Universidade Estadual de Londrina

ANDRÉ COSTA SANTOS

UM ESTUDO PLURIDIMENSIONAL EM ATLAS URBANO E RURAIS:

ASPECTOS DE VARIAÇÃO NA BAHIA E NO PARANÁ

Londrina
2025

ANDRÉ COSTA SANTOS

UM ESTUDO PLURIDIMENSIONAL EM ATLAS URBANO E RURAIS:

ASPECTOS DE VARIAÇÃO NA BAHIA E NO PARANÁ

Trabalho de Conclusão de Curso apresentado à Universidade Estadual de Londrina - UEL, como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem.

Orientador: Prof.ª. Dr.ª Vanderci de Andrade Aguilera.

Londrina
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Santos, André Costa .
Um estudo pluridimensional em Atlas Urbano e Rurais: : aspectos de variação lexical na Bahia e no Paraná / André Costa Santos. - Londrina, 2025.
140 f. : il.

Orientador: Vanderci de Andrade Aguilera .
Dissertação (Mestrado em Estudos da Linguagem) - Universidade Estadual de Londrina, Centro de Letras e Ciências Humanas, Programa de Pós-Graduação em Estudos da Linguagem, 2025.
Inclui bibliografia.

1. Dialetologia - Tese. 2. Variantes Agropastoris - Tese. 3. ALiB - Tese. 4. ALPR e APFB - Tese. 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 8

ANDRÉ COSTA SANTOS

UM ESTUDO PLURIDIMENSIONAL EM ATLAS URBANO E RURAIS:

ASPECTOS DE VARIAÇÃO LEXICAL NA BAHIA E NO PARANÁ

Trabalho de Conclusão de Curso apresentado à Universidade Estadual de Londrina - UEL, como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem

BANCA EXAMINADORA

_____

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

_____

Prof.ª Dr.ª Fabiane Cristina Altino
Universidade Estadual de Londrina - UEL

_____

Prof. Dr. Valter Pereira Romano
Universidade Federal de Santa Catarina - UFSC

Londrina, 28 de agosto de 2025.

AGRADECIMENTOS

A Deus, por me abençoar com sabedoria e resiliência para alcançar meus objetivos com coragem, persistência e determinação. Foi pela fé, meu Onipotente e Onisciente, que suplantei o medo, a insegurança e a fraqueza. De joelhos dobrados, supliquei-lhe o discernimento necessário para obter a vitória neste mestrado, e o Senhor, ser de bondade e amor, atendeu plenamente aos meus clamores. Obrigado, Pai Eterno, por estar sempre presente na minha vida e por me tornar um cristão mais assíduo em seus desígnios.

À minha mãe, Maria Helena Costa Santos, fonte de sabedoria e ternura, por ter me ensinado o caminho do bem, da bondade e da honestidade. O legado deixado pela senhora, que partiu muito cedo desta vida para servir ao Pai, está presente em meu coração e nas minhas atitudes. Muito obrigado, minha amada e estimada mãe, pelos ensinamentos que me fizeram um filho melhor, um pai mais presente, um esposo mais carinhoso. Gratidão.

A meu pai, Dalvo Silva dos Santos, homem apaixonado pela terra, por ter me ensinado a importância de cuidar do meio ambiente, das plantas e dos animais com a sabedoria do homem simples do campo, que acorda cedo para produzir nossos alimentos. Obrigado, meu amado pai, por cultivar, em nossa família, esses valores.

A meus irmãos, Marcelo, Leonardo e Yonara, por me incentivarem, aconselharem e apoiarem nesta trajetória do mestrado e da vida, sempre almejando o meu crescimento intelectual e profissional. Obrigado, meus amados irmãos. Esta conquista não é apenas minha, mas de vocês também. Juntos vencemos com amor, respeito e carinho.

À minha esposa, Carla Schneider, mulher guerreira, destemida e otimista, por ter estado ao meu lado nos momentos de tranquilidade e de aflição, ajudando-me a superar, com sabedoria e fé, as adversidades da vida, promovendo o meu crescimento familiar e profissional. Obrigado por você fazer parte da minha vida, minha amada.

Ao meu filho, Rafael Schneider Costa, por ser a minha maior fonte de inspiração e de superação, pois todas as batalhas por mim travadas e vencidas tiveram como propósito o seu bem-estar, visando sempre a um futuro melhor para você, com muita dignidade, prosperidade e presença divina. Obrigado por você ser o meu maior presente, meu amado filho.

À professora Vanderci de Andrade Aguilera, por ter me orientado com muita competência, sabedoria e entusiasmo em suas aulas de Dialetologia, Pesquisa Sociolinguística e na dissertação do mestrado. Devo-lhe todo o meu respeito, gratidão e carinho. Agradeço a Deus por tê-la posto no meu caminho, pois, sem os seus ensinamentos, eu jamais teria chegado tão longe. Durante esse percurso, que trilhamos juntos, a senhora soube me ouvir e me aconselhar, nos momentos de dificuldades e aflições pelos quais passei e, com discernimento e generosidade, ministrou-me os conselhos certos para eu superar os problemas pessoais sem, entretanto, perder o foco da pesquisa. Obrigado, professora, por me transformar hoje em um educador e em uma pessoa melhor do que fui ontem.

Aos professores Valter Romano e Fabiane Altino, por me acolherem prontamente no Congresso da Abralin, no Sedata, no estágio, na qualificação e na defesa. O atendimento às minhas solicitações, as correções textuais e as sugestões foram essenciais para o exitoso desfecho desta pesquisa científica. Obrigado, professores, por me transformarem em um pesquisador mais preparado e apaixonado por Dialetologia.

Ao professor Marcelo Silveira, por ter me auxiliado incessantemente na conclusão deste trabalho. Sua calma, bom humor e competência vieram na dosagem certa, contribuindo significativamente para a melhoria desta dissertação. Obrigado, ilustre professor Marcelo.

Ao Conselho Nacional de Desenvolvimento (CNPq), por financiar minha pesquisa por meio da bolsa de estudos, viabilizando o desenvolvimento deste estudo e minha formação acadêmica.

À minha sogra, Edina, por ter sido atenciosa e pacienciosa comigo durante esse ciclo de desafios, superações e vitórias. À minha tia Edna, por me incentivar e me apoiar sempre para eu cursar o mestrado e aprimorar minha formação. Obrigado, minhas estimadas, pelo incentivo e pela confiança na minha capacidade.

Aos meus diletos amigos João, Willian e Sócrates, por terem me auxiliado ininterruptamente nos trabalhos do mestrado, socorrendo-me inclusive em fins de semana, domingos e feriados, com apoio técnico e motivação. Obrigado, meus amigos, pela prontidão com que me atenderam.

RESUMO

SANTOS, André Costa. **Um estudo pluridimensional em atlas urbano e rurais:** aspectos de variação lexical na Bahia e no Paraná. 2025. 140 f. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2025.

Este trabalho apresenta um estudo diacrônico, diatópico e diastrático das variantes lexicais para *sabugo, soca, vagem* e *cangalha*, coletadas em quatro *corpora* distintos relativos a três sincronias: (i) *Atlas Prévio dos Falares Baianos* – APFB – (Rossi, 1963); (ii) *Atlas Linguístico do Paraná* – ALPR (Aguilera, 1994) e volume II (Altino, 2007); e (iii) dados inéditos do *Projeto Atlas Linguístico do Brasil* (ALiB), coletados nos estados da Bahia e do Paraná, no início da década de 2010. Os objetivos desta dissertação consistem em: (i) colaborar com a realização do ALiB no que tange ao levantamento e descrição dos dados para fins de cartografação; (ii) comparar os dados baianos e paranaenses dos respectivos atlas estaduais com os obtidos pelo ALiB nesses estados; (iii) verificar a correlação rural × urbana na elicitação das respostas; e (iv) avaliar a influência de outras variáveis, como sexo e faixa etária, sobre os acervos lexicais constituídos. O estudo tem como base os pressupostos teóricos da Dialetologia Tradicional ou Monodimensional (Trudgill; Campoy, 2007), Bidimensional (Cardoso, 2010) e Pluridimensional (Thun, 1998). De modo geral, a variável diatópica foi a mais produtiva, com reflexos diretos nos dados pesquisados, e a variável faixa etária foi responsável pela manutenção de lexias mais conservadoras e tradicionais. Já a variável sexo apresentou pouca relevância para a pesquisa. Os resultados deste estudo contribuem para o entendimento da variação lexical em contextos rural e urbano, na Bahia e no Paraná, revelando maior índice de apagamento e diversidade na fala dos informantes da zona urbana e maior apropriação de termos técnicos e específicos na fala dos informantes da zona rural.

Palavras-chave: Dialetologia; ALiB; ALPR; APFB; variantes agropastoris.

ABSTRACT

SANTOS, André Costa. **A multidimensional study in urban and rural atlases:** aspects of lexical variation in Bahia and Paraná. 2025. 140 f. Dissertation (Master's in Language Studies) – Universidade Estadual de Londrina, Londrina, 2025.

This paper presents a diachronic, diatopic, and diastratic study of the lexical variants for *sabugo*, *soca*, *vagem* and *cangalha*, collected in four distinct *corpora* related to three synchronies: (i) *Atlas Prévio dos Falares Baianos* – APFB – (Rossi, 1963); (ii) *Atlas Linguístico do Paraná* – ALPR (Aguilera, 1994) and volume II (Altino, 2007); and (iii) unpublished data from the *Projeto Atlas Linguístico do Brasil* (ALiB) [Linguistic Atlas of Brazil Project], collected in Bahia and Paraná in the early 2010s. The objectives of this research are to: (i) collaborate with ALiB in terms of collecting and describing the data for mapping purposes; (ii) compare the data from Bahia and Paraná from the respective state atlases with those obtained by ALiB in these states; (iii) verify the rural × urban correlation in the elicitation of responses; and (iv) assess the influence of other variables, such as gender and age group, on the lexical collections constituted. The study is based on the theoretical assumptions of Monodimensional, Bidimensional, and Pluridimensional Dialectology, presented mainly by Trudgill and Campoy (2007), Cardoso (2010), and Thun (1998). Overall, the diatopic variable was the most productive, directly impacting the data studied, while the age group variable was responsible for maintaining more conservative and traditional lexicons. The gender variable, however, showed little relevance to the research. The results of this study contribute to the understanding of lexical variation in rural and urban contexts in Bahia and Paraná, revealing a greater rate of erasure and diversity in the speech of informants from urban areas, and greater appropriation of technical and specific terms in the speech of informants from rural areas.

Keywords: Dialectology; ALiB; ALPR; APFB; agropastoral variants.

LISTA DE FIGURAS

Figura 1 – Carta V do APFB – Pontos linguísticos 46
Figura 2 – Carta IV do ALPR – Pontos linguísticos do ALPR 48
Figura 3 – Rede de pontos do ALiB-BA 50
Figura 4 – Rede de pontos do ALiB-PR 51
Figura 5 – Ilustração para *sabugo*. 58
Figura 6 – Carta 28 do APFB com a distribuição das variantes para sabugo 60
Figura 7 – Carta 28 do APFB com a distribuição das variantes para sabugo 61
Figura 8 – Mapa da Bahia com as mesorregiões 62
Figura 9 – Carta experimental para *sabugo* nos dados do ALiB-BA 66
Figura 10 – Carta 364 do ALPR II (Altino, 2007) com as variantes para *sabugo.* 67
Figura 11 – Ilustração para soca de cana-de-açúcar 72
Figura 12 – Relatório das variantes para *soca* nos dados do ALiB-BA 73
Figura 13 – Carta experimental das variantes de soca no ALiB-BA 74
Figura 14 – Relatório das variantes para soca no ALiB-BA, segundo a variável sexo 76
Figura 15 – Relatório das variantes para soca no ALiB-BA segundo a variável faixa etária 77
Figura 16 – Carta das variantes para soca nos dados do ALiB-BA, segundo as variáveis sexo e faixa etária 78
Figura 17 – Relatório das variantes para soca nos dados do ALPR II 79
Figura 18 – Relatório das variantes para soca nos dados do ALPR II, segundo a variável sexo 80
Figura 19 – Carta para soca no ALPR II, segundo a variável sexo 81
Figura 20 – Carta isoléxica para soca elaborada com os dados do ALPR II 82
Figura 21 – Relatório das variantes para *soca* nos dados do ALiB-PR. 87
Figura 22 – Carta monodimensional das variantes de *soca* nos dados do ALiB-PR 88
Figura 23 – Relatório das variantes para *soca* nos dados do ALiB-PR, segundo a variável sexo 90

Figura 24 – Relatório das variantes para *soca* nos dados do ALiB-PR segundo a variável faixa etária 92
Figura 25 – Carta experimental com a distribuição das variantes para soca no ALiB-PR segundo as variáveis sexo e faixa etária 93
Figura 26 – Ilustração de *vagem* 97
Figura 27 – Relatório com a distribuição das variantes para *vagem* no ALiB-BA segundo a variável diatópica 98
Figura 28 – Carta 83 do *Atlas Linguístico do Paraná* (ALPR) com as variantes de vagem 99
Figura 29 – Relatório com a distribuição das variantes de vagem no ALPR 100
Figura 30 – Relatório com a distribuição das variantes de *vagem* no ALPR segundo a variável sexo 101
Figura 31 – Carta de isoléxicas das variantes de *vagem* no ALPR 102
Figura 32 – Relatório com a distribuição das variantes para *vagem* no ALiB-PR segundo a variável diatópica 104
Figura 33 – Carta experimental das variantes de vagem no ALiB-PR segundo a variável diatópica 105
Figura 34 – Relatório com a distribuição das variantes para *vagem* no ALiB-PR segundo a variável sexo 106
Figura 35 – Relatório com a distribuição das variantes para vagem no ALiB-PR segundo a variável faixa etária 107
Figura 36 – Carta experimental com as variantes de vagem no ALiB-PR segundo as variáveis sexo e faixa etária 108
Figura 37 – Ilustração do referente da *cangalha* 110
Figura 38 – Carta 61 do ALPR - Variantes para o triângulo que se coloca no pescoço de animais (Questão 123 – ALPR) 112
Figura 39 – Carta experimental com as variantes de cangalha nos dados do ALiB-PR segundo as variáveis sexo e faixa etária 116
Figura 40 – Carta experimental com a distribuição das variantes de cangalha registradas no ALiB-PR segundo a variável sexo 119
Figura 41 – Carta experimental com a distribuição das variantes para cangalha registradas pelo ALiB-PR segundo a variável faixa etária 121

LISTA DE GRÁFICOS

Gráfico 1 – Evolução da População Urbana e Rural no Brasil (1940-2022)..... 28
Gráfico 2 – Variantes de cangalha nos dados do *Atlas Linguístico do Paraná* – ALPR..... 111
Gráfico 3 – Distribuição das variantes segundo a variável sexo nos dados do ALPR..... 113
Gráfico 4 – Recorte temporal da população rural x urbana no Paraná (1940 a 2010)..... 124

LISTA DE TABELAS

Tabela 1 – Variantes lexicais para sabugo nos dados do APFB 59
Tabela 2 – Variantes para sabugo na fala dos informantes do ALiB-BA..... 64
Tabela 3 – Distribuição das variantes segundo a variável sexo nos dados do ALiB-BA..... 65
Tabela 4 – Distribuição das variantes, segundo a variável faixa etária, nos dados do ALiB-BA..... 65
Tabela 5 – Ocorrências das variantes para sabugo nos dados do ALPR II 68
Tabela 6 – Distribuição das variantes para sabugo nos dados do ALiB-PR, segundo o perfil dos informantes 69
Tabela 7 – Distribuição das variantes para sabugo nos dados do ALiB-PR, segundo as variáveis sexo e faixa etária..... 69
Tabela 8 – Percentual das variantes de sabugo no APFB, ALiB-BA, ALPR e ALiB-PR 70
Tabela 9 – Percentual das variantes de soca no ALiB-BA, ALPR e ALiB-PR 96
Tabela 10 – Percentual das variantes de vagem no ALiB-BA, ALPR e ALiB-PR 109
Tabela 11 – Percentual das variantes de cangalha no ALPR e ALiB-PR 123
Tabela 12 – Lista com os verbetes dicionarizados..... 129

LISTA DE ABREVIATURAS E SIGLAS

ALiB-BA	*Atlas Linguístico do Brasil* - Bahia
ALiB-PR	*Atlas Linguístico Brasil* - Paraná
ALPR	*Atlas Linguístico do Paraná*
QSL	Questionário Semântico- Lexical

SUMÁRIO

1 INTRODUÇÃO 16

2 UMA ABORDAGEM SOBRE O FALAR URBANO E O RURAL 19

2.1 Um levantamento sobre traços do rural e do urbano no Brasil do descobrimento até o século XX 19

2.2 Um estudo sobre a formação do Brasil sob uma perspectiva étnica, antropológica, histórica e geográfica 21

3 REFERENCIAL TEÓRICO 31

3.1 Dialetologia 31

3.2 Dialetologia tradicional ou monodimensional 32

3.3 Um breve panorama da dialetologia no Brasil 33

3.4 Dialetologia bidimensional 34

3.5 Dialetologia pluridimensional 34

3.6 Geolinguística na Europa e no Brasil 37

3.7 Uma abordagem panorâmica da dialetologia no Brasil 38

3.8 O pioneirismo do dialeto caipira de Amaral 39

3.9 Um retrato da diatopia lexical nas regiões Sul e Nordeste do Brasil 40

3.10 Descrição das fontes 44

3.11 *Atlas Prévio dos Falares Baianos* 44

3.12 Descrição do *Atlas Linguístico do Paraná* e do *Atlas Linguístico do Paraná II* 46

3.13 A importância do ALPR para os estudos dialetológicos no Brasil e no mundo 48

3.14 Fontes do *Projeto Atlas Linguístico do Brasil*: Bahia e Paraná 49

3.15 Contribuições do Projeto ALiB Para os Estudos Dialetológicos 51

4 PROCEDIMENTOS METODOLÓGICOS 53

4.1 A Escolha das Questões do ALPR e do ALIB 53

4.2 Seleção das fontes geolinguísticas 54
4.3 Tabulação dos dados das fontes 54
4.4 Procedimentos para elaboração das cartas no SGVCLIN 55
4.5 Descrição e análise quantitativa e qualitativa das fontes pesquisadas 56

5 ANÁLISE DOS DADOS 57
5.1 QSL 046 quando se tiram os grãos de milho da espiga, sobra o quê? 57
5.1.1 Análise das variantes de *Sabugo* no APFB 58
5.1.2 Análise das variantes para sabugo no projeto ALiB-BA 64
5.1.3 Análise da variável sexo para sabugo nos dados do ALiB-BA 65
5.1.4 Análise da variável faixa etária para sabugo nos dados do ALiB-BA 65
5.1.5 Análise das variantes de *sabugo* no ALPR II 67
5.1.6 Análise das variantes de *sabugo* no ALiB-PR 68
5.1.7 Análise da variável sexo para sabugo nos dados do ALiB-PR 69
5.1.8 Análise comparativa das variáveis para sabugo nos dados da Bahia e do Paraná 70
5.2 QSL 277 Do ALPR - Depois que se Corta o Pé de Arroz ou de Fumo, fica uma parte Embaixo assim na Terra. Você já Chegou a Ver Plantação Assim? 72
5.2.1 Análise das variantes lexicais de *soca* nos dados do ALiB-BA 73
5.2.2 Análise da variável sexo nos dados do ALiB-BA referentes às variantes de *soca* 75
5.2.3 Análise da variável faixa etária nos dados do ALiB-BA para as variantes de *soca* 76
5.2.4 Análise das variantes lexicais de *soca* nos dados do ALPR II 79
5.2.5 Análise da variável sexo de *soca* nos dados do ALPR II 80
5.2.6 Análise das variantes linguísticas de *soca* nos dados do ALiB-PR 85
5.2.7 Análise da variável sexo nos dados do ALiB-PR 89
5.2.8 Análise da variável faixa etária nos dados do ALiB-PR 91
5.2.9 Análise comparativa dos dados do ALiB-BA, ALPR e ALiB-PR 96

5.3 **QSL 049 onde ficam os grãos do feijão, no pé, antes de serem colhidos?** ... 96
5.3.1 Análise das variantes de *vagem* nos dados do ALiB-BA ... 97
5.3.2 Análise das variantes de *vagem* nos dados do ALPR ... 98
5.3.3 Análise da variável sexo de vagem nos dados do ALPR ... 100
5.3.4 Análise das variantes de vagem nos dados do ALIB-PR ... 103
5.3.5 Análise das variantes de *vagem* no ALiB-PR segundo a variável sexo ... 106
5.3.6 Análise das variantes de *vagem* no ALiB-PR segundo a variável faixa etária ... 107
5.3.7 Análise comparativa dos dados do ALiB-BA, ALPR e ALiB-PR acerca das variantes de *vagem* ... 109
5.4 **QSL 054 Armação de Madeira que se Coloca no Pescoço do Animal [...] Para não Atravessar a Cerca** ... 110
5.4.1 Análise das variantes de cangalha nos dados do ALPR ... 111
5.4.2 Um retrato da variação lexical para *cangalha* nos dados do ALPR ... 111
5.4.3 Descrição dos dados na perspectiva da variável sexo no ALPR ... 113
5.4.4 Variação lexical para *cangalha* nos dados do ALIB-PR ... 114
5.5.5 Um panorama dos dados para cangalha no ALiB-PR ... 115
5.4.6 Uma descrição dos dados segundo a variável sexo no ALiB-PR ... 118
5.4.7 Descrição dos dados segundo a variável faixa etária no ALiB-PR ... 120
5.4.8 Análise comparativa dos dados do ALPR e do ALiB-PR ... 122

6 **DICIONARIZAÇÃO DAS VARIANTES ESTUDADAS** ... 125
6.1 **Dicionarização das principais variantes de sabugo** ... 125
6.2 **Dicionarização das principais variantes de *soca*** ... 125
6.3 **Dicionarização das variantes de vagem** ... 127
6.4 **Dicionarização das variantes de *canga*** ... 127

7 **CONSIDERAÇÕES FINAIS** ... 130

REFERÊNCIAS ... 136

1 INTRODUÇÃO

A língua reflete a história e a identidade de seus falantes, constituindo-se em um patrimônio dinâmico e diversificado. No Brasil, país de dimensões continentais e de rica pluralidade linguística, a variação do português é evidente em diferentes níveis – fonético, lexical, morfológico e sintático – configurando um campo fértil para estudos que buscam compreender a relação entre língua, espaço e sociedade. Essa diversidade linguística é o resultado de múltiplos fatores, como região, faixa etária, escolaridade, sexo e contexto rural ou urbano, que interagem para moldar as diferentes formas de expressão verbal.

No contexto da Dialetologia, as variações lexicais, em particular, revelam importantes aspectos das dinâmicas sociolinguísticas e culturais de um povo. As diferentes formas usadas para o mesmo referente – como *sabugo* ou *capuco*, *soca* ou *soqueira*, *girassol* ou *mirassol*, *vagem* ou *bainha* – ilustram como o léxico pode variar em função de fatores geográficos, históricos e sociais.

Nesse sentido, a presente dissertação tem como objetivo central analisar a distribuição das variantes lexicais em dados geolinguísticos de dois estados brasileiros: Bahia e Paraná. São objetivos específicos: (i) descrever os dados coletados pelo ALiB para fins de cartografação; (ii) comparar os dados baianos e paranaenses dos respectivos atlas estaduais com os obtidos pelo Projeto ALiB nesses estados; (iii) verificar a correlação rural × urbana na elicitação das respostas; e (iv) avaliar a influência das variáveis sexo e faixa etária.

A análise baseia-se em dados coletados de diferentes *corpora*, abrangendo três sincronias: o *Atlas Prévio dos Falares Baianos* (Rossi, 1963), o *Atlas Linguístico do Paraná* (Aguilera, 1994) e o *Projeto Atlas Linguístico do Brasil* (ALiB) em ambos os estados. A comparação entre esses acervos lexicais permitirá identificar padrões de manutenção ou apagamento das variantes em diferentes contextos sociogeográficos.

Partimos da hipótese de que as variantes lexicais relacionadas ao vocabulário agropastoril tendem a apresentar um processo de apagamento mais acentuado em contextos urbanos, devido ao afastamento das práticas agrícolas.

Considerando que as palavras mudam de acordo com fatores como o tempo, o lugar e o sexo, esta dissertação contribui, diretamente, para a preservação da diversidade linguística brasileira e, especialmente, paranaense e baiana. O respeito

por essas formas de manifestação auxilia na preservação do patrimônio social e lexical, combatendo, assim, muitos preconceitos e, em especial, o linguístico.

Ressaltamos ainda que esse fenômeno da variação linguística interessa tanto aos dialetólogos, que investigam a distribuição geográfica das variantes conservadoras ou inovadoras, quanto aos sociolinguistas, que analisam sua correlação com fatores sociais.

O presente estudo fundamenta-se nos pressupostos teórico-metodológicos da Geolinguística (Cardoso, 2010) e da Dialetologia Pluridimensional, proposta por Thun (1998), que, no Brasil, vem se consolidando paulatinamente no *Projeto Atlas Linguístico do Brasil* (ALiB), iniciado na Universidade Federal da Bahia, em 1996, em parceria com outras universidades, a fim de retratar a diversidade linguística do português falado no Brasil sob os aspectos fonético, morfossintático, lexical, entre outros. Dessa forma, os procedimentos metodológicos estão pautados na Dialetologia Tradicional, na Dialetologia Bidimensional e na Dialetologia Pluridimensional.

Esta pesquisa fornece uma visão mais abrangente sobre a variabilidade e os processos de mudança lexical no português brasileiro, ressaltando a importância de preservar o patrimônio linguístico nacional diante das transformações sociais e culturais que impactam a língua.

A escolha do tema desta dissertação teve como mote nossas vivências em dois ambientes geográficos e culturais distintos. Filho de um agricultor e de uma professora e cartorária do interior da Bahia, desde os tenros anos de idade acompanhei meu pai na lida áspera e, ao mesmo tempo, aprazível da vida simples e sofrida da roça, ora arando a terra, plantando e colhendo arroz, algodão, milho e feijão; ora apartando o gado, curando um bezerro ou rebanhando o rebanho a cavalo, de domingo a domingo, sob sol ou chuva. Graças a essa proximidade com o trabalho agrícola, pude conhecer a *soca* de algodão e de arroz, a *vagem* de feijão, o *capuco* de milho e a *canga* da cabra roceira. Já adolescente, fiz curso de datilografia e fui trabalhar no cartório de minha saudosa mãe, com quem aprendi esse ofício e os principais valores que norteiam a minha essência como filho, irmão, pai, esposo e cidadão.

Concluído o magistério em minha pacata e hospitaleira cidade de Caturama, na Bahia, almejando cursar Direito, migrei, em julho de 1996, para o Paraná, a convite de parentes, a fim de estudar, dada a escassez, à época, de ensino superior no estado baiano. Com muita dedicação, fui aprovado na Faculdade de Direito de Jacarezinho, onde, para prover o meu sustento, lecionava, nas horas vagas, Língua Portuguesa no

ensino médio e em cursinhos pré-vestibulares dessa acolhedora cidade e em suas adjacências, expandindo esse projeto para o interior de São Paulo, onde trabalhei em escolas particulares e cursinhos de Ourinhos, Bauru e Marília. Após essa etapa, incentivado por parentes e amigos, advoguei por um período exíguo sem, entretanto, abdicar do magistério. Em 2020, concluí a faculdade de Letras, na modalidade EaD, e, para minha felicidade, fui aprovado no mestrado da Universidade Estadual de Londrina (UEL), estando agora prestes a realizar o sonho de me tornar mestre em Estudos da Linguagem.

Esta dissertação é composta de oito capítulos. O capítulo 1 trata da introdução, que contextualiza a importância da variação lexical para os estudos e delineia os objetivos, descrevendo as fontes sobre as quais incide a pesquisa. O capítulo 2 versa sobre os pressupostos teórico-metodológicos em que se fundamenta a investigação, dividida em três eixos principais, e apresenta uma abordagem sobre o método geolinguístico, além de traçar um panorama sobre a dialetologia no Brasil e descrever as fontes. O capítulo 4, subdividido em seções, faz um levantamento sobre o rural e o urbano no Brasil. O capítulo 5 aborda os procedimentos metodológicos e descreve as etapas percorridas para a consecução deste trabalho. O capítulo 6 apresenta a descrição e a discussão dos dados, elucidados por meio de relatórios, tabelas e cartas linguísticas. No capítulo 7, é feita uma imersão lexicográfica em quatro dicionários de edições e suportes diferentes para averiguar quais verbetes foram registrados. O capítulo 8, intitulado *Considerações finais*, traz as reflexões acerca da pesquisa.

2 UMA ABORDAGEM SOBRE O FALAR URBANO E O RURAL

Este capítulo versa sobre a relação entre o falar rural e o falar urbano numa perspectiva histórica, social e geográfica, com base em Isquerdo (2010), Ribeiro (2015), Diégues Júnior (1960) e Santos (2007), retratando as transformações ocorridas no espaço e seus reflexos na variação linguística da Bahia e do Paraná.

2.1 UM LEVANTAMENTO SOBRE TRAÇOS DO RURAL E DO URBANO NO BRASIL DO DESCOBRIMENTO ATÉ O SÉCULO XX

A grande extensão territorial do Brasil é marcada pelo contraste entre áreas rurais e urbanas. Cada uma delas, com suas peculiaridades históricas, culturais e econômicas, constitui um verdadeiro retrato da diversidade do país. É certo que tanto no campo quanto nas cidades há problemas sociais que demandam soluções enérgicas para melhorar a vida de seus habitantes, quer seja no acesso à educação, à saúde, à segurança, ao emprego ou ao lazer. Por outro lado, não se pode negar que a participação de ambos é imprescindível para o crescimento do país. O setor rural produz os alimentos (feijão, arroz, legumes, carne, leite) que abastecem a população da zona urbana. Esta, por sua vez, é responsável pelo setor de serviços (médicos, advogados, engenheiros, mecânicos etc.) e pelo comércio (lojas, padarias, supermercados, farmácias), atendendo às necessidades vitais de seus habitantes.

Para compreender essa dicotomia rural × urbano, é preciso primeiro responder a duas indagações: o que é rural? e o que é urbano? Por muito tempo, o rural foi rotulado pejorativamente como símbolo do que é caipira, retrógrado, atrasado, ignorante, cafona, do interior, entre outras denominações. Assim, podia-se concluir que o espaço rural estava intrinsecamente relacionado às atividades voltadas ao cultivo da terra e à criação de animais.

No que diz respeito aos costumes e tradições de seus habitantes, as comunidades rurais, isoladas dos grandes centros e com baixa densidade demográfica, eram constituídas por laços familiares mais próximos, em um ambiente em que todos se conheciam e se ajudavam nas atividades da lida do campo. Seus moradores, tidos por muitos como matutos, eram pessoas simples, com pouco ou nenhum grau de instrução formal; a maioria possuía baixo nível de escolaridade, e os poucos privilegiados talvez chegassem ao ensino fundamental, completo ou incompleto.

Já o urbano era compreendido como o espaço com alta densidade demográfica, constituído por aglomerações de casas, ruas, avenidas e praças pavimentadas e bem cuidadas, com eletricidade, água encanada, escolas, hospitais, comércio e prefeitura. Portanto, um espaço geográfico com infraestrutura, voltado ao progresso, ao moderno e ao dinâmico. Era também o espaço do cinema, do teatro, dos restaurantes e lanchonetes, onde seus moradores eram considerados civilizados e falavam corretamente o português culto, uma vez que, geograficamente, tinham maior acesso às boas escolas e às universidades. Diante dessas diferenças, pode-se concluir que a população rural é mais homogênea e a urbana mais heterogênea, cada qual com suas peculiaridades em relação aos costumes, à vida social, ao trabalho e à escolaridade.

A respeito dessa dicotomia rural × urbano, vale salientar que até o século XVIII o Brasil era um país agrário, com alta produção de grãos para abastecer o mercado interno e externo. Contudo, a partir do século XIX, esse modelo perde força e cede espaço ao crescimento das cidades, dando ensejo a uma nova configuração do país. Nesse sentido, Isquerdo (2010, p. 137) esclarece:

> O Brasil-Colônia foi marcado pelo seu caráter eminentemente rural. As cidades e vilas, em sua maioria, localizadas em zonas costeiras, pouca influência exerciam sobre as povoações do interior, uma vez que as cidades dependiam muito do meio rural, situação que perdurou por mais de três séculos (XVI a XVIII). No século XIX ocorreu o declínio do patriarcalismo rural e o surgimento do urbano, quando a cidade passou a influir normativamente na vida do País e uma nova geografia humana e urbana começou a ser desenhada na sociedade brasileira, com a fundação de inúmeras cidades e o crescimento vertiginoso de outras.

Desde os primórdios do descobrimento até o século XIX, o espaço rural ocupou lugar de destaque na economia nacional, sendo uma verdadeira fonte geradora de riquezas, principalmente durante os períodos colonial e imperial. Sua economia tinha por base a exportação de cana-de-açúcar, pedras preciosas e café para o mercado externo. Para atender às demandas internacionais e ampliar o lucro, tornou-se necessário importar mão de obra da África. Nesse contexto, o escravo africano ganha relevância no incremento da produção desses produtos. Contudo, as condições impostas a eles eram degradantes e desumanas, desde a insalubridade dos navios negreiros até as péssimas condições das senzalas, onde descansavam depois de jornadas exaustivas nas lavouras e na mineração. As bases dessa prosperidade estavam na monocultura, na mão de obra escravocrata e no latifúndio.

2.2 Um Estudo Sobre a Formação do Brasil sob uma Perspectiva Étnica, Antropológica, Histórica e Geográfica

Ribeiro (2015) retrata a composição étnica e cultural do Brasil, elencando a miscigenação de três raças na constituição da nossa gênese: o europeu, o indígena e o africano. A obra é dividida em cinco partes – *O Novo Mundo*, *Gestão Étnica*, *Processo Sociocultural*, *Os Brasis na História* e *O Destino Nacional* –, todas de grande importância para a compreensão da nossa história, dada a variedade dos assuntos, que abordam aspectos como a chegada do colonizador português, as missões jesuíticas, os governadores-gerais, as sesmarias e os ciclos econômicos, como, por exemplo, a cana-de-açúcar e a mineração. Para este trabalho, interessa principalmente a quarta parte, intitulada *Os Brasis na História*, na qual Ribeiro (2015) expõe:

> Uma copiosa documentação histórica mostra que, poucas décadas depois da invasão, já se havia formado no Brasil uma protocélula étnica neobrasileira diferenciada tanto da portuguesa como das indígenas. Essa etnia embrionária, multiplicada e difundida em vários núcleos - primeiro ao longo da costa atlântica, depois transladando-se para os sertões interiores ou subindo pelos afluentes dos grandes rios -, é que iria modelar a vida social e cultural das ilhas-Brasil. Cada uma delas singularizada pelo ajustamento às condições locais, tanto ecológicas quanto de tipos de produção, mas permanecendo sempre como um renovo genésico da mesma matriz. (Ribeiro, 2015, p. 201)

Dadas as peculiaridades de cada região, o autor divide *Os Brasis* em cinco: *O Brasil crioulo*, *O Brasil caboclo*, *O Brasil sertanejo*, *O Brasil caipira* e *Brasis sulinos: gaúchos, matutos e gringos*. Em linguagem acessível e até poética, Ribeiro fala, com minúcias, sobre essas ilhas-Brasil, ressaltando as características de cada uma delas, seu povo, economia, clima, relevo e cultura.

Ribeiro (2015, p. 203) retrata essa diversidade da formação do povo brasileiro:

> Elas são representadas pela cultura crioula, que se desenvolveu nas comunidades da faixa de terras frescas e férteis do Nordeste, tendo como instituição coordenadora fundamental o engenho açucareiro. Pela cultura caipira, da população das áreas de ocupação dos mamelucos paulistas, constituída, primeiro, através das atividades de preia de índios para a venda, depois, da mineração de ouro e diamantes e, mais tarde, com as grandes fazendas de café e a industrialização. Pela cultura sertaneja, que se funde e difunde através dos currais de gado, desde o Nordeste árido até os cerrados do Centro-Oeste. Pela cultura cabocla das populações da Amazônia, engajadas na coleta de drogas da mata, principalmente nos seringais. Pela cultura gaúcha do pastoreio nas campinas do Sul e suas duas variantes, a matuta-açoriana (muito parecida com a caipira) e a

> gringocaipira das áreas colonizadas por imigrantes, predominantemente alemães e italianos.

Interessa-nos principalmente a análise historiográfica e antropológica do terceiro e do quinto *Brasis*, tendo em vista a convergência de assuntos neles abordados e as fontes de pesquisa desta dissertação.

Para o autor, esse Brasil é retratado como uma vasta extensão territorial composta por vegetação rala, solo pobre e chuvas irregulares, principalmente no semiárido brasileiro, delimitada de um lado pela floresta da costa atlântica e de outro pela floresta amazônica. Diferentemente das terras úmidas do Brasil crioulo, com suas chuvas abundantes, onde a cana-de-açúcar prosperava em torno dos engenhos, no semiárido esse cultivo não vingava, obrigando o sertanejo a buscar sua subsistência na criação de gado caprino, ovino e, principalmente, bovino.

Esse território contempla o sertão nordestino, região de clima quente e seco, com chuvas irregulares, onde a caatinga, com suas plantas espinhosas, resiste a essas adversidades. Nesse cenário, o rio São Francisco ganha enorme importância na difusão da criação de gado e no povoamento da região. Ressalte-se que os rebanhos eram comprados, mas as terras pertenciam à Coroa, que as concedia gratuitamente em sesmarias.[1] A figura do vaqueiro foi determinante para a criação de gado. Eram homens simples, trajando roupas de couro (gibão) durante a lida com o gado, para se proteger dos espinhos da caatinga. Moravam com suas famílias em casas modestas nas fazendas e, para prover o sustento, plantavam roças e criavam vacas para a produção de leite. Já o Brasil Central caracteriza-se pelos arbustos tortuosos e gramíneas, plantas típicas do cerrado, bioma propício para a criação de gado.

> Para além da faixa nordestina das terras frescas e férteis do massapé, com rica cobertura florestal, onde se implantaram os engenhos de açúcar, desdobram-se as terras de uma outra área ecológica. Começam pela orla descontínua ainda úmida do agreste e prosseguem com as enormes extensões semi-áridas das caatingas. Mais além, penetrando já o Brasil Central, se elevam em planalto como campos cerrados que se estendem por milhares de léguas quadradas (Ribeiro, 2015, p. 250-267).

Os *Brasis sulinos* surgem a partir da expansão dos antigos paulistas, mas se distanciaram culturalmente destes, adquirindo características peculiares em relação

1 Lotes de terras distribuídas pela Coroa a um sesmeiro, pessoa beneficiada.

aos outros povos brasileiros (crioulos, caboclos, sertanejos e caipiras). Essa heterogeneidade tão ímpar ganhou contornos internos, dando ensejo a três grupos étnicos distintos dentro da Região Sul: gaúchos, matutos e gringos.

Os matutos-açorianos, imigrantes açorianos que ocuparam os estados do Paraná, Santa Catarina e Rio Grande do Sul, representam a identidade do homem simples, constituída pela composição do camponês português que trouxe consigo valores, costumes e experiência na lida com a terra.

Originários dos pampas, os gaúchos têm características próprias em seu modo de falar, vestir e se comportar, adquirindo sua identidade por meio de costumes e tradições. Usando botas, bombacha, guaiaca e chapéu e munidos de um bom chimarrão, simbolizam o homem hábil, que, montado a cavalo, campeava o gado que se multiplicava nas pradarias do rio da Prata com desenvoltura e competência.

Os gringos eram formados por imigrantes europeus, principalmente alemães e italianos, que trouxeram sua cultura, costumes, danças e festas típicas, como a Oktoberfest de Blumenau, em Santa Catarina. Além disso, inovaram a agricultura por meio de práticas europeias em pequenas propriedades, cultivando a terra de modo intensivo, em regime de policultura.

Ribeiro (2015, p. 299) descreve didaticamente as três vertentes dos *Brasis sulinos*:

> Tais são os lavradores matutos de origem principalmente açoriana, que ocupam a faixa litorânea do Paraná para o sul; os representantes atuais dos antigos gaúchos da zona de campos da fronteira rio-platense e dos bolsões pastoris de Santa Catarina e do Paraná, e, finalmente, a formação gringo-brasileira dos descendentes de imigrantes europeus, que formam uma ilha na zona central, avançando sobre as duas outras áreas.

A diversidade identitária do povo brasileiro poderia ter sido uma fonte de desagregação do país, dando ensejo à formação de outras nações, como ocorreu na colonização espanhola da América do Sul. Todavia, essa heterogeneidade, fruto da miscigenação entre indígenas, africanos e europeus, foi responsável pelo sentimento de brasilidade que uniu esse mosaico de *Brasis* em um só povo: o brasileiro.

O surgimento e o crescimento das vilas e cidades no Brasil estão intrinsecamente relacionados com os ciclos econômicos (cana-de-açúcar, mineração, café, entre outros), que trouxeram consigo operários, comerciantes, funcionários públicos e imigrantes para atender ao mercado interno, incipiente e sedento de expansão. Muitas dessas cidades surgiram por ato expresso da Coroa, como

Salvador, Belém e Florianópolis, que exerciam funções de comércio, importação e exportação de açúcar e, mais tarde, de pedras preciosas.

De acordo com Ribeiro (2015), as cidades brasileiras mais antigas foram Salvador, Rio de Janeiro, João Pessoa, Olinda, Recife, São Luís e Ouro Preto, entre outras. Dada a riqueza do açúcar, Recife e Salvador prosperaram bastante, a primeira principalmente em virtude dos empreendimentos dos holandeses. São Luís cresceu vertiginosamente graças à Guerra de Secessão nos Estados Unidos, a ponto de, em 1872, ultrapassar São Paulo em população e riqueza.

> No curso desses séculos as cidades cresceram e se ornaram como portentosos centros de vida urbana, só comparáveis aos do México. Os holandeses enriqueceram Recife. A riqueza das minas se exibiu em Ouro Preto e outras cidades do ouro, engalanou a Bahia e depois, o Rio (Ribeiro, 2015, p. 146).

A urbanização no Brasil colônia esteve sempre atrelada ao monopólio e à monocultura, que expulsaram o homem do campo, dando ensejo ao êxodo rural e deslocando o eixo econômico para as grandes cidades. Todavia, o fator determinante para o crescimento urbano foi a industrialização, uma vez que uma não existe sem a outra. A necessidade de mão de obra para trabalhar nas fábricas corroborou a mudança desse eixo econômico. Além disso, a busca por melhor qualidade de vida transformou esse contexto. Assim, água encanada, energia e atendimento médico foram determinantes para a formação dos núcleos urbanos no Brasil.

De igual relevância para esta dissertação é a obra de Diégues Júnior (1960), que mergulha na diversidade cultural brasileira, iluminando especificidades que compõem a pluralidade de manifestações em cada recanto do país. O autor trata das influências históricas, sociais e econômicas que perpassam as regiões, revelando como tradições, crenças e práticas populares se entrelaçam. O trabalho de Diégues Júnior propõe um olhar atento sobre o papel da colonização, da imigração e das interações entre as culturas indígena, africana e europeia, ao delinear a construção das identidades culturais que permeiam a nação brasileira.

Ao tratar da formação, das características e das condições atuais, o autor examina regiões do Brasil em que a cultura voltada para a criação de gado, em contexto rural, se sobressai. Ao trazer à tona a noção de *Mediterrâneo Pastoril*, o autor estabelece analogia com a tradição de territórios secos e de pastoreio, típicos da região mediterrânea. A análise aborda como essa atividade, essencialmente focada no pastoreio, impacta a organização social e cultural dessas áreas, refletindo uma

fusão entre as condições ambientais, as heranças da colonização portuguesa e os desafios do clima semiárido. Diégues Júnior (1960) destaca também como as relações de trabalho e o modo de vida rural se entrelaçam com a resistência das comunidades locais, que preservam suas particularidades apesar das adversidades e da estrutura socioeconômica em constante transformação.

Segundo Diégues Júnior (1960), a introdução da criação de gado na região mediterrânea do Brasil ocorreu por meio de expedições que, seguindo o curso do rio São Francisco, se espalharam pelo interior da Bahia e de Pernambuco. O autor destaca a expedição de Navarro-Espinoza, realizada entre 1553 e 1555, e, posteriormente, as de Cristóvão da Rocha e Domingos Martins, que partiram de Pernambuco após 1590. Muitas e sucessivas foram as expedições posteriores, todas com o propósito de desbravar e expandir as fronteiras do sertão.

O objetivo inicial dessas expedições, além de estabelecer a criação de gado, era também a busca por pedras preciosas. Contudo, como aponta o autor, nem nas regiões do Baixo, nem do Médio, nem do Alto São Francisco se encontraram as riquezas minerais almejadas.

O fracasso inicial na busca pelas tão almejadas jazidas de ouro deu lugar a um novo ciclo: o comércio de gado nas regiões interiores desses estados, propiciando o surgimento de grandes pecuaristas.

> No desenvolvimento do criatório foi importante o papel de Garcia d'Ávila, importando gado de Cabo Verde. Dos seus primeiros currais, foram nascendo outros, cresceu a pecuária e o gado foi penetrando, afastando-se do litoral. Os Garcias d'Ávila e os Guedes de Brito transformaram-se em verdadeiros e poderosos donos das sesmarias de criação de gado, ao longo do São Francisco, expandindo-se daí em outras direções (Diégues Júnior, 1960, p. 144).

Portanto, a criação de gado foi responsável pelo povoamento do sertão nordestino, a partir do rio São Francisco, adentrando o interior do Ceará, do Rio Grande do Norte e do Piauí. Em ritmo mais lento, essa ocupação chegou ao Maranhão e a Goiás, espalhando-se mais tarde por outros estados, alcançando inclusive o Triângulo Mineiro.

Em virtude do gado, o homem adentrou o interior, explorou suas terras e ali construiu moradias. Segundo o autor, os rebanhos não apenas foram responsáveis pela ocupação humana, mas também pela cultura e pela economia da região, integrando o quotidiano das pessoas a ponto de o período ser cognominado *era do couro*, pois essa matéria-prima estava presente em quase tudo, inclusive na porta das

casas, na cama dura, nas cordas para a lida com os bois, no *mocó* (saco de peles de animais), na bainha de faca, entre outros – verdadeira referência à importância da pecuária na vida daquelas comunidades do sertão.

Para Diégues Júnior (1960), a relação entre patrão e vaqueiro assume novo arranjo de trabalho, no qual a remuneração pelos serviços de manejo do gado era feita na proporção de uma cria para cada grupo de quatro animais nascidos na propriedade. A hostilidade inerente ao trabalho também contribuiu para estreitar os laços entre o vaqueiro e o fazendeiro, conferindo à relação um caráter mais humanizado e menos rigidamente hierarquizado.

Conforme descreve o autor, a sociedade brasileira do período era escravagista, mas, no sertão, a proporção de escravos diminuía drasticamente em comparação com os engenhos de açúcar. Essa diferença se devia ao fato de a atividade pecuária demandar quantidade significativamente menor de mão de obra, uma vez que o vaqueiro, sozinho, era capaz de manejar todo o rebanho.

Sobre a colonização estrangeira e o papel do imigrante, Diégues Júnior (1960) explora a contribuição dos imigrantes na formação socioeconômica do Brasil. O autor analisa diferentes grupos que chegaram ao país, principalmente a partir do século XIX, os quais influenciaram substancialmente a agricultura, a cultura e a indústria.

Com o Bloqueio Continental perpetrado por Napoleão Bonaparte, a Família Real foi obrigada a abandonar Portugal e vir para o Brasil, onde fixou moradia no Rio de Janeiro em 1808. Para Diégues Júnior (1960), decisões tomadas pelo Príncipe Regente contribuíram significativamente para a imigração no Brasil, como a abertura dos portos, no mesmo ano, e o decreto de 25 de novembro, concedendo terras a estrangeiros.

A imigração alemã, iniciada em 1824 no Rio Grande do Sul, facilitou a expansão para outros estados, especialmente Santa Catarina, Paraná e São Paulo. No Sul do país, os imigrantes encontraram condições climáticas similares às de seus países de origem, o que contribuiu para o estabelecimento e desenvolvimento dessas comunidades. Conforme destacado pelo autor (Diégues Júnior, 1960), os primeiros grupos estrangeiros a se fixarem na região foram compostos por italianos e alemães, seguidos por russos, poloneses, ucranianos e japoneses. Diferentemente de São Paulo, em terras meridionais houve formação de colônias.

O auge da imigração ocorreu a partir de 1850, com correntes que aportaram em São Paulo, constituídas principalmente por italianos, recrutados para substituir a

mão de obra escrava nas lavouras de café, intensificando-se a partir de **1888**, com a abolição da escravatura (Diégues Júnior, 1960).

Diversos fatores foram decisivos para que imigrantes escolhessem o Sul como destino. Entre outros aspectos, destacam-se a concessão de lotes de terras devolutas para a formação de pequenas propriedades e a exígua presença de trabalho escravo, em contraste com outras regiões do país. Ademais, a ausência de latifúndios e de grandes propriedades contribuiu para essa preferência. De acordo com o autor, fatores naturais, como terras férteis e clima semelhante ao europeu, incentivaram o povoamento dessa região pelos colonos, que encontraram condições ideais para fixar moradia.

O Brasil rural, durante os períodos colonial e imperial, foi o cerne da economia do país. Até o século XX, o espaço rural foi o maior gerador de riquezas, com economia pujante exportadora de produtos primários, como cana-de-açúcar e café. Entretanto, com a industrialização iniciada a partir de 1930, com a indústria de base da Era Vargas (1930 a 1945) e, em especial, com a criação da Companhia Siderúrgica Nacional, em Volta Redonda, a indústria assumiu papel importante na economia brasileira, com substituição de importações e alto investimento de capital público.

A consolidação viria nas décadas seguintes, sobretudo no fim dos anos 1950, no governo de Juscelino Kubitschek (1956 a 1961). Ao assumir o poder, JK implementou o Plano de Metas que, segundo seu lema, faria o Brasil “crescer cinquenta anos em cinco”. Nesse período, com abertura ao capital externo, chegaram as multinacionais e, em especial, a indústria automobilística.

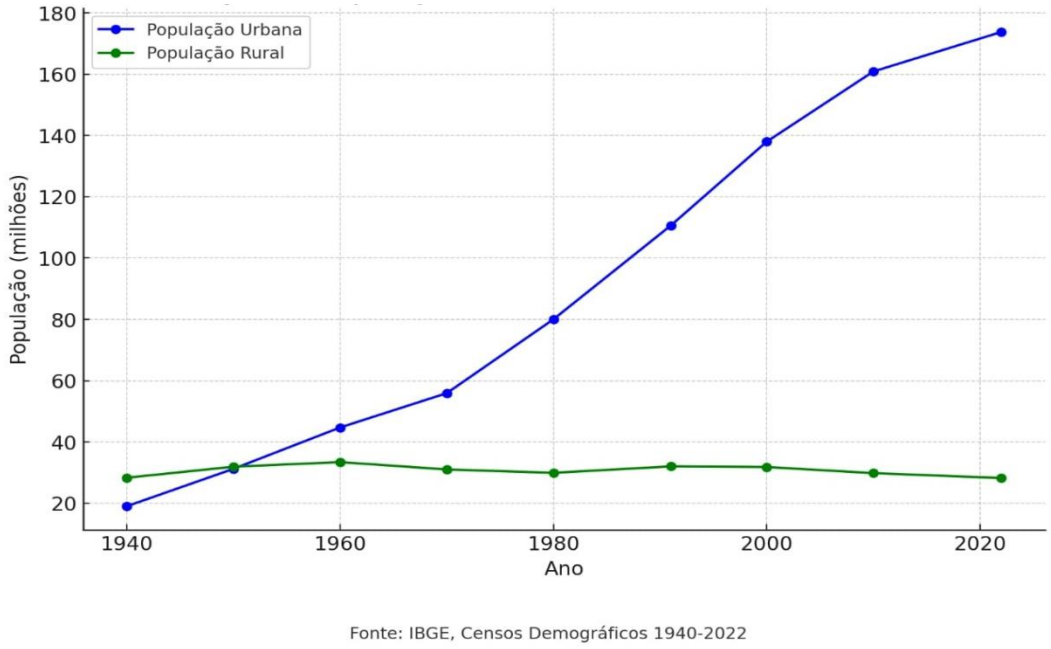
Esse plano audacioso incluiu a construção de Brasília – sonho antigo de José Bonifácio[2] –, com a consequente mudança da capital do Rio de Janeiro para a Região Centro-Oeste. Para pôr em prática esse plano, criou-se a estatal Novacap, que realizou concurso para o projeto arquitetônico, vencido por Lúcio Costa e com arquitetura de Oscar Niemeyer.

Para a consecução desse projeto, o governo contou com grande contingente de trabalhadores nordestinos que migraram em massa para a futura capital e ficaram conhecidos como candangos. Concluídas as obras, muitos tiveram de deixar Brasília; parte deles se alojou no entorno da nova capital, dando origem às cidades-satélites. Com o intuito de demonstrar o processo de urbanização no Brasil, elaboramos o

[2] Paulista de Santos (1763 a 1838), José Bonifácio de Andrada e Silva, conhecido como Patriarca da Independência, exerceu papel decisivo na independência política do Brasil.

Gráfico 1, com a evolução da população rural × urbana ao longo das décadas (1940 a 2022).

Gráfico 1 – Evolução da População Urbana e Rural no Brasil (1940-2022)



Fonte: IBGE (2022).

De acordo com o Gráfico 1, podemos concluir que, até o fim da década de 1940, a população rural era a mais expressiva, pois o Brasil era um país agrário, movido pela exportação de matérias-primas (açúcar, café, pedras preciosas etc.). Todavia, a partir da década de 1950, com a industrialização impulsionada principalmente pelo Plano de Metas, do governo JK, a população urbana ultrapassa, pela primeira vez, a rural, que entra em declínio nas décadas seguintes.

É imprescindível ressaltar que a urbanização no Brasil não ocorreu de forma homogênea em todo o território, tratando-se de um processo centralizado na Região Sudeste, sobretudo em São Paulo e no Rio de Janeiro, onde a industrialização estava em constante ascensão e, consequentemente, recebia migrantes de outras regiões, principalmente nordestinos que vinham em busca de trabalho.

Assim, o contraste entre as realidades urbana e rural no Brasil também se reflete nas relações de trabalho e nas dinâmicas econômicas. No ambiente urbano, predominam atividades associadas aos setores secundário e terciário, como a indústria, o comércio e os serviços, enquanto o meio rural está vinculado predominantemente ao setor primário, com foco na agropecuária. Contudo, atualmente, esse modelo apresenta grande diversidade: enquanto o agronegócio se

expande, impulsionado por tecnologias avançadas e altos índices de produtividade, os pequenos agricultores ainda enfrentam obstáculos significativos, como a limitação no acesso a crédito, a inovações tecnológicas e a mercados. Esse cenário favorece a concentração de terras e de renda, intensificando as disparidades regionais e sociais.

A obra *O espaço do cidadão* (Santos, 2007) realiza uma análise das relações de interdependência entre os espaços rural e urbano, evidenciando como essas dimensões territoriais, apesar de suas especificidades, articulam-se e promovem transformações recíprocas. O autor explora criticamente os impactos dessas interações no exercício pleno da cidadania, enfatizando as disparidades socioespaciais resultantes de dinâmicas históricas e econômicas que tendem a privilegiar as lógicas urbanas em detrimento das rurais. Santos argumenta que a organização territorial brasileira reflete um modelo de desenvolvimento assimétrico, no qual as interações entre o campo e a cidade são atravessadas por contradições estruturais e desafios significativos à inclusão e à equidade social. Nas palavras do autor:

> Esse homem do campo é menos titular de direitos que a maioria dos homens da cidade, já que os serviços públicos essenciais lhe são negados, sob a desculpa da carência de recursos para lhe fazer chegar saúde e educação, água e eletricidade, para não falar de tantos outros serviços essenciais (Santos, 2007, p. 42).

Essas diversidades históricas, antropológicas, étnicas e geográficas refletem-se nos costumes, nas tradições e, sobretudo, no falar rural × urbano, uma vez que suas peculiaridades se traduzem no repertório linguístico do homem do campo em contraste com o da cidade. No primeiro caso, há uma linguagem mais tradicional e conservadora, que preserva sua origem latina, indígena ou africana (Almeida-Baronas, 2012, p. 15). No segundo, observa-se uma linguagem mais flexível e dinâmica, sujeita a alterações e a incorporações de estrangeirismos, como, por exemplo, anglicismos (palavras de origem inglesa), galicismos (palavras de origem francesa) e italianismos (palavras de origem italiana), resultado do contato, nos grandes centros urbanos, com pessoas de diferentes nações, em especial os imigrantes que aqui aportaram em busca de melhor qualidade de vida, sobretudo no período entre guerras.

Pesquisas comprovam que o falar rural e o falar urbano, dentro da perspectiva da variação linguística, constituem vertentes diferentes da mesma língua, cada qual com suas peculiaridades, que podem ocorrer na fonética, na morfossintaxe e no

léxico, considerando-se os aspectos sociais que as consubstanciam, a exemplo dos costumes, festividades e tradições, elementos indissociáveis de uma sociedade em parte rural e em parte urbana.

Em síntese, o acervo literário de que dispusemos neste capítulo – em especial as obras de Ribeiro (2015) e Diégues Júnior (1960) – ajudou-nos a compreender como se deram os processos de ocupação e povoamento (rural × urbano) na Bahia e no Paraná, apresentando peculiaridades históricas, geográficas e econômicas de cada um desses estados, o que permitirá um melhor entendimento dos desdobramentos da variação linguística em ambos os *corpora* analisados neste trabalho científico.

3 REFERENCIAL TEÓRICO

Os pressupostos teóricos desta dissertação estão alicerçados em quatro ramos da Linguística, denominados, respectivamente, dialetologia tradicional, dialetologia bidimensional, dialetologia pluridimensional e geolinguística, com base nos estudos de Trudgill e Campoy (2007), Cardoso (2010) e Thun (1998).

3.1 DIALETOLOGIA

Para compreender a definição de Dialetologia, cabe primeiramente estudar o que é um dialeto. Parte-se da premissa de que o dialeto é a variação que ocorre dentro de uma língua, marcada por alterações no vocabulário, na gramática e na pronúncia, podendo ser observada em razão de fatores históricos, geográficos ou sociais.

Chambers e Trudgill (1994) estabelecem uma correlação entre os conceitos de dialeto e de sotaque. Segundo os autores, o sotaque refere-se especificamente à maneira como um falante articula os sons da língua, abrangendo, assim, variações de natureza fonética e fonológica, gramaticais e lexicais, configurando-se como uma categoria mais ampla de variação linguística.

Cabe à Dialetologia estudar os dialetos, entendidos como o conjunto de características fonéticas, morfossintáticas ou lexicais distribuídas por determinado espaço geográfico, demonstrando a variabilidade que se reflete na constituição de uma língua em uso.

De acordo com Trudgill e Campoy (2007), a Dialetologia constitui-se como um campo da Linguística que se dedica ao estudo dos dialetos, com foco nas variações fonológicas, morfológicas e lexicais que caracterizam diferentes formas de fala. Inicialmente, a disciplina teve como principal objeto de estudo os dialetos tradicionais de áreas rurais, analisados sob uma perspectiva diatópica, isto é, levando-se em conta a distribuição espacial dessas variantes linguísticas. Essa abordagem possibilita uma compreensão detalhada das formas dialetais, suas transformações e os fatores regionais que influenciam sua formação e perpetuação. Assim, o estudo dialetológico transcende a mera descrição das variações, englobando também a investigação dos elementos que determinam e sustentam os dialetos.

Nessa perspectiva, a Dialetologia, segundo Cardoso (2010, p. 15), configura-se como uma subárea da Linguística voltada à identificação, descrição e análise da variação linguística em seus diferentes usos, considerando as dimensões geográficas,

socioculturais e cronológicas que influenciam essa diversidade. Esse campo de investigação visa mapear as variações internas de uma língua, explorando as modificações que ocorrem em função de aspectos espaciais, sociais e temporais, com o intuito de compreender os processos de variação e transformação linguística ao longo do tempo e do espaço.

3.2 Dialetologia Tradicional ou Monodimensional

A Dialetologia Tradicional, também denominada monodimensional, horizontal ou areal, preocupa-se prioritariamente em descrever os aspectos fonológicos, morfológicos, sintáticos e lexicais, junto a informantes analfabetos de origem rural. Seu pioneirismo desempenha papel relevante na geração de mapas linguísticos, priorizando o aspecto diatópico na descrição dos dados, sem se preocupar, a princípio, com os aspectos sociais (sexo e faixa etária).

Historicamente, a Dialetologia Monodimensional firmou-se no início do século XX com os estudos de Gilliéron e Edmont, na publicação do *Atlas Linguistique de la France*, que aplicava uma metodologia voltada especificamente ao mapeamento espacial das variantes linguísticas. Gilliéron realizou o levantamento em 639 localidades, utilizando um questionário que inicialmente continha cerca de 1.400 perguntas, número que chegou a aproximadamente 1.900 ao final dos inquéritos (Cardoso, 2010). Esse trabalho, entretanto, recebeu críticas por não considerar os aspectos sociais.

Esse modelo dedicou-se sobretudo ao estudo do léxico, mas também abrangeu fatores fonéticos e morfossintáticos, oferecendo um panorama estático da língua em um momento histórico. Os dados obtidos eram transformados em cartas linguísticas, nas quais eram desenhadas linhas de isoglossa para indicar os limites geográficos de certas variantes.

No Brasil, vários trabalhos se incumbiram de retratar a realidade do português aqui falado, iniciando-se por meio de estudos monográficos até a elaboração dos primeiros atlas regionais, que possibilitaram a cartografação das variantes linguísticas em alguns estados da federação, como Minas Gerais e Paraíba.

3.3 Um Breve Panorama da Dialetologia no Brasil

Feitas essas considerações, apresentamos um breve panorama da Dialetologia no Brasil. A história desse ramo da Linguística em nosso país tem como ponto de partida a contribuição de Domingos Borges de Barros, o Visconde de Pedra Branca, que realizou um estudo comparativo entre o português do Brasil e o de Portugal, em 1826. Entretanto, o marco inicial da Dialetologia no Brasil consolidou-se com a publicação da obra *O Dialeto Caipira*, de Amadeu de Amaral, em 1920.

A publicação de *O Dialeto Caipira* (Amaral, 1920) confrontou o estereótipo preconceituoso vigente à época, buscando legitimar a fala do interior do Brasil, sobretudo dos moradores da zona rural, em comparação à variante de maior prestígio, a norma culta da língua portuguesa. A edição dessa obra suscitou um movimento liderado por linguistas e estudiosos com o propósito de mapear os falares de todas as regiões do país sob o ponto de vista fonético, morfossintático e lexical.

Trabalhos monográficos, como o de Amaral (1920), prepararam o caminho para a elaboração de um atlas nacional. Esse desejo só se tornou possível, contudo, com a publicação do Decreto n. 30.643, de 20 de março de 1952, que contou com a participação de Serafim da Silva Neto, Celso Cunha e Antenor Nascentes. Esses estudiosos empenharam-se em estabelecer as orientações metodológicas necessárias para concretizar o projeto. Nesse sentido, Cardoso (1999, p. 233-255) argumenta:

> Posta na letra da lei, não bastaria isso para que se viesse, de fato, a ter um outro momento na história dos estudos dialetais no Brasil. Era necessário que uma nova visão se introduzisse na abordagem dos fenômenos da variação lingüística no país. E isso aconteceu graças ao trabalho de figuras pioneiras, das quais destaco Antenor Nascentes, Serafim da Silva Neto, Celso Cunha e Nelson Rossi, que se empenharam na implantação de um novo momento para a Dialectologia brasileira: o início dos estudos de geografia lingüística.

Baseados nas orientações metodológicas de Silva Neto (1957) e Nascentes (1958, 1961), surgiram os primeiros atlas monodimensionais publicados no Brasil, que se dedicaram à cartografia de dados exclusivamente na dimensão diatópica: o *Esboço de um Atlas Lingüístico de Minas Gerais* (Ribeiro *et al.*, 1977), o *Atlas Lingüístico da Paraíba* (Aragão; Menezes,1984) e o *Atlas Lingüístico e Etnográfico da Região Sul do Brasil* (Koch; Klassmann; Altenhofen, 2002). Este último abrange os estados do Paraná, Santa Catarina e Rio Grande do Sul em sua totalidade.

3.4 Dialetologia Bidimensional

Os primeiros estudos dialetológicos preocupavam-se apenas com os aspectos diatópicos, principalmente na Europa, isto é, a distribuição espacial das variantes. Por essa razão, sua dimensão é denominada Dialetologia Monodimensional (variação diatópica, horizontal ou espacial), sem consideração inicial de outros fatores extralinguísticos. Percebendo a necessidade de investigar outras variáveis, estudiosos passaram a explorar novas possibilidades, dando ensejo à Dialetologia Bidimensional.

No Brasil, a Dialetologia Bidimensional inicia-se com o *Atlas Prévio dos Falares Baianos* – APFB (Rossi, 1963), ainda de forma não sistematizada, já que não havia regularidade da variável sexo em todos os pontos. Assim, havia localidades em que ambos os informantes eram do sexo masculino e outras em que os dois eram do sexo feminino. Também se registraram pontos com apenas um informante e outros com mais de dois, de acordo com a Carta 28 do APFB (Figura 6, p. 59).

Também integram essa classificação o *Atlas Linguístico de Sergipe* – ALS (Ferreira *et al.*, 1987) e o *Atlas Linguístico do Paraná* – ALPR (Aguilera, 1994), nos quais, além da diatopia, contempla-se a dimensão diassexual, com a inserção sistemática da fala masculina e feminina em todos os pontos.

Conforme Cardoso (2010, p. 154), o *Atlas Linguístico de Sergipe* deu continuidade ao projeto dialetológico iniciado na Bahia, em razão da proximidade geográfica e da facilidade de acesso. Esse projeto foi conduzido pela mesma equipe responsável pelo APFB e integrou-se à região do falar baiano, segundo a classificação proposta por Nascentes ([1922] 1953).

O *Atlas Linguístico do Paraná* (Aguilera, 1994) constitui uma pesquisa geolinguística que examina a variação linguística no estado, adotando, na seleção dos informantes, além da variável diatópica, a variável sexo.

3.5 Dialetologia Pluridimensional

Os estudos dialetológicos, historicamente, pautaram-se em três métodos: o monográfico, o lexicográfico e o cartográfico. Este último teve extraordinário desenvolvimento durante o século XX, sobretudo nos estudos românicos, e pressupõe o registro em mapas especiais, com número relativamente elevado de formas

linguísticas (fônicas, léxicas ou gramaticais), comprovadas mediante pesquisa direta e unitária em uma rede de pontos de determinado território (Coseriu, 1991, p. 103).

O avanço das pesquisas linguísticas mostrou que o escopo da Dialetologia não poderia se limitar apenas à descrição diatópica, restrita aos falares rurais, uma vez que o homem, atualmente, está inserido em um contexto mais dinâmico e urbanizado, no qual estabelece múltiplas relações sociais.

Segundo Thun (2000, p. 408), a Dialetologia Pluridimensional diferencia-se pela ampliação do campo de estudo da Dialetologia Tradicional, promovendo uma análise mais profunda das variações linguísticas e investigando, além da variável diatópica, variáveis sociais (sexo, faixa etária, escolaridade, classe social, entre outras). Essa abordagem multidimensional possibilita uma visão mais completa e detalhada do espaço linguístico, enriquecendo a compreensão das variações e dinâmicas presentes nas línguas.

No Brasil, a Dialetologia Pluridimensional teve início com o *Projeto Atlas Linguístico do Brasil* – ALiB, no final da década de 1990, após a edição de atlas estaduais que abriram caminhos para a elaboração de um trabalho de amplitude nacional, capaz de mapear os falares de diversas regiões do país.

Com o advento do ALiB, a dialetologia brasileira entrou em nova fase, incorporando outras variáveis, como sexo, faixa etária e escolaridade. Desse modo, projetos estaduais passaram a empregar novas tecnologias e a ampliar o perfil dos informantes.

Para Altino (2007), o *Atlas Linguístico do Pará* – ALISPA, que integra o *Atlas Geo-sociolinguístico do Pará* (Razky, 2005), é classificado como pluridimensional, pois congrega, além da variável diatópica, as variáveis sexo, faixa etária e escolaridade. Sua abordagem é inovadora na organização dos dados, utilizando formatos textuais e audiovisuais.

No ALISPA, a pesquisa foi conduzida com 40 informantes, divididos entre duas faixas etárias (18 a 30 anos e 40 a 70 anos), de ambos os sexos e com escolaridade até a 4ª série do ensino fundamental, abrangendo dez localidades no estado do Pará. Esse formato permite uma análise aprofundada da variação linguística, especialmente nos aspectos fonéticos, e abre novas possibilidades para o estudo dialetológico. A combinação de dados textuais e sonoros representa uma evolução metodológica significativa, promovendo uma compreensão mais precisa dos fenômenos linguísticos

e consolidando uma nova fase nas investigações da variação e mudança em contextos regionais e globais.

Surgiu, então, a necessidade de agregar aos estudos dialetológicos os princípios da Sociolinguística variacionista, iniciada por William Labov e outros estudiosos, na década de 1960, que definiram, como objeto de estudo, a língua em uso em sua vertente falada, permeada por aspectos socioculturais que refletem, direta ou indiretamente, no modo de falar das comunidades.

Para Labov ([1972] 2008, p. 22), a língua precisa ser estudada em sua essência, diversidade e realidade, pois não é homogênea, estável ou imutável. O estudo da variação e mudança deve considerar também os fatores sociais (idade, lugar, classe social, sexo), determinantes na consolidação da diversidade linguística, consubstanciada em sua heterogeneidade intrínseca.

Os sociolinguistas advertem, porém, que os fatores sociais não exercem a mesma influência em todas as comunidades. Assim, nível social, situação econômica, idade e sexo podem ter pesos distintos, de modo que, em uma localidade, prevaleça a *variável idade* e, em outra, a *variável sexo*. Nesses termos, esclarece Moreno Fernández (1998, p. 34):

> Geralmente, os fatores sociais que mostram uma maior capacidade de influência sobre a variação linguística são o sexo, a idade, o nível de instrução, o nível sociocultural e a etnia, entre outros que também hão de interessar.[3]

Moreno Fernández descreve o grau de influência dessas *variáveis sociais* dentro de uma comunidade linguística e ressalta o maior peso de uma em relação a outra, de acordo com o contexto. Isso demonstra que a heterogeneidade de uma língua decorre tanto de fatores linguísticos (classe gramatical da lexia, natureza do fonema, tonicidade da vogal em análise, entre outros) quanto de fatores extralinguísticos (idade, região, sexo). Tais elementos têm papel fundamental para a Sociolinguística e para a Dialetologia Pluridimensional, já que língua e sociedade estão intrinsecamente relacionadas, de modo que uma reflete diretamente a outra. Em outras palavras, a língua é um fenômeno social e traz consigo características inerentes à sociedade a que serve. Assim, se a sociedade é mais *urbana* ou *rural*,

[3] “Generalmente, los factores sociales que muestran una mayor capacidad de influencia sobre la variación lingüística son el sexo, la edad, el nivel de instrucción, el nivel sociocultural y la etnia, entre otros que también han de interesar.”

mais *conservadora* ou *liberal*, mais *letrada* ou *analfabeta*, assim também será a linguagem de seus falantes

3.6 GEOLINGUÍSTICA NA EUROPA E NO BRASIL

A Geolinguística surge como um campo interdisciplinar que integra a Linguística laboviana, a Dialetologia e a geografia humana e física. Com foco exclusivo nos objetivos linguísticos, esse ramo busca compreender as interações entre linguagem e espaço, ressaltando a influência desses aspectos na manifestação e distribuição das variações linguísticas. Trudgill e Campoy (2007, p. 160) a descrevem como

> [...] uma área de estudo multidisciplinar e relativamente recente que se ocupa das relações existentes entre a linguagem e a geografia, representando uma síntese da linguística secular laboviana, a dialetologia e a geografia humana e cujos objetivos são puramente linguísticos.

Os estudos dialetais, fundamentais para a compreensão da diversidade linguística, tiveram início no século XIX, com metodologias pioneiras, mas também com limitações significativas. Esse panorama inicial da Geolinguística revela os desafios e avanços que moldaram a pesquisa linguística moderna.

Segundo Chambers e Trudgill (1994), os estudos dialetais com métodos da geografia linguística tiveram início na Alemanha, em 1876, sob a liderança de Georg Wenker. O pesquisador enviou frases no alemão padrão a professores do norte do país, solicitando que as adaptassem a seus respectivos dialetos locais. Entre 1877 e 1887, Wenker reuniu aproximadamente 45.000 respostas de quase 50.000 participantes, identificando variantes regionais específicas em cada frase. Contudo, o imenso volume de dados tornou a análise inviável, obrigando-o a restringir o estudo a áreas delimitadas e a palavras selecionadas.

De acordo com Cardoso (2010, p. 41), a pesquisa de Wenker recebeu críticas pela ausência de metodologia apropriada para retratar a diversidade linguística da Alemanha. Entre as limitações apontadas, destacam-se a não consideração das variáveis sociais (faixa etária, sexo e escolaridade) e o fato de os informantes serem apenas professores letrados, desconsiderando grupos com menor instrução. Além disso, a coleta indireta dos dados, feita por correspondência, comprometeu a fidedignidade de respostas de natureza fonética.

Para Cardoso (2010, p. 42), o *Atlas Linguistique de la France* (ALF), desenvolvido entre 1902 e 1910 por Jules Gilliéron, com apoio de Edmond Edmont, representou um avanço no campo da Geolinguística. A pesquisa abrangeu 639 localidades e utilizou um questionário inicialmente com 1.400 perguntas, ampliado para 1.900, com ênfase na morfologia verbal.

Mesmo sem formação em Linguística e utilizando apenas uma bicicleta, Edmond Edmont percorreu as 639 localidades e, em quatro anos, concluiu a pesquisa, cujo prazo inicial era de cinco. Nesse período, entrevistou informantes e registrou minuciosamente os dialetos franceses, provençais e franco-provençais (Altino, 2007, p. 20).

Apesar de sua relevância histórica, o ALF foi acusado de incompletude, sobretudo pela omissão de termos importantes para a compreensão mais ampla das variedades linguísticas. Além disso, a ausência de considerações sobre diferenças sociais que influenciam o uso da língua limitou o alcance das conclusões do atlas.

No Brasil, o marco da Geolinguística ocorreu com a publicação do Decreto n. 30.643, de 20 de março de 1952, que determinou a elaboração de um *Atlas Linguístico do Brasil* (ALiB). A partir dessa publicação, vários atlas regionais foram produzidos, abrindo caminho para que, em 1996, o Projeto ALiB fosse retomado em nível nacional, com o objetivo de mapear as variações linguísticas em todas as regiões do país.

Portanto, a Geolinguística – ramo da Dialetologia que se ocupa da representação geográfica das variantes linguísticas por meio de cartas – possui importância fundamental no Brasil e no mundo, mostrando-se imprescindível para compreender a relação entre língua, espaço e sociedade. Esse campo permite identificar, descrever e representar fenômenos fonéticos, morfossintáticos e lexicais de uma região, país ou continente.

3.7 Uma Abordagem Panorâmica da Dialetologia no Brasil

De acordo com Mota e Cardoso (2006), a Dialetologia no Brasil pode ser dividida em três fases distintas.

A primeira fase perdurou por aproximadamente um século, de 1826 a 1920. Como já salientado, seu protagonismo deve-se a Domingos Borges de Barros, Visconde de Pedra Branca, que buscou descrever o português falado no Brasil e suas peculiaridades lexicais.

A segunda fase iniciou-se em 1920, com a publicação da obra *O Dialeto Caipira*, de Amadeu Amaral, que analisou o dialeto caipira falado principalmente no interior de São Paulo e em localidades vizinhas, ressaltando aspectos fonéticos, sintáticos e lexicais. Pouco depois, em 1922, Antenor Nascentes publicou *O Linguajar Carioca* e, em 1934, Mário Marroquim lançou *A Língua do Nordeste*. Nesse contexto, em 1952, foi editado o Decreto n. 30.643, viabilizando a elaboração de um atlas linguístico que retratasse o português falado no Brasil, o que se consolidaria mais tarde com o Projeto ALiB.

A terceira fase corresponde à concretização do mencionado decreto, com a publicação dos primeiros atlas estaduais, como o *Atlas Prévio dos Falares Baianos* – APFB (1963), o *Atlas Linguístico de Minas Gerais* – ALMG (1977), o *Atlas Linguístico da Paraíba* (1984), o *Atlas Linguístico de Sergipe* – ALSE (1987), o *Atlas Linguístico do Paraná* – ALPR (1994) e o *Atlas Linguístico e Etnográfico da Região Sul do Brasil* – ALERS (2002), entre outros.

A elaboração de atlas estaduais ou regionais serviu de parâmetro para a criação de um projeto de amplitude nacional: o *Atlas Linguístico do Brasil* (ALiB), implementado no final do século XX na Universidade Federal da Bahia (UFBA), por uma comissão formada por dialetólogos e sociolinguistas. O projeto contempla todas as regiões do país, com uma rede de pontos composta por 250 localidades, promovendo um estudo geolinguístico pluridimensional, ao investigar particularidades fonéticas, morfossintáticas e lexicais do português falado no Brasil junto a informantes estratificados segundo sexo, escolaridade e faixa etária.

3.8 O Pioneirismo do Dialeto Caipira de Amaral

O Dialeto Caipira (Amaral; Duarte, 1955, p. 16), marco dos estudos linguísticos no Brasil, retrata o português falado na zona rural do interior de São Paulo. Essa obra apresenta as principais características dessa variação diatópica, assim distribuídas: 1. *aspectos fonéticos*: a exemplo dos rotacismos (troca do *l* pelo *r*); 2. *estrutura morfossintática*: incidência sobre a morfologia e a sintaxe, como a flexão verbal e o uso dos pronomes; 3. *aspectos lexicais*: ocorrências no vocabulário, em que se encontra um verdadeiro inventário de palavras, muitas delas de origem indígena e africana, além de arcaísmos ainda presentes.

Esta última característica merece destaque, por ser objeto da presente pesquisa. Acerca do vocabulário, o autor assim se exprime

:

> São em grande número, relativamente à extensão do vocabulário dialetal, as formas esquecidas ou desusadas na língua. Lendo-se certos documentos vernáculos dos fins do século XV e de princípios e meiados (sic) do século XVI, fica-se impressionado pelo ar de semelhança da respectiva linguagem com a dos nossos roceiros e com a linguagem tradicional dos paulistas de boa família, que não é senão o mesmo dialeto um pouco mais polido (Amaral; Duarte, 1955, p. 16).

Essa obra é de grande importância para os estudos dialetológicos no Brasil, dado o seu pioneirismo, sua compreensão da variação linguística e da identidade cultural do interior de São Paulo. Ela evidencia as peculiaridades do falar rural, mais conservador em virtude do isolamento geográfico em relação aos grandes centros urbanos e do insuficiente acesso a meios de comunicação (jornal, rádio e revista), precários e pouco difundidos nessas localidades. Esse falar, que resiste ao tempo com seus costumes, cultura e linguagem específica, foi muitas vezes rechaçado por intelectuais da época, por ser considerado um desacato à norma culta da língua portuguesa. O autor, entretanto, condena veementemente tal análise preconceituosa e enaltece o falar rural como fenômeno natural de toda língua.

Em síntese, o Brasil revela marcas contrastantes entre o rural e o urbano, decorrentes de fatores históricos, econômicos e étnicos, que se refletem na formação cultural e social de seus “Brasis”, fruto da miscigenação e dos processos de imigração, mais intensos no Sul do país. Tais fatores repercutem na variação linguística de seus falantes.

3.9 Um Retrato da Diatopia Lexical nas Regiões Sul e Nordeste do Brasil

Tratamos, nesta seção, da diatopia lexical no Sul e no Nordeste, com base em estudos de Aguilera (2015), Almeida-Baronas (2012), Rodrigues (2007) e Oliveira (2001), retratando as transformações no espaço e suas influências na variação linguística, que refletem a identidade e os modos de vida dos falantes.

Aguilera (2015) analisa o falar no estado do Paraná, discutindo as peculiaridades das variantes rural e urbana para revelar as diferenças, sob o ponto de vista lexical, na mudança do eixo rural–urbano.

A fim de verificar as alterações diatópicas no Paraná, a autora apresenta um quadro comparativo de questões do campo semântico da fauna presentes no *Atlas Linguístico do Brasil* (ALiB-PR), que têm cartas correspondentes no *Atlas Linguístico*

do Paraná (ALPR): QSL nº 64 “urubu”, QSL nº 65 “beija-flor”, QSL nº 78 “boi sem chifre” e QSL nº 86 “bicho-de-fruta”. Conforme já salientado, o ALPR é um atlas rural, pois seus informantes têm origem no campo (agricultores, pecuaristas) e, consequentemente, maior familiaridade com o acervo linguístico agropastoril.

Para ilustrar essas diferenças, a autora lista uma série de lexias, em forma de texto, relacionadas ao homem do campo, a fim de demonstrar que dificilmente um jovem da cidade compreenderia o seu conteúdo.

> O piá, meio lonanco, que ficou com jojoca depois de comer as bananas inconhas e a coruja embrulhada em folhas de bananeira, atravessou, de vereda, a sanga com água pelo garrão, querendo alcançar o arco-da-velha, mas, com medo dos fuzilos e dos calmariões, voltou para casa. À noite, caminhando devarde pelos carreiros, buscou no céu o caminho-de-adão-e-eva, mas só conseguiu ver o criso da luma que ia chorungar toda a plantação de feijão (Aguilera, 2015, p.19).

Segundo Aguilera (2015), o Paraná, desde a década de 1950, vem passando por um processo de urbanização, deixando de ser um estado agrícola para se tornar urbano. Para a autora, a sociedade mudou: os campos estão vazios e as periferias, cheias. Sustenta ainda que a informática já está presente na vida das pessoas, desde as casas até as escolas. Assim, a linguagem sintetiza os costumes, as tradições e a vida de uma comunidade.

As transformações pelas quais passou a sociedade paranaense repercutem no léxico de seus informantes e, consequentemente, palavras antes correntes tornaram-se desconhecidas, mesmo no vocabulário passivo.

> Assim é que, nos falares paranaenses, lexias antes tão produtivas como caminho de Santiago, arco-da-velha, arco-da-aliança, calmarião, tormenta, fuzilo, burcão, sanga, arroio, cuitelo, corvo, galinha polaca, galinha peca, banana inconha, lonanco, náfico, capela do olho, jojoca, pechada, ramona, e até mesmo o luvisome e o boitatá estão desaparecendo do vocabulário até mesmo passivo de nossos contemporâneos (Aguilera, 2015, p. 31).

Para Aguilera (2015), uma versão atual da linguagem do homem do campo, transposta para a fala urbana, ficaria assim:

> O piá ou menino, meio manco, que ficou com soluço depois de comer as bananas gêmeas (ou felipes) e a pamonha (não mais embrulhada em folha de bananeira, mas saboreada na pamonharia), atravessou, apressadamente o riozinho, com água nos calcanhares, querendo alcançar o arco-íris, mas, com medo dos raios e trovões, voltou para casa. À noite, caminhando sossegado pelos trilhos, buscou, no céu, a via-láctea, mas só conseguiu ver o eclipse da lua, que iria murchar toda a plantação de feijão (Aguilera, 2015, p. 31-32).

Sob os pontos de vista diacrônico e diatópico, a autora destaca o apagamento de muitas variantes linguísticas, outrora correntes no falar rural paranaense, em razão do processo de urbanização a partir de 1950. Esse processo trouxe reflexos para o acervo linguístico, cada vez mais inovador, dado o dinamismo e as mutações próprias das línguas e, em especial, do falar paranaense.

Almeida-Baronas (2012), em *Dialeto Caipira e Urbanização: uma química que não se mistura (?)*, apresenta traços linguísticos que caracterizam o linguajar caipira como tradicional e inovador ao mesmo tempo. Acerca das marcas tradicionais, a autora afirma:

> O falar rural apresenta algumas marcas que evidenciam sua relação com o passado, pois traz, em suas formas diversas, alguns traços da língua portuguesa em sua formação, os quais revelam resquícios de outras línguas, como o latim, línguas africanas e línguas indígenas (Almeida-Baronas, 2012, p. 15).

Todavia, a variante caipira atual apresenta marcas de urbanização, em razão de transformações como a chegada dos meios de transporte e comunicação (rádio, televisão, celular, jornal), bem como o acesso a recursos médicos e à educação – este último determinante para a aproximação com a norma culta da língua portuguesa.

> Constatou-se que o falar dos informantes do corpus apresenta marcas do chamado dialeto caipira, evidenciando as transformações ocorridas na língua portuguesa a partir de sua origem latina, mas que também se compõem de pistas de urbanização, com elementos lexicais referentes ao cotidiano urbano e elementos gramaticais que se aproximam da norma culta (Almeida-Baronas, 2012, p. 21).

Portanto, para a autora, o dialeto caipira não se separa totalmente da norma culta. O que existe é uma aproximação, uma vez que foram encontradas marcas de urbanização no falar caipira.

De igual relevância é a tese de Rodrigues (2007), que analisa a interação do vocabulário rural do Paraná com a sua história social, a partir da Geografia Linguística. O *corpus* do trabalho é composto por noventa cartas lexicais do *Atlas Linguístico do Paraná* (Aguilera, 1994). A autora associa registros do léxico a fontes históricas e a relatos de viagens do século XIX, revelando três áreas lexicais bem definidas: no Norte, a influência de mineiros, paulistas e nordestinos; no Sul, a de catarinenses e gaúchos; e no Centro-Norte, uma zona de transição. Além disso, no Paraná Tradicional encontram-se marcas do português antigo e variantes de origem tupi.

Segundo Rodrigues (2007), a linguagem rural vem passando por transformações em virtude da escolarização, da maior proximidade com o meio urbano e do acesso à televisão. Esses fatores acarretam alterações na pronúncia, incorporações e desuso de lexias. Por isso, a documentação dessa linguagem é fundamental para a sua preservação.

Até os anos 1970, a população rural do Paraná era expressiva, sobretudo em razão das lavouras de café, que demandavam intensa mão de obra. A cafeicultura, enquanto monocultura, foi responsável por manter o homem no campo e promover o desenvolvimento, especialmente na região Norte do estado.

Entretanto, na segunda metade dessa década, o cenário se transformou, culminando em intenso êxodo rural, agravado pelas fortes geadas, em especial a de 1975, que derrubou a lucratividade e a produtividade da cafeicultura. A destruição dos cafezais levou os produtores a diversificar os investimentos em outras lavouras, como a soja, que, diferentemente do café, não demandava tanta mão de obra (Oliveira, 2001).

Consequentemente, essa mudança do campo para a cidade trouxe reflexos para a linguagem paranaense, já que alguns termos são mais frequentes no meio rural e outros no urbano, retratando o peso da variação diatópica. Nesse sentido, esclarece Rodrigues:

> Foi possível entender que, das trinta cartas preparadas a partir do questionário do Atlas Lingüístico do Brasil – ALiB (1998), quatorze coincidem com as do Atlas Lingüístico do Paraná, quer dizer, podem ser comparadas. Entre as não-coincidentes, estão algumas que podem ser relacionados mais estreitamente ao mundo rural, como *jacá, gambá, anca* e *borralho;* aquelas que fazem parte da realidade urbana: *sinaleiro, lombada, lote (terreno/data)* e os que permeiam as duas dimensões: *mosca varejeira, seios, fantasma, aguardente, toco de cigarro, isqueiro, bêbado, bodega, grampo (sem pressão)* (Rodrigues, 2007, p. 134).

A partir da análise desse acervo sobre a variação linguística no Paraná, inferimos que fatores climáticos, históricos, geográficos e econômicos incidiram na consolidação do falar paranaense, ora prevalecendo aspectos de ruralidade, ora refletindo o peso da urbanização. Agora, analisaremos como essa variação se processou na Bahia, estado pertencente ao Nordeste, região com características e formação distintas em relação ao Sul.

Oliveira (2014) também contribui para esta dissertação, pois sua pesquisa tem como *corpus* dados do Projeto ALiB, referentes aos estados da Bahia e do Paraná,

consubstanciados pelas perguntas de 1 a 33 do Questionário Semântico-Lexical (QSL) do ALiB, relativo às áreas semânticas acidentes geográficos, fenômenos atmosféricos, astros e tempo em ambas as localidades. A autora salienta:

> Descritas as realidades da Bahia e do Paraná, neste item procuramos analisar, de forma comparativa as duas áreas estudadas. O Quadro 22 mostra as lexias e expressões fraseológicas coincidentes nas cidades do interior da Bahia e Paraná, bem como lexias fornecidas apenas por informantes do interior da Bahia ou do interior do Paraná [...].

A pesquisa de Oliveira discute a diversidade linguística entre dois estados, um do Nordeste e outro do Sul, consolidando nossa expectativa de que aspectos diatópicos e sociais moldam o acervo linguístico de comunidades pertencentes a um país de dimensões continentais como o Brasil.

Portanto, trabalhos dialetológicos sobre a variação rural–urbana, como os de Aguilera (2015), Almeida-Baronas (2012), Rodrigues (2007) e Oliveira (2014), corroboram nossa pesquisa no sentido de que a diatopia exerce papel determinante no léxico, uma vez que a linguagem do campo preserva mais marcas do passado, da tradição e dos costumes, diferentemente da linguagem da cidade, mais dinâmica, inovadora e propensa a mudanças.

3.10 Descrição das Fontes

Nesta seção, procedemos à descrição de cada uma das fontes, priorizando suas principais características, como a origem, a rede de pontos, os informantes e as variáveis envolvidas. Inicialmente, analisamos os dados dos atlas estaduais – o *Atlas Prévio dos Falares Baianos* (APFB) e o *Atlas Linguístico do Paraná* (ALPR). Em seguida, tratamos dos dados inéditos do *Projeto Atlas Linguístico do Brasil* relativos à Bahia e ao Paraná.

3.11 Atlas Prévio dos Falares Baianos

Seguindo as diretrizes metodológicas propostas por Silva Neto (1957) e Nascentes (1958; 1961), surgiram os primeiros atlas linguísticos estaduais, com destaque para dois estudos fundamentais que compõem o escopo desta pesquisa: o *Atlas Prévio dos Falares Baianos* (APFB), publicado por Rossi em 1963, e os *Atlas Linguísticos do Paraná* (Aguilera, 1994; Altino, 2007).

O APFB representa um marco importante no estudo da variação linguística na Bahia, destacando-se como o primeiro projeto voltado ao mapeamento dos falares regionais do estado. A pesquisa conduzida por Nelson Rossi baseou-se na metodologia desenvolvida por Silva Neto e Nascentes, com coleta de dados realizada por meio de inquéritos diretos para registrar as particularidades linguísticas de diferentes localidades. Seu objetivo consistiu em descrever as variações lexical e fonética presentes no português falado na Bahia.

O APFB contemplou informantes de origem rural. Foram coletados dados em 50 localidades, com 100 informantes, entre 25 e 84 anos, analfabetos ou semianalfabetos, de ambos os sexos. Sobre esse atlas, Aragão (2008, p. 5) descreve:

> O Questionário contém 164 questões nos Campos Semânticos: agricultura, pecuária, anatomia e fisiologia humana, culinária e alimentação, geografia e astronomia. As cartas são em número de 209, compreendendo 11 de identificação, 154 fonéticas e léxicas e 44 cartas resumo. Os termos vêm transcritos no interior da própria carta ou com legendas e símbolos, em preto e branco e coloridos [...].

Os pontos do APFB distribuem-se de leste a oeste e de norte a sul. Assim, o ponto 1 – Abadia – situa-se no Nordeste do estado, enquanto o ponto 50 – Ibiranhém – localiza-se no Sul baiano. A Figura 1 mostra a distribuição dos pontos investigados pelo APFB.

Figura 1 – Carta V do APFB – Pontos linguísticos

Fonte: Banco de dados do APFB.

3.12 Descrição do Atlas Linguístico do Paraná e do Atlas Linguístico do Paraná II

O ALPR é resultado da tese de doutoramento de Aguilera (defendida em 1990 e publicada em 1994). O questionário, composto por 318 perguntas indiretas, contempla dois grandes campos: a Terra e o Homem. O primeiro subdivide-se em: a) natureza, fenômenos atmosféricos, astros, tempo etc.; b) flora: árvores, frutos etc.; c) plantas medicinais; e d) fauna: aves, pássaros, animais etc. O segundo campo envolve: a) partes do corpo, funções, doenças etc.; b) vestuário e calçados; c) agricultura, instrumentos agrícolas etc.; e d) brinquedos e jogos infantis. Trata-se de um atlas bidimensional, uma vez que, no momento da coleta dos dados, foram consideradas duas variáveis: a diatópica (o lugar) e a diassexual (sistematicamente, um homem e uma mulher em cada localidade).

Os dados do ALPR foram coletados entre 1987 e 1990, em 65 pontos do estado do Paraná. Em cada um deles, foram ouvidos dois informantes adultos da zona rural, analfabetos ou semianalfabetos, sendo um homem e uma mulher (portanto, bidimensional), com idade entre 30 e 60 anos. O questionário, baseado em Caruso (1983), compõe-se de 318 questões que versam sobre a Terra e o Homem. Nas palavras de Aguilera (2015, p. 101-119),

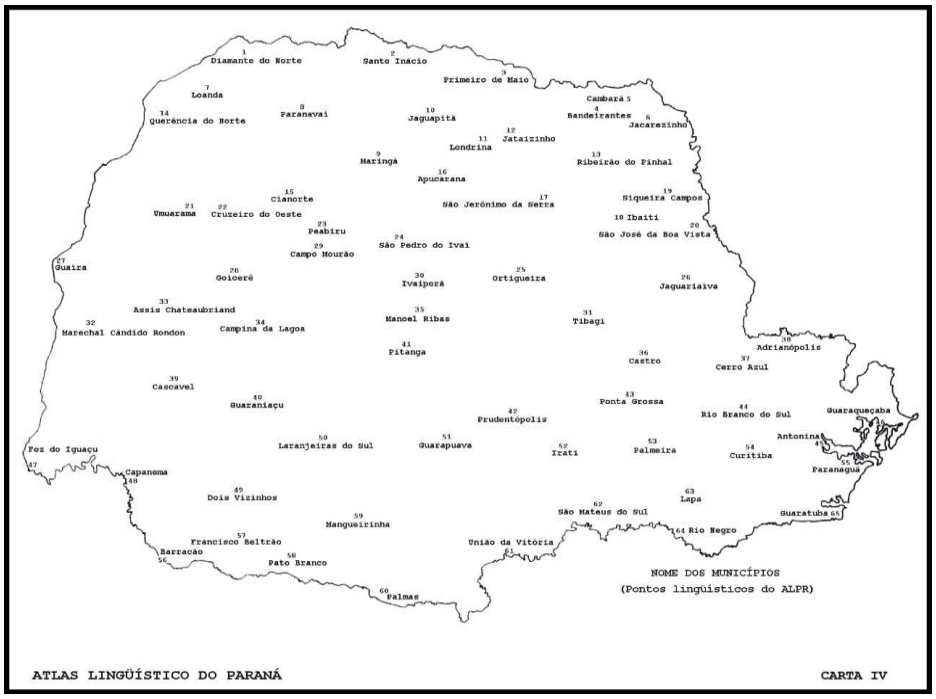
> [...] após alguns ajustes, o questionário, composto por 318 questões, diretas e indiretas, distribuídas em dois campos semânticos principais: Homem e Terra que, por sua vez, envolvem subcampos: natureza (57 questões), flora (37); plantas medicinais (15); vida selvagem (50); o corpo e as doenças (105); vestuário (14); agricultura e implementos agrícolas (29); brinquedos e jogos (13). Além do questionário, há cinco histórias sobre traços culturais regionais, crenças populares e experiências pessoais: o corpus registrado responde a um total de 325 questões. As gravações foram transcritas e impressas: estão reunidas em 13 volumes inéditos, com aproximadamente 2.000 páginas.

Já em relação ao *Atlas Linguístico do Paraná II*, Altino deu continuidade aos trabalhos do ALPR, mapeando as respostas das questões coletadas por Aguilera que não haviam sido contempladas no primeiro atlas. Nesse sentido, Altino (2012, p. 812-832) informa:

> A exemplo do primeiro volume, as cartas apresentam informações semântico-lexicais, fonético-fonológicas e notas explicativas para reproduzir as variantes registradas nos 65 pontos de estudo. Além de fornecer subsídios para outras pesquisas, o volume dois do ALPR oferece aos estudantes e pesquisadores da língua portuguesa um instrumento de acesso a dados dialetológicos e geolinguísticos mais amplos sobre a linguagem rural paranaense.

Para o ALPR, Aguilera (1994) e Altino (2007) elaboraram cartas analíticas, mistas e sintéticas. Os pontos distribuem-se do oeste para o leste numa ordem crescente, isto é, do menor para o maior. Assim, o ponto 1 – Diamante do Norte – situa-se no Noroeste do estado, e o ponto 65, em Guaratuba, na Região Leste Paranaense. A Figura 2 mostra a distribuição dos pontos investigados pelo ALPR.

Figura 2 – Carta IV do ALPR – Pontos linguísticos do ALPR



Fonte: Aguilera (1994).

3.13 A IMPORTÂNCIA DO ALPR PARA OS ESTUDOS DIALETOLÓGICOS NO BRASIL E NO MUNDO

O ALPR (Aguilera, 1994) foi o quinto atlas linguístico regional publicado no Brasil, depois do *Atlas Prévio dos Falares Baianos* (Rossi, 1963), do *Esboço de um Atlas Linguístico de Minas Gerais* (Ribeiro *et al.*, 1977), do *Atlas Linguístico da Paraíba* (Aragão; Menezes, 1984) e do *Atlas Linguístico de Sergipe* (Ferreira, 1987). Pioneiro na Região Sul, dedicou-se à descrição pormenorizada da fala rural, com coleta de dados em 65 localidades paranaenses, escolhidas com base em critérios de relevância histórica, geográfica e sociocultural. Seu estudo promoveu uma verdadeira imersão na diversidade linguística, abrangendo aspectos fonético-fonológicos e semântico-lexicais.

O ALPR apresenta uma fotografia da diversidade linguística no Paraná, retratando diferentes facetas étnico-raciais (indígenas, imigrantes europeus, migrantes de outras regiões) na consolidação do falar paranaense. O estudo demonstrou que o português brasileiro apresenta variação linguística mesmo dentro de um único estado, rompendo, assim, com a visão de homogeneização da língua em espaços limitados. Serviu, ainda, de base e fonte metodológica para inúmeros trabalhos científicos, dissertações, teses e outros atlas estaduais.

Nessa perspectiva, o ALPR II (Altino, 2007) teve como finalidade dar continuidade aos estudos iniciados por Aguilera (1994) e não contemplados no primeiro atlas, cartografando dados inéditos referentes a mais de 50% do questionário original. À semelhança de Aguilera (1994), as cartas descrevem informações fonético-fonológicas e semântico-lexicais, além de notas sobre ocorrências únicas. As respostas dos informantes são representadas pelos símbolos da biologia: (♀) e (♂), para identificar, respectivamente, o sexo feminino e o masculino.

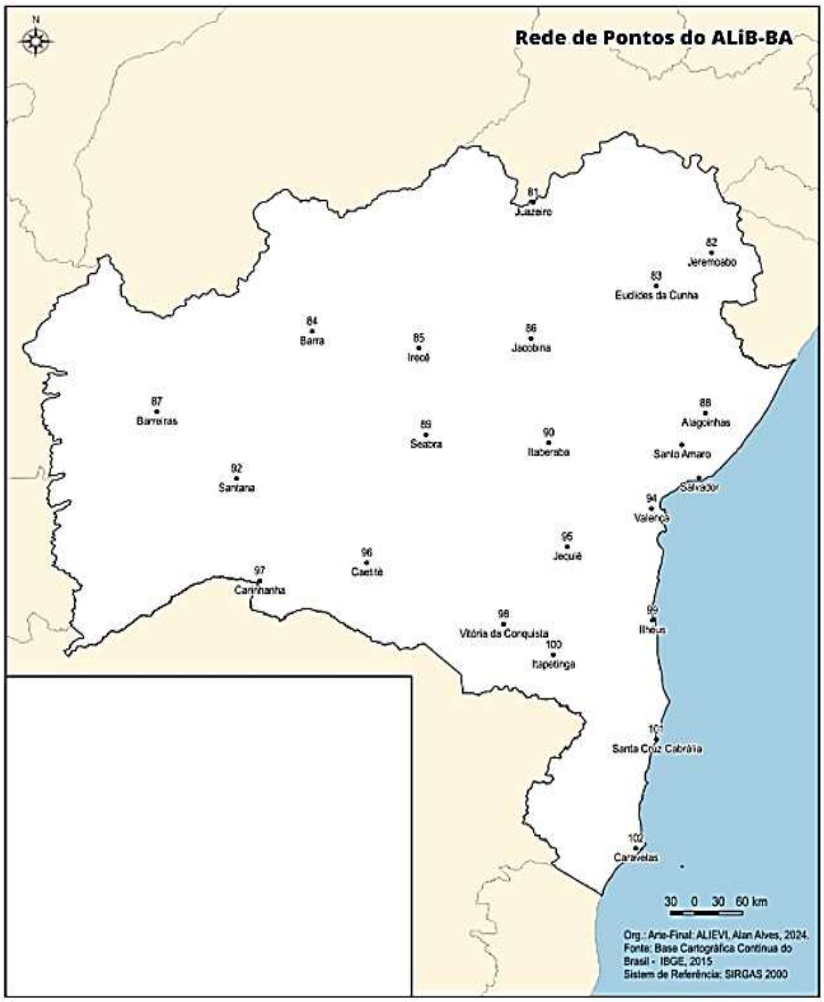
Portanto, o ALPR constitui para estudantes de Letras, professores, mestrandos, doutorandos e dialetólogos uma rica fonte de pesquisa acerca da variação linguística no Paraná, nos aspectos fonético-fonológico e semântico-lexical. Sua relevância reside também na preservação de um acervo lexical que revela as transformações pelas quais passou a sociedade paranaense na mudança do eixo rural-urbano promovida pelo êxodo rural, reverberando no léxico de seus falantes.

3.14 FONTES DO *PROJETO ATLAS LINGUÍSTICO DO BRASIL*: BAHIA E PARANÁ

O *Atlas Linguístico do Brasil* – ALiB, projeto de amplitude nacional lançado em 1996 na Universidade Federal da Bahia e coordenado por um comitê formado por dialetólogos e sociolinguistas de várias universidades brasileiras, tem entre seus principais objetivos mapear as variantes linguísticas do português falado no Brasil e fornecer subsídios para outros estudiosos, como lexicógrafos, historiadores, sociólogos e pedagogos. O ALiB dispõe de uma rede de pontos composta por 250 localidades, distribuídas por todas as regiões do país.

A rede de pontos da Bahia para o ALiB inclui a capital e 21 localidades do interior, conforme mostra a Figura 3. Os dados foram coletados entre 2003 e 2009 pela equipe do ALiB-BA.

Figura 3 – Rede de pontos do ALiB-BA



Fonte: Banco de dados do ALiB-BA.

A rede de pontos do Paraná para o Projeto ALiB é formada pela capital e 16 localidades do interior. Os dados, ainda inéditos, foram coletados pela equipe do ALiB-PR entre dezembro de 2001 e fevereiro de 2004.[4] A Figura 4 apresenta a rede de pontos do ALiB-PR.

[4] Excetuam-se a entrevista da informante 2 de Terra Boa e a dos informantes 1 e 2 de Adrianópolis, feitas em fevereiro de 2009 e janeiro de 2010, respectivamente, devido a problemas técnicos ocorridos nas coletas anteriores.

Figura 4 – Rede de pontos do ALiB-PR

Fonte: Banco de dados do ALiB-PR.

3.15 Contribuições do Projeto ALiB Para os Estudos Dialetológicos

O ALiB, atlas pluridimensional de amplitude nacional, constitui uma fonte de pesquisa relevante para os estudos comparativos do português brasileiro falado em diversas regiões do país – estados, capitais e comunidades do interior –, revelando a variação linguística em seus aspectos fonético-fonológicos, morfossintáticos e lexicais, em uma vasta rede de pontos distribuídos pelo território nacional.

Sua elaboração representa um avanço nas pesquisas dialetológicas iniciadas em meados do século XX, quando surgiram os primeiros trabalhos cartográficos de amplitude regional, que tiveram como precursor o *Atlas Prévio dos Falares Baianos* (Rossi, 1963). O ALiB inovou ao adotar o método pluridimensional na coleta de dados, contemplando variáveis sociais (sexo, faixa etária, escolaridade) e corroborando aspectos outrora não considerados.

Os resultados de suas pesquisas podem subsidiar políticas públicas educacionais voltadas à elaboração de materiais didáticos mais fiéis à realidade linguística brasileira, com maior representatividade das variantes linguísticas das classes menos favorecidas, contribuindo de forma significativa para uma educação mais inclusiva e livre de preconceito linguístico em um país que ainda prioriza a variante de maior prestígio em detrimento de outras variantes, consideradas inferiores e estigmatizadas.

As cartas lexicais do ALiB, com dados das capitais, retratam de modo pormenorizado como uma palavra ou expressão é representada regionalmente na fala de informantes de sexo, idade e escolaridade distintas, em contextos formais e informais, descrevendo um traço identitário da pluralidade linguística da sociedade brasileira como reflexo de sua miscigenação. Como exemplo, destaca-se a distribuição espacial, nas capitais, de *penca*, *palma*, *cacho* e *concha* na carta L 06 (Cardoso *et al.*, 2014).

Sob uma perspectiva fonético-fonológica, estudos realizados pelo ALiB e apresentados na carta F04 C 6 (Cardoso *et al.*, 2014), referentes a dados das capitais, revelam que existem pelo menos quatro tipos de realização do /r/: o r retroflexo, também conhecido como *r caipira*, registrado principalmente nas capitais do Centro-Oeste; o r tepe, mais frequente em Porto Alegre, Curitiba e São Paulo; o r carioca; e o r glotal, predominante nas capitais das regiões Norte e Nordeste, além de Belo Horizonte (Sudeste) e Florianópolis (Sul).

Em uma dimensão morfossintática, a carta M 02 (Cardoso *et al.*, 2014) mostra a distribuição das formas pronominais *tu* e *você*. A primeira predomina em Porto Alegre e Florianópolis, além de ocorrer em capitais do Norte, como Macapá, Belém e Rio Branco, e em São Luís, no Nordeste. Já *você* distribui-se por todas as capitais: ora como variante categórica (Campo Grande, São Paulo, Salvador), ora predominante (Natal, Maceió, Rio de Janeiro), ora em concorrência com *tu* (Belém, Macapá, São Luís), evidenciando a diversidade linguística em um país de dimensões continentais.

Dessa forma, o ALiB, idealizado na década de 1950, concretizou-se a partir da iniciativa do Instituto de Letras da Universidade Federal da Bahia (UFBA). Amparado pelos atlas regionais que abriram os caminhos para as pesquisas dialetológicas em vários estados brasileiros, consolidou-se como um acervo vasto, eclético e de grande relevância para alunos, professores e pesquisadores do Brasil e do exterior.

4 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo, apresentamos os procedimentos metodológicos que norteiam esta dissertação.

Para a consecução dos objetivos, percorremos os seguintes passos: (i) escolha das questões a serem analisadas; (ii) seleção das fontes geolinguísticas; (iii) constituição e descrição das fontes; (iv) levantamento das variantes em planilhas de Excel, com a indicação dos espaços geográficos (estado da Bahia, localidades rurais e urbanas; estado do Paraná, localidades rurais e urbanas) e da distribuição diastrática (variável sexo no ALiB-BA, ALPR e ALiB-PR; e variável faixa etária no ALiB-BA e ALiB-PR); (v) descrição e análise quantitativa e qualitativa dos dados dos *corpora*; e (vi) comparação dos dados nas quatro fontes.

4.1 A ESCOLHA DAS QUESTÕES DO ALPR E DO ALIB

Para esta dissertação, optamos por analisar as questões semântico-lexicais dentro da subárea agropastoril. A motivação para a escolha do objeto de análise decorre de nossa vivência em diferentes fases da vida, nos contextos rural e urbano, e da observação das dissemelhanças entre as linguagens em uso nesses dois ambientes. Essas diferenças linguísticas, que emergem de interações culturais, sociais e históricas próprias de cada contexto, oferecem um campo fértil para o estudo da variação linguística e das transformações semânticas e lexicais que ocorrem quando se comparam práticas linguísticas de regiões distintas.

Dessa forma, analisamos as respostas dadas às seguintes questões do mundo agrícola: a primeira fonte refere-se às respostas que constam das cartas 28 e 28 R do Atlas Prévio dos Falares Baianos (Rossi, 1963) para *sabugo*; a segunda diz respeito às cartas 83 (*vagem*-lexical) e 139 (*girassol*) do Atlas Linguístico do Paraná (Aguilera, 1994) e às cartas 298 (*soca*) e 364 (*sabugo*) do Atlas Linguístico do Paraná II (Altino, 2007); a terceira e a quarta baseiam-se nas respostas dadas a cinco questões semântico-lexicais (QSL) do *Projeto Atlas Linguístico do Brasil* (ALiB) referentes à agricultura: *espiga, sabugo, soca, girassol* e *vagem*.

Para fins didáticos, reproduzimos a formulação dessas questões no Questionário Semântico-Lexical do ALiB (Comitê Nacional do ALiB, 2001):

i) (045) Quando se vai colher o milho, o que é que se tira do pé?;

ii) (046) E quando se tira da espiga, todos os grãos de milho, sobra o quê? Você pega a espiga, tira todos os grãos, sobra o quê?;
iii) (047) Depois que se corta o pé do arroz ou do fumo, fica uma parte em baixo assim na terra. Você já chegou a ver plantação assim?
iv) (048) Uma flor grande amarela, redonda que... tem umas sementes assim?;
v) (049) Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?;
vi) (054) Como se chama a armação de madeira, que se coloca no pescoço de animais, pode ser um porco, um bezerro, um carneiro, para ele não atravessar a cerca?.

A análise contrastiva dos itens comuns ao Projeto ALiB, ao ALPR e ao APFB pode revelar a multiplicidade de lexias para o mesmo vocábulo em contextos diferentes, de modo a referendar a heterogeneidade do português brasileiro no âmbito lexical. Assim, esperamos que este estudo revele a diversidade do nosso vernáculo em sincronias diferentes, evidenciando a identidade linguística dos seus falantes e as peculiaridades intrínsecas a cada localidade investigada no Paraná e na Bahia.

4.2 Seleção das Fontes Geolinguísticas

Sob uma perspectiva que valoriza a interação entre os fatores históricos, culturais e geográficos na conformação dos falares, os atlas constituem instrumentos indispensáveis para compreender a organização espacial das variantes linguísticas. No caso da Bahia e do Paraná – regiões marcadas por trajetórias históricas, culturais e processos migratórios distintos –, o entrecruzamento dos dados permite revelar não apenas peculiaridades locais, mas também pontos de convergência que atestam a dinâmica viva e multifacetada da língua portuguesa no Brasil.

4.3 Tabulação dos Dados das Fontes

Nesta etapa metodológica, foram elaboradas planilhas no Excel e no Software para Geração e Visualização de Cartas Linguísticas (SGVCLin) (Romano; Seabra; Oliveira, 2014) para sistematizar os dados coletados, com os resultados apresentados em gráficos que ilustram as porcentagens das variantes lexicais identificadas em cada uma das fontes analisadas (APFB, ALiB-BA, ALPR e ALiB-PR). Esses gráficos permitem uma visualização clara das tendências linguísticas observadas.

Para evidenciar a variação lexical de forma detalhada, as tabelas foram organizadas com base em variáveis relevantes, como localidade (variação diatópica), sexo (variação diassexual) e idade (variação diageracional). Essa estruturação busca retratar as dinâmicas linguísticas, facilitando a identificação de padrões e peculiaridades em cada contexto.

Além disso, a categorização dos dados seguiu princípios metodológicos que garantem a comparabilidade entre as fontes analisadas. O uso de ferramentas visuais, como tabelas e gráficos, não apenas facilita a interpretação dos resultados, mas também contribui para a acessibilidade do estudo.

Com essa abordagem metodológica, o estudo adota uma perspectiva ampla e integradora, que destaca a diversidade das variantes linguísticas e os fatores sociais e geográficos que moldam sua distribuição.

4.4 Procedimentos para Elaboração das Cartas no SGVCLIN

Nessa fase, elaboramos as cartas-base e utilizamos o *software* SGVCLin (Romano; Seabra; Oliveira, 2014), ferramenta virtual para gerar cartas linguísticas. As etapas de cadastramento foram as seguintes: (i) registro do questionário semântico-lexical; (ii) questões e variantes (*espiga*, *sabugo*, *soca*, *girassol*, *vagem* e *cangalha etc.*); (iii) instalação dos mapas da Bahia (APFB e ALiB-BA) e do Paraná (ALPR e ALiB-PR); (iv) rede de pontos com 154 localidades, distribuídas em quatro fontes – APFB, ALiB-BA, ALPR e ALiB-PR; (v) ficha dos informantes, com sexo e faixa etária; (VI) respostas dos informantes; (vi) elaboração das cartas lexicais.

No tocante às cartas, buscamos delinear as peculiaridades inerentes a cada fonte, classificando-as conforme os seguintes critérios: *cartas diatópicas*, destinadas exclusivamente à identificação das variantes mais recorrentes, desconsiderando outros aspectos linguísticos; *cartas diassexuais*, concebidas para cotejar os resultados obtidos na fala masculina e feminina; e *cartas pluridimensionais*, estruturadas com o propósito de correlacionar as variáveis sociais – sexo e faixa etária –, a fim de mensurar a relevância de cada dimensão no âmbito da presente investigação.

4.5 Descrição e Análise Quantitativa e Qualitativa das Fontes Pesquisadas

Nesta etapa, procedemos à descrição e análise quantitativa e qualitativa de cada uma das fontes, separadamente, e, na sequência, estabelecemos comparações entre elas.

A análise quantitativa é importante para interpretar padrões gerais de ocorrência das variantes linguísticas em determinadas localidades, sem a preocupação com os fatores que norteiam as respostas dos informantes. Já a análise qualitativa tem por finalidade averiguar o peso das variáveis sociais, como sexo e faixa etária.

5 ANÁLISE DOS DADOS

Neste capítulo, apresentamos a análise dos resultados obtidos ao investigar a variação lexical das variantes agropastoris (*sabugo, soca, vagem* e *cangalha*) referentes às questões semântico-lexicais do ALiB e do ALPR em épocas e ambientes diferentes – APFB (Rossi, 1963) e ALPR (Aguilera, 1994), elaborados com dados rurais, e ALiB-BA e ALiB-PR (coletas a partir de 2003), com dados urbanos.

Para a análise das variantes aqui propostas, recorremos às seguintes fontes: para *sabugo*, três fontes – APFB, carta 28, dados do ALiB-BA e do ALiB-PR, uma vez que o ALPR não investigou esse conceito; para *soca*, o ALPR II e os dados do ALiB-BA e do ALiB-PR; e, para *vagem* e *cangalha*, o ALPR e os dados do ALiB em ambos os estados.

No plano diacrônico, a análise efetivou-se do seguinte modo: com *sabugo*, a partir dos dados do APFB; com *soca*, pelos dados do ALPR II; e com *vagem* e *cangalha*, pelos dados do ALPR. Quanto ao parâmetro diassexual, embora o APFB não tenha incluído essa variável, foi possível realizar a análise recorrendo às fichas de identificação dos informantes (Rossi, 1963). A análise diageracional foi possível apenas com os dados do ALiB.

Este capítulo está estruturado pela análise das variantes segundo a ordem em que aparecem no ALiB: QSL 046 (*sabugo*), QSL 047 (*soca*), QSL 049 (*vagem*) e QSL 054 (*cangalha*). Em relação aos *corpora*, optamos por analisar primeiro os dados da Bahia (APFB e ALiB-BA) e, depois, os do Paraná (ALPR e ALiB-PR). Por fim, estabelecemos as comparações possíveis entre as quatro fontes estudadas.

Para cada questão analisada, apresentamos um relatório elaborado com o auxílio da ferramenta SGVCLin, contendo os números absolutos e os percentuais de ocorrência, seguido de uma carta lexical com a distribuição espacial das variantes.

5.1 QSL 046 QUANDO SE TIRAM OS GRÃOS DE MILHO DA ESPIGA, SOBRA O QUÊ?

Nesta seção, averiguamos as ocorrências para *sabugo*, entendido como a espiga de milho debulhada sem os grãos, nas quatro fontes: APFB, ALiB-BA, ALPR e ALiB-PR. Ressaltamos que esta é a única questão comum às quatro fontes consultadas para este trabalho. A Figura 5 mostra a parte da espiga onde ficam alojados os grãos do milho.

Figura 5 – Ilustração para *sabugo*.

Fonte: Facebook.

5.1.1 Análise das variantes de *Sabugo* no APFB

Na análise dos dados do APFB (Rossi, 1963), embora não esteja sistematizado, trata-se de um atlas bidimensional com informantes de origem rural. Foram registradas dez variantes: *capuco, sabuco, papuco, sabugo, cascabulho, batueira, sapuco, capucho, pipuco* e *baguço*. Contabilizamos 105 respostas válidas, das quais 46 pertencem à fala masculina e 59 à feminina. O número de não respostas foi de apenas cinco (uma entre os homens e quatro entre as mulheres). A aparente discrepância entre o número de informantes (100) e o de respostas (105) se explica pelo fato de que, em alguns pontos, um único informante apresentou mais de uma variante para o mesmo referente.

A Tabela 1 apresenta a quantificação das variantes para *sabugo*, em números absolutos e por sexo, juntamente com a distribuição das ocorrências na carta 28 do APFB.

Tabela 1 – Variantes lexicais para *sabugo* nos dados do APFB

Variante	Homens	Mulheres	Total
Capuco	19	20	39
Sabuco	12	15	27
Papuco	6	9	15
Sabugo	4	3	7
Cascabulho	3	3	6
Batuera	1	4	5
Sapuco	0	2	2
Outros	1	3	4
NR	1	4	5
Total	47	63	110

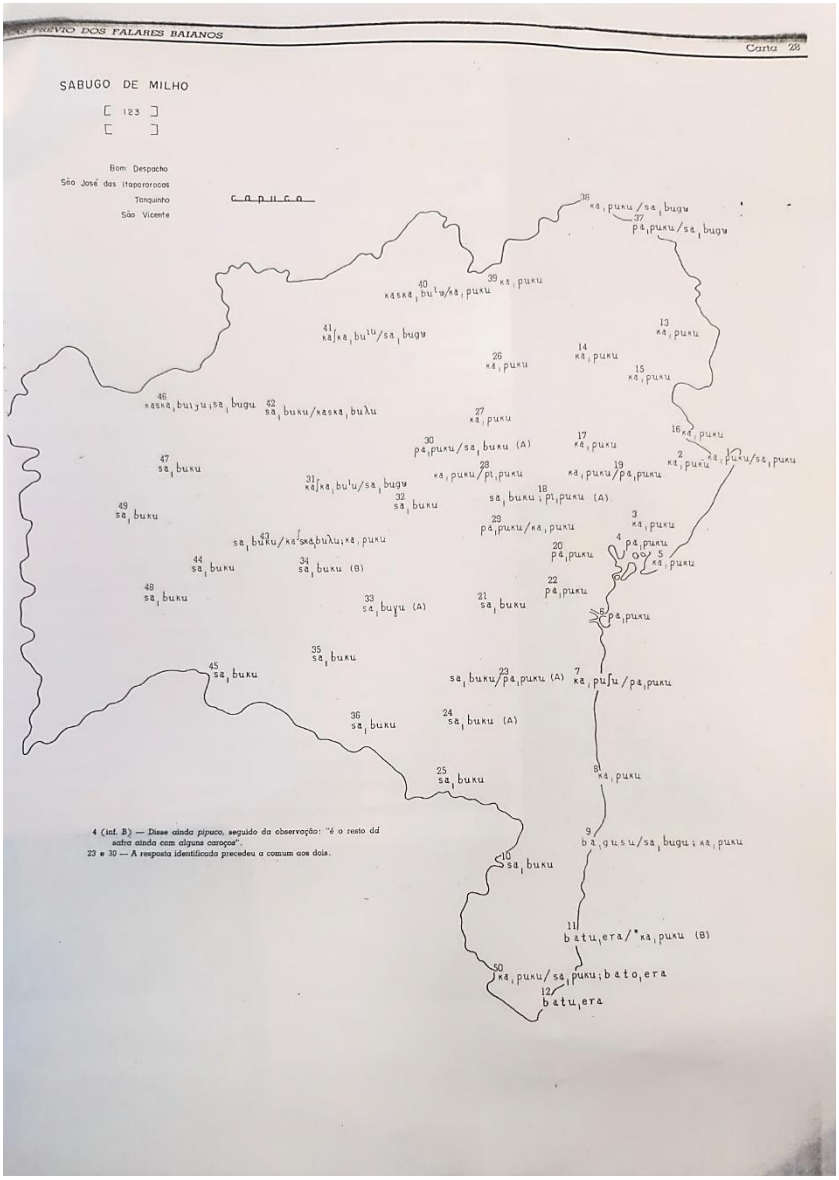
Fonte: Banco de dados do APFB.

Observamos que as variantes fonéticas (*sabugo, sabuco* e *sapuco*), que se distinguem apenas pela sonorização das consoantes ([k] > [g] e [p] > [b]), não foram agrupadas neste trabalho, a fim de averiguarmos, oportunamente, se devem ser analisadas conjuntamente.

De acordo com os dados da Tabela 1, foram registradas as seguintes distribuições: *capuco* (37%), *sabuco* (26%), *papuco* (14%), *sabugo* (7%) e *cascabulho* (6%). As demais variantes obtiveram valores abaixo de 5%. Quanto à variável sexo, embora o APFB não tenha sistematizado esse dado, foi possível recuperá-lo por meio das fichas de identificação das localidades e dos informantes.

Para melhor compreensão dessas ocorrências, apresentamos a Carta 28 do APFB (Figura 6) e a carta isoléxica (Figura 7).

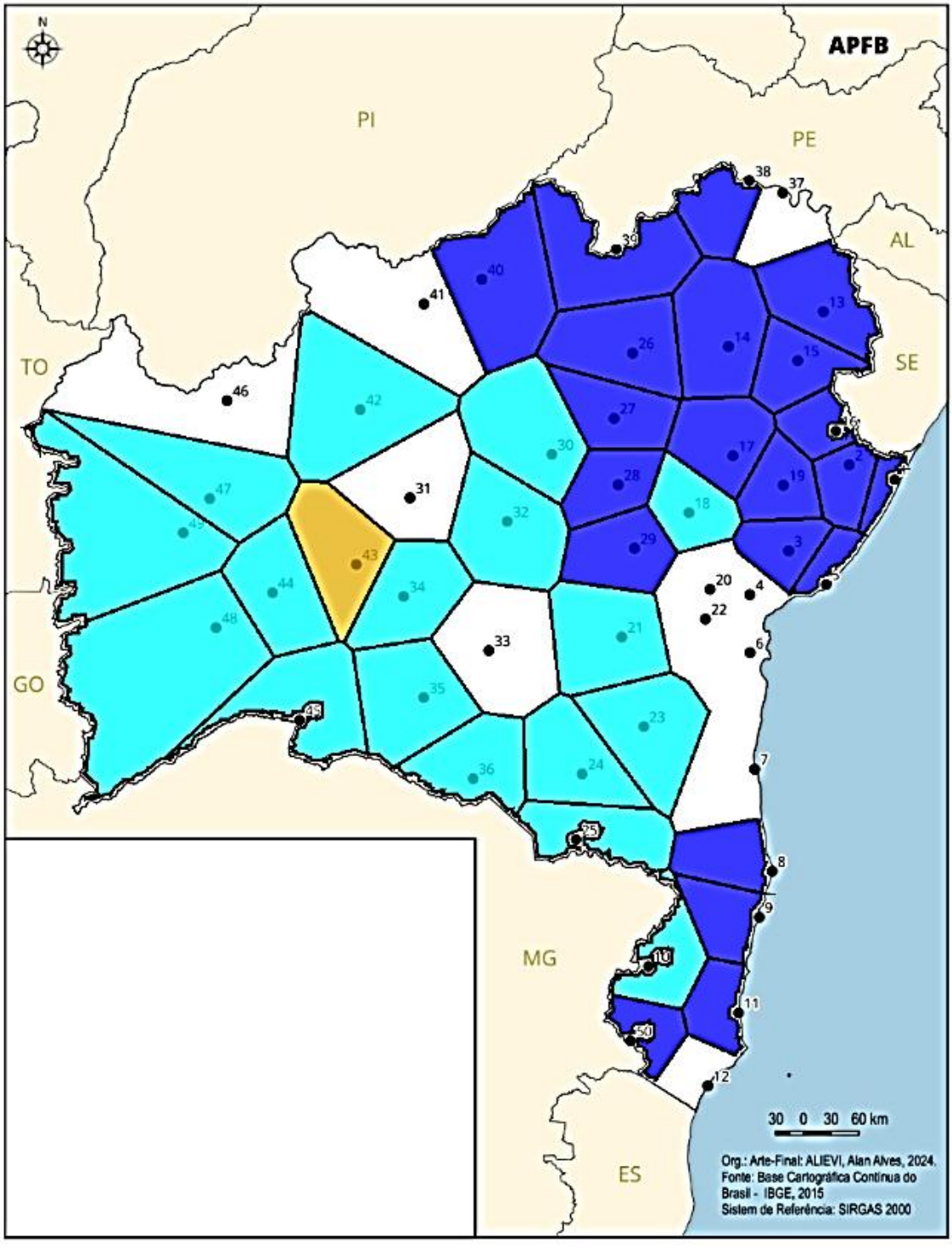
Figura 6 – Carta 28 do APFB com a distribuição das variantes para sabugo



Fonte: Banco de dados do APFB.

Figura 7 – Carta 28 do APFB com a distribuição das variantes para sabugo

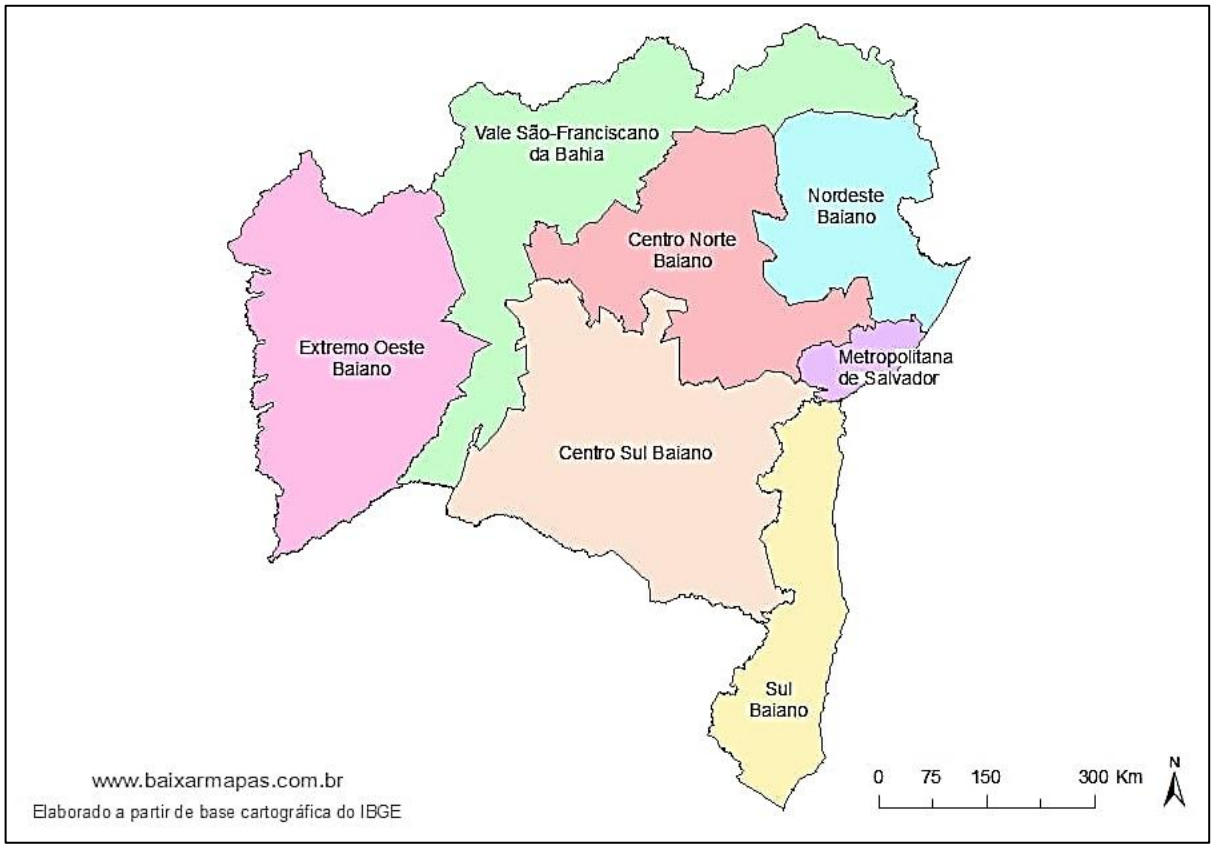
Carta isoléxica para *capuno* do APFB



Fonte: Banco de dados do APFB.

A Carta 28 mostra a seguinte irradiação das variantes de acordo com as mesorregiões da Bahia: (i) a forma mais frequente, *capuco*, concentra-se nas regiões Centro-Norte Baiano, Nordeste Baiano, nos limites com Sergipe e Pernambuco, e na Região Metropolitana de Salvador, apresentando ainda dispersão descontínua em pontos do Sul; (ii) *sabuco* ocupa o Centro-Sul Baiano, a divisa com Minas Gerais, a parte sul do Vale São-Franciscano e o Extremo-Oeste Baiano, na divisa com Goiás e Tocantins; (iii) *papuco* ocorre em área mais direcionada para o Sul Baiano, próxima ao litoral, mas também adjacente ao Centro-Sul Baiano e à Região Metropolitana de Salvador; (iv) *sabugo* não apresenta concentração definida, embora sua trajetória se delineie do Sul para o Norte em áreas difusas; (v) o mesmo ocorre com as seis ocorrências de *cascabulho*, registradas em localidades próximas ao Oeste Baiano. Para melhor compreensão da distribuição geolinguística das variantes lexicais, a Figura 7 retrata as mesorregiões da Bahia, áreas com características físicas, econômicas e sociais semelhantes.

Figura 8 – Mapa da Bahia com as mesorregiões



Fonte: Elaborado a partir da base cartográfica do IBGE, disponível em https://www.baixarmapas.com.br/mapa-da-bahia-mesorregioes/.

A diversidade linguística do falar baiano está intrinsecamente relacionada à sua cultura e à sua diversidade étnica, fruto da miscigenação do colonizador português com o indígena e o africano, o que deu ensejo à formação de uma população com características linguísticas próprias na fonética, na morfossintaxe e, principalmente, no léxico. Nesse nível da linguagem, registramos uma profusão de variantes para nomear o mesmo referente (*sabugo, sabuco, capuco, papuco, cascabulho* etc.). Essa heterogeneidade pode estar relacionada a dois fatores: (i) a grande extensão territorial e (ii) o processo de ocupação e povoamento da Bahia. Dessa forma, para compreender essas manifestações da linguagem, é imprescindível conhecer como ocorreram os fatos históricos.

Documentalmente, a Bahia constitui o ponto de partida da história brasileira. Isso se explica pelo fato de as caravelas de Pedro Álvares Cabral terem atracado primeiramente no Sul da Bahia, onde foi celebrada a primeira missa pelo frei Henrique Soares de Coimbra, na Praia da Coroa Vermelha, em Porto Seguro, no dia 26 de abril de 1500. Esse acontecimento consolidou os primeiros atos lusitanos logo após o descobrimento ou achamento da pátria.

A história da Bahia consiste, pois, em uma sucessão de acontecimentos relevantes que consolidaram a formação desse estado. Assim, a chegada de Cabral ao Sul da Bahia marca o nascimento do Brasil e o surgimento de um território heterogêneo, com vasta cultura, tradições, costumes e festas. A instituição das Capitanias Hereditárias, divisão territorial e administrativa criada pela Coroa Portuguesa como mecanismo de colonização, fomentou uma economia voltada a suprir as necessidades da metrópole. As entradas, responsáveis pela interiorização e pelo povoamento do estado em busca de pedras preciosas, proporcionaram o surgimento de vilas e cidades e de uma economia voltada principalmente para a mineração.

Todos esses fatores contribuíram, direta ou indiretamente, para consubstanciar a cultura, as tradições e a história baiana. Portanto, essa estrutura colonial foi determinante para o linguajar local em suas várias facetas e, sobretudo, em seu aspecto lexical, por abrigar uma quantidade significativa de variantes para nomear um mesmo referente.

5.1.2 Análise das Variantes para Sabugo no Projeto ALiB-BA

Ao analisar os dados do ALiB-BA para *sabugo*, constatamos a presença de 15 variantes entre as respostas dos 88 informantes (44 homens e 44 mulheres), distribuídos por 22 pontos do estado baiano. O número de respostas válidas foi de 88 (43 na fala masculina e 45 na fala feminina). Contabilizamos apenas três não respostas.

As variantes mais representativas, em ordem decrescente, foram: *sabugo*, com 28 respostas; *capuco*, com 21; e *sabuco*, com 19. As respostas únicas, que somaram 20 ocorrências, foram agrupadas sob a denominação "outras": *sapuco, cascabugo, espiga, pituco, papuco, bagaço, miolo, tungado, batuela, cascabú (sic), cascaboi, cascabulho*, entre outras. Para melhor visualização, elaboramos a Tabela 2 com os resultados das variantes para *sabugo*, de acordo com a estratificação dos informantes.

Tabela 2 – Variantes para *sabugo* na fala dos informantes do ALiB-BA

variantes	total	%
sabugo	28	30,77%
capuco	21	23,08%
outras	20	21,98%
sabuco	19	20,88%
não respostas	3	3,30%
total	91	

Fonte: Dados da pesquisa.

Sob o ponto de vista das variáveis diatópicas (rural x urbano) e diacrônicas (1963 x 2000-2004), expressas nas duas fontes (APFB x ALiB-BA), observamos o seguinte: (i) ao contrário dos resultados rurais do APFB (1963), os dados atuais e urbanos do ALiB mostram a expansão de *sabugo* (32%) como variante preferencial, seguida de *capuco* (24%) e *sabuco* (22%); (ii) o número de *outras denominações* (23%) aumentou consideravelmente, apontando para possíveis resquícios da fala rural armazenada no inconsciente do falante ou para variantes regionais que permaneceram no vocabulário ativo do falante urbano; (iii) há uma tendência para o uso de formas padrão em detrimento das regionais. Os Gráficos 2 e 3 apresentam os resultados das variantes em percentual de acordo com as variáveis sexo e faixa etária.

5.1.3 Análise da Variável Sexo para Sabugo nos Dados do ALiB-BA

Tabela 3 – Distribuição das variantes segundo a variável sexo nos dados do ALiB-BA

variantes	homens	mulheres	soma	%
sabugo	13 (28%)	15 (33%)	28	30,77%
capuco	09 (20%)	12 (26%)	21	23,08%
outras	10 (22%)	10 (22%)	20	21,98%
sabuco	11 (24%)	08 (18%)	19	20,88%
não respostas	02 (6%)	01 (2%)	3	3,30%
total	45 (100%)	46 (100%)	91	100,00%

Fonte: Banco de dados do ALiB-BA.

Considerando que um intervalo de 5% é relativamente revelador, podemos inferir que as mulheres ora assumem a variante de prestígio *sabugo*, ora a variante *capuco*, conservadora e regional. A maior diferença em favor da fala masculina, no entanto, está na variante *sabuco*, forma mais próxima da etimologia latina: *sabūcus.*

5.1.4 Análise da Variável Faixa Etária para Sabugo nos Dados do ALiB-BA

A variável faixa etária, conforme a Tabela 4, revelou-se importante diante da maior frequência da forma padrão *sabugo* (32%) entre os falantes da Faixa I (18 a 30 anos), o que pode indicar uma mudança em progresso. As variantes *sabuco*, “outras” (*sapuco, cascabulho, espiga, pituco, papuco, bagaço, miolo, tungado, batuela* etc.) e, principalmente, *capuco*, foram mais frequentes na fala da Faixa II e menos na Faixa I, o que pode apontar para a possibilidade de abandono gradativo dessas formas em favor da variante padrão *sabugo*.

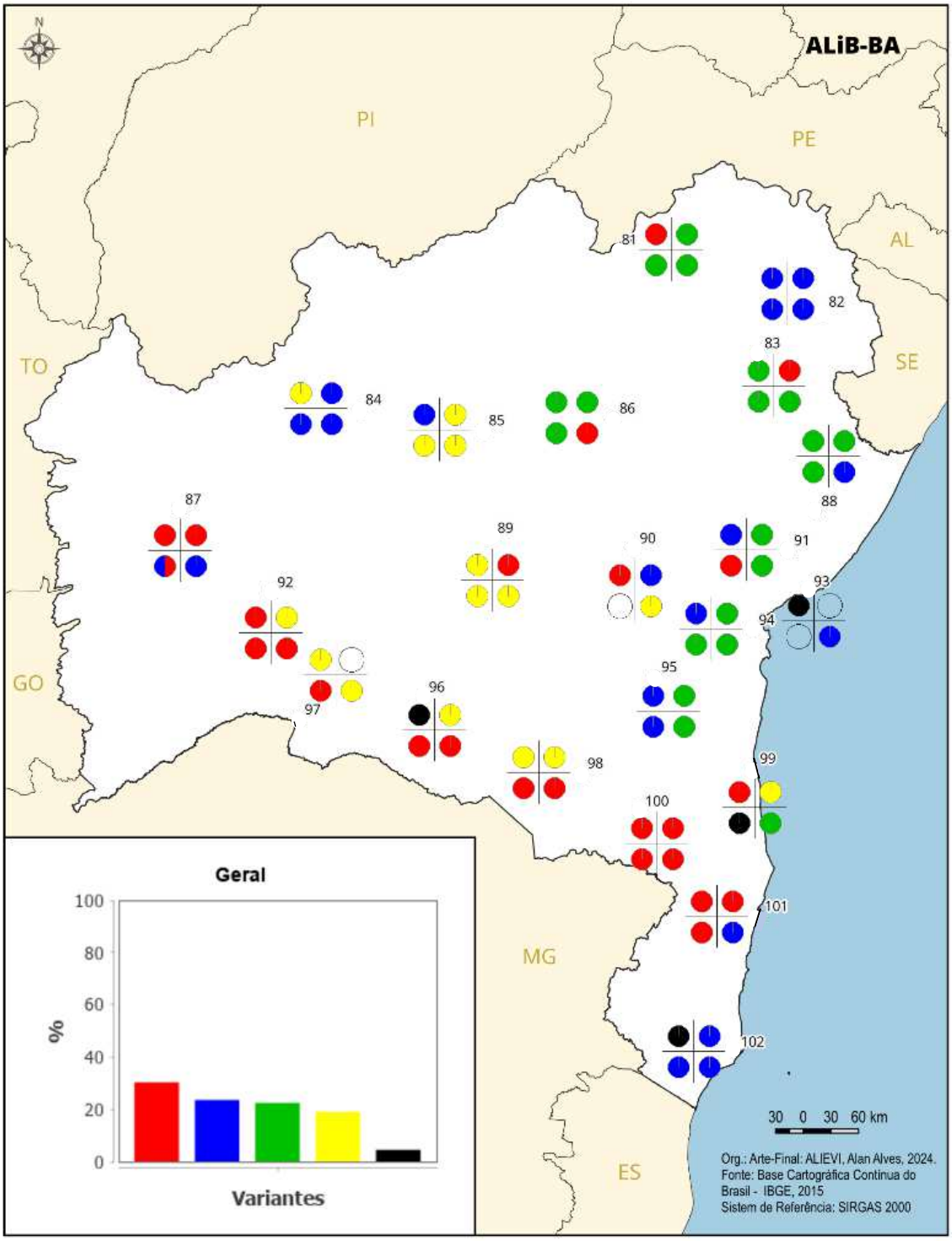
Tabela 4 – Distribuição das variantes, segundo a variável faixa etária, nos dados do ALiB-BA

Variantes	Faixa I	Faixa II	Soma	%
sabugo	17 (36,2%)	11 (25%)	28	30,77%
capuco	08 (17%)	13 (29,5%)	21	23,08%
outras	10 (21,3%)	10 (23%)	20	21,98%
sabuco	09 (19,1%)	10 (22,7%)	19	20,88%
não respostas	03 (6,4%)	00 (0%)	3	3,30%
total	47 (100%)	44 (100%)	91	100,00%

Fonte: Banco de dados do ALiB-BA.

Para sintetizar o que acabamos de descrever, elaboramos a carta experimental (Figura 8) com a distribuição espacial de *sabugo* nos dados do ALiB-BA.

Figura 9 – Carta experimental para *sabugo* nos dados do ALiB-BA



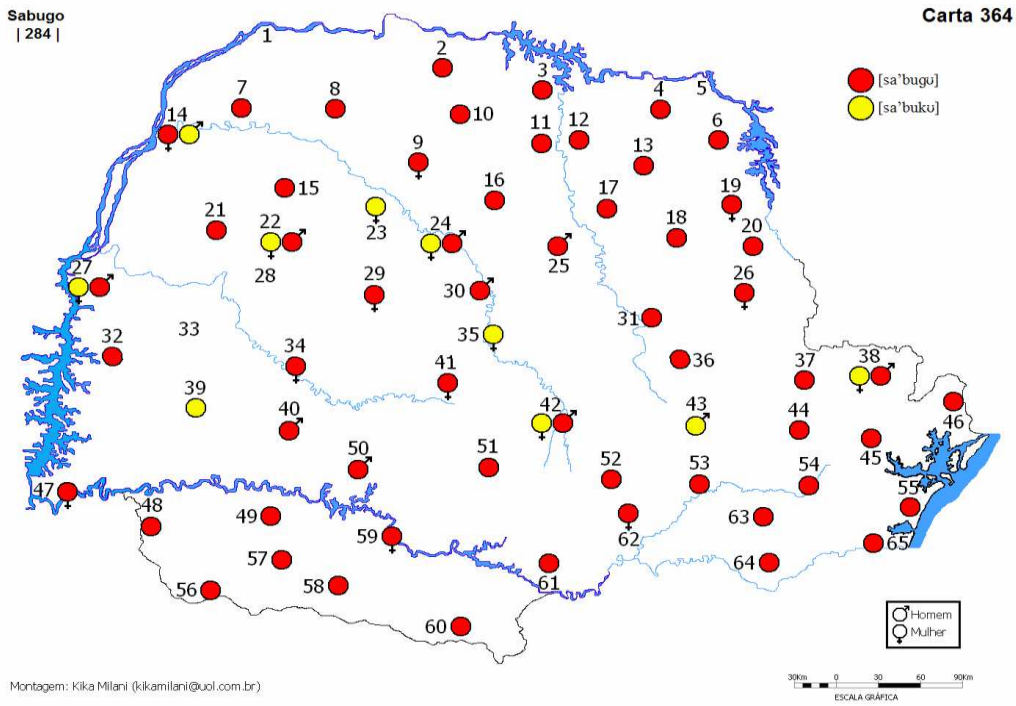
Fonte: Banco de dados do ALiB-BA.

De acordo com a carta experimental (Figura 9), a variante majoritária *sabugo* distribui-se geograficamente por toda a Bahia, mas é mais frequente nas regiões Centro-Sul Baiano e Extremo-Oeste Baiano.[5] Todavia, a lexia *capuco*, assim como no APFB (Figura 6), concentra-se com maior densidade nas regiões Centro-Norte Baiano e Nordeste Baiano, nos limites com Sergipe e Pernambuco, bem como na Região Metropolitana de Salvador, estendendo-se também a pontos do norte e do sul baianos.

5.1.5 Análise das Variantes de *Sabugo* no ALPR II

A Carta 364 do ALPR II (Altino, 2007) traz 128[6] respostas válidas (65 na fala feminina: informante A; 63 na fala masculina: informante B), correspondentes a duas variantes fonéticas (*sabugo* e *sabuco*). Para melhor compreensão da análise, apresentamos a Carta 364 do ALPR II (Altino, 2007), conforme a Figura 10.

Figura 10 – Carta 364 do ALPR II (Altino, 2007) com as variantes para *sabugo.*



Fonte: ALPR II (Altino, 2007).

[5] Nas cartas foram representadas apenas as variantes mais produtivas, acho que fica legal..

[6] Observamos que, no ponto 19, não foi possível registrar a resposta do informante B (homem) por problemas técnicos e no ponto 28, o informante B (homem) não soube responder.

A Tabela 5 retrata as ocorrências para *sabugo* nos dados do ALPR II, segundo a variável sexo, em números absolutos, e o Gráfico 4, em percentuais, respectivamente.

Tabela 5 – Ocorrências das variantes para *sabugo* nos dados do ALPR II

Variantes	Mulheres	Homens	Soma	%
sabugo	57 (88%)	60 (94%)	117	90,00%
sabuco	8 (12%)	3 (5%)	11	8,46%
Não respostas	0 (0%)	2 (1%)	2	1,54%
total	65 (100%)	65 (100%)	130	100,00%

Fonte: Dados da pesquisa.

Os resultados expostos na Tabela 5 revelam que tanto os homens quanto as mulheres utilizaram majoritariamente a variante padrão *sabugo*, com resultados muito próximos, o que demonstra que a *variável sexo* não foi determinante para os resultados. Observa-se, entretanto, que a forma conservadora *sabuco* é ligeiramente mais frequente na fala feminina.

A Figura 10 evidencia as duas variantes presentes no Paraná: a predominância de *sabugo*[7] com 117 ocorrências (91%) e *sabuco* com 11 registros (9%). *Sabugo* pode ser considerada uma variante fonética de *sabuco* devido à sonorização da consoante [k] > [g], portanto inovadora e *sabuco*, a conservadora, procedente do latim *sabūcus*.

5.1.6 Análise das Variantes de *Sabugo* no ALiB-PR

Em relação aos resultados para *sabugo* no ALiB-PR, cujos informantes são oriundos da zona urbana, constatamos, na fala dos 68 informantes, 72 respostas válidas nas 17 localidades paranaenses.[8] A seguir, apresentamos a Tabela 6 com os resultados em números absolutos e a Tabela 7 com números absolutos e percentuais para as variáveis sociais.

---

[7] De acordo com Ferreira (2004), *sabugo* significa, originalmente, a planta sabugueiro e procede da forma latina *sabucu*.

[8] Apenas no ponto 209 (Terra Boa), devido a problemas técnicos, não foi elicitada a resposta da informante 4

Tabela 6 – Distribuição das variantes para sabugo nos dados do ALiB-PR, segundo o perfil dos informantes

Variante	inf. 1	inf.2	inf.3	inf. 4	total	%
sabugo	16	16	16	17	65	90,28%
sabuco	0	0	1	0	1	1,39%
outras	2	2	1	1	6	8,33%
total	18	18	18	18	72	100,00%

Fonte: Dados da pesquisa.

A Tabela 6 mostra que a maioria dos informantes respondeu *sabugo*, com 65 ocorrências (90%). Consta apenas um registro de sabuco (1%) e seis registros de *outras* variantes (*talo*, *caroço*, *espiga*, *bagaço*, *palha*, *cabelinho*), cada uma com apenas uma ocorrência, totalizando 8%

5.1.7 Análise da Variável Sexo para Sabugo nos Dados do ALiB-PR

A respeito da variável sexo, há 32 respostas para *sabugo* na fala masculina (89%) e 33 na fala feminina (92%), assim distribuídas: três na fala dos homens (50%) e três na fala das mulheres (50%). Tal resultado se reflete na faixa etária, com 32 respostas na faixa I (89%) e 33 na faixa II (92%). Consta apenas um registro para *sabuco* na fala do informante 3 - mulher mais velha – representando apenas 1,3%; outras variantes com quatro ocorrências na faixa I (67%) e duas na faixa II (33%).

Esses resultados demonstram que, em relação às duas *variáveis sociais* (sexo e faixa etária), a distribuição das respostas é praticamente equivalente, com discreta predominância da variante *sabugo* em todas as categorias.

Tabela 7 – Distribuição das variantes para sabugo nos dados do ALiB-PR, segundo as variáveis sexo e faixa etária

Variantes	Variável Sexo		Variável Faixa Etária	
	Homens	Mulheres	Faixa I	Faixa II
Sabugo	32 (89%)	33 (92%)	32 (89%)	33 (92%)
Sabuco	0 (0%)	1(1%)	0 (0%)	1 (1%)
Outras	3 (50%)	3 (50%)	4 (67%)	2 (33%)

Fonte: Dados da pesquisa.

5.1.8 Análise Comparativa das Variáveis para Sabugo nos Dados da Bahia e do Paraná

Para contrastar os resultados das quatro fontes, elaboramos a Tabela 8 com números absolutos e percentuais.

Tabela 8 – Percentual das variantes de *sabugo* no APFB, ALiB-BA, ALPR e ALiB-PR

Variante	Capuco	Sabuco	Sabugo	Papuco	Outras	NR
APFB	35%	25%	7%	14%	4%	5%
ALiB-BA	23%	21%	31%	22%	22%	3%
ALPR	0%	8%	90%	0%	0%	2%
ALiB-PR	0%	1%	90%	0%	8%	0%

Fonte: Dados da pesquisa.

Portanto, sob o viés diatópico, as fontes estaduais (APFB e ALPR, ambos rurais) revelaram resultados diferentes, apontando para as divergências sócio-históricas na ocupação de ambos os espaços geográficos. Diante dos dados apresentados, podemos concluir que as diferenças entre o falar baiano e o paranaense já haviam sido registradas por Nascentes ([1922] 1953) em relação às regiões dialetais do Brasil, retratando cada uma delas por meio dos seus subfalares. Essa variação dialetal retratada pelo autor reverbera no léxico de ambos os estados, revelando uma variação maior na Bahia e menor no Paraná.

No APFB, dados do início da década de 1960, *capuco* foi a variante mais produtiva, mas outras lexias também foram registradas (*sabuco, papuco, sabugo, cascabulho, batuera, sapuco* etc.), o que revela a heterogeneidade do falar rural baiano naquela década. No ALPR, a variante *sabugo* (inovadora) foi a mais expressiva, seguida de sua variante fonética *sabuco* (conservadora), com baixa incidência, o que demonstra a tendência de homogeneização do léxico rural paranaense em favor da variante padrão.

Sob o ponto de vista diassexual, os resultados do ALiB em ambos os estados foram próximos, divergindo apenas na escolha da variante lexical (*capuco* na Bahia e *sabugo* no Paraná), fato que demonstra o baixo peso dessa variável, sendo o sexo fator prescindível para o presente estudo.

Em relação à variável diageracional, os dados demonstram que, no ALiB-BA, a variante majoritária para a Faixa I (18 a 30 anos) foi *sabuco*, e para a Faixa II (50 a 65

anos), *capuco*. Essa diferença de resultados em razão da idade credencia a idade como fator relevante na escolha das variantes lexicais nesse *corpus*. A mesma constatação não se observa nos dados do ALPR, em que o número de ocorrências para ambas as faixas é bastante próximo.

Sob o ponto de vista diacrônico, podemos notar que, na Bahia, houve variação na escolha das variantes obtidas em sincronias diferentes (APFB, Rossi 1963; ALiB-BA, 2000/2004), merecendo esclarecimento mais pormenorizado. Dessa forma, percebemos que houve pouca variação lexical quando comparados os resultados do APFB (de origem rural) com os do ALiB-BA (de origem urbana), mas com mudança da variante mais produtiva: *capuco* no primeiro caso e *sabuco* no segundo.

Ao contrastarmos os números do APFB/ALiB-BA com os do ALPR/ALiB-PR, depreendemos uma maior heterogeneidade lexical na Bahia, supostamente em razão de alguns fatores, como a sua dimensão espacial, com uma área de 564.760,429 km² (IBGE, 2022), fato que a credencia como quinto maior estado brasileiro em extensão territorial, bem como os processos de ocupação e povoamento, consubstanciados pela miscigenação de três grupos étnicos distintos cultural, social e linguisticamente: o colonizador português, os povos indígenas e os escravos africanos. Dessa forma, a junção desses fatores possivelmente corroborou para uma diversidade linguística maior que no Paraná.

Todavia, nos dados do Paraná, os resultados foram próximos em ambas as fontes, nas cartas do ALPR II (Altino, 2007) e do Projeto ALiB-PR (2000/2004): predomínio de *sabugo*, com alta incidência, e sua variante conservadora *sabuco*, com baixa incidência. Logo, essa variável não foi responsável pela escolha das variantes no estado sulista. Acreditamos que a homogeneidade dos resultados em ambas as fontes do Paraná (rural e urbana) decorre também de fatores sócio-histórico-geográficos, já que o estado foi colonizado principalmente por imigrantes europeus na região mais ao sul, apesar da intensa migração de paulistas e mineiros para o Norte paranaense durante o ciclo de expansão do café, bem como de gaúchos para as regiões Oeste e Sudoeste, transformando-as em um grande polo de suinocultura e de cereais. Além disso, destacamos sua menor extensão territorial, de 199.298,981 km², ocupando a décima quinta colocação no *ranking* nacional (IBGE, 2022).

Portanto, não apenas o lapso temporal de aproximadamente 40 anos, mas também a origem distinta dos informantes (rural x urbana) e as mudanças sócio-

históricas nesse período foram responsáveis pela mudança na escolha da variante mais produtiva, revelando a importância do peso dessa variável.

5.2 QSL 277 Do ALPR - Depois que se Corta o Pé de Arroz ou de Fumo, fica uma parte Embaixo assim na Terra. Você já Chegou a Ver Plantação Assim?

Nesta seção, analisamos as ocorrências para *soca* (Figura 11), rebrota de algumas culturas, como a do arroz, do fumo, do algodão e da cana-de-açúcar, em três *corpora*: ALPR, carta 298, questão 278; e nos dados inéditos do ALiB-BA e do ALiB-PR.

Figura 11 – Ilustração para soca de cana-de-açúcar

Fonte: Acervo do autor.

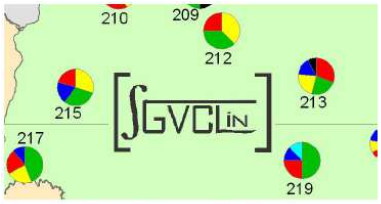
5.2.1 Análise das Variantes Lexicais de *Soca* nos Dados do ALiB-BA

Nos dados do ALiB-BA, registramos 58 respostas válidas referentes às denominações para *soca*, 38 não respostas, além de dois áudios com problemas técnicos: ponto 94-inf. 4 (Valença) e ponto 95-inf. 3 (Jequié). A análise dos dados revelou um elevado grau de variabilidade lexical associado à variante em questão.

A análise dos dados revelou um elevado grau de variabilidade lexical associado à variante em questão. A forma lexical mais recorrente foi *toco*, com 22 ocorrências (23%). Numa ordem decrescente, seguem as demais variantes: soca/soqueira, com nove (9%); *tronco*, com oito (8%); *talo*, com cinco (5%); *broto/brota*, com quatro (4%); *raiz,* com três (3%); *cepa,* com duas (2%); *outras* variantes, com apenas uma ocorrência (*caule, froco, poda, renovo* e *tambora*), totalizaram seis registros (6%).

A Figura12 apresenta o relatório de produtividade das variantes para a questão em pauta.

Figura 12 – Relatório das variantes para *soca* nos dados do ALiB-BA



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

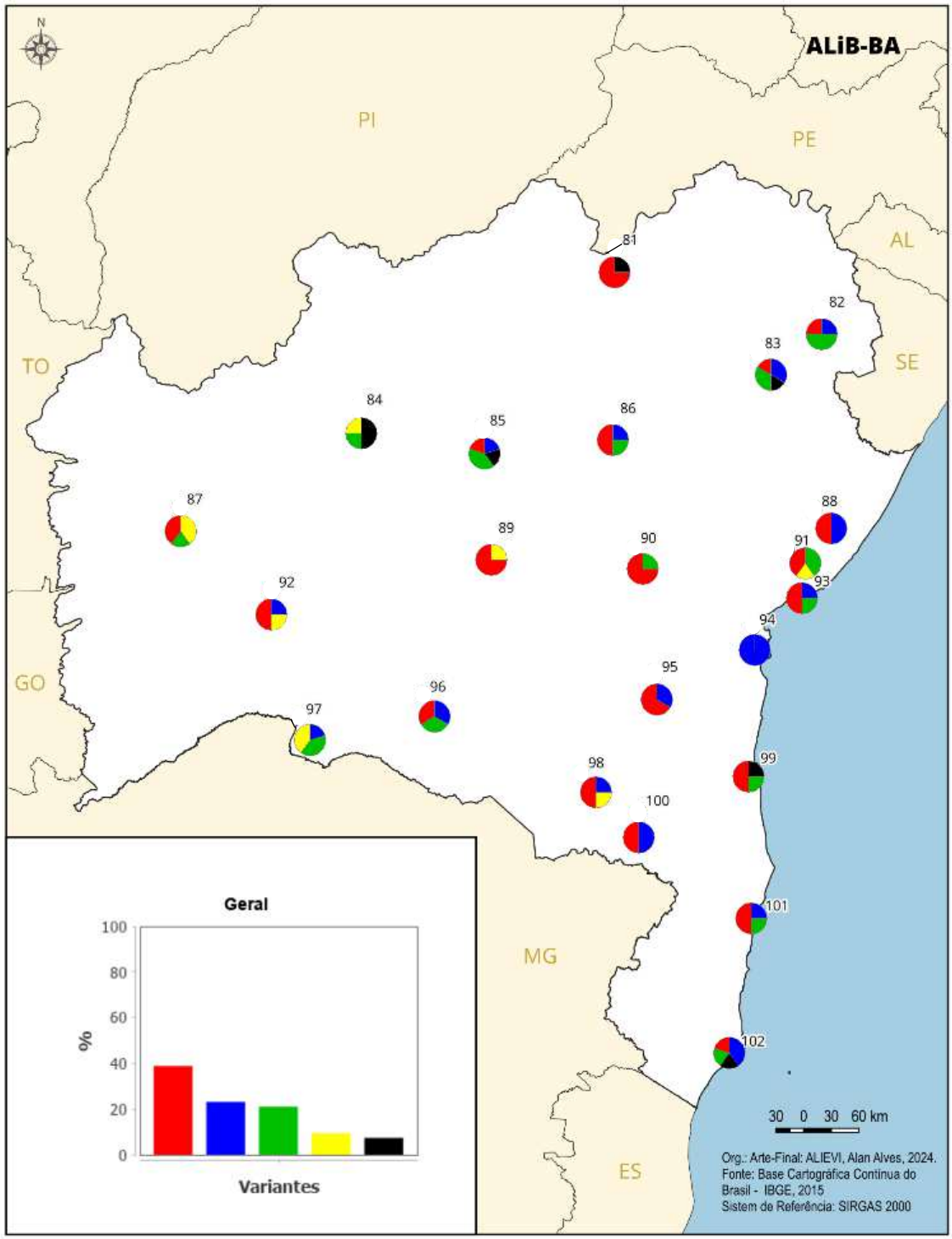
Variantes	Número de ocorrências	%
não respostas	37	38.54%
toco	22	22.92%
soca/ soqueira	9	9.38%
tronco	8	8.33%
outras	6	6.25%
talo	5	5.21%
brota/ broto	4	4.17%
raiz	3	3.12%
cepa	2	2.08%
	96	

PT: 2PT: 94/4, 95/3

Fonte: Banco de dados do ALiB-BA.

Para elucidar essas ocorrências, elaboramos a carta monodimensional (Figura 13), com a distribuição das variantes de *soca* no ALiB-BA.

Figura 13 – Carta experimental das variantes de soca no ALiB-BA



Fonte: Banco de dados do ALiB-BA.

5.2.2 Análise da Variável Sexo nos Dados do ALiB-BA Referentes às Variantes de *Soca*

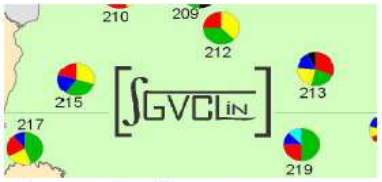
A análise da variável sexo evidencia que, das 58 ocorrências documentadas como respostas válidas, 37 foram provenientes da fala masculina e 22 da feminina, denotando maior amplitude de variabilidade lexical entre os homens. Verificamos também um número expressivo de não respostas em ambos os grupos, totalizando 37 registros. Contudo, a proporção de *não respostas* foi maior entre as mulheres (24 ocorrências), em contraste com os homens (13 ocorrências). Esses resultados podem ser interpretados como reflexos de uma organização social ainda pautada por padrões patriarcais, nos quais os homens desempenham atividades de maior exigência física e complexidade, como aquelas vinculadas à agricultura e à pecuária, enquanto às mulheres são atribuídas funções predominantemente associadas ao espaço doméstico.

A variante lexical de maior produtividade, na fala masculina, foi *toco*, com 15 registros (30%); *soca/soqueira*, com oito (16%); *talo*, com quatro (8%). Menos frequentes, temos: *brota/broto*, com três ocorrências (6%); *tronco*, com duas (6%); *cepo/cepa* e *raiz*, com apenas um (2%). As variantes com ocorrência única foram agrupadas em *outras* (*caule, renova* e *cama*) para análise conjunta, totalizando registros (6%).

Na fala das mulheres, registramos: *toco*, com sete ocorrências (15%); *tronco*, com seis (13%); *raiz*, com duas ocorrências (4%); e as hápax *broto/brota, cepa, talo* e *soca/soqueira* (9%). Outros itens lexicais (*poda, froco* e *tambora*), com um registro de cada, apesar de não recobrirem o conceito de *soca*, foram contabilizados em *outras* (7%).

Para melhor visualização dos dados, a Figura 14 traz o relatório das variantes para *soca* no ALiB-BA, segundo a variável sexo, e a carta experimental (Figura 15) apresenta a distribuição das variantes na rede de pontos.

Figura 14 – Relatório das variantes para soca no ALiB-BA, segundo a variável sexo



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

Variantes	Número de ocorrências	%
masculino		
toco	15	30.00%
não respostas	13	26.00%
soca/ soqueira	8	16.00%
talo	4	8.00%
brota/ broto	3	6.00%
outras	3	6.00%
tronco	2	4.00%
raiz	1	2.00%
cepa	1	2.00%
	50	
feminino		
não respostas	24	52.17%
toco	7	15.22%
tronco	6	13.04%
outras	3	6.52%
raiz	2	4.35%
brota/ broto	1	2.17%
soca/ soqueira	1	2.17%
cepa	1	2.17%
talo	1	2.17%
	46	

PT: 2PT: 94/4, 95/3

Fonte: Banco de dados do ALiB-BA.

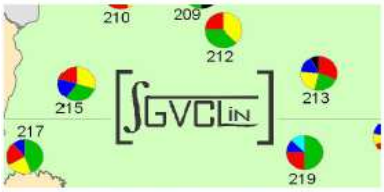
Os dados da Figura 14 evidenciam que a variável sexo não exerceu um papel preponderante na variação linguística observada na QSL 047 do ALiB-BA. Portanto, tal resultado revela certa homogeneidade nas respostas, independentemente do sexo dos participantes.

5.2.3 Análise da Variável Faixa Etária nos Dados do ALiB-BA para as Variantes de *Soca*

Entre os informantes da Faixa I (18 a 30 anos), foram registradas 44 ocorrências. O número de não respostas foi expressivo, já que 24 informantes (55%) desconhecem o nome do referente. A forma *toco* foi a mais frequente, com oito registros (18%). Numa ordem decrescente, as demais variantes obtiveram os seguintes resultados: *tronco*, com quatro respostas (9%); *raiz*, com três (7%); *brota/broto* e *talo*, com apenas uma menção cada uma (5%); *outras variantes* (*caule* e *renovo*), com um registro apenas (7%), foram agrupadas.

Na Faixa II (50 a 65 anos), foram registradas 47 ocorrências. A variante mais produtiva foi *toco*, com 13 registros (28%). Aqui também o número de não respostas foi alto, tendo em vista que 12 informantes (26%) desconheceram o nome do referido objeto. As demais variantes obtiveram os seguintes valores em ordem decrescente: *soca/soqueira*, com oito respostas (17%); *tronco*, com quatro (9%); *talo*, com três (6%); *cepa* e *brota/broto*, duas cada uma (8%); outras variantes (*poda, tambora* e *froco*) foram agrupadas, dada a baixa incidência (6%). A Figura 15 traz o relatório das ocorrências das variantes por ponto linguístico, considerando ambas as faixas etárias.

Figura 15 – Relatório das variantes para soca no ALiB-BA segundo a variável faixa etária



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

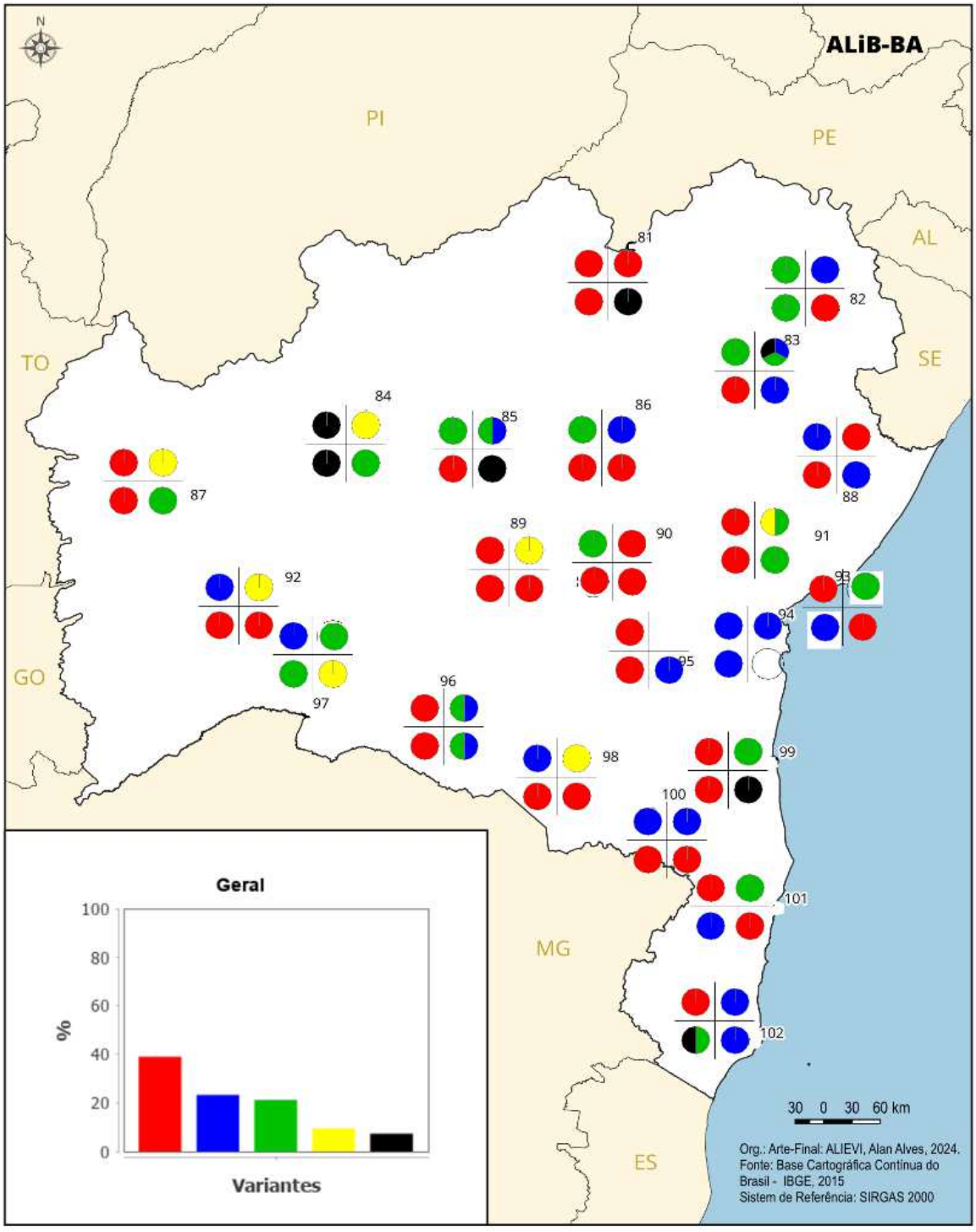
Variantes	Número de ocorrências	%
Faixa I		
não respostas	24	54.55%
toco	8	18.18%
tronco	4	9.09%
raiz	3	6.82%
outras	3	6.82%
talo	1	2.27%
brota/ broto	1	2.27%
	44	
Faixa II		
toco	13	27.66%
não respostas	12	25.53%
soca/ soqueira	8	17.02%
tronco	4	8.51%
talo	3	6.38%
outras	3	6.38%
cepa	2	4.26%
brota/ broto	2	4.26%
	47	

PT: 2PT: 94/4, 95/3

Fonte: Banco de dados do ALiB-BA.

A Figura 16 retrata as variantes para *soca* nos dados do ALiB-BA, de acordo com as variáveis sexo e faixa etária agrupadas.

Figura 16 – Carta das variantes para soca nos dados do ALiB-BA, segundo as variáveis sexo e faixa etária



47 – Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

Legenda

- Não respostas
- toco
- outras
- soca/ soqueira
- tronco

①②
③④

1: masculino – Faixa I
2: masculino – Faixa II
3: feminino – Faixa I
4: feminino – Faixa II

SGVCLin]® - 2015

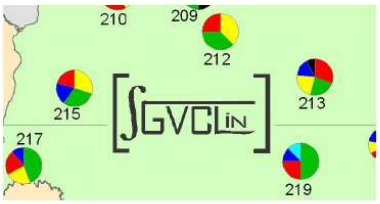
Fonte: Banco de dados do ALiB-BA.

5.2.4 Análise das Variantes Lexicais de *Soca* nos Dados do ALPR II

O ALPR, que deu origem ao ALPR II, entrevistou 130 informantes em 65 localidades da rede de pontos, selecionando em cada localidade um homem e uma mulher, com idades entre 18 e 60 anos. O levantamento dos dados da carta 298 resultou em 137 respostas, com quatro não respostas (informante 1 dos pontos 4, 46 e 60; informante 2 do ponto 19). A baixa incidência de *não respostas* decorre da natureza do ALPR: um atlas que privilegiou informantes e questões relativas à área rural.

A forma linguística mais recorrente foi *soca*, juntamente com sua variante morfológica (acréscimo do sufixo *-eira*) *soqueira*, que somaram 72 registros (51%)[9]. Em ordem decrescente, obtivemos: *toco*, com 27 ocorrências (19%); *resteva* e *cepa*, com 11 ocorrências cada uma (8%); *raiz*, com oito registros (6%); *touceira*, com seis registros (4%); e, por fim, a expressão *palha do arroz*, identificada na fala de três participantes (2%). A Figura 17 apresenta o relatório das variantes em números absolutos e percentuais.

Figura 17 – Relatório das variantes para soca nos dados do ALPR II



Geral

Número da questão: 277

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

Variantes	Número de ocorrências	%
soca/ soqueira	72	50.70%
toco	27	19.01%
cepa	11	7.75%
resteva	11	7.75%
raiz	8	5.63%
touceira	6	4.23%
não respostas	4	2.82%
palha do arroz	3	2.11%
	142	

PT: 2PT: 38/1, 38/2

Fonte: Banco de dados do ALPR II.

Os dados da Figura 17 mostram que esse referente é polimórfico nos dados do ALPR II. Tal variabilidade pode ser explicada pela origem rural dos informantes, que,

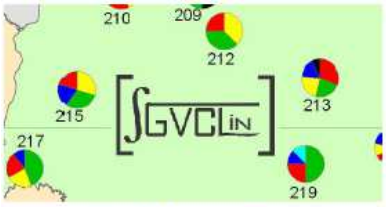
[9] Apresentamos aqui os percentuais pelos números inteiros, sem as frações.

consequentemente, apresentam maior familiaridade com as denominações do campo semântico campesino.

5.2.5 Análise da Variável Sexo de *Soca* nos Dados do ALPR II

Ao comparar as respostas elicitadas por informantes de ambos os sexos, constatamos que os resultados são bastante semelhantes, conforme ilustrado na Figura 18, relatório das variantes para *soca* no ALPR II.

Figura 18 – Relatório das variantes para soca nos dados do ALPR II, segundo a variável sexo



Geral

Número da questão: 277

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

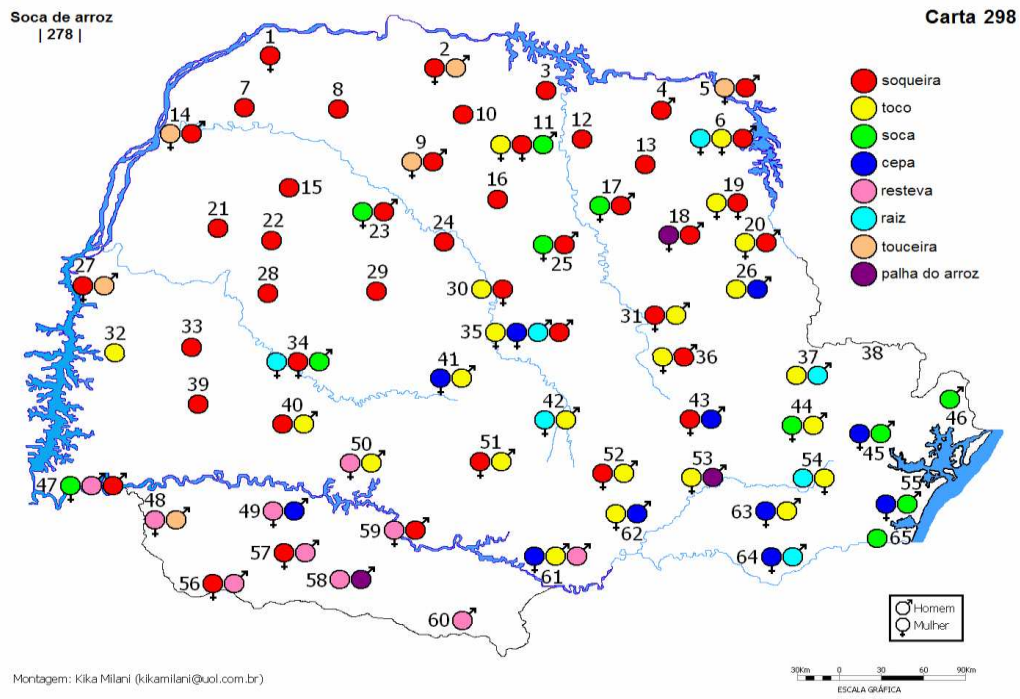
Variantes	Número de ocorrências	%
masculino		
soca/ soqueira	37	52.11%
toco	14	19.72%
resteva	6	8.45%
raiz	4	5.63%
cepa	4	5.63%
touceira	3	4.23%
palha do arroz	2	2.82%
não respostas	1	1.41%
	71	
feminino		
soca/ soqueira	35	49.30%
toco	13	18.31%
cepa	7	9.86%
resteva	5	7.04%
raiz	4	5.63%
touceira	3	4.23%
não respostas	3	4.23%
palha do arroz	1	1.41%
	71	

PT: 2PT: 38/1, 38/2

Fonte: Banco de dados do ALPR II.

Apresentamos, na Figura 19, a carta 277 (Altino, 2007), que descreve as ocorrências das variantes lexicais para *soca* nos dados do ALPR II, segundo a variável sexo.

Figura 19 – Carta para soca no ALPR II, segundo a variável sexo

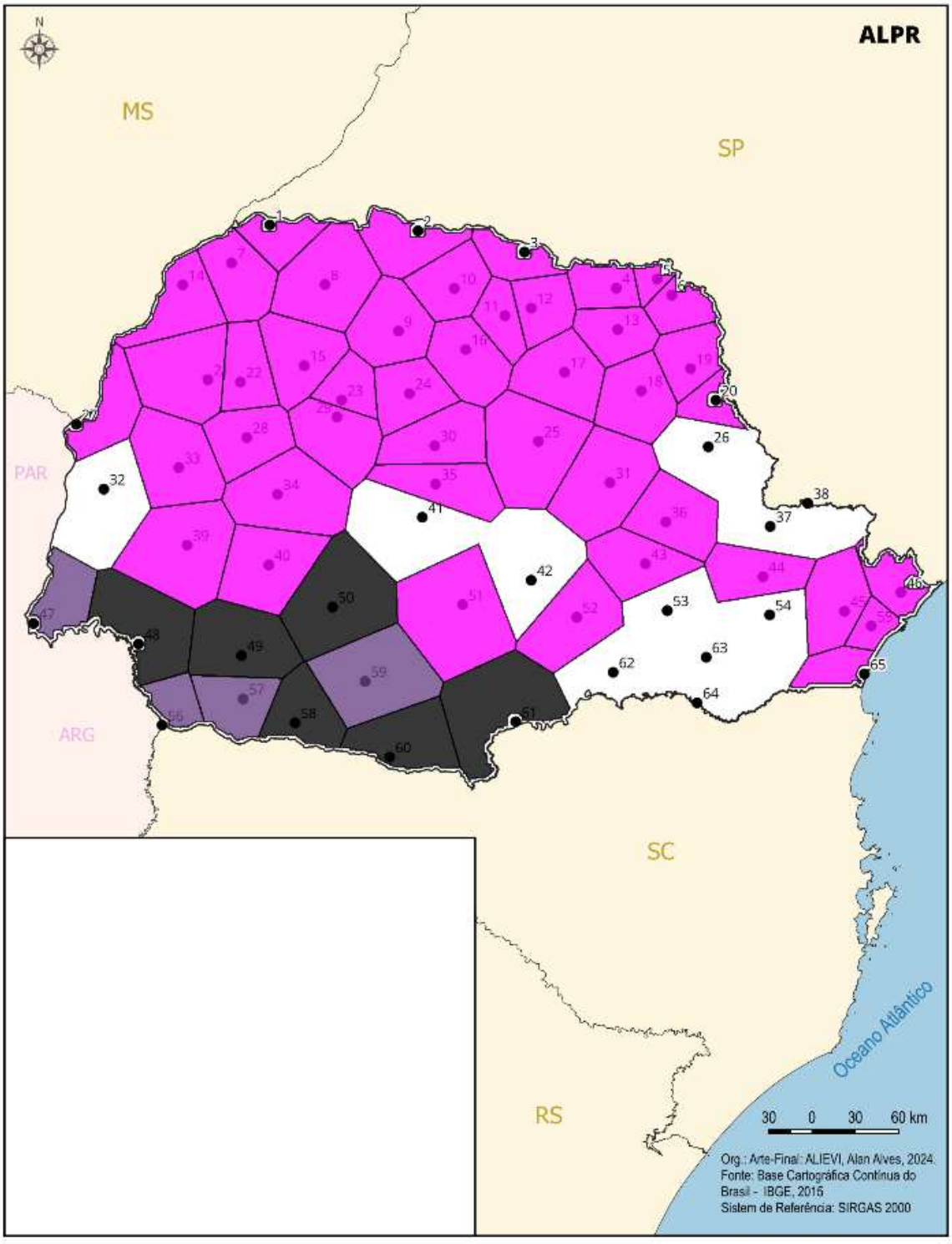


Fonte: Banco de dados do ALPR II.

Conforme observamos na Figura 19, as variantes *touceira* e *resteva* apresentam contornos bem delimitados em espaços geolinguísticos contínuos. A primeira distribui-se por todas as áreas do Paraná; contudo, encontra-se mais densamente concentrada no Leste Paranaense e, principalmente, no Norte e Noroeste do estado, nas divisas com São Paulo e Mato Grosso do Sul. A segunda estende-se pelo Oeste e Sudoeste, na fronteira com o Paraguai e a Argentina, bem como na divisa com Santa Catarina.

Para melhor compreender esse fenômeno linguístico, elaboramos a carta de arealidade que delimita as áreas de maior concentração de determinadas variantes, que, na presente seção, são *soca/soqueira* e *resteva* (Figura 20), bem como as áreas de coocorrência de ambas as formas.

Figura 20 – Carta isoléxica para soca elaborada com os dados do ALPR II



Fonte: Banco de dados do ALPR-II.

De acordo com a Figura 20, concluímos que há duas isoglossas bem definidas: *soca/soqueira* e *resteva*. Tais ocorrências podem estar associadas aos movimentos históricos responsáveis pela formação do Paraná, conforme se depreende do levantamento historiográfico que passamos a abordar a respeito do surgimento e do povoamento do estado.

Em meados do século XVII, a procura por pedras preciosas incentivou a vinda de bandeirantes paulistas para a região do Leste Paranaense, dando ensejo ao surgimento de cidades como Curitiba, Morretes, Antonina e Paranaguá, sendo esta última fundada por padres jesuítas (Aguilera, 1994). De acordo com a autora, consolidando essa fase pioneira, Curitiba foi elevada à categoria de cidade em 29 de março de 1693. O surgimento dos primeiros povoamentos nessa região pode explicar a presença de *soca* em espaços territoriais distintos, supostamente transportada do litoral para o Norte do Paraná pelos bandeirantes.

Após o povoamento dessa região, localizada mais a leste, a escassez de pedras preciosas e a necessidade de demarcação fomentaram a marcha para o oeste. Nesse contexto, as missões tiveram papel importante no processo de interiorização e povoamento da área. Para viabilizar essa ocupação, a pecuária e a extração da erva-mate tornaram-se a base da economia regional.

Alguns acontecimentos internacionais desencadearam um novo ciclo econômico no Paraná. O fim da escravidão, promovido pela Lei Áurea em 13 de maio de 1888, e a consequente escassez de mão de obra ocasionaram a vinda de imigrantes para o Brasil, a fim de substituir os escravos na produção agrícola. Soma-se a isso outro componente: o período entre guerras, que dizimou a economia europeia, deixou milhares de famílias sem emprego e sem casa em um continente devastado. Nesse contexto beligerante, chegou ao Paraná um contingente expressivo de imigrantes italianos, alemães, poloneses e ucranianos. Tais povos encontraram condições propícias para estabelecer suas residências, tais como clima semelhante ao europeu, terras férteis e incentivos para aquisição de propriedades rurais. Segundo Priori, a respeito da imigração no Paraná, é importante salientar que

> […] os imigrantes vinham sob a condição de colonos e tinham de formar as lavouras a partir da abertura das picadas na mata, bem como fazer suas próprias moradias. Neste Estado, o objetivo primordial era fomentar o cultivo da agricultura de subsistência, para manter abastecidos os centros urbanos que já passavam por raleamento de produtos agrícolas (Priori *et al,* 2012, p. 36).

Outro fator relevante para o povoamento do interior do Paraná foi a erva-mate, planta nativa da região sul do estado, consumida pelos povos indígenas antes mesmo da chegada dos europeus. Vendo nesse produto uma fonte de renda, os imigrantes europeus passaram a cultivá-la abundantemente. O primeiro ciclo da erva-mate terminou pouco antes de 1820, época em que ainda não era industrializada. O segundo ciclo estendeu-se de 1820 a 1875, período em que a planta ganhou importância econômica e dimensão internacional. O terceiro ciclo ocorreu entre 1875 e 1880, quando novas técnicas ampliaram sua produtividade e expansão territorial, alcançando outros estados. Essa cultura foi especialmente explorada por imigrantes poloneses e ucranianos, notadamente em pequenas propriedades. Após o seu ápice, o cultivo entrou em declínio, cedendo lugar a um novo ciclo econômico no Paraná: a monocultura do café.

No fim do século XIX, a produção excessiva de café no Oeste Paulista exigiu do governo estadual medidas restritivas para valorização do produto na balança comercial internacional. Esse acontecimento, aliado ao aprimoramento de novas técnicas de cultivo dessa *commodity*, desestimulou os produtores paulistas, que viram no Norte do Paraná condições favoráveis para ampliar seus lucros. A esse respeito, Priori *et al.* explicam:

> No fim do século XIX e no início do século XX a extensão dos cafeeiros paulistas atingiu o Paraná. Com o aumento excessivo da produção e da oferta, em 1902 o Estado de São Paulo adotou políticas de restrição da produção do café, proibindo o seu plantio por um período de cinco anos. Em função desse problema os fazendeiros paulistas optaram em procurar terras no Norte do Paraná, impulsionando a expansão cafeeira no Estado (Priori *et al.*, 2012, p. 93).

A chegada do café a essa região foi responsável pelo povoamento, primeiramente, do Norte Pioneiro ou Norte Velho, abrangendo municípios como Jacarezinho, Santo Antônio da Platina, Andirá e Cambará. Jacarezinho, por exemplo, transformou-se na primeira capital do café no Paraná, com cerca de 500.000 pés plantados no início do século XX (Wachowicz, 1987). Após a ocupação do Norte Pioneiro por paulistas e mineiros, o café expandiu-se para o Norte Novo, região de Londrina e Maringá, e para o Norte Novíssimo, região de Paranavaí. Graças à riqueza gerada pela cultura cafeeira, Londrina e Maringá prosperaram, tornando-se grandes centros comerciais com economia promissora.

Segundo Aguilera (1994), com base em Cardoso e Westphalen, o Sudoeste e o Oeste do estado, mais ligados ao Rio Grande do Sul, tiveram seu processo de

ocupação a partir da década de 1940. A atividade econômica desses territórios teve como polos cidades como Cascavel, Marechal Cândido Rondon, Pato Branco e Francisco Beltrão. Nesses municípios, desenvolveram-se a suinocultura e a produção de grãos, corroborando o povoamento de uma região até então pouco habitada. Sobre esse aspecto, Vescovi e Seide reiteram:

> O Oeste paranaense foi sendo constituído a partir da década de 50. As principais ascendências da região são a italiana e a germânica, tendo em vista que os pioneiros, assim nomeados os desbravadores do oeste, são, em sua maioria, oriundos do Rio Grande do Sul e de Santa Catarina e filhos ou netos de imigrantes alemães e italianos que migraram tal região (Vescovi; Seide, 2013, p. 3).

Dado que a ocupação e o povoamento do Oeste e do Noroeste paranaenses foram protagonizados por gaúchos e catarinenses, e considerando que a variante *resteva* apresenta alta frequência em ambas as regiões, na divisa com Santa Catarina, deduzimos que essa lexia tenha chegado ao Paraná por meio desses migrantes.

A análise da carta evidencia a presença de uma marcante diversidade linguística característica do dialeto paranaense, resultante dos processos históricos e culturais que compõem a identidade do estado.

A prevalência da variante *soca/soqueira* demonstra um traço de conservadorismo linguístico em regiões onde a homogeneidade histórica e cultural é mais evidente: *soca/soqueira* no Norte do estado, em razão do fluxo migratório dos bandeirantes do litoral; e *resteva* no Oeste e Noroeste, em virtude da ocupação a partir da década de 1940 por gaúchos e catarinenses. Esse fato indica práticas agrícolas profundamente enraizadas – o café no primeiro caso e a suinocultura no segundo – que contribuem para a consolidação de uma norma regional estratificada.

5.2.6 Análise das Variantes Linguísticas de *Soca* nos Dados do ALiB-PR

Os dados do ALiB-PR referentes a *soca* somam 77 respostas válidas, correspondentes a 25 variantes lexicais, algumas das quais foram agrupadas considerando variações de ordem morfofonêmica (*soca, soque, soqueira*) ou de gênero gramatical (*cepo* e *cepa*). Registramos também seis casos de ausência de resposta, decorrentes do desconhecimento do referente ou de sua denominação, além de sete respostas prejudicadas devido a problemas técnicos.

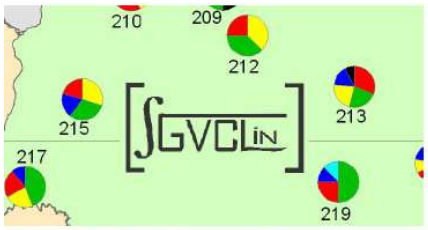
A variante predominante foi *soca*, incluindo suas realizações morfofonêmicas *soque* e *soqueira*, com 15 registros (19%). Em ordem decrescente, obtivemos: *raiz*,

com 12 ocorrências (16%); *toco*, com 11 registros (14%); *touceira* e *toça*, com sete cada uma (9%). Entre as menos recorrentes, estão: *tronco* e *talo*, com cinco ocorrências cada uma (6%), e *palha*, com quatro (5%). Já *cepa* e *cepo* tiveram três registros (4%), enquanto *semente* e *tiguera* apareceram duas vezes cada uma (3%). As demais variantes – *bagaço, capim, estalo, rama, muda, sobras, restos, broto, grãos, centelho do trigo* e *caule* –, citadas apenas uma vez, integram o grupo "outras".

Verificamos que o número de não respostas foi pouco expressivo, uma vez que apenas seis informantes deixaram de reconhecer a denominação atribuída ao referente da QSL 47. Esse resultado pode indicar a crescente influência do agronegócio moderno – caracterizado pela rotação de culturas, mecanização agrícola, adoção de tecnologias avançadas e elevados índices de produtividade – no cotidiano urbano. Tal disseminação favorece a popularização de termos e expressões anteriormente restritos a grupos minoritários de origem rural, inserindo-os no vocabulário corrente de uma sociedade progressivamente conectada às dinâmicas do campo. No entanto, a ausência de uma denominação predominante e o distanciamento da maioria dos informantes da lida agrícola cotidiana podem ter levado ao polimorfismo e à extensão do significado de um referente para outro com o qual mantém algum traço semântico, como *cepa, raiz, bagaço, tronco* e *talo*, entre outros.

Para melhor compreensão da variação lexical de *soca* nos dados do ALiB-PR, elaboramos a Figura 21 com o número de ocorrências e os respectivos percentuais.

Figura 21 – Relatório das variantes para *soca* nos dados do ALiB-PR.



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

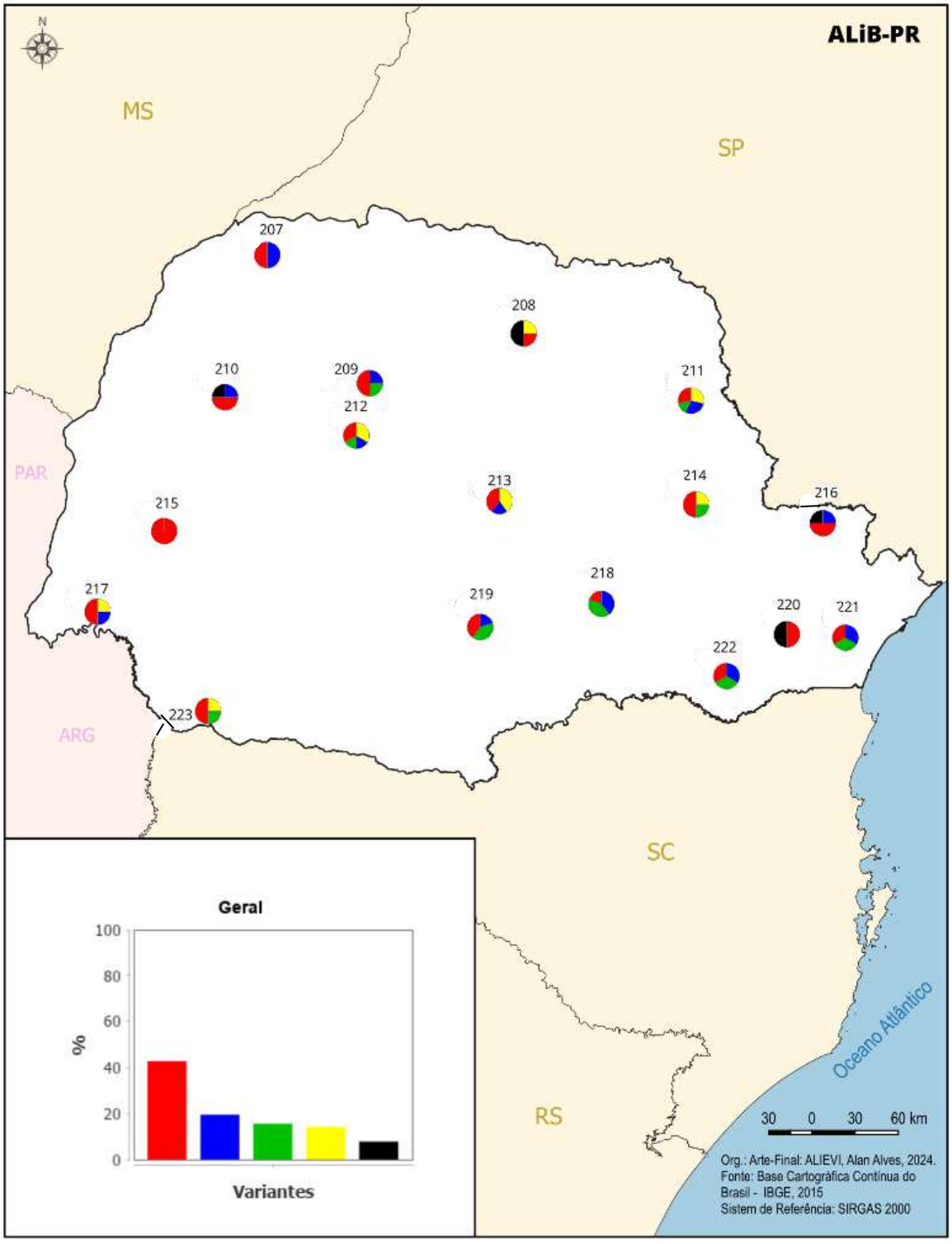
Variantes	Número de ocorrências	%
outras	17	22.08%
soca/ soque/ soqueira	15	19.48%
raiz	12	15.58%
toco	11	14.29%
não respostas	6	7.79%
touceira/ touça	6	7.79%
talo	5	6.49%
tronco	5	6.49%
	77	

PT: 7PT: 209/2, 215/1, 215/2, 216/4, 219/4, 221/1, 222/1

Fonte: Banco de dados do ALiB-PR.

A Figura 22 ilustra a distribuição diatópica das variantes da QSL 47 (*soca*) nos dados do ALiB-PR.

Figura 22 – Carta monodimensional das variantes de *soca* nos dados do ALiB-PR



Fonte: Banco de dados do ALiB-PR.

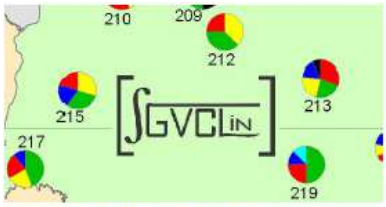
5.2.7 Análise da Variável Sexo nos Dados do ALiB-PR

Das 42 ocorrências para os homens, *soca* e suas variantes (*soque* e *soqueira*) foram as mais produtivas, com 13 registros (31%). As demais lexias, com frequência menor, distribuem-se em ordem decrescente: *toco*, com sete (17%); *touceira* e *raiz*, com quatro cada uma (10%); *não respostas* e *talo*, com três cada uma (7%); e *tronco*, com apenas uma (2%).

Das 35 ocorrências para as mulheres, *raiz* foi a variante majoritária, com oito registros (23%). As demais lexias apresentaram baixa produtividade, em ordem decrescente: *tronco* e *toco*, com quatro cada uma (11%); *não respostas*, com três (9%); *touceira, soca* e *talo*, com dois registros cada (6%).

Para melhor compreensão da variável sexo, elaboramos a Figura 23 com os números absolutos e os percentuais. Na sequência, apresentamos a carta experimental (Figura 24), que mostra a disposição das variantes para *soca* conforme a variável faixa etária, pelos pontos linguísticos.

Figura 23 – Relatório das variantes para *soca* nos dados do ALiB-PR, segundo a variável sexo



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

Variantes	Número de ocorrências	%
masculino		
soca/ soque/ soqueira	13	30.95%
outras	7	16.67%
toco	7	16.67%
touceira	4	9.52%
raiz	4	9.52%
não respostas	3	7.14%
talo	3	7.14%
tronco	1	2.38%
	42	
feminino		
outras	10	28.57%
raiz	8	22.86%
tronco	4	11.43%
toco	4	11.43%
não respostas	3	8.57%
touceira	2	5.71%
soca/ soque/ soqueira	2	5.71%
talo	2	5.71%
	35	

PT: 7PT: 209/2, 215/1, 215/2, 216/4, 219/4, 221/1, 222/1

Fonte: Banco de dados do ALiB-PR.

A distribuição dos dados segundo a variável sexo indica algumas tendências, ainda que discretas. A presença de variantes diretamente relacionadas ao objetivo da pergunta mostra que os homens, embora moradores urbanos, estão mais familiarizados com denominações específicas para o referente, elegendo, em ordem decrescente: *soca* (31%), *toco* (17%), *touceira* (10%), *raiz* (10%), *talo* (7%) e *não respostas* (7%). As mulheres, por sua vez, elencaram: *raiz* (23%), *tronco* (11%), *toco* (11%), *não respostas* (9%) e *touceira* (6%).

5.2.8 Análise da Variável Faixa Etária nos Dados do ALiB-PR

Na faixa etária 1 (informantes mais jovens: 18 a 30 anos), registramos 37 ocorrências distribuídas por 17 denominações, além de cinco *não respostas* (14%). A variante mais frequente foi *toco*, com oito registros (22%); seguida de *raiz*, com sete (19%); *talo*, com quatro (11%); *tronco*, com dois (5%). Já *soca* (e suas variantes *soque* e *soqueira*) teve apenas um registro (3%). Outras denominações – *estalo, muda, bagaço, broto, tiguera, rama, caule, restos, semente, sobras, cepa* e *cepo* – foram citadas uma única vez cada (27%).

Na faixa etária 2 (informantes mais velhos: 50 a 65 anos), os dados revelaram 40 registros e uma *não resposta*. A variante mais recorrente foi *soca* e suas variantes morfofonêmicas (*soque* e *soqueira*), com 14 ocorrências (35%); *touceira* (e a forma *toça*), com seis registros (15%); *raiz*, com cinco (12%); *tronco*, com três (7%). Outras variantes, como *grãos, talo, tiguera, capim, semente* e *centelho de trigo*, foram mencionadas por apenas um informante cada (17%), o que sugere possível desconhecimento do referente e de sua denominação no contexto urbano.

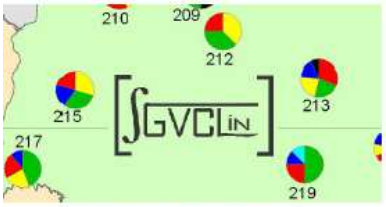
Os resultados apontam maior apagamento lexical na faixa 1, enquanto a faixa 2 apresenta maior manutenção. Isso indica que a variável idade desempenha papel determinante nas escolhas linguísticas observadas nesta pesquisa. Pode-se inferir ainda que os participantes da faixa 2 mantiveram contato mais direto com o meio rural ou residiram em cidades cuja economia se baseia no agronegócio, com reflexos na linguagem. Esse vínculo com práticas agrárias parece ter influenciado a escolha da variante predominante *soca* (e suas variantes *soque* e *soqueira*). Ressaltamos a baixa incidência de *toco* na faixa 1 (22%) em comparação com *soca* na faixa 2 (35%).

Tais resultados revelam que os informantes mais velhos preservam em sua memória afetiva marcas de ruralidade associadas a um contexto mais campesino, tradicional e conservador, social e linguisticamente. Apesar da baixa incidência de *não*

respostas em ambos os grupos, elas foram mais expressivas entre os informantes mais jovens, apontando para um cotidiano indissociável das práticas urbanizadas, no qual o dinamismo e a fluidez das relações intersociais reverberam na língua, tornando-a mais flexível, mutável e intercambiável, favorecendo apagamentos ou alterações.

Para sintetizar, apresentamos a Figura 24, que mostra o relatório com as ocorrências das variantes de *soca*, em números absolutos e percentuais, bem como a carta (Figura 25), que apresenta a distribuição espacial das variantes nos dados do ALiB-PR segundo a variável faixa etária.

Figura 24 – Relatório das variantes para *soca* nos dados do ALiB-PR segundo a variável faixa etária



Geral

Número da questão: 47

Questão Depois que se corta o pé de arroz ou de fumo, fica uma parte embaixo assim na terra. Você já chegou a ver plantação assim?

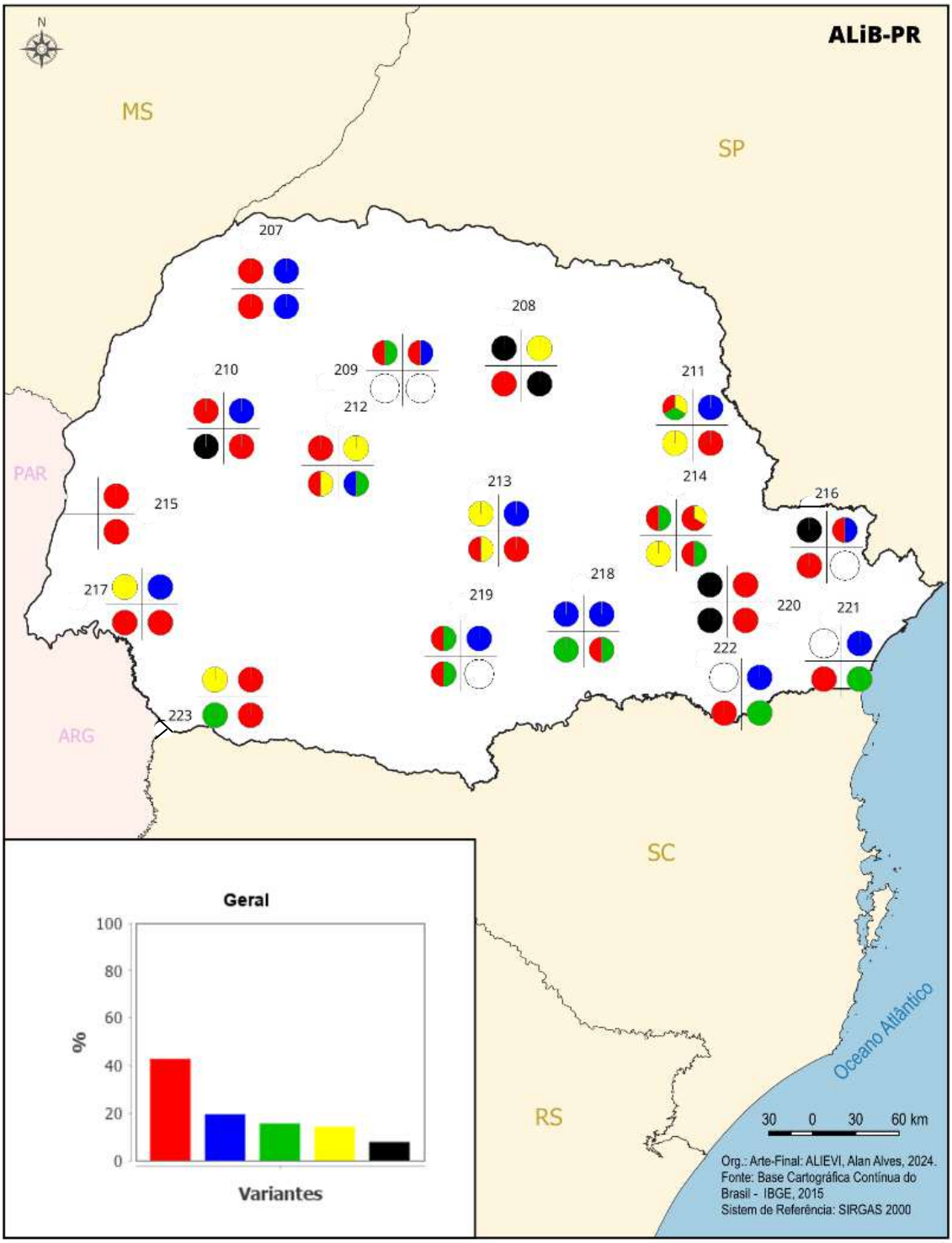
Variantes	Número de ocorrências	%
Faixa I		
outras	10	27.03%
toco	8	21.62%
raiz	7	18.92%
não respostas	5	13.51%
talo	4	10.81%
tronco	2	5.41%
soca/ soque/ soqueira	1	2.70%
	37	
Faixa II		
soca/ soque/ soqueira	14	35.00%
outras	7	17.50%
touceira	6	15.00%
raiz	5	12.50%
tronco	3	7.50%
toco	3	7.50%
talo	1	2.50%
não respostas	1	2.50%
	40	

PT: 7PT: 209/2, 215/1, 215/2, 216/4, 219/4, 221/1, 222/1

Fonte: Banco de dados do ALiB-PR.

A Figura 25 retrata, por meio da carta pluridimensional, a distribuição das variantes de *soca* nos dados do ALiB-PR, segundo as variáveis sexo e faixa etária agrupadas.

Figura 25 – Carta experimental com a distribuição das variantes para soca no ALiB-PR segundo as variáveis sexo e faixa etária



Fonte: Banco de dados do ALiB-PR.

Em relação à variação diacrônica, os dados da Figura 25 demonstram que, no ALiB-BA, *toco* foi a variante mais frequente, representando 23%, seguido de *soca/soqueira* (9%), *tronco* e *talo*, evidenciando também uma ampliação da diversidade lexical, mas com predominância distinta, associada a uma realidade geocultural diferente. Todavia, as *não respostas* foram bastante expressivas (39%), o que sugere um maior apagamento dessas variantes agropastoris no falar urbano da Bahia.

Já no ALPR, a variante *soca* predominou com 51% das ocorrências, sugerindo um contexto rural conservador, com menor diversidade lexical. No ALiB-PR, embora *soca* ainda se mantenha relevante (19%), observa-se uma ampliação das variantes lexicais, com destaque para *raiz, toco, touceira/toça*, além de *talo* e *tronco*, refletindo maior diversidade. As *não respostas* no Paraná apresentaram baixa incidência – apenas 3% no ALPR e 8% no ALiB-PR –, o que aponta para uma marca de ruralidade mais consolidada no falar paranaense, ainda que em contexto urbano. Essa comparação ilustra a variação linguística diacrônica, com mudanças no uso e na diversidade lexical ao longo do tempo.

Quanto à variação diatópica, notam-se alterações de acordo com a localização (rural x urbana). No ALPR, há predominância de termos técnicos e específicos, como *soca/soqueira, resteva* e *cepa*, que refletem práticas agrícolas tradicionais e contato direto com o meio natural. Essas variantes estão profundamente ligadas à atividade produtiva e às interações cotidianas com o ambiente rural. Além disso, o uso continuado de expressões locais demonstra preservação cultural, frequentemente incentivada pela oralidade e pela transmissão geracional, o que reforça a identidade linguística das comunidades rurais.

No ALiB-BA e no ALiB-PR, observa-se maior ampliação lexical e predominância de variantes genéricas, fenômeno frequentemente associado ao desconhecimento do referente ou de sua denominação específica. Esse comportamento reflete o distanciamento das práticas agrícolas e a proximidade com o cotidiano urbano. No ALiB-BA, por exemplo, o apagamento lexical se mostra mais evidente: 37 informantes declararam desconhecer o nome relativo ao objeto da QSL 047, corroborando a expectativa de maior incidência desse fenômeno em ambientes urbanos. Em contrapartida, no ALiB-PR, o número reduzido de *não respostas* (apenas seis informantes) pode ser explicado pela forte tradição agrária do estado, reconhecido

como um dos maiores produtores de grãos do Brasil, mesmo após a urbanização consolidada na década de 1980.

Quanto à variável sexo, os dados evidenciam diferenças significativas. No ALiB-BA, *toco* é a variante mais recorrente na fala masculina (30%), seguido por *não respostas* (26%) e por *soca/soqueira*. Na fala feminina, as *não respostas* foram majoritárias (52%), revelando maior afastamento dessas variantes agropastoris. Entre as respostas válidas, *toco* (15%) foi a mais frequente, seguida de *tronco*, o que indica baixa produtividade dessa variável na Bahia, dada a similaridade entre homens e mulheres. No ALPR, a variante *soca* predominou, com leve preferência masculina (52%) em comparação à feminina (49%), refletindo uso mais enraizado em contexto rural. No ALiB-PR, embora *soca* seja a mais frequente na fala masculina (31%), é seguida de *toco, touceira* e *toça*, demonstrando maior diversidade lexical. Já entre as mulheres, a variante mais produtiva foi *raiz* (23%), seguida de *tronco* e *toco*. Tal divergência sugere que a variável sexo teve papel relevante na análise dos dados do ALiB-PR.

Nos dados do ALiB-PR da faixa etária I (18 a 30 anos), *toco* foi a mais frequente (22%), seguido de *raiz* (19%), *não respostas* (14%), *talo* (11%) e *tronco* (5%). Isso indica maior variação lexical e uso de formas mais gerais entre os jovens. Na faixa etária II (50 a 65 anos), *soca* e suas variantes fonéticas (*soque* e *soqueira*) foram as mais recorrentes (35%), seguidas de *touceira* e *raiz*. Outras variantes, como *tronco* (7%) e *toco* (7%), também foram mencionadas, refletindo a manutenção de termos mais específicos, possivelmente vinculados ao contato dos informantes mais velhos com o meio rural.

Nos dados do ALiB-BA da faixa etária I (18 a 30 anos), as *não respostas* corresponderam a mais da metade (55%). A variante *toco* foi a mais frequente (18%), seguida de *tronco* (9%) e *raiz* (7%). Já na faixa etária II (50 a 65 anos), *toco* foi novamente a mais comum (28%), seguida de *não respostas* (26%), *soca/soqueira* (17%), *tronco* e *talo*. Esse padrão reforça a tendência de maior uso de variantes específicas na faixa etária mais velha, sugerindo vínculo mais próximo com práticas agrícolas.

Esses dados confirmam que a faixa etária exerce papel determinante nas escolhas lexicais: as gerações mais velhas preservam termos específicos, enquanto as mais jovens recorrem a variantes mais genéricas.

5.2.9 Análise Comparativa dos Dados do ALiB-BA, ALPR e ALiB-PR

Para compreendermos melhor a diversidade lexical, elaboramos a Tabela 9, com os percentuais de ocorrência das variantes mais produtivas em cada uma das fontes pesquisadas.

Tabela 9 – Percentual das variantes de soca no ALiB-BA, ALPR e ALiB-PR

Variante	Toco	Soca	Resteva	Cepa	Outras	NR
ALiB-BA	23%	9%	0%	2%	6%	39%
ALPR	19%	51%	8%	8%	0%	3%
ALiB-PR	14%	19%	16%	4%	22%	8%

Fonte: Dados da pesquisa.

Da análise comparativa das fontes, ressaltamos o elevado índice de *não respostas* no ALiB-BA (39%) e a baixa incidência no Paraná, sobretudo no contexto rural (3%). As variantes mais produtivas foram: *toco* (23%) no ALiB-BA; *soca* (51%) no ALPR, refletindo maior domínio de termos tradicionais pelos informantes do meio rural; e *soca* (19%) no ALiB-PR, revelando um decréscimo dessa lexia na fala urbana paranaense. Destacamos ainda a ocorrência de *resteva* apenas no ALPR (8%) e no ALiB-PR (16%), o que pode ser interpretado como um regionalismo intrínseco ao Paraná. A maior heterogeneidade lexical (*outras variantes*) ocorreu nos dados do ALiB e, sobremaneira, no ALiB-PR (22%), confirmando a expectativa de que o distanciamento das práticas campesinas repercute no uso de termos mais genéricos.

5.3 QSL 049 Onde ficam os grãos do feijão, no pé, antes de serem colhidos?

Nesta seção, analisamos as ocorrências da QSL 049 do ALiB, referente à *vagem*, fruto do feijoeiro amplamente utilizado na alimentação brasileira, ilustrada na Figura 26, em três fontes: ALiB-BA, ALPR e ALiB-PR.

Figura 26 – Ilustração de *vagem*

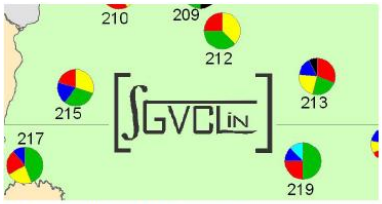
Fonte: Getty Images.

5.3.1 Análise das Variantes de *Vagem* nos Dados do ALiB-BA

Na fala dos informantes do ALiB-BA, houve 93 ocorrências, das quais 84 foram respostas válidas, dois problemas técnicos (ponto 94, informante 4; ponto 100, informante 3) e nove *não respostas* (ponto 82, informante 2; ponto 87, informante 2; ponto 91, informante 2; ponto 93, informantes 1 e 4; ponto 96, informante 1; ponto 101, informante 2; ponto 102, informante 3).

A variante majoritária foi *vagem* e sua realização fonética *bagem*, com 72 ocorrências (77%); *não respostas*, com nove (10%); *casca*, com oito registros (9%); e demais lexias agrupadas sob a denominação *outras* (*capote, coras de feijão, na vara* e *palha*), cada uma com apenas uma menção, totalizando juntas quatro ocorrências (4%). A Figura 27 mostra a distribuição das variantes em números absolutos e percentuais.

Figura 27 – Relatório com a distribuição das variantes para *vagem* no ALiB-BA segundo a variável diatópica



Geral

Número da questão: 49

Questão Onde que ficam os grãos de feijão, no pé, antes de serem colhidos?

Variantes	Número de ocorrências	%
bagem/ vagem	72	77.42%
não respostas	9	9.68%
casca	8	8.60%
outras	4	4.30%
	93	

PT: 2PT: 94/4, 100/3

Fonte: Banco de dados do ALiB-BA.

De acordo com a Figura 27, observamos que *vagem/bagem* está presente na maioria dos pontos da rede de localidades da Bahia, seja como variante única (pontos 83, 84, 85, 86, 89 e 98), seja em concorrência com outras variantes ou com ausência de respostas (nos demais pontos).

Chama atenção o número relativamente baixo de *não respostas*, considerando que, em contextos urbanos, entrevistados afastados do vínculo direto com o campo e suas atividades tendem a desconhecer variantes ligadas ao universo agropastoril.

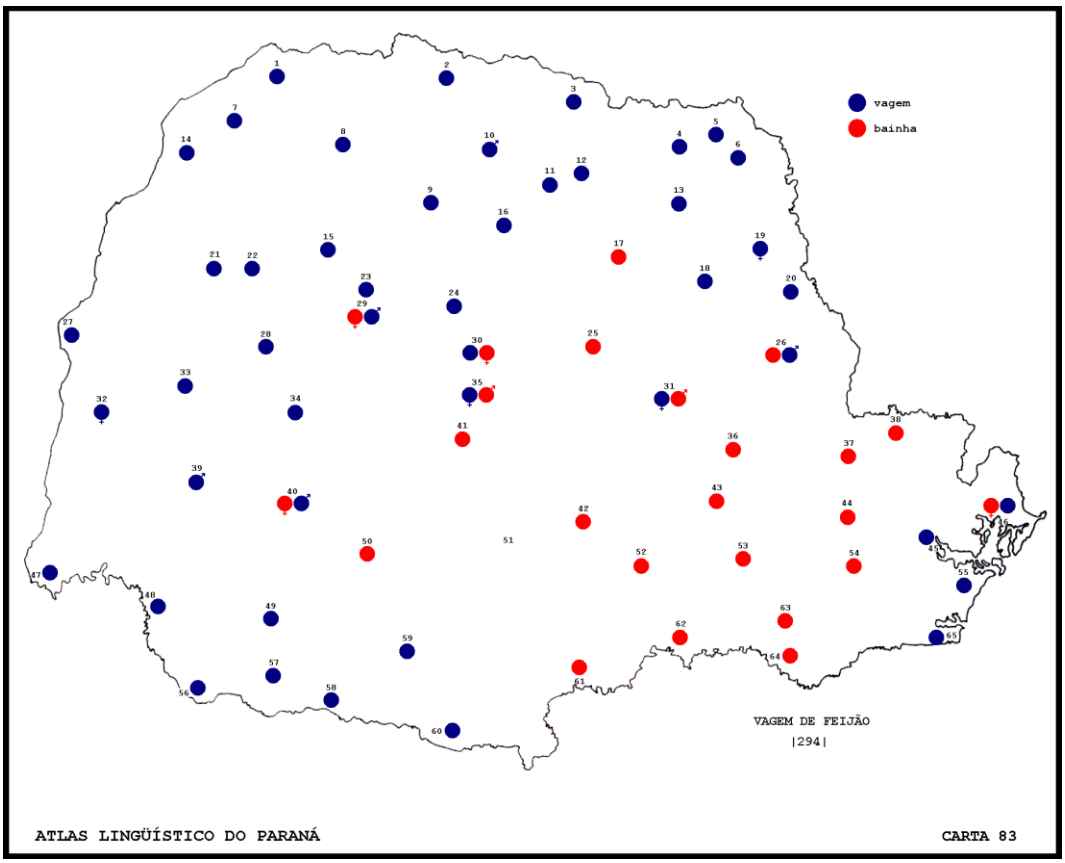
O termo *vagem/bagem*, integrado ao vocabulário rural baiano, pode ter sido amplamente perpetuado ao longo das gerações, consolidando-se em comunidades rurais e, posteriormente, expandindo-se para os centros urbanos. Apesar de a população urbana da Bahia alcançar 67,12% (IBGE, 2022), o estado preserva marcas linguísticas ligadas ao campo, especialmente em pequenas localidades do interior. Essas áreas, com feiras tradicionais realizadas em espaços públicos, constituem polos de interação sociolinguística, nos quais itens hortifrutigranjeiros e agropecuários são comercializados. Essa dinâmica favorece o intercâmbio vocabular entre moradores da zona rural e da zona urbana, evidenciando processos de manutenção e ampliação dessas variantes linguísticas no acervo lexical baiano.

5.3.2 Análise das Variantes de *Vagem* nos Dados do ALPR

As variantes para *vagem* no ALPR estão representadas na Carta 83 (Aguilera, 1994), que demonstra a concorrência de apenas duas formas: *vagem* e *bainha*. As

variantes fonéticas de *vagem* constam da Carta 137. A Figura 28 apresenta essas variantes no ALPR.

Figura 28 – Carta 83 do Atlas Linguístico do Paraná (ALPR) com as variantes de vagem



Fonte: Aguilera (1994).

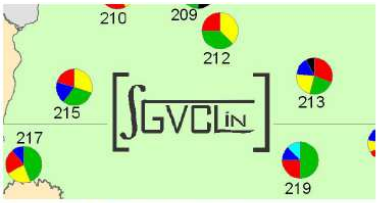
Verificamos que, em todos os 65 pontos, foi registrada pelo menos uma variante.[10] A distribuição areal de ambas deixa claro que a variável predominante é a diatópica, haja vista a concentração de *vagem* no espaço denominado Paraná Moderno e de *bainha* no Paraná Tradicional, conforme será detalhado adiante.

Nos dados do ALPR, houve 127 respostas para a QSL 294 e quatro ausências.[11] A variante mais produtiva foi *vagem*, com 85 registros (65%), seguida de *bainha*, com 42 respostas (32%). A Figura 29 mostra o relatório com a distribuição dos dados em números absolutos e percentuais, conforme apresentado pelo SGVCLin.

[10] No ponto 51- não constam, por um lapso na cartografação, as respostas da Informante A: na **bage**... dentro do chão...da **bainha;** e do informante B: da **bainha**.

[11] A questão não foi formulada para os informantes 10A, 32B e 39A e houve um problema na gravação do informante 19B.

Figura 29 – Relatório com a distribuição das variantes de vagem no ALPR



Geral

Número da questão: 294

Questão Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?

Variantes	Número de ocorrências	%
vagem	85	64.89%
bainha	42	32.06%
não respostas	4	3.05%
	131	

PT: 2PT: 51/1, 51/2

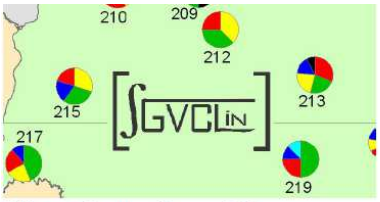
Fonte: Banco de dados do ALPR.

5.3.3 Análise da Variável Sexo de Vagem nos Dados do ALPR

Dos 85 registros de *vagem*, 42 correspondem à fala feminina (64%) e 43 à fala masculina (66%). Já das 42 ocorrências de *bainha*, 22 foram registradas na fala de mulheres (33%) e 20 na de homens (31%). Os índices não são muito discrepantes, embora se perceba uma tendência das mulheres deste *corpus* a reproduzirem com maior frequência a forma *bainha*.

Para melhor compreensão dos resultados, elaboramos a Figura 30, que traz o relatório das ocorrências das variantes de *vagem* no ALPR, com as respectivas porcentagens.

Figura 30 – Relatório com a distribuição das variantes de *vagem* no ALPR segundo a variável sexo



Geral

Número da questão: 294

Questão Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?

Variantes	Número de ocorrências	%
feminino		
vagem	42	63.64%
bainha	22	33.33%
não respostas	2	3.03%
	66	
masculino		
vagem	43	66.15%
bainha	20	30.77%
não respostas	2	3.08%
	65	

PT: 2PT: 51/1, 51/2

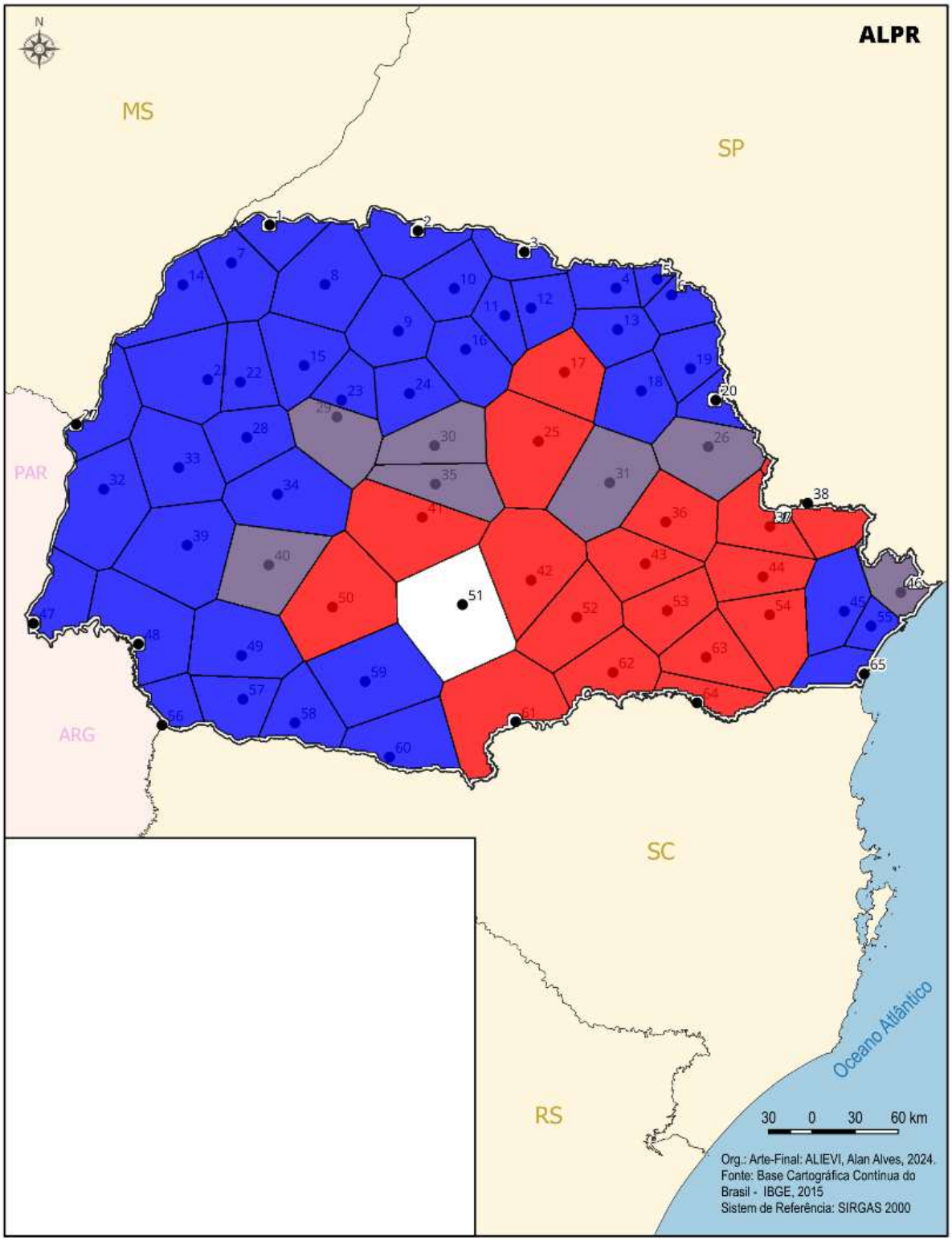
Fonte: Banco de dados do ALPR.

Conforme observamos na Carta 83 do ALPR (Figura 31), *bainha* foi registrada no território paranaense com distribuição mais densa no espaço denominado por Cardoso e Westphalen (1986) como Paraná Tradicional, representado pelas localidades mais antigas do estado, dispostas no eixo litoral–centro–nordeste. Já no Paraná Moderno, a variante mais frequente é *vagem*.

Os fatores sociais e econômicos associados aos processos de modernização parecem ter desempenhado papel significativo na consolidação do termo *vagem* no léxico do Paraná Moderno. Em contrapartida, a predominância de *bainha* no Paraná Tradicional pode ser atribuída à persistência de práticas agrícolas mais convencionais, refletindo um repertório lexical conservador preservado em comunidades linguísticas com menor exposição a influências externas. Assim, concluímos que a variável sexo não apresentou relevância significativa no cenário analisado, uma vez que os dados coletados de informantes de ambos os sexos demonstraram equivalência nos padrões de uso.

Para melhor interpretação desse fenômeno linguístico, elaboramos a carta isoléxica (Figura 31), com duas isoglossas bem delimitadas e distribuídas pelo território paranaense: *vagem* e *bainha*.

Figura 31 – Carta de isoléxicas das variantes de *vagem* no ALPR



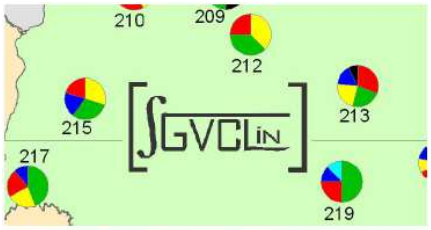
Fonte: Banco de dados do ALPR.

A delimitação das áreas de uso lexical nesta carta fundamenta-se na identificação de isoglossas que marcam os limites geográficos entre as variantes. Nota-se uma distribuição territorial que reflete a dinâmica geolinguística da região: *vagem* predomina na porção norte e oeste do estado, enquanto *bainha* prevalece no centro e sudeste. As áreas em cinza indicam zonas de transição, nas quais ambas coexistem em frequências próximas. Essa coexistência lexical aponta para processos de interferência ou adaptação linguística, comuns em contextos de contato cultural e histórico.

5.3.4 Análise das Variantes de Vagem nos Dados do ALIB-PR

Na fala dos informantes da zona urbana do Paraná, houve 73 ocorrências, duas *não respostas* (ponto 216, informante 1; ponto 220, informante 2) e três problemas técnicos (ponto 209, informante 2; ponto 219, informante 1; ponto 221, informante 2). A variante mais produtiva foi *vagem*, juntamente com sua variante fonética *bagem*, com 49 ocorrências (65%). A segunda foi *bainha*, com 20 registros (27%). Outras variantes (*casca, na ponta, palha* e *talo*), com apenas uma ocorrência cada, foram agrupadas em razão da baixa incidência, totalizando 5%. Essas formas podem sinalizar seu gradativo apagamento ou, ainda, revelar o desconhecimento do referente e de sua denominação por parte dos informantes. A Figura 32 mostra o relatório da distribuição das variantes de *vagem* no ALiB-PR.

Figura 32 – Relatório com a distribuição das variantes para *vagem* no ALiB-PR segundo a variável diatópica



Geral

Número da questão: 49

Questão Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?

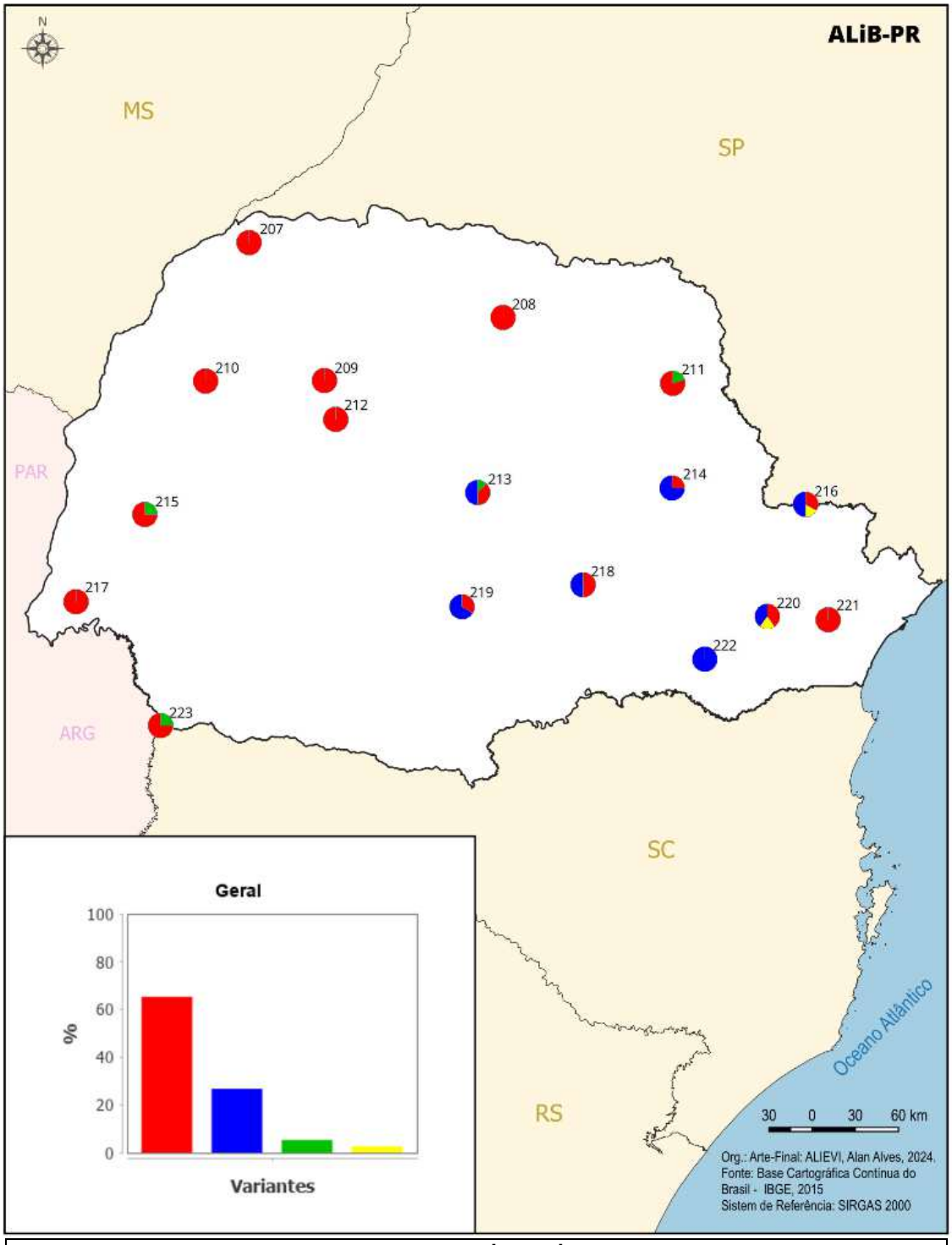
Variantes	Número de ocorrências	%
bagem/ vagem	49	65.33%
bainha	20	26.67%
casca/ na ponta/ palha/ talo	4	5.33%
não respostas	2	2.67%
	75	

PT: 3PT: 209/2, 219/1, 221/2

Fonte: Elaborada pelo autor com dados do ALiB-PR.

Para melhor visualização das ocorrências de *vagem* no ALiB-PR, elaboramos a carta monodimensional (Figura 33).

Figura 33 – Carta experimental das variantes de vagem no ALiB-PR segundo a variável diatópica



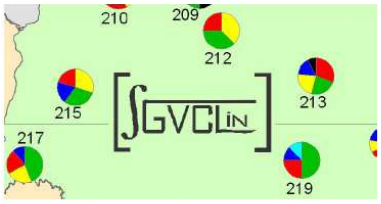
Fonte: Banco de dados do ALiB-PR.

De acordo com os dados analisados e a carta monodimensional (Figura 33), notamos que *vagem* e sua variante fonética *bagem* foram as formas mais produtivas, registradas em quase todos os pontos, com exceção do ponto 222, em que foi elicitada apenas *bainha*. Esse dado pode indicar a baixa incidência de *bainha* nos centros urbanos, confirmando sua maior frequência em pontos mais isolados e menos sujeitos à variação.

5.3.5 Análise das Variantes de *Vagem* no ALiB-PR Segundo a Variável Sexo

Das 38 ocorrências na fala masculina, 26 foram para *bagem/vagem* (68%), dez para *bainha* (26%), apenas uma para *não resposta* (3%) e uma para *outras* (3%). Das 37 ocorrências femininas, 23 foram para *bagem/vagem* (62%), dez para *bainha* (27%), três para outras variantes com apenas um registro cada (8%) e uma para *não resposta* (3%). A Figura 34 apresenta esses dados em números absolutos e percentuais.

Figura 34 – Relatório com a distribuição das variantes para *vagem* no ALiB-PR segundo a variável sexo



Geral

Número da questão: 49

Questão Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?

Variantes	Número de ocorrências	%
masculino		
bagem/ vagem	26	68.42%
bainha	10	26.32%
não respostas	1	2.63%
casca/ na ponta/ palha/ talo	1	2.63%
	38	
feminino		
bagem/ vagem	23	62.16%
bainha	10	27.03%
casca/ na ponta/ palha/ talo	3	8.11%
não respostas	1	2.70%
	37	

PT: 3PT: 209/2, 219/1, 221/2

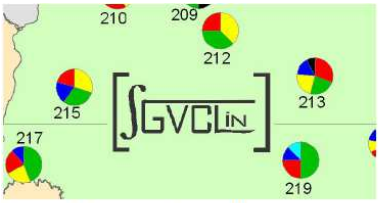
Fonte: Banco de dados do ALPR.

Ao confrontar os resultados, concluímos que a variável sexo não exerceu influência significativa sobre os dados do ALiB-PR, já que as proporções entre homens e mulheres são bastante semelhantes. Dessa forma, julgamos desnecessário um aprofundamento maior dessa variável na presente dissertação.

5.3.6 Análise das Variantes de *Vagem* no ALiB-PR Segundo a Variável Faixa Etária

Na faixa 1 (informantes mais jovens), houve 36 ocorrências: 24 para *vagem* (67%), sete para *bainha* (19%), três para *outras* (8%) e duas *não respostas* (6%). Na faixa 2 (informantes mais velhos), registraram-se 39 ocorrências: 25 para *vagem* (64%), 13 para *bainha* (33%) e apenas uma para *outras* (3%). Nessa faixa, não houve *não respostas*. A Figura 35 sintetiza os resultados em números e porcentagens.

Figura 35 – Relatório com a distribuição das variantes para vagem no ALiB-PR segundo a variável faixa etária



Geral

Número da questão: 49

Questão Onde que ficam os grãos do feijão, no pé, antes de serem colhidos?

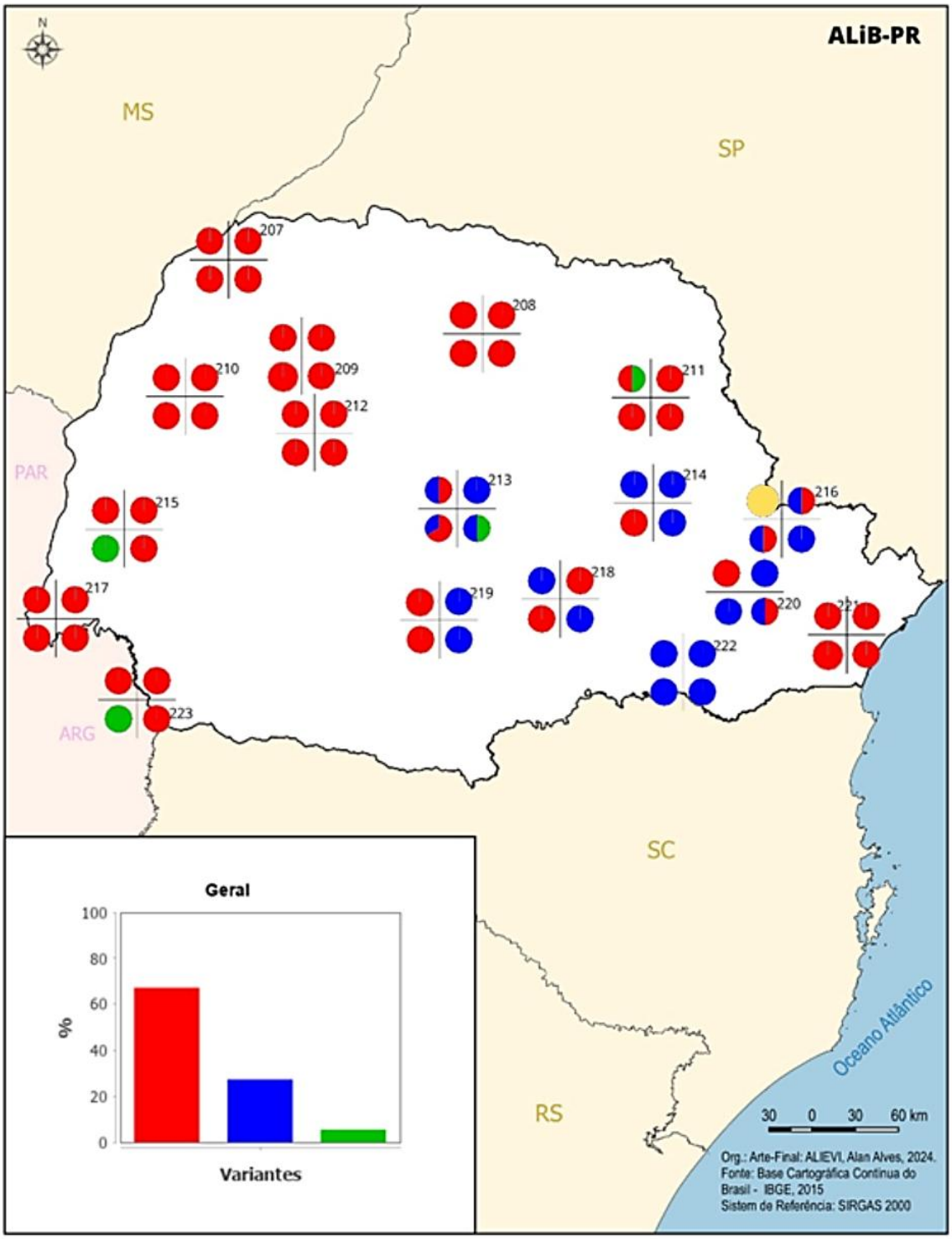
Variantes	Número de ocorrências	%
Faixa I		
bagem/ vagem	24	66.67%
bainha	7	19.44%
casca/ na ponta/ palha/ talo	3	8.33%
não respostas	2	5.56%
	36	
Faixa II		
bagem/ vagem	25	64.10%
bainha	13	33.33%
casca/ na ponta/ palha/ talo	1	2.56%
	39	

PT: 3PT: 209/2, 219/1, 221/2

Fonte: Banco de dados do ALPR.

A carta experimental (Figura 36) apresenta a distribuição das variantes por localidade, sexo e faixa etária dos informantes.

Figura 36 – Carta experimental com as variantes de vagem no ALiB-PR segundo as variáveis sexo e faixa etária



Fonte: Banco de dados do ALiB-PR.

A análise da carta experimental (Figura 36) evidencia a predominância de *vagem/bagem* na maior parte da rede de pontos do ALiB-PR, ora presente em ambos os sexos e faixas etárias, ora restrita a uma dessas variáveis. Tal presença significativa sugere que fatores como redes sociais, mídia e o fluxo contínuo nas cidades podem ter contribuído para a popularização dessa lexia, especialmente na forma padrão *vagem*, possivelmente motivada pela escolarização dos informantes. Já *bainha* foi registrada de forma residual em localidades do sudeste e do centro do estado, geralmente em associação a outra lexia, com exceção do ponto 222, em que se apresentou isoladamente.

5.3.7 Análise Comparativa dos Dados do ALiB-BA, ALPR e ALiB-PR acerca das Variantes de *Vagem*

Os dados dos três atlas linguísticos analisados – ALPR, ALiB-PR e ALiB-BA – revelam como fatores geográficos (urbano x rural) e sociais (idade e sexo) moldam o vocabulário. Em todos os casos, observamos que a variante *vagem* está mais associada a contextos urbanos e modernizados, enquanto *bainha* aparece vinculada a um vocabulário rural e conservador, típico das práticas agrícolas tradicionais. A análise sugere que o ambiente urbano exerce impacto significativo na evolução das variantes linguísticas, com *vagem* prevalecendo em áreas urbanas e *bainha*, na zona rural.

Esses padrões indicam que, além das influências geográficas, aspectos como urbanização, escolarização e dinâmicas sociolinguísticas desempenham papéis essenciais na manutenção ou perda de determinadas lexias. A resistência de termos como *bainha* em áreas rurais também aponta para a persistência de uma cultura agrária que, mesmo diante da globalização, preserva práticas e vocabulário característicos.

A comparação entre os *corpora* permite compreender melhor as variações lexicais relativas a *vagem*, conforme descrito na Tabela 10.

Tabela 10 – Percentual das variantes de vagem no ALiB-BA, ALPR e ALiB-PR

Variante	Vagem	Bainha	Casca	Outras	NR
ALiB-BA	77%	0%	9%	4%	10%
ALPR	65%	32%	0%	0%	3%
ALiB-PR	65%	27%	1%	4%	3%

Fonte: Dados da pesquisa.

Os dados apontam *vagem* como a variante mais produtiva em todas as fontes: 77% no ALiB-BA e 65% no ALPR e no ALiB-PR. Na Bahia, houve pouca variação, já que 77% das ocorrências corresponderam a *vagem*, 9% a *casca* e 4% a *outras*. *Bainha* não aparece nos registros da Bahia; como já mencionado, essa variante ocorre com maior frequência no Paraná Tradicional, rarefazendo-se nas demais regiões do estado. As *não respostas* foram mais comuns na fala dos baianos (10%) e residuais no Paraná (3%).

5.4 QSL 054 Armação de Madeira que se Coloca no Pescoço do Animal [...] para não Atravessar a Cerca

A Figura 37 ilustra o referente da QSL 054 – *cangalha*, triângulo normalmente de madeira que se coloca no pescoço do porco, do bezerro ou do bode para evitar que atravesse a cerca.

Figura 37 – Ilustração do referente da *cangalha*

Fonte: Dados da pesquisa.

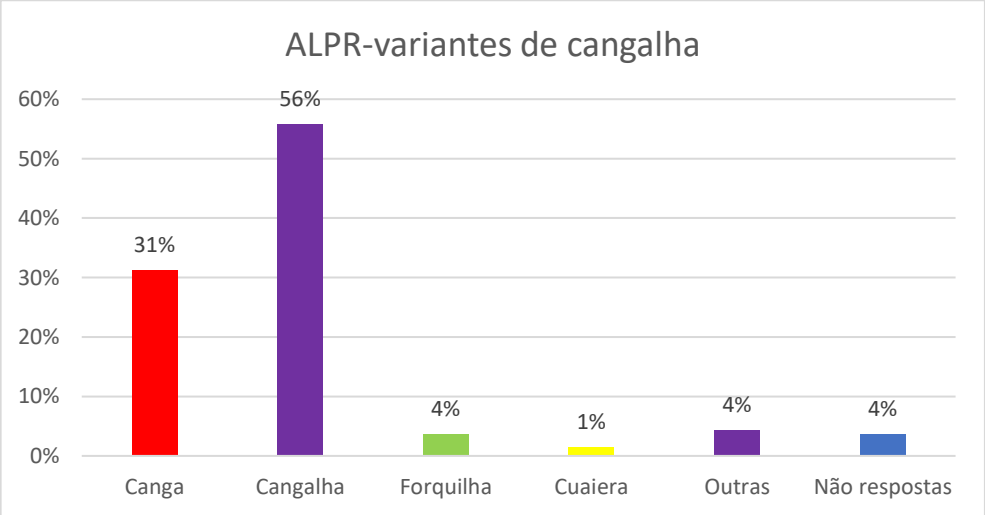
5.4.1 Análise das Variantes de Cangalha nos Dados do ALPR

Iniciamos pela descrição dos dados da Carta 61 do ALPR (Aguilera, 1994), verificando, na fala dos 130 informantes (65 homens e 65 mulheres), as variantes registradas. Na sequência, apresentamos os dados gerais e por sexo por meio de gráficos. Para o cálculo dos percentuais, consideramos o número total de respostas e de *não respostas*, isto é, todo o universo de dados.

5.4.2 Um Retrato da Variação Lexical para *Cangalha* nos Dados do ALPR

A distribuição das variantes para o conceito *cangalha* contou com 138 ocorrências, sendo 133 respostas válidas e cinco *não respostas*. A variante mais produtiva foi *cangalha*, com 77 respostas (56%), seguida de *canga*, com 43 ocorrências (31%). As demais variantes tiveram valores menos expressivos: *forquilha*, mencionada por cinco informantes (4%); *cuaiera*, por dois informantes (1%). Outras seis respostas não foram mapeadas na carta, mas constam nas notas: *varal, gancho, trava, cabresto, jangada* e *focinheira*, cada uma com uma ocorrência. Essas variantes foram agrupadas sob o rótulo *outras* para efeito de análise, totalizando 4% dos dados. Cinco informantes (4%) não conheciam ou não lembravam o nome da peça. O Gráfico 2 apresenta esses resultados em percentuais.

Gráfico 2 – Variantes de *cangalha* nos dados do Atlas Linguístico do Paraná – ALPR

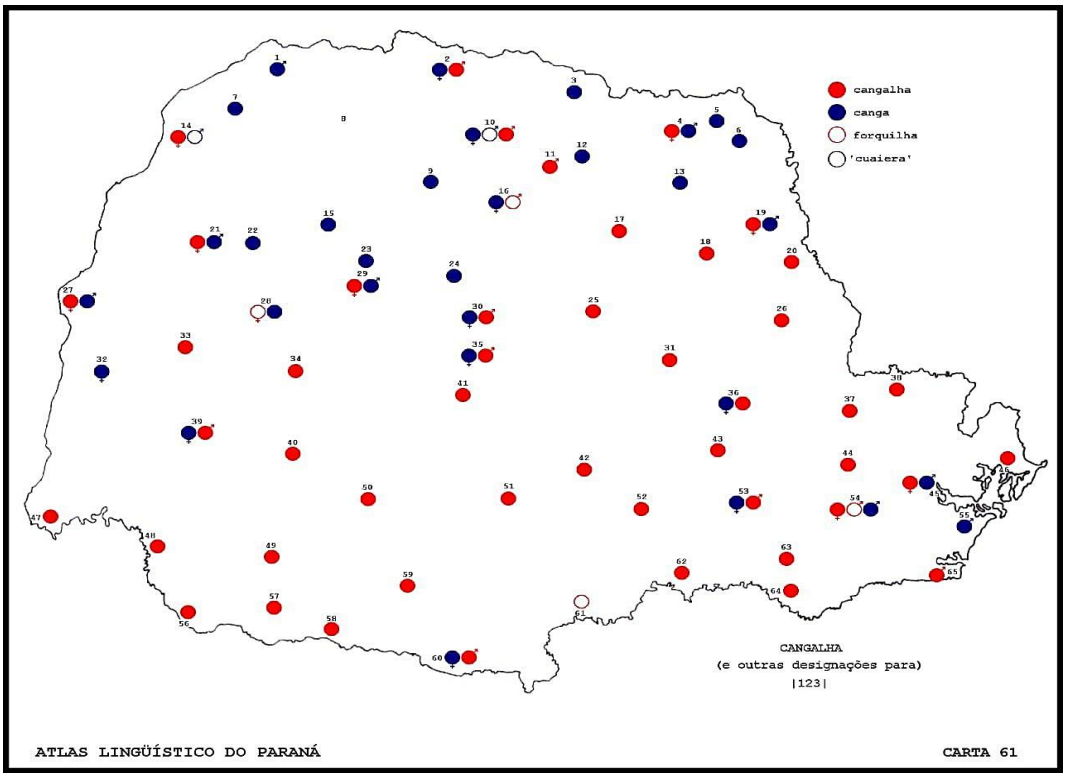


Fonte: Elaborado pelos autores com os dados do ALPR (Aguilera, 1994).

Para visualizar a distribuição, apresentamos a Carta 61 (Aguilera, 1994). Por se tratar de atlas bidimensional, a distinção dos registros femininos (♀) e masculinos

(♂) (diassexual) faz-se por símbolos da biologia. Nos pontos em que não há identificação de sexo, a resposta foi dada por ambos os informantes. Trata-se de carta do tipo misto, na qual as variantes são representadas na legenda por círculos coloridos preenchidos ou vazios, distribuídos pelos pontos da rede. A Figura 38 retrata as ocorrências para *cangalha* no ALPR.

Figura 38 – Carta 61 do ALPR - Variantes para o triângulo que se coloca no pescoço de animais (Questão 123 – ALPR)



Fonte: Aguilera (1994).

Em síntese, o Gráfico 2 revela que, nos dados rurais da década de 1990, prevaleceram as variantes *cangalha* (56%) e *canga* (31%). Por se tratar de um atlas rural, o número de *não respostas* foi reduzido, pois a maioria reconhecia o referente e sua denominação.

Do ponto de vista diatópico, conforme a Carta 61 do ALPR (Figura 38), *cangalha* foi registrada em quase todo o território paranaense, com maior concentração no espaço denominado por Cardoso e Westphalen (1986) de Paraná Tradicional, que abrange as localidades mais antigas do estado (litoral > centro > nordeste). Já no Paraná Moderno, prevalece a variante *canga*.

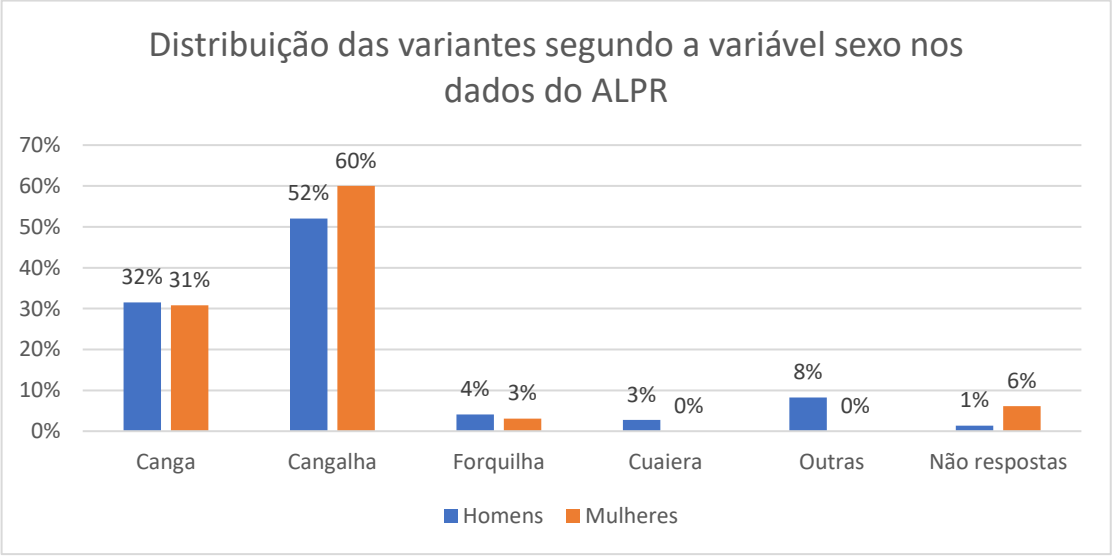
Sobre essa divisão, adotamos o que estabelecem Cardoso e Westphalen (1986). Para os autores, a história do Paraná compreende a formação de três

comunidades regionais: *Paraná Tradicional*, que se esboçou no século XVII com a busca de ouro, estruturou-se no século XVIII com a criação e comércio de gado e se consolidou no século XIX com a extração e comércio da erva-mate e da madeira; e *Paraná Moderno*, correspondente à ocupação do Norte com a cultura do café no início do século XX, ligada a São Paulo, e do *Sudoeste e Oeste* com a criação de suínos e o cultivo de cereais, ligada ao Rio Grande do Sul.

5.4.3 Descrição dos Dados na Perspectiva da Variável Sexo no ALPR

Por se tratar de um atlas bidimensional (variáveis diatópica x sexo), o ALPR, composto de 65 pontos distribuídos pelo estado, entrevistou 130 informantes (65 homens e 65 mulheres), de modo que, em cada localidade, fossem investigados um homem e uma mulher. Das 138 ocorrências, 133 foram respostas válidas e cinco, *não respostas*. O Gráfico 3 descreve as porcentagens de *cangalha* de acordo com a variável sexo no ALPR.

Gráfico 3 – Distribuição das variantes segundo a variável sexo nos dados do ALPR



Fonte: Elaborado pelo autor com os dados do ALPR.

O Gráfico 3 mostra que *cangalha* se distribui de forma praticamente equitativa entre os sexos: 38 respostas na fala masculina (52%) e 39 na feminina (60%). Situação semelhante ocorre com *canga*: das 43 respostas, 23 estão na fala masculina (32%) e 20 na feminina (31%). Quanto a *forquilha*, embora tenha sido registrada em apenas cinco ocorrências, três foram na fala de homens (4%) e duas na de mulheres (3%), o que sugere leve predominância masculina, possivelmente relacionada à maior

familiaridade com o manejo de animais. Essa hipótese também se aplica a *cuaiera*, que aparece apenas na fala masculina (duas ocorrências, 3%). Outras variantes somaram seis respostas (8%), todas registradas entre os homens. As *não respostas* (cinco no total) foram mais incidentes na fala feminina, com quatro ocorrências (6%), contra apenas uma (1%) entre os homens.

A relevância da variável sexo pode ser observada nos casos das quatro *não respostas*: três mulheres afirmaram desconhecer o objeto, enquanto apenas um homem não soube nomeá-lo. Tais resultados podem estar relacionados à divisão de papéis: cabia à mulher o cuidado com a casa, os filhos e as aves no entorno da residência, enquanto ao homem competia a responsabilidade pela terra e pelos animais de maior porte (ovinos, caprinos, equinos e bovinos).

Quanto à influência do sexo sobre a mudança linguística, sociolinguistas como Silva-Corvalán (1989) e Paiva (2004) observaram, em diversos estudos, que ora as mulheres empregam mais as formas padrão e de maior prestígio do que os homens (LABOV, 2008 [1972]); ora o sexo exerce pouca influência sobre a mudança; ora, ainda, são as mulheres que rompem mais a norma do que os homens.

5.4.4 Variação Lexical para *Cangalha* nos Dados do ALIB-PR

A equipe do Atlas Linguístico do Brasil – Regional PR investigou 68 informantes (34 homens e 34 mulheres) em 17 localidades da rede paranaense, computando 70 respostas[12] para análise, incluindo as de quatro informantes com nível fundamental da capital. Para o cálculo dos percentuais, consideramos o total de respostas e de *não respostas*, isto é, todo o universo de dados.

Trata-se de um atlas pluridimensional, que analisa não apenas a variável diatópica, mas também as variáveis sexo (masculino x feminino) e faixa etária (Faixa I: 18 a 30 anos; Faixa II: 50 a 65 anos). Iniciamos com uma análise panorâmica do ALiB-PR e, em seguida, discutimos os resultados das variáveis sociais.

Por razões metodológicas e didáticas, descrevemos separadamente as variáveis sociais, o que permite melhor interpretação dos dados, inclusive com a elaboração de cartas lexicais. Assim, primeiramente analisamos a *variável sexo* e, depois, a *variável faixa etária*.

[12] Dois informantes deram mais de uma resposta.

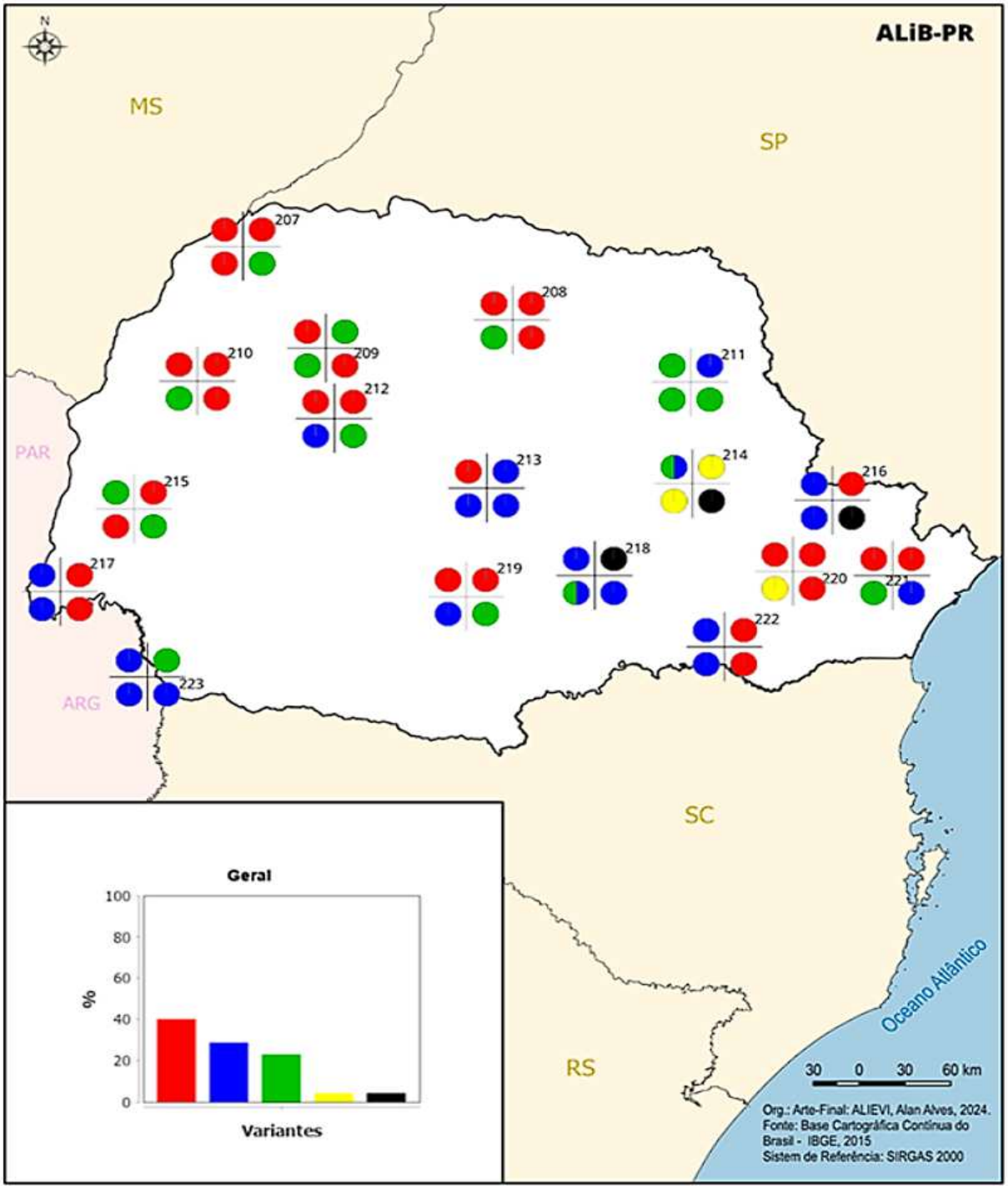
5.5.5 Um Panorama dos Dados para Cangalha no ALiB-PR

Para demonstrar a influência das variáveis em questão (espaço rural e urbano, sexo e faixa etária[13]), elaboramos a carta experimental da Figura 39. Nela, os informantes 1 e 3 representam o sexo masculino e os informantes 2 e 4, o sexo feminino.

Quanto à faixa etária, convencionamos dispor os informantes mais jovens (1 e 2 – Faixa I: 18 a 30 anos) na parte superior do ponto e os mais velhos (3 e 4 – Faixa II: 50 a 65 anos) na parte inferior (Figura 39).

[13] A variável faixa etária foi incluída apenas na coleta do ALiB.

Figura 39 – Carta experimental com as variantes de cangalha nos dados do ALiB-PR segundo as variáveis sexo e faixa etária



54 – Como se chama a armação de madeira, que se coloca no pescoço dos animais, pode ser um porco, um bezerro, um carneiro para ele não atravessar a cerca?

Legenda

- não resposta
- cangalha
- canga
- forquilha
- outras

①②
③④

1: masculino – Faixa I
2: feminino – Faixa I
3: masculino – Faixa II
4: feminino – Faixa II

SGVCLin]® - 2015

Fonte: Banco de dados do ALiB-PR.

Conforme mostra a Carta experimental (Figura 39), a ausência de respostas no ALiB-PR é bastante significativa, não ocorrendo apenas nos pontos 211-Tomazina, 214-Piraí do Sul, 218-Imbituva e 223-Barracão. Esse fato pode estar ligado à história de criação e povoamento desses municípios. Piraí do Sul teve início no século XVII, como pouso de tropeiros no bairro da Lança, no Caminho do Viamão-RS. Imbituva surgiu em 1871, a partir do arraial Cupim, também pouso de tropeiros e marchantes. Essa origem comum pode justificar a presença de respostas em todos os falantes, dada a tradição pastoril. Barracão, fundado em meados do século XX, originou-se de hospedagem para tropeiros de erva-mate e viajantes no início do século. Já Tomazina se formou a partir de terras doadas em 1867 pelo sertanista Joaquim Thomaz Pereira da Silva, vindo de Minas Gerais, atraído pela fertilidade para cultivo do café.

Quanto à distribuição diatópica das formas válidas, *cangalha* e *canga* foram registradas na maioria dos pontos. *Cangalha* aparece preferencialmente no Paraná Tradicional (pontos 214, 216, 217, 218, 219, 220, 221, 222 e 223), mas também em três pontos do Paraná Moderno (211-Tomazina, 212-Campo Mourão, 213-Cândido de Abreu). Já *canga* é mais frequente no Paraná Moderno (207, 208, 209, 210, 211, 212, 213 e 215), embora também ocorra em pontos do Paraná Tradicional (214, 218, 219, 221 e 223). As três ocorrências de *forquilha* foram registradas em dois pontos do Paraná Tradicional: 214-Piraí do Sul e 220-Curitiba.

De acordo com os dados coletados, podemos observar que grande parte dos informantes, totalizando 28, não soube responder à pergunta sobre a armação de madeira em forma de triângulo (40%). Pode-se justificar a significativa ausência de respostas pelo fato de se tratar de atlas que privilegia a linguagem urbana, ao contrário do ALPR que investigou a fala rural. Os informantes, em sua maioria, exercem funções ligadas à vida na cidade, como chapeiro, dona de casa, pedreiro, taxista, costureira etc., por isso desconhecem o nome do referido artefato utilizado para conter o rompimento da cerca pelos quadrúpedes de menor porte.

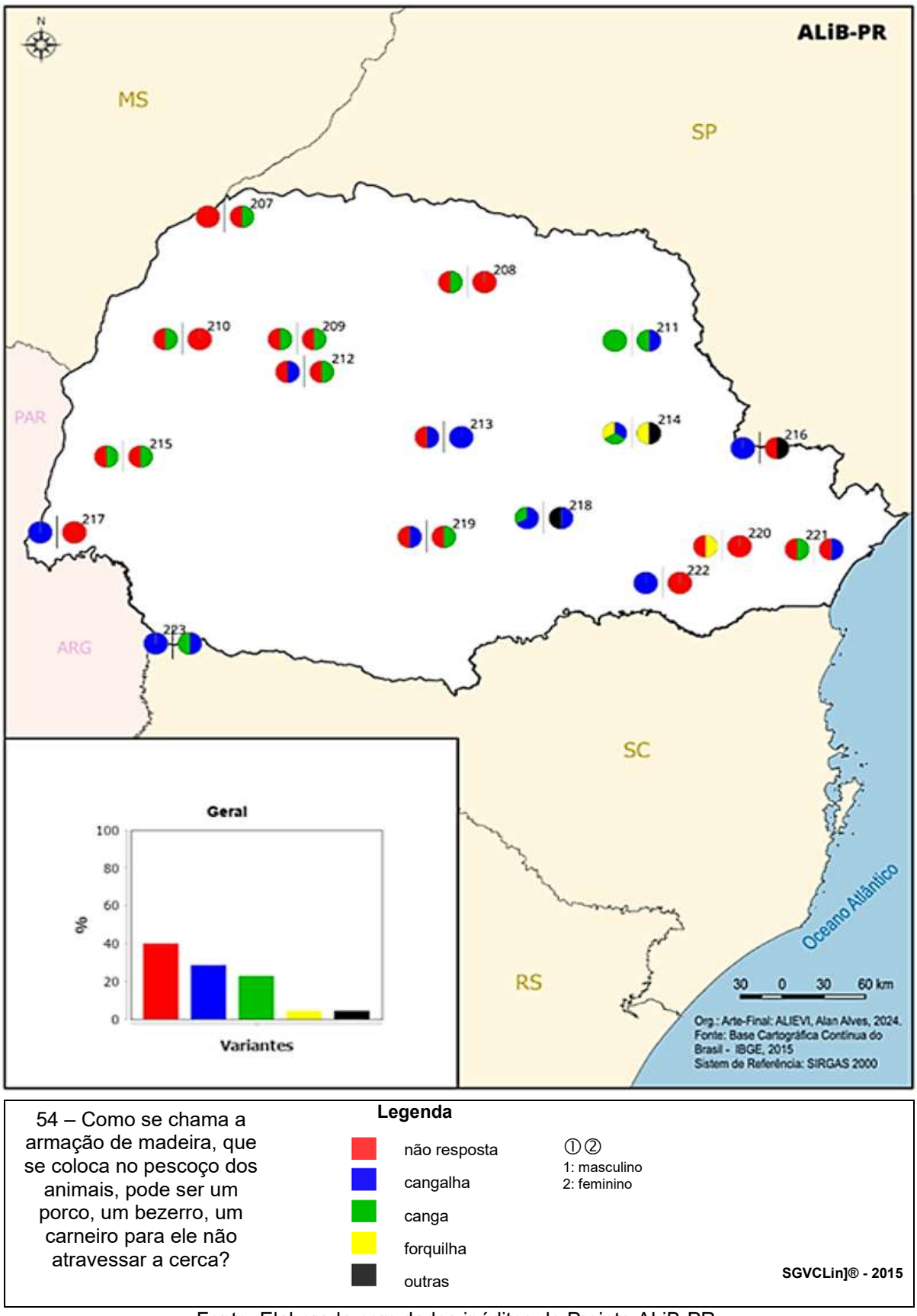
Das 70 ocorrências, 42 foram respostas válidas e 28, *não respostas*. Entre as válidas, a variante mais produtiva foi *cangalha* (20 ocorrências, 29%), seguida de *canga* (16 ocorrências, 23%) e *forquilha* (3 ocorrências, 4%). Embora próximas, as frequências de *cangalha* e *canga* são significativas, considerando-se que o *corpus* válido foi de apenas 39 registros frente aos 68 esperados. Sob o rótulo *outras* (4%), agrupamos ocorrências únicas: *quadro, pescoceira* e *cabresto*.

A análise mostrou que o número de *não respostas* foi bastante expressivo (40%), o que reforça o peso da mudança rural-urbano na elicitação, já que informantes urbanos tendem a desconhecer instrumentos típicos das atividades rurículas.

5.4.6 Uma Descrição dos Dados Segundo a Variável Sexo no ALiB-PR

Para melhor compreender a importância da variável sexo na elicitação das denominações do referente em pauta, no ALiB-PR, elaboramos a Carta experimental, representada na Figura 40, com as variantes para *cangalha* distribuídas pelo território paranaense. Vale ressaltar que se trata de carta mista, pois vem com a descrição espacial das lexias, que são representadas por círculos preenchidos (coloridos) na legenda. Na carta, o número 1 (à esquerda) representa os informantes do sexo masculino; o número 2 (à direita), os do sexo feminino.

Figura 40 – Carta experimental com a distribuição das variantes de cangalha registradas no ALiB-PR segundo a variável sexo



Fonte: Elaborada com dados inéditos do Projeto ALiB-PR.

De acordo com o levantamento dos dados do ALiB-PR para *cangalha*, do número total de ocorrências (70), obtivemos 42 respostas válidas (60%) – 25 na fala masculina e 17 na fala feminina – e 28 de *não respostas* (40%) – 11 na fala dos homens e 17 na fala das mulheres. Do total de 25 respostas válidas para o sexo masculino, *cangalha* foi a mais produtiva, com 14 ocorrências (39%); *canga*, com nove respostas (25%); e *forquilha*, com duas ocorrências (6%). Da totalidade de 17 respostas válidas para o sexo feminino, *canga* prevaleceu, com sete ocorrências (21%); *cangalha*, com seis (18%); *forquilha* mencionada apenas uma vez (3%); e outras variantes (*pescoceira, quadro* e *cabresto*), com três ocorrências no somatório, registradas somente na fala feminina (9%).

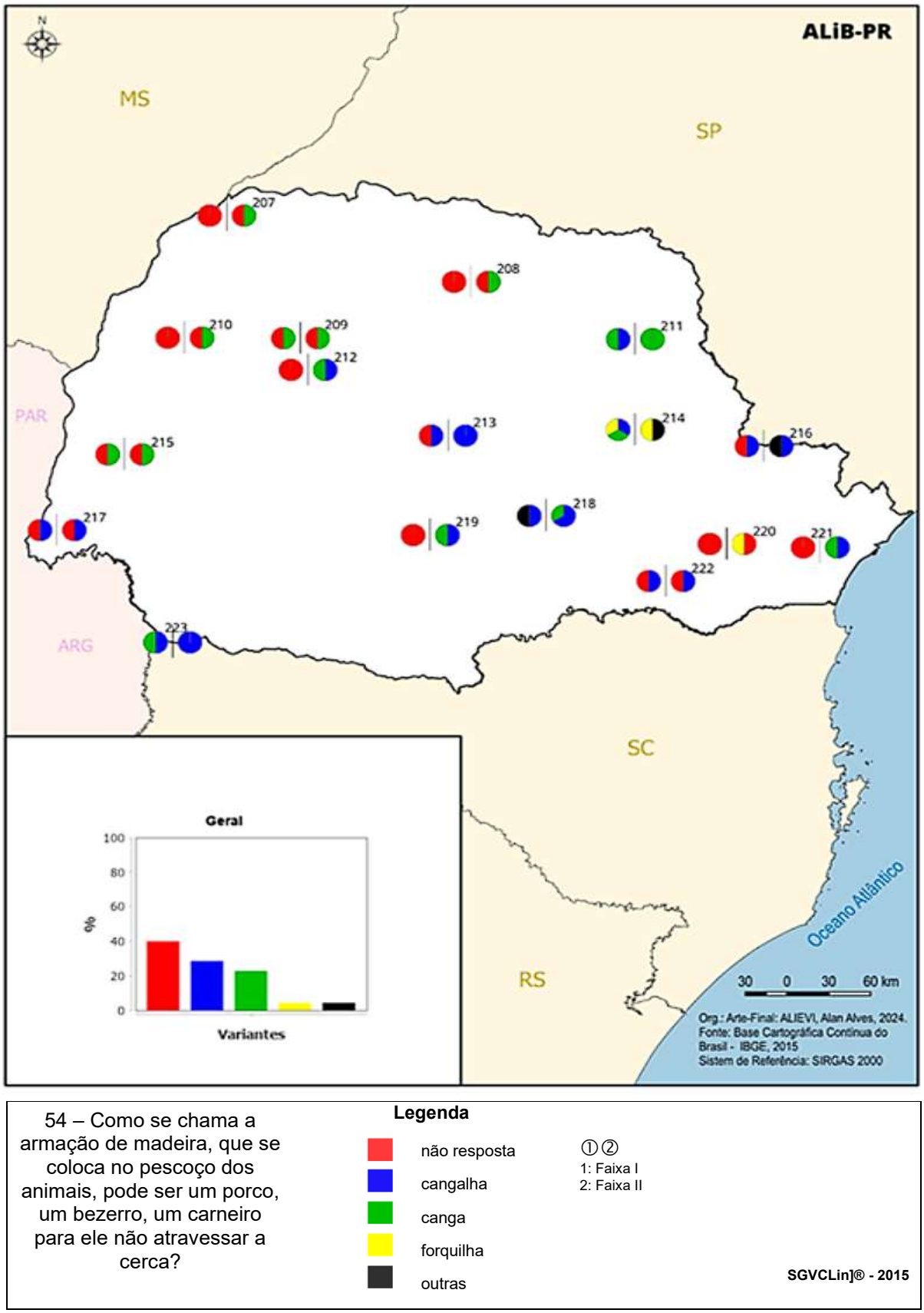
Em síntese, a ausência de respostas nos dados do ALiB-PR foi bastante significativa, principalmente na fala feminina. Dos 64 informantes, obtivemos 70 itens para análise, sendo 42 de respostas (60%) e 28 de *não respostas* (40%). Desse total de 28 ausências de resposta, 11 estão na fala masculina (39%) e 17 na feminina (61%), apontando para o maior distanciamento da mulher da vida no campo em relação ao homem. Vale lembrar que o desconhecimento do nome e do objeto pelo homem urbano também é bastante expressivo.

Além disso, verificamos que, diante do desconhecimento do nome e/ou do objeto, as mulheres são mais criativas, inovam estendendo o significado de um referente para outro (*pescoceira, quadro* e *cabresto*). Os dados ratificam a observação de Silva-Corvalán (1989, p. 69), ao afirmar que é bem sabido que, em igualdade de condições, a fala feminina é diferente da fala masculina.

5.4.7 Descrição dos Dados Segundo a Variável Faixa Etária no ALiB-PR

A variável faixa etária tem por finalidade descrever as diferenças linguísticas entre dois grupos com idades distintas: informantes mais jovens (Faixa I – 18 a 30 anos) e informantes mais velhos (Faixa II – 50 a 65 anos). Para melhor compreender as ocorrências para a *variável faixa etária* no ALiB-PR, elaboramos a Carta experimental, representada na Figura 41, com as variantes para *cangalha* distribuídas pelo território paranaense.

Figura 41 – Carta experimental com a distribuição das variantes para cangalha registradas pelo ALiB-PR segundo a variável faixa etária



Fonte: Banco de dados do ALiB-PR.

A Figura 41 mostra que as *não respostas* ocorrem majoritariamente na fala dos mais jovens, pertencentes à Faixa I. Dos 34 entrevistados, com idade entre 18 e 30 anos, foram extraídas 35 ocorrências. Desse total, 20 deles (57%) declararam não conhecer ou não saber o nome do artefato usado para o animal não atravessar a cerca. Dentre as respostas válidas na Faixa I, a mais produtiva foi *cangalha*, com oito ocorrências (23%); *canga*, com cinco respostas (14%); *forquilha* e outras (*cabresto*) com apenas um registro (3%) cada uma.

Em relação à Faixa II, com 35 ocorrências, o número de *não respostas* foi bem menor, com apenas oito ausências (23%). A variante mais frequente foi *cangalha*, com 12 respostas (34%); seguida de *canga*, com 11 registros (31%); *forquilha* e outras, com dois registros cada (6%).

Constatamos que o maior número de *não respostas* ocorreu entre os informantes da Faixa I, com 20 ausências (57%); já na Faixa II, apenas oito entrevistados não souberam o nome do triângulo que se coloca no pescoço de animais para não atravessar a cerca (23%). Verifica-se também o predomínio das variantes *cangalha* e *canga* em ambos os intervalos de idade, mas com certa equivalência numérica entre os entrevistados mais velhos (34% e 31%, respectivamente).

Os dados apontam para a influência da variável faixa etária em dois aspectos: i) os informantes da Faixa II podem ter guardado na memória as situações rurais vivenciadas ou transmitidas por ancestrais; ii) os falantes da Faixa I não possuem em sua memória passiva o referente ou o signo linguístico que o identifica.

5.4.8 Análise Comparativa dos Dados do ALPR e do ALiB-PR

Quanto à diacronia, é importante ressaltar que os dados do ALPR foram coletados entre 1987 e 1990 e os do ALiB-PR, majoritariamente, entre 2001 e 2004. Esse lapso temporal, embora curto, de menos de duas décadas, foi objeto de investigação para verificar se houve alteração, manutenção ou apagamento das variantes para *cangalha* em ambas as fontes.

Diante dos resultados obtidos, observamos que, no ALPR, o índice de não respostas foi pouco expressivo (4%). Os dados demonstram que, do ponto de vista diatópico, os informantes do ALPR, em sua maioria, por viverem da terra (agricultores, lavradores, pecuaristas), souberam nomear o referido objeto. Todavia, os do ALiB, por serem de origem urbana e não terem vínculo direto com a terra (chapeiro, dona de casa, pedreiro, taxista, costureira etc.), demonstraram desconhecer tanto o referente

quanto suas designações (40%). Desse modo, é viável creditar à variável diatópica (rural x urbano) essa variação que parece tender para a mudança, conforme demonstra a Tabela 11.

Tabela 11 – Percentual das variantes de cangalha no ALPR e ALiB-PR

Fonte/Variante	cangalha	canga	Forquilha	Cuaiera	outras	NR
ALPR (rural)	56%	31%	4%	1%	4%	4%
ALiB-PR (urbano)	29%	23%	4%	0%	4%	40%

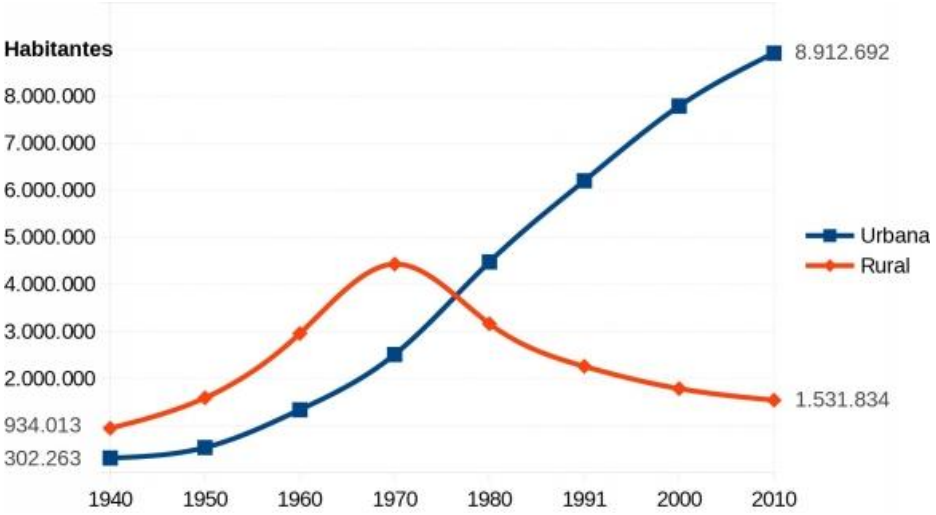
Fonte: Dados da pesquisa.

A Tabela 11 mostra que, no ALPR, destaca-se como mais produtiva a variante *cangalha*, representando 56% das ocorrências, seguida de *canga*, com 31%. As demais obtiveram resultados pouco expressivos: *forquilha*, com 4%; *cuaiera*, com 1%; e outras, com 4%. No ALiB-PR, *cangalha* obteve 29% das respostas, sendo a mais elicitada, seguida por *canga*, com 23%. *Forquilha* e outras variantes tiveram apenas 4% cada uma, e *cuaiera* não foi mencionada por nenhum informante.

De acordo com a Tabela 11, verificamos que a variável tempo é um fator preponderante para as mudanças registradas, uma vez que se observa o decréscimo de registros de *cangalha* (de 56% para 29%), a diminuição de ocorrências de *canga* (de 31% para 23%), a estabilização de *forquilha* (em 4%) e a manutenção do grupo “outras” (4%). No entanto, o que parece exercer maior influência é a relação rural x urbano entre os dois estudos paranaenses.

Para Monteiro (2016, p. 34), até a década de 1970, o Paraná era um estado predominantemente rural; a partir desse marco temporal, sua população entrou em declínio, dando ensejo a um ciclo de urbanização ascendente. O Gráfico 4 ajuda a compreender melhor essa dinâmica populacional.

Gráfico 4 – Recorte temporal da população rural x urbana no Paraná (1940 a 2010)



Fonte: IBGE – Censos 1940-2010 (Monteiro, 2016).

O Gráfico 4 permite acompanhar a queda da população rural e o aumento gradual da população urbana no período que compreende a coleta dos dados do ALPR (atlas de natureza rural, 1987-1990) e do ALiB-PR (atlas de natureza urbana, 2001-2004). Os dados demonstram que o índice de ruralidade decresce significativamente na década de 1980 e de forma menos acentuada no início dos anos 2000.

O decréscimo da população rural e o crescimento da população urbana entre as décadas de 1980 e 2010, motivados pelo fluxo migratório do campo para a cidade, certamente trouxeram consequências para o léxico, haja vista a mudança expressiva na cultura, nos hábitos e nas ocupações dos indivíduos. Os reflexos desse êxodo rural foram documentados pela geolinguística paranaense e permitem acompanhar a variação linguística, registrando a redução do número e da frequência das variantes agropastoris *cangalha* e *canga* e o aumento do número de *não respostas* nos dados do ALiB-PR.

6 DICIONARIZAÇÃO DAS VARIANTES ESTUDADAS

Este capítulo trata do levantamento lexicográfico dos verbetes em dicionários com estruturas diferentes em relação à cronologia e ao suporte (impresso x *on-line*): Silva (1823), Houaiss (2001), Ferreira (2010) e Michaelis (*on-line* [2025]).

6.1 DICIONARIZAÇÃO DAS PRINCIPAIS VARIANTES DE SABUGO

Quanto à dicionarização das variantes, verificamos que *capuco*, presente no APFB e no ALiB-BA, não está registrado em Silva (1823), em Ferreira (2010) nem em Houaiss (2001). Contudo, consta no Michaelis (*Capuco*, [2025]) com a seguinte acepção: "s.m. sabugo de milho; batuera".

O item lexical *sabugo*, em Silva (1823), dentre outras acepções, aparece como: "§ Sabugo do milho – a parte onde o gráo está embebido nos alvados, ou alveolos". Em Ferreira (2010), temos: "[Do lat. *sabucu*, 'sabugueiro'.] S.m. 1. Bot. Eixo da inflorescência e da infrutescência do milho, que se hipertrofia, onde se inserem as flores e os frutos (popularmente denominadas sementes) [...] 6. Espiga de milho sem grãos". No dicionário Michaelis (*Sabugo*, [2025]), na quarta acepção, encontra-se: "Espiga de milho após a retirada dos grãos; carolo". Em Houaiss (2001): "4 espiga de milho a que se retiram os grãos; carolo [...]".

Já os itens lexicais *sabuco* e *papuco* não constam dos dicionários consultados.

6.2 DICIONARIZAÇÃO DAS PRINCIPAIS VARIANTES DE *SOCA*

O item lexical *soca* foi catalogado por Silva (1823), em linguagem arcaica, mas oficial da época, com a seguinte acepção: "No Brasil planta-se a cana de assucar, e a primeira producçáo se diz planta; cortada ella, dos pés que ficão em terra brota outra novidade que se diz soca; e desta cortada torna a brotar a ressoca. Insul. 10. 82. § Não ter nem soca, i.e., nem branca, nem hum seitil".

Ferreira (2010) também registra *soca*[1] "[Do tupi = 'renovo', 'pimpolho'.] *S.f.* 1. *Bot*. Nome vulgar de rizoma ou caule subterrâneo. 2. *Bras*. A segunda produção da cana depois de cortada a primeira. [A primeira é planta; a segunda, soca; a terceira, ressoca; a quarta, contrassoca.] 3. *Bras*. Touceira de capim. 4. Bras. N.E. A segunda colheita do fumo. 5. Bras. ES A segunda colheita do arroz. 6. Bras. RS Entre os ervateiros, a árvore de mate quando podada". A unidade lexical *soca* foi registrada

também por Michaelis (Soca, [2025]), mas com outra acepção e, por isso, não a transcrevemos.

De modo similar, Houaiss (2001) também registra o termo, com três chamadas: "**1** *soca* s.f.(1935 cf insul) **1** infrm. rizoma ou caule subterrâneo **2** (1711) brotação após o primeiro corte da cana-de-açúcar cf. planta, ressoca e contrassoca **3** B moita de capim **4** B N.E. a segunda colheita do fumo **5** ES a segunda colheita de arroz **6** MG tabaco de inferior qualidade **7** RS a árvore de mate, quando podada [...]; **2** *soca* s.f."

O item lexical ***soqueira*** não consta do dicionário de Silva (1823), porém, consta dos demais. Para Ferreira (2010) *soqueira* é descrita da seguinte forma: "[De *soca*[1] + *eira*.] S.f. 1. O raizame das canas, e doutras plantas, após o corte". Em Houaiss (2001), *soqueira* está registrada com a mesma acepção: "1 emaranhado de raízes que ficam na terra após o corte das canas e de outras plantas". Em Michaelis (Soqueira [2025]) significa "Raizame que permanece na terra após o corte de canas e de outras plantas".

O item lexical *toco* aparece em Silva (1823) com a seguinte semântica: "tronco de arvore, cepa. Alarte". Ferreira (2010) assim o descreve em sua primeira acepção: "1. Parte do tronco vegetal que permanece ligada à terra depois de cortada a árvore". Para Michaelis (Toco, [2025]), significa "Parte do tronco que fica preso ao solo, após uma árvore ser cortada ou quebrada [...]". Em Houaiss (2001), constam várias acepções, todavia registramos aquela com o mesmo traço semântico da QSL 047: "1 no corte de qualquer planta, parte que permanece presa ao solo".

O verbete *raiz* consta em todos os dicionários selecionados para esta dissertação. Em Silva (1823), consta esta seguinte acepção "a parte da planta, que fica embaixo da terra, e que absorve para a nutrir os sucos appropríados". Em Ferreira (2010), encontra-se definida como "Porção inferior do eixo de uma planta, que cresce, ger., para dentro do solo, e cuja função fundamental é fixar o organismo vegetal e retirar do substrato os nutrientes e a água necessários à vida da planta". Em Michaelis (Raiz, [2025]), está registrada como termo da botânica: "Órgão de uma planta vascular, desenvolvido a partir da radícula, geralmente fixo ao solo, crescendo com ramificações secundárias, com as funções de fixar a planta a um substrato, além de absorver e conduzir água e nutrientes; estirpe". O dicionário Houaiss (2001, p. 2379) registra esse item lexical como "eixo de uma planta vascular, que se desenvolve a partir da radícula [...]".

O item lexical *resteva* aparece nos dicionários pesquisados. Em Silva (1823), consta com o significado de s.f. rastolho, que aparece no mesmo dicionário (Silva, 1823) com a acepção de s.m. a cana do trigo segado, que fica com a raiz na terra. O lexicógrafo Ferreira (2010) assim o descreve (ê) [Do lat. vulg. * restipa.] S.f. V. restolhal. No dicionário Michaelis (Resteva, [2025]), há o verbete com remissão à restolhal, que está registrado com a denotação de Extensão de terra em que há restolho; pajonal, resteva. Para Houaiss (2001) *resteva* /ê/ designa “palha, colmo (de telhado) [...] resteba [...]”.

6.3 DICIONARIZAÇÃO DAS VARIANTES DE VAGEM

O verbete *vagem* foi registrado no dicionário Silva (1823) com a seguinte acepção “a bainha em que estáo os legumes, como feijões, hervilhas, Scc”. Para Ferreira (2010), “[Do lat. vagina, ‘bainha’, pela f. dialetal baginha, tomada como dim. De *bage, f. ainda hoje pop.] S.f.**1**. Bot. Legume (1). **2**. Feijão verde [Var., nessas aceps.: bagem.]. Os lexicógrafos Houaiss (2001) e Michaelis (Vagem, [2025]) reiteram os conceitos presentes em Ferreira e acrescentam que esse verbete tem como variante ‘bagem’.

O item lexical *bagem* não consta no dicionário de Silva (1823), porém foi registrado em Ferreira (2010), Michaelis (Bagem, [2025]) e Houaiss (2001) com remissão à sua variante *vagem*.

Já a unidade léxica *bainha* foi documentada em Silva (1823) “Baje de legume [...]”. Para Ferreira (2010), o verbete *bainha* designa “Qualquer formação que circunda órgão ou parte deste [...]. Em Michaelis (Bainha, [2025]) e Houaiss (2001), há o registro de ***bainha,*** mas sem a respectiva acepção do referente da QSL 049.

6.4 DICIONARIZAÇÃO DAS VARIANTES DE *CANGA*

Os lexicógrafos Silva (1823), Michaelis (Canga [2025]) e Houaiss (2001) registram o item lexical *canga* com várias acepções. Todavia, nenhuma delas retrata o referente da QSL 054 – “armação de madeira que se coloca no pescoço do animal [...] para não atravessar a cerca”. Em Ferreira (2010), *canga* é registrada como “[Do celta *cambica, ‘madeira curva’, poss.] S.f. 1. Peça de madeira que prende os bois pelo pescoço e os liga ao carro, ou ao arado; jugo: um boi magro ... esticava o pescoço esfolado pela canga e mugia (Coelho Neto, Sertão, p. 79)”.

A unidade lexical *cangalha* foi catalogada em todos os dicionários pesquisados. Entretanto, em Silva (1823), as acepções registradas não preservam o sentido da QSL 054.

Para o dicionarista Ferreira (2010), *cangalha* é designativo de "[De canga[1] (1) + alha.] S.f. 1. Cangalhas (1 e 2). 2. Bras. S. Peça de três paus, unidos em triângulo, que se enfia no pescoço dos porcos para não destruírem hortas cultivadas".

Michaelis (Cangalha [2025]) registrou esse vocábulo com várias acepções, mas apenas a primeira representa o triângulo de madeira que se coloca no pescoço do animal, conforme se depreende da transcrição do verbete "Triângulo de madeira que se coloca no pescoço dos suínos para não devastarem hortas cultivadas".

Houaiss também contemplou *cangalha* com essa mesma acepção "s.f. [...] **2** triângulo de madeira que se coloca no pescoço dos suínos para impedir que fucem canteiros".

A unidade léxica *forquilha* consta do dicionário de Silva (1823), mas com outras acepções. Ferreira (2010) documentou *forquilha* "[Do esp. Horquilla.] S.f. **1**. Pequeno forcado de três pontas". Em Michaelis (Forquilha, [2025]) e Houaiss (2001), há várias acepções para *forquilha*, mas nenhuma delas retrata o objeto de madeira em formato de triângulo.

Elaboramos a Tabela 12 para verificar em quais dicionários os verbetes, objeto desta dissertação, foram registrados com a mesma acepção das questões semântico-lexicais.

Tabela 12 – Lista com os verbetes dicionarizados

Variantes	Silva (1823)	Ferreira (2010)	Houaiss (2001)	Michaelis (*on-line*)
capuco	não	não	não	sim
sabugo	sim	sim	sim	sim
soca	sim	sim	sim	não
soqueira	não	sim	sim	sim
toco	sim	sim	sim	sim
raiz	sim	sim	sim	sim
resteva	sim	sim	sim	sim
vagem	sim	sim	sim	sim
bagem	não	sim	sim	sim
bainha	sim	sim	não	não
canga	não	sim	não	não
cangalha	não	sim	sim	sim
forquilha	não	sim	não	não

Fonte: Dados da pesquisa.

De acordo com a Tabela 12, percebemos que alguns verbetes (*sabugo*, *toco*, *raiz*, *resteva* e *vagem*) foram registrados em todos os dicionários. Por outro lado, outros itens (*capuco*, *soqueira*, *bagem*, *bainha* e *canga*) constam apenas em algumas obras lexicográficas. Esses resultados servirão de parâmetro para averiguarmos os desdobramentos do léxico catalogado na variação lexical de certas variantes, tema que será retomado nas considerações finais.

7 CONSIDERAÇÕES FINAIS

A pesquisa, pautada nos pressupostos teóricos da Dialetologia, resultou neste trabalho que apresenta um estudo das variantes agropastoris do Questionário Semântico-Lexical do Projeto Atlas Linguístico do Brasil – QSL 046 (*sabugo*), QSL 047 (*soca*), QSL 049 (*vagem*) e QSL 054 (*cangalha*) – analisadas em quatro fontes distintas referentes a três sincronias: (I) *Atlas Prévio dos Falares Baianos* (Rossi, 1963); *Atlas Linguístico do Paraná* (Aguilera, 1994; Altino, 2007) e dados inéditos do Projeto Atlas Linguístico do Brasil em ambos os estados no início dos anos 2000. Objetivamos com esta dissertação: (i) descrever os dados coletados pelo ALiB para fins de cartografação; (ii) comparar os dados baianos e paranaenses dos respectivos atlas estaduais com os obtidos pelo Projeto ALiB nesses estados; (iii) verificar a correlação *rural x urbano* na elicitação das respostas; (iv) avaliar a influência das variáveis sexo, faixa etária e escolaridade.

Em relação ao objetivo (i) - descrever os dados coletados pelo ALiB para fins de cartografação – pudemos elaborar cartas linguísticas com a apresentação das ocorrências em cada rede de pontos investigada, o que nos possibilitou compreender a distribuição geolinguística das variantes em cada localidade.

Quanto ao objetivo (ii) - comparar os dados baianos e paranaenses dos respectivos atlas estaduais com os obtidos pelo Projeto ALiB nesses estados – a análise contrastiva entre os dois estados corroborou nossas expectativas, reafirmando a produtividade da variação diatópica, na medida em que *capuco*, *sapuco* e *pipuco* foram registradas apenas no APFB e no ALiB-BA. Já a correlação *rural x urbano* ratificou o alto índice de *não respostas* na fala dos informantes das cidades, o que foi comprovado nos resultados do ALiB-BA, com 40% de abstenções, mesmo sem esse resultado se repetir no ALiB-PR.

Quanto ao objetivo (iii) - verificar a correlação *rural x urbano* na elicitação das respostas – verificamos que, na zona rural, os informantes detêm um acervo lexical mais técnico, preciso e específico para as variantes agropastoris, uma vez que esse léxico está intrinsecamente relacionado à lida do campo, integrando o seu cotidiano, a sua cultura e os seus costumes. Além disso, notamos que o vocabulário rural é mais conservador e tradicional, diferentemente do urbano, mais dinâmico, inovador e propenso à variação.

No que diz respeito ao objetivo (iv) - avaliar a influência das variáveis sexo e faixa etária – observamos que ora uma variável se revela mais produtiva para uma variante, ora menos produtiva para outra. Nesse sentido, a variável sexo apresentou algumas oscilações pontuais na fala masculina e na feminina. Já a variável faixa etária foi responsável pela manutenção de termos mais técnicos, específicos e tradicionais, conforme revelado na elicitação de *capuco*, *soca/soqueira*, *cepa* e *resteva*.

A constatação inicial da hipótese deste trabalho, qual seja, a de que as variantes lexicais relacionadas ao vocabulário agropastoril tendem a apresentar um processo de apagamento mais acentuado em contextos urbanos, retratou resultados diferentes, ora confirmando, ora não confirmando tal hipótese, conforme verificaremos agora na descrição de cada uma das variantes separadamente. Para *sabugo* e *vagem*, houve baixa incidência de *não respostas* em todos os *corpora*. Porém, para *cangalha*, as *não respostas* foram produtivas, chegando a 40% no ALiB-PR e a apenas 4% no ALPR. Tal resultado foi similar para *soca* no ALiB-BA, atingindo 39% dos informantes, diversamente do Paraná, onde houve baixa incidência, dada a ruralidade desse estado – um dos maiores produtores de grãos do país –, o que se reflete no léxico de seus informantes, mesmo naqueles dos centros urbanos.

Dada a diversidade de variantes estudadas em três *corpora*, passamos a analisar cada uma delas separadamente para compreendermos melhor os resultados da pesquisa e as peculiaridades de cada uma sob o viés das variáveis contempladas em cada *corpus*, de acordo com a Dialetologia Tradicional e a Dialetologia Pluridimensional. Iniciamos por *sabugo*, depois *soca*, *vagem* e, por último, *cangalha.*

Para a QSL 046 (*sabugo* e suas variantes), as fontes estaduais (APFB e ALPR, ambos rurais) revelaram resultados diferentes na diatopia, apontando para as divergências sócio-históricas na ocupação de ambos os espaços geográficos. No APFB, *capuco* foi a variante mais produtiva, mas outras lexias também foram registradas (*sabuco, papuco, sabugo, cascabulho, batuera, sapuco* etc.), o que revela a heterogeneidade do falar rural baiano. No ALPR, a variante *sabugo* (inovadora) foi a mais expressiva, seguida de sua variante fonética *sabuco* (conservadora), com baixa incidência, o que demonstra certa homogeneidade do falar rural paranaense. Creditamos essa distinção no léxico de ambos os estados a fatores socioculturais e, principalmente, à localização geolinguística (Bahia no Nordeste e Paraná no Sul). O processo de ocupação e povoamento ajuda-nos a entender, já que se trata de “Brasis” diferentes (Ribeiro, 2015), sendo a Bahia representada pela cultura sertaneja, forjada

às margens do Rio São Francisco, onde surgiram as primeiras fazendas de gado, administradas pela figura do vaqueiro, dando ensejo ao surgimento das primeiras comunidades. Para o autor, o Paraná se insere no denominado Brasil sulino, com forte presença do imigrante europeu, que viu nessa região, de clima parecido com o seu, a oportunidade de recomeçar a vida depois de a Europa ser devastada pela Segunda Guerra (1939 a 1945). Além disso, outros grupos migratórios exerceram papel fundamental na constituição da cultura e do falar paranaense: mineiros e paulistas no Norte do Paraná e gaúchos no Oeste.

Já em relação à variação diassexual, foco do objetivo (iv), os resultados em ambos os estados foram próximos entre homens e mulheres, divergindo apenas na escolha da variante lexical (*capuco* na Bahia e *sabugo* no Paraná), fato que demonstra o baixo peso dessa variável, sendo o sexo fator prescindível para o presente estudo.

Sob o aspecto diageracional, os dados revelaram que, no ALiB-BA, a variante majoritária para a Faixa I (18 a 30 anos) foi *sabuco* e, para a Faixa II (50 a 65 anos), *capuco*. Dessa forma, pudemos concluir que a idade se revelou decisiva na escolha das variantes lexicais nessa fonte. No ALPR, em ambas as faixas etárias, *sabugo* foi a variante mais produtiva, com valores equivalentes; portanto, essa variável não foi relevante na escolha das variantes. Ressaltamos ainda que o número de *não respostas* foi baixo em ambos os estados, contrariando a hipótese inicial de que, nos centros urbanos, as abstenções de respostas são elevadas.

Todavia, creditamos essa baixa incidência das *variantes de sabugo*, metonimicamente associadas ao milho, a fatores como o êxito do agronegócio brasileiro, com sua rotação de culturas e mecanização da lavoura; a música sertaneja, em expansão no país; a culinária e, em especial, as festas juninas em todas as regiões do Brasil; além dos meios de comunicação, a exemplo do programa *Globo Rural* (dominical, na TV Globo) e do *Sítio do Picapau Amarelo*, na personagem Visconde de Sabugosa. Acreditamos, pois, que esses fatores, juntos, proporcionaram maior disseminação das variantes agropastoris na maioria das cidades brasileiras, consolidando a apropriação dos falantes em relação a *sabugo* e suas variantes.

Na diacronia, pôde-se notar que, na Bahia, houve alteração na escolha das variantes obtidas em sincronias diferentes (APFB, Rossi, 1963; ALiB-BA, 2000/2004). Não só o lapso temporal de aproximadamente 40 anos, mas também a origem distinta dos informantes (*rural x urbana*) e as mudanças sócio-históricas nesse período foram responsáveis pela mudança na variante mais produtiva, revelando a importância do

peso dessa variável. Nos dados do Paraná, todavia, os resultados foram próximos em ambas as fontes – ALPR (Aguilera, 1994; cartas em Altino, 2007) e ALiB-PR (2000/2004): predomínio de *sabugo*, com alta incidência, e sua variante conservadora *sabuco*, com baixa incidência. Logo, essa variável não foi responsável pela escolha das variantes no estado sulista.

Em relação à QSL 047 (*soca* e suas variantes), obtivemos resultados diferentes de acordo com a localização (*rural x urbana*). No ALPR, houve maior concentração de termos mais pertinentes ao contexto rural, como *soca/soqueira, resteva* e *cepa*, demonstrando maior convívio e afinidade com as práticas agrícolas, o que encontra respaldo no fato de o Paraná ser um dos maiores produtores de grãos do Brasil, só ficando atrás do Mato Grosso, de acordo com estudos realizados pela Embrapa. Já em relação aos dados do ALiB em ambos os estados, as variantes elicitadas na fala dos informantes revelam certo grau de generalidade e afastamento do convívio rural, dada a origem urbana de seus participantes. Notamos um alto índice de variabilidade nos dados do ALiB na Bahia e no Paraná. O número de *não respostas* foi baixo nesse estado sulista, representando apenas 3% dos dados no ALPR e 8% no ALiB-PR. Porém, no ALiB-BA, a maioria dos informantes não soube a denominação da planta que produz de novo após ser cortada.

A pesquisa confirmou que a variação léxica é fortemente influenciada por fatores diatópicos e diastráticos. No primeiro caso, termos como *soca, soqueira, resteva* e *cepa* atestam o vínculo entre o léxico rural e práticas agrícolas tradicionais, preservadas pela oralidade e pela transmissão geracional.

As variáveis sociais também exerceram papel importante. O sexo influenciou escolhas lexicais, como no caso do ALiB-BA, com maior frequência de *toco* na fala masculina, enquanto, na fala feminina, prevaleceram as *não respostas*; e do ALiB-PR, em que observamos maior diversidade no léxico, com *raiz* mais frequente entre mulheres e *soca* entre homens. A idade, por sua vez, mostrou que os mais velhos tendem a conservar variantes mais específicas, reforçando a ligação com atividades agrícolas.

Em relação à QSL 049 (*vagem* e suas variantes), as variáveis sexo, faixa etária e a diageracional não trouxeram contribuições para a pesquisa, independentemente da fonte analisada. Contudo, a variável diatópica influenciou decisivamente as respostas dos informantes do ALPR, prevalecendo *bainha* no Paraná Tradicional, ligado a práticas agrícolas convencionais e a menor influência externa, ao passo que

vagem/bagem se destacam no Paraná Moderno, associado a transformações sociais e econômicas decorrentes da modernização e interiorização do estado.

No tocante à QSL 054 (*cangalha* e suas variantes), o estudo comparativo entre o ALPR e o ALiB-PR demonstrou que a variação lexical paranaense se constitui pela interação entre as variáveis diatópica e diastrática. Nas áreas rurais, o léxico agropastoril mantém-se preservado, com destaque para *cangalha* no Paraná Moderno e *canga* distribuída por todo o estado, porém mais concentrada no Paraná Tradicional; já nos centros urbanos, observamos um decréscimo das variantes agropastoris e um apagamento progressivo, com alto índice de *não respostas*.

A diacronia ratificou esse cenário, mas demonstrou-se associada à diatopia, em um contexto de urbanização acelerada e perda de referências ligadas ao campo. A variável sexo demonstrou equilíbrio no ALPR, mas diferenças sutis no ALiB-PR, com prevalência de *cangalha* na fala masculina e *canga* na feminina. Já a faixa etária confirmou que os informantes mais velhos preservam, em sua memória, um acervo lexical mais específico e conservador, diversamente dos informantes mais jovens, entre os quais o apagamento das lexias agropastoris foi elevado.

A imersão lexicográfica, realizada nos dicionários selecionados, revelou que o verbete *capuco*, registrado apenas em Michaelis (*Capuco*, [2025]) e refletido nas fontes da Bahia (APFB e ALiB-BA), pode ser interpretado como um regionalismo restrito a esse estado. De modo similar, *resteva*, apesar de constar de todos os dicionários, aparece apenas no Paraná (ALPR e ALiB-PR), denotando também seu caráter regional. Já os verbetes *cangalha* e, principalmente, *canga* apresentaram outros semas. Para *canga*, é muito comum a acepção de “peça de madeira que une uma junta de bois para o arado; jugo”. Já em relação a *cangalha*, também foi frequente a acepção de “artefato, geralmente de madeira, que se coloca no lombo das cavalgaduras para pendurar carga”. Esses vocábulos parecem ganhar novos significados, mas, mesmo assim, ainda preservam o seu sentido agropastoril. Todavia, esses léxicos não foram registrados no corpo da dissertação por representarem outros referentes não contemplados nesta pesquisa.

Esta dissertação permitiu-nos registrar que a língua vai além de transmitir significados, pois carrega a memória coletiva de uma comunidade, disseminando vivências, costumes e tradições, exprimindo sua identidade linguística, social e cultural. Dessa forma, sustentamos que o idioma é inseparável da sociedade a que serve, já que reflete transformações e, simultaneamente, preserva expressões que

poderiam desaparecer gradativamente, dando voz aos sujeitos que as usam, muitas vezes silenciados e estigmatizados, em um país onde ainda impera o preconceito linguístico. Portanto, o legado desta dissertação não se limita ao campo lexical, mas alcança outras manifestações sociais e culturais aqui representadas.

Acreditamos que novas pesquisas acerca das variantes aqui estudadas possam agregar conhecimento científico para os estudos linguísticos e confirmar os resultados apurados, especialmente no âmbito da Dialetologia, ampliando o *corpus* deste trabalho para incluir novas variantes pertencentes ao mesmo campo semântico agropastoril, sobremaneira na rede de pontos coincidente entre os pontos do ALiB e os dos atlas estaduais – APFB e ALPR.

Portanto, esperamos, com esta dissertação, não apenas auxiliar estudantes, professores e dialetólogos no desenvolvimento de pesquisas similares por meio da Dialetologia Tradicional e da Dialetologia Pluridimensional, mas também contribuir para os estudos do Projeto ALiB e do ALPR com os resultados aqui apresentados.

REFERÊNCIAS

AGUILERA, Vanderci de Andrade. **Atlas Lingüístico do Paraná – ALPR.** Curitiba: Imprensa Oficial do Estado, 1994.

AGUILERA, Vanderci de Andrade. O léxico paranaense: uma viagem pelas veredas rurais e pelos caminhos urbanos. *In:* FAGUNDES, Edson Domingos; LOREGIAN-PENKAL, Loregian-Penkal; MENON, Odete Pereira da Silva (org.). **O falar paranaense.** Curitiba: UTFPR, 2015. p. 19-34.

ALMEIDA-BARONAS, Joyce Elaine de. Dialeto caipira e urbanização: uma química que não se mistura (?). *In*: ALTINO, Fabiane Cristina (org.). **Múltiplos olhares sobre a diversidade linguística**: nos caminhos de Vanderci de Andrade Aguilera. Londrina: Midiograf, 2012. p. 11-23.

ALTINO, Fabiane Cristina. Atlas Linguístico do Paraná – v. II: comentários sobre a Dialetometria. **Estudos Linguísticos,** São Paulo, v. 41, n. 2, p. 818-832, 2012.

ALTINO, Fabiane Cristina. **Atlas linguístico do Paraná**. 2007. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2007. Disponível em: https://repositorio.uel.br/bitstreams/2f11daa1-d540-4931-981f-2cbac085dee7/download. Acesso em: 07 fev. 2025.

AMARAL, Amadeu. **O dialeto caipira:** gramática, vocabulário. São Paulo: O Livro, 1920.

AMARAL, Amadeu; DUARTE, Paulo. **O dialeto caipira**: gramática – vocabulário. São Paulo: Casa Editora O Livro, 1955.

ARAGÃO, Maria do Socorro Silva de. Os estudos dialetais e geolingüísticos no Brasil. **Revue roumaine de linguistique**, p. 126-140, 2008.

ARAGÃO, Maria do Socorro Silva de; MENEZES, Cleuza Palmeira Bezerra de. **Atlas Lingüístico da Paraíba.** Brasília: Universidade Federal da Paraíba; CNPq, 1984.

BAGEM. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/bagem/. Acesso em: 01 mar. 2025.

CANGA. *In*: HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001.

CANGA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/palavra/BlDO/canga%3CEi%3E1%3C/Ei%3E/. Acesso em: 3 mar. 2025.

CANGALHA. *In*: HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001.

CANGALHA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/cangalha/. Acesso em: 04 mar. 2025.

CAPUCO. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/capuco/. Acesso em: 4 mar. 2025.

CARDOSO, Jayme Antonio; WESTPHALEN, Cecília Maria. **Atlas histórico do Paraná**. 2. ed. Curitiba: Livraria do Chain, 1986.

CARDOSO, Suzana Alice Marcelino *et al.* **Atlas Lingüístico do Brasil: cartas linguísticas**, v. 2. Londrina: EDUEL, 2014.

CARDOSO, Suzana Alice Marcelino. A dialectologia no Brasil: perspectivas. **DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada**, v. 233-255, 1999.

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

CARUSO, Pedro. **Atlas lingüístico do estado de São Paulo: questionário.** Assis: UNESP; Prefeitura Municipal de Assis, 1983.

CHAMBERS, Jack K.; TRUDGILL, Peter. **La Dialetologia.** Madrid: Visor Libros, 1994.

COMITÊ NACIONAL DO ALIB. **Projeto do Atlas Lingüístico do Brasil**: Questionários 2001. Londrina: EDUEL, 2001.

COSERIU, E. La geografia linguística. *In*: COSERIU, E. **El hombre y su lenguaje**. 2. Ed. Madrid: Gredos, 1991, p. 103-158.

DIÉGUES JÚNIOR, Manuel. **Regiões culturais do Brasil**. Centro Brasileiro de Pesquisas Educacionais, INEP, Ministério da Educação e Cultura, 1960.

FERREIRA, Aurélio Buarque de Holanda. **Dicionário Aurélio da língua portuguesa**. 5. ed. Curitiba: Positivo, 2010.

FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário Aurélio da língua portuguesa**. 3. ed. Curitiba: Positivo, 2004.

FERREIRA, Carlota *et al.* **Atlas lingüístico de Sergipe**. Salvador: Universidade Federal da Bahia, 1987.

FORQUILHA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/forquilha/. Acesso em: 4 mar. 2025.

HOUAISS, Antônio. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001. Disponível em: https://michaelis.uol.com.br/palavra/mNa2/capuco/. Acesso em: 28 de fev. 2025.

IBGE. **Cidades**. Rio de Janeiro: IBGE, 2022. Disponível em: https://www.ibge.gov.br/cidades-e-estados/pr.html. Acesso em: 22 de jun. 2025.

ISQUERDO, Aparecida Negri. Revisitando os conceitos de rural e urbano no português do Brasil: contribuição do Projeto ALiB. *In*: LLIESCU, Maria; DANLER, Paul. **Actes du XXVe Congrés international de linguistique et de philologie romanes**. Berlin; New York: Walter de Gruyter, 2010. v. 4, p. 137-145.

KOCH, Walter; KLASSMANN, Mário Silfredo; ALTENHOFEN, Cléo. **Atlas Lingüístico-etnográfico da Região Sul do Brasil.** v. 1-2. Porto Alegre; Florianópolis; Curitiba: UFRGS; UFSC; UFPR, 2002.

LABOV, William. **Padrões sociolinguísticos**. São Paulo: Parábola, [1972] 2008.

MIRASSOL. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/palavra/9odme/mirassolense%3CEi%3E1%3C/Ei%3E/. Acesso em: 28 fev. 2025.

MONTEIRO, Ricardo Rodrigues. A cartografia do fenômeno urbano e econômico no Paraná: uma leitura com auxílio da semiótica. **Revista Confins**, n. 27, p. 1, 2016. Disponível em: https://doi.org/10.4000/confins.10966. Acesso em: 21 fev. 2025.

MORENO FERNÁNDEZ, Francisco. **Principios de sociolingüística y sociología del lenguaje**. Barcelona: Ariel, 1998.

MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino. Sobre a dialectologia no Brasil: para uma nova divisão dos estudos dialetais brasileiros. *In*: MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino (org.). **Documentos 2**: projeto do atlas lingüístico do Brasil. Salvador: Quarteto, 2006. p. 15-26.

NASCENTES, Antenor. **Bases para a elaboração de um atlas lingüístico do Brasil I**. Rio de Janeiro: Casa de Rui Barbosa, 1958.

NASCENTES, Antenor. **Bases para a elaboração de um atlas lingüístico do Brasil II**. Rio de Janeiro: Casa de Rui Barbosa, 1961.

NASCENTES, Antenor. **O linguajar carioca.** 2. ed. Rio de Janeiro: Organização Simões, 1953.

OLIVEIRA, Dennison de. Transformações recentes da sociedade paranaense: indústrias e cidades. In: __. **Urbanização e Industrialização no Paraná**. Curitiba: SEED, 2001. (Coleção História do Paraná. Textos introdutórios) p. 27-37.

OLIVEIRA, Genivaldo da Conceição. **O léxico nosso de cada dia na Bahia e no Paraná**: acidentes geográficos, fenômenos atmosféricos, astros e tempo. Salvador: UFBA, 2014.

PAIVA, Maria da Conceição de. A variável gênero/sexo. *In:* MOLLICA, Maria Cecilia; BRAGA, Maria Luiza. **Introdução à sociolinguística:** o tratamento da variação. São Paulo: Contexto, 2004. p. 33-42.

PAPUCO. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/papuco/. Acesso em: 4 mar. 2025.

PRIORI, Angelo *et al.* **História do Paraná: séculos XIX e XX**. Maringá, Eduem, 2012.

RAIZ. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/raiz/. Acesso em: 23 fev. 2025.

RAZKY, Abdelhak. Atlas lingüístico sonoro do Pará: uma nova perspectiva para a organização de corpus geolinguístico. *In*: AGUILERA, Vanderci de Andrade (org.). **A geolingüística no Brasil**: trilhas seguidas, caminhos a percorrer. Londrina: Eduel, 2005. p. 209-228.

RESTEVA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/restolhal/. Acesso em: 26 fev. 2025.

RIBEIRO, Darcy. **O povo brasileiro**: a formação e o sentido do Brasil. 3. ed. São Paulo: Global Editora, 2015.

RIBEIRO, José *et al.* **Esboço de um atlas lingüístico de Minas Gerais.** Rio de Janeiro: MEC/Casa de Rui Barbosa, 1977.

RODRIGUES, Rosa Belli. **Em busca de uma história social para o léxico rural paranaense.** 2007. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2007.

ROMANO, Valter Pereira; SEABRA, Rodrigo Duarte; OLIVEIRA, Nathan [SGVCLin] – Software para geração e visualização de cartas linguísticas. RELin Revista de Estudos da Linguagem. Belo Horizonte, UFMG, v. 22, n.1, p. 119-151, 2014.

ROSSI, Nelson. **Atlas Prévio dos Falares Baianos.** Rio de Janeiro: Instituto Nacional do Livro, 1963.

SABUCO. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/sabuco/. Acesso em: 7 mar. 2025.

SANTOS, Milton. **O espaço do cidadão**. São Paulo: Edusp, 2007.

SILVA NETO, Serafim da. **Guia para estudos dialectológicos.** Belém: Conselho Nacional de Pesquisas; Instituto Nacional de Pesquisas da Amazônia, 1957.

SILVA, Antônio de Moraes. **Diccionario da língua portugueza**. 3. ed. Lisboa: [Officina de Simão Thaddeo Ferreira], 1823. v. 2. Disponível em: https://www2.senado.leg.br/bdsf/handle/id/562936. Acesso em: 21 fev. 2025.

SILVA-CORVALÁN, Carmen. **Sociolingüística:** teoría y análisis. Madrid: Alhambra, 1989.

SOCA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/soca/. Acesso em: 21 fev. 2025.

SOQUEIRA. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, 2025. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/soqueira/. Acesso em: 21 fev. 2025.

THUN, Harald. Introduction à la table ronde. *In*: INTERNATIONAL CONGRESS OF ROMANCE LINGUISTICS AND PHILOLOGY ROMANES, 22, Tübingem. **Actes** […]. Tübingen: Niemeyer, 1998. p. 407-409. Disponível em: https://archive.org/details/actesduxxiiecong0000inte. Acesso em: 27 fev. 2025.

THUN, Harald. La géographie linguistique romane à la fin du XX siècle. In: CONGRÈS INTERNATIONAL DE LINGUISTIQUEET DE PHILOLOGIE ROMANES, 22., Bruxelles, 1998. Actes...Tübingen: Max Niemeyer, 2000. p. 367-388. v. III.

TOCO. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/toco/. Acesso em: 22 fev. 2025.

TRUDGILL, Peter; CAMPOY, J. M. Hernández. **Diccionario de sociolingüística**. Madrid: Gredos, 2007.

VAGEM. *In*: MICHAELIS, Carolina; MICHAELIS, Henriette. **Michaelis**: dicionário brasileiro da língua portuguesa. São Paulo: Melhoramentos, [2025]. Disponível em: https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/vagem/. Acesso em: 28 fev. 2025.

VESCOVI, Jéssica Paula; SEIDE, Márcia Sipavicius. **Antroponímia no oeste paranaense.** Revista Philologus, v. 55, p. 225-232, 2013.

WACHOWICZ, Ruy Christovam. **Norte Velho, Norte Pioneiro**. Gráfica Vicentina: Curitiba, 1987.