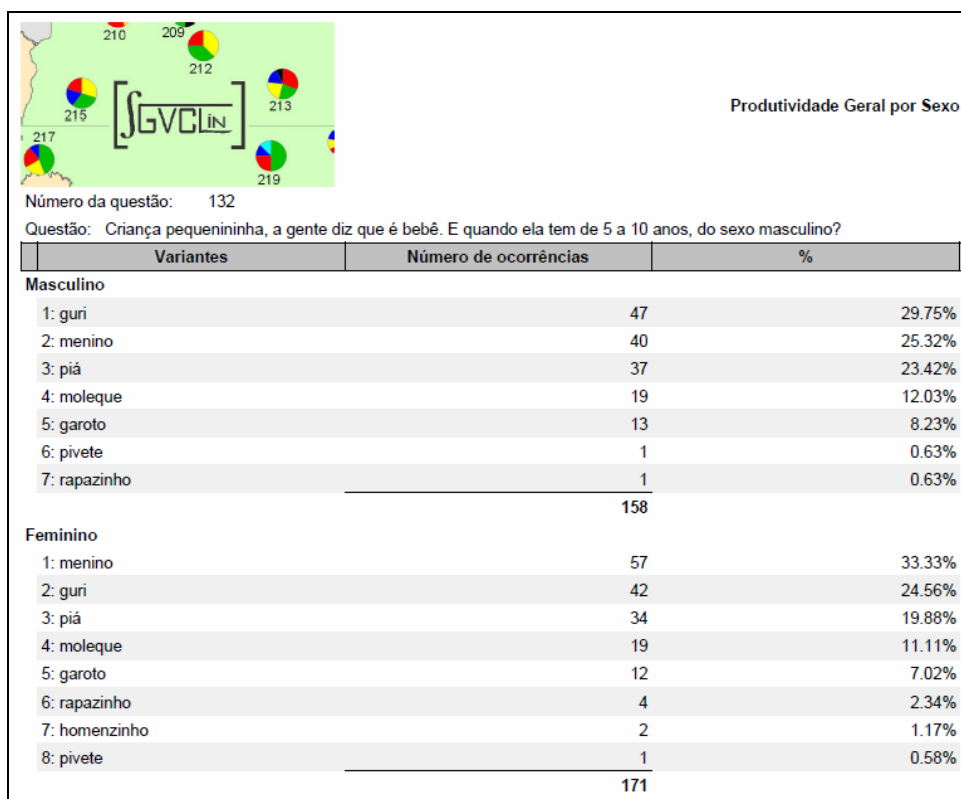
⁶⁶ Neste trabalho, não são utilizadas todas as funcionalidades e tipos de relatórios do [SGVCLin] devido aos objetivos já explicitados.

Neste momento, é importante ressaltar que cada categoria de relatório, quando selecionada pelo usuário, gera um arquivo em formato PDF com as informações tabuladas, por questão, e contabilizadas em função da produtividade decrescente das respostas dos informantes. As informações numéricas que constam dos relatórios são as frequências absolutas (número de ocorrências) e relativas (porcentagem) de cada variante, bem como o total das ocorrências. Além disso, as observações relativas à ocorrência de NS (não soube) e PT (problema técnico) são impressas no rodapé do relatório, com a indicação de seu total, bem como o número do ponto e o informante.

A Figura 44 exibe um exemplo de relatório ilustrativo com a produtividade geral segundo o sexo. Após o usuário selecionar a opção, automaticamente o software filtra os resultados segundo a variável de interesse e realiza a contagem da produtividade de cada variante.

As cartas geradas pela ferramenta apresentam a distribuição das respostas dos informantes, para cada questão, exibidas graficamente sobre um dos diversos mapas base que podem ser inseridos no software. A representação das informações é realizada por gráficos no formato pizza (*pie chart*). Além do mapa com os gráficos, a janela apresenta ainda a legenda da carta (com base na quantidade de variantes informada pelo usuário) usando como referência a produtividade decrescente de cada variante e um histograma com a produtividade geral (Figura 45).

Figura 44 - Exemplo de relatório geral por sexo



Fonte: [ΣGVCLin].

Figura 45 - Exemplo de carta diatópica



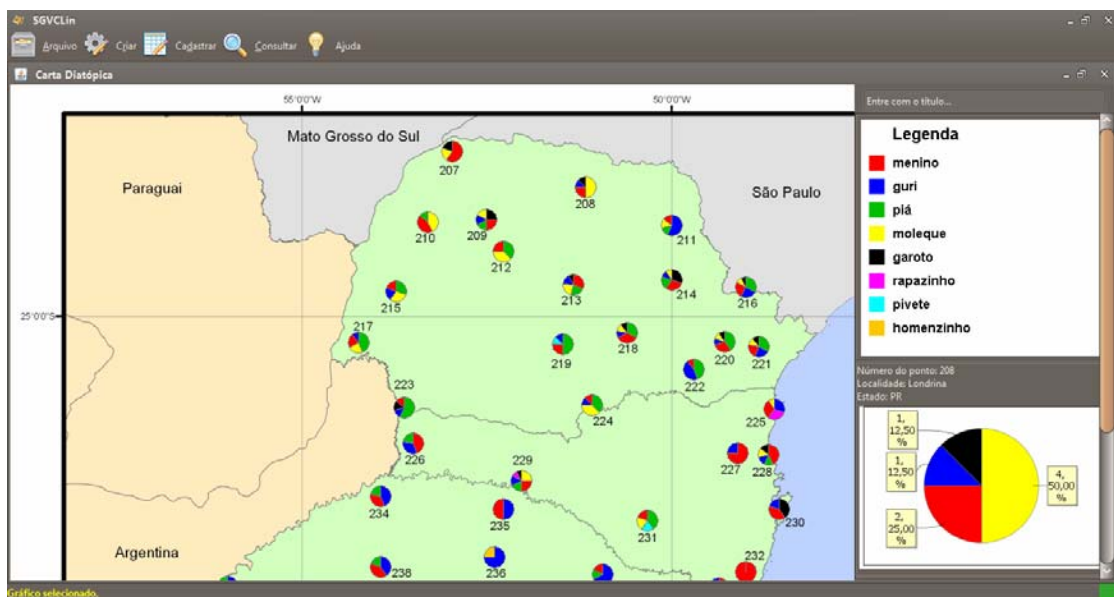
Fonte: [ΣGVCLin]

O software permite ainda a seleção dos gráficos individualmente, de modo que o usuário possa manipular sua posição de acordo com a preferência. Para

selecionar um gráfico, basta clicar sobre ele usando o botão esquerdo do mouse. Após a seleção, cada pizza pode ser transladada mediante o uso das teclas direcionais (setas) presentes no teclado do usuário. Isso permite melhor ajuste dos gráficos durante o processo de confecção da carta. Ao selecionar um gráfico, o [ΣGVCLin] exibe uma cópia ampliada na lateral direita inferior da janela, de modo que o usuário possa visualizar mais facilmente a distribuição das variantes naquele ponto linguístico, bem como comentários específicos daquela localidade, por exemplo, o relato dos informantes, quando houver. A Figura 46 mostra a distribuição das variantes com base na seleção do ponto 208 do mapa.

Ao clicar em qualquer área do mapa, o gráfico em questão perde a seleção, e ela é repassada à imagem. Com o mapa selecionado, o usuário pode movimentá-lo por meio do mesmo princípio aplicado aos gráficos, de acordo com sua preferência. Assim, pode explorar outras regiões do mapa que possivelmente não estejam visíveis em determinado momento da interação.

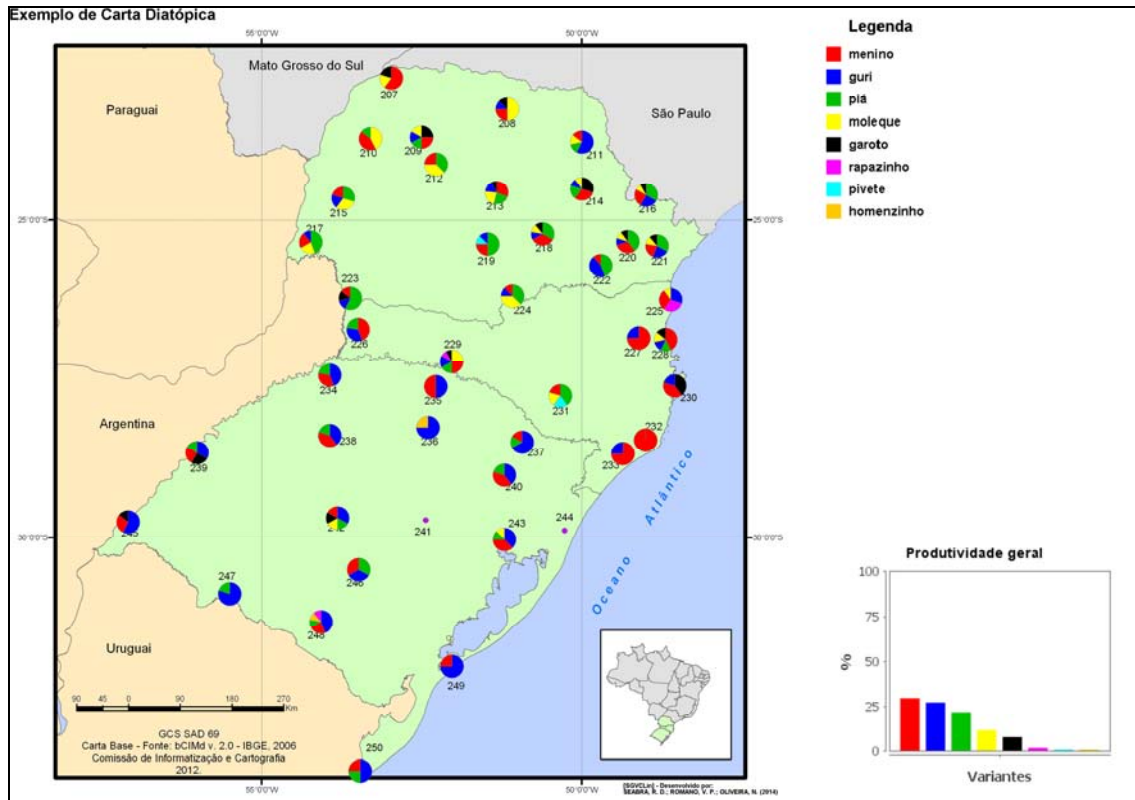
Figura 46 - Exemplo de seleção de gráfico



Fonte: [ΣGVCLin]

Após produzir sua carta, o usuário deve atribuir-lhe um título, que será adicionado à imagem gerada para posterior impressão (Figura 47).

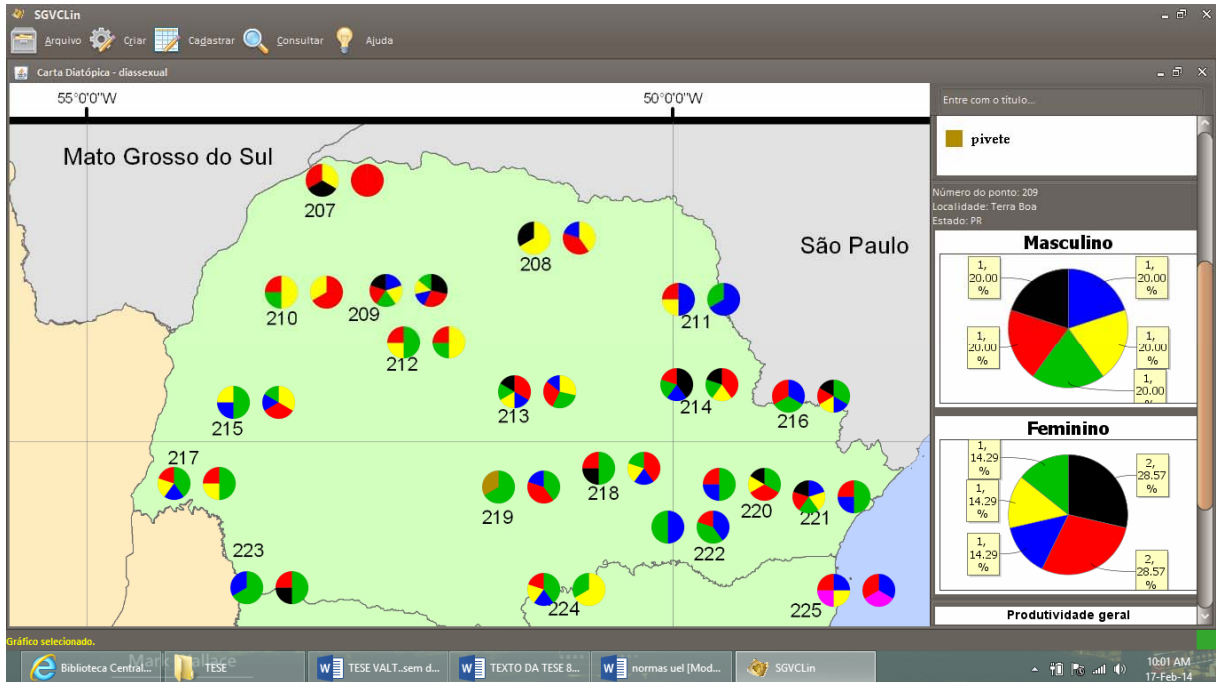
Figura 47 - Exemplo de carta diatópica gerada para impressão



Fonte: [ΣGVCLin]

No caso das cartas bidimensionais, os gráficos são exibidos lado a lado. Na carta diatópica/diassexual, o gráfico da esquerda representa as respostas do perfil masculino e o da direita, as do perfil feminino. Na carta diatópica/diageracional, o gráfico da esquerda representa as respostas do perfil jovem, enquanto o da direita, as do perfil idoso. A Figura 48 ilustra o processo de geração de uma carta diatópica/diassexual. Ao selecionar um gráfico, o par correspondente a ele é automaticamente identificado. A partir daí, qualquer movimento aplicado pelo usuário será atribuído ao conjunto seleção (par de gráficos).

Figura 48 - Exemplo de carta diatópica/diassexual

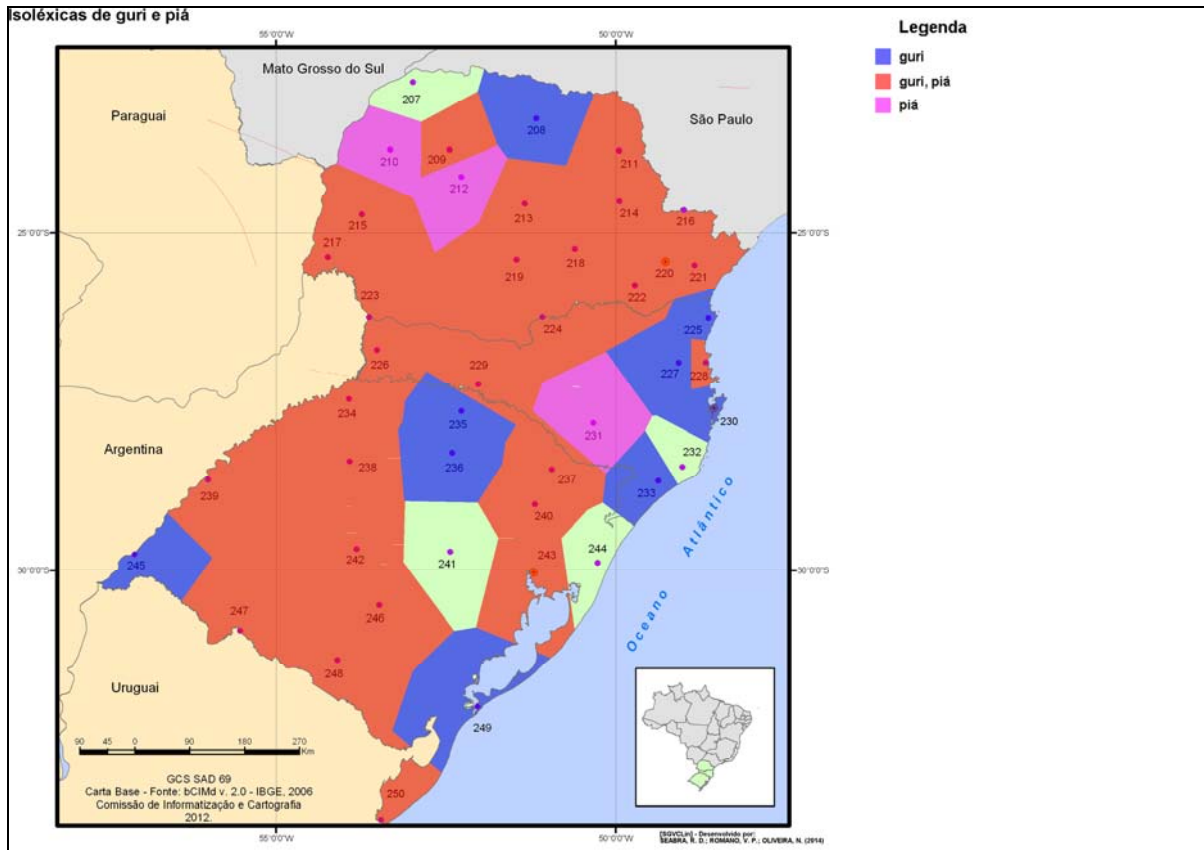


Fonte: [ΣGVCLin]

A última funcionalidade disponível na opção cartas refere-se à geração de cartas de arealidade. O usuário pode acessar a opção de geração dessas cartas e selecionar quais variantes deseja visualizar. A Figura 49 mostra a arealidade das variantes *guri* e *piá* na região Sul do Brasil. Conforme a legenda, as regiões delimitadas pela cor azul indicam as ocorrências da variante *guri*, a variante *piá*, por sua vez, está representada pela cor magenta. Finalmente, as regiões em vermelho apontam as áreas do mapa em que ocorrem as duas variantes concomitantemente, as chamadas áreas de heteroglossas, conforme a nomenclatura de Chambers; Trudgill (1994, p. 140).

Durante sua execução, o [ΣGVCLin] permite o processamento de até quatro variantes simultaneamente, selecionadas a partir de uma lista de opções. O usuário também pode alterar as cores da legenda clicando com o botão direito do mouse sobre a cor que deseja alterar. Em seguida, basta selecionar a opção 'alterar cor'.

Figura 49 - Exemplo de carta isoléxica



Fonte: [ΣGVCLin]

Para o estabelecimento dos limites de isoglossas, foram considerados aspectos do método dialectométrico no que se refere ao estabelecimento de um interponto entre duas localidades, que, segundo Goebel (1981 e 1983), “constitui-se um ponto médio dos segmentos que une cada par de pontos” (GOEBL apud ALTINO, 2012, p. 823). Vale lembrar que não foi empreendido o método dialectométrico na elaboração de cartas de arealidade do [ΣGVCLin], nem este foi o objetivo, apenas foi utilizado o conceito de interponto para o desenvolvimento de um algoritmo de reconhecimento da área de abrangência de determinada variante.

3.3 ABORDAGEM ESTATÍSTICA

Uma vez fornecidos relatórios diversos pelo [ΣGVCLin], os dados foram submetidos a um tratamento estatístico no software SPSS⁶⁷ (*Statistical Package for the Social Sciences*) versão 17.0 (SPSS, 2008), o que possibilitou estabelecer as análises descritivas e inferenciais pertinentes ao estudo.

Ademais, nos estudos lexicais, carecem pesquisas que abordam tratamentos estatísticos sofisticados na análise dos dados adentrando à formulação de hipóteses, e não somente pautando-se na descrição de percentuais em função de frequências absolutas e relativas. Aspectos teóricos estatísticos relevantes para o estudo da variação linguística, por exemplo, podem ser encontrados em Oliveira (2009) e mais informações sobre o tratamento estatístico em dados geolinguísticos podem ser obtidas em Romano e Seabra (2014b).

A análise estatística dos resultados obtidos constitui um importante instrumento na validação desses dados. Resumidamente, o objetivo consiste em extrapolar os resultados da análise para a população estudada, neste caso, os informantes representativos do falar de cada região brasileira. Este trabalho apresenta alguns testes de hipóteses explorados visando a inferir conclusões sobre esta população. Na prática, os testes de hipóteses são utilizados para realizar inferências estatísticas. No caso das inferências, por meio do intervalo de confiança, busca-se delimitar o parâmetro populacional desconhecido. Assim, formula-se uma hipótese quanto ao valor do parâmetro e, com base nos elementos amostrais, aplica-se um teste que indicará a aceitação ou a rejeição da hipótese formulada. Para cada uma das questões estudadas nesta tese, foram formuladas hipóteses específicas para cada caso, de acordo com as análises empreendidas nas cartas linguísticas.

Em um teste de hipóteses, podem-se manipular duas categorias de hipóteses: a hipótese nula e a hipótese alternativa. A primeira refere-se à hipótese que se deseja provar como sendo verdadeira, ao passo que a segunda consiste na negação da primeira. O teste pode conduzir à aceitação ou à rejeição da hipótese nula, que corresponde, respectivamente, à negação ou à afirmação da hipótese alternativa. Neste âmbito, encontra-se aplicado o nível de significância de um teste, que consiste na probabilidade máxima de rejeitar a hipótese nula. Ou seja, o nível de significância de 5% (amplamente utilizado em pesquisas experimentais) corresponde

⁶⁷ Software de análise estatística amplamente utilizado em pesquisas experimentais aplicadas nas mais variadas áreas do conhecimento humano.

à probabilidade máxima de se rejeitar a hipótese nula quando, de fato, ela é verdadeira, caracterizando o erro do tipo I (MAGALHÃES; LIMA, 2001).

As variáveis numéricas, ou seja, o número de ocorrências de cada variante, foram expressas em função da média e desvio padrão. Quanto às variáveis categóricas, aqui representadas pelas variantes lexicais, foram utilizadas, para sumariá-las, frequências simples e relativas (números e porcentagens).

As análises que foram realizadas são:

- **Descritivas:** frequência, média e desvio-padrão;
- **Avaliação da aderência das variáveis à distribuição normal:** o teste de Kolmogorov-Smirnov (KS) foi utilizado para testar a hipótese de normalidade na distribuição. Valores de p-valor maiores que 0,05 indicam que a variável tem adesão à distribuição normal (COSTA NETO, 1977);
- **Comparação de médias entre grupos:** a comparação foi realizada considerando a adesão das variáveis à distribuição normal. Para as questões analisadas nesta oportunidade, como nem todas as variáveis não apresentaram adesão a essa distribuição, foram utilizados: (i) Teste de Mann-Whitney, para comparar dois grupos independentes; (ii) Teste de Kruskal-Wallis, para comparar mais que dois grupos independentes.

Os resultados serão considerados estatisticamente significantes se p-valor < 0,05, com intervalo de confiança de 95%.

3.3.1 Tratamento dos Dados⁶⁸

⁶⁸ Dada a extensão do trabalho, considerando-se os objetivos e a metodologia adotada para o tratamento das questões selecionadas, optou-se em restringir a análise dos dados à dimensão diatópica da variação. A discussão dos resultados sobre a variação diagenérica e diageracional acrescentaria um número substancial de páginas ao Capítulo 4, e embora de extrema importância para inferências acerca da variação e da mudança linguística, não contribuiria diretamente aos objetivos traçados, tocando “os veios sociolinguísticos” (CARDOSO, 2010). Para as análises e para a defesa da existência de áreas lexicais no território do subfalar sulista, delimitado por Nascentes (1953), optou-se estritamente pela variação diatópica associada a aspectos sócio-históricos. Todavia, os resultados atinentes às variáveis sexo e faixa etária poderão ser explorados em outra oportunidade, mais especificamente em futuros trabalhos, como o de pós-doutoramento.

Considerando os objetivos delimitados para o estudo, a apresentação e a análise dos dados compreendem duas abordagens: a primeira, em que se empreende o tratamento da estatística descritiva e a segunda em que se fazem testes estatísticos para a comprovação de hipóteses.

A primeira parte da análise apresenta a estatística descritiva em números absolutos e relativos a partir dos relatórios e das cartas linguísticas, bem como considerando as áreas de ocorrência das variantes documentadas no *corpus* e a sua correlação à sócio-história das localidades/região em que ocorrem.

Tanto os relatórios quanto as cartas linguísticas, referenciadas no capítulo 4, constam do volume 2 deste trabalho, seguindo a ordem numérico-alfabética de identificação. Por exemplo, o relatório 1 apresenta o conjunto geral das respostas para a primeira questão analisada e o relatório 1a refere-se a essa mesma questão, porém, considerando a produtividade por Estado. Para os relatórios, então, há dois tipos, o relatório 1, relatório 1a, o relatório 2, relatório 2a e, assim, sucessivamente.

Quanto às cartas, a carta de número 1 documenta dados da primeira questão analisada e as cartas decorrentes desta primeira são as que apresentam isoléxicas e/ou heteroléxicas evidenciando áreas lexicais, também identificadas com letras do alfabeto. Ex.: Carta 1; Carta 1a; Carta 1b; Carta 1c e, assim, progressivamente, conforme as especificidades apresentadas pelo conjunto de respostas documentado.

Vale salientar que a análise das questões não segue a ordem em que se apresenta no QSL, uma vez que não se trabalha com uma área semântica específica. Assim, as questões são analisadas de acordo com o levantamento e revisão dos dados. Ressalte-se também que esta primeira análise vale-se de gráficos, tabelas, cartas linguísticas e da comparação dos itens estudados aos documentados em outros trabalhos de cunho geolinguístico que contemplam pelo menos uma parte do território do falar sulista. Buscam-se também, nesta primeira análise, informações lexicográficas pertinentes ao estudo em dois dos principais dicionários da Língua Portuguesa: Houaiss e Villar (2001) e Ferreira (2004). Quando necessário, para esclarecer informações sobre a etimologia de determinada variante, recorre-se ao dicionário de Cunha (2010).

A segunda parte da análise corresponde ao tratamento estatístico realizado no programa SPSS a partir dos resultados obtidos. Realizam-se, assim,

testes de hipóteses para os itens de acordo com a análise. As hipóteses são formuladas para cada questão, seguindo a ordem em que foram elaboradas no sentido de provar a sua veracidade. Variantes que apresentaram baixa produtividade (pouco difundidas) ou formas lexicais que ocorrem exclusivamente em determinada área não necessitaram do emprego de testes de hipóteses. São 21 as hipóteses testadas, referentes a cinco questões focalizadas, as questões 001, 177, 156, 132 e 039, todas pertencentes ao QSL.

- Hipótese nula 1: a variante *córrego* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 2: a variante *riacho* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 3: a variante *riozinho* apresenta distribuição de produtividade heterogênea nos Estados de SP, PR, SC e RS.
- Hipótese nula 4: a variante *ribeirão* apresenta distribuição de produtividade heterogênea nos Estados de MG, SP, PR e SC.
- Hipótese nula 5: a variante *geleia* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 6: a variante *chimia/michia* apresenta distribuição de produtividade heterogênea nos Estados do PR, SC e RS.
- Hipótese nula 7: “*formas genéricas*” apresentam distribuição de produtividade heterogênea nos Estados do PR e SC.
- Hipótese nula 8: a variante *bolinha de gude* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 9: a variante *bolita* apresenta distribuição de produtividade homogênea nas regiões Centro-Oeste e Sul.
- Hipótese nula 10: a variante *bolita* apresenta distribuição de produtividade homogênea nos Estados de MT, MS e RS.
- Hipótese nula 11: a variante *búrica* apresenta distribuição de produtividade heterogênea nos Estados de SP, PR e SC.
- Hipótese nula 12: a variante *menino* apresenta distribuição de produtividade heterogênea em todos os Estados.

- Hipótese nula 13: a variante *moleque* apresenta distribuição de produtividade heterogênea nos Estados de SP, PR, MG, GO, MS e MT.
- Hipótese nula 14: a variante *guri* apresenta distribuição de produtividade homogênea nos Estados do RS, MS e MT.
- Hipótese nula 15: a variante *garoto* apresenta distribuição de produtividade homogênea nos Estados de SP, PR, MS, MG e GO.
- Hipótese nula 16: a variante *piá* apresenta distribuição de produtividade heterogênea nos Estados do PR, SC e RS.
- Hipótese nula 17: a variante *mexerica* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 18: a variante *poncã* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 19: a variante *tangerina* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 20: a variante *bergamota* apresenta distribuição de produtividade homogênea nos Estados de SC e RS.
- Hipótese nula 21: a variante *mimosa* apresenta distribuição de produtividade heterogênea nos Estados do PR e SC.

CAPÍTULO 4

DESCRIÇÃO E ANÁLISE DOS DADOS

Neste capítulo, apresentam-se a descrição e análise das variantes lexicais registradas para cinco questões do QSL: Questão 001 – *córrego*; 177 – *geleia*; 156 – *bolinha de gude*, 132 – *menino* e 039 - *tangerina*. Cada questão é apresentada e analisada isoladamente considerando-se os objetivos traçados para este trabalho, uma vez que elas apresentam especificidades que são tratadas individualmente, seja com a criação de rótulos para alguns itens, seja com agrupamento de variantes quando necessário.

Não obstante, são realizados testes de hipóteses específicos para cada caso. Desse modo, nas próximas seções, são apresentados os resultados com o foco principal na distribuição diatópica dos itens lexicais registrados. Para a primeira questão – *córrego* – apresenta-se um modelo de análise, com detalhes sobre o tratamento estatístico empreendido. Para as questões subsequentes, com a finalidade de reduzir informações estatísticas que não são detalhadas considerando-se o foco da análise, são ocultadas as tabelas, abordando-se apenas os dados relevantes para a descrição das hipóteses.

4.1 QUESTÃO 001 - CÓRREGO

A questão 001 do QSL, pertencente à área semântica *Acidentes Geográficos*, 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, como, por exemplo, as variantes *rio*, *rio pequeno*, *rio estreito* e *rio raso*.
- **“formas inadequadas”**: contempla os itens lexicais que, apesar de possuírem traços semânticos comuns ao referente, não designam, especificamente, o item buscado, como, por exemplo,

as formas *lagoa*, *lago*, *poço*, *açude*. Essas variantes denominam outro referente, que tem como traço característico a presença de água parada, represada.

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

Além dos rótulos criados, foram necessários agrupamentos de variantes morfofonêmicas⁶⁹, como é o caso de itens em que houve:

- (i) redução da proparoxítone (*córrego* > *corgo*);
- (ii) formas no diminutivo (*corgo* > *corguinho*; *riacho* > *riachinho/riachozinho*; *sanga* > *sanguinha*, *rego* > *reguinho*, *valeta* > *valetinha*)
- (iii) formas que apresentam o mesmo radical (*valo*, *vala*, *valão*)
- (iv) 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). 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 foi questionado sobre tal designativo.

Ainda para a análise das respostas, quando necessárias as informações sobre a etimologia dos vocábulos, recorreu-se também ao Dicionário Etimológico da Língua Portuguesa, de Antônio Geraldo da Cunha (2010).

Assim, realizados os devidos agrupamentos, o Relatório 1 (vol. 2) apresenta os 644 registros, distribuídos em 14 itens para essa questão.

As cinco variantes mais produtivas no *corpus* são: *córrego* e suas variantes morfofonêmicas (261 ocorrências – 40,53%), *riacho* e variantes (122

⁶⁹ 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.”

ocorrências – 18, 94%), *riozinho* (82 – 12,73%), *ribeirão* (38 ocorrências – 5,90%) e *sanga/sanguinha*, com 37 registros (5,75%). Juntas essas cinco variantes representam 83,85% das respostas.

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

Com menor índice de produtividade, encontram-se oito “respostas inadequadas” (*lago, lagoa, açude e poço*). A variante *corixo* obteve apenas cinco registros, representando menos de 1% do *corpus* (0,78%). Ainda em menor produtividade, 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%. Somam-se a esse conjunto três abstenções de resposta e seis inquéritos em que houve problema técnico, ou por falha na gravação ou por falta de exploração adequada da resposta obtida.

A partir da carta 1 (vol. 2) é possível visualizar a distribuição diatópica de seis das 14 formas registradas no *corpus*. Embora o [ΣGVCLin] 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 que não haja comprometimento na distinção das cores nos gráficos de pizzas. As demais formas lexicais estão agrupadas no item *outras*, representadas pela cor cinza. Desse modo, este item da legenda contempla as formas sob o rótulo *sugestão na pergunta, rego* e suas variantes morfofonêmicas, *arroio, formas inadequadas, canal, grota d’água* e as *hápx legomena: bocaina, braço de rio e lajeado*.

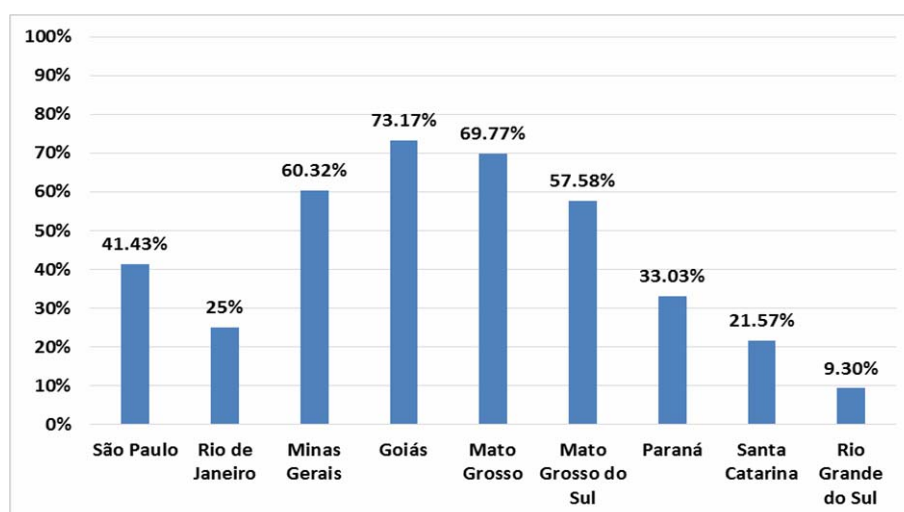
Conforme se observa na carta 1, há uma concentração maior da variante *córrego* com suas variantes morfofonêmicas⁷⁰ em localidades situadas no Estado de São Paulo (não ocorrendo em 5 dos 38 pontos paulistas), e sua presença é registrada em todos os pontos dos Estados selecionados no Rio de Janeiro, Minas Gerais, Mato Grosso do Sul, Mato Grosso e Goiás. À medida que se avança no

⁷⁰ Para tratar das formas agrupadas, ao longo do texto, opta-se pelo uso da forma padrão. Assim, ao se usar o vocábulo *córrego*, necessariamente, está se referindo às suas variantes morfofonêmicas: *corgo, corquinho*. Esse mesmo critério se aplica às outras variantes/questões e, conseqüentemente, para os títulos das respectivas cartas linguísticas.

sentido meridional do Brasil, *córrego* vai perdendo produtividade, apesar de ainda continuar sendo a forma mais frequente no Paraná (33,03%), conforme se observa no Relatório 1a (vol. 2). Já em Santa Catarina, a variante *córrego* encontra-se como a terceira mais produtiva (21,57%) e no Rio Grande do Sul está em quinto lugar (9,30%).

Partindo dos valores constantes do referido relatório, a Figura 50 apresenta a produtividade da variante *córrego* por Estado. Vale salientar que estes números são proporcionais ao total de ocorrências das variantes no Estado, bem como ao número de localidades investigadas, ou seja, para o Estado do Rio de Janeiro, cujos dados referem-se a duas localidades (ponto 205 - Barra Mansa e 206 - Paraty), os 25% calculados têm como parâmetro apenas os oito informantes dessas duas localidades (quatro em cada ponto).

Figura 50 - Produtividade da variante *córrego* por Estado



Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

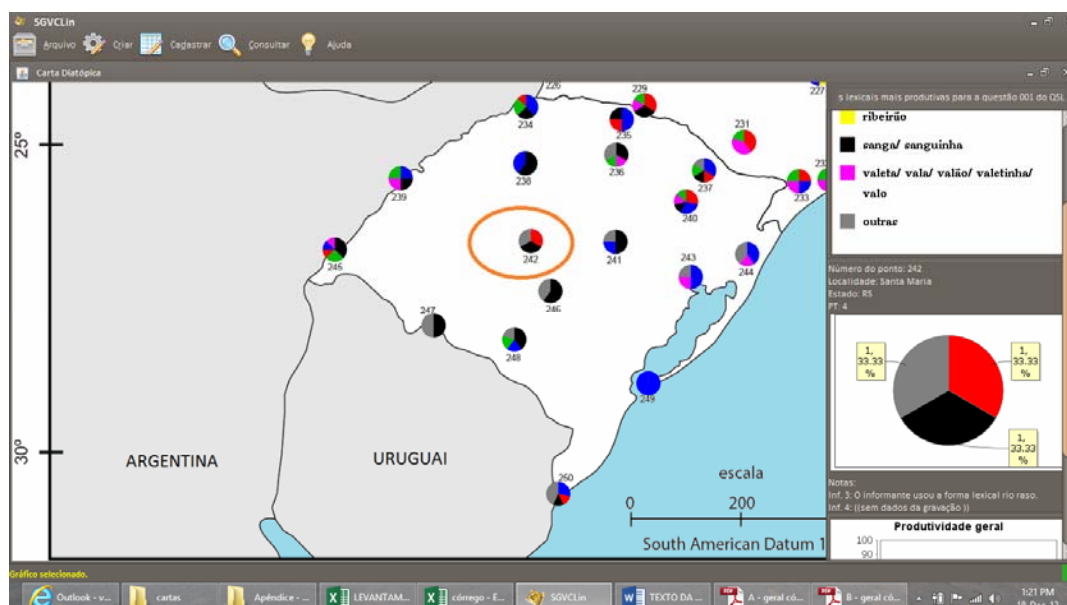
É interessante notar a partir deste gráfico que o índice de ocorrência da forma mais difundida entre os informantes – *córrego* – diminui nos Estados da região Sul (PR, SC e RS), o que, como será descrito posteriormente, se justifica pela presença de variantes regionais nestes Estados, ao passo que, nos Estados da região Centro-Oeste, o item obteve produtividade acima de 50%. Pela carta 1a, verifica-se que *córrego* contempla grande área do território investigado, principalmente, em Estados das regiões Sudeste e Centro-Oeste. Na região Sul, a variante abrange grande parte do território do PR e de SC. Nota-se, entretanto, que

em cada um dos Estados em que ocorre esta variante, há diferentes índices de produtividade, conforme se verifica no Relatório 1a.

De acordo com a carta 1a, observa-se que, no Estado de São Paulo, essa variante não está registrada apenas em cinco pontos: (174 – Bragança Paulista, 175 – Taubaté, 182 – Capão Bonito, 183 – Itanhaém e 187 – Cananéia). No Paraná, somente em duas localidades não há a ocorrência de *córrego* (ponto 218 – Imbituva e 223 – Barracão). Em Santa Catarina, esta variante não ocorre nos pontos 227 – Blumenau e 230 – Florianópolis.

Nas localidades gaúchas, *córrego* obtém pouca produtividade, apresentando-se em uma ocorrência nos pontos linguísticos do extremo norte do Estado (pontos 234 – Três Passos, 235 – Erechim, 237 – Vacaria), em uma localidade do centro do RS – 242 (Santa Maria) e nos pontos 245 – Uruguaiana e 250 – Chuí, no extremo sul do país. Apenas em Flores da Cunha – ponto 240 – *córrego* apresenta-se com duas ocorrências. Em Santa Maria, ponto 242, essa forma lexical representa 33,33% das respostas⁷¹, conforme se visualiza com mais detalhes na Figura 51.

Figura 51 - Zoom do Rio Grande do Sul - foco na produtividade do ponto 242 (Santa Maria) carta 1

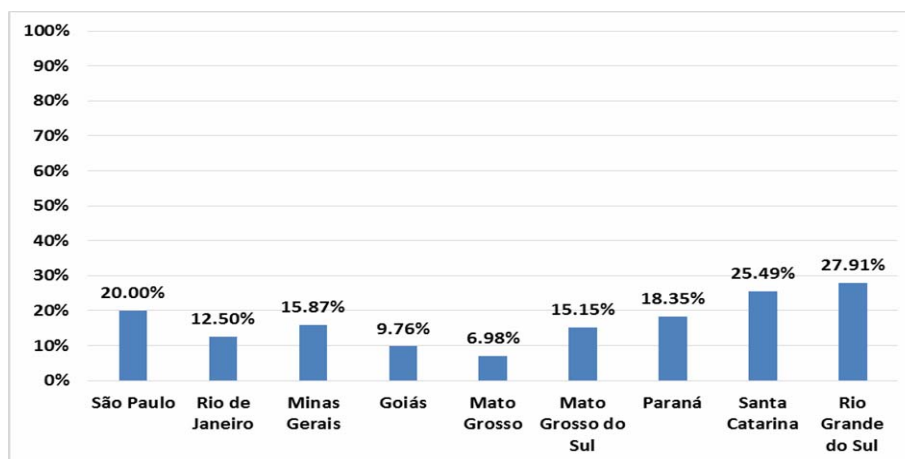


Fonte: [SOVCLin]

⁷¹ Cabe lembrar que, por motivos técnicos (falhas na gravação), não foram contemplados os dados referentes à informante 4 (mulher da segunda faixa etária).

A segunda variante mais produtiva nos dados analisados é *riacho* que representa, juntamente com as formas morfofonêmicas, 18,94% do *corpus*. Esta variante obtém diferentes percentuais em cada Estado, conforme se observa na Figura 52, sendo o Estado do Rio Grande do Sul e Santa Catarina os que apresentam maior percentual, 27,91% e 25,49%, e Mato Grosso e Goiás os que figuram com os menores índices, 6,98% e 9,76%, respectivamente.

Figura 52 - Produtividade da variante *riacho* por Estado



Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

Quanto à distribuição diatópica, observa-se que a forma lexical *riacho* teve uma abrangência esparsa contemplando todos os Estados que fazem parte desta pesquisa (Carta 1b). A maior concentração de *riacho* está nos Estados da região Sul (PR, SC e RS) e no Estado de SP, conforme se visualiza na Figura 52. Nota-se, pela carta 1b, que no MT esta variante ocorre em duas localidades isoladas (pontos 110 – Cáceres e 104 – São Félix do Araguaia), portanto, área descontínua. Em GO, a variante prossegue do sul para o centro do Estado contemplando quatro cidades goianas (pontos 124 – Jataí, 126 – Quirinópolis, 123 – Goiânia e 125 – Catalão). Ainda vindo de Goiás, esta isoléxica atinge dois pontos de controle no território mineiro, pontos 130 – Unai e 132 – Pirapora e, procedente do nordeste paulista, Vale do Paraíba e sul de MG, a abrangência de *riacho* atinge o ponto de controle 145 – São João Del Rei (MG) e 205 – Barra Mansa (RJ).

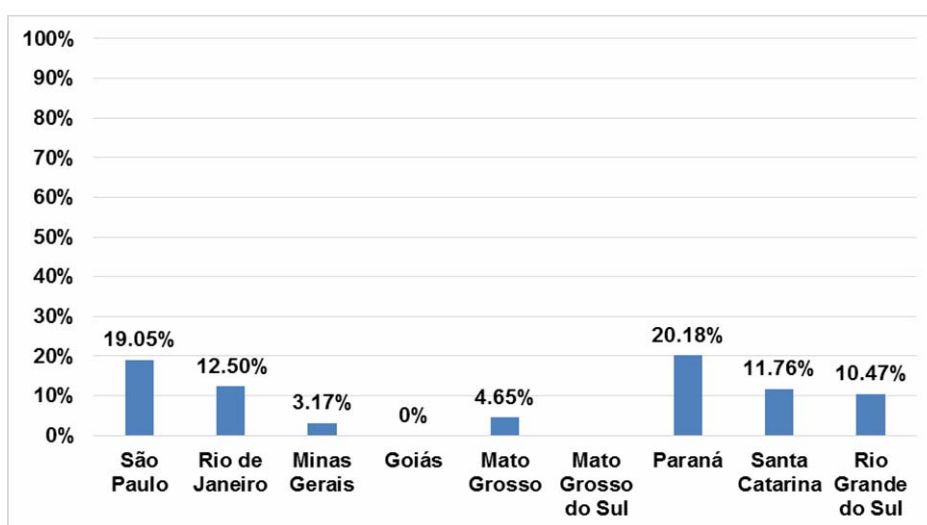
Riacho não ocorre em 11 das 38 localidades paulistas (cinco no noroeste), quatro no centro do Estado e duas no vale do Paraíba – ponto 174 (Bragança Paulista) e 176 (Guaratinguetá). A área de isoléxica vinda do oeste paulista adentra-se o MS contemplando dois pontos de inquérito, Campo Grande

(ponto 115) e Nioaque (116). *Riacho* ainda ocorre na região do Vale do Ribeira de onde se irradia para o litoral e interior do PR. Em território paranaense, essa variante não está registrada em quatro cidades, duas no norte do Estado (pontos 209 – Terra Boa e 210 – Umuarama), uma no centro – ponto 213 (Cândido de Abreu) e uma no oeste paranaense – ponto 217 (São Miguel do Iguaçu).

A área de ocorrência de *riacho* avança rumo à fronteira política com SC, contemplando o litoral norte: ponto 225 – São Francisco do Sul, 227 – Blumenau e 228 – Itajaí, além da capital – 230 (Florianópolis). No norte catarinense, *riacho* ocorre em Porto União (ponto 224) e pelo oeste – ponto 226 (São Miguel do Oeste) expande-se ao RS. Desse modo, o interior catarinense não é contemplado com a variante *riacho* que prossegue no RS na porção oeste e leste ligadas por uma localidade do norte rio-grandense (ponto 235 – Erechim). Essa forma lexical não ocorre em quatro localidades do interior gaúcho. Desse modo, compreende-se que *riacho* é uma forma lexical comum a diferentes áreas do território investigado, ocorrendo, inclusive, em pontos de controle.

A variante *riozinho*, comum a algumas áreas, representa 12,73% das respostas, ocorrendo em sete dos nove Estados contemplados pela pesquisa. Quantitativamente, esta forma lexical é mais significativa no Estado de SP e PR, conforme se observa na distribuição percentual do item por Estado (Figura 53).

Figura 53 - Produtividade da variante *riozinho* por Estado



Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

A carta 1c mostra que *riozinho* é uma forma lexical que se distribui pelo Estado de SP, adentra o PR e, a partir do interior de SC, atinge o norte e oeste

do RS, além de um ponto isolado no extremo sul do país – 248 (Bagé). Em MG, a variante não avança além de um ponto localizado no sul do Estado, 149 – Itajubá. No MT, há ocorrência de *riozinho* em dois pontos isolados – Aripuanã (ponto 103), no noroeste, e Cáceres – ponto 110 – situado no sudoeste mato-grossense.

É interessante notar que *riozinho* não está presente nas localidades sul-mato-grossenses e goianas, indicando, provavelmente, que a forma não tenha se propagado do Estado de SP para os Estados limítrofes, haja vista a influência dos paulistas nos Estados circunvizinhos, conforme já apontado no Capítulo 1, por meio de movimentos de ocupação humana, principalmente as Bandeiras dos séculos XVI e XVII. Pressupõe-se que se esta variante tivesse SP como foco de irradiação, haveria ocorrência em quantidade significativa em todo o sul de MG e Triângulo Mineiro, bem como no MS e no sul de GO.

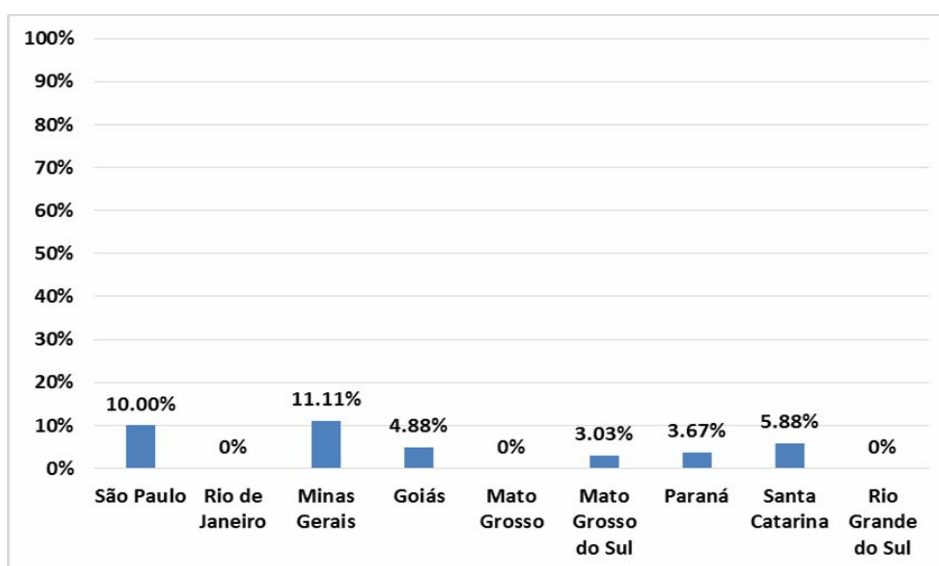
Para confirmar essa premissa, verifica-se que o Atlas Linguístico do Mato Grosso do Sul (OLIVEIRA, 2007) na carta intitulada *córrego* (p. 93) documenta apenas uma ocorrência de *riozinho* no extremo sul do Estado (fronteira política com o PR) no município chamado Iguatemi, o que permite afirmar que, possivelmente, esta forma tem como foco irradiador o Estado do Paraná, estendendo-se, a partir deste Estado, tanto para o interior de SP quanto para SC, e pelo interior catarinense atinge o norte rio-grandense e outros pontos.

O Atlas Linguístico do Paraná (AGUILERA, 1994), na carta nº 2, não documenta a variante *riozinho*, porém, o Atlas Linguístico Etnográfico da Região Sul do Brasil (ALTENHOFEN; KLASSMANN, 2011), na carta 007, apresenta esta forma como a quinta variante mais produtiva, representando 9% das respostas em toda a região Sul do Brasil. No ALERS, *riozinho* está presente em 14 localidades paranaenses, em três localidades catarinenses e em uma localidade gaúcha. No PR, a variante ocorre, principalmente, no oeste do Estado, além de pontos isolados no norte, leste e centro. Assim, pela carta do ALERS, verifica-se que *riozinho* representa, aproximadamente, 18% do *corpus* paranaense, 7% do catarinense e 1% do sul-rio-grandense. Analogamente, este panorama se repete no *corpus* do ALiB, porém com valores diferentes. Assim, nota-se que no PR o índice de ocorrência de *riozinho* é maior (20,18%), decrescendo em SC (11,76%) e no RS – 10,47% (Figura 53). Seria *riozinho* uma possível variante que evidenciaria a presença de um *falar paranaense* na área investigada, haja vista sua produtividade no PR e distribuição

diatópica em Estados circunvizinhos (SC, MS e SP)? Em busca desta e de outras respostas é que se empreendem testes estatísticos e análises de outras questões.

Ribeirão é a quarta variante no universo das respostas, representando, no entanto, 5,94%. Observa-se que esta forma lexical não ocorre em três dos nove Estados investigados e obtém diferentes índices percentuais nos demais, conforme se verifica na Figura 54.

Figura 54 - Produtividade da variante *ribeirão* por Estado



Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

A variante é mais produtiva nos Estados de MG e SP, seguindo-se SC. Nos demais (GO, PR e MS), a produtividade foi menor que 5% das respostas. Apesar de a representatividade da variante em números absolutos e relativos ser baixa, pela carta de arealidade 1d é possível acompanhar a sua distribuição diatópica na rede de pontos desta pesquisa. Partindo do norte pioneiro do PR – ponto 211 (Tomazina) – *ribeirão* segue em um feixe que perpassa o Estado de SP, do sudoeste ao nordeste, contemplando cidades do Vale do Ribeira (pontos 181 – Itararé, 182 – Capão Bonito, 177 – Itapetininga), passando pelo interior do Estado e atingindo algumas cidades do Vale do Paraíba (pontos 168 – Mogi Mirim, 174 – Bragança Paulista, 175 – Taubaté e 176 – Guaratinguetá). Em território mineiro, esta forma lexical se espalha pelo sul alcançando dois pontos de controle (pontos 145 – São João Del Rei e 130 – Unaí), além de três cidades do noroeste paulista (pontos 150 – Jales, 153 – Barretos e 154 – Franca).

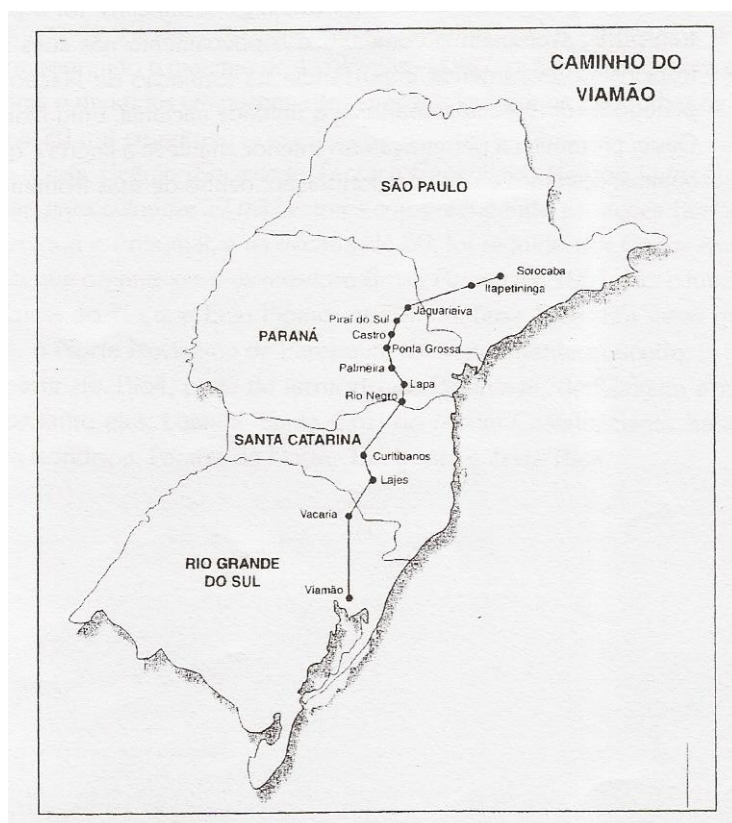
Nota-se também que *ribeirão* ocorre em dois pontos isolados de GO (118 – Porangatu e 126 – Quirinópolis) e em um ponto do MS (112 – Coxim). Na região Sul do país, a variante ocorre em dois pontos do noroeste paranaense (210 – Umuarama e 212 – Campo Mourão) e em duas localidades do nordeste de SC (pontos 227 – Blumenau e 228 – Itajaí).

Observando-se a carta lexical 2 do ALPR (AGUILERA, 1994), verifica-se que *ribeirão* é uma forma que ocorre em grande parte do Estado, com exceção do noroeste, do sul e oeste paranaense. Na região do Norte Central e do Norte Pioneiro, a variante ocorre, por exemplo, em Londrina, Cambará e Ibaiti. Na carta 007 do ALERS (ALTENHOFEN; KLASSMANN, 2011), a área de ocorrência de *ribeirão* coincide com a do ALPR, no que se refere ao nordeste do PR. Ainda no ALERS, a ocorrência de *ribeirão* em cidades catarinenses localizadas na região de Blumenau coincide com o registrado no *corpus* do ALiB (cartas 1 e 1d).

A partir da análise da carta do ALERS, no que se refere ao PR, e da carta do ALPR, em contraste com as isoléxicas encontradas na carta 1d (*corpus* do ALiB), identifica-se que a variante *ribeirão*, possivelmente, teria se irradiado do sul de MG, uma vez que é a região do falar sulista que apresenta maior produtividade e por meio de um feixe que corta o Estado de SP, chega ao norte do PR. Sabe-se que o Estado do PR, sobretudo, o norte paranaense, recebeu a influência direta de colonizadores mineiros e paulistas durante o seu processo de formação e, possivelmente, essa seria uma das muitas marcas deixadas no léxico dessa região do Estado. No ALERS e no ALPR, a área de ocorrência de *ribeirão* amplia-se ainda mais, abrangendo outras cidades do Norte Pioneiro paranaense, como Siqueira Campos, Jacarezinho, Santo Antônio da Platina e Londrina, o que, pode confirmar essa premissa sobre a influência de mineiros e paulistas nessa região do Estado.

Sabe-se que a rota dos tropeiros não chegou, pelo menos em um primeiro momento, ao norte do PR, pois, conforme salientado no capítulo 1, a principal rota do tropeirismo no Brasil foi a do denominado caminho de Viamão (Figura 55). Desse modo, os tropeiros saíam de Viamão-RS, cruzavam SC passando por Lajes, Curitibaanos e adentravam o Estado do PR por Rio Negro, Lapa, Palmeira, Ponta Grossa, Castro, Piraí do Sul e Jaguariaíva (STECA; FLORES, 2002) de onde seguiam rumo ao Estado de São Paulo até chegar a Sorocaba, para as feiras de muares.

Figura 55 - Principal rota dos tropeiros – Caminho de Viamão



Fonte: Cigolini; Melo e Lopes (1998, p. 66)

No entanto, *ribeirão* não é uma forma documentada nas principais cidades paranaenses situadas na rota do antigo Caminho de Viamão (pontos 222 – Lapa e 214 – Piraí do Sul) conforme se verifica na carta 1d. Ressalte-se também que a ocorrência de *ribeirão* no Estado de SC, no *corpus* do ALiB, restringiu-se aos pontos 227 – Blumenau e 228 – Itajaí. Na carta 007 do ALERS, além de sua ocorrência no território paranaense, *ribeirão* apresenta-se também em 11 cidades catarinenses localizadas no nordeste do Estado. Portanto, na mesma região em que foi documentada pelo ALiB, inclusive, nos dois pontos de inquérito do atlas nacional.

Uma pesquisa sobre a história das localidades no site do IBGE CIDADES⁷² mostra que os municípios catarinenses em que ocorre a variante *ribeirão* foram colonizados, em sua maioria, por europeus (italianos, poloneses e, principalmente, alemães). Não obstante, buscando aspectos demográficos da população dessas localidades, pelo filtro “Migração”, observa-se que o número de migrantes oriundos do Sudeste do país não é tão significativo naquela região. Isso

⁷² Disponível em: <http://cidades.ibge.gov.br/xtras/uf.php?lang=&coduf=42&search=santa-catarina>. Acesso: 17 dez. 2013.

evidencia que o comportamento de *ribeirão*, nesta área do falar sulista, não se justifica pela premissa levantada acerca da influência de mineiros, ou seja, possivelmente, a variante *ribeirão* em SC não é um reflexo da migração de mineiros ou paulistas para a região; antes, trata-se, no caso em questão, de uma área lexical isolada, assim como se observa no oeste paranaense e nas localidades goianas (carta 1d).

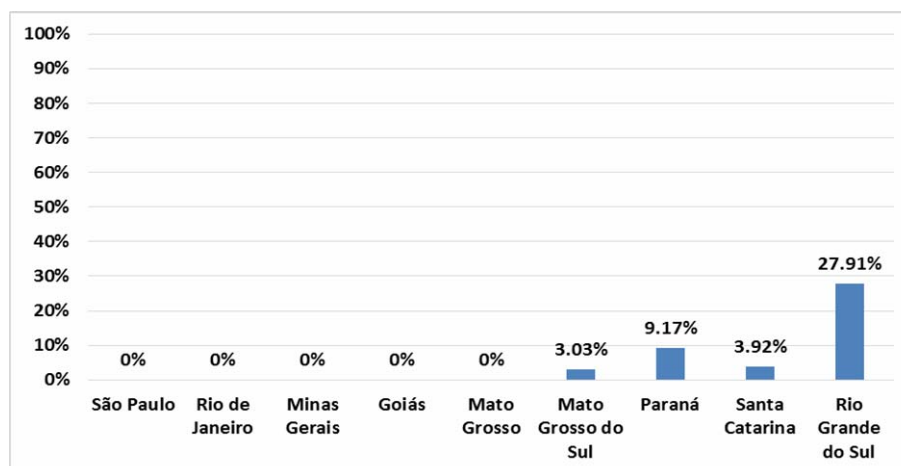
Conforme atestam Margotti e Vieira (2006), essa região em que ocorre *ribeirão* (litoral norte catarinense) seguramente não recebeu influências sul-rio-grandenses (ao contrário do que ocorreu no sul, no planalto e no oeste do Estado), nem de paulistas em razão do tropeirismo. Possivelmente, a presença de *ribeirão* nesta área, inclusive coincidente com o registro na carta do ALERS, deve-se a origens ainda mais antigas. Não seria a presença de *ribeirão* um resquício dos primeiros movimentos migratórios de vicentistas e açorianos deixados no léxico?

Sabe-se que esses grupos foram responsáveis pela primeira ocupação do litoral catarinense, conforme apontado no capítulo 1, e, posteriormente, vieram a ocupar áreas um pouco mais afastadas do litoral, no vale do rio Itajaí-Açu ou no vale do rio Itapocu. Ressalte-se ainda que em Florianópolis encontra-se uma localidade tipicamente açoriana denominada *Ribeirão da Ilha*. Essa seria uma possível explicação para a presença de *ribeirão* nesta região de SC.

Vale lembrar que as regiões em que ocorre esta variante em SC, PR, MS e GO não têm ligação diatópica com a área de isoléxica identificada no Sul de MG que, por um feixe, corta o Estado de SP até atingir o norte do PR, região comprovadamente identificada como uma área do Paraná Moderno, colonizada, em geral, durante o auge do café no Brasil (início do século XX).

A quinta variante mais produtiva entre os informantes é *sanga* que obteve 37 ocorrências, equivalendo a 5,75% do *corpus* de acordo com o relatório geral de produtividade (Relatório 1). Esta variante apresenta-se em quatro dos nove Estados que compõem a pesquisa: MS, PR, SC e RS, com diferentes índices de produtividade conforme se observa na Figura 56.

Dentre esses quatro Estados, a variante é mais produtiva no RS, onde obtém 27,91% de representatividade (24 ocorrências), igualando a *riacho* que também apresenta o mesmo percentual (Relatório 1a).

Figura 56 - Produtividade da variante *sanga/sanguinha* por Estado

Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

No Estado do PR, *sanga* apresenta 10 ocorrências (9,17%) e com menor produtividade encontra-se em SC, com apenas dois registros do item (3,82%) e no MS, com somente uma ocorrência (3,03%).

Essa variante permite o traçado de uma isoléxica (Carta 1e) 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 Iguçu). Embora não prossiga numa faixa contínua (sem corte), a variante *sanga* ocorre também em uma localidade sul-mato-grossense – ponto 117 – (Ponta Porã). No Paraná, *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 de *sanga* no subfalar sulista evidenciam traços da história social da região, sobretudo ao se associar esse fato linguístico ao caminho dos tropeiros, nos séculos XVIII e XIX e, mais recentemente, com a reimigração de gaúchos de origem eslava e germânica.

Conforme salientado no capítulo 1, o Tropeirismo consistiu em um importante movimento econômico e de ocupação humana que contribuiu para o surgimento de inúmeras cidades. 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 da variante *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.

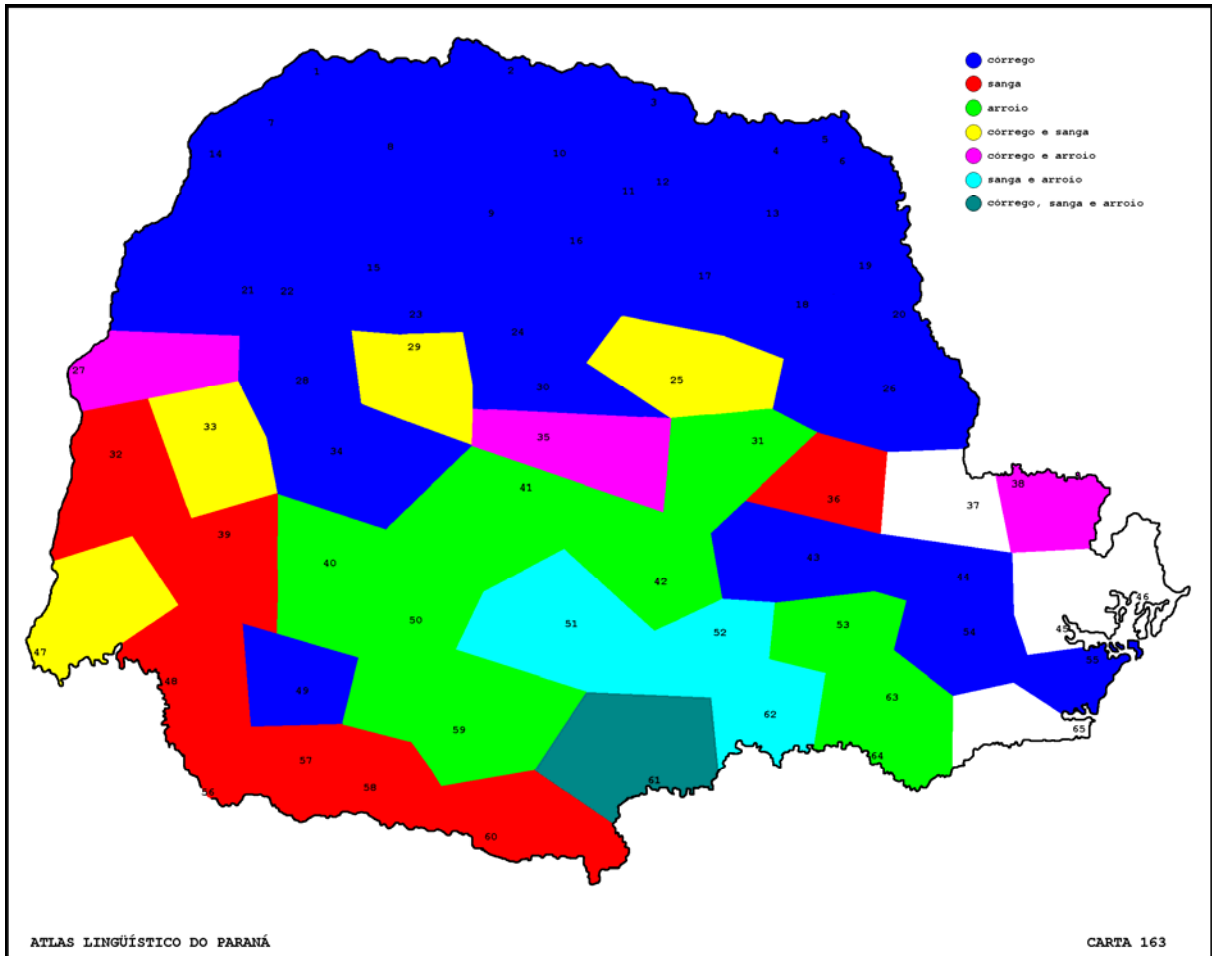
Abandona-se o traçado que cortava o Rio Grande transversalmente (Uruguaiana, Alegrete, Viamão, Santo Antônio da Patrulha, São Francisco de Paula, Bom Jesus, Vacaria), adotando-se o traçado São Borja, Santo Ângelo, Cruz Alta, Carazinho, Passo Fundo, Lagoa Vermelha, Vacaria. Com exceção das duas primeiras que já existiam como povoações missioneiras, as demais citadas surgiram como consequência do novo traçado que ficou conhecido como o Caminho das Missões. Esse caminho, com o mesmo nome, viria a ser modificado pelo tropeiro paranaense Francisco da Rocha Loures, avançando ainda mais na direção do Oeste (...). (BRUM apud ROCHA, 2008, p. 61)

Segundo Ferreira (2004), o verbete *sanga* provém do espanhol platino *zanja* e está dicionarizado, em uma de suas acepções, como um brasileirismo do RS e de SC para designar “um pequeno regato que seca facilmente” (FERREIRA, 2004). Desse modo, a presença de *sanga* nessa região revela um traço 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 ALiB têm mostrado que a variante 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. Porém, cabe notar que apesar de no *corpus* do 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 encontra documentada tal variante, o que não possibilita afirmar, por ora, a influência de gaúchos no MS, pelo menos no que se refere ao uso de *sanga*.

O ALPR (AGUILERA, 1994), na carta 163, apresenta as isoléxicas das variantes para o “rio pequeno” (Figura 57) e no que se refere à variante *sanga* (cor vermelha), nota-se que esta forma lexical ocorre, principalmente, em localidades do sudoeste paranaense, além de outros pontos em que coocorre com outra variante.

Figura 57 - Carta linguística 163 do ALPR

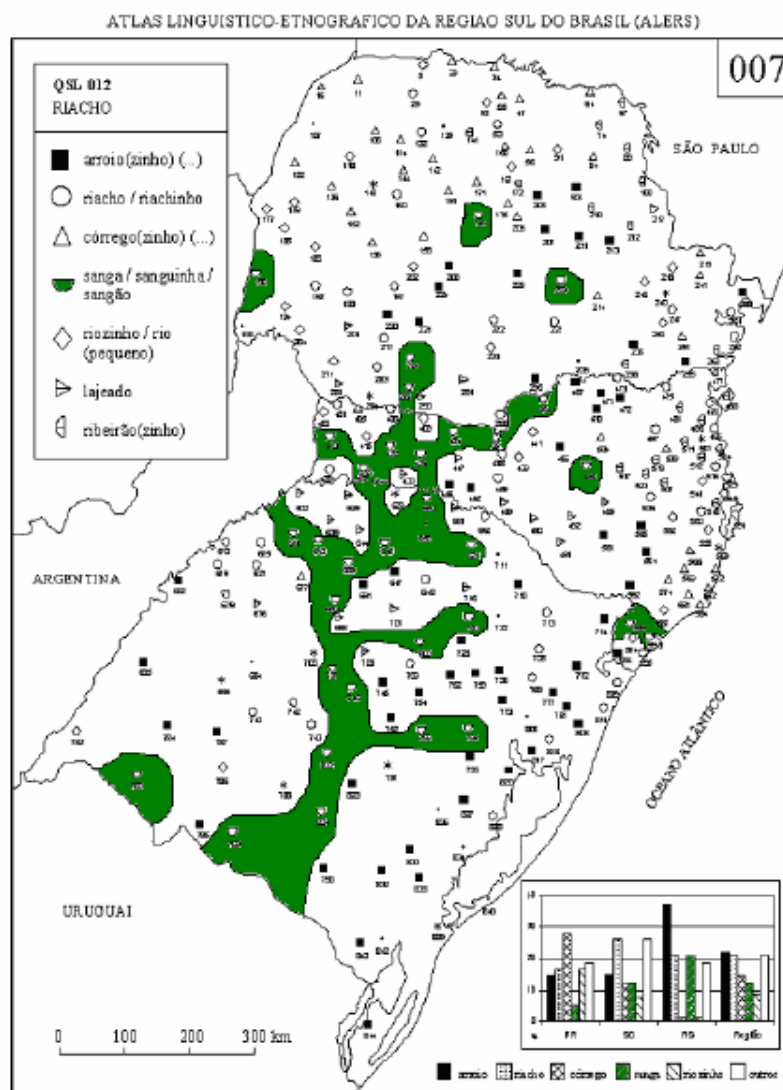


Fonte: Atlas Linguístico do Paraná (AGUILERA, 1994)

Rocha (2008), em estudo realizado a partir de cartas lexicais do ALERS, apresenta a carta linguística do referido atlas (com adaptações) que documenta as variantes lexicais para o “rio pequeno”, com destaque ao item *sanga* (Figura 58), evidenciando também 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 ALiB.

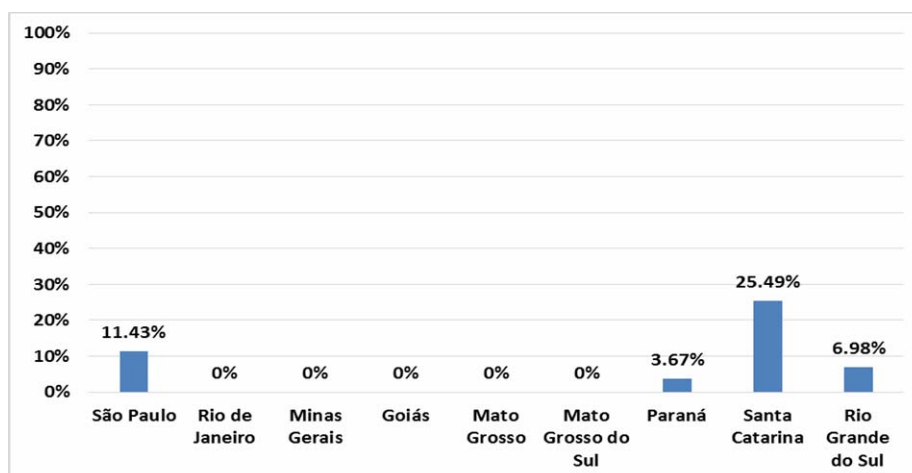
Figura 58 - Carta linguística 007 do ALERS

59



Fonte: Atlas Linguístico Etnográfico da Região Sul do Brasil (adaptação Rocha, 2008, p.59)

O item lexical *valeta* foi documentada como a sexta forma mais produtiva, conforme se observa na Carta 1 e no Relatório 1 (vol. 2). No *corpus*, foram documentadas 26 ocorrências, equivalendo a 4,06% das respostas, presentes em quatro dos nove Estados selecionados pela pesquisa: SP (1,43%), PR (3,67%), SC (25,49%) e RS (6,98%) (Figura 59).

Figura 59 - Produtividade da variante *valeta* por Estado

Fonte: [ΣGVCLin] – Relatório 1a (vol.2)

Conforme se observa, a variante *valeta* foi mais produtiva no Estado de SC, irradiando-se por todo o litoral e adentrando o interior catarinense, atingindo os pontos 231 – Lages, 229 – Concórdia e 224 – Porto União, no norte do Estado, conforme se observa na carta 1f. Essa mesma arealidade alcança o Estado do PR, contemplando o ponto 221 – Morretes, no litoral, e a capital – ponto 220 – Curitiba, além de uma área isolada a oeste – ponto 217 – São Miguel do Iguçu.

No RS, *valeta* apresenta-se com baixa produtividade (6,98%), localizada em duas áreas distintas. A primeira delas vai de Porto Alegre (ponto 243), atingindo o ponto 244 (Osório) no nordeste do Estado, e num feixe rumo ao norte chega a Flores da Cunha (240), e em sentido oeste atinge Passo Fundo – ponto 236. A segunda área de ocorrência localiza-se no sudoeste do Estado, pontos 239 (São Borja) e 245 (Uruguaiana), na fronteira do Brasil com a Argentina. Vale lembrar, no entanto, que essas áreas não são identificáveis no ALERS, uma vez que o referido atlas não documenta esta forma lexical na carta 007.

No Estado de SP, *valeta* ocorre em três localidades também com baixa produtividade: na capital, ponto 179 – São Paulo, no ponto 186 – Registro (região do Vale do Ribeira) e em uma localidade no litoral norte – ponto 180 – Caraguatatuba. É interessante notar que o Atlas Semântico-Lexical da Grande ABC (CRISTIANINI, 2007) registra no cartograma 01 – Córrego – uma ocorrência da variante *valeta* no ponto linguístico Rio Grande da Serra. No trabalho de Encarnação (2010) – Atlas Semântico-Lexical de Caraguatatuba, Ilha Bela, São Sebastião e Ubatuba – municípios do litoral norte de São Paulo – encontram-se duas ocorrências

da variante *valo* (uma em Caraguatatuba e uma em Ubatuba), conforme se visualiza no cartograma 1 do referido trabalho. No entanto, em ambos os trabalhos, considera-se apenas a primeira resposta dos informantes e as variantes lexicais foram registradas cartograficamente sem os devidos agrupamentos ou esclarecimentos sobre o referente. Ressalta-se, entretanto, que os trabalhos de Cristianini (2007) e Encarnação (2010) revelam a presença de *valeta* e variantes na região metropolitana de São Paulo (CRISTIANINI, 2007) e no litoral norte do Estado (ENCARNAÇÃO, 2010), inclusive coincidindo com a mesma localidade selecionada pelo ALiB – Caraguatatuba.

Tanto em Cristianini (2007) e em Encarnação (2010), quanto nos dados do ALiB, *valeta* e variantes apresentam-se de forma pouco produtiva, não podendo se configurar como uma forma regional, mas sim como ocorrências isoladas de ocorrência desta variante no Estado de SP.

Em suma, com base no Relatório 1a, *valeta* encontra-se registrada com diferentes índices em SC, RS, PR e SP, sendo mais produtiva, no entanto, no primeiro desses Estados. Pode-se, de certo modo, afirmar que a forma se irradiou para as demais áreas de forma irregular, figurando além do território catarinense, no litoral paranaense, no território paulista (principalmente no litoral) e em duas áreas gaúchas (no sudoeste – fronteira com a Argentina e na porção leste do RS).

A partir do relatório 1 (vol. 2), observa-se que as formas que obtêm menos de 4% de produtividade no universo das respostas estão agrupadas na categoria *outras*, o que facilita a visualização e representação na carta linguística 1.

Retomando o que já foi dito no início dessa análise, a categoria *outras* contempla as ocorrências de formas agrupadas no rótulo “*sugestão na pergunta*” (*rio, rio raso, rio estreito, rio pequeno*), *arroio, rego* (variantes morfofonêmicas), as “*formas inapropriadas*” (*lago, lagoa, açude*), *corixo, canal*, “*formas pouco produtivas*” (*braço de rio, bocaina e lajeado*) e *grota d’água*. Deste rol de variantes, merecem destaque para discussão, dada a sua distribuição diatópica e isoléxicas que permitem traçar, as variantes *arroio, rego* e *corixo*.

A forma lexical *arroio* apresenta-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 – 5,50%). Assim como a variante *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 desta variante nos Estados da Região Sul (Carta 1g), verificam-se novamente as marcas deixadas no léxico pelo movimento dos gaúchos rumo ao Estado do PR, possivelmente durante o ciclo do Tropeirismo e, posteriormente, da reemigração gaúcha de descendentes alemães, eslavos e italianos.

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 (Piraí do Sul). Nesta distribuição, a variante acompanha os antigos caminhos dos tropeiros, demonstrando mais uma vez as marcas linguísticas deixadas por este movimento humano que se reflete, sobretudo, no léxico da região. Desse modo, ao lado da ocorrência de *sanga*, pode-se observar a sobreposição dessas isoléxicas na carta 1h – Arealidade *sanga* e *arroio* – o que confirma as premissas aventadas anteriormente sobre a influência do tropeirismo e da situação de contato do português com o espanhol, formando, de certo modo, uma área lexical no território investigado.

Distinguem-se, pela carta 1h, a variante *sanga* no oeste rio-grandense e a variante *arroio* no leste do Estado. Na região central, há a sobreposição de formas ocorrendo tanto *arroio* quanto *sanga*, revelando áreas lexicais delimitadas por linhas de heteroléxicas⁷³. A carta 1h permite visualizar também uma área em Santa Catarina pela qual se liga ao Paraná, a qual Romano e Aguilera (2014, no prelo) denominam como *área lexical do interior catarinense*⁷⁴. No Paraná, as isoléxicas se sobrepõem em Lapa e Imbituva, pontos 222 e 218, respectivamente. *Sanga* segue rumo o MS, refletindo-se em Ponta Porã – ponto 117 e *arroio* avança rumo ao Norte Pioneiro paranaense, chegando até o ponto 214 –

⁷³ Termo usado, nesse trabalho, por extensão ao conceito de heteroglossa, de Chambers; Trudgill (1994, p. 140).

⁷⁴ Os autores fazem uma revisão dos trabalhos lexicais com vistas a discutir os padrões de variação lexical no sul do Brasil. Assim, partindo dos trabalhos de Koch (2000) e Altenhofen (2005), analisam a distribuição diatópica das variantes lexicais para *bolinha de gude* nesta região do país, considerando-se o *corpus* do ALiB. Com o trabalho, os autores confirmaram três das oito ‘áreas dialetais’ estabelecidas por Altenhofen (2005) e propõem outras três áreas, entre elas 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.

Pirai 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 deste possível *falar sulista* é que se empreendem as análises deste trabalho.

Rego e variantes apresentam-se com 17 ocorrências no *corpus* (2,64%). Conforme os dados do relatório 1a e a carta 1i, esta variante ocorre em MG (quatro registros – 6,35%), GO (3 registros, 7,32%), MT (seis registros – 13,95%), MS (três ocorrências – 9,09%) e apresenta uma ocorrência em SP (0,48%). Neste último Estado, o item *regó* foi registrado apenas em Assis – ponto 169. No MS, apresenta-se em Ponta Porã (ponto 117) e Nioaque (116). No MT, ocorre, principalmente, no sudeste do Estado: pontos 106 – Poxoréu, 109 – Barra do Garças e 111 – Alto Araguaia. No Estado de GO, *regó* apresenta-se em Goiânia (ponto 123) e em Catalão (125). No território mineiro, a variante ocorre em dois pontos: 135 – Uberlândia e 136 – Patos de Minas e em dois pontos de controle: Unaí (ponto 130) e Formiga (141).

Embora ocorra com baixa representatividade, esta forma lexical revela traços diatópicos que diferenciam as áreas em que ocorre em relação à Região Sul, onde não foi registrado o item. No entanto, não é possível afirmar que a forma *regó* constitui uma área lexical específica dentro do subfalar sulista, de Nascentes (1953), uma vez que em termos de distribuição espacial, a área de ocorrência não apresenta contiguidade territorial.

A variante *corixo*, porém, apesar de ocorrer com baixa produtividade (cinco registros – 0,78%), está documentada em duas áreas específicas do subfalar sulista, no ponto 113 – Corumbá-MS e em Vila Bela da Santíssima Trindade-MT (ponto 107) – carta 1j. Nesta última, *corixo* apresenta-se como ocorrência única e, em Corumbá, está presente na fala dos quatro informantes. De acordo com o relato de um dos informantes dessa localidade, *corixo* é uma forma típica para denominar o *córrego* na região do pantanal, conforme se atesta pela transcrição *ipsis litteris* apresentada a seguir⁷⁵.

INF.- Geralmente fala córrego.

⁷⁵ Informante 4 de Corumbá (ponto 113) – mulher, segunda faixa etária (50 a 65 anos). Siglas INQ. (inquiridor/a) e INF. (informante).

INQ.- E algum outro nome aqui pelo Pantanal?

INF.- No Pantanal não, no Pantanal... aqui num tem córrego, aqui nós falamos curixo (...)

INQ.- E agora, nesses corixos, a água é corrente?

INF.- É a água, é.

INQ.- E aí quando, ele tem o quê? Mais ou menos uns dois metros de largura, o corixo?

INF.- Ah tem, tem.

De acordo com Houaiss e Villar (2001) e Ferreira (2004), *corixo* é uma variante de *corixa*, forma considerada por este último como um brasileirismo do MT e GO para designar o “canal por onde as águas das lagoas, dos brejos e dos campos baixos se escoam para os rios vizinhos” (FERREIRA, 2004). Quanto à etimologia, não há indícios nestas duas obras lexicográficas de se tratar de um vocábulo de origem castelhana ou portuguesa. No dicionário de etimologia de Cunha (2010), não consta entrada para *corixo/curixo (a)*. O que se sabe é que *corixo* é uma variante típica daquela região do país pois o ALMS (OLIVEIRA, 2007) documenta seis ocorrências de *corixu/curixu* distribuídas em três localidades: uma ocorrência em Corumbá, duas em Nhecolândia e três em Porto Esperança, o que forma uma área lexical deste item localizada no Pantanal sul-mato-grossense.

Observa-se diante deste cenário o polimorfismo acentuado para a questão em pauta que merece uma visão ampla a partir de um mapa de sobreposição de isoléxicas, formando áreas lexicais, conforme se verifica na carta 1k. Considerando apenas as formas lexicais, *córrego*, *sanga*, *arroio* e *ribeirão*, verifica-se pela carta em questão que, na área correspondente ao subfalar sulista, de Nascentes (1953), os Estados pertencentes ao Sudeste e Centro-Oeste apresentam menor número de sobreposição de isoléxicas em comparação aos Estados da Região Sul, onde se observa o maior número de formas regionais e combinação dessas formas com outras variantes.

Ou seja, considerando essas quatro formas lexicais, 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, portanto, para se referir a (i), opta-se pela expressão *falar paulista* e a (ii), por *falar sulista*, nestes casos 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.

4.1.1 Tratamento Estatístico

A análise dos dados em números absolutos e relativos (ocorrência e percentual), além das análises diatópico-históricas considerando-se as áreas de ocorrência das variantes e a correlação dos registros com outros atlas linguísticos, encaminha para a formulação de quatro hipóteses que mereceram tratamento estatístico mais refinado na ferramenta SPSS. Quatro são as variantes às quais foram aplicados os tratamentos estatísticos atinentes à rede de pontos da área geográfica do subfalar sulista⁷⁶, devido à sua produtividade significativa no *corpus*, a saber: *córrego*, *riacho*, *riozinho* e *ribeirão*.

As hipóteses formuladas para o tratamento estatístico são:

- Hipótese nula 1: a variante *córrego* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 2: a variante *riacho* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 3: a variante *riozinho* apresenta distribuição de produtividade heterogênea nos Estados de São Paulo, Paraná, Santa Catarina e Rio Grande do Sul.
- Hipótese nula 4: a variante *ribeirão* apresenta distribuição de produtividade heterogênea nos Estados de Minas Gerais, São Paulo, Paraná e Santa Catarina.

Antes, porém, da validação dessas hipóteses, fez-se necessário analisar o comportamento das referidas variantes caracterizadas como variáveis

⁷⁶ Para o tratamento estatístico, não foram considerados os dados referentes aos pontos de controle para que não haja interferência nos resultados.

aleatórias dependentes em diferentes contextos. Para isso, foi utilizado o teste KS⁷⁷ objetivando verificar se os dados apresentam aderência à distribuição de probabilidade normal. A Tabela 1 apresenta os resumos estatísticos segundo as variantes e os Estados.

Tabela 1 - Resumo estatístico segundo o Estado para a questão 001 do QSL (verificação da normalidade dos dados)

		<i>córrego</i>	<i>riacho</i>	<i>riozinho</i>	<i>ribeirão</i>
MT	<i>Respostas: média (N*)</i>	3,43 (7)	0,29 (7)	0,14 (7)	-
	<i>Desvio padrão</i>	0,976	0,756	0,378	-
	<i>KS (p-valor)**</i>	1,151 (0,141)	1,335 (0,057)	1,335 (0,057)	-
MS	<i>Respostas: média (N*)</i>	3,17 (6)	0,83 (6)	-	0,17 (6)
	<i>Desvio padrão</i>	0,983	1,602	-	0,408
	<i>KS (p-valor)**</i>	0,739 (0,646)	0,895 (0,4)	-	1,205 (0,110)
GO	<i>Respostas: média (N*)</i>	3,67 (6)	0,67 (6)	-	0,17 (6)
	<i>Desvio padrão</i>	0,516	0,516	-	0,408
	<i>KS (p-valor)**</i>	0,998 (0,272)	0,998 (0,272)	-	1,205 (0,110)
MG	<i>Respostas: média (N*)</i>	3,29 (7)	0,86 (7)	0,29 (7)	0,71 (7)
	<i>Desvio padrão</i>	0,951	0,9	0,756	0,756
	<i>KS (p-valor)**</i>	0,913 (0,375)	0,683 (0,739)	1,335 (0,057)	0,678 (0,748)
SP	<i>Respostas: média (N*)</i>	2,29 (38)	1,11 (38)	1,05 (38)	0,55 (38)
	<i>Desvio padrão</i>	1,25	0,953	0,837	0,86
	<i>KS (p-valor)**</i>	1,651 (0,009)	1,569 (0,015)	1,615 (0,011)	2,288 (p<0,001)
PR	<i>Respostas: média (N*)</i>	2,12 (17)	1,18 (17)	1,29 (17)	0,24 (17)
	<i>Desvio padrão</i>	1,409	0,883	0,92	0,562
	<i>KS (p-valor)**</i>	0,816 (0,518)	0,933 (0,349)	1,028 (0,242)	2,003 (p<0,001)
SC	<i>Respostas: média (N*)</i>	1,1 (10)	1,3 (10)	0,6 (10)	0,3 (10)
	<i>Desvio padrão</i>	0,738	1,16	0,699	0,675
	<i>KS (p-valor)**</i>	0,803 (0,539)	0,639 (0,809)	0,963 (0,312)	1,491 (0,023)
RS	<i>Respostas: média (N*)</i>	0,47 (17)	1,41 (17)	0,53 (17)	-
	<i>Desvio padrão</i>	0,624	1,004	0,717	-
	<i>KS (p-valor)**</i>	1,496 (0,023)	1,033 (0,236)	1,476 (0,026)	-

Fonte: Banco de dados do ALiB (2013) **Nota:** *N = Quantidade de localidades ** Teste Kolmogorov-Smirnov

⁷⁷ Modelo estatístico desenvolvido pelos matemáticos russos Andrei Kolmogorov e Vladimir Ivanovich Smirnov (LILIEFORS, 1967).

Neste momento, é importante destacar que a distribuição normal é uma das mais importantes distribuições da estatística, sendo também conhecida como distribuição de Gauss ou Gaussiana, e tem grande uso na estatística inferencial.

A partir dos dados apresentados na Tabela 1, nota-se que no MT, MS, GO e MG as variantes apresentam aderência à distribuição normal (p -valor $> 0,05$). Já no Estado de SP, as quatro variantes mais produtivas do *corpus* não apresentam aderência à normalidade. No PR e SC, as variantes que apresentam distribuição normal são *córrego*, *riacho* e *riozinho* e, no Estado do RS, apenas a variante *riacho* apresenta aderência à normal. Portanto, como nem todas as formas comportam-se com aderência a essa distribuição de probabilidade, considera-se que os itens lexicais estudados não apresentam aderência à normal, sendo necessária a aplicação de testes estatísticos não-paramétricos para testar as hipóteses formuladas no estudo.

Com a finalidade de verificar a homogeneidade da distribuição de produtividade das variantes lexicais mais produtivas nos Estados contemplados pela pesquisa, foi aplicado o teste estatístico Kruskal-Wallis⁷⁸.

Tabela 2 - Teste de homogeneidade da distribuição de produtividade da variante *córrego* com base nos Estados

	Média do rank (N)	df	Teste*(p-valor)
Córrego	MT	83 (7)	7
	MS	76,83 (6)	
	GO	88 (6)	
	MG	79,57 (7)	
	SP	57,51 (38)	
	PR	54,21 (17)	
	SC	32,15 (10)	
	RS	19,44 (17)	
			48,954 ($p < 0,001$)

Fonte: Banco de dados do ALiB (2013) *Teste de Kruskal-Wallis N = quantidade de localidades

Desse modo, para a variante *córrego*, o resultado (Tabela 2) mostra que o teste alcançou nível significativo, podendo-se assumir que as variâncias dos grupos não são homogêneas. A partir da análise de variância, nota-se que há

⁷⁸ Modelo estatístico desenvolvido pelos matemáticos Willian H. Kruskal e W. Allen Wallis (KRUSKAL; WALLIS, 1952).

diferença significativa entre as médias de todos os Estados ($p < 0,001$). Assim, conclui-se que os Estados não apresentam distribuição de produtividade com comportamento homogêneo e a hipótese nula 1 pode ser aceita, ou seja, *córrego* e suas variantes morfofonêmicas, apesar de apresentarem ampla distribuição diatópica nos Estados, com exceção do Rio Grande do Sul (carta 1a), esta distribuição não é homogênea, o que evidencia o polimorfismo lexical da questão em pauta.

Comparando-se a carta 1a (Arealidade da variante *córrego*) com a carta 1l (Arealidade gradual da variante *córrego*), observa-se que, de fato, esta forma lexical, embora esteja amplamente difundida pelo território investigado, com exceção do RS, apresenta variações quanto à produtividade na rede de pontos. Desse modo, a carta 1l confirma os resultados do teste estatístico evidenciando que a variante *córrego* e formas morfofonêmicas não apresentam comportamento homogêneo. Este item lexical é mais representativo, sobretudo, em grande parte do território dos Estados da região Centro-Oeste e também no Triângulo Mineiro, onde a produtividade chega a 100% de incidência. Por outro lado, no Estado de SP há uma oscilação entre os índices de 75% a 25%, e à medida que se 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 desta variante no território do subfalar sulista.

A Tabela 3 apresenta os resultados do teste de homogeneidade de *riacho* e variantes morfofonêmicas nos Estados contemplados pelo estudo, evidenciando que o teste Kruskal-Wallis não alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são homogêneas.

Tabela 3 - Teste de homogeneidade da variante *riacho* com base nos Estados

	Média do rank (N)	df	Teste*(p-valor)	
<i>riacho</i>	MT	28,64 (7)	7	10,831 (0,146)
	MS	40 (6)		
	GO	43,67 (6)		
	MG	48,86 (7)		
	SP	55,86 (38)		
	PR	59,03 (17)		
	SC	60,4 (10)		
	RS	65,38 (17)		

Fonte: Banco de dados do ALiB (2013) *Teste de Kruskal-Wallis N = quantidade de localidades

A partir da análise de variância, nota-se que não há diferença significativa entre as médias de todos os Estados ($p = 0,146$). Deste modo, de acordo com os dados apresentados, conclui-se que os Estados apresentam distribuição de produtividade com comportamento homogêneo e a hipótese nula 2 pode ser rejeitada. A carta linguística 1m ratifica os resultados encontrados na aplicação do teste estatístico, uma vez que se observa o predomínio do item na maior parte do território, com 25% de produtividade, havendo em apenas um ponto a predominância de 100% (ponto 115 – Campo Grande/MS), e pontos isolados com 75%. Ou seja, na maioria dos pontos em que ocorre a variante *riacho*, o índice de representatividade oscila de 25 a 50%. Desse modo, apesar de ocorrer com diferentes índices, o que motivou a hipótese de que a variante apresentava comportamento heterogêneo (hipótese nula 2), o comportamento do item, de fato, é parecido na maioria dos pontos, conforme ilustra a carta linguística 1m. Ou seja, *riacho* não caracteriza, propriamente, uma área lexical, antes, comporta-se como item comum a diferentes áreas envolvidas pelo estudo.

A Tabela 4 apresenta os resultados do mesmo teste estatístico para a variante *riozinho* nos Estados do PR, SP, SC e RS, a fim de comprovar a premissa estabelecida anteriormente sobre a irradiação dessa variante a partir do PR para SP e SC e, daí, para o interior do Rio Grande do Sul.

Tabela 4 - Teste de homogeneidade da variante *riozinho* com base nos Estados do SP, PR, SC e RS

		Média do rank (N)	df	Teste*(p-valor)
<i>riozinho</i>	SP	44,61 (38)	3	9,215 (0,027)
	PR	50,71 (17)		
	SC	32,75 (10)		
	RS	30,5 (17)		

Fonte: Banco de dados do ALiB (2013) *Teste de Kruskal-Wallis N = quantidade de localidades

O resultado mostra que o teste alcançou nível significativo nesses Estados, podendo-se assumir que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias dos Estados ($p = 0,027$). Assim, de acordo com os dados apresentados, conclui-se que os Estados apresentam distribuição de produtividade com comportamento heterogêneo e a hipótese nula 3 pode ser aceita.

A carta linguística 1n ratifica os resultados do teste estatístico, ilustrando comportamento heterogêneo na maioria dos pontos linguísticos quanto à distribuição diatópica de *riozinho*, uma vez que apenas em três pontos há 100% de produtividade desta variante (pontos 153 – Barretos/SP, 168 – Mogi Mirim/SP e 222 – Lapa/PR). Nas demais localidades, o índice de incidência oscila de 33 a 66%, predominando as áreas com o menor índice, ou seja, na maioria dos pontos a variante representou 33% das respostas (ocorrência única na localidade). Todavia, observa-se na carta que em grande parte do território gaúcho e do litoral catarinense não há o registro deste item, contribuindo, portanto, para a confirmação da heterogeneidade da distribuição da variante nos Estados supramencionados. Desse modo, os resultados possibilitam confirmar a premissa aventada sobre a difusão dessa variante a partir do PR, ou seja, a existência de uma subárea lexical que compreende localidades tanto da parte meridional (*falar sulista*) quanto da parte setentrional do território investigado (*falar paulista*), contemplando o Estado do PR de onde se irradia para o interior de SP, SC e RS.

O mesmo teste foi aplicado para a variante *ribeirão* com o objetivo de verificar se as ocorrências das variâncias são heterogêneas nos dados de MG, SP, PR e SC (Tabela 5).

Tabela 5 - Teste de homogeneidade da variante *ribeirão* com base nos Estados de MG, SP, PR e SC

		Média do rank (N)	df	Teste*(p-valor)
<i>ribeirão</i>	MG	44,79 (7)	3	4,210 (0,24)
	SP	38,38 (38)		
	PR	31,26 (17)		
	SC	32,45 (10)		

Fonte: Banco de dados do ALiB (2013) *Teste de Kruskal-Wallis N = quantidade de localidades

O resultado mostra que o teste não alcançou nível significativo em termos de produtividade, podendo-se assumir que as variâncias dos grupos são homogêneas, ou seja, têm comportamento similar. A partir da análise de variância, nota-se que não há diferença significativa entre as médias dos Estados ($p = 0,24$). Portanto, conclui-se que os Estados apresentam distribuição de produtividade com comportamento homogêneo e a hipótese nula 4 pode ser rejeitada.

Acreditava-se que, pelo fato de os números absolutos indicarem maior concentração do item em localidades mineiras e paulistas, *ribeirão* apresentaria distribuição heterogênea no *corpus*. Porém, pela carta 1o, verifica-se comportamento homogêneo do item, predominando no território investigado, com 33% de representatividade (uma ocorrência por ponto). Apenas em dois pontos verifica-se 100% de representatividade de *ribeirão* (pontos 170 – Bernardino de Campos/SP e 174 – Bragança Paulista/SP). Nas demais localidades, ou ocorre com 66% (seis pontos) ou com 33% de representatividade. Verifica-se, desse modo, comportamento homogêneo na distribuição de produtividade do item.

Em síntese, o panorama das hipóteses para a questão 001 do QSL é apresentado no Quadro 3.

Quadro 3 - Síntese das hipóteses para a questão 001 do QSL

	ACEITA	REJEITADA
HIPÓTESE NULA 1 [#]	X	
HIPÓTESE NULA 2 ^{##}		X
HIPÓTESE NULA 3 ^{###}	X	
HIPÓTESE NULA 4 ^{####}		X

córrego é heterogêneo em todos os Estados

riacho é heterogêneo em todos os Estados

riozinho é heterogêneo nos Estados de PR, SP, SC e RS

ribeirão é heterogêneo nos Estados de MG, SP, PR e SC

De acordo com o Quadro 3, duas das quatro hipóteses formuladas podem ser aceitas por meio do teste estatístico empreendido. De posse da análise estatística e descritiva (números absolutos e relativos), bem como considerando a área de ocorrência das variantes documentadas, chega-se às seguintes conclusões acerca das variantes lexicais para questão 001 do QSL na região considerada por Nascentes (1953) como subfalar sulista.

- (i) *Córrego* é uma variante comum no território constituindo-se como a forma mais produtiva, presente em todos os Estados, inclusive nos pontos de controle. A produtividade dessa variante está condicionada à região em que ocorre no território investigado, pois, à medida que se avança em sentido meridional, fica menor a sua representatividade, uma vez que nos Estados do PR, SC e RS ocorrem formas lexicais consideradas como regionais, que delimitam áreas lexicais e possibilitam o traçado de isoléxicas e heteroléxicas. O teste de hipótese para a variante *córrego* bem como a carta 11 evidenciam que o item não apresenta distribuição homogênea em termos de produtividade, o que ratifica a heterogeneidade e o polimorfismo para a questão, embora, espacialmente, a distribuição da variante seja amplamente difundida. Possivelmente, portanto, o item caracteriza um possível *falar paulista* que se difunde pela parte setentrional do território investigado, principalmente, pelo uso da forma considerada padrão.
- (ii) *Riacho* também se mostra comum aos Estados contemplados, ocorrendo, inclusive, nos pontos de controle no limite desse subfalar sulista com os demais (mineiro e fluminense), além da área correspondente ao território denominado incharacterístico (Norte do Mato Grosso). Em termos de produtividade, o teste estatístico e a carta 1m confirmam a homogeneidade, rejeitando a hipótese nula 2. Este item, portanto, mostra-se comum às duas grandes áreas encontradas no território investigado, não sendo exclusivo de nenhum dos dois falares aqui defendidos.

- (iii) *Riozinho* é uma variante mais produtiva no Estado do PR, no que se refere à produtividade e à distribuição espacial. Esta variante expande sua área de ocorrência aos Estados limítrofes (SP e SC) e, por um corredor central no Estado de Santa Catarina, atinge o norte e oeste do RS. A ausência desse item no sul de GO e MS enfraquece a hipótese de que esta forma lexical tenha se irradiado a partir do Estado de SP, por motivos sócio-históricos já conhecidos. O teste estatístico e a carta 1n evidenciam diferenças significativas entre o PR em comparação com SP, SC e RS, demonstrando que a distribuição desta variante é heterogênea nestes Estados e a possível existência de subáreas lexicais comuns tanto ao *falar paulista* quanto ao *falar sulista* localizadas, principalmente, no Estado do PR de onde se irradia para SP, SC e RS.
- (iv) *Ribeirão* forma uma área lexical que contempla grande parte do sul de MG (inclusive pontos de controle em MG e GO) e, por um feixe central, corta o Estado de SP até chegar ao norte paranaense. Desta área, depreende-se a influência dos mineiros e paulistas no léxico da região, principalmente pelo fato de essa variante ser atestada com mais representatividade nas cartas do ALERS e do ALPR, onde a rede de pontos é mais densa.. Nota-se, no entanto, a ocorrência de *ribeirão* também em SC (região de Blumenau e Itajaí), localidades que não tiveram a influência direta desses fluxos migratórios durante o processo de formação. Desse modo, as outras regiões em que ocorre a variante não podem ser interpretadas como reflexo de fluxos migratórios, mas, possivelmente, remete a processos migratórios mais antigos, possivelmente, pela influência de vicentinos e açorianos naquela área do território catarinense durante o século XVIII. Verificando a distribuição de produtividade nos Estados de MG, SP, PR e SC, encontrou-se comportamento homogêneo na distribuição dessa variante, o que não revela diferença significativa em termos de produtividade, embora, diatopicamente, encontrem-se essas áreas bem delimitadas, ao

se confrontar as cartas 1d e 1o. Ribeirão, desse modo, seria uma variante que se enquadraria mais no *falar paulista*, porém, com ocorrência não exclusiva em Estados das Regiões Centro-Oeste e Sudeste.

Quanto às variantes às quais não foram aplicados testes de hipóteses, considerando-se a baixa representatividade no *corpus* e as regiões em que ocorrem, sinteticamente, fazem-se os seguintes apontamentos:

- (v) *Valeta* ocorre de forma produtiva formando uma área de isoléxica que contempla o litoral catarinense, atinge uma localidade paranaense e adentra o interior de SC. Esta forma ocorre também em alguns pontos isolados em SP, PR e RS, portanto, não pertence a um falar específico, mas evidencia subáreas lexicais comuns aos dois grandes falares: *paulista* e *sulista*.
- (vi) As variantes *sanga* e *arroio* constituem isoléxicas que evidenciam traços da sócio-história do RS, SC e PR a ponto de delimitar uma área lexical no subfalar sulista em localidades que tiveram a influência do movimento dos tropeiros no processo de formação. Estas variantes dividem o RS em duas partes: a porção oeste (caracterizada pela presença de *sanga*) e a porção leste (por *arroio*). No interior gaúcho, há uma confluência das duas variantes, uma área de heteroléxica, que adentra o PR por meio do corredor central do interior catarinense espalhando-se no território paranaense por todo centro-sul (principalmente na região do município de Lapa). *Sanga* segue rumo ao MS atingindo o município de Ponta Porã e *arroio* segue em direção ao norte, sem, contudo, ultrapassar a fronteira paranaense. Ambas as variantes revelam aspectos do contato do português com o espanhol no sul do Brasil e sua área de abrangência também se justifica pela influência dos descendentes de migrantes europeus que se deslocaram para as fronteiras agrícolas, sobretudo, do oeste de SC e sudoeste do PR, além de

outras áreas. Sanga e arroio, portanto, caracterizam o *falar sulista*, de influência sul-rio-grandense no oeste de SC e sudoeste do PR e também revela contato do português com línguas de colonização, no caso o espanhol.

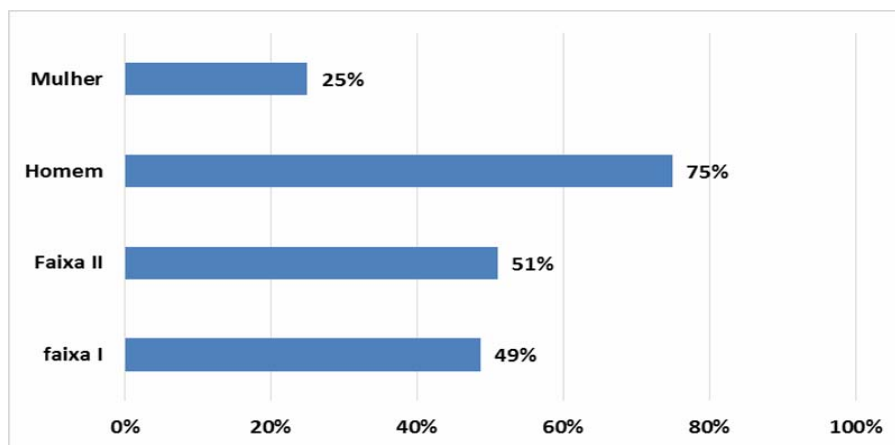
Desse modo, a área correspondente ao subfalar sulista, de Antenor Nascentes (1953), quanto à distribuição das variantes para o conceito de *córrego*, pode ser dividida em duas partes: (i) parte setentrional, em que ocorre uma ampla área lexical da forma considerada padrão (*córrego*), principalmente, em Estados da região Sudeste e Centro-Oeste; e (ii) parte meridional que apresenta polimorfismo acentuado com a presença de formas lexicais típicas dos Estados do Sul do Brasil, resultantes do contato do português com línguas de colonização.

4.2 QUESTÃO 177 – GELEIA

A questão 177 do QSL, pertencente à área semântica Alimentação e Cozinha, objetiva registrar os designativos que recobrem o conceito da “pasta feita de frutas para passar no pão, biscoito” (COMITÊ NACIONAL DO PROJETO ALiB, 2001).

Conforme o relatório 2 (vol.2), foram documentados 420 registros, além de 85 abstenções de respostas e quatro problemas técnicos na gravação. A Figura 60 apresenta os resultados das abstenções de resposta em dados percentuais, na qual se observa a sua predominância entre os informantes do sexo masculino (75%) e entre os da segunda faixa etária (51%).

Figura 60 - Resultado da abstenção de resposta para a questão 177 do QSL, segundo o perfil do informante.



Fonte: [ΣGVCLin] – Relatório 2 (vol.2)

Uma informante do ponto 165 (Presidente Prudente-SP), embora conhecesse a variante *geleia* para denominar o referente, ao ser indagada sobre o uso de tal iguaria no cardápio cotidiano, apresentou o seguinte relato:

INF.- Geleia [Ζεελεφ].

INQ.- Você já viu geleia?

INF.- Já.

INQ.- E você usa?

INF.- Não. Eu não tenho condição (risos) porque é muita criança, se fô comprá não vence. (risos)⁷⁹

Ou seja, o próprio perfil dos informantes do ALiB contribui para esse índice de abstenção de resposta, muitas vezes, por conta do desconhecimento do referente, uma vez que os entrevistados possuem baixa escolaridade que, em geral, se associa a profissões que não possibilitam poder aquisitivo elevado. Este critério, possivelmente, justifica o índice de abstenção.

Quanto às variantes documentadas, são três as formas lexicais apresentadas e analisadas (*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*.

⁷⁹ Informante 2 do ponto 165 (Mulher, Faixa Etária I, de Presidente Prudente-SP).

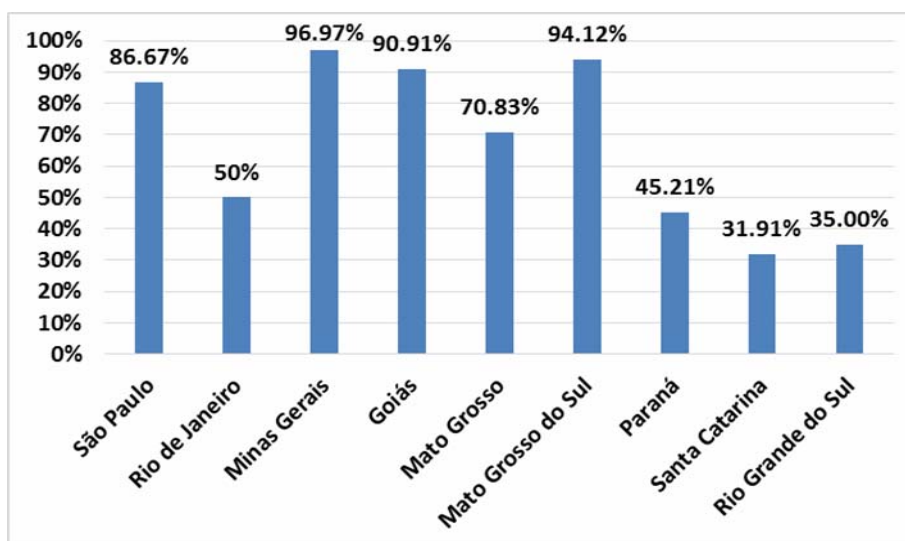
Observando o relatório 2, verifica-se que a forma lexical *geleia* predomina em 267 das 420 ocorrências (63,57%), seguida de *chimia* – 71 registros (16,90%), *formas genéricas* (53 – 12,62% do *corpus*) e, em menor produtividade, encontram-se as *formas inadequadas* (11 ocorrências – 2,62%), itens em que a própria pergunta sugere a resposta – 10 registros (2,38%) e, por fim, a variante *musse*, com oito ocorrências (1,90%).

Conforme se observa, para esta questão, não é documentado o polimorfismo lexical, principalmente pelo fato de o índice de abstenção de resposta ser alto. No entanto, as ocorrências das variantes delimitam áreas lexicais no território investigado, conforme se verifica na carta 2. Nesta carta, 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 a partir da forma lexical *doce*, enquanto *musse* se apresenta significativamente em SC e *chimia* se difunde pelo Estado do RS.

Quanto à produtividade da variante mais difundida na área do subfalar sulista (*geleia*), observa-se que este item obtém diferentes índices de representatividade nos nove Estados contemplados pelo estudo. Conforme o relatório geral de produtividade por Estado (relatório 2a – vol. 2), sistematizado na Figura 61, verifica-se que este item lexical apresenta-se com alto índice percentual em MG (96,97%), MS (94%), GO (90,91%), SP (86,67%) e MT (70,83%). No RJ, a variante ocorre em 50% das respostas válidas e, nos Estados da região Sul, *geleia* coocorre com outras formas, ficando abaixo de 50% de representatividade. No PR, o item lexical representa 45,21% das respostas, em SC chega ao índice de 31,91% e, no RS, representa 35% do *corpus*.

A carta 2a documenta a distribuição de *geleia* na região contemplada pelo estudo, evidenciando que esta forma lexical está amplamente difundida pelo território, com ocorrência, inclusive, nos pontos de controle, exceto em uma localidade fluminense, ponto 205 – Barra Mansa. Na área do subfalar sulista de Nascentes (1953), *geleia* não ocorre no oeste do PR e de SC, bem como em uma localidade do sudeste catarinense, ponto 233 – Criciúma, e em três pontos gaúchos, 240 – Flores da Cunha, 243 – Porto Alegre e 245 – Uruguaiana.

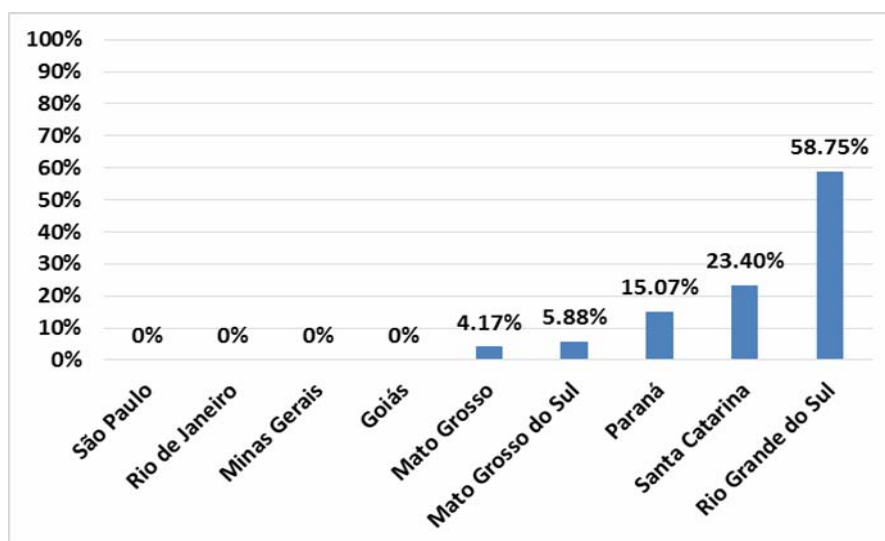
Figura 61 - Produtividade de *geleia* por Estado



Fonte: [ΣGVCLin] – Relatório 2a (vol.2)

A segunda variante mais produtiva é *chimia* e a sua variante morfofonêmica *michia*. Essas formas representam juntas 16,90% do *corpus* apresentando-se em 71 das 420 ocorrências. *Chimia* é uma variante fonética de *chimiê* e *chimier*, caracterizadas pelo dicionário Houaiss (2001) como um regionalismo do Rio Grande do Sul para designar a "geleia para passar no pão da merenda; chimiê" (HOUAISS; VILLAR, 2001). De acordo com esta obra lexicográfica, este vocábulo provém do alemão "Schimiere" que significa, lubrificante, graxa, proveniente do verbo *schmieren* 'untar, lubrificar, borrar, sujar.' Assim, para Houaiss; Villar (2001), "o uso do vocábulo *chimier* para designar 'geleia' é uma adaptação do al. *Schmiere* ao port., criada pelos brasileiros de ascendência alemã, residentes no Sul do Brasil, a partir da ideia de que, por extensão de sentido, a geleia serve para 'lubrificar' pão, biscoito etc." (HOUAISS; VILLAR, 2001).

Desse modo, considerando-se a etimologia do vocábulo, bem como a história social, sobretudo, da Região Sul, a presença de *chimia*, variante difundida pela área correspondente ao *falar sulista*, evidencia traços de bilinguismo e empréstimo do português de línguas de colonização. Dos nove Estados contemplados pelo estudo, este item lexical está presente em cinco com diferentes índices de representatividade, conforme se observa na Figura 62.

Figura 62 - Produtividade de *chimia* por Estado

Fonte: [ΣGVCLin] – Relatório 2a (vol.2)

De acordo com os dados constantes da Figura 62 e do relatório 2a (vol.2), observa-se que no MT e no MS *chimia* apresenta baixa representatividade, figurando como ocorrência única no ponto 117 – Ponta Porã/MS, 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, né [əðoσI]: “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* [ΣIεμι□], né, tudo quando é feito de doce pa passá no pão é *chimia*⁸⁰.

Ou seja, 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*.

De fato, a distribuição diatópica de *chimia* no território investigado (carta 2b) revela uma área que contempla o Estado do RS quase em sua totalidade,

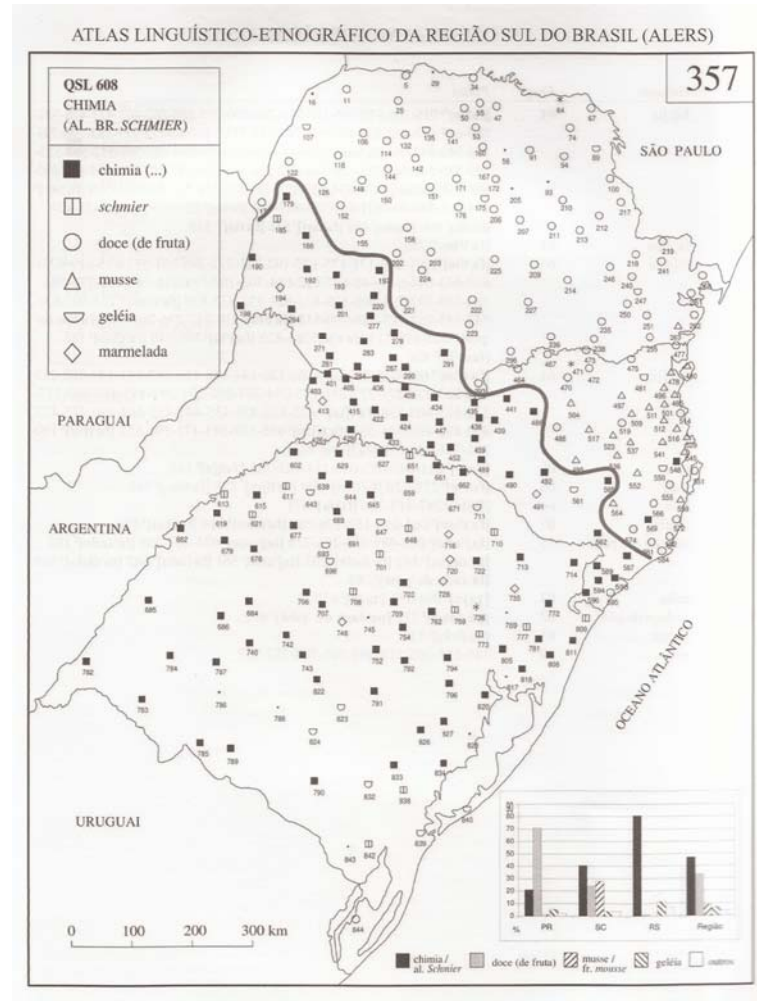
⁸⁰ Informante 3 do ponto 103 (Homem, Faixa etária II, de Aripuanã-MT).

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).

De acordo com o relatório 2a, no PR, a variante *chimia* representa 15,07% de produtividade (11 ocorrências). Em SC, este item lexical ocorre em cinco das dez localidades (11 ocorrências – 23,40%), formando uma área que diferencia o oeste catarinense do leste do Estado, onde não ocorre esta variante. *Chimia*, entretanto, é significativa no RS, apresentando-se como variante majoritária (58,75%), não ocorrendo somente em uma localidade – ponto 249 – São José do Norte.

O ALPR (AGUILERA, 1994) e o ALPR II (ALTINO, 2007) não apresentam a carta linguística para os designativos da “geleia”, no entanto, o ALERS (ALTENHOFEN; KLASSMANN, 2011), na carta 357 (Figura 63), documenta cinco formas lexicais para a pasta feita de fruta que se usa para passar no pão e no biscoito.

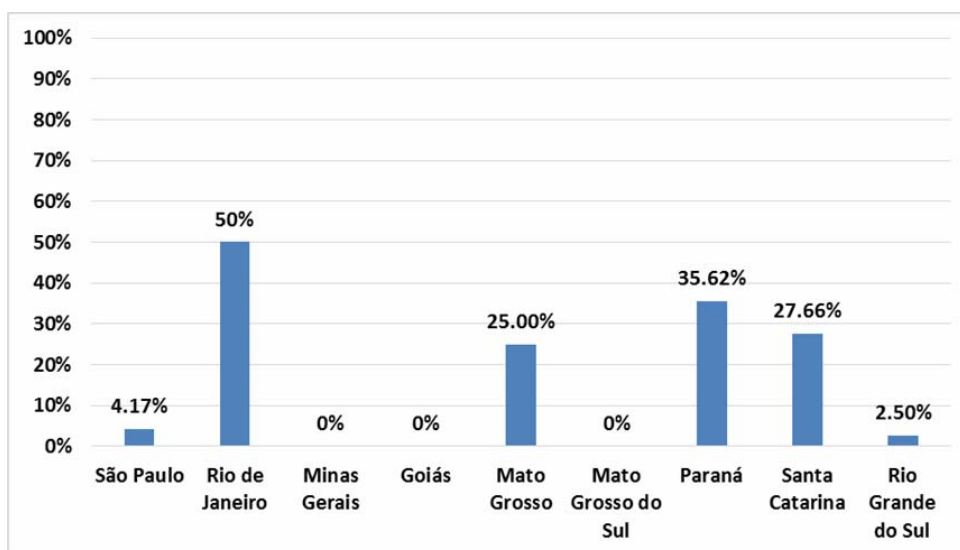
Figura 63 - Carta linguística 357 do ALERS



Fonte: Atlas Linguístico Etnográfico da Região Sul do Brasil (ALTENHOFEN; KLASSMANN, 2011)

De acordo com a referida carta, *chimia* é a forma lexical que predomina no *corpus* daquele atlas, evidenciando uma área lexical que abrange todo o território gaúcho, adentra o sul e oeste catarinense e contempla o sudoeste paranaense, apresentando uma distribuição semelhante àquela encontrada no *corpus* do ALiB (Carta 2b).

Conforme se observa na carta 2c, as áreas de ocorrência de formas lexicais agrupadas sob o rótulo “*formas genéricas*”, formadas a partir do vocábulo “doce”, abrangem pontos de controle no MT (ponto 103 – Aripuanã e 104 – São Félix Araguaia) e uma localidade fluminense (ponto 205 – Barra Mansa) também de controle, cuja ocorrência na localidade é de 50% (Figura 64). Na área investigada, essas formas obtêm diferentes índices de representatividade em cada Estado, não ocorrendo somente em MG, GO e MS.

Figura 64 - Produtividade das “formas genéricas” relativas à *geleia* por Estado

Fonte: [ΣGVCLin] – Relatório 2a (vol.2)

No território do subfalar sulista (carta 2c), as *formas genéricas* se difundem em grande parte pelo território paranaense e, em uma faixa contínua, adentram o litoral norte de SC e, por um corredor central, atingem o litoral sul deste Estado, ocorrendo também em uma localidade do oeste catarinense – ponto 226 (São Miguel do Oeste). Observa-se também a presença de *formas genéricas* em pontos localizados na fronteira política entre SP e PR (pontos 170 – Bernardino de Campos, 181 – Itararé e 185 – Ribeira), bem como em dois pontos isolados no Estado paulista (156 – Araçatuba e 178 – Sorocaba). No RS, foram documentadas *formas genéricas* no centro do Estado – ponto 242 (Santa Maria) e no nordeste (ponto 240 – Flores da Cunha).

Como se verifica, nos Estados da região Sul, o uso de *formas genéricas* é mais produtivo no PR – 35,62% das respostas (Figura 64), decaindo em SC (27,66%) e chegando apenas a 2,5% das respostas entre os gaúchos. Panorama análogo encontra-se no *corpus* do ALERS (ALTENHOFEN; KLASSMANN, 2011), no qual o uso da variante *doce*, aqui considerada como uma *forma genérica*, é predominante no PR, decaindo em SC e atingindo baixo índice de ocorrência no RS (Figura 63). Nesse atlas, *doce* em contraste à *chimia* define áreas lexicais separando a região Sul em duas grandes áreas, uma localizada na porção setentrional, de influência paranaense, caracterizada pela presença de *doce* e a outra na parte meridional, de influência rio-grandense, com a presença de *chimia*.

A variante menos produtiva no conjunto das respostas, *musse*, por sua vez, delimita uma área lexical no território investigado (Carta 2d). Apesar de obter apenas oito ocorrências (1,90% do *corpus*), esta variante apresenta-se em uma área específica do subfalar sulista, exclusivamente, 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 no ponto 231 – Lages, no sul de SC. Entretanto, no ALERS, *musse* ocorre também no RS e no PR, embora com baixa produtividade. No território catarinense, *musse* ocorre na mesma área em que se apresenta no *corpus* do ALiB (Figura 63).

Por fim, as formas consideradas inadequadas (*creme, creme de fruta, melado, patê, patê de fruta e polpa de fruta*), as quais não foram questionadas pelo inquiridor, juntamente com formas em que a própria formulação da questão sugere a resposta obtida (*pasta/pasta de fruta*), representam aproximadamente 5% do *corpus* e não se apresentam em uma área específica do subfalar sulista.

Para uma visão ampla do comportamento das variantes lexicais registradas para a questão 177, a carta 2e apresenta a sobreposição de isoléxicas, formando áreas de heteroléxicas e possibilita verificar a distribuição das principais formas registradas na área geográfica investigada: *geleia, chimia, formas genéricas e musse*. Desse modo, situação análoga à que foi apresentada na carta 1k repete-se para os designativos da *geleia*. Verifica-se que, na área correspondente ao subfalar sulista, os Estados pertencentes às regiões Sudeste e Centro-Oeste apresentam um menor número de sobreposição de isoléxicas em comparação com os Estados da Região Sul, onde se observa o maior número de formas regionais e combinação dessas formas com outras variantes, formando heteroléxicas. Observando a carta 2e, 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 nos Estados da região Sul, que apresentam maior número de combinação de variantes em áreas de heteroléxicas.

Ou seja, considerando essas quatro formas lexicais, a região do subfalar sulista, de Antenor Nascentes, mais uma vez 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.

4.2.1 Tratamento estatístico

Para o tratamento estatístico das variantes mais representativas da questão 177 do QSL, foram averiguadas três hipóteses, a saber:

- Hipótese nula 5: a variante *geleia* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 6: a variante *chimialmichia* apresenta distribuição de produtividade heterogênea nos Estados do PR, SC e RS.
- Hipótese nula 7: “*formas genéricas*” apresentam distribuição de produtividade heterogênea nos Estados do PR e SC.

A escolha dessas hipóteses pautou-se na distribuição espacial das referidas formas, uma vez que itens que não se apresentam significativamente em determinado Estado não se mostram relevantes para os testes estatísticos. Tome-se como exemplo a inclusão do Estado de SP na hipótese nula 7. Observando-se o relatório geral de produtividade (relatório 2), nota-se que o grupo *formas genéricas*, neste Estado, obteve apenas cinco ocorrências, evidenciando que não há homogeneidade desta forma naquele Estado devido à sua baixa incidência. O mesmo se dá com a inclusão de outros Estados para aquela hipótese. Por outro lado, a ocorrência exclusiva de determinada variante em um único Estado também inviabilizou o teste e a formulação de uma nova hipótese. Por exemplo, a partir da carta 2d, nota-se que *musse* não ocorre em outra área do território investigado, não necessitando, portanto, de um teste estatístico comparando essa variante com outras áreas de ocorrência.

A aplicação do teste KS revelou que as variantes *geleia* e *formas genéricas*, no Estado de SP, não apresentam aderência à distribuição normal ($p = 0,007$ e $p < 0,001$, respectivamente). O mesmo panorama se verifica para a variante *chimia* no PR ($p < 0,001$) e para as variantes *formas genéricas* no RS ($p < 0,001$). Vale destacar que a variante *chimia* não apresentou ocorrência em GO, MG e SP. Por sua vez, *formas genéricas* não se apresentaram em GO, MS e MG. Portanto, como nem todos os itens lexicais apresentam aderência a essa distribuição de

probabilidade, similarmente à questão anterior, foi necessária a aplicação de testes estatísticos não-paramétricos para testar as hipóteses formuladas.

Para o teste da hipótese nula 5, o resultado revela que o teste Kruskal-Wallis alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias de todos os Estados ($p = 0,004$), e a hipótese nula 5 pode ser aceita.

Ou seja, o teste estatístico ratifica a hipótese de que para a variante padrão e mais produtiva – *geleia* – apesar de ocorrer em quase todas as localidades, formando uma grande área lexical (carta 2a), não há homogeneidade na sua distribuição de produtividade, apresentando comportamento heterogêneo, sobretudo por cocorrer em determinados Estados com formas regionais. Observa-se, desse modo, pela carta 2f, que grande parte do território investigado apresenta 50% de ocorrência do item, principalmente, em Estados da região Centro-Oeste. Em SP, entretanto, o índice de ocorrência de *geleia* atinge 75% na porção leste desse Estado, chegando ao sul de MG e ao PR, na região por onde passavam as antigas rotas de tropeiros – localidades como Piraí do Sul (ponto 2014), Ibituva (218), Lapa (222), entre outras.

O índice de 75% de ocorrência do item compreende também uma faixa territorial que perpassa GO, MT e chega ao MS. Observa-se, por outro lado, que o índice de 100% de ocorrência de *geleia* concentra-se, basicamente, em três áreas: (i) oeste do MS (pontos 113 – Corumbá e 116 – Nioaque); (ii) sul de MG (pontos 140 – Passos, 144 – Lavras, 147 – Poços de Caldas) e também em duas localidades no extremo norte de SP (ponto 154 – Franca e 157 – Ribeirão Preto) e (iii) no litoral e sudeste de SP (pontos 182 – Capão Bonito, 183 – Itanhaém, 184 – Santos e 186 – Registro). Além dessas áreas, nota-se também ocorrência de 100% de produtividade em pontos isolados de GO, PR, SP e RS. O menor índice de *geleia* (25%) ocorre, sobretudo, no território gaúcho, catarinense e mato-grossense.

Com o objetivo de verificar se as variâncias das ocorrências de *chimia* são heterogêneas nos Estados do PR, SC e RS, foi aplicado o teste Kruskal-Wallis que, novamente, alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias dos Estados comparados ($p < 0,001$), e a hipótese nula 6 pode ser aceita.

A carta 2g confirma o resultado do teste estatístico quanto à distribuição heterogênea de *chimia*. Esta forma, de influência gaúcha, conforme se tem verificado, está amplamente difundida pelo interior do RS, com diferentes índices de representatividade, adentrando o oeste catarinense e paranaense até atingir, inclusive, uma localidade sul-mato-grossense (ponto 117 – Ponta Porã). A única localidade em que houve o maior índice desta variante (100%) é o ponto 217 (São Miguel do Iguazu/PR). Nos demais pontos linguísticos, a produtividade do item varia, apresentando-se em uma área que vai do RS ao oeste de SC, com 80% de representatividade. Entretanto, na maior parte do território em que ocorre, a variante *chimia* apresenta entre 25 e 60% de representatividade. Desse modo, *chimia* caracteriza um possível *falar sulista* de influência sul-rio-grandense no Estado de SC e PR e também do contato do português com línguas de colonização, no caso, o alemão.

Quanto à diferença significativa na distribuição de produtividade de *formas genéricas* nos Estados do PR e SC, o teste de Mann-Whitney não alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são homogêneas. A partir da análise de variância, nota-se que não há diferença significativa entre as médias dos Estados comparados ($p = 0,657$), e a hipótese nula 7 pode ser rejeitada.

Apesar de distribuídas em uma ampla faixa territorial no centro-sul paranaense em que alcançam 75% de representatividade (carta 2h), as *formas genéricas* não apresentam, de fato, grandes diferenças em relação ao Estado de SC, pois, embora haja menos ocorrências dos itens genéricos neste Estado, observa-se que em duas localidades (pontos 224 – Porto União, extremo norte catarinense) e ponto 233 (Criciúma – no litoral sul), o uso de *formas genéricas* atingiu 100% de representatividade. Já o PR figurou com índices iguais ou abaixo de 75%. Essa explicação justifica o resultado do teste estatístico, pois se observa que estas duas cidades catarinenses, onde houve 100% de incidência, ajudam a equiparar os dados ao Estado do PR. Desse modo, a carta 2h mostra comportamento semelhante, constituindo-se, portanto, como áreas que apresentam certa homogeneidade quanto à distribuição destes itens genéricos.

Sumarizando esses resultados, o Quadro 4 indica que duas das três hipóteses averiguadas para a questão 177 do QSL podem ser aceitas.

Quadro 4 – Síntese das hipóteses para a questão 177 do QSL

	ACEITA	REJEITADA
HIPÓTESE NULA 5 [#]	X	
HIPÓTESE NULA 6 ^{##}	X	
HIPÓTESE NULA 7 ^{###}		X

geleia é heterogênea em todos os Estados

chimia é heterogênea nos Estados de PR, SC e RS

formas *genéricas* são heterogêneas nos Estados de PR e SC

De posse desses resultados, formulam-se as seguintes conclusões:

- (i) A questão 177 do QSL não apresentou polimorfismo lexical, possivelmente em decorrência do alto índice de abstenção de resposta, registrado, sobretudo, entre os informantes da segunda faixa etária, majoritariamente do sexo masculino. Nota-se, inclusive, o desconhecimento do referente por parte desses informantes, uma vez que o perfil social (informantes de baixa instrução, que conseqüentemente, na maioria das vezes, coincide com a classe social) acaba não possibilitando a inclusão de *geleia* no cardápio cotidiano da família devido ao custo desta iguaria.
- (ii) A forma padrão *geleia*, embora com ampla distribuição diatópica no território investigado (carta 2a), não apresenta distribuição de produtividade homogênea na área investigada (carta 2f). Essa variante, conforme “caminha” em direção à Região Sul, vai perdendo representatividade. Desse modo, a heterogeneidade no comportamento dessa variante, comprovada por meio do teste estatístico, possibilita a aceitação da hipótese nula 5. Esta forma lexical, provavelmente, caracteriza o denominado falar paulista que se expande também em direção ao norte do PR e aos Estados da região Centro-Oeste do país.
- (iii) Em decorrência dos resultados da hipótese nula 5, registram-se *chimia* e sua variante morfofonêmica *michia*, principalmente no RS. Essa variante, ao contrário do que ocorre com *geleia*, aumenta sua representatividade conforme se expande em sentido meridional do país, representando, aproximadamente, 58% das respostas no Estado do RS. O teste estatístico

empreendido confirmou a heterogeneidade no comportamento de *chimia* nos Estados do PR, SC e RS, demonstrando que essa forma lexical não apresenta comportamento homogêneo, o que possibilita a aceitação da hipótese nula 6, evidenciando a influência sul-rio-grandense. Além disso, a presença de *chimia*, nesta região do território, define uma área de isoléxica que evidencia traços do contato do português com as línguas de colonização, no caso, a alemã, uma vez que as obras lexicográficas consultadas apontam essa procedência do vocábulo. Em comparação com os dados do ALERS, os dados do ALiB confirmam esta mesma área de ocorrência, além da produtividade em outras localidades. Assume-se, portanto, que *chimia* é uma forma lexical que evidencia a existência de um possível *falar sulista* de influência sul-rio-grandense, pois é mais produtiva no RS e se caracteriza também pela influência de línguas de colonização no léxico português.

- (iv) Em contraste com a ocorrência da variante *chimia*, mais representativa no RS, as formas lexicais obtidas a partir do item *doce* (de fruta), agrupadas sob o rótulo “*formas genéricas*”, figuram, principalmente, no PR e em SC, além de ocorrências ínfimas em SP e RS e na região de pontos de controle do MT e RJ. O teste da hipótese 7 revelou que a presença de *formas genéricas* em SC e PR é homogênea, embora sua representatividade seja mais produtiva no PR, conforme se verifica no relatório 2a (vol. 2), não podendo, desse modo, se assumir como verdadeira a hipótese nula 7.
- (v) A variante *musse* ocorre exclusivamente no Estado de SC, evidenciando uma área lexical que parte do nordeste em direção ao sul do Estado. *Musse*, nas duas obras lexicográficas consultadas (HOUAISS; VILLAR, 2001) e (FERREIRA, 2004), não é uma forma dicionarizada na acepção buscada pela questão 177 do QSL. No entanto, observa-se que este item lexical comporta-se como uma variante regional que recobre o conceito do referente em pauta na mesma área geográfica

documentada na carta 357 do ALERS. Até o momento, uma pergunta pode ser formulada: Esta área de ocorrência de *musse* seria um indício da existência de uma subárea lexical no território investigado? Em busca dessa e de outras respostas é que se empreendem as próximas análises.

- (vi) A presença de *formas inadequadas* e formas lexicais sugeridas pela própria formulação da questão são documentadas com baixo índice de representatividade que, no entanto, são apresentadas, considerando-se o método quantitativo utilizado. Esses rótulos possibilitam a inclusão de formas não compreendidas adequadamente pelo informante que, no entanto, foram aceitas como resposta válida pelo inquiridor. Considerar essas formas como abstenção seria um equívoco.

Por fim, o mesmo panorama encontrado para a questão 001 se repete nos dados da questão 177. Ou seja, a área correspondente ao subfalar sulista, de Antenor Nascentes, para as variantes lexicais que recobrem o conceito da “pasta feita de fruta que se passa no pão”, pode ser dividida em duas grandes partes: (i) parte setentrional abrangendo os Estados das regiões Centro-Oeste e Sudeste, que apresentam menor número de heteroléxicas e uma ampla área em que ocorre exclusivamente a forma padrão do português “*geleia*”; e (ii) a parte meridional, principalmente, nos Estados da região Sul, em que o polimorfismo se acentua, ocorrendo maior número de sobreposição de formas lexicais. Esses seriam, portanto, os dois grandes falares: (i) o paulista e o (ii) sulista, de influência sul-riograndense e línguas de colonização. A esses dois falares, encontram-se duas subáreas lexicais no PR e em SC, com presença de “*formas genéricas*” e de “*musse*”.

4.3 QUESTÃO 156 – BOLINHA DE GUDE

A pergunta 156 do QSL – *coisinhas redondas de vidro com que os meninos gostam de brincar*, constante da área semântica Jogos e Diversões infantis, apresenta um total de 40 formas incluindo as variantes fonéticas e morfofonêmicas.

Desse modo, são necessários alguns 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 etc.;*
- (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.*

Observa-se, no relatório 3, que para algumas das formas agrupadas utiliza-se mais de um critério. Não obstante, fez-se necessária também a inclusão do rótulo “formas pouco produtivas” para agrupar as *hápax legomena*: *biribinha, bili, biroquê, boleja, pelota e pedrinha de vidro.*

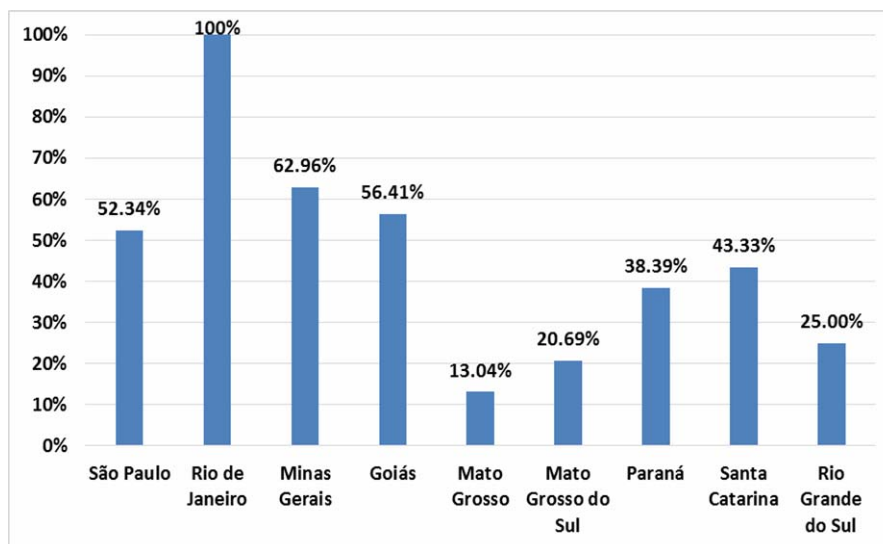
Assim, para a questão 156 do QSL, são documentados 663 registros distribuídos em 11 itens, além de duas abstenções de resposta e cinco respostas prejudicadas devido a problemas técnicos. Observa-se no relatório 3 que a variante mais produtiva é *bola de gude* (43,44%), seguida por *bolita* (18,70%), *bolinha/ bola de vidro* (12,97%) e *búrica* (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 carta linguística 3 apresenta a distribuição diatópica de sete das 12 formas lexicais registradas. Os itens que representam menos de 1% de produtividade, na cartografia, estão documentadas pelo item *outras* que compreende as variantes: *birola*, *fubeca*, *peca*, *peteca*.

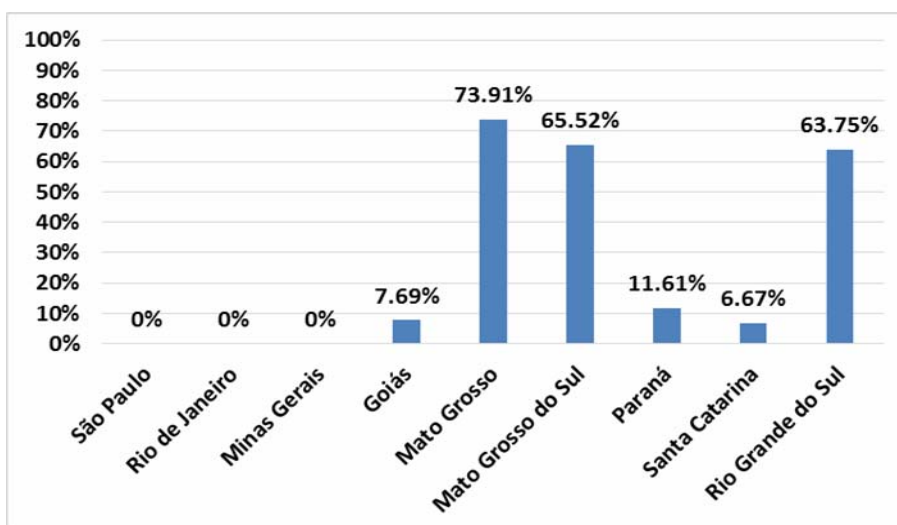
Comparando-se a carta 3 com a carta de arealidade da variante mais produtiva no *corpus* (Carta 3a) – *bolinha de gude* – observa-se que este item lexical está amplamente difundido pelo território do subfalar sulista e adjacências com diferentes índices de ocorrência. Observando-se o relatório geral por Estado (relatório 3a), verifica-se que *bolinha de gude* é a variante mais produtiva em seis dos nove Estados que compõem a pesquisa: GO, MG, PR, RJ, SC e SP, atingindo, no entanto, em apenas quatro desses Estados, produtividade acima de 50%, conforme se verifica na Figura 65. Os Estados em que há o menor índice percentual de *bolinha de gude* são o MT (13,04%), o MS (20,69%) e o RS (25%).

Figura 65 - Produtividade da variante *bolinha de gude* por Estado



Fonte: [□GVCLin] – Relatório 3a (vol.2)

A segunda variante mais produtiva – *bolita* – apresenta-se em seis dos nove Estados (RS, SC, PR, MS, MT e GO) com diferentes índices de produtividade, conforme se verifica na Figura 66.

Figura 66 - Produtividade da variante *bolita* por Estado

Fonte: [ΣGVCLin] – Relatório 3a (vol.2)

Esta forma lexical não ocorre nos Estados da região Sudeste, sendo representativa no Centro-Oeste e Sul. *Bolita* predomina em localidades matogrossenses e sul-mato-grossenses, apresentando-se como a variante majoritária. Entretanto, em Goiás, a produtividade de *bolita* atinge apenas 7,69%. Já na região Sul, SC apresenta o menor índice de produtividade de *bolita* (6,67%), e no PR, o item representa 11,61% das respostas. Porém, nas localidades gaúchas, *bolita* é a forma mais produtiva (63,75%).

A carta 3b mostra a difusão desta variante pelo território investigado. A área correspondente ao uso da variante *bolita* contempla quase todo o RS, principalmente, uma porção central, adentrando SC e PR pela região oeste destes Estados, ou seja, pela área fronteiriça do Brasil com a Argentina e o Paraguai.

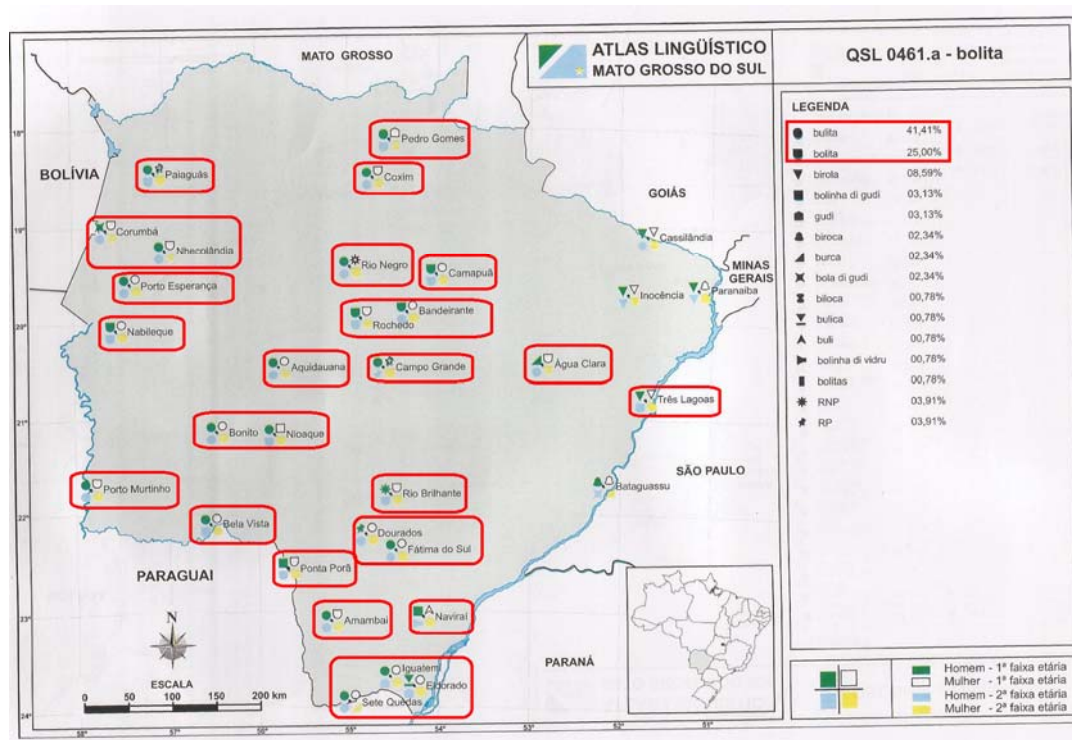
No Estado do MS, *bolita* ocorre em cinco dos seis pontos (com exceção do ponto 114 – Paranaíba – localizado próximo à fronteira política do MS com SP, MG e GO). Em GO, a variante é registrada apenas em um ponto linguístico no sul do Estado (124 – Jataí) e, em território mato-grossense, este item lexical está amplamente difundido, ocorrendo, inclusive, nos pontos de controle (103 – Aripuanã e 104 – São Félix do Araguaia), localizados no extremo norte do Estado.

Ferreira (2004) e Houaiss e Villar (2001) registram o verbete *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 Mesoregião Sudeste do Mato Grosso (CUBA,

2009) encontra-se registrada esta variante em 100% das respostas. O Atlas Linguístico do Mato Grosso do Sul (OLIVEIRA, 2007) também documenta os designativos para o referente na carta identificada como *bolita*. Reproduz-se, na Figura 67, essa carta (com adaptações), destacando-se as localidades em que houve a ocorrência de *bolita*.

Observa-se, desse modo, a ampla distribuição deste item pelo Estado, não ocorrendo, neste atlas, também na área correspondente à fronteira do MS com MG e GO, assim como no *corpus* do ALiB (Carta 3b). Ou seja, a cidade de Paranaíba novamente não é contemplada pela presença de *bolita*.

Figura 67 - Carta linguística BOLITA do ALMS (com adaptações)



Fonte: ALMS (OLIVEIRA, 2007)

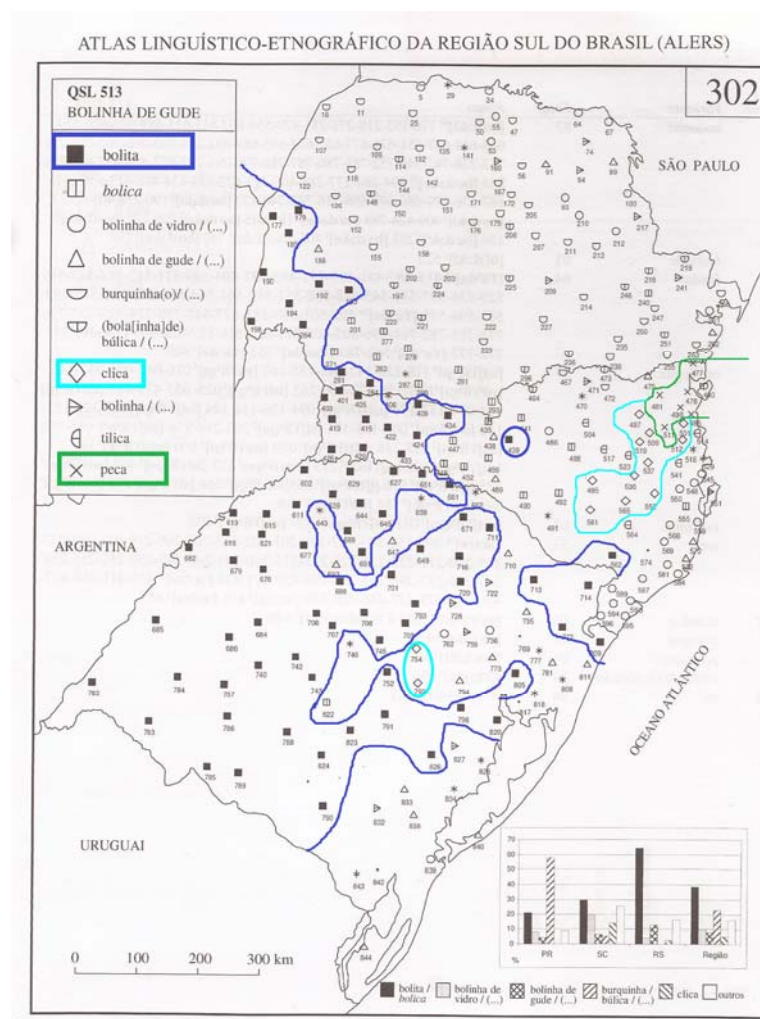
Reis (2006), na carta 218, apresenta a distribuição das variantes para *bolita* no município de Ponta Porã/MS, cidade localizada na fronteira do Brasil com o Paraguai. Nesse trabalho, encontra-se este item lexical como forma mais produtiva na maioria dos pontos linguísticos, concorrendo apenas em três localidades com a variante *bola/bolinha de gude*.

Augusto (2012) também apresenta um cartograma para os designativos da questão 156 do ALiB⁸¹. No cartograma de mesmo número, a autora documenta cinco ocorrências da variante *bolita*, sendo quatro delas no ponto 6 – Mineiros – localizado próximo à fronteira política de GO com o MT e MS e uma ocorrência no ponto 7 – Morrinhos –, localizado no centro-sul do Estado. Novamente, assim como nos dados do ALiB, observa-se a mesma área de ocorrência de *bolita* em GO, ou seja, no sul do Estado.

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 carta (Figura 68), com adaptações, verifica-se que a área correspondente à ocorrência da variante *bolita* coincide, em grande parte, com a documentada pelo ALiB (carta 3b), sobretudo no que se refere 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 de *bolita* diminui, figurando na área fronteira destes dois últimos Estados com a Argentina e o Paraguai.

⁸¹ Augusto (2012) baseou-se em muitos dos critérios metodológicos do ALiB para o desenvolvimento de seu trabalho, dentre eles, a opção pelo QSL na coleta dos dados. Mais detalhes sobre este e outros atlas podem ser obtidos em Romano (2013).

Figura 68 - Carta linguística 302 do ALERS (com adaptações)



Fonte: ALERS (ALTENHOFEN; KLASSMANN, 2011)

Esse panorama geolinguístico, tanto no ALERS quanto nos dados do ALiB, revela traços de empréstimos lexicais do português com línguas de contato fronteiriço, no caso o espanhol. No Dicionário da Real Academia Española (2001), o verbete *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 dura⁸²” ou mesmo, “cada una de estas bolas⁸³”. 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, com suas variantes fonéticas e morfofonêmicas, evidencia traços das línguas em contato nesta região do país. Nem Houaiss; Villar (2001) nem Ferreira (2004) apresentam a datação deste vocábulo na

⁸² “Jogo de crianças que se faz com bolas pequenas de barro, vidro ou outra matéria dura” (TN)

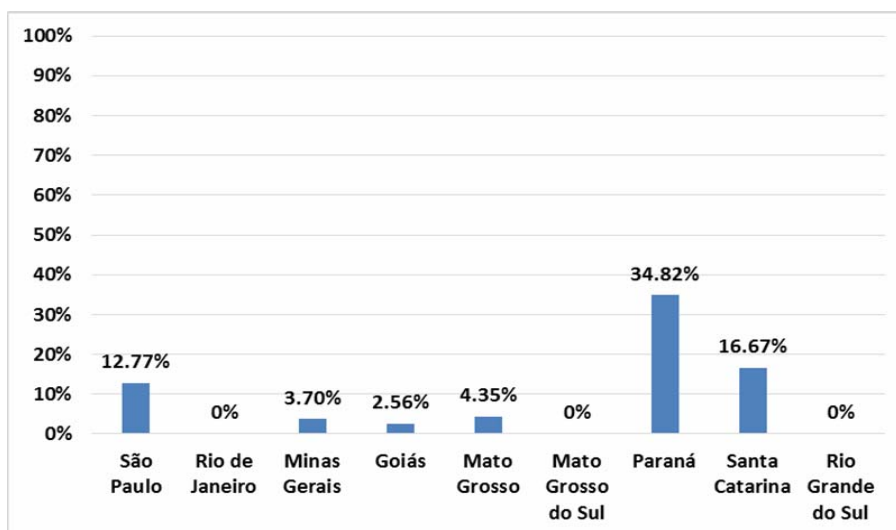
⁸³ “Cada uma destas bolas” (TN)

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. 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 carta 3c apresenta as áreas de ocorrência das duas variantes mais produtivas para a questão em pauta – *bolinha de gude* e *bolita*. Observa-se, desse modo, 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 dos Estados do PR e SC. Já a porção oeste do território investigado apresenta dois panoramas: (i) uma faixa territorial correspondente à cor roxa no mapa que revela linhas de heteroléxica, ou seja, que delimitam a coocorrência das duas formas lexicais (*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 este item, compreendendo o RS, numa faixa que percorre do sudoeste deste Estado, adentra o interior e chega ao norte gaúcho. No MS, esta 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 mato-grossense; e no ponto de controle 103 (Aripuanã), localizado fora da área do subfalar sulista, no noroeste do Estado.

O item *búrica* e suas variantes morfofonêmicas representam 12,67% do *corpus*. Trata-se de uma forma lexical que tem ampla distribuição diatópica concentrando-se principalmente no PR, no interior de SC e oeste do Estado de SP, além de ocorrer em uma localidade no sul de GO e nos pontos de controle 103 (Aripuanã/MT) e 145 (São João Del Rei/MG), conforme se verifica pela carta 3d.

Dos Estados em que ocorre *búrica*, o PR é o que apresenta o maior índice de representatividade (34,82%), seguindo-se SC (16,67%) e SP (12,77%), conforme se visualiza na Figura 69.

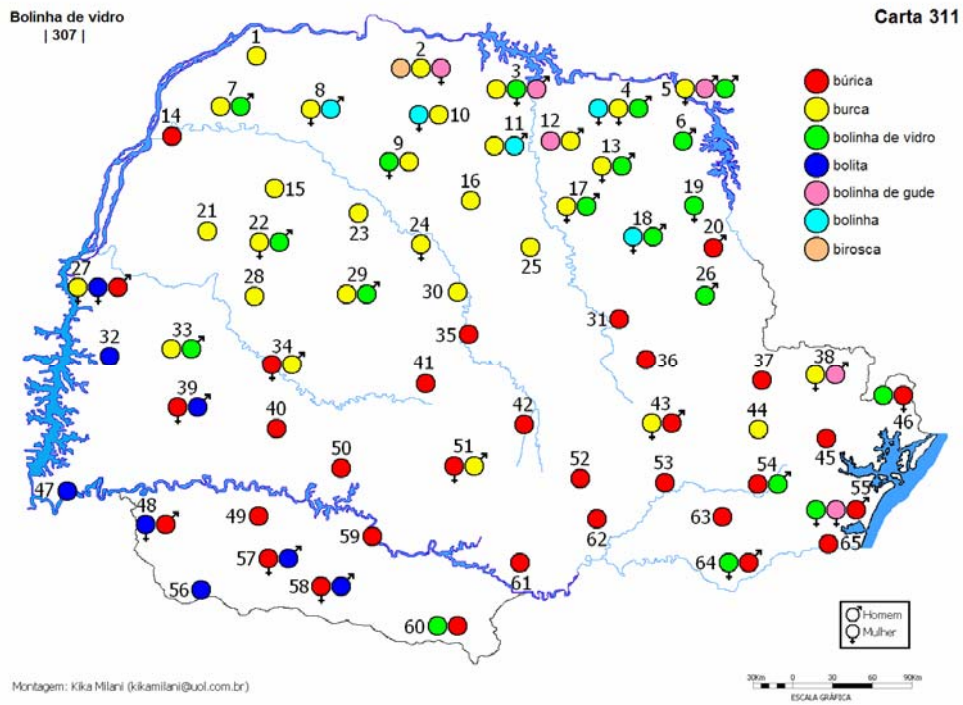
Figura 69 - Produtividade da variante *búrica* por Estado

Fonte: [ΣGVCLin] – Relatório 3a (vol.2)

O ALPR II (ALTINO, 2007) documenta na carta 311 as variantes lexicais para o referente em pauta. Observa-se que foi considerado um total de sete formas lexicais com diferentes índices de ocorrência (Figuras 70 e 71). Dentre outras formas lexicais, a pesquisadora não considerou as formas *búrica* e *burca* como variantes fonéticas do mesmo item. Assim, verifica-se que a variante *burca* ocorre, principalmente, na região norte do Estado (Paraná Moderno), enquanto *búrica* apresenta-se amplamente distribuída pelo centro-sul paranaense (Paraná Tradicional). Juntando as duas formas lexicais, pode-se notar que *búrica* predomina no território paranaense.

Tomando-se como referência o percentual de produtividade de *búrica/burca*, observa-se que juntas essas variantes representam 67,63% de frequência no *corpus* do ALPR II.

Figura 70 - Carta linguística 311 do ALPR II



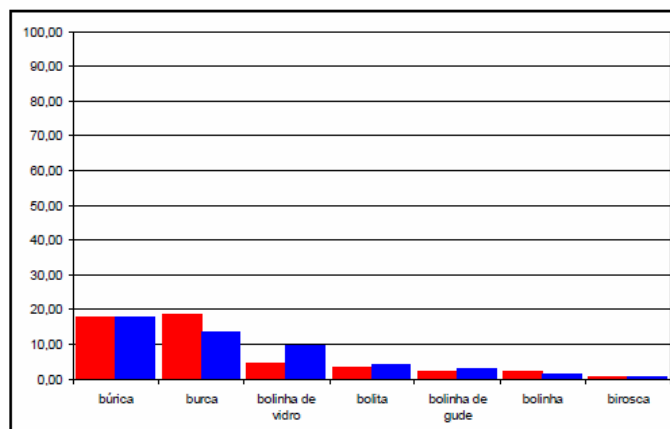
Fonte: ALPR II (ALTINO, 2007)

Figura 71 - Notas da carta linguística 311 do ALPR II

Ocorrências Únicas – Carta 311

Ponto 64 – Inf. A: Peça
 Ponto 64 – Inf. B: Pelote

Carta 311 (145 ocorrências)	Mulher	Homem
búrica	26 (17,93%)	26 (17,93%)
burca	27 (18,62%)	20 (13,78%)
bolinha de vidro	7 (4,83%)	14 (9,66%)
bolita	5 (3,45%)	6 (4,14%)
bolinha de gude	3 (2,07%)	4 (2,76%)
bolinha	3 (2,07%)	2 (1,38%)
biroasca	1 (0,69%)	1 (0,69%)



Fonte: ALPR II (ALTINO, 2007)

A carta linguística 3e apresenta a área lexical em que ocorrem as variantes *burca/burquinha/buque* sem agrupá-las às outras formas fonéticas e morfofonêmicas. Comparando-se os dados do ALPR II (carta 311) com a carta 3e, é possível identificar a mesma área de ocorrência de *burca* no território paranaense. Ainda nesta carta, observa-se que *burca* adentra as localidades do oeste do Estado de SP, além de apresentar ocorrências isoladas no sul de GO (ponto 124 – Jataí) e MG (ponto 145 – São João Del Rei).

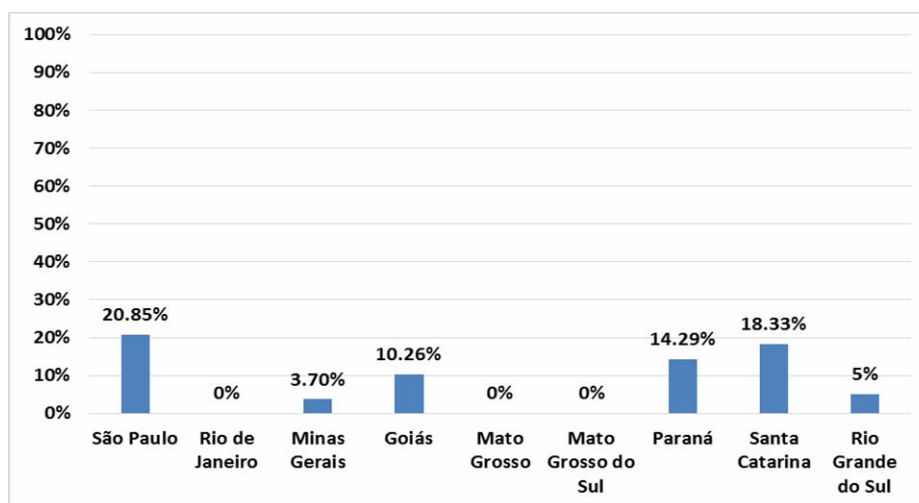
Separando-se *burca/burquinha/buque* das outras formas e representando-as na mesma carta linguística, verificam-se, conforme a carta 3f, áreas em que se usa tanto *burca* quanto *búrica* (e formas agrupadas) representadas pela cor verde. Nesta carta, a cor azul representa a área em que se usa somente a variante *búrica* (e formas agrupadas) correspondente aos pontos localizados no centro-sul do PR e o uso de *burca* no noroeste desse Estado, situação análoga ao que registra o ALPR II. Ou seja, há variantes fonéticas do item *búrica* que revelam áreas lexicais no Estado do Paraná⁸⁴.

Ademais, sem adentrar a especificidades na análise desta carta do ALPR II, cabe notar a presença de *bolita* no sudoeste do Estado (Figura 70), assim como ocorre nos dados do ALERS e do ALiB, o que confirma esta área lexical dentro do território investigado, como rota de passagem dos gaúchos rumo à região Centro-Oeste, e como área no Estado do PR que mais recebe influência do português em contato com o espanhol.

*Bolinha e Bola/bolinha de vidro*⁸⁵ representam 12,97% das respostas e estão presentes em seis dos nove Estados que compõem a pesquisa, com diferentes índices de representatividade, restringindo-se aos limites do subfalar sulista, ou seja, não ocorreram em localidades de controle (Figura 72).

⁸⁴ A área de ocorrência de *burca* no norte paranaense caracteriza-se como uma área lexical na Região Sul do Brasil, de acordo com Romano; Aguilera (2014, no prelo). A esta área Koch (2000) e Altenhofen (2005) chamaram de *zona lateral do Paraná do norte*. Romano; Aguilera (2014 no prelo) apresentam uma descrição detalhada dos padrões de variação lexical na região Sul do país, associando aspectos sócio-históricos e a relação dos dados analisados aos trabalhos análogos com vistas a identificar áreas lexicais nesta região administrativa.

⁸⁵ Pelo fato de *bolinha* ocorrer na mesma área geográfica de *bolinha de vidro*, optou-se por agrupar ambas as variantes em um mesmo item.

Figura 72 - Produtividade da variante *bolinha de vidro* por Estado

Fonte: [ΣGVCLin] – Relatório 3a (vol.2)

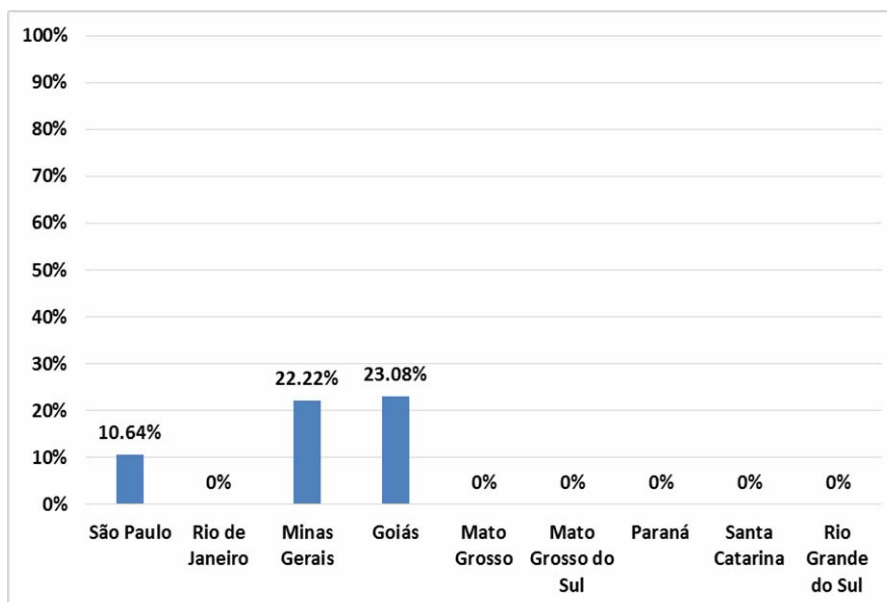
Dentre os Estados em que a variante é mais produtiva, encontram-se SP (20,85%) e SC (18,33%). Seguem-se os Estados do PR (14,29%), de GO (10,26%) e, com menor índice, RS (5%) e MG (3,70%). A distribuição diatópica desta variante é esparsa, concentrando-se, principalmente, no Estado de SP e PR, conforme se visualiza na carta 3g. Uma primeira conclusão que se poderia tirar observando a carta é que a variante se distribui a partir do Estado de SP, adentrando o PR até a parte central de SC, podendo-se, inclusive, afirmar que *bolinha (de vidro)*, possivelmente, constituiria uma variante que tipifica o falar paulista. Entretanto, a ausência desta forma lexical em grande parte do território do sul de MG e Triângulo Mineiro e também a ausência de uma maior distribuição do item no sul de GO enfraquecem tal assertiva, uma vez que, comprovadamente, são atestadas as influências de paulistas nestas regiões.

A variante *bilosca* e formas agrupadas ocorrem, na área investigada, somente nos Estados de GO, MG e SP, com diferentes índices de representatividade, conforme a Figura 73.

Em MG e GO, esta variante é a segunda forma lexical mais produtiva e em SP, apresenta-se como o quarto item mais recorrente. Esta variante forma uma área lexical que abrange o Triângulo Mineiro, conforme a carta 3h, adentra o sul e oeste de GO, além de se expandir para as localidades do noroeste paulista, seguindo até o oeste deste Estado na fronteira com o PR. Fora da área do subfalar sulista, *bilosca* e variantes ocorrem também em três pontos de controle:

ponto 121 (Formosa-GO), e, em MG, nos pontos 130 (Unaí) e 145 (São João Del Rei).

Figura 73 - Produtividade da variante *bilosca* por Estado



Fonte: [ΣGVCLin] – Relatório 3a (vol.2)

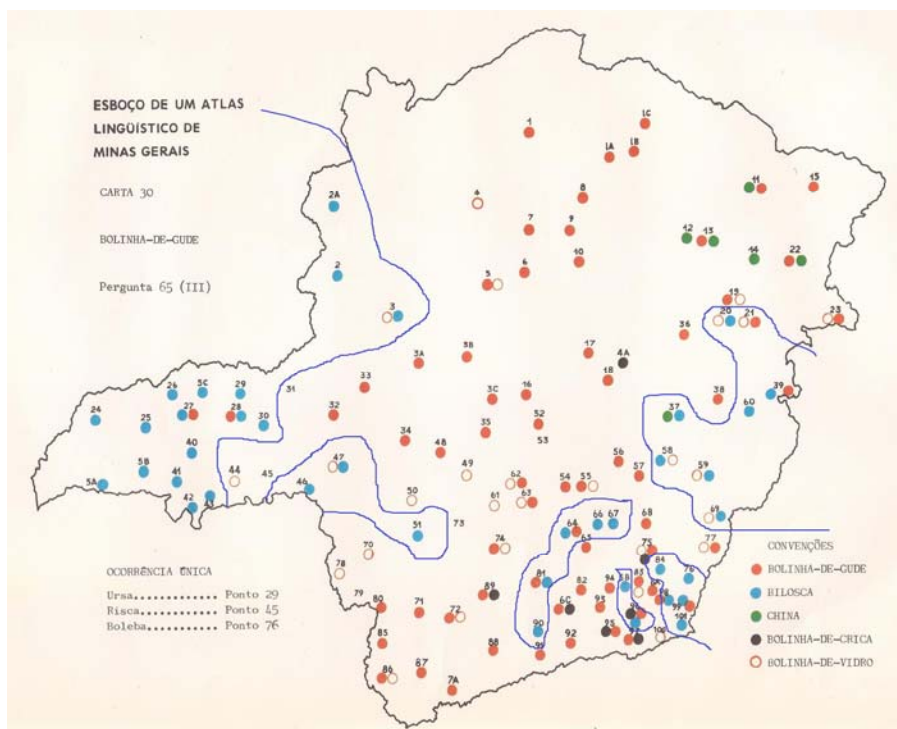
Em Ferreira (2004), o verbete *bilosca* encontra-se registrado como um brasileirismo de MG na acepção de *gude*, ao passo que em Houaiss e Villar (2001) consta como um brasileirismo de GO. Desse modo, os dados do ALiB parecem confirmar as informações de Houaiss e Villar (2001), pois a variante tem presença significativa não apenas em MG, mas também nas localidades goianas localizadas na porção oeste do Estado. Não obstante, os dados revelam a influência de mineiros no Estado de SP evidenciando a semelhança linguística do noroeste e oeste deste Estado com o Triângulo Mineiro, pelo menos no que se refere ao uso da variante *bilosca*.

Nota-se que esta forma típica de MG é encontrada no *Esboço de um Atlas Linguístico de Minas Gerais* – EALMG (RIBEIRO et al., 1977) como a segunda variante mais produtiva, conforme se verifica na Figura 74 que apresenta a carta nº 30 do referido atlas.

No EALMG (RIBEIRO et al. 1977), *bilosca* encontra-se em duas áreas bem definidas: (i) no oeste e no Triângulo Mineiro e (ii) no leste do Estado, além de pontos localizados no sudeste de MG. Isso prova que, na década de 70, o item já fazia parte do vocabulário ativo dos mineiros. Nos dados do ALiB, essas mesmas áreas de ocorrência se confirmam, uma vez que há grande distribuição de

bilosca na fronteira do Estado com GO e também a presença do item no ponto de controle 145 – São João Del Rei⁸⁶.

Figura 74 - Carta 30 do EALMG (com adaptações)



Fonte: EALMG (RIBEIRO et al., 1977)

A ocorrência e a expansão deste item no território paulista, poderiam ser explicadas pelo movimento interno de migração ocorrido durante o ciclo do café no Brasil, em que os mineiros rumavam às novas fazendas de café no oeste paulista e norte do PR. A ocorrência deste item em localidades como Presidente Epitácio (ponto 161), Teodoro Sampaio (164) e Presidente Prudente (Ponto 165) confirma esta hipótese, que pode ser ratificada pela carta 311 do ALPR II (Figura 70), em que se verifica a presença da variante *birosca* no ponto 2 – Santo Inácio – localizado no extremo norte paranaense.

A forma *clica* e variantes agrupadas apresentam baixo índice de ocorrência no *corpus* geral (1,81%) e ocorrem, esparsamente, em quatro Estados, com diferentes índices: RS (5%), SC (8,33%), PR (0,89%) e MG (3,70%). Conforme se verifica na carta 3i, *clica* apresenta-se em pontos isolados na área do subfalar sulista. No território gaúcho, ocorre exclusivamente no ponto 241 (Santa Cruz do

⁸⁶ Dados os objetivos do ALiB, atlas nacional, a rede de pontos não é tão densa quanto a dos atlas estaduais. Considerando esta particularidade, pode-se deduzir que se a rede fosse mais densa, a presença do item no território, possivelmente, seria maior.

Sul). Em SC, a variante está presente nos pontos 227 – Blumenau e 228 – Itajaí. No PR, apresenta-se no ponto 217 – São Miguel do Iguaçu e em MG está registrada em Lavras (ponto 144).

Vieira (2004, p.8), em trabalho desenvolvido a partir dos dados do ALERS com vistas a discutir traços de bilinguismo no léxico catarinense, aventa a hipótese de *clica*, forma não dicionarizada ser uma derivação de *clique* (clicar) que, em alemão, é *klicken*. Essa hipótese da autora considera o fato de *clica* ocorrer em Blumenau, cidade de colonização alemã.

De fato, verifica-se a presença germânica na colonização dessas localidades em que ocorre *clica* no *corpus* do ALiB (Santa Cruz do Sul/RS, Blumenau e Itajaí/SC e São Miguel do Iguaçu-PR),. Desse modo, possivelmente, por extensão de sentido, os informantes utilizem esta forma para denominar o referente dado o caráter onomatopaico do vocábulo. Na carta linguística do ALERS (Figura 68), *clica* ocorre na mesma área geográfica em que foi registrada pelo ALiB, no que se refere ao Estado de SC. Ou seja, área em que a maioria das localidades foi fundada sob a influência do processo de imigração germânica. Naquele atlas regional, porém, a variante não é registrada no oeste paranaense.

Por outro lado, verifica-se que a área de ocorrência de *clica* não se restringe às localidades do sul do país. Sabe-se que a imigração germânica foi mais significativa na Região Sul do Brasil, porém, esta não foi a única que atraiu estes imigrantes, pois como se sabe, o ES também foi um dos destinos dos colonos à época da grande imigração alemã, além, obviamente, de localidades específicas no RJ, MG e SP.

Um resquício da influência de imigrantes alemães em MG, no que se refere ao léxico, pode ser encontrado na presença de *clica*, em Lavras. Todavia, na década de 70, o EALMG já documentara esta variante, no sul mineiro e na região da Zona da Mata (Figura 74), em localidades, como Juiz de Fora, município que recebeu influência direta da imigração germânica, como documentam os registros históricos (LESSA, 1985). Nos dados do ALiB, que tem uma rede de pontos menos densa, a variante ainda está presente nesta região linguística, embora com menor representatividade.

A carta 3 apresenta o item “outras” como rótulo que agrupa as variantes *birola*, *peca*, *peteca*, *fubeca* e *formas pouco produtivas*. Fazendo a disjunção dessas variantes sem incluir formas pouco produtivas, a carta 3j apresenta

as áreas de ocorrência das quatro primeiras variantes. Dentro da área do subfalar sulista, registra-se a forma lexical *peca* no ponto 225 (São Francisco do Sul-SC), ou seja, na mesma área linguística em que ocorre nos dados do ALERS (Figura 68). Uma consulta ao banco de dados deste trabalho permite verificar que o item é bastante produtivo nesta localidade representando 66,7% das respostas válidas, num total de quatro ocorrências. *Peca* apresenta semelhança formal com o item *peteca*, variante típica da região Norte do Brasil, conforme demonstra o trabalho de Romano e Isquierdo (2015 no prelo)⁸⁷. Desse modo, seria *peca*, forma típica de uma única localidade catarinense, uma redução de *peteca* com a síncope da sílaba tônica? Ou um item não teria relação com o outro?

No *corpus* sob análise, *peteca* ocorre somente no ponto de controle 104 – São Félix do Araguaia-MT (Carta 3j). A presença deste item nessa localidade justifica-se pela proximidade geográfica e, conseqüentemente, linguística com o Subfalar Amazônico, conforme a nomenclatura de Nascentes (1952). Vale notar que *peteca* também é atestada nesta região por Portilho (2013), o que caracterizaria uma possível expansão do Subfalar Amazônico às localidades adjacentes.

A variante *peca* não está dicionarizada nem em Houaiss e Villar (2001) nem em Ferreira (2004), ao passo que *peteca* já se encontra na acepção de *gude* em ambas as obras lexicográficas. Há de se fazer estudo mais aprofundado para responder a estes questionamentos, porém, cabe ressaltar que *peca* ainda continua presente no litoral norte de SC, confirmando a área de ocorrência do item no ALERS (ALTENHOFEN; KLASSMANN, 2011) (Figura 68).

As *hápax legomena* (*biribinha*, *bili*, *biroquê*, *boleja*, *pedrinha de vidro* e *pelota*) apresentam baixa incidência no *corpus*. Distribuem-se em pontos do RS, SP, MG e MT.

Em Bagé/RS (ponto 248), registra-se a variante *boleja*, conforme o relato transcrito *ipsis litteris*⁸⁸:

⁸⁷ Trata-se de um artigo elaborado na disciplina Lexicologia e Lexicografia cursada no ano de 2011 no curso de mestrado em Estudos da Linguagem da UEL. Dada a qualidade do texto e a relevância dos dados, a professora da disciplina Dr^a. Aparecida Negri Isquierdo propôs ao Comitê do ALiB a inclusão deste trabalho no terceiro volume do Atlas Linguístico do Brasil. O texto discute a distribuição diatópica das variantes lexicais para a questão 156 do QSL na rede de pontos das capitais.

⁸⁸ Informante 3 do ponto 248 (homem, faixa etária II, de Bagé/RS).

INF.- Bolinha de gude.

INQ.- Tem outros nomes aqui?

INF.- Boleja. Aqui na fronteira é boleja ou bolita, lá no Uruguai é bolita ou boleja, aqui no Brasil é bolinha de gude.

Ou seja, o informante fornece informações diatópicas acerca do uso das variantes utilizadas na região de fronteira do Brasil com o Uruguai. Já no ponto 172 (Bragança Paulista/SP), é registrado o item *bili*. O informante que mencionou esta variante não teceu comentários e apresentou outros designativos mais comuns no uso. Possivelmente, *bili* é uma variante fonética de *bila*, forma comum na fala de informantes de Fortaleza-CE e dicionarizada na acepção de *bolinha de gude*, conforme atestam na carta L18b do ALiB (CARDOSO et al., 2014) e Romano e Isquerdo (2015 no prelo). É interessante notar que *bila* é uma forma comum na Paraíba, como bem retrata o Atlas Linguístico da Paraíba (ARAGÃO; BEZERRA de MENEZES, 1987) na carta linguística 106. No ALiB (carta l18b), o item também está registrado em João Pessoa. A presença de *bila* no corpus investigado, embora como ocorrência única, seria um resquício do processo migratório de nordestinos em direção ao Estado de São Paulo? Essa pergunta carece de mais estudos e aponta para a influência dos processos de movimento humano no léxico do PB.

Ainda no Estado de SP, no ponto 153 – Barretos –, está registrada a variante *biribinha*. Em MG (ponto 137 – Uberlândia), documenta-se *pedrinha de gude* e no ponto 136 (Patos de Minas) registra-se uma ocorrência de *biroquê*. Por fim, no MT, no ponto 111 (Alto Araguaia), documenta-se *birola*.

4.3.1 Tratamento Estatístico

Foram três as variantes submetidas aos tratamentos estatísticos devido à produtividade significativa no *corpus* e à distribuição diatópica observada nas cartas linguísticas analisadas, a saber: *bolinha de gude*, *bolita* e *búrica*.

As hipóteses formuladas para validação estatística são:

- Hipótese nula 8: a variante *bolinha de gude* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 9: a variante *bolita* apresenta distribuição de produtividade homogênea nas regiões Centro-Oeste e Sul.

- Hipótese nula 10: a variante *bolita* apresenta distribuição de produtividade homogênea nos Estados de MT, MS e RS.
- Hipótese nula 11: a variante *búrica* apresenta distribuição de produtividade heterogênea nos Estados de SP, PR e SC.

A exemplo das questões anteriores, foi necessário verificar o comportamento da normalidade das três variantes nos Estados contemplados pelo estudo. Desse modo, de acordo com os resultados do teste KS, a variante *bolinha de gude* não apresenta aderência à distribuição normal apenas no Estado de SP. No caso da variante *bolita*, foi verificada aderência à normal apenas nos Estados de MS, MT e GO. Nos Estados de MG e de SP, não houve ocorrência deste item lexical. Para a variante *búrica*, a normalidade foi verificada nos Estados de GO, PR e SC, não ocorrendo em MT, MS, MG e RS.

Este panorama evidencia a aplicação de testes não-paramétricos para a validação das hipóteses. Dessa maneira, foi aplicado o teste Kruskal-Wallis para as hipóteses 8, 10 e 11 e o teste Mann-Whitney para validação da hipótese nula 9, considerando-se os grupos necessários para a validação de cada hipótese.

Para o teste da hipótese 8, a aplicação de Kruskal-Wallis revela que as variâncias das ocorrências de *bolinha de gude* são heterogêneas em todos os Estados analisados. A partir da análise de variância, nota-se que há diferença significativa entre as médias de todos os Estados ($p < 0,001$), e a hipótese nula 8 pode ser aceita.

Ou seja, o tratamento estatístico dos dados revela que a variante mais produtiva no *corpus* e que apresenta ampla distribuição diatópica, conforme se verifica pela carta 3a, não tem distribuição de produtividade homogênea, o que evidencia o polimorfismo lexical para a questão em pauta. A carta 3k ilustra a heterogeneidade na distribuição do item pelo território investigado. Desse modo, verifica-se que *bolinha de gude* é mais frequente na porção leste do território investigado numa área que corresponde ao Estado de SP, adentra grande parte do nordeste e centro do PR, expande-se para o sul de MG, contemplando os pontos de controle (adjacências), refletindo-se também no sudeste e centro de GO. Esta área geográfica, em que o item é mais frequente, apresenta de 60 a 100% de produtividade, e caracteriza o possível *falar paulista*. Por outro lado, em Estados como RS, SC, sul e sudoeste do PR, nos Estados do MS e MT, verifica-se que a

variante apresenta menor índice de produtividade (0 a 40%), comportando-se, portanto, como uma área do território que não apresenta grandes influências do *falar paulista*. Desse modo, verifica-se que o uso da forma padrão diminui em frequência em sentido meridional, sobretudo devido a outras formas lexicais que caracterizam o possível *falar sulista* de influência sul-rio-grandense e do português em contato com o espanhol.

A análise da carta 3b levou à formulação das hipóteses nulas 9 e 10. Na hipótese 9, o objetivo é comparar dois grupos (regiões Sul e Centro-Oeste) quanto à produtividade de *bolita*. Assim, para verificar se as variâncias das ocorrências são homogêneas nas regiões analisadas, foi aplicado o teste de Mann-Whitney. O resultado mostra que o teste alcançou nível significativo, podendo-se assumir que as variâncias dos grupos não são homogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias das regiões ($p = 0,049$), e a hipótese nula 9 pode ser rejeitada. Desse modo, os resultados do teste não ratificam os apontamentos feitos na primeira parte desta análise sobre a semelhança linguística do Centro-Oeste como um todo em relação à região Sul. Analisando o conjunto das ocorrências, verifica-se que não há homogeneidade na distribuição do item *bolita* nas duas regiões. Esse comportamento se justifica pelo baixo índice de ocorrências da variante *bolita*, principalmente, nos Estados de GO, PR e SC, conforme atesta o relatório 3a (vol.2). Ou seja, os Estados das regiões envolvidas nesta hipótese apresentam comportamento diferenciado recebendo influências também do *falar paulista*.

Todavia, a hipótese nula 10 verifica a possível homogeneidade de *bolita* nos Estados de MT, MS e RS. O teste Kruskal-Wallis não alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são homogêneas. A partir da análise de variância, nota-se que não há diferença significativa entre as médias dos Estados ($p = 0,151$), e a hipótese nula 10 pode ser aceita, o que ratifica a semelhança linguística do MT e MS com o Estado do RS, sobretudo em decorrência de fluxos migratórios.

A carta 3l traz a distribuição de produtividade da variante *bolita* e formas morfofonêmicas pelo território investigado e ratifica as hipóteses testadas anteriormente, mostrando que: (i) os Estados do MS e MT apresentam comportamento parecido em relação ao RS, com 80% de produtividade do item na maior parte do território desses Estados; (ii) em GO, a variante não adentra o interior

do Estado figurando apenas no sudoeste goiano – ponto 124 (Jataí), onde foi registrado com 60% de produtividade; (iii) em SC e PR, esta variante ocorre no oeste dos Estados atingindo 100% de ocorrência em Toledo/PR (ponto 215). Depreende-se, desse modo, que o *falar sulista* está amplamente difundido, não em toda a Região Sul, mas pelos Estados do MS e MT, por meio de um corredor lateral do oeste catarinense e paranaense. Essa área de ocorrência recebe influência do contato do português com o espanhol, haja vista que a predominância do item ocorre na região de fronteira do Brasil com o Uruguai, Argentina, Paraguai e Bolívia.

Assim, para a distribuição das variantes que recobrem o conceito da *bolinha de gude*, o território investigado pode ser dividido em duas partes. Na porção setentrional encontra-se o *falar paulista* (presença da forma padrão) contemplando a maior parte dos Estados e, na porção meridional, verifica-se o *falar sulista*, contemplando RS, oeste paranaense e catarinense e Estados do MS e MT, conforme se observa, por exemplo, na carta 3c.

Os resultados relativos ao teste da última hipótese desta questão mostram que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias ($p = 0,009$), e a hipótese nula 11 pode ser aceita. Desse modo, o teste estatístico confirma as inferências feitas a partir da análise descritiva. Ao observar as médias entre SP (29) e SC (29,95), nota-se que se mantêm próximas, apresentando significativa diferença em relação à média das ocorrências encontradas no PR (43,74).

De fato, observando-se a carta 3m, verifica-se que *búllica* e suas variantes morfofonêmicas apresentam maior representatividade no PR, onde contemplam maior número de pontos, com o percentual oscilando, na maioria dos pontos, entre 60 e 80%, enquanto em SP, embora duas localidades apresentem 100% de ocorrência do item (ponto 162 – Adamantina e 167 – Bauru), há localidades em que o índice atingiu apenas 20% de representatividade. Visivelmente, observa-se uma distribuição heterogênea de *búllica* nos Estados investigados. *Búllica* teria, assim, expandido para as outras regiões a partir do PR, o que caracterizaria uma subárea lexical, localizada, principalmente, na região centro-sul deste Estado.

Sumariamente, tem-se o seguinte panorama acerca dos testes de hipóteses empreendidos (Quadro 5).

Quadro 5 – Síntese das hipóteses para a questão 156 do QSL

	ACEITA	REJEITADA
HIPÓTESE NULA 8 [#]	X	
HIPÓTESE NULA 9 ^{##}		X
HIPÓTESE NULA 10 ^{###}	X	
HIPÓTESE NULA 11 ^{####}	X	

bolinha de gude é heterogênea em todos os Estados

bolita é heterogênea nas regiões Centro-Oeste e Sul

bolita é homogênea nos Estados de MT, MS e RS

búrica é heterogênea nos Estados de SP, PR e SC

De posse desses resultados, algumas considerações podem ser feitas acerca da distribuição diatópica das variantes lexicais para a questão 156 do QSL:

- (i) A variante mais produtiva (*bola/bolinha de gude*) ocorre amplamente pelo território investigado. Porém, apesar da ampla distribuição diatópica (carta 3a) que atinge todos os Estados envolvidos, os resultados do teste estatístico e a carta 3k evidenciam que esta variante não apresenta comportamento homogêneo, indicando, portanto, que há heterogeneidade na distribuição do item, sobretudo devido à sua concorrência com outras formas lexicais, confirmando, desse modo, a hipótese nula 8.
- (ii) A segunda variante mais produtiva no *corpus* – *bolita* – revela a área de ocorrência de um *falar sulista* que se caracteriza pelo contato linguístico do português com outras línguas, no caso o espanhol, uma vez que, de acordo com os lexicógrafos consultados, o item é de origem platina. Todavia, o teste estatístico comparando a Região Sul e Centro-Oeste evidencia heterogeneidade na distribuição, não podendo confirmar a hipótese 9, em decorrência dos baixos índices do item em SC, no PR e em GO. Já comparando os Estados de MT e MS com o RS, confirma-se a hipótese nula 10, ou seja, não há uma diferença significativa na distribuição dos itens em cada um dos Estados. Desse modo, a estatística inferencial confirma as hipóteses formuladas ratificando a influência de gaúchos no

MT e no MS, como bem ilustra a carta 3l, além, obviamente, do contato do português com o espanhol em situação de fronteira com países como Bolívia, Paraguai, Argentina e Uruguai (carta 3b). Os resultados de outros trabalhos geolinguísticos têm apontado essas mesmas conclusões.

- (iii) No conjunto dos dados, o PR comporta-se como o Estado em que mais ocorre a variante *búrica* e formas morfofonêmicas, seguindo-se de SP e SC. Desse modo, o teste estatístico para verificar a diferença significativa na distribuição de produtividade de *búrica* nestes três Estados revela que o comportamento é diferenciado, ou seja, heterogêneo, podendo-se assumir como verdadeira a hipótese nula 11, apontando, desse modo, para a existência de subáreas lexicais que ora se expandem em direção ao denominado *falar paulista*, ora ao *sulista*.

Os testes estatísticos não foram necessários para todos os itens documentados, seja pela configuração na distribuição diatópica das variantes, seja pelo baixo índice de ocorrência no *corpus*. Desse modo, sobre as variantes que não foram tratadas com testes de hipóteses, pode-se concluir que:

- (iv) *Bolinha/ Bolinha de vidro* apresentam distribuição esparsa na área investigada não caracterizando uma área lexical, considerando-se os traçados das linhas de isoléxicas, pois o item é comum a diferentes regiões.
- (v) *Bilosca* forma uma área lexical que envolve três Estados: MG, GO e SP. A presença dessa variante no Estado de SP possivelmente pode ser justificada por fluxos migratórios de mineiros rumo ao oeste paulista, principalmente a partir da segunda fase do ciclo do café no Brasil. Não obstante, *bilosca* é uma forma amplamente documentada no EALMG e possivelmente pertence ao falar mineiro, que tem se inserido nos falares sulistas por meio de fluxos migratórios.

- (vi) *Clica*, variante de motivação onomatopaica, apresenta baixo índice de produtividade no *corpus* e figura em quatro localidades da região Sul que, coincidentemente, receberam influências germânicas no processo de formação, além de ocorrer em uma localidade mineira. Vieira (2004) atribui a ocorrência do vocábulo ao bilinguismo no sul do Brasil, região para qual migrou grande contingente de germânicos durante o período de ocupação e colonização do território. Entretanto, os dados geolinguísticos apontam que *clica*, já na década de 70, era uma forma bastante difundida pelo território de MG, como atesta a carta 30 do EALMG (RIBEIRO et al., 1977). Os dados do ALiB confirmam a presença de *clica* em MG revelando outra região brasileira que também recebeu influência dos povos germânicos, influências estas que, mesmo timidamente, se manifestam no léxico. São necessários, no entanto, estudos mais aprofundados sobre outro subfalar de Nascentes (1953), principalmente o mineiro, para verificar de forma mais precisa a influência de imigrantes alemães no Estado de MG, que, por ventura, se reflete no léxico. Não obstante, pode-se afirmar que *clica* evidencia a existência de subárea lexical no território investigado.
- (vii) *Peca* é uma forma recorrente e hegemônica no litoral norte de SC. Nos dados do ALiB, o item ocorre na mesma área geográfica documentada na carta 302 do ALERS. Supõe-se que *peca* seja uma variante fonética de *peteca* devido à semelhança formal de um item com outro.
- (viii) *Peteca*, por sua vez, ocorre no ponto de controle 104 – São Félix do Araguaia/MT –, ou seja, numa área linguística além do território do subfalar sulista. *Peteca*, como bem retratam obras lexicográficas e o trabalho em desenvolvimento de Romano e Isquerdo (2015, no prelo), é uma forma típica da região Norte do país. A ocorrência do item nas adjacências do falar sulista revela a importância do ponto de controle como parâmetro para uma delimitação mais precisa dos limites e alcances deste falar. Ou seja, a presença do item *peteca*, neste ponto de

controle, revela influências do falar Amazônico nesta área geográfica do MT.

- (ix) Acerca das *hápx legomena*, merece ser mencionada a ocorrência do item *bili*, que figura, exclusivamente, no Estado de SP. Possivelmente, *bili* é uma variante fonética de *bila*, forma comum na fala de informantes de Fortaleza-CE e dicionarizada na acepção de *bolinha de gude*, conforme atestam Romano e Isquierdo (2015, no prelo) e as obras lexicográficas de Ferreira (2004) e Houaiss e Villar (2001). É interessante notar também que *bila* é uma forma comum na Paraíba, como atesta o *Atlas Linguístico da Paraíba* (ARAGÃO; BEZERRA de MENEZES, 1987) na carta linguística 106. A presença de *bila*, embora como ocorrência única, possivelmente se deve à influência dos nordestinos no Estado de SP em decorrência de fluxos migratórios.

Os dados evidenciam, portanto, heterogeneidade linguística dentro do território investigado no que se refere às variantes para a *bolinha de gude* e apontam, novamente, a existência de dois falares: um *paulista (bolinha de gude)* e o *sulista (bolita)*. Ainda dentro desses dois grandes falares verifica-se a presença de subáreas lexicais, aqui caracterizadas pela variante *búrica*, que se localiza, principalmente, no centro-sul do PR, expandindo-se em um corredor central de SC. Outras variantes também delimitam subáreas lexicais intermediárias entre os dois falares, por exemplo, *clica* e *peca*, principalmente, na faixa leste e litoral norte de SC. Ademais, a configuração diatópica do item *bilosca* também dá indícios da possível expansão do subfalar mineiro, de Antenor Nascentes, que apresenta suas influências também na área geográfica delimitada neste trabalho como *falar paulista*.

O *falar paulista*, desse modo, distribui-se para a 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. Esta distribuição deste falar pode ser justificada pelos fluxos migratórios dos gaúchos rumo às novas fronteiras agrícolas do Centro-Oeste.

Todavia, para este item lexical específico, este falar está caracterizado pelo contato do português com outras línguas, no caso, o espanhol. Entretanto, os limites desses falares têm se mostrado fluidos e também cabe notar que cada um deles apresenta outras áreas lexicais. Por exemplo, na área do *falar paulista*, verificam-se áreas lexicais em que ocorre *bilosca* (*falar mineiro?*) e *búrica* (*falar paranaense?*); ao passo que, no *falar sulista*, encontram-se áreas em que ocorrem *peca* e *clica* (*falar catarinense?*).

4.4 QUESTÃO 132 - MENINO

A questão 132 do QSL, que visa a documentar as variantes lexicais para a criança de 5 a 10 anos do sexo masculino, apresenta um conjunto de oito variantes lexicais, além de formas inadequadas, abstenção de resposta e três gravações com problemas técnicos.

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.

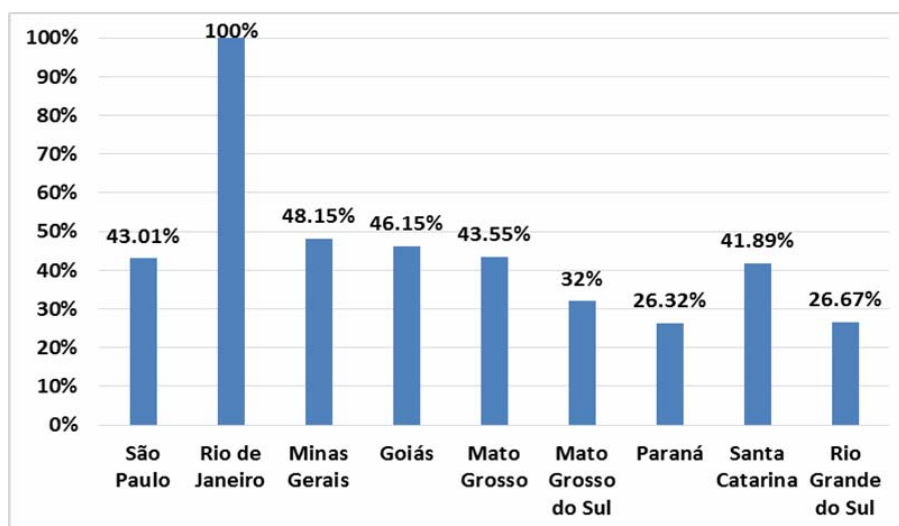
De acordo com o relatório 4 (vol.2), a variante mais produtiva, *menino*, representa 38,55% do *corpus*, seguida por *moleque* (20,14%), *guri* (17,15%), *garoto* (10,47%) e *piá* (9,55%). Em menor índice, encontram-se *pivete* (1,84%), *rapazinho* (1,73%) e com menos de 1% de produtividade ocorrem “formas inadequadas” (0,46%) e a *hápax legomena*, *bambino*, com apenas 0,12% de representatividade.

A carta 4 apresenta a distribuição diatópica das nove formas documentadas no *corpus*, mostrando maior concentração da variante *moleque* no território do Estado de SP, estendendo-se ao Estado de MG e GO. Por outro lado, a variante *guri* está amplamente distribuída pelo RS, adentrando os Estados do MS e MT. No PR e SC, nota-se maior distribuição da variante *piá*.

Como desmembramento desta carta geral, a carta linguística subsequente (4a) apresenta a área de ocorrência da variante mais produtiva no *corpus – menino*. Visualiza-se, desse modo, que a variante padrão apresenta-se em quase todas as localidades, não ocorrendo somente em três cidades gaúchas (ponto 236 – Passo Fundo, 247 – Santana do Livramento e 249 – São José do Norte) e em um município mato-grossense, ponto 105 (Diamantino).

Dentre os Estados investigados, com exceção do RJ cujos dados referem-se a duas localidades, de acordo com o relatório 4a, nenhuma outra unidade federativa atinge 50% de produtividade para a variante *menino*, conforme se verifica pela Figura 75.

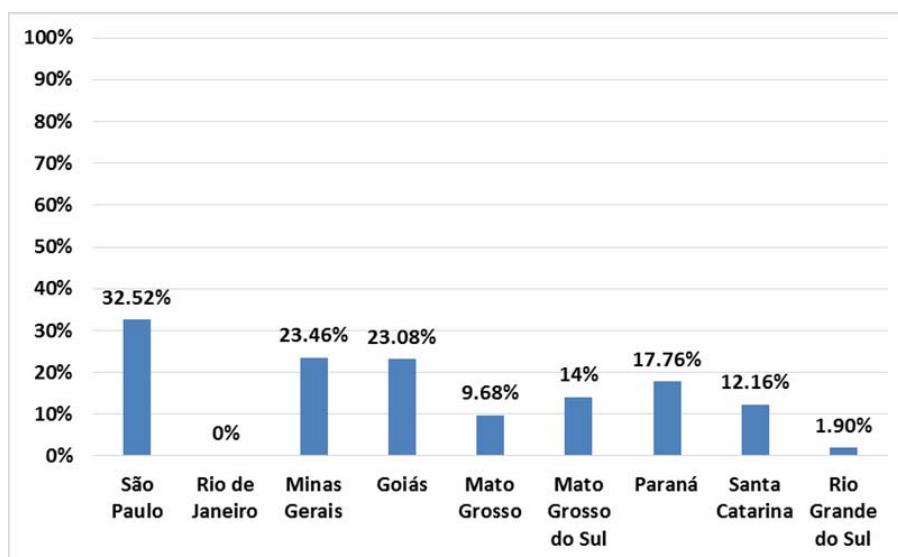
Figura 75 - Produtividade da variante *menino* por Estado



Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

Entre os Estados, MG e GO apresentam os maiores índices, 48,15% e 46,15%, respectivamente, seguindo-se o MT (43,55%), SP (43,01%), SC (41,89%), e, com menor percentual, encontram-se MS (32%), RS (26,67%) e PR (26,32%).

Quanto à variante *moleque*, segunda forma lexical mais produtiva no conjunto das respostas, observa-se que este item não ocorre no Estado do RJ, e, nos demais, apresenta-se com índices percentuais diferenciados, conforme se verifica na Figura 76, obtendo baixa produtividade no MT (9,68%) e, principalmente, no RS (1,90%).

Figura 76 - Produtividade da variante *moleque* por Estado

Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

A carta linguística 4b ilustra a área de ocorrência desta variante, que está distribuída em grande parte do território do subfalar sulista. No entanto, *moleque* não é uma forma difundida no RS, o que permite inferir que é uma variante que, possivelmente, tem como foco irradiador o Estado de SP, constituindo-se, dessa forma, em uma variante que caracteriza o *falar paulista*.

Desse modo, a partir de SP, a forma se expande para o sul de MG e Triângulo Mineiro, atingindo o Estado de GO e a porção oeste e sul do MT, além de adentrar, em grande parte, o território do MS. No PR, *moleque* segue, contemplando amplamente o território paranaense, não atingindo, no entanto, o sudoeste e o ponto 222 – Lapa, no sudeste deste Estado. A partir de um feixe central, a variante adentra o Estado de SC dividindo-o ao meio, além de ocorrer em cidades localizadas no litoral norte de SC e em dois pontos gaúchos (242 – Santa Maria e 243 – Porto Alegre).

No ALERS, a variante *moleque*, documentada nas cartas 270 e 271, encontra-se como o quarto item da legenda, não se constituindo, desse modo, como uma forma tão produtiva, o que é compensado pela ocorrência de variantes regionais nos Estados da região Sul.

Romano e Seabra (2014a) confirmam que o item lexical *moleque* 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, das quais se destacam as quatro primeiras: “1.Negrinho. 2.Bras. Indivíduo sem palavra, ou sem gravidade. 3.Bras. Canalha, patife, velhaco. 4.Bras. Menino de pouca idade” (FERREIRA, 2004). Houaiss e Villar (2001), por sua vez, apresentam 17 acepções para o verbete “moleque”, das quais 13 são para emprego como substantivo masculino e quatro como adjetivos. Desta última obra lexicográfica, 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).

O uso da variante moleque caracteriza também, para alguns informantes, desprestígio em relação à denominação dada à criança, recebendo, em algumas vezes ou em determinados contextos, conotação pejorativa, conforme se observa, por exemplo, na fala de um dos informantes⁸⁹.

INF.- Minino [μινινΥ].

INQ.- Tem outro nome pra menino, aqui?

INF.- Tem. Muleque [μυελεκι], pivete [πιεπετσι]. Depende do Estado da criança. Se a gente vê que é minino, a gente fala que é minino, né, mai' se a gente vê que é mai' crescidin, aí já... pivete, é muleque.

INQ.- Chamar uma criança de moleque, isso é ofensivo?

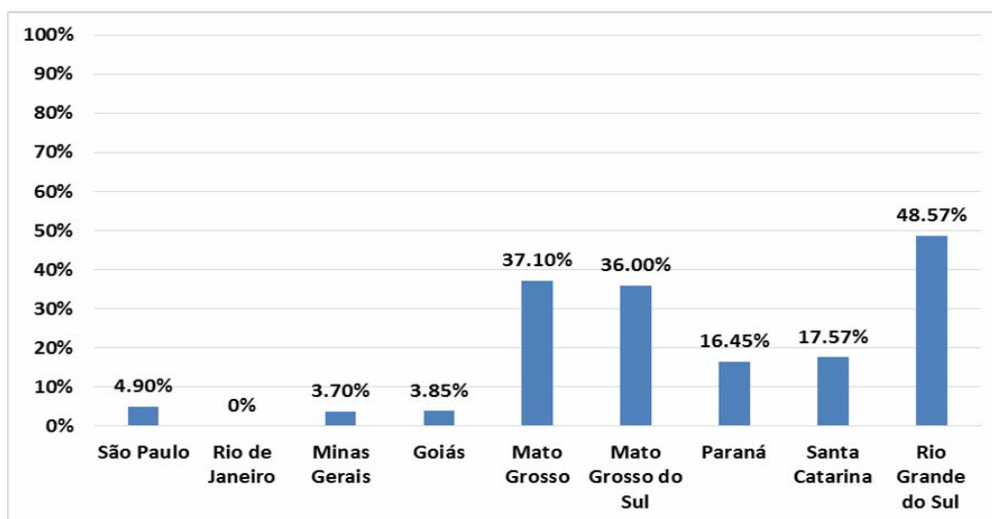
INF.- Ah, pra mim sei lá né, é falta de educação né, criança, criança é criança.

INQ.- Se você por exemplo tivesse um filho e falar “eu tenho um moleque lá em casa”.

INF.- Ai já não. Depende do lugar né, depende da hora, você não vai sair na rua e chamar o minino de muleque né, daí tem o pai do menino, o pai vai vim querer tirar satisfação, daí já tá ofendendo.

A terceira variante mais produtiva é *guri*, forma majoritária no RS (48,57%), o que compensa a baixa produtividade da variante *moleque* nas cidades gaúchas, conforme se verifica no relatório 4a. A Figura 77 ilustra esta distribuição indicando a produtividade dessa variante nos Estados. Verifica-se que *guri* se concentra, além do RS, sobretudo, nos Estados do MS e MT, 36% e 37,10%, respectivamente. Nos demais Estados, há decréscimo na produtividade da variante, representando 17,57% das respostas dos catarinenses e 16,45% das dos paranaenses. Em SP, MG e GO, a ocorrência do item é pequena, ficando abaixo de 5%.

⁸⁹ Informante 1 do ponto 135 (Homem, faixa etária I: 18 a 30 anos, de Uberlândia/MG).

Figura 77 - Produtividade da variante *guri* por Estado

Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

A carta 4c apresenta a distribuição espacial desta forma lexical no território investigado, apontando que esta 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, a presença de *guri* está em quase todo o Estado, não contemplando, entretanto, localidades do noroeste paranaense, região de colonização mais recente sob influência de mineiros, paulistas e nordestinos na primeira metade do século XX.

Por outro lado, esta 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):

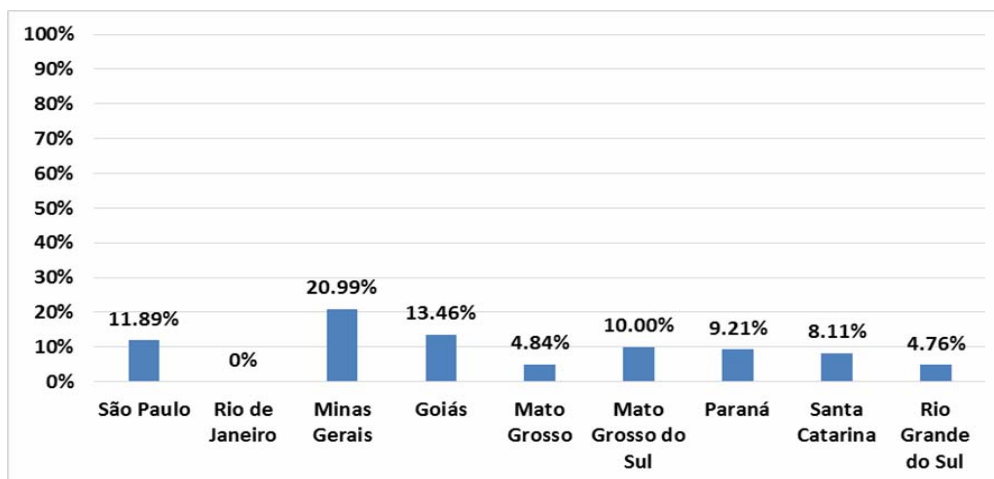
a presença da variante “guri” em determinadas localidades paulistas, tais como Itararé, Itapetininga e Sorocaba, pode ser explicada pela influência vinda do sul do país, haja vista que algumas dessas localidades eram pontos pertencentes à antiga rota proveniente do Caminho das Tropas, pela consequência do Tropeirismo no Paraná.

Não obstante, 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 de migração é decorrente das novas fronteiras

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

A variante *garoto*, por sua vez, não apresenta alto percentual de representatividade no RS, mas, em cada um dos Estados, traz índices de ocorrência diferenciados, conforme se observa na Figura 78.

Figura 78 - Produtividade da variante *garoto* por Estado



Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

Entre os Estados, MG (20,99%) e SP (11,89%) são os que apresentam o maior índice percentual da variante *garoto*, seguidos por GO (13,46%) e MS (10%). No PR e em SC, a forma ocorre em 9,21% e 8,11% das respostas, respectivamente, e, com menor produtividade, encontram-se o MT e o RS, cujo comportamento é semelhante, com 4,84% e 4,76%, respectivamente.

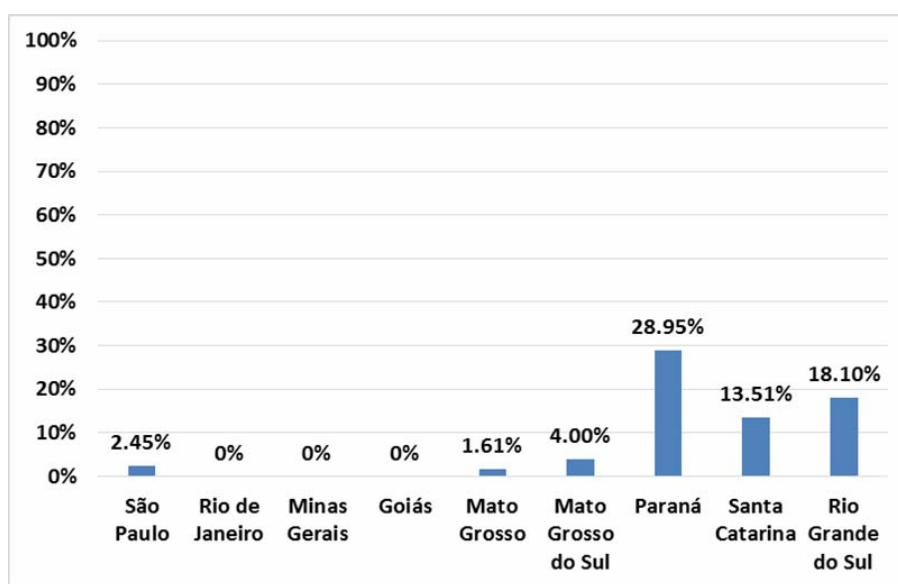
Na carta 4d, observa-se uma ampla área lexical que se estende pelo Estado de SP, e contempla toda a região de MG (inclusive pontos de controle) e grande parte do território goiano, com exceção do sul deste Estado. Esta área alcança MS e, em linha contínua, atinge duas cidades do sul do MT, pontos 106 e 108 (Poxoréu e Cuiabá). No PR, a área de isoléxica contempla as regiões central e norte, além da área do litoral.

A variante *garoto* ocorre também, isoladamente, em três pontos de SC: um localizado no oeste do Estado (ponto 229 – Concórdia) e dois no litoral – 228 (Itajaí) e 230 (Florianópolis). No RS, há a ocorrência isolada de *garoto* em quatro localidades, uma no norte do Estado (ponto 240 – Flores da Cunha), uma ocorrência no centro (ponto 242 – Santa Maria) e duas ocorrências no sudoeste

(ponto 239 – São Borja e 245 – Uruguaiana). Ainda, isoladamente, observa-se uma ocorrência no ponto de controle 103 – Aripuanã – situado no noroeste mato-grossense.

A variante *piá*, por sua vez, é mais frequente nos Estados da região Sul, principalmente, no PR, conforme se observa pela distribuição de produtividade na Figura 79.

Figura 79 - Produtividade da variante *piá* por Estado



Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

Em Estados como SP, MT e MS, a frequência deste item lexical é pequena, ao passo que, no PR, *piá* ocorre como variante majoritária obtendo 28,95% de produtividade, seguindo-se o RS (18,10%) e SC (13,51%).

A carta linguística 4e apresenta a distribuição deste item na constituição de uma ampla área lexical na região Sul, contemplando os três Estados. Esta área, no entanto, não contempla algumas cidades gaúchas, pontos como Erechim (235) e Passo Fundo (236), no norte do Estado, além dos pontos 245 – Uruguaiana (sudoeste gaúcho) 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 catarinense.

A distribuição espacial de *piá* também não contempla o noroeste paranaense, de colonização mineira e nordestina, conforme já foi apontado

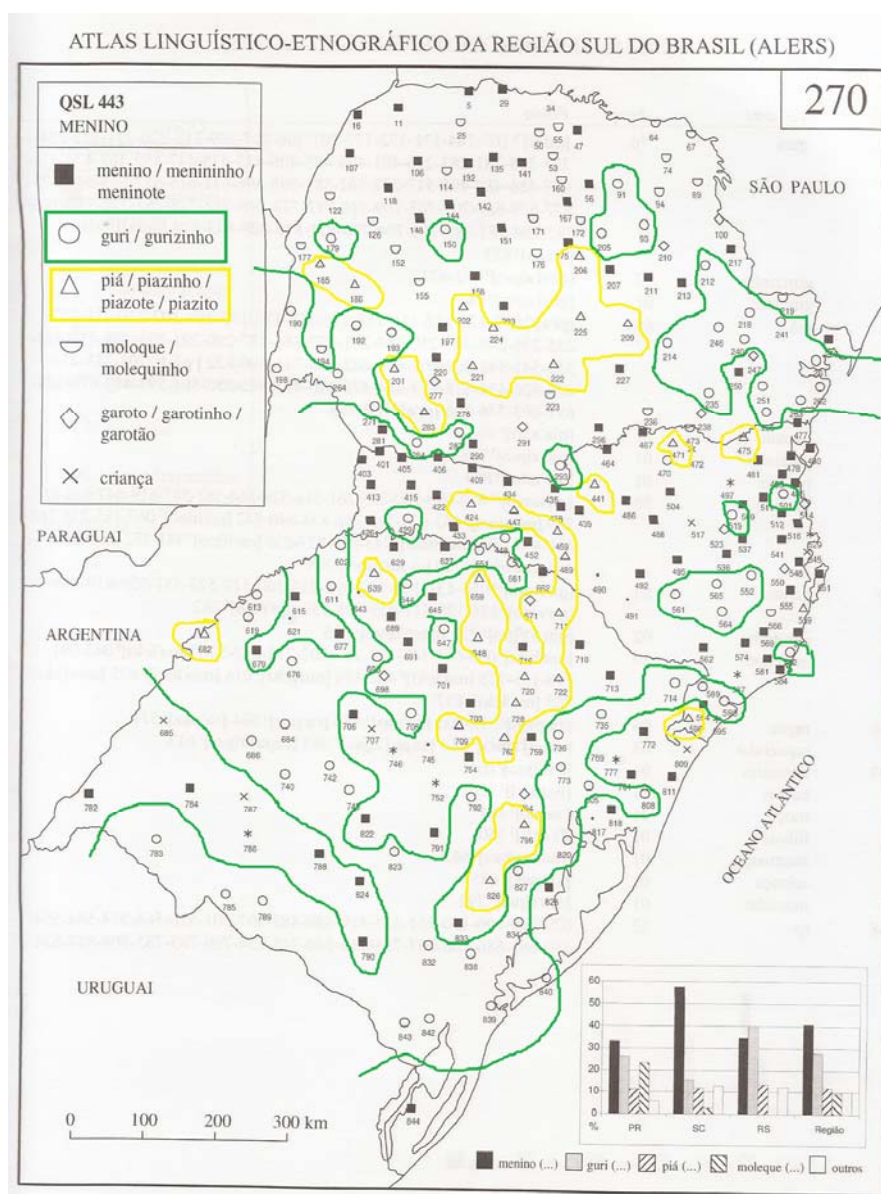
anteriormente, expandindo-se, no entanto, para uma localidade sul-mato-grossense – 117 (Ponta Porã) e cidades paulistas localizadas no Vale do Ribeira. Ou seja, assim como ocorre com a variante *guri*, a presença de *piá* nesta área do Estado de SP evidencia traços da história social das localidades, uma vez que esta forma típica da região Sul, possivelmente, teria sido trazida para essa região paulista por meio dos tropeiros, durante o ciclo do Tropeirismo no Paraná, persistindo até hoje no vocabulário ativo dos naturais desta região paulista.

O ALPR I e o ALPR II não apresentam cartas linguísticas que documentem os designativos para o referente em pauta, entretanto, o ALERS, nas cartas 270 e 271, registra as principais variantes e outras denominações, respectivamente. Na Figura 80, reproduz-se a carta 270 do referido atlas (com adaptações).

As isoléxicas desenhadas para a variante *guri* (verde) e *piá* (amarelo) permitem visualizar de forma mais clara as áreas de ocorrência dessas duas variantes regionais nos Estados da região Sul, demonstrando que ambas se apresentam nos Estados, ora formando amplas áreas lexicais, ora ocorrendo em pontos isolados.

Como bem retratam 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~ĩri* com a acepção que designa o “bagre novo” (tipo de peixe), por extensão de sentido, a criança. Ainda nesta entrada, o etimologista remete o usuário ao verbete “guiri - 'sm bagre' / curi 1587, guori datada aproximadamente 1631”. Houaiss e Villar (2001), em contrapartida, registram a datação de 1890, documentada no Dicionário Etimológico da Língua Portuguesa, de José Pedro Machado (1952). Para os lexicógrafos, dentre outras acepções, trata-se de um regionalismo brasileiro para designar o “menino”, a “criança”. Ferreira (2004), apesar de apresentar a mesma acepção para o vocábulo, traz outra etimologia. Segundo o dicionarista, o vocábulo “guri” também vem do tupi, porém com o sentido de pequeno, não fazendo, portanto, alusão ao peixe.

Figura 80 - Carta linguística 270 do ALERS



Fonte: ALERS (ALTENHOFEN; KLASSMANN, 2011) – com adaptações

Quanto à variante *piá*, de acordo com Houaiss e Villar (2001), o verbete refere-se ao menino indígena e, analogamente, é empregado para se referir ao menino mestiço de indígena com branco ou mesmo qualquer criança do sexo masculino. Ainda segundo os lexicógrafos, trata-se de um regionalismo de Santa Catarina e Rio Grande do Sul para designar o peão menor de idade que não é de raça branca (HOUAISS; VILLAR, 2001). Ferreira (2004) traz no verbete *piá* a informação de que se trata um brasileirismo para designar o índio jovem ou mestiço jovem de branco com índio; pequeno caboclo. Nessa mesma entrada, remete o usuário ao verbete “menino”, tratando-se, pois, de um sinônimo. Em Santa Catarina e Rio Grande do Sul 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). Assim, em consonância com as ideias de Romano e Seabra (2014a), as informações constantes dos dicionários apontam a influência de fatores sociais e históricos na constituição do léxico português.

Para alguns informantes que não são da região Sul, o uso da variante *piá*, em geral, é acompanhado de informações diatópicas, conforme se observa na fala de um informante mato-grossense⁹⁰.

INF.- Ah... nós já chama ele de Maciel: “ôh, Marcielzinho, vem cá” ele já cresceu...

INQ.- É, mas e se o senhor não sabe o nome dele, na rua...

INF.- Ah, não nói chamamo é na nossa, na nossa language, nós chamava assim: “ôh, minino, vem cá” né, nós chamava “ôh, minino [μινινΥ], vem cá.” Já na língua do pessoal aqui, daqui pra frente, principalmente, da região do Paraná chama piá, né: “ôh, piá [πια], vem cá”.

INQ.- Mas aqui é menino?

INF.- Aqui é menino, na nossa tradição (=tradição)...

Já a variante *guri*, para alguns informantes que declaram conhecer o item, em geral, a sua ocorrência está associada ao MT, conforme se observa nos inquéritos reproduzidos *ipsis litteris* a seguir.

INF.- Lá no Mato Grosso, eles chama de guri [γυρι].

INQ.- E aqui?

INF.- Aqui nós fala menino. Lá no Mato Grosso já chama de guri.

INQ.- Pra menino, só menino mesmo? Moleque não?

INF.- Ah, tem moleque [μυελεκι] também, falam, um fala moleque: “Ah, sai daí moleque!”... né, moleque, menino.

INQ.- E chamar alguém de moleque ofende?

INF.- Não.

INQ.- É normal.

INF.- É normal.

INQ.- Aqui também usa muito falar.

INF.- Aqui também, usa⁹¹.

INF.- Minino [μινινΥ], assim num tem, né...

INQ.- Menino?

INF.- É.

INQ.- Conhece por algum outro nome?

INF.- Parece que no Mato Grosso fala guri [γυρι], né.

INQ.- É?

INF.- É, o guri. Chama de guri.

⁹⁰ Informante 3 do ponto 103 (Homem, Faixa etária II: 50 a 65 anos, de Aripuanã/MT).

⁹¹ Informante 4 do ponto 165 (Mulher, Faixa etária II: 50 a 65 anos, de Teodoro Sampaio/SP).

INQ.- Aqui fala guri?

INF.- Não, aqui num chama de guri não, já vi falá em guri i é... no Mato Grosso, que eu tenho uma cumade lá e a gente ia pra lá, ela chamava as minina dela de guri.

INQ.- Aqui não?

INF.- Aqui num... nunca vi falá não⁹².

Ou seja, enquanto *piá*, para os informantes, caracteriza-se como uma forma lexical tipicamente paranaense, à variante *guri* é atribuído um caráter diatópico que não se associa, exclusivamente, aos gaúchos, uma vez que é reconhecida como forma legítima da fala de informantes sul-mato-grossenses e mato-grossenses, conforme se observa:

INF.- É piá [πιᾱ], é um monte de coisa.

INQ.- Aqui se usa piá?

INF.- Eh, sempre fala né, mas sempre quem usa é a pessoa de Santa Catarina né, os gaúcho que...

INQ.- Que moram aqui, né?

INF.- É.

INQ.- E o mato-grossense fala como?

INF.- Menino, é guri [γυᾱPι].

INQ.- Fala mais.

INF.- É guri, né.

INQ.- É⁹³.

Observando-se a carta 4f, nota-se uma ampla área na Região Sul em que ocorrem, indistintamente, as variantes *guri* e *piá*, o que ratifica o caráter regional atribuído a essas formas lexicais. *Guri* apresenta-se como forma exclusiva em áreas isoladas, principalmente, em SP, MG e GO. Em grande parte do território dos Estados do MS e MT, observa-se a ampla distribuição deste item. *Piá*, por sua vez, não apresenta distribuição exclusiva na área do território investigado e não há formação de grandes áreas lexicais, uma vez que ocorre, isoladamente, em pontos do RS, SC, PR e SP.

Observando-se, portanto, o comportamento e a distribuição de quatro variantes lexicais: *moleque*, *garoto*, *guri* e *piá*, podem-se fazer algumas considerações acerca da área em que ocorrem, caracterizando dois padrões de variação lexical na área correspondente ao subfalar sulista, de Antenor Nascentes: uma de influência paulista, denominada como *falar paulista* e outra de influência sul-

⁹² Informante 3 do ponto 123 (Homem, Faixa etária II: 50 a 65 anos, de Goiânia/GO).

⁹³ Informante 2 do ponto 115 (Mulher, Faixa etária I: 18 a 30 anos), de Campo Grande/MS).

rio-grandense e paranaense, denominado como *falar sulista*. Para comprovar a presença desses dois falares na área investigada, apresentam-se algumas cartas de arealidade, que pelo traçado de linhas de isoléxicas e heteroléxicas evidenciam áreas lexicais distintas.

Por exemplo, a carta 4g documenta a distribuição de *piá* e *garoto* em que se verificam áreas de coocorrências das duas formas (cor vermelha), principalmente, no PR. A ampla área lexical de *piá* (cor preta) contempla uma faixa que vai do RS ao PR, ao passo que a área de *garoto* (cor amarela), a partir do Estado de SP, difunde-se pelo noroeste paranaense, MS, sul de MG, de onde se expande para GO.

Piá e *moleque* apresentam comportamento semelhante aos resultados encontrados na carta 4g e também evidenciam a presença dos dois falares (*o paulista e o sulista*). Assim, de acordo com a carta 4h, o *falar paulista* é caracterizado pela maciça presença de *moleque* no Estado de SP (cor azul), que adentra o norte do PR, o Estado do MS, MG, GO e MT. Uma área de coocorrência dos itens lexicais se forma, principalmente, no território paranaense e catarinense, enquanto a área lexical de *piá* apresenta-se, sobretudo, no território gaúcho e sudoeste paranaense (cor preta).

Quanto a *guri* e *garoto*, observa-se pela carta 4i que há maior número de heteroléxicas, embora ainda se verifique a influência do *falar sulista* (isoléxica de *guri*), sobretudo no Estado do MT. O *falar paulista*, por sua vez, caracterizado pela ocorrência de *garoto* reflete-se, principalmente, em MG e em GO. Panorama análogo acontece com as variantes *guri* e *moleque* (carta 4j). Esta última caracteriza o *falar paulista* formando ampla área lexical, principalmente, em SP, MG e GO, ao passo que *guri* apresenta-se no RS, MS e MT. Somam-se a estas duas áreas, as regiões em que ocorrem indistintamente as duas formas lexicais delimitadas por linhas de heteroléxicas.

Com isso, o *falar paulista* se faz presente nas designações para o referente por meio das variantes *garoto* e *moleque*, ao passo que o *falar sulista*, de influência sul-rio-grandense e paranaense, traz as variantes *guri* e *piá*, expandindo-se com *guri* ao Estado do MS e MT. É necessário, no entanto, salientar que os limites que separam esses dois grandes falares são fluidos, ora apresentando maior número de heteroléxicas (com a sobreposição de isoléxicas), ora menor número.

Quanto às variantes menos produtivas no *corpus*: *pivete/pivetinho*, *rapazinho/rapazote*, formas inadequadas e a *hápax legomena bambino*, verifica-se que foram obtidas como segunda ou terceira resposta ao lado de variantes principais, como *moleque*, *menino*, *guri* e que não ocorreram em área específica no território investigado. O Quadro 6 apresenta o percentual de produtividade dessas variantes que, do ponto de vista quantitativo, foram pouco significativas.

Quadro 6 - Formas pouco produtivas para a questão 132 do QSL

Variantes/Estados	SP	RJ	MG	GO	MT	MS	PR	SC	RS
<i>pivete/pivetinho</i>	3,15%	-	2,97%	3,85%	1,61%	4,00%	0,66%	1,35%	-
<i>rapazinho/rapazote</i>	1,75%	-	-	7,69%	-	4%	-	4,05%	-
<i>formas inadequadas</i>	0,35%	-	1,23%	1,92%	-	-	-	1,35%	-
<i>Bambino</i>	-	-	-	-	-	-	0,66%	-	-

Fonte: [ΣGVCLin] – Relatório 4a (vol.2)

Considerando o uso, por outro lado, a variante pouco produtiva *pivete*, revela aspectos que a qualificam como uma forma lexical desprestigiada, conforme se observa na fala de alguns informantes⁹⁴.

INF.- Moleque [μoελEκI].

INQ.- Que mais?

INF.- Hum, pivete [πιεπεEτΣI].

INQ.- É ofensivo?

INF.- Pivete é.

INQ.- E moleque?

INF.- Moleque também, eu acho que é.

INQ.- Você, por exemplo, se você tivesse um, “ah eu tenho um molequinho lá em casa.” Isso é ofensivo?

INF.- Ah, dependendo da forma que fala não, se você fala assim: Ah eu tenho um molequinho lá em casa, e se você fala: O moleque sai daí, já tô ofendendo, são as formas de falá.

Esse desprestígio atribuído à variante *pivete* explicita-se na fala da maioria dos informantes que a mencionaram, não se justificando, pois, uma representação cartográfica desses itens.

De posse destes resultados, foram realizados testes estatísticos para a análise das principais variantes: *menino*, *moleque*, *guri*, *garoto* e *piá*, com a finalidade de comprovar as inferências feitas na análise descritiva em termos

⁹⁴ Informante 2 do ponto 137 (Mulher, faixa etária I: 18 a 30 anos, de Campina Verde/MG).

absolutos e relativos. Ou seja, elaboraram-se hipóteses acerca das variantes mais produtivas no *corpus* com a finalidade de comprová-las ou não.

4.4.1 Tratamento estatístico

]

A partir da análise dos relatórios e do comportamento das variantes observados nas cartas, foram levantadas cinco hipóteses, a saber:

- Hipótese nula 12: a variante *menino* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 13: a variante *moleque* apresenta distribuição de produtividade heterogênea nos Estados de SP, PR, MG, GO, MS e MT.
- Hipótese nula 14: a variante *guri* apresenta distribuição de produtividade homogênea nos Estados do RS, MS e MT.
- Hipótese nula 15: a variante *garoto* apresenta distribuição de produtividade homogênea nos Estados de SP, PR, MS, MG e GO.
- Hipótese nula 16: a variante *piá* apresenta distribuição de produtividade heterogênea nos Estados do PR, SC e RS.

A aplicação do teste KS revela que as cinco principais variantes (*menino*, *moleque*, *guri*, *garoto* e *piá*) manifestaram comportamento diferenciado, ou seja, nem todas as formas apresentaram aderência à normal. Por exemplo, enquanto no MS todas as formas apresentam adesão à normalidade, em SP, apenas a variante *moleque* apresenta este comportamento (p-valor > 0,05). Desse modo, a partir destes resultados, torna-se necessária a aplicação de testes não-paramétricos para verificar as hipóteses aventadas. Vale ressaltar, contudo, que a variante *piá* não ocorreu em MT, GO e MG.

A hipótese nula 12 consiste em comprovar se a variante mais produtiva no *corpus* apresenta comportamento heterogêneo no território investigado, embora a carta 4a evidencie ampla distribuição diatópica deste item. Com o objetivo de verificar se as variâncias dos registros da variante *menino* são heterogêneas em

todos os Estados analisados, foi aplicado o teste de Kruskal-Wallis. O resultado mostra que o teste alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são heterogêneas, demonstrando que há diferença significativa entre as médias de todos os Estados ($p < 0,001$), e a hipótese nula 12 pode ser aceita.

A carta 4k ratifica os resultados do teste estatístico evidenciando que, de fato, há diferentes índices de produtividade da variante *menino* de acordo com a região do subfalar sulista em que ocorre. De acordo com a referida carta, verifica-se que a variante predomina em grande parte do território, sobretudo, nos Estados do Centro-Oeste e Sudeste com o índice de 75% de ocorrência. Em uma faixa territorial que vai do norte de SP, adentra o Triângulo Mineiro e sul de MG e chega ao sul de GO, a incidência da variante *menino* se acentua atingindo 100% de produtividade. Na região Sul do país, entretanto, a incidência de 100% dessa variante está registrada somente em quatro localidades: pontos 238 – Ijuí/RS, 132 – Criciúma/SC, 233 – Tubarão/SC e 226 – São Miguel do Oeste/SC. Ou seja, na região Sul, em comparação com as outras, observa-se uma diminuição na produtividade do item, predominando a margem de 25 a 75% de incidência. A área de ocorrência de *menino*, portanto, apesar de se apresentar amplamente distribuída pelo território investigado, não se comporta de forma homogênea, ratificando a presença de áreas lexicais no território investigado.

Acredita-se que *moleque*, possivelmente de influência paulista, conforme vem se afirmando ao longo desta análise, apresenta diferença significativa nos Estados de SP, PR, MG, GO, MS e MT, uma vez que por meio da carta linguística 4b, deduz-se que, possivelmente, esta forma lexical tenha se irradiado a partir de SP para os demais Estados. O resultado estatístico do teste empreendido confirma esta hipótese inicial, pois o teste Kruskal-Wallis alcançou nível significativo, podendo-se, portanto, assumir que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se que há diferença significativa entre as médias de todos os Estados mencionados ($p = 0,002$), e a hipótese nula 13 pode ser aceita.

A carta de isoléxica 4l confirma o resultado do teste, evidenciando que há diferenças nos índices de ocorrência da variante *moleque* nos Estados supracitados, com maior concentração visual significativa em SP. Observa-se, desse modo, que o índice de 100% de produtividade esteve em um ponto do PR (208 – Nova Londrina), em uma localidade mineira (ponto 147 – Poços de Caldas) e em

seis localidades paulistas (pontos, 152 – São José do Rio Preto, 157 – Ribeirão Preto, 166 – Marília, 177 – Itapetininga, 178 – Sorocaba e 183 – Itanhaém). Além disso, 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, como, por exemplo, localidades gaúchas e pontos situados ao norte do MS, MT e GO. De acordo com esses dados, a variante *moleque* caracterizaria um possível *falar paulista* dentro do território investigado.

Com o objetivo de verificar se a variante *guri* apresenta comportamento semelhante nos Estados do MT, MS e RS foi aplicado o teste estatístico para verificação da hipótese nula 14. Os resultados mostram que o teste não alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são homogêneas. A partir da análise de variância, nota-se que não há diferença significativa entre as médias dos Estados ($p = 0,764$), e a hipótese nula 14 pode ser aceita.

A carta 4m confirma os resultados encontrados para essa hipótese. De fato, observando-se as áreas de ocorrência deste item no território do subfalar sulista, registram-se diferenças nos índices de SC e PR, onde predominam as áreas com 25 a 50% de ocorrência. Ao passo que, em grande parte dos territórios do RS, MS e MT, predomina a margem de 75 a 100% de ocorrência, ou seja, há um comportamento diferenciado, evidenciando heterogeneidade de ocorrência do item nos Estados pertinentes. No entanto, comparando-se a distribuição pelo RS em relação ao MS e MT, observa-se comportamento semelhante, uma vez que os índices de ocorrência do item estão na mesma margem, de 50 a 100% de incidência, apresentando, portanto, certa homogeneidade na distribuição. Desse modo, pressupõe-se que os resultados do teste estatístico e a carta 4m ratificam as inferências feitas acerca da influência de gaúchos, sobretudo no Estado do MT em decorrência de processos de migração interna em busca das novas fronteiras agrícolas, principalmente, em meados das décadas de 1970 e 80.

Os fluxos migratórios, desse modo, não interferiram somente na economia da região, mas, principalmente, na cultura e na língua, o que se reflete no léxico. Esta variante, portanto, caracteriza a existência de um possível *falar sulista*

(de influência sul-rio-grandense) que se irradia a partir do RS e se expande, principalmente, em direção aos Estados do MS e MT, por um corredor do oeste catarinense e paranaense.

Com o objetivo de verificar se as variâncias dos registros de *garoto* são homogêneas nos Estados de SP, PR, MS, MG e GO foi aplicado o teste de Kruskal-Wallis, cujo resultado não alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são homogêneas. A partir da análise de variância, nota-se que não há diferença significativa entre as médias dos Estados ($p = 0,117$), e a hipótese nula 15 pode ser aceita.

A carta 4n ilustra esta distribuição pela qual é possível observar que em grande parte do território o item se apresenta com 25% de produtividade. Este índice sobe para 50%, sobretudo nas localidades mineiras e em pontos isolados nos demais Estados, enquanto 75% de incidência de *garoto* figuram somente em três localidades (pontos 209 – Terra Boa/PR, 214 – Piraí do Sul/PR e 274 – Bragança Paulista/SP). Nota-se, porém, que apenas em Florianópolis (ponto 230), *garoto* obtém 100% de representatividade.

Desse modo, com base no resultado observado e na distribuição diatópica, provavelmente *garoto* não é uma variante que caracterizaria o possível *falar paulista*. Quatro motivos justificam tal posicionamento: (i) a baixa incidência do item no interior do Estado de SP (em geral 25%), (ii) o alto índice de *garoto* em Florianópolis (100% das ocorrências), (iii) ausência da variante no sul do Estado de GO e, por fim, (iv) a incidência de 50% do item em uma faixa territorial de MG, inclusive em pontos de controle. Seria *garoto* uma forma lexical típica de outros falares, como o mineiro, por exemplo, uma vez que sua distribuição diatópica recobre mais significativamente as localidades de MG? A resposta a essa pergunta não se busca neste trabalho, pois a pesquisa aponta para a necessidade de investigar outros falares estabelecidos por Nascentes (1953) para se fazer observações mais pontuais sobre a ocorrência do item, bem como identificar em que região do país esta variante é mais significativa.

A análise descritiva em termos absolutos e relativos levou à formulação da última hipótese testada para esta questão: a variante *piá* apresenta distribuição de produtividade heterogênea nos Estados do PR, SC e RS. O resultado mostra que o teste alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são heterogêneas. A partir da análise de variância, nota-se

que há diferença significativa entre as médias dos Estados ($p = 0,003$), e a hipótese nula 16 pode ser aceita. A heterogeneidade justifica-se pelos números absolutos e relativos que, de fato, mostram que a variante é mais produtiva no PR, conforme se aponta na primeira parte desta análise. O mesmo panorama se repete na carta 4o, na qual é possível verificar a distribuição do item neste Estado, evidenciando que a porção centro-sul do Estado apresenta uma ampla faixa territorial em que o índice de produtividade se mantém entre 75 a 100%. Este índice diminui no sentido norte do PR e nos Estados de SC e RS. Assim, pressupõe-se que *piá* não é uma variante de influência gaúcha, nem de influência paulista. Seria, então, esta variante um indício da existência de um falar paranaense? Mais itens lexicais merecem ser analisados para se ratificar o que caracterizaria este possível falar paranaense, porém, até o momento, observa-se que a presença de *piá* caracteriza esta região do país e evidencia subáreas lexicais que ora se enquadram no *falar sulista*, ora no *falar paulista*.

Em suma, o Quadro 7 apresenta uma síntese das hipóteses aventadas para a questão 132.

Quadro 7 - Síntese das hipóteses para a questão 132 do QSL

	ACEITA	REJEITADA
HIPÓTESE NULA 12 [#]	X	
HIPÓTESE NULA 13 ^{##}	X	
HIPÓTESE NULA 14 ^{###}	X	
HIPÓTESE NULA 15 ^{####}	X	
HIPÓTESE NULA 16 ^{#####}	X	

menino é heterogênea em todos os Estados

moleque/molequinho é heterogênea nos Estados de MT, MS, GO, MG, SP e PR

guri/gurizinho é homogênea nos Estados de RS, MS, MT

garoto/garotinho é homogênea nos Estados de MS, GO, MG, SP e PR

piá/piazinho é heterogênea nos Estados de PR, SC e RS

A partir da análise empreendida para a questão 132 do QSL, pode-se concluir que:

- (i) A variante padrão – *menino* – encontra-se amplamente distribuída pelo território investigado, embora não atinja percentual acima de 50% nos Estados, exceto no RJ, cujos dados referem-se apenas a uma localidade. Mesmo estando esta variante amplamente distribuída pelo território, conforme se observa na carta linguística 4a, os testes estatísticos e a carta

linguística 4k evidenciam que seu comportamento é heterogêneo, confirmando, desse modo, a hipótese nula 12, em decorrência do polimorfismo lexical para a denominação do referente em pauta.

- (ii) *Moleque* – segunda variante de maior ocorrência no *corpus* – apresenta-se de forma mais produtiva nos Estados de SP, MG e GO, evidenciando que, possivelmente, este item teria se irradiado a partir de SP para outras regiões. Entretanto, o teste estatístico empreendido para verificação da hipótese nula 13 nos Estados em que o item ocorre (SP, PR, MG, GO, MS, MT) identifica diferença significativa deste item nos Estados supramencionados. A carta linguística 4l confirma a premissa de que *moleque* caracteriza o denominado *falar paulista*. Além do mais, por meio desta carta, identifica-se que o item predomina com 75% de produtividade na maioria dos pontos linguísticos em que ocorre, não figurando em área de influência do *falar sulista*, principalmente no Estado do RS e MT. Não obstante, de acordo com estudos já realizados acerca deste item, sabe-se que esta variante revela a influência da língua africana no português e, em algumas situações, pode receber caráter disfêmico e insultuoso.
- (iii) A variante *guri* comporta-se como um item que evidencia aspectos do movimento de migração interna de gaúchos em direção ao Centro-Oeste, principalmente, no MS e MT, uma vez que, nestes dois Estados, a variante se apresenta de forma mais produtiva. Ao considerar os Estados do RS, MS e MT, o resultado do teste revela o comportamento homogêneo da variante, ratificando a semelhança linguística entre os três Estados, o que confirma a hipótese nula 14. Portanto, *guri* caracteriza o *falar sulista*, que demonstra influências do contato do tupi com a língua portuguesa e se estende além dos limites políticos do RS.
- (iv) O item *garoto* – quarta variante mais produtiva no conjunto de respostas – apresenta distribuição de produtividade homogênea

no território em que ocorre, confirmando a hipótese nula 15. Essa variante é mais representativa em MG, evidenciando que este item lexical não é uma variante característica do *falar paulista* e representa, possivelmente, uma forma lexical que pertence também a outros falares, como o mineiro, por exemplo.

- (v) A variante *piá* caracteriza-se como uma forma lexical típica da região Sul do país, em particular, associada ao Estado do PR, segundo relatos dos informantes. O item, no entanto, ocorre também em outros Estados tais como em SP, MS e MT, conforme se verifica na carta linguística 4e. O teste estatístico evidencia que *piá* apresenta distribuição heterogênea no PR, SC e RS, confirmando, desse modo, a hipótese nula 16. Observando-se a carta 4o, nota-se maior concentração do item no centro-sul paranaense o que, possivelmente, dá indícios da existência de subáreas lexicais ou mesmo um possível falar paranaense. *Piá*, assim como *guri*, apresenta etimologia que evidencia traços do tupi com o português e ocorre em cidades paulistas que foram pontos de passagem nos antigos caminhos das Tropas, como Itararé, Itapetininga e Sorocaba. Desse modo, a variante aponta também para aspectos sócio-históricos na constituição do léxico daquela região de SP, Vale do Ribeira, bem como caracteriza uma subárea lexical na região do *falar sulista*. Novamente, salienta-se que são necessárias mais pesquisas para identificação mais precisa e validação se esta subárea não representaria um possível *falar paranaense*.
- (vi) As variantes menos produtivas no *corpus* (*pivete*, *rapazinho*, *formas inadequadas* e *bambino*) não ocorrem em uma área específica do território investigado, figurando, em geral, como segunda ou terceira resposta ao lado de variantes principais como *moleque*, *menino*, *guri* e *piá*.

Em síntese, os dois possíveis grandes falares aqui defendidos, para os designativos da questão 132, estariam, assim, representados: (i) na porção setentrional, o *falar paulista*, caracterizado pelo uso da variante padrão *menino* que

diminui sua incidência à medida que ‘caminha’ para a Região Sul, além do uso da variante *moleque* – muito significativa no Estado de SP; (ii) na porção meridional do território investigado, encontra-se o *falar sulista* de influência sul-rio-grandense em direção aos Estados do MS e MT por meio de um corredor lateral do oeste catarinense e paranaense, neste caso, com a difusão da variante *guri* e também pela variante *piá*, esta última dando indícios da existência de subáreas lexicais localizadas no território paranaense. Tanto *piá* quanto *guri* figuram no Estado de SP em localidades que foram pontos de passagem dos tropeiros vindos do Sul do país. Desse modo, evidenciam aspectos sócio-históricos que marcam o léxico dessa região paulista (Vale do Ribeira).

4.5 QUESTÃO 039 – TANGERINA

A questão 39 do QSL, que objetiva documentar os designativos para as frutas menores que a laranja que se descascam com a mão, no *corpus* selecionado, apresenta um total de 1106 registros distribuídos em 29 formas, incluindo-se variantes fonéticas e morfofonêmicas. Considerando as formas que apresentam processos fonéticos e morfofonêmicos diversos, foram feitos os seguintes agrupamentos:

- (i) Alternância de fonema inicial de /b/ > /v/ (bergamota/vergamota);
- (ii) Alternância da vogal átona final /a/ > /e/ (bergamota/bergamote, moricota/moricote, mangota/mangote, morocota/morocote, muricota/muricote, morgota/morgote);
- (iii) Sonorização da consoante velar de /k/ > /g/: (morcote/ morgote;, morcota/morgota, marcota/margota);
- (iv) Casos de assimilação, seja pela influência da vogal tônica seguinte {mar-} > {mor-}, {ma-} > {mo-} (maricota/moricota); (margota/morgota/morgote, marcota/morcote); pela consoante nasal bilabial sobre a vogal {ma-} > {mu-} (maricota/muricota); seja pela influência da consoante nasal bilabial sobre a consoante da sílaba seguinte (morgota/mormota);
- (v) Inserção de vogal epentética [i] ou [o] (marcota/Maricota, morcote/moricote, morcota/moricota);

- (morikota/morocota/morocote); (vi) Síncopa de sílaba pretônica (muricote/mucote);
- (vii) Adequação ao gênero do primeiro nome: {a} < {o} (laranja-crava/laranja-cravo).

Assim, foram consideradas nove formas lexicais distintas, registradas com diferentes índices de produtividade, conforme o relatório 5, figurando, entre as formas mais produtivas, as variantes: *mexerica* (333 ocorrências, 30,11%), *poncã* (314, 2,39%), *tangerina* (244, 22,06%) e *bergamota* com suas variantes morfofonêmicas (112 registros, 10,13%). Com menor produtividade, encontram-se *morcote* e variantes morfofonêmicas (54 ocorrências, 4,88%), *mimosa* (33, 2,98), *laranja-cravo(a)* (13 registros, 1,18%), *mandarina* (duas ocorrências, 0,18) e a hápax legomena *carioquinha* (0,09%).

Além deste conjunto de realizações lexicais, observa-se a ocorrência de adjetivos que especificam a variedade dessa fruta ou atribuem características à *mexerica* ou à *bergamota*, uma vez que esta questão do QSL prevê a apuração de detalhes sobre o referente. Este detalhamento se justifica pelo fato de a fruta ser amplamente conhecida em todo o território nacional e apresentar diversidade de tipos decorrentes ou da hibridização natural ou de experimentos realizados pelo homem, como bem assevera Costa (2011).

Assim, considerando os adjetivos que “denunciam quem comeu o fruto”, foram registradas 17 formas de *mexerica enredeira*, 15 de *mexerica fuxiqueira*, uma de *mexerica fofoqueira*. Ainda considerando o cheiro característico que delata quem manuseou a fruta, ocorrem três registros de *mexerica fedidinha*, uma ocorrência de *mexerica cheirosinha* e sete registros de *mexerica bode/bodinha*, motivação lexical justificada, possivelmente, por analogia ao odor do animal (bode)⁹⁵.

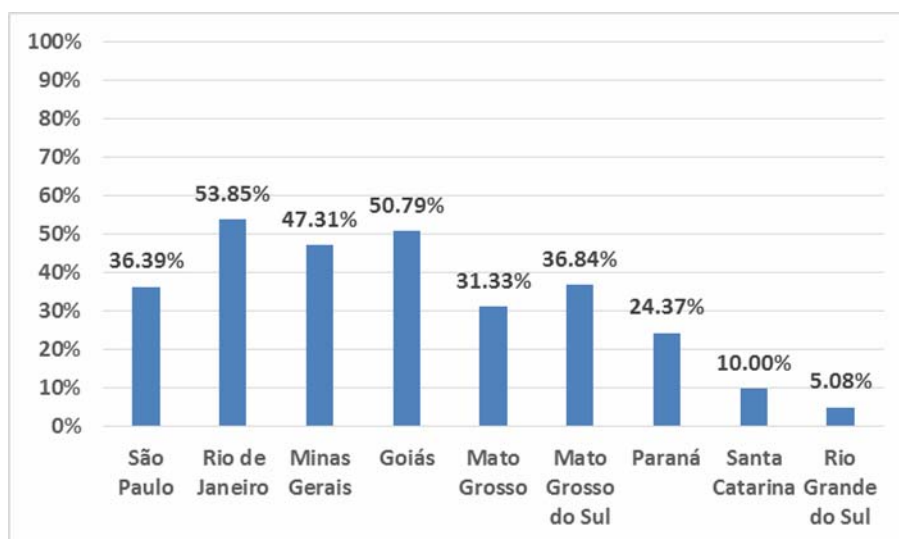
Quanto às variedades de *mexerica* ou *bergamota*, encontram-se quatro ocorrências de *vergamota do céu*, sete registros de *mexerica caipira*, dez ocorrências de *mexerica cravo*, duas, de *mexerica comum* e um registro de *mexerica*

⁹⁵ Para um detalhamento sobre o uso dos adjetivos *enredeira* e *fuxiqueira* em localidades da Região Centro-Oeste, recomenda-se a leitura de Costa (2011), uma vez que a autora tratou, separadamente, estes itens, considerando a produtividade de cada um em algumas localidades como Paranaíba-MS, Aruanã-GO e Jataí-GO. Dados os objetivos desta tese, consideram-se esses designativos como uma qualidade atribuída à variante *mexerica* e não uma variante específica, conforme relato dos informantes apresentados adiante.

rosa. Portanto, essas 68 ocorrências que especificam ou qualificam o tipo de *mexerica* ou *bergamota* não foram consideradas dentro dos 1106 registros, uma vez que os informantes que forneceram esses detalhes já haviam respondido, como principais variantes ou *mexerica* ou *bergamota*. Não obstante, consta também deste quadro uma abstenção de resposta (ponto 113/1 – Corumbá/MS) e dois problemas técnicos por dificuldades de acesso ao áudio: ponto 242/4 (Santa Maria - RS) e 249 (São José do Norte - RS).

A carta linguística 5 apresenta a distribuição diatópica das nove formas lexicais consideradas como designativos para o referente. Nesta carta, de um modo geral, é possível visualizar uma ampla distribuição das três variantes mais produtivas (*mexerica*, *poncã* e *tangerina*) nos Estados de SP, sul de MG, sul de GO, no MT e MS. A variante, *poncã*, por sua vez, apresenta-se também amplamente difundida pelo Estado do PR e também por SP, MS e MT. *Bergamota* distribui-se pelo Estado do RS e SC, além de ocorrer no oeste paranaense. *Morcote* apresenta-se, sobretudo, no Estado de SP, enquanto a variante *mimosa* restringe-se às localidades paranaenses e ocorre em um município do norte catarinense (ponto 224-Porto União).

A carta 5a, por meio da arealidade da variante *mexerica*, evidencia os limites da isoléxica na porção meridional do subfalar sulista, de Nascentes (1953). Observa-se que no litoral central de SC, no interior deste Estado, no litoral paranaense (ponto 221 – Morretes) e em grande parte do Estado do RS, não há ocorrência da variante *mexerica*, que, por sua vez, se distribui, amplamente, por todo o território do Estado de SP, MS, PR, MT, sul de GO, sul de MG, ocorrendo, inclusive, nas adjacências do subfalar sulista. A Figura 81 apresenta esta distribuição em termos quantitativos.

Figura 81 - Produtividade da variante *mexerica* por Estado

Fonte: [ΣGVCLin] - Relatório 5a (vol. 2)

Observa-se que, nos dados das duas localidades de controle do RJ e no Estado de GO, *mexerica* apresenta-se com percentual acima de 50%, seguido pelos dados de MG (47,31%), MS (36,84%), SP (36,39%) e MT (31,33). Nos Estados da região Sul, o índice é ainda menor (24,37%, no PR), 20% em SC e apenas 5,08% no RS.

Como se verifica, o item está amplamente difundido pelo território investigado e está registrado nos dois principais dicionários da língua portuguesa na acepção buscada para o referente da questão 39 do QSL. Para Ferreira (2004), esta variante lexical provém do substantivo *mexerico* que, por sua vez, significa o “ato de mexericar; enredo; intriga” e, segundo Houaiss e Villar (2001), *mexerica* é uma forma regressiva do verbo *mexericar*. Desse modo, as obras lexicográficas e o uso dos informantes atribuem ao nome da fruta características que lhe são peculiares, ou seja, o fato de o odor forte denunciar quem o comeu justifica, por extensão de sentido, o uso do designativo *mexerica*, assim como adjetivos que intensificam esta característica, conforme se observa em relatos dos informantes.

INF.- Mixirica, tem uma mixiriquinha piquena que tem um chero forte na casca, então a gente chama ela de irredera porque se você robá do vizinho e chegá perto do vizinho, o vizinho sabe que você pegô se cê não tivé lá na tua casa, normalmente⁹⁶.

⁹⁶ Informante 3 do ponto 158 (Homem, faixa etária II: 50 a 65 anos, de Barretos/SP).

INF.- Irridera (risos), irredera é uma pessoa fofo... é como se fosse fofoquera, porque a pessoa chupa não tem, num pode escondê de ninguém, porque o chero dela é muito forte (...)⁹⁷

INF.- Muita gente fala mixirica enredera, porque quem chupá ela todo mundo sabe, né, porque o chero é muito forte.⁹⁸

INF.- Não, tem a mixirica, a pocã que é grande e tem uma pequenininha que chama mixirica fuxiquera aqui.

INQ.- Ah, por quê?

INF.- É porque na hora que você descasca ela chera longe, é por isso que eles geralmente chama ela de fuxiquera (...)⁹⁹

Considerando a produtividade do item *mexerica*, bem como a sua distribuição diatópica na área investigada, possivelmente, esta variante lexical seria uma das características do denominado falar paulista, que se difunde do Estado de SP para o sul de MG, para os Estados da região Centro-Oeste, interior do PR, e por um corredor central atinge campos de Lages em SC, parte do território catarinense já identificada como uma área de influência paulista, por Margotti; Vieira (2006), e denominada, por Romano e Aguilera (2014, no prelo), como *área do interior central de Santa Catarina*.

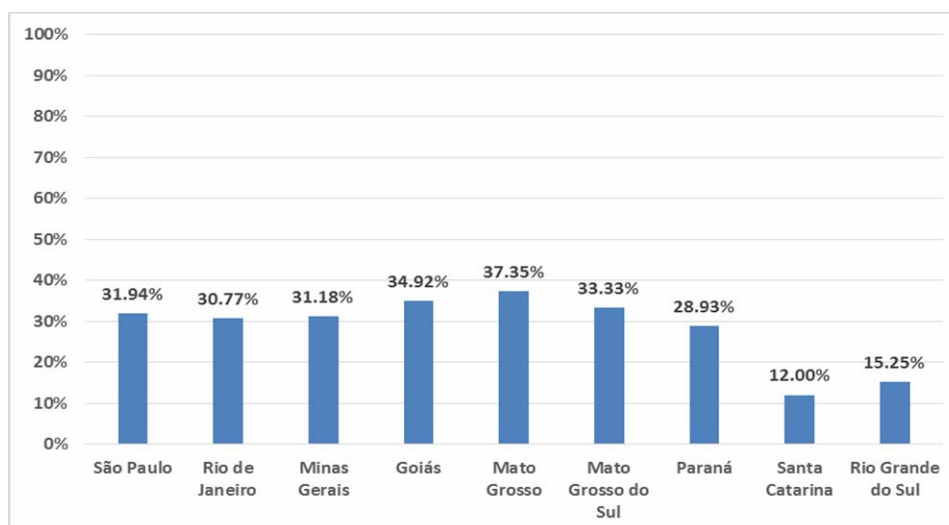
Observando-se as áreas de ocorrência de *poncã*, segunda variante mais produtiva (Carta 5b), nota-se que este item lexical está amplamente distribuído pela área investigada, não ocorrendo, principalmente, no extremo sul do território investigado, em uma área que corresponde ao nordeste do RS (ponto 237 – Vacaria), litoral sul de SC (pontos 233, 232- Criciúma e Tubarão) e em uma localidade do nordeste catarinense (Blumenau - ponto 227), além de não ocorrer também em uma localidade do sul de GO (ponto 126 – Quirinópolis).

Em termos quantitativos, a Figura 82 apresenta a distribuição percentual deste item nos Estados investigados, evidenciando valores próximos entre eles. Nota-se, entretanto, que, no RS e em SC, o índice de ocorrência de *poncã* diminui para 15,25% e 12%, respectivamente.

⁹⁷ Informante 4 do ponto 140 (Mulher, faixa etária II: 50 a 65 anos, de Passos/MG).

⁹⁸ Informante 4 do ponto 114 (Mulher, faixa etária II: 50 a 65 anos, de Paranaíba/MS).

⁹⁹ Informante 1 do ponto 120 (homem, faixa etária I: 18 a 30 anos, de Aruanã/GO).

Figura 82 - Produtividade da variante *poncã* por Estado

Fonte: [ΣGVCLin] - Relatório 5a (vol. 2)

Essa diminuição da incidência de *poncã* e também de *mexerica* (Figura 81), em SC e RS, pode ser explicada pela presença de uma forma considerada típica desses Estados, *bergamota*. Essas informações diatópicas podem ser observadas, inclusive, na fala de informantes goianos e paranaenses.

INF – É... porque... é complicado o negócio, é tangerina, poncã, mixirica, no sul ela chama, é, bergamota, vai... Tanto nome, né? ¹⁰⁰

INF.- A vergamota como os gaúcho dizia, no tempo que eles passavam por aqui, né, e tinha o hotel Guarani que até eu trabalhei cum dezesseis ano. Então, eles dizia vergamota, “nós compramo vergamota”. Daí no começo eu não sabia o que era. “Tem vergamota?” Tinha a laranja baiana né, tinha a mixirica e tangirina, nós dizia mixirica, nem tangirina num dizia, né. Daí que fomo aprendendo a tangirina e a vergamota. Vergamota era a mesma tangirina.

INQ.- E são iguais?

INF.- São tudo igual. O gaúcho dizia vergamota. ¹⁰¹

INF.- Bom a poncã é uma fruta que... mais fácil de descascá né, fica assim desgrudado da casca, e aqui a gente chama de poncã, mas só que tem outros lugar que chama bergamota.

INQ.- Aqui não se usa a palavra bergamota?

INF.- Aqui não.

INQ.- Não. Mas a senhora ouviu onde?

INF.- Eu ouvi, eu ouvi uma mulher que veio de Santa Catarina, fala que lá chama bergamota.

INQ.- Isso. ¹⁰²

¹⁰⁰ Informante 3 do ponto 120 (Homem, faixa etária II: 50 a 65 anos, de Formosa/GO).

¹⁰¹ Informante 4 do ponto 216 (Mulher, faixa etária II: 50 a 65 anos, de Adrianópolis/PR).

Conforme se observa, há uma associação da variante *bergamota* aos falantes catarinenses e gaúchos. Para Houaiss e Villar (2001), *poncã* é um substantivo masculino, na acepção de uma “rubrica da agricultura para a variedade de tangerina, grande e de casca frouxa, originária do Japão” (HOUAISS; VILLAR, 2001). Afirmam os lexicógrafos sobre a etimologia desta palavra que: “segundo Michaelis, vem do japonês ponkan”. Em Ferreira (2004), constam as mesmas informações acerca desta variante, porém, além desta acepção, o dicionarista traz algumas informações adicionais, por exemplo: o lugar onde geralmente se cultivava esse fruto, dizendo que esse “tipo de tangerina é hoje cultivada no Brasil, sobretudo em São Paulo, por japoneses, e que se caracteriza pelas dimensões avantajadas e casca muito frouxa” (FERREIRA, 2004). Quanto às características - avantajada, de casca frouxa - são as mesmas especificações dadas pelos informantes ao descrever o fruto.

INF.- A pocã tem casca grossa, e é bem graúda, bem grossa graúda, bem cascuda com a casca grossa e tem a tangerina que é a casquinha mais fininha, mais grudadinha, e tem a morgota também que tem a casca mais fininha ainda, e ela também a graúda, a morgota, e a bergamota é a tradicional.¹⁰³

INQ.- A poncã é aquela grande?

INF.- Aquela grande.

INQ.- Que a casca solta?

INF.- Solta, isso.¹⁰⁴

INF.- A diferença daqui da terra a senhora qué dizê né? A da terra, a polcã é que é uma casca grossa né, aquela mais fácil de descascá, aquela chamada pocã aí tem a mexerica que é um pouquinho mais difíci, aí tem a maricota que é um mais difíci inda. A maricota quela de casca lisa, bem lisinha memo, bem fininha também assim, na hora de descascá é té perigoso machucá os gomo.¹⁰⁵

Poncã, assim como a variante *mexerica*, é uma forma lexical que, possivelmente, caracteriza o *falar paulista*, que avança a partir de SP e se estende a toda a região Centro-Oeste, ao Estado do PR, parte do território catarinense, chegando até o RS.

A terceira variante mais produtiva, no conjunto das respostas, é *tangerina* e, conforme a distribuição do item nos Estados (Figura 83), observa-se

¹⁰² Informante 4 do ponto 114 (Mulher, faixa etária II: 50 a 65 anos, de Nova Londrina/PR).

¹⁰³ Informante 2 do ponto 240 (mulher, faixa etária I: 18 a 30 anos, de Flores da Cunha/RS).

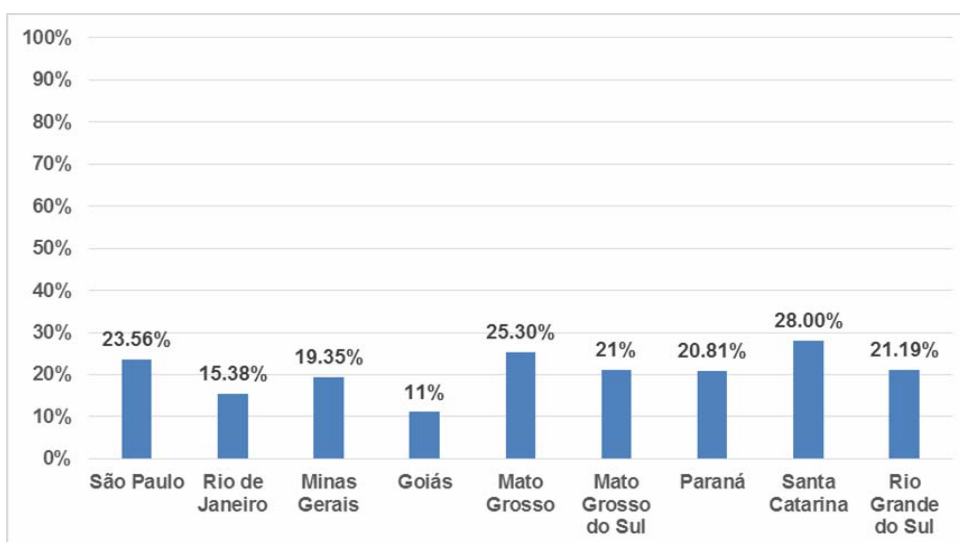
¹⁰⁴ Informante 4 do ponto 176 (mulher, faixa etária II: 50 a 65 anos, de Guaratinguetá/SP).

¹⁰⁵ Informante 1 do ponto 115 (homem, faixa etária I: 18 a 30 anos, de Campo Grande/MS).

que esta variante mantém seu percentual em até 28% das respostas de cada Estado. Dentre as unidades federativas, as que apresentam o menor índice de *tangerina* são GO (11%) e RJ (15,38%), sendo que os dados do RJ referem-se apenas a duas localidades de controle.

A distribuição diatópica de *tangerina* abrange todo o território investigado (Carta 5c), não ocorrendo em dois pontos do oeste e dois do nordeste do RS (ponto 234 – Três Passos e 239 – São Borja e ponto 243 – Porto Alegre e 244 – Osório, respectivamente). Em SC, o item só não foi registrado em uma localidade do litoral sul (ponto 232 – Tubarão) e no PR, não se apresentou em Lapa (ponto 222). A variante também não ocorreu em Campinas/SP (ponto 173), em Parnaíba-MS (ponto 114) e em duas localidades do interior de GO (ponto 126 – Quirinópolis, e 122 – Goiás). Nos pontos de controle, *tangerina* não foi elicitada em Porangatu-GO (ponto 118), Unaí-MG (130) e Paraty-RJ (206). Ou seja, dos 118 pontos investigados neste trabalho, em apenas 13 a variante não ocorreu, o que ratifica o quanto a forma considerada padrão está difundida pelo território investigado.

Figura 83- Produtividade da variante *tangerina* por Estado



Fonte: [ΣGVCLin] - Relatório 5a (vol. 2)

O estudo de Romano e Aguilera (2008), que se pautou em obras lexicográficas visando a discutir a dicionarização dos designativos para a questão

em pauta, revelou que, dentre os dicionários investigados naquela pesquisa¹⁰⁶, Houaiss (2001) parece ser o mais completo na definição do verbete *tangerina*, pois, além da acepção - fruto da tangerineira - registra as variantes regionais - *bergamota*, *laranja-cravo*, *laranja-mimosa*, *mandarina*, *mexerica*, *mimosa*, *tangerina-cravo*, *tangerina-do-rio*, *vergamota* - sem, contudo, indicar em que região estas formas são mais frequentes.

De fato, *tangerina* não se refere a uma área geográfica específica, trata-se de uma forma comum a todo território e, como se observa em outro estudo de Romano e Aguilera (2009), *tangerina* predomina em capitais das regiões Norte e Nordeste do Brasil, o que também pode ser observado na carta L05 do Atlas Linguístico do Brasil (CARDOSO et al. 2014b). No que se refere ao território investigado, verifica-se a presença de *tangerina* na maioria dos pontos, ou seja, em 105 das 118 localidades selecionadas, embora com índices que não ultrapassam 28% das ocorrências em cada Estado.

A quarta variante mais produtiva que se apresenta, aproximadamente, em 10% do *corpus*, conforme o relatório 5, é *bergamota*. Ferreira (2004) anota que *bergamota* vem do turco *beg armudi*, 'pêra do príncipe', pelo italiano *bergamota*, indicando como brasileirismo de SC e RS com remissiva à *tangerina*. Romano e Aguilera (2008), com base nas obras lexicográficas consultadas afirmam que:

Para Nascentes (1966), teria vindo através do italiano *bergamotta* ou do francês *bergamotte*. Segundo Bueno (1964), não se compreende como *beg armudi* passou a *bergamotta* em italiano e como o nome de uma pêra possa designar uma mexerica. Para Fontinha (s.d), *bergamota* veio mais especificamente de Pêrgamo, na Itália e é uma espécie de cidra aromática, com que se fazem cosméticos; planta labiada, de perfume muito agradável; casta de pêra muito sumarenta (ROMANO; AGUILERA, 2008, p.7).

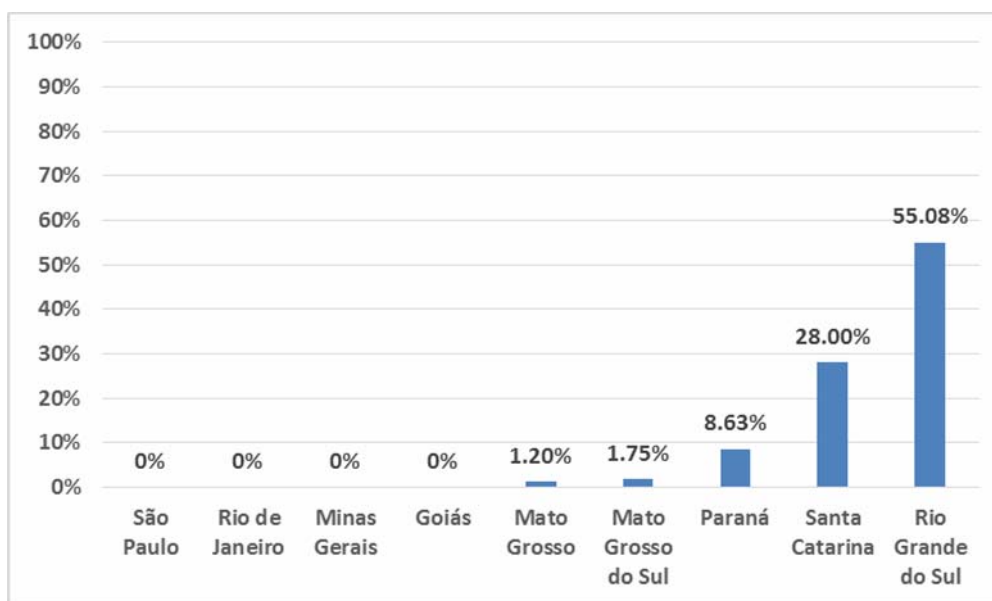
Houaiss e Villar (2001), dentre outras acepções, indicam que é pequena árvore (*Citrus aurantium* subespécie *bergamia*) da família das rutáceas, de flores muito aromáticas e fruto piriforme com casca fina, lisa e amarela;

¹⁰⁶ Os autores consultaram seis dicionários de Língua Portuguesa: Ferreira (2004), Houaiss (e Villar (2001), Mirador Internacional (1975), Silveira Bueno (1969), Aulete (1964), Morais Silva (1945) e quatro dicionários particularmente designados como etimológicos: Cunha (1986), Nascentes (1966), Bueno (1964) e Fontinha (s.d).

bergamoteira, e ainda como fruto dessa árvore com remissiva para *tangerina* e como um regionalismo do RS e SC.

Em termos quantitativos, observa-se que *bergamota* obtém representatividade considerável nas localidades gaúchas, ultrapassando metade das respostas válidas (55,08%), seguido-se SC (28%) e PR (8,63%). Com menor produtividade, registram-se como ocorrência única no MT (1,20%) e MS (1,75%), conforme se verifica na Figura 84.

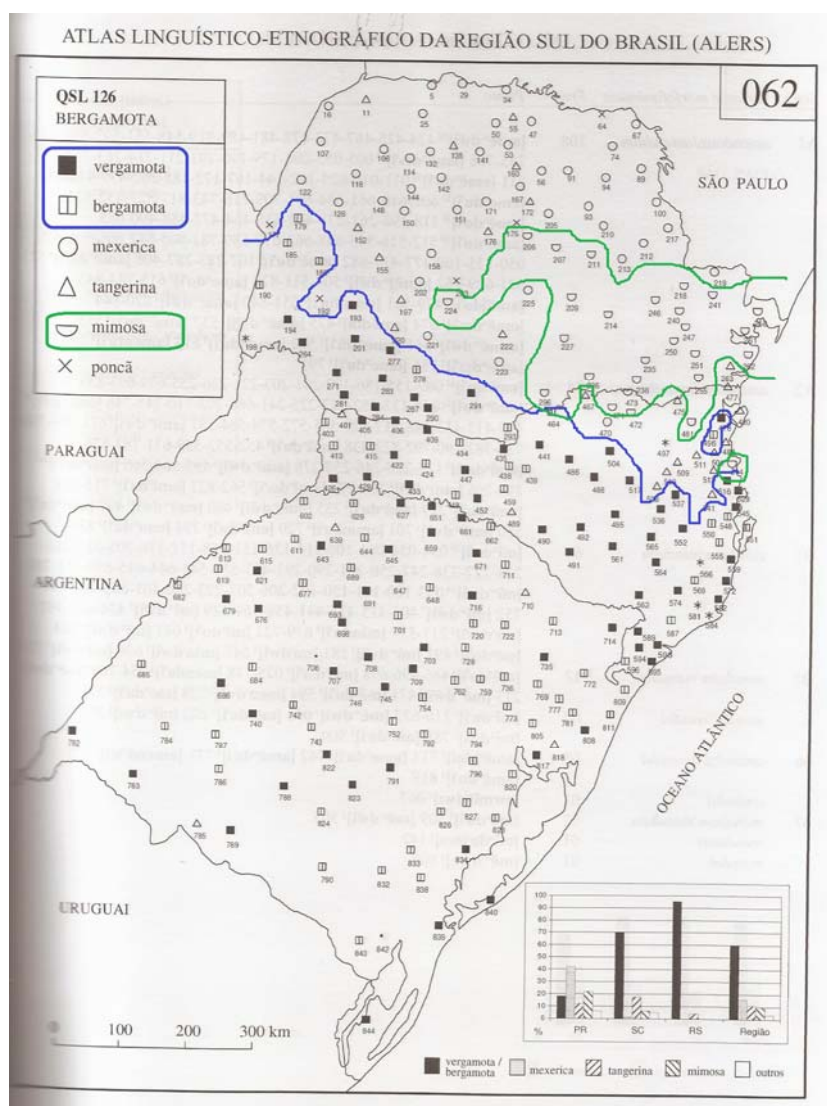
Quanto à distribuição diatópica, *bergamota*, possivelmente, caracterizaria o que se vem denominando neste trabalho como *falar sulista*, de influência sul-riograndense e das línguas em contato com o português, no caso, o italiano, uma vez que na carta 5d pode-se observar a distribuição diatópica desta variante, definindo áreas de ocorrências bem delimitadas. Assim, a variante se difunde por todo o Estado do RS, de SC, atingindo também cidades localizadas no sul do PR (222-Lapa) até o ponto 214 (Piraí do Sul/PR), além de Morretes (221), no litoral. A ocorrência de *bergamota* na região de Lapa, que se difunde até Piraí do Sul, possivelmente, seria uma marca deixada pelos gaúchos no léxico da região, uma vez que se sabe, conforme o capítulo 1, que a cidade de Lapa foi fundada a partir dos antigos ranchos dos tropeiros vindos do sul do País em direção ao Estado de SP. Portanto, este município localiza-se em uma das rotas do tropeirismo (Caminho de Viamão). Por outro lado, essa área se estende também por um corredor do oeste paranaense, área que Altenhofen (2005) denominou como *Corredor oeste de projeção rio-grandense*, a partir do qual *bergamota* atinge o ponto 117 (Ponta Porã), no sudoeste do MS e ocorre isoladamente em Diamantino/MT (ponto 105).

Figura 84 - Produtividade da variante *bergamota* por Estado

Fonte: [ΣGVCLin] - Relatório 5a (vol. 2)

Na carta linguística 31 do Atlas Linguístico do Município de Ponta Porã (REIS, 2006), há o registro de duas ocorrências de *vergamota*, além de variantes como *mandarina*, de origem castelhana, e outras variantes do português e do guarani. O ALERS (ALTENHOFEN; KLASSMAN, 2011), na carta 62, documenta as cinco principais variantes para o referente (*bergamota*, *mexerica*, *tangerina*, *mimosa* e *poncã*), separando, na legenda, as formas *bergamota* e *vergamota*. Na Figura 85, apresenta-se a referida carta, incluindo esses dois itens como variantes do mesmo item lexical.

Figura 85 - Carta linguística 062 do ALERS



Fonte: ALERS (ALTENHOFEN; KLASSMANN, 2011) – com adaptações

Adaptando a carta linguística do ALERS com o traçado de linhas de isoléxicas para as variantes *bergamota/vergamota* e *mimosa*, observa-se que essas duas linhas revelam três áreas lexicais na Região Sul do país. A primeira delas na parte meridional (linha azul), com ocorrência de *bergamota/vergamota* no RS, SC (com exceção do nordeste deste Estado), atingindo também o sudoeste e oeste do PR. A segunda área revelada pela isoléxica em verde localiza-se, principalmente, na porção leste do PR, com ocorrência da variante *mimosa*. Por fim, a terceira área revelada pela delimitação dessas duas linhas de isoléxicas apresenta a distribuição de *mexerica*, localizada na porção central e norte do PR. Não obstante, há ocorrência esparsa de *tangerina* e *poncã* em algumas localidades deste atlas regional.

A área de ocorrência de *bergamota* nos dados do ALiB em comparação com a área registrada nos dados do ALERS revela certa semelhança quanto à distribuição, principalmente no que se refere ao corredor do sudoeste paranaense, confirmando a hipótese de que o falar sulista de influência sul-riograndense e do português em contato com outras línguas se expande além do território gaúcho, conforme se observa, por exemplo, na distribuição de outro itens como *sanga/arroio* (Carta 1h), *chimia* (Carta 2b), *bolita* (Carta 3b), *piá/guri* (Carta 4f). *Bergamota*, desse modo, evidencia características desse falar sulista de influência sul-riograndense e de línguas de colonização.

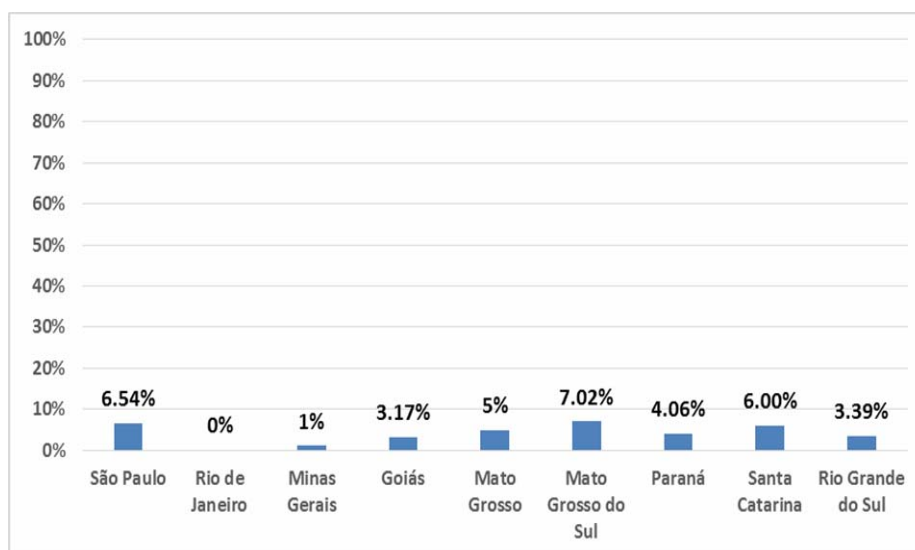
Quanto a quinta variante lexical registrada no *corpus*, *morcote*, sabe-se que diversos grupos híbridos fazem parte do grupo das *tangerinas*, “entre eles, o tangor Murcote, o mais conhecido, é o resultado do cruzamento de laranja doce com tangerina” (PIO, 2000). De acordo com informações do site CEASA-Campinas, a “variedade murcott (murcote) denominado de Tangor é resultante do cruzamento da tangerina com a laranja doce (*Citrus reticulata* Blanco x *Citrus sinensis* L. Osbeck)”. Esta variedade de tangerina foi descoberta por Walter T. Swingle, em 1913, na Flórida, EUA, e propagada pelos agricultores, também da Flórida, Charles Murcott Smith e J. Ward Smith (GONÇALVES, 2014). No português, obviamente, o designativo, motivado pelo sobrenome de um dos pioneiros no cultivo da fruta, sofreu alterações fonéticas.

No *corpus* investigado, dentre as variantes lexicais documentadas, a que apresenta maior número de variantes fonéticas e morfofonêmicas (18 formas) é *morcote*, caracterizada pela maioria dos informantes como uma variedade do fruto que tem como característica principal a casca mais grudada à polpa, tornando-se difícil descascá-la.

INF.- Tem, mixirica tem vários tipo né, tem umas duras, a gente chama mangote, essas é... casca dura, pra ‘rancá ela tem que sê co’ a mão mesmo, mais é bem preguenta, num sai fácil¹⁰⁷.

Quanto à produtividade de *morcote*, observa-se pela Figura 86 que o item não foi registrado no RJ, pelo menos nas duas localidades de controle, esse apresenta com baixo percentual nos demais Estados.

¹⁰⁷ Informante 3 do ponto 123 (homem, faixa etária II: 50 a 65 anos, de Goiânia/GO).

Figura 86 - Produtividade da variante *morcote* por Estado

Fonte: [ΣGVCLin] - Relatório 5a (vol. 2)

Quanto à distribuição diatópica, a carta linguística 5e mostra a ocorrência do item principalmente em uma faixa territorial que vai do noroeste paranaense, passando pelo interior do MS até o ponto 105 (Diamantino-MT). Há também a distribuição do item pelo interior de SP e também na região do Vale do Ribeira, de onde avança para o centro do PR em uma faixa contínua.

Uma terceira área de ocorrência de *morcote* localiza-se no sudoeste do PR – ponto 223 (Barracão), atinge uma localidade do oeste de SC (229 – Concórdia) e três localidades do noroeste gaúcho (234 – Três Passos, 235 - Erechim e 236 – Passo Fundo). Há também ocorrências da variante em pontos isolados no litoral norte de SC (225 – São Francisco do Sul), litoral norte de SP (180 – Caraguatatuba), uma localidade no sul de MG (140 – Passos) e em Goiânia-GO (ponto 123). A baixa incidência de *morcote*, na maioria das vezes como ocorrência única na localidade, bem como sua distribuição na área do território investigado não permite enquadrá-la nem em um nem em outro subfalar aqui apresentado, uma vez que se mostrou comum a diferentes Estados, assim como ocorre *tangerina*.

A variante *mimosa*, no cômputo geral das ocorrências, apresenta 33 registros representando 2,98% das respostas no *corpus*. No relatório 5a, verifica-se que este item lexical ocorre apenas em três Estados, PR (12,18%), SC (8%) e SP como ocorrência única, o que representa apenas 0,26%. Embora se apresente como variante pouco produtiva, do ponto de vista diatópico *mimosa* comporta-se como

uma variante regional, não figurando em outras áreas do território investigado além das mencionadas, conforme se verifica na carta linguística 5f.

Em geral, a variante encontra-se registrada no PR e é reconhecida por alguns informantes como uma forma típica para denominar o referente na cidade de Curitiba:

INQ.- A senhora ouviu falar que a mixirica tenha algum outro nome em algum outro lugar? Curitiba...

INF.- Parece que em Curitiba eles chamam de mimosa, agora eu num... num lembro porque faiz muito tempo que eu num vô pra lá, né, então... porque lá... pra lá o palavriado é diferente, né, do nosso daqui do norte.

INQ.- É é diferente sim.

INF.- Do sul é bem mais diferente.¹⁰⁸

Mimosa está presente no litoral paranaense, de onde atinge o litoral norte de SC, avança as cidades paranaenses até o ponto 212 (Campo Mourão), onde foi registrada na fala da informante de nº 2 (Mulher, faixa etária I). Ocorre em uma cidade do norte pioneiro (211 – Tomazina), de onde atinge, inclusive, uma cidade paulista no Vale do Ribeira (ponto 185 – Ribeira). Em direção sul, a variante avança à porção central do PR, passando pelo ponto 224 (Porto União, norte de SC), avança até o ponto 231 (Lages/SC) e chega até uma localidade no litoral sul catarinense (232 – Tubarão). Em comparação com a carta do ALERS (Figura 85), *mimosa* se apresenta na mesma área linguística registrada no atlas regional, dando indícios de um território que apresenta especificidades lexicais que o distinguem do falar de influência paulista e o de influência sul-riograndense.

Ferreira (2004) traz duas entradas para *mimosa*, sendo a segunda a que melhor se encaixa ao referente, pois o lexicógrafo remete este item a *tangerina* e ainda diz que é “forma reduzida de laranja-mimosa”. Esta, por sua vez pesquisada, ratifica esta designação, aparecendo como um “substantivo feminino, brasileirismo do sul de SP, PR e norte de Santa Catarina, remetendo para tangerina” (FERREIRA, 2004), portanto, sinônimo do termo em questão. Houaiss e Villar (2001), dentre outras acepções, também trazem *laranja-mimosa* como um regionalismo do Sul do Brasil sinônimo de *tangerina*.

O verbete *mimo*, para Cunha (2010), refere-se à “coisa delicada que se oferece ou dá”, do qual derivam por criação expressiva termos como ‘amimar, mimar, mimosa’, Desse modo, a denominação *mimosa* para a fruta em questão

¹⁰⁸ Informante 4 do ponto 212 (Mulher, faixa etária II: 50 a 65 anos, de Campo Mourão/PR).

parece estar ligada às qualidades da casca, fina, macia e delicada, estendida pelos informantes à denominação do fruto.

Visualiza-se na carta 5g a distribuição diatópica das três últimas variantes menos produtivas *laranja-cravo/a* (1,18%), *mandarina* (0,18%) e *carioquinha* (0,09%). Nesta carta, observa-se a presença de *laranja-cravo/a* no litoral sul de SC (pontos 233 e 232, Criciúma e Tubarão) e em duas localidades do litoral norte deste Estado (pontos 225 e 228 - São Francisco do Sul e Itajaí). No Estado de SP, o item ocorre, isoladamente, em três pontos (186 – Registro, 173 – Campinas e 176 – Guaratinguetá). Fora dos limites do subfalar sulista, *laranja-cravo* apresenta-se também no ponto de controle 130 – Unaí/MG.

A variante *laranja-cravo/a* encontra-se dicionarizada na acepção buscada para o referente nos dois dicionários consultados - Houaiss (2001) e Ferreira (2004) – indicada pelo primeiro como um regionalismo do Brasil e no segundo como um brasileirismo, sem, contudo, indicarem a região a que se referem. Sabe-se, todavia, que *laranja-cravo* é uma forma recorrente no Nordeste do país, conforme Romano e Aguilera (2009) e está amplamente documentada na carta L05 do Atlas Linguístico do Brasil (CARDOSO et al., 2014b).

A variante *mandarina* apresenta-se em Barracão/PR (ponto 223) e Terra Boa/PR (ponto 209) e evidencia traços do espanhol no português. Em Houaiss e Villar (2001) e Ferreira (2004), encontram-se duas entradas para o verbete. Na segunda entrada, a variante *mandarina* é apresentada como um regionalismo do RS (HOUAISS; VILLAR, 2001) e como um brasileirismo do RS (FERREIRA, 2004), sinônimo de *tangerina*. Todavia, a ocorrência de *mandarina* no *corpus* restringe-se a duas localidades paranaenses, uma na fronteira do Brasil com Paraguai e Argentina (Barracão – ponto 223) e outra no norte do Estado (209 – Terra Boa). Há indícios de que esta variante pertence ao *falar sulista*, entretanto, pelo fato de o ALERS não documentar a forma *mandarina* em nenhum dos três Estados da Região Sul e por sua baixa representatividade no *corpus* sob análise, não seria conveniente incluí-la no conjunto de variantes que tipificam o *falar sulista*.

De acordo com Cunha (2010), *mandarina* provém de *mandarim*, “alto funcionário público, antigamente na China” e *mandarina* tanto pode se referir à esposa do mandarim, quanto à tangerina, datando o ano 1873. O Dicionário Real Academia Espanhola (DRAE), no verbete *mandarina*, dentre outras acepções, apresenta a remissiva para *naranja mandarina* (tangerina): “Variedad que se

distingue en ser pequeña, aplastada, de cáscara muy fácil de separar y pulpa muy dulce”¹⁰⁹ e, de acordo com Corominas, o nome é dado ao fruto “por alusão à cor do traje dos mandarins” (COROMINAS apud HOUAISS; VILLAR, 2001).

A *hápx legomena carioquinha* ocorre apenas na fala de um informante da cidade de São Paulo como uma segunda resposta e não se encontra registrada em nenhuma das obras consultadas.

4.5.1 Tratamento estatístico

Foram cinco as variantes submetidas aos tratamentos estatísticos devido à sua distribuição diatópica, bem como a sua produtividade no *corpus*: *mexerica*, *poncã*, *tangerina*, *bergamota* e *mimosa*.

As hipóteses formuladas para validação estatística são:

- Hipótese nula 17: a variante *mexerica* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 18: a variante *poncã* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 19: variante *tangerina* apresenta distribuição de produtividade heterogênea em todos os Estados.
- Hipótese nula 20: a variante *bergamota* apresenta distribuição de produtividade homogênea nos Estados de Santa Catarina e Rio Grande do Sul.
- Hipótese nula 21: a variante *mimosa* apresenta distribuição de produtividade heterogênea nos Estados do Paraná e Santa Catarina.

A aplicação do teste KS revela que das cinco principais variantes (*mexerica*, *poncã*, *tangerina*, *bergamota* e *mimosa*) apenas *tangerina* apresenta aderência à distribuição normal (p-valor > 0,05). Desse modo, a partir destes resultados, torna-se necessária a aplicação de testes não-paramétricos para verificar as hipóteses aventadas. Vale ressaltar, contudo, que a variante *bergamota* não

¹⁰⁹ “Variedade que se distingue por ser pequena, achatada, de casca muito fácil de se retirada e polpa muito doce.” (TN)

ocorre em GO, MG e SP, ao passo que *mimosa* não se apresenta nos Estados de MT, MS, GO, MG e RS.

A hipótese nula 17 consiste em comprovar se a variante mais produtiva no *corpus* apresenta comportamento heterogêneo no território investigado, uma vez que a carta 5a revela uma ampla distribuição diatópica deste item, que não ocorre, todavia, em parte do RS e de SC. Com o objetivo de verificar se as variâncias dos registros da variante *mexerica* são heterogêneas em todos os Estados analisados, foi aplicado o teste de Kruskal-Wallis cujo resultado mostra que o teste alcançou nível significativo, podendo-se assumir que as variâncias dos grupos são heterogêneas, ou seja, há diferença significativa entre as médias de todos os Estados ($p < 0,001$) e a hipótese nula 17 pode ser aceita.

A carta 5h ratifica os resultados do teste estatístico evidenciando que, de fato, há diferentes índices de produtividade da variante *mexerica* de acordo com a área do território investigado. Na referida carta, verifica-se que a variante predomina em grande parte do território, o que caracteriza um possível falar paulista, principalmente nas áreas em que ocorre com 100% de produtividade (Estado de SP e adjacências). Desse modo, o falar paulista estaria em todo o Estado de SP, sul de MG e Triângulo Mineiro, por onde se expande para GO, e para a porção leste do MT e MS, além de alcançar todo o norte paranaense. As áreas em que a ocorrência de *mexerica* se apresenta com outros índices percentuais evidenciam os limites fluidos deste possível falar, que pode avançar para o restante do Estado do MT e MS, além de adentrar o centro-sul do PR e pelo oeste deste Estado alcançar o oeste catarinense. Por um corredor central, o *falar paulista* perpassa cidades que foram pontos de passagem dos tropeiros, como Lapa-PR (ponto 222) e chega à região dos campos de Lages (pontos 231 – Lages/SC), área de SC em que, reconhecidamente, ocorreu um intenso intercâmbio de gaúchos e paulistas durante o Ciclo do Tropeirismo no Brasil. Esta região de SC que recebe influências do *falar paulista* já fora reconhecida em trabalhos de cunho dialetal como os de Altenhofen (2005), Margotti e Vieira (2006) e Romano e Aguilera (2014, no prelo) e, nesta pesquisa, novamente é atestada.

A hipótese nula 18 consiste em verificar a heterogeneidade da variante *poncã* nos Estados e o teste Kruskal-Wallis novamente indica que o resultado alcançou nível significativo ($p < 0,001$). Com isso, assume-se que as variâncias dos grupos são heterogêneas, demonstrando que há diferença

significativa entre as médias de todos os Estados e a hipótese nula 18 pode ser aceita.

A carta 5i apresenta esta heterogeneidade, evidenciando a distribuição areal de *poncã*, que, embora se apresente em grande parte do território investigado (carta 5b), concentra-se com maior incidência (100%), no oeste e sudoeste paulista, expandindo-se para o norte do PR e para três localidades do MS (ponto 114 – Paranaíba, 115 – Campo Grande, 116 – Nioaque). Há também ocorrência de 100% de *poncã* no norte paulista (157 – Ribeirão Preto e 158 – Barretos), região do Vale do Paraíba (168 – Mogi Mirim, 175 – Taubaté, 176 – Guaratinguetá), além do litoral norte do Estado (ponto 180 – Caraguatatuba). No sul de MG, há 100% de produtividade de *poncã* em Itajubá (ponto 149) e Passos (140). No Triângulo Mineiro, *poncã* está em 100% das respostas dos habitantes de Patos de Minas (ponto 136) e, em GO, apresenta-se no ponto 122 (Goiás-GO). Ademais, o item mostra-se com produtividade significativa também em quatro localidades MT: Poxoréu (ponto 106), São Félix do Araguaia (107), Cuiabá (ponto 108), e Cáceres (110). Além das ocorrências pontuais de 100% de *poncã* no PR, em Toledo (ponto 215), Lapa (222), Morretes (221), e no ponto 234 (Três Passos/RS), o item mostra-se significativo em três dos 10 pontos de controle (103 – Aripuanã/MT, 118 – Porangatu/GO, 121 – Formosa/GO). Nas demais áreas de ocorrência, a incidência de *poncã* varia de 25 a 75% de produtividade.

A concentração do item no interior paulista com influência também no norte do PR e em cidades do MS ratifica a inferência que se faz sobre essa variante caracterizar um possível *falar paulista* cujos limites se apresentam fluidos, não só para os designativos em pauta, mas também para outras variantes referentes a outras questões, levando a inferir que o *falar paulista* se irradia a partir do Estado de SP, onde apresenta maior concentração de determinada forma lexical e se expande para os Estados limítrofes com diferentes índices de ocorrência.

A mesma situação foi verificada no teste da hipótese nula 19, que buscou averiguar a distribuição da variante *tangerina* em todo o território investigado. A partir da aplicação do teste Kruskal-Wallis, verifica-se que as variâncias dos grupos são heterogêneas ($p = 0,021$), evidenciando, portanto, a aceitação desta hipótese.

Mais uma vez a carta de arealidade gradual ratifica este comportamento heterogêneo do item no território investigado. Na carta 5j, verifica-se

que *tangerina* ocorre em quase todo o território, porém com diferentes índices percentuais, o que não recomenda enquadrá-la como uma forma típica de nenhum dos dois falares aqui defendidos, o *sulista* e o *paulista*. Ademais, verifica-se, por meio de outros trabalhos como o de Romano; Aguilera (2008), Romano; Aguilera (2009), bem como pela carta L05 do ALiB (CARDOSO et al., 2014), o quanto *tangerina* é uma forma produtiva nas capitais da região Norte e Nordeste do país. Um estudo que contemplasse os pontos do interior destas duas regiões brasileiras em comparação com os dados apresentados nas cartas 5, 5c e, principalmente 5j, poderia evidenciar a maior ou menor produtividade do item no Norte e no Nordeste em relação ao território aqui investigado. Todavia, como o processo de cartografia de dados do interior do ALiB está em sua fase inicial (planejamento), por ora, acredita-se que não seria conveniente considerar *tangerina* como uma forma lexical que tipifica o *falar paulista* ou o *sulista*, antes, trata-se de uma forma comum a outras regiões brasileiras.

A hipótese nula 20 consiste em verificar a homogeneidade do item lexical *bergamota* nos Estados de SC e RS. Todavia, a aplicação do teste Mann-Whitney revela o comportamento heterogêneo desta variante, uma vez que o resultado do teste alcançou nível significativo ($p = 0,007$). Face a esses resultados, a hipótese nula 20 pode ser rejeitada.

Apesar de ocorrer em todo o território do RS e de SC, expandindo-se para o oeste do PR e também no interior por localidades situadas nos antigos Caminhos das Tropas (ponto 222- Lapa e 214 -Piraí do Sul), conforme se verifica na carta 5d, o teste estatístico evidencia que *bergamota* apresenta uma distribuição de produtividade heterogênea o que fica mais evidente na carta 5k. *Bergamota* e suas variantes fonéticas e morfofonêmicas, de fato, são mais produtivas no Estado do RS (Figura 84), onde ultrapassam 50% das respostas consideradas. Portanto, é uma forma regional que, assim como a *mexerica*, possibilita o traçado de isoléxicas para a definição do denominado *falar sulista*. Este falar compreende, portanto, o Estado do RS, adentra o centro e o oeste catarinense até atingir o oeste paranaense, onde a produtividade do item atinge o percentual de 100%, área da região sul já identificada por Altenhofen (2005) e Romano; Aguilera (2014, no prelo). Todavia, os limites do *falar sulista*, no caso de influência sul-riograndense em direção ao oeste catarinense e paranaense e de línguas de colonização (italiano), são fluidos e, nesta carta, se evidenciam a partir de expansão para o restante do Estado de SC até

atingir localidades do centro-sul paranaense, onde o índice de produtividade cai para 75%, 50% e 25%, regressivamente.

No caso da variante *mimosa*, o teste da hipótese nula 21, que se refere à distribuição de produtividade heterogênea nos Estados do Paraná e Santa Catarina, revela que o resultado da aplicação de Mann-Whitney não alcançou nível significativo ($p = 0,682$), podendo-se assumir que as variâncias dos grupos são homogêneas e a hipótese nula 21 pode ser rejeitada.

Verifica-se na carta 5I que *mimosa* localiza-se tanto no PR quanto em SC com diferentes índices, predominando, todavia, a produtividade de 100% em cidades paranaenses, enquanto em apenas uma cidade catarinense observa-se a ocorrência máxima de *mimosa*. A área lexical identificada para *mimosa* coincide com a encontrada nos dados do ALERS e dá indícios de que há especificidades lexicais nesta região do subfalar sulista de Nascentes (1953) que não se identificam nem com o falar do RS nem com o de SP. Seriam evidências estas de um possível falar paranaense ou catarinense? Há que se aprofundar ainda sobre este fato para verificar tal assertiva, porém, neste trabalho, identificam-se itens característicos de subáreas lexicais constantes do *falar sulista* de influência sul rio-grandense e línguas de colonização.

Em síntese, o Quadro 8 apresenta um resumo das hipóteses aventadas para a questão 039.

Quadro 8 - Síntese das hipóteses para a questão 039 do QSL

	ACEITA	REJEITADA
HIPÓTESE NULA 17 [#]	X	
HIPÓTESE NULA 18 ^{##}	X	
HIPÓTESE NULA 19 ^{###}	X	
HIPÓTESE NULA 20 ^{####}		X
HIPÓTESE NULA 21 ^{#####}		X

mexerica é heterogênea em todos os Estados

poncã o é heterogênea em todos os Estados

tangerina é heterogênea em todos os Estados

bergamota é homogênea nos Estados de RS e SC

mimosa é heterogênea nos Estados de PR e SC

A partir da análise empreendida para a questão 039 do QSL, podem-se estabelecer as seguintes conclusões:

- (i) A variante mais produtiva, *mexerica*, denominação motivada pelo cheiro característico que denuncia quem manuseou ou comeu o fruto, está amplamente distribuída pelo território investigado e pode ser considerada uma forma que tipifica o falar paulista, uma vez que o teste estatístico empreendido para hipótese nula 17 confirma a heterogeneidade do item que ocorre com diferentes índices, diminuindo de produtividade nos Estados da Região Sul do país (Carta 5h).
- (ii) *Poncã* é uma forma lexical que também tipifica o falar paulista, embora esteja amplamente distribuída por todo o território, todavia com diferentes índices de produtividade, o que confirma a hipótese nula 18 acerca da heterogeneidade do item no território investigado.
- (iii) A variante *tangerina*, forma considerada padrão pelos dicionaristas consultados, encontra-se distribuída em grande parte do território investigado com diferentes índices, o que confirma a hipótese nula 19. Todavia, *tangerina* não caracteriza nenhum dos dois falares aqui defendidos, pois é uma forma identificada em outras regiões brasileiras, onde atinge grande produtividade, conforme se verifica na carta L05 do ALiB (CARDOSO et al., 2014b) e descrevem detalhadamente Romano e Aguilera (2008) e Romano e Aguilera (2009)
- (iv) A variante *bergamota* revela contato do português com línguas de colonização, no caso o italiano, e tipifica o falar sulista de influência sul rio-grandense que avança em sentido oeste catarinense e sudoeste do PR. O teste sobre a homogeneidade do item nos Estados de SC e RS não confirma a hipótese nula 20, ou seja, a variante apresenta comportamento heterogêneo na região em que ocorre, pois se apresenta mais produtiva, principalmente no RS (Carta 5k), o que evidencia a influência rio-grandense no falar sulista.
- (v) A variante *morcote* é um item lexical motivado pelo sobrenome de um dos pioneiros no cultivo deste híbrido da fruta, e apresenta o maior número de variantes fonéticas e

morfofonêmicas. A produtividade de *morcote* no *corpus* sob análise é baixa e sua distribuição espacial é esparsa, não se concentrando em determinada área ou território, o que não justifica teste de hipótese para o item. Ou seja, *morcote* não tipifica nem o *falar sulista* nem o *paulista*.

- (vi) *Mimosa*, por sua vez, embora se apresente com baixa produtividade, sob o ponto de vista diatópico, revela uma subárea lexical no *falar sulista* que compreende a parte leste do PR e de SC, cuja área de ocorrência coincide, inclusive, como o que se observa na carta 62 do ALERS. Esta área de ocorrência apresenta-se de forma homogênea, o que não confirma a hipótese nula 21. De fato, há indícios de subáreas lexicais dentro dos dois falares aqui defendidos.
- (vii) As variantes menos produtivas *laranja cravo/a*, *mandarina* e *carioquinha* ocorrem esparsamente no território investigado e estão atestadas em outros trabalhos geolinguísticos.

Do exposto, conclui-se que, para a questão 039, o território investigado novamente apresenta comportamento semelhante ao que se verificou nas demais questões analisadas neste estudo. Ou seja, a área delimitada por Antenor Nascentes como subfalar sulista, de acordo com as cartas de arealidade aqui analisadas e também com o que se observa nas cartas 5m e 5n (vol. 2), pode ser dividida em dois grandes falares: (i) o *falar paulista*, localizado na porção setentrional (predomínio de *mexerica* e *poncã*) que se expande para Estados da região Centro-Oeste e (ii) o *falar sulista*, situado na porção meridional do Brasil, caracterizado pelo maior número de formas regionais e variantes de origens não-lusas, evidenciando traços de língua de colonização.

Vale salientar, todavia, que os limites e abrangência de ambos os falares são fluidos, ora havendo a interinfluência de um em outro, ora prevalecendo a delimitação de áreas lexicais distintas. Ademais, somam-se subáreas lexicais que se localizam tanto em um falar quanto em outro, caso de *morcote*, na área do *falar paulista* e de *mimosa*, na área do *falar sulista*.

CONSIDERAÇÕES FINAIS

Voltando-se aos objetivos desta pesquisa, observa-se que o presente trabalho, além de discutir a divisão dialetal de Antenor Nascentes (1953) no que se refere à área denominada pelo estudioso como subfalar sulista, propõe uma revisão desta divisão com base em dados lexicais. Desse modo, após o estudo minucioso do comportamento diatópico das variantes lexicais documentadas para as cinco questões analisadas, chegou-se à carta 6 (vol. 2), que evidencia a localização geográfica dos dois grandes possíveis falares aqui defendidos: o *falar paulista* e o *falar sulista*.

O *falar paulista* difunde-se a partir do Estado de SP, com influência na região norte do PR, oeste do MS, sudoeste e interior do Estado de GO, sul de MG e Triângulo Mineiro. Ademais, este falar, até onde se verificou, alcança também os pontos de controle, ou seja, não se limita à área geográfica do *subfalar sulista* definida por Nascentes (1953) e diferencia-se da região meridional do território investigado. 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 selecionadas.

Pelo que se verificou, o *falar paulista* caracteriza-se por uma maior homogeneidade lexical na área geográfica, revelando um menor número de coocorrência de variantes lexicais. Ou seja, na área do *falar paulista*, há o predomínio das formas mais produtivas que são consideradas padrão: *córrego*, *geleia*, *bolinha de gude*, *menino*, *moleque*, *mexerica*, *poncã*, que, à medida que avançam para a porção sul do país, apresentam uma diminuição em seu índice de produtividade (cartas 1l, 2f, 3k, 4l, 5h, 5i e 5j).

Este *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 interior adentro, chegando a Minas Gerais e, por esse caminho, alcançou o Estado de GO e o Estado do MT.

Os limites que definem o *falar paulista* são virtuais e fluidos, ora alcançando para 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.

Não obstante, ainda nesta área setentrional do território investigado, verifica-se a ocorrência de uma subárea lexical que, possivelmente, recebe influência de outro subfalar de Nascentes (1953), o mineiro, conforme se verifica na carta 3h, que evidencia a presença de *bilosca* na área do Triângulo Mineiro, oeste de GO e noroeste do Estado de SP. Nesse sentido, a pesquisa aponta para a necessidade de mais estudos acerca de outros subfalares, considerando a mesma metodologia aqui empreendida, o que poderá contribuir para uma definição mais precisa dos limites desse possível *falar paulista*.

O *falar sulista*, por sua vez, localiza-se principalmente na porção meridional do Brasil, contemplando o Estado do RS e, por um corredor do oeste catarinense e sudoeste paranaense, atinge o Estado do MS. Este corredor lateral em SC e PR já fora identificado sob o ponto de vista fonético e morfossintático, inicialmente, por Koch (2000) e, posteriormente, por Altenhofen (2005) e mais, recentemente, sob a perspectiva lexical, por Romano e Aguilera (2014, no prelo). 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 como o alemão e o italiano, o que se evidencia por variantes presentes na norma lexical dos informantes, como *sanga*, *arroio*, *chimia*, *bolita* e *bergamota*.

Este falar também apresenta outras variantes típicas que revelam o contato do português com o tupi, como, por exemplo, *guri* e *piá*. No tocante a esses itens lexicais, o *falar sulista* reflete-se em dois Estados da região Centro-Oeste, o MT e o MS, em consequência de correntes migratórias mais recentes. Os limites desse falar também são virtuais e fluidos, e esse falar, além de se expandir para o Centro-Oeste, adentra o Estado de SC e, por um corredor central do PR, passando em localidades como Lapa e Pirai do Sul, atinge cidades paulistas localizadas no Vale do Ribeira até chegar a Itapetininga e Sorocaba.

Essa configuração diatópica de determinadas variantes lexicais evidencia traços da história social dessas localidades, uma vez que tais variantes revelam rastros do tropeirismo, ou seja, os tropeiros gaúchos além de heranças deixadas na cultura material de determinadas cidades, deixaram também suas

marcas na norma lexical que até hoje ecoam no vocabulário ativo de informantes naturais daquela região.

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* e *mimosa* que merecem atenção mais detalhada. São denominadas como subáreas pelo fato de ora se identificarem com o *falar paulista*, ora com o *falar sulista*, figurando, portanto, em uma área geográfica comum aos dois grandes falares.

Por considerar como *corpus* de análise um expressivo conjunto de dados, a pesquisa possibilitou a comparação de variantes lexicais pertencentes a regiões administrativas diferentes. Sob uma mesma metodologia, o trabalho apresentou o comportamento diatópico de variantes em uma área geográfica que ultrapassa limites políticos (estaduais e regionais).

Essa comparação somente foi possível devido à ferramenta computacional desenvolvida no âmbito da tese, o [SGVCLin], uma vez que o grande rol de informações linguísticas foi passível de quantificação e, sobretudo, de cartografia sistemática por meio dessa ferramenta. Foram abordadas cinco questões, que originaram 10 relatórios, em um total de 3672 registros, que perfazem, sumariamente, 59 itens lexicais. Dentre as cartas geradas na ferramenta, para o estudo com foco na diatopia, as questões selecionadas resultaram num conjunto de 71 mapas que retratam o comportamento espacial das variantes lexicais no território do subfalar sulista e adjacências.

Ademais, o caráter inovador do [SGVCLin] contribui para o desenvolvimento da Geolinguística brasileira, possibilitando a geração de cartas linguísticas e relatórios diversos de forma automática com economia de tempo e investimento financeiro. Salienta-se também a grande receptividade da ferramenta pela comunidade científica no âmbito nacional (ROMANO; SEABRA; OLIVEIRA, 2014) e internacional (SEABRA; ROMANO; OLIVEIRA, 2014).

Outro aspecto inovador apresentado no trabalho refere-se ao tratamento estatístico empreendido, utilizando-se o programa SPSS para os testes das hipóteses. A estatística inferencial, desse modo, vem contribuir para os estudos linguísticos, possibilitando afirmações respaldadas em critérios matemáticos que auxiliam a interpretação das cartas linguísticas. Foi possível, a partir desses testes, verificar a homogeneidade ou heterogeneidade de determinada variante não apenas

pela análise visual das cartas, mas também pela consideração de variáveis numéricas que extrapolam números absolutos e relativos. Por meio do trabalho de Romano e Seabra (2014b), foi possível verificar a aplicabilidade desses testes para esta tese e, das 21 hipóteses testadas, 15 foram aceitas e 6, rejeitadas, o que possibilita afirmar que a área correspondente ao território investigado apresenta, de fato, comportamento heterogêneo.

Retomando as perguntas de pesquisa que nortearam o desenvolvimento do trabalho (p. 23), chega-se às seguintes respostas:

- Por meio do léxico é possível delinear ou estabelecer áreas dialetais?

As áreas lexicais são possíveis de serem definidas, desde que sejam considerados os aspectos sociais e históricos envolvidos na configuração diatópica de determinada região, levando em conta fatores tais como migração e imigração, que, indubitavelmente, interferem na disseminação de variantes.

- A divisão de Antenor Nascentes (1953), no que se refere ao subfalar sulista, ainda se sustentaria no século XXI sob o ponto de vista lexical?
-

Considerando que o PB, em meados da segunda década do século XXI, já se encontra documentado em áudio pela equipe do Projeto ALiB, pode-se afirmar que, sob o ponto de vista do léxico, a divisão de Nascentes (1953), para o subfalar sulista, não é mais válida. Parece inadequado considerar como pertencente ao mesmo falar Estados linguística e sócio-historicamente distintos como SP e RS, uma vez que apresentam aspectos que os definem e os diferenciam.

- Atualmente, qual seria a definição e quais os limites e abrangência deste subfalar ou falar sulista?

O território investigado pode ser dividido em duas grandes áreas: (i) a meridional (*falar sulista*) e (ii) a setentrional (*falar paulista*), caracterizadas anteriormente e que apresentam limites virtuais e fluidos.

- A área geográfica correspondente ao subfalar sulista, sob ponto de vista do léxico, apresenta mais semelhanças ou diferenças?

Verificou-se que o comportamento das variantes apresenta diferenças significativas, ratificadas pelas cartas de arealidade gradual e também por meio dos testes de hipóteses.

- Há subáreas lexicais na área geográfica investigada? Como elas estão configuradas?

Há indícios da existência de subáreas lexicais em espaços comuns tanto ao *falar sulista* quanto ao *falar paulista*, o que indica a necessidade de estudos mais apurados sobre essas áreas, bem como de outros subfalares definidos por Nascentes (1953), a fim de delimitar com mais precisão qual a abrangência desses dois grandes falares por ora identificados, bem como empreender análises considerando outros níveis linguísticos como o fonético e o morfossintático com vistas a traçar as áreas dialetais para o PB.

- A distribuição das variantes lexicais reflete os processos de ocupação e povoamento, bem como os fluxos de migração observados na área em foco?

Concluindo e parafraseando a epígrafe que introduz o primeiro capítulo desta tese: estudar o português do Brasil é estudar a história de nossa formação. O léxico propicia este tipo de estudo, pois evidencia aspectos históricos e sociais. Desse modo, pelas análises empreendidas, comprovam-se que as variantes se disseminam de um lugar para o outro e a presença de formas lexicais em determinado território indica sua origem e os rastros deixados pelo homem no decorrer de sua história.

REFERÊNCIAS

- ADAS, M. **Panorama geográfico do Brasil: contradições, impasses e desafios**. São Paulo: Moderna, 2004.
- AGUILERA, V. de A. **Atlas Linguístico do Paraná**. Curitiba: Imprensa Oficial, 1994.
- _____. Léxico e áreas dialetais: o que podem demonstrar os dados do ALIB. **ANAIIS - VI Congresso Internacional da Abralín / Dermeval da Hora (org.)**. - João Pessoa: Ideia, 2009. Vol. 2, p.4219-4233.
- _____. De onde vieram por onde anda as nossas libélulas e jacintas? Um estudo da etimologia popular com base em dados do Atlas linguístico do Brasil (ALIB). **Estudos linguísticos e literários**, n 41. Salvador: Programa de Pós-graduação em Língua e Cultura, Programa de Pós-graduação em Literatura e Cultura, Universidade Federal da Bahia, jan-jun. 2010, p. 291-309.
- ALGATÃO, F. C. de S. O tropeiro como propagador cultural e mola mestra da cultura cafeeira no século XIX. **Histórica** – Revista Eletrônica do Arquivo Público do Estado de São Paulo. N. 41. Ano 6. Abril, 2010. Disponível em:
<http://www.arquivoEstado.sp.gov.br/historica/edicoes_ateriores/pdfs/historica41.pdf
> Acesso em 26 mai 2013.
- ALMEIDA, A. de. **Vida e morte do tropeiro**. São Paulo: Martins; EDUSP, 1981.
- 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. (orgs.) **Atlas Linguístico-Etnográfico da Região Sul do Brasil – ALERS: Cartas Semântico-Lexicais**. Porto Alegre: Editora da UFRGS; Florianópolis: Ed. UFSC, 2011.
- ALTINO, F. C. **Atlas Linguístico do Paraná II** . 2007. 2v. Tese. (Doutorado em Estudos da Linguagem), Universidade Estadual de Londrina, Londrina, 2007.
- ALTINO, F. C. Atlas Linguístico do Paraná – v.II: comentários sobre a dialectometria. **Estudos Linguísticos**. São Paulo, v. 41, n.2 , 2012, p. 818-832.
- ALVAR, M. Hacia los conceptos de lengua, dialecto y hablas. Alicante : Biblioteca Virtual Miguel de Cervantes, 2007 [1961]. **Nueva Revista de Filología Hispánica**, Año 15 (1961), pp. 51-60. Disponível em:
<http://bib.cervantesvirtual.com/servlet/SirveObras/p184/02583952000205051154480/p0000001.htm#l_0_>. Acesso 07 set. 2013.
- AMARAL, A. **O dialeto capira: gramática, vocabulário**. 4. Ed. São Paulo: HUCITEC; Brasília: INL, 1981 [1920].
- AMARAL, L. **História geral da agricultura brasileira**. 2. ed. v. 2. São Paulo:

Companhia Ed. Nacional, 1958. Disponível em:
<http://www.brasiliana.com.br/obras/historia-geral-da-agricultura-brasileira-v2> .
 Acesso em: 07 jul. 2013.

ANDRADE, O. de. **Poesias reunidas**. 2. ed. Rio de Janeiro: Civilização Brasileira, 1972.

_____. O manifesto antropófago [1924]. In: TELES, G. M. **Vanguarda européia e modernismo brasileiro**: apresentação e crítica dos principais manifestos vanguardistas. 3ª ed. Petrópolis: Vozes; Brasília: INL, 1976. Disponível em:
<http://www.ufrgs.br/cdrom/oandrade/oandrade.pdf>. Acesso em 21 ago. 2013.

ANTONIL, A. J. **Cultura e opulência do Brasill, por suas drogas e minas**. 2. ed. Rio de Janeiro : Typ. Imp. e Const. de J. Villeneuve e Comp, 1837.

ARAGÃO, M. do S. S. de; BEZERRA DE MENEZES, C. **Atlas Linguístico da Paraíba**. Brasília: UFPB; CNPq, Coordenação Editorial, 1984.

AUGUSTO, V. L. dos S. **Atlas Semântico-Lexical de Goiás**. 2012. 3 v. Tese (Doutorado em Linguística) Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2012.

BENKE, V.C.M. **Tabus linguísticos nas capitais do Brasil**: um estudo baseado em dados geossociolinguísticos. 2012. 313 p. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2012.

BIDERMAN, M. T. C. Léxico, testemunho de uma cultura. In.: **Actas do XIX Congreso Internacional de Linguística e Filoxía Románicas**. A Coruña: Fundación Pedro Barrié de la Maza, Conde de Fenosa, 1992, p. 397-405.

_____. O Português Brasileiro e o Português Europeu: Identidade e contrastes. In: **Revue belge de philologie et d'histoire**. Tome 79, fasc. 3, 2001. Langues et littératures modernes - Moderne taal- en letterkunde. pp. 963-975. Disponível em:
http://www.persee.fr/web/revues/home/prescript/article/rbph_0035-0818_2001_num_79_3_4556>. Acesso em: 07 ag. 2013.

BOLÉO, M. de P. **Brasileirismos**: problemas de método. Separta da Brasilia, vol. III. Coimbra: Coimbra Ed.Limitada, 1943.

BONADIO, G O tropeirismo e a formação do Brasil. In.: BONADIO, G. (org.). **O Tropeirismo e a formação do Brasil**. Sorocaba: Academia Sorocabana de Letras; Fundação Ubaldino do Amaral; Skol; Momesso; Caracu, 1984, p. 45-47.

BRANDÃO, S. F. **A geografia linguística no Brasil**. São Paulo: Ática, 1991.

BUENO, E. **Brasil**: uma história, cinco séculos de um país em construção. Rio de Janeiro, Leya, 2012.

CALLAGE, R. **Vocabulário gaúcho**. Porto Alegre: [s/ed.], 1964 [1926].

CAMARA Jr. J. M. **Dicionário de Linguística e gramática**: referente à língua portuguesa. 26 .ed. Petrópolis: Vozes, 2007.

CAMARGO, J. F. de. **Crescimento da população no Estado de São Paulo e seus aspectos econômico**: ensaio sobre as relações entre a demografia e a economia. v.1. São Paulo: FFLCH, 1952.

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

_____. 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.

_____. A Geolinguística no Terceiro Milênio: Monodimensional ou pluridimensional? **Revista do GELNE**. Ano 4, N. 2, 2002. Disponível em: <http://www.gelne.org.br/RevistaGelne/arquivos/artigos/art_34a01e3a7b2f8deaa71b52a3df2d54c0_12.pdf>. Acesso em 02 set. 2013.

_____. **Geolinguística**: tradição e modernidade. São Paulo: Parábola Editorial, 2010.

CARDOSO, S. A. M. S. et al. **Atlas Linguístico do Brasil - Vol. 1 (Introdução)**. Londrina: EDUEL, 2014a.

CARDOSO, S. A. M. S. et al. **Atlas Linguístico do Brasil - Vol. 2 (cartas Linguísticas)**. Londrina: EDUEL, 2014b.

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

CARVALHO, D. F. Café, ferrovias e crescimento populacional: o florescimento da região noroeste paulista. **Histórica** – Revista Eletrônica do Arquivo Público do Estado de São Paulo. N. 27. Ano 3. Nov, 2007. Disponível em: <<http://www.historica.arquivoEstado.sp.gov.br/materias/anteriores/edicao27/materia02/>> Acesso em 06 jul. 2013.

CASTILHO, A. T. de. Variação dialetal e ensino institucionalizado da língua portuguesa. In.: BAGNO, M. (org.). **Linguística da Norma**. São Paulo: Loyola, 2002, p. 27-36.

_____. O português do Brasil. In.: ILARI, R. **Linguística românica**. 3. Ed. São Paulo: Ática, 2004, p. 269-281.

_____. **Nova Gramática do Português brasileiro**. São Paulo: Contexto, 2010.

CASTRO, E. de. **Ensaio de geographia linguística**. 2. ed. São Paulo: Companhia Editora Nacional, 1941. Disponível em: <<http://www.brasiliana.com.br/obras/ensaios-de-geografia-linguistica/pagina/3/texto>> Acesso em: 05 mai 2013.

CASTRO, Y. P. de. **Falares africanos na interação social do Brasil Colônia**. Salvador: Centro de Estudos Baianos, n. 89, 1980.

CECHIN, J. **A construção e a operação das rodovias no Brasil do século XIX**. 1978. 107 p. Dissertação (Mestrado). Instituto de Economia. Universidade Estadual de Campinas, Campinas, 1978.

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

CIGOLINI, A; MELO, L. de. LOPES, N. **Paraná: quadro natural, transformações territoriais e economia**. Curitiba: Renascer, 1998.

CINTRA, L. F. L. **Estudos de Dialectologia Portuguesa**. 2. ed. Lisboa: Sá da Costa, 1995.

COELHO, F. A. **A língua portuguesa**. Porto: Universal, 1881.

_____. Os dialectos românicos ou neolatinos na África, Ásia e América. **Estudos linguísticos crioulos**. Reedição de artigos publicados no Boletim da Sociedade de Geografia de Lisboa. Lisboa: Academia Internacional de Cultura Portuguesa, 1967.

COHEN, M. A. A. de M. et al. Filologia bandeirante. **Filologia e linguística portuguesa**, n. 1, 1997, p. 79-94. Disponível em: <
<http://www.fflch.usp.br/dlcvlport/site/images/arquivos/FLP1/megale1997.pdf>>. Acesso em: 11 ago. 2013.

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

CORTESÃO, J. **Raposo Tavares e a formação territorial do Brasil**. Vol. 1. Lisboa: Portugália Editora, 1966.

COSERIU, E. **Lições de linguística geral**. Trad. Evanildo Bechara. Rio de Janeiro: ao Livro Técnico, 2004 [1980].

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

COSTA, D. de S. S. Designações para “tangerina” no Brasil central: reflexões com base em dados do projeto ALiB. XX Seminário do CELLIP – Centro de Estudos Linguístico e Literários do Paraná. **Anais...** CD-ROM, Londrina: UEL, 2011, p. 1-12.

COSTA, D. de S. S. **O léxico indígena nas capitais brasileira: um estudo geolinguístico**. 209, p. 2012. Dissertação (Mestrado em Estudos de Linguagens). Universidade Federal do Mato Grosso do Sul, Campo Grande, 2012.

COSTA, D. de S. S; ISQUERDO, A. N. Um estudo etnolinguístico de designativos para ‘gambá’ no Brasil Central: contribuições do Projeto ALiB. **Estudos Linguísticos**, São Paulo, n. 41.2, 2012, p. 779-792.

COSTA, F. A. P. da. **O vocabulário pernambucano**. Recife: Secretaria da Educação e Cultura; MEC, 1976 [1937].

COSTA, G. P. da. Mudanças no espaço geográfico do Brasil. In.: MURRIE, Z. de F. (coord.). **História e Geografia: livro do estudante - ensino fundamental**. 2. ed. Brasília: MEC/INEP, 2006, p. 28-40.

COSTA NETO, P. L. O. **Estatística**. São Paulo: Edgard Blucher Ltda., 1977.

CRISTIANINI, A. C. **Atlas Semântico-lexical da região do Grande ABC**. 2007. 3. v. Tese (Doutorado em Linguística) Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2007.

CUBA, M. A. **Atlas Linguístico da Mesorregião Sudeste de Mato Grosso**. 2009 .2.v.. (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.

CUNHA, C. **Língua portuguesa e realidade brasileira**. 2. ed. Rio de Janeiro: Tempo Brasiliense, 1970.

_____. **Língua, nação, alienação**. Rio de Janeiro: Nova Fronteira, 1981.

_____. Conservação e inovação no português do Brasil. **O eixo e a roda**. n.5, Belo Horizonte: 1986, p. 199-230. Disponível em: <http://www.letras.ufmg.br/poslit/08_publicacoes_txt/CELSOCUNHA.pdf>. Acesso em: 11 ago. 2013.

DIÉGUES Jr. M. **Regiões Culturais do Brasil**. Rio de Janeiro: INEP/MEC, 1960.

DIETRICH, W.; NOLL, V. O papel do tupi na formação do português brasileiro. In.: NOLL, V.; DIETRICH, W. (orgs.). **O português e o tupi no Brasil**. São Paulo: Contexto, 2010, pp. 81-103.

DUARTE, A. C. O Centro-Oeste na organização regional do Brasil. In: **Fundação Instituto Brasileiro de Geografia e Estatística, Geografia do Brasil – Região Centro-Oeste**. v.1. Rio de Janeiro: IBGE, 1988.

ELIA, S. **O problema da língua brasileira**. Rio de Janeiro: INL, 1961 [1940].

_____. **A unidade linguística do Brasil: condicionamentos geoeconômicos**. Rio de Janeiro: Padrão Livraria Editora LTDA, 1979.

ELLIS Jr. A. **Raça de Gigantes: a civilização no planalto paulista**. São Paulo: Editora Hélios Limitada, 1926.

ENCARNAÇÃO, M. R. T. da. **Atlas semântico-lexical de Caraguatatuba, Ilhabela,**

São Sebastião e Ubatuba – municípios do litoral norte de São Paulo. 741 f. 2010. Tese (Doutorado em Linguística) Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2010.

FARACO, C. A. Norma-padrão brasileira. Desembaraçando alguns nós. In: BAGNO, M. (org.). **Linguística da norma**. São Paulo: Loyola, 2002, p. 37-61.

FAUSTO, B. **História do Brasil**. 11.ed. São Paulo: Editora da Universidade de São Paulo, 2003.

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

FERREIRA, C. et al. **Atlas linguístico de Sergipe**. Salvador: Universidade Federal da Bahia; Fundação de Cultura de Sergipe, 1987.

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

FRANCO, F de A. C. **Bandeiras e bandeirantes de São Paulo**. São Paulo: Companhia Editôra Nacional, 1940. Disponível em: <
<http://www.brasiliana.com.br/obras/bandeiras-e-bandeirantes-de-sao-paulo/pagina/3/texto>>. Acesso em: 01 mai 2013.

FRASSON, A. C; GOMES, S. A. **Tropeirismo: Processo civilizatório da região sul do Brasil**, 2010 Disponível em: <
http://www.educadores.diaadia.pr.gov.br/arquivos/File/2010/artigos_teses/2010/Historia/artigos/frasson_artigo.pdf> Acesso em: 25 mai 2013.

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

FRIOLI, A. O cotidiano do Tropeirismo pelo Brasil. In.: BONADIO, G. (org.). **O Tropeirismo e a formação do Brasil**. Sorocaba: Academia Sorocabana de Letras; Fundação Ubaldino do Amaral; Skol; Momesso; Caracu, 1984, pp. 25-29.

FURLAN, O. **Influência açoriana no português do Brasil em Santa Catarina**. Florianópolis: EDUFSC, 1989.

GARCIA, R. Dicionário de brasileirismos, **Revista do Instituto Histórico e Geográfico Brasileiro**, 76, 1912.

GARCÍA MOUTON, P. Dialectología y geografía lingüística. In.: ALVAR, M. (dir.). **Manual de dialectología hispánica: el español de España**. Barcelona: Ariel, 1996, p.63-77.

GHIRARDELLO, N. A companhia estrada de ferro Noroeste do Brasil. In.: GHIRARDELLO, N. **À beira da linha: formações urbanas da Noroeste Paulista** [online]. São Paulo: Editora UNESP, 2002. p. 17-65. Disponível em:

<<http://books.scielo.org/id/z3/pdf/ghirardello-9788539302420-02.pdf>>. Acesso 25 ago. 2014.

GOMES, A. C. Imigrantes italianos: entre a italianità e a brasilidade. In: Instituto Brasileiro de Geografia e Estatística. **Brasil: 500 anos de povoamento**. Rio de Janeiro, 2000. Disponível em: <<http://brasil500anos.ibge.gov.br/territorio-brasileiro-e-povoamento/italianos>>. Acesso em 04 set. 2014.

GONÇALVES, A. J. Migrações Internas: evoluções e desafios. **Estudos Avançados**. Universidade de São Paulo. v. 15, n. 43. São Paulo, set.-dez. 2001. Disponível em: <http://www.scielo.br/scielo.php?pid=S0103-40142001000300014&script=sci_arttext>. Acesso em: 23 jul. 2013.>. Acesso em: 23 jul. 2013.

GONÇALVES, J. M. **Tangerina murcote**. Disponível em: <<http://www.ceasacampinas.com.br/novo/DicasVer.asp?id=948>>. Acesso 27 out. 2014.

GUIMARÃES, E. A língua portuguesa no Brasil. **Revista Ciência e Cultura**. [online]. 2005, vol.57, n.2, pp. 24-28. Disponível em: <<http://cienciaecultura.bvs.br/pdf/cic/v57n2/a15v57n2.pdf>>. Acesso em: 27 ago 2013.

GUY, G. **Linguistic Variation in Brazilian Portuguese: Aspects of the phonology, syntax and language history**. Philadelphia: University of Pennsylvania, PhD. dissertation, 1981.

HOLANDA, S. B. de. **Raízes do Brasil**. 13. Ed., Rio de Janeiro: Editora José Olympio, 1979.

HOLLOWAY, T. H. **Imigrantes para o café: café e sociedade em São Paulo (1886-1934)**. Rio de Janeiro: Paz e Terra, 1984.

HOUAISS, A.; VILLAR, M. de S. **Dicionário eletrônico Houaiss da língua portuguesa**. Versão 1.0. Rio de Janeiro: Objetiva, 2001.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Brasil: 500 anos de povoamento**. Rio de Janeiro, 2000. Disponível em: <<http://brasil500anos.ibge.gov.br/territorio-brasileiro-e-povoamento/italianos/regioes-de-origem>>. Acesso em: 03 set. 2014.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Censo demográfico 2010: nupcialidade, fecundidade e migração - resultados da amostra**. Rio de Janeiro, 2010. Disponível em: <ftp://ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Nupcialidade_Fecundidade_Migracao/censo_nup_fec_mig.pdf>. Acesso em: 25 jul. 2013.

ISQUERDO, A. N. Léxico em tempo e espaço: a questão dos regionalismos. In.: MARIN, J. R.; VASCONCELOS, C. A. de. (orgs.). **História, região e identidades**. Campo Grande: Ed. UFMS, 2003, p. 165-181.

_____. **Atlas regionais brasileiros publicados e em curso: percursos metodológicos.** In.: MOTA, J. A.; CARDOSO, S. A. M (orgs.) **Documentos 2: Projeto Atlas Linguístico do Brasil.** Salvador: Quarteto, 2006, p. 67-94.

_____. Designações para estilingue em atlas lingüísticos brasileiros: perspectivas diatópica e sócio-histórica. In.: TROTTER, David (Ed.) **Actes du XXIV^e Congrès International de Linguistique et de Philologie Romanes.** Tome I. Tübingen: Max Niemeyer Verlag, 2007. p. 533-546.

_____. O caminho do rio, o caminho do homem, o caminho das palavras. In.: RIBEIRO, S. S. C.; COSTA, S. B. B.; CARDOSO, S. A. M. (orgs.) **Dos sons às palavras: nas trilhas da língua portuguesa.** Salvador: EDUFBA, 2009, p. 41-59.

JACINO, R. **O negro no mercado de trabalho em São Paulo pós-abolição 1912/1920.** 2012, 204 p. Tese (Doutorado em Ciência Econômicas). Universidade de São Paulo, São Paulo, 2012.

JOB, V. R. Algumas considerações sobre o Ciclo do Ouro e o Tropeirismo. In.: BONADIO, G. (org.). **O Tropeirismo e a formação do Brasil.** Sorocaba: Academia Sorocabana de Letras; Fundação Ubaldino do Amaral; Skol; Momesso; Caracu, 1984, p. 9-15.

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

KOCH, W.; ALTENHOFEN, C. V.; KLASSMANN, M. S. (orgs.) **Atlas Linguístico-Etnográfico da Região Sul do Brasil – ALERS: Cartas Fonéticas e Morfossintáticas.** 2. Ed. Porto Alegre: Editora da UFRGS; Florianópolis: Ed. UFSC, 2011.

KRUSKAL, W. H.; WALLIS, W. A. Use of Ranks in One-Criterion Variance Analysis. **Journal of the American Statistical Association**, v. 47, n. 260, dez. 1952, p. 583-621.

LACERDA, J. F. de. **Estudos sobre a crise e o preço do café e a sua solução.** São Paulo: Thyphographia do Diário Oficial, 1901.

LAMOUNIER, M. L. Entre a escravidão e o trabalho livre. Escravos e imigrantes nas obras de construção das ferrovias no Brasil no século XIX. ANPEC 2008 - ENCONTRO NACIONAL DE ECONOMIA. **Anais...** Disponível em: <<http://www.anpec.org.br/encontro2008/artigos/20080721117010-.pdf>>. Acesso: 06 jul. 2013.

LAYTANO, D. de. **Vocabulário do nordeste do Rio Grande do Sul: linguagem dos praeiros.** Almanaque Globo. Porto Alegre: Globo, 1933.

LEITE DE VASCONCELOS, J. **Esquisse d'une dialectologie portugaise**. 2. ed. Lisboa: Centro de Estudos Filológicos, 1970.

LESSA, J. **Juiz de Fora e seus pioneiros: do caminho novo à proclamação**. Universidade Federal de Juiz de Fora e Fundação Cultural Alfredo Ferreira Lage, 1985.

LILIEFORS, H. W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. **Journal of the American Statistical Association**, v. 62, n. 318, jun. 1967, p. 399-402.

LOBO, T. C. F. **Variantes nacionais do português: sobre a questão da definição do português do Brasil**. 2013. Disponível em: <http://www.prohpor.ufba.br/variantes.html#_ftn1>. Acesso em: 23 ago. 2013.

LUCCHESI, D. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. **Revista internacional de Língua portuguesa**, Lisboa, AULP, n. 12, 1994, p. 17-28.

_____. As duas grandes vertentes da história sociolinguística do Brasil (1500-2000). **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, São Paulo, PUC. V. 17, n.1. 2001, p. 97-130. Disponível em: <http://www.scielo.br/scielo.php?pid=S0102-44502001000100005&script=sci_arttext>. Acesso em: 23 ago. 2013.

_____. Norma linguística e realidade social. In: BAGNO, M. (org.). **Linguística da norma**. São Paulo: Loyola, 2002, 63-92.

_____. Parâmetros Sociolinguísticos do Português Brasileiro. **Revista da ABRALIN**, v. 5, n.1. dez. 2006, p.83-112. Disponível em: <<http://www.vivalinguaportuguesa.com.br/admin/images/cientifico/1451040420120.PDF>>. Acesso em: 23 ago. 2013.

_____. História do contato entre línguas no Brasil. In.: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (orgs.) **O português afro-brasileiro**. Salvador: UFBA, 2009, p. 41-71.

_____. **A polarização sociolinguística do Brasil**. 2013. Disponível em: <<http://www.vertentes.ufba.br/projeto/tema>>. Acesso 23 ago. 2013.

MACHADO, J. P. **Dicionário etimológico da língua portuguesa**. Lisboa, 1952.
MAGALHÃES, M. N.; LIMA, A. C. P. **Noções de Probabilidade e Estatística**. 3. ed., São Paulo: IME-USP, 2001.

MANDEL, E. **O significado da segunda guerra mundial**. São Paulo: Ática, 1989.

MARGOTTI, F. W. **Difusão Sócio-geográfica do português em contato com o italiano no sul do Brasil**. 2004. 332 p. Tese de Doutorado (Programa de pós-graduação em Letras) Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.

MARGOTTI, F. W. Geolinguística pluridimensional: desafios metodológicos. VIII ENCONTRO DO CELSUL. **Anais...** 2008. Disponível em: <http://celsul.org.br/Encontros/08/geolinguistica_pluridimensional.pdf>. Acesso em: 02 set. 2013.

MARGOTTI, F. W.; VIEIRA, H. Características de uma área lexical heterogênea na Região Sul do Brasil. In.: VANDRESEN, P. (org.). **Variação, mudança e contato linguístico no português da Região Sul**. Pelotas: EDUCAT, 2006, p.245-260.

MARROQUIM, M. **A Língua do Nordeste**. 3. ed. Curitiba: HD Livros, 1996 [1934].

MARTINS, D.; VANALLI, S. **Migrantes**. 6. ed. São Paulo: Contexto, 2004.

MATTOS e SILVA, R. V. **Ensaio para uma sócio-história do português brasileiro**. São Paulo: Parábola Editorial, 2004.

_____. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo: Parábola Editorial, 2008.

MATTOS, M. Fases de prosperidade e de declínio do Tropeirismo. In.: BONADIO, G. (org.). **O Tropeirismo e a formação do Brasil**. Sorocaba: Academia Sorocabana de Letras; Fundação Ubaldino do Amaral; Skol; Momesso; Caracu, 1984, p. 17-24.

MAUCH, C.; VASCONCELOS, N.(Orgs.). **Os alemães no sul do Brasil: cultura, etnicidade e história**. Canoas: Ed. Ulbra, 1994.

MELO, G. C. de. **A língua do Brasil**. 4. ed. Rio de Janeiro: Padrão Editora, 1981 [1946].

MENDONÇA, R. **A influência africana no português do Brasil**. São Paulo: Nacional, 1935.

_____. **O português do Brasil**. São Paulo: Nacional, 1937.

MENEZES, M. L. P. Tendências atuais das migrações internas no Brasil. **Scripta Nova - Revista Electrónica de Geografía y Ciencias Sociales**. Universidad de Barcelona, n. 69 (45), 1 ago. 2000. Disponível em: <<http://www.ub.edu/geocrit/sn-69-45.htm>>. Acesso em: 23 jul. 2013.

MILLIET, S. **O roteiro do café e outros ensaios**. 4. ed. São Paulo: HUCITEC; Instituto Nacional do Livro, 1982.

MIRANDA, V. C. de. **Glossário Paraense** ou Coleção de vocábulos peculiares à Amazônia e especialmente à ilha do Marajó. Pará: Liv. Maranhense, 1968 [1905].

MONTEIRO, C. **A Linguagem dos Cantadores**. Rio de Janeiro: Borsoi, 1933.

MONTES GIRALDO, J. J. **Dialectología general e hispanoamericana: orientación Teórica, metodológica y bibliográfica**. 2.ed. Bogotá: Instituto Caro y Cuervo, 1987.

- MORAES DE CASTILHO, C. M. **O processo de redobramento sintático no português medieval**: formação das perífrases com *estar*. 2005, 272 p. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, 2005.
- 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.
- MOTA, J. A.; CARDOSO, S. A. M. Sobre a Dialetoлогия no Brasil. In.: MOTA, J. A.; CARDOSO, S. A. M. **Documentos 2**: Projeto Atlas Linguístico do Brasil. Salvador: Quarteto, 2006, p. 15-34.
- MUSSA, A. **O papel das línguas africanas na história do português do Brasil**. Rio de Janeiro: UFRJ, (Dissertação de mestrado), 1991, mimeo.
- NARO, A. J. A study on the origins of pidginization. **Language**, LSA, v. 54, n.2, 1978, pp. 314-347.
- NARO, A. J.; SCHERRE, M. M. P. (orgs.). **Origens do português brasileiro**. São Paulo: Parábola Editorial, 2007.
- NASCENTES, A. **O Linguajar Carioca**. 2. ed. Rio de Janeiro: Organizações Simões, 1953 [1922].
- _____. **O idioma nacional**. 3. ed. Rio de Janeiro: Livraria Acadêmica, 1960.
- _____. **Bases para elaboração do Atlas Linguístico do Brasil**. Rio de Janeiro: Ministério da Educação e Cultura, Casa Rui Barbosa, 1958.
- NEIVA, A. **Estudos de língua nacional**. Rio de Janeiro: Companhia Ed. Nacional, 1940. Disponível em: <<http://www.brasiliana.com.br/obras/estudos-da-lingua-nacional/preambulo/8/texto>>. Acesso em 01 set. 2013.
- NOLL, V. **O português brasileiro**: formação e contrastes. Trad. Mário Eduardo Viaro. São Paulo: Globo, 2008.
- OLIVEIRA, A. J de. Análise quantitativa no estudo da variação linguística: noções de estatística e análise comparativa entre Varbrul e SPSS. **Revista de Estudos da Linguagem**, Belo Horizonte, UFMG. v.17, n.2, 2009, p-93-119.
- OLIVEIRA, A. T. R. de; ERVATTI, L. R.; O'NEILL, M. M. V. C. O panorama dos deslocamentos populacionais no Brasil: PNADs e Censos Demográficos. In.: OLIVEIRA, L. A. P. de; OLIVEIRA, A. T. R. de. **Reflexões sobre os deslocamentos populacionais no Brasil**. n. 1 Estudos e Análises, informação demográfica e Socioeconômica. Rio de Janeiro: IBGE, 2011. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/reflexoes_deslocamentos/deslocamentos.pdf>. Acesso em: 24 jul. 2013.

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

ORLANDI, E. P.. A língua brasileira. **Ciencia e Cultura**. [online]. 2005, vol.57, n.2, pp. 29-30. Disponível em: <<http://cienciaecultura.bvs.br/pdf/cic/v57n2/a16v57n2.pdf>>. Acesso em: 26 ago. 2013.

PAIM, M. M. T. A variação lexical nos campos semânticos corpo humano e ciclos da vida: o que revelam os dados do Projeto ALiB. **Diadorim: Revista de Estudos Linguísticos e Literários**, n. 8. Rio de Janeiro: UFRJ, 2011, p. 143-159.

PAIXÃO CÔRTEZ, J. C. **Danças Birivas do Tropeirismo Gaúcho**. Porto Alegre: CORAG, 2000.

PARKVALL, M.; ÁLVARES LÓPEZ, L. Português vernáculo brasileiro e a hipótese da semi-crioulização. **Revista da ABRALIN**. v. 2, n.1, jun. 2003, p.111-152. Disponível em: <http://www.abralin.org/revista/RV2N1/debates/RV2N1_debates.pdf>. Acesso em: 08 ago. 2013.

PAULI, E.. Italianos em Santa Catarina In.: **Enciclopédia Simpozio: Discurso catarinense**. n.10. 1997. Disponível em: <http://www.cfh.ufsc.br/~simpozio/Catarinense/Discurso_catarinense_texto_mega/98sc00010.html>. Acesso em 02 set. 2014.

PENHA, J. A. P. **A teoria do português do Brasil como modalidade arcaizantes**. Tese (Doutorado em Língua portuguesa). Araraquara: Faculdade de Filosofia Ciências e Letras, 2 v. 1970.

PENHA, J. A. P. **Português rural de Minas numa visão tridimensional: na fala, nos textos regionais e nos textos antigos**. Franca: Universidade Estadual Paulista, 1997.

PESTANA, P. R. **O Café em São Paulo: notas históricas**. São Paulo: Secretaria da Agricultura, Indústria e Comércio do Estado de São Paulo, 1927.

PETRI, K. C. Terras de imigração em São Paulo: política fundiária e trabalho rural. **Histórica – Revista Eletrônica do Arquivo Público do Estado de São Paulo**. N. 2. Ano 1. Jun, 2005. Disponível em<<http://www.historica.arquivoEstado.sp.gov.br/materias/anteriores/edicao02/materia01/Terralmigracao.pdf>>. Acesso em: 10 jul. 2013.

PINTO, E. P. **A gramatiquinha de Mário de Andrade: texto e contexto**. São Paulo: Livraria Duas Cidades, 1991.

PINTO, L. M. da S. **Dicionário da Língua Brasileira**. Ouro Preto: Typographia da Silva, 1832. Disponível em: <<http://www.brasiliana.usp.br/pt-br/dicionario/edicao/3>> Acesso em: 22 de jun. 2013.

PIO, R. M. **Classificações das tangerinas**. Folder. São Paulo: CEAGESP, 2000. Disponível em: <www.ceagesp.gov.br/produtor/classific/fc_tangerina>. Acesso

em 17 out. 2014.

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

RADTKE, E.; THUN, H. (eds.). **Neue Wege der romanischen Geolinguistik**: Akten des Symposiums zur empirischen Dialektologie. Kiel : Westensee - Verl, 1996.

RAIMUNDO, J. **O elemento afro-negro na língua portuguesa**. Rio de Janeiro, Renascença, 1933.

RAZKY, A. A dimensão sociodialetal do léxico no Projeto Atlas Linguístico do Brasil. **Signum**: Estudos da Linguagem, Londrina, UEL, v. 13, n.2, 2013, p.247-270.

REAL ACADEMIA ESPAÑOLA. **Diccionario de la lengua española**. 22 ed. Madrid: Real Academia Española, 2001. Disponível em: <
<http://lema.rae.es/drae/?val=bolita>> Acesso em: 28 jan. 2014.

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. 2. v. 2006. (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, J. et al. **Esboço de um atlas linguístico de Minas Gerais**. Rio de Janeiro: Casa de Rui Barbosa, 1977.

RIBEIRO, S. S. C. **Brinquedos e brincadeiras infantis na área do falar baiano**. 2012, 466 p. Tese de Doutorado (Programa de Pós-graduação em Letras e Linguística). Universidade Federal de Bahia, Salvador, 2012.

RICARDO, C. **Marcha para Oeste**. 4. ed. Rio de Janeiro: Editôra da Universidade São Paulo; Livraria José Olympio Editôra, 1970.

ROCHA, P. G. da. **O português de contato com o espanhol no sul do Brasil**: empréstimos lexicais. 2008. 148 p. Dissertação (Mestrado em Linguística). Programa de Pós-graduação em Linguística do Centro de Comunicação e Expressão. 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.

RODRIGUES, A. D. As línguas gerais sul-americanas. **Papia**: Revista Brasileira de estudos crioulos e Similares. V. 4, n. 2, 1996, p.6-18. Disponível em: <
<http://www.abecs.net/ojs/index.php/papia/article/view/143/210>>. Acesso em: 07 ago. 2013.

_____. Panorama das línguas indígenas da Amazônia. In.: QUEIXALÓS, F; RENALT-LESCURE, O.(orgs.). **As línguas amazônicas hoje**. São Paulo: IDR;ISA; MPEG, 2000, p. 15-28.

ROMANO, V. P. Percurso historiográfico e metodológico da Geolinguística. **Papéis**, Campo Grande, UFMS, 2014 (no prelo).

ROMANO, V. P. **Atlas Geossociolinguístico de Londrina**: um estudo em tempo real e tempo aparente. 2012. 2v. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2012. Volume 1: 206 f. Volume 2: três cartas introdutórias e 71 cartas linguísticas.

_____. Balanço crítico da Geolinguística brasileira e a proposição de uma divisão. **Entretextos**. Londrina, UEL, v. 13, n.2, jul.- dez., 2013, p. 203-242.

ROMANO, V. P.; AGUILERA, V. de A. Um estudo lexicológico sobre as variantes lexicais para a tangerina. IV ERIC - Encontro Regional de Iniciação Científica, 2008, Mandaguari - PR. **Anais...** Mandaguari: FAFIMAN, 2008.

_____ A. Distribuição diatópica para as variantes para tangerina: um estudo geossociolinguístico: In: ISQUERDO, A. N.; ALTINO, F. C.; AGUILERA, V. de A. (orgs.) **Atlas linguístico do Brasil**: descrevendo a língua, formando jovens pesquisadores. CD-ROM, vol. 1. Londrina: UEL, 2009, p. 148-157.

_____. Padrões de variação lexical no sul do Brasil a partir dos dados do Projeto ALiB. **Estudos Linguísticos**. São Paulo, v. 43, 2014, (no prelo).

ROMANO, V. P.; ISQUERDO, A. N. Designações para 'bolinha de gude' nas capitais brasileiras. In.: CARDOSO, S. A. M et al. **Atlas Linguístico do Brasil** -Vol. 3 (Estudos). Londrina, EDUEL, 2015 (no prelo).

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, 2014a, p. 463-497.

_____. Dados geolinguísticos sob uma perspectiva estatística: a variação lexical no Centro-Oeste, Sudeste e Sul do Brasil. **RELin Revista de Estudos da Linguagem**. Belo Horizonte, UFMG, v. 22, n.2, 2014b, p. 59-92.

ROMANO, V. P.; SEABRA, R. D.; OLIVEIRA, N. [ΣGVCLin] – Software para geração e visualização de cartas linguísticas. **RELin Revista de Estudos da Linguagem**. Belo Horizonte, UFMG, v. 22, n.1, 2014, p. 119-151.

ROSSI, N. et al. **Atlas Prévio dos Falares Baianos**. Rio de Janeiro: INL, 1963.

RUBIM, B. da C. **Vocabulário brasileiro para servir de complemento aos dicionários da língua portuguesa**. Rio de Janeiro: Typ. Dous de dezembro de Paula Brito; Imprensa da Casa Imperial, 1853. Disponível em: <http://www.brasiliana.usp.br/bbd/bitstream/handle/1918/01634100/016341_COMPL ETO.pdf>. Acesso em 02 set. 2013.

SAES, F. A. M. de. **As ferrovias de São Paulo (1870-1940)**. São Paulo: HUCITEC, 1981.

SAINT-HILAIRE, A. de. **Viagens pelas províncias de Rio de Janeiro e Minas Gerais**. Tomo I. Trad.: Clado Ribeiro de Lessa. São Paulo: Companhia Editora Nacional, 1938. Disponível em: <<http://www.brasiliana.com.br/obras/viagem-pelas-provincias-do-rio-de-janeiro-e-minas-gerais-t-1/pagina/7/texto>> Acesso em: 07 jun 2013.

SAMPAIO, T. **O Tupi na Geographia Nacional**. Memoria lida no Instituto Historico e Geographico de S. Paulo. São Paulo: Typ. da Casa Eclectica, 1901. <<http://biblio.etnolinguistica.org/sampaio-1901-tupi/p/1>>. Acesso em 22 ago. 2013.

SANTOS, R. B. **Migrações no Brasil**. São Paulo: Scipione, 1994.

SAPIR, E. Língua e Ambiente. **Linguística como Ciência: Ensaios**. Trad. J. Mattoso Câmara Jr. Rio de Janeiro: Livraria Acadêmica, 1969, p. 43-62.

SEABRA, R. D. ; ROMANO, V. P.; OLIVEIRA, N. . A Brazilian Contribution for Automated Linguistic Cartography. **American Journal of Linguistics**, v. 3, p. 27-40, 2014.

SILVA, L. M. da; AGUILERA, V. de A. Variantes lexicais para o gambá: um estudo geolinguístico nas capitais brasileiras. 1º CIELLI – COLÓQUIO INTERNACIONAL DE ESTUDOS LINGUÍSTICOS E LITERÁRIOS. **Anais...** Maringá: Universidade Estadual de Maringá, 2010. p. 1-13.

SILVA NETO, S. da. **Guia para estudos dialectológicos**. 2. ed. Belém: Instituto Nacional de Pescis da Amazônia, 1957.

_____. **Introdução ao estudo da língua portuguesa no Brasil**. 2.ed. Rio de Janeiro: INL, 1963 [1951].

_____. **História da Língua Portuguesa**. 3. ed. Rio de Janeiro: Presença; Brasília: INL, 1979 [1952].

SILVA, J. J. P da. **O idioma do hodierno Portugal comparado com o do Brasil por um brasileiro**. Rio de Janeiro: Lourenço Winte, 1879.

SIMIELLI, M. E. **Geoatlas**. 32. ed. São Paulo: Ática, 2006.

SIMONSEN, R. **História econômica do Brasil (1500-1820)**. Tomo I. São Paulo: Companhia Editora Nacional, 1937. Disponível em: <<http://www.brasiliana.com.br/obras/historia-economica-do-brasil-1500-1820>> Acesso em: 01 jun. 2013.

_____. **Evolução industrial do Brasil e outros estudos**. São Paulo: Companhia Ed. Nacional, Editora da Universidade de São Paulo, 1973. Disponível em: <<http://www.brasiliana.com.br/obras/evolucao-industrial-do-brasil-e-outros-ensaios/pagina/163/>>. Acesso em: 07 jul. 2013.

SMITH, H. H. Uma Fazenda de café no tempo do império. **Separata de Revista DNC**, ago, 1940. Rio de Janeiro: D.N.C., 1941.

SOARES, A. J. de M. **Diccionario brasileiro de língua portugueza**. Rio de Janeiro: Typ. De G. Leuzinger & Filhos, 1888.

SOBRINHO, A. M. **A civilização do café (1820-1920)**. São Paulo: Brasiliense, 1967.
SOUTHEY, R. **História do Brasil**. Tomo 5. Trad.: Luiz Joaquim de Oliveira e Castro. Rio de Janeiro: Livraria de B.L. Garnier, 1862. Disponível em: <<http://www.brasiliana.usp.br/node/411>> Acesso: 09 jun. 2013.

SPSS for Windows. Version 17.0. [S.l.]: IBM, 2008. 1 CD-ROM.

STECA, L. C.; FLORES, M. D. **História do Paraná: do século XVI à década de 1950**. Londrina: Editora UEL, 2002.

STEHLING, L. J. **Juiz de Fora, a companhia união e indústria e os alemães**. Edição da Prefeitura de Juiz de Fora, FUNALFA, 1979.

TARALLO, F. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In.:ROBERTS, I.; KATO, M. (orgs.). **Português brasileiro: uma viagem diacrônica, homenagem a Fernando Tarallo**. 2. Ed. Campinas: Editora da UNICAMP, 1996, p. 35-68.

TAUNAY, A. de E. **História geral das bandeiras paulistas**. 11 volumes. São Paulo: Typ. Ideal, 1924-1950.

_____. **História do café no Brasil: no Brasil Imperial 1822-1872**. Rio de Janeiro, Departamento Nacional do Café, 1939. v.3.

TEIXEIRA, J. A. O falar mineiro. **Separata da Revista do Arquivo Municipal de São Paulo**, São Paulo: Prefeitura Municipal de São Paulo. Departamento de Cultura, 1938.

_____. **Estudos de dialectologia portuguesa: linguagem de Goiás**. São Paulo: Edit. Anchieta, 1944.

TELES, A. R. T. F.; RIBEIRO, S. S. C. Apresentando a cartografia aos linguistas: o Projeto ALiB. In: MOTA, J. A.; CARDOSO, S. A. M. (Orgs.). **Documentos 2**. Projeto Atlas Lingüístico do Brasil. Salvador: Quarteto, 2006, p. 207-226.

TEYSSIER, P. **História da língua portuguesa**. Trad. Celso Cunha. 6. Ed. Lisboa: Livraria Sá da Costa, 1994.

THUN, H. La géographie linguistique romane à la fin du XX siècle. In.: RAENDONCK, D. V. et all. (Orgs). **Actes du XXII Congrès International de Linguistique e Philologie Romanes**. Bruxelles, 1998, p. 367-409.

VALLONE, G. Crise de 1929 atingiu a economia e mudou a ordem política no Brasil. **Folha de São Paulo**, 24 out. 2009. Folha Online. Disponível em: <http://www1.folha.uol.com.br/folha/dinheiro/ult91u642391.shtml>. Acesso em: 23 jul. 2013.

VENY, J. **Introducció a la dialectologia catalana**. Barcelona: Enciclopèdia Catalana, 1986.

VIANNA, H. **História do Brasil**. 7. ed. São Paulo: Melhoramentos, 1970.

VIEIRA, H. G. Traços de bilinguismo no léxico catarinense: um estudo pluridimensional. 6º ENCONTRO DO CELSUL. **Anais...** Florianópolis: UFSC, 2004. Disponível em: <http://celsul.org.br/Encontros/06/Individuais/175.pdf>. Acesso: 05 fev. 2014.

VIEIRA, R. Feira de mueres de Sorocaba. In.: BONADIO, G. (org.). **O Tropeirismo e a formação do Brasil**. Sorocaba: Academia Sorocabana de Letras; Fundação Ubaldino do Amaral; Skol; Momesso; Caracu, 1984, p. 31-37.

WILLEMS, E. **A aculturação dos alemães no Brasil**: estudo antropológico dos imigrantes alemães e seus descendentes no Brasil. 2 ed. São Paulo: Ed. Nacional, 1980.

YIDA, V. **O campo semântico da alimentação e cozinha no Atlas Linguístico do Brasil**: um estudo lexical nas capitais. 2011, 191 p. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2011.

ZAMBERLAM, J. **O processo migratório no Brasil e os desafios da mobilidade humana na globalização**. Porto Alegre: Palloti, 2004.

ANEXOS

ANEXO A
(Decreto de 1952)

Decreto nº 30.643, de 20 de Março de 1952

Institui o Centro de Pesquisas da Casa de Rui Barbosa e dispõe sobre o seu funcionamento.

O PRESIDENTE DE REPÚBLICA, usando da atribuição que lhe confere o artigo 87, item I, da Constituição,
DECRETA.

Art. 1º Fica instituído, na Casa de Rui Barbosa, o Centro de Pesquisar da Casa de Rui Barbosa.

Art. 2º O Centro ora instituído realizara seus estudos e trabalhos no domínio do direito e da filosofia, e terá por campo de pesquisar a biblioteca e os arquivos da aludida Casa de Rui Barbosa e novas aquisições necessárias ao desenvolvimento dos seus trabalho.

Art. 3º O Centro em referencia compreenderá, inicialmente, duas Secções: a de Direito e a de Filologia, dirigidas cada qual por uma Comissão de especialistas convidados pelo Ministro da Educação e Saúde, mediante parecer do Diretor da Casa de Rui Barbosa.

§1º Cada Comissão de que trata este artigo estabelecerá anualmente, um plano de trabalho em cuja elaboração poderão colaborar os professôres universitários do Brasil, providos em cátedras de direito e de filologia ou em cadeiras afins.

§2º A Comissão de Direito planejará publicações de bibliografia jurídica, de jurisprudência e de história do direito, organizando catálogos de publicações jurídicas, legislativas parlamentares e jurisprudência do Brasil; boletins de bibliografia brasileira e estrangeira, estudos sistemáticos de bibliografia e hemerografia de jurisprudência federal e das unidades da Federação.

§3º A Comissão de Filologia promoverá pesquisar em todo o vasto campo de filologia portuguesa-fonologia, morfológicas, sintáticas léxicas, etimológicas, métricas, onomatológicas, dialetológicas bibliograficas, históricas literárias, problemas de texto, de fontes, de autoria, de influências, sendo sua finalidade principal a elaboração do "Atlas Lingüístico do Brasil".

Art. 4º O Centro de Pesquisar da Casa de Rui Barbosa promoverá a publicação de seus trabalhos em arquivos e boletins periódicos, visando especialmente aos resultados das pesquisas que terão caráter estritamente objetivo.

Art. 5º Além da Biblioteca da Casa de Rui Barbosa e das aquisições imprescindíveis para atualiza-la, o Centro em causa disporá de instalações apropriadas, cujo projeto será oportunamente apresentado.

Art. 6º O Ministério da Educação e Saúde contratara os técnicos em bibliografia, arqueologia e revisão necessários aos trabalhos em cursos.

Art. 7º Os casos omissos neste Decreto serão resolvidos por meio de instruções do Ministro da Educação e Saúde.

Art. 8º Êste Decreto entrara em vigor na data de sua publicação.

Rio de Janeiro, em 20 de março de 1952, 131º da Independência e 64º da Republica.

GETÚLIO VARGAS

E. Simões Filho

Publicação:

- Diário Oficial da União - Seção 1 - 22/03/1952 , Página 4665 (Publicação Original)
- Coleção de Leis do Brasil - 1952 , Página 170 Vol. 2 (Publicação Original)

Disponível em: <<http://www2.camara.gov.br/legin/fed/decret/1950-1959/decreto-30643-20-marco-1952-339719-publicacaooriginal-1-pe.html>> Acesso: 04/12/13.

ANEXO B

(Autorização do Comitê Nacional para uso do *Corpus* do Projeto Alib)

DECLARAÇÃO

Ao utilizar como referencial empírico do trabalho de Pós-Graduação, intitulado *Os falares sulistas a partir dos dados do Projeto Atlas Linguístico do Brasil* que desenvolvo sob a orientação da Dr.^a Vanderci de Andrade Aguilera, Diretora Científica do Projeto Atlas Linguístico do Brasil (Projeto ALiB), dados do *corpus* desse Projeto, declaro:

1. Estar ciente de que os materiais do Banco de Dados do Projeto ALiB a mim facultados não podem ser repassados, enquanto conjunto de dados, a outro(s) pesquisador(es) e/ou interessado(s) na matéria.
2. Ter pleno conhecimento de que a divulgação parcial ou final do trabalho deve ser sempre acompanhada da indicação da fonte (Banco de Dados do Projeto ALiB) e da citação do nome do orientador.
3. Autorizar que os resultados da análise por mim efetuada sejam utilizados nas publicações do Atlas Linguístico do Brasil, em quaisquer dos volumes que venham a integrar a coleção, mediante a indicação da fonte e a citação do meu nome.
4. Oferecer a minha contrapartida ao Atlas Linguístico do Brasil colaborando, se requerido, na transcrição de dados, catalogação e cópia de materiais e em outras atividades que não impliquem a pesquisa de campo.

E por estar de acordo, firmo a presente DECLARAÇÃO que tem, também, o CIENTE da Orientadora.

Londrina, 02 de maio de 2013.



Valter Pereira Romano

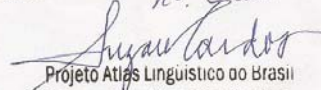
RG 41048894-X (SSP-SP)

CPF 351916288-10

REGISTRADO no Projeto ALiB sob n° 02980

CIENTE:


Orientadora Prof.^a Vanderci de Andrade Aguilera
Diretora Científica


Registrado no
Projeto ALiB sob
n° 020
Projeto Atlas Linguístico do Brasil
Suzana Alice Marcelino Cardoso
Diretora-Presidente