

Análise dialetométrica das cartas lexicais do EALMG: uma contribuição para a descrição dos falares mineiros

Dialectometric Analysis of the Lexical Maps from EALMG: A Contribution to the Description of Minas Gerais Speech

Valter Pereira Romano

Universidade Federal de Santa Catarina
(UFSC) | Florianópolis | SC | BR
vatler.pereira.romano@ufsc.br
<https://orcid.org/0000-0002-8882-3188>

Fernando Brissos

Universidade de Lisboa (UL) | Lisboa | PT
fernandobrisos@campus.ul.pt
<https://orcid.org/0000-0002-2525-1987>

Resumo: O texto apresenta a análise dialetométrica do segundo atlas linguístico publicado no Brasil, o *Esboço de um Atlas Linguístico de Minas Gerais – EALMG* (Ribeiro et al., 1977), com vistas a discutir a divisão dialetal mineira proposta por Zágari (1998, 2005): falar paulista, falar mineiro e falar baiano. A partir do conjunto de 22 cartas lexicais do atlas, emprega-se um método estatístico-matemático para a análise do conjunto do material, utilizando-se a ferramenta on-line DiaTech®, *software* de análise dialetométrica amplamente validado para os parâmetros da Escola Dialetométrica de Salzburgo (Goebel, 2012) (Brissos, Gillier, Saramago, 2017). Como resultado, apresentam-se análises de *clusters* e mapas sinópticos de similaridade e de assimetria, que confirmam parcialmente a tripartição referida do dialeto mineiro. O trabalho contribui com a Dialetologia brasileira pois avança em abordagens dialetométricas que, no que tange à aplicação do método em dados do Português Brasileiro, ainda tem pouca tradição.

Palavras-chave: atlas linguísticos; Minas Gerais; falares mineiros; dialetometria.

Abstract: This paper presents a dialectometric analysis of the second linguistic atlas published in Brazil, the Draft of a Linguistic Atlas of Minas Gerais – EALMG (Ribeiro et al. 1977), with the aim of discussing the dialectal division of Minas as proposed by Zágari (1998, 2005): São Paulo speech, Minas speech, and Bahia speech. Based on a set of 22 lexical maps from the atlas, a statistical-mathematical method is employed to analyze the material, using the online tool DiaTech®, a well-



validated software for dialectometric analysis according to the parameters of the Salzburg Dialectometric School (Goebel, 2012) (Brissos, Gillier, Saramago, 2017). The results include cluster analyses and synoptic maps of similarity and asymmetry, which partially confirm the aforementioned tripartition of the Minas dialect. This study contributes to Brazilian Dialectology by advancing dialectometric approaches, which, concerning their application to Brazilian Portuguese data, are still relatively underexplored.

Keywords: linguistic atlases; Minas Gerais; Minas speech; dialectometry.

1 Introdução

Minas Gerais é um estado-chave para as pesquisas geolinguísticas no Brasil (Romano, Seabra, 2017), não apenas pela sua importância sociocultural e histórica para a formação do Português Brasileiro (PB), como também porque foi a segunda Unidade Federativa que teve seu atlas linguístico estadual publicado que, juntamente com quatro outros atlas estaduais (Bahia, Paraíba, Sergipe e Paraná), instalou a mentalidade dialetológica apregoada por Silva Neto (1957) e deu bases sólidas para outros projetos de atlas que o sucederam, culminando, no final do século XX, no início das atividades do Projeto Atlas Linguístico do Brasil.¹

Um dos grandes responsáveis, e por que não o grande idealizador e divulgador da Geolinguística no Brasil, foi o pesquisador Mário Roberto Lobuglio Zágari, ex-professor da Universidade Federal de Juiz de Fora que, juntamente com outros colegas, lançou-se ao empreendimento da elaboração do que denominaram como *Esboço de um Atlas Linguístico de Minas Gerais* (EALMG), resultando a sua publicação no final da década de 1970 (Ribeiro et al. 1977).

O EALMG é um atlas de extrema importância para a Dialetologia brasileira, e uma obra já estudada por outras pesquisas, Martins (2006), Mendonça e Romano (2020), Zágari (2005[1998])², Rocha e Ramos (2010), Rocha (2012), Rocha e Antunes (2012), Romano e Seabra (2017) e Romano e Cruz (2020), seguindo os princípios de análises da Dialetologia e da Geolinguística clássicas, na comparação e correlação entre as áreas geográficas e aspectos linguísticos correlacionados ao perfil dos informantes do EALMG.

Este artigo soma-se ao conjunto de estudos realizados com base no EALMG, avançando em direção aos anseios do professor Zágari (2005) de desenvolver um estudo dialetométrico das cartas do atlas, segundo o autor:

¹ Conf. <https://alib.ufba.br/>.

² A primeira versão do texto de Zágari com a proposta de delimitação das áreas dialetais de Minas Gerais é de 1998. Utiliza-se neste texto a segunda versão, publicada em 2005, conf. lista de referências.

O Atlas mineiro, em seu último volume (IV), segue esse modelo, usando a medição dos graus de distanciamento entre dois pontos, o Índice Relativo de Identidade (IRI) para as semelhanças e o Índice Relativo de Distância (IRD) para as diferenças de tal modo que $IRI + IRD = 100$, acompanhando Goebel (1981: 361-3) e o Atlas Lingüístico do Litoral Português. (Zágari, 2005, p.62)

Embora tenha sido visionário à sua época, o falecimento do professor Zágari e a dissolução da equipe³ não permitiram que o EALMG tivesse continuidade e hoje se tem apenas o primeiro volume do atlas publicado. Segundo os autores, o primeiro e único volume do EALMG objetivou “de forma simples e visual, apresentar a variação lexical mineira em dois campos semânticos definidos: *tempo e folguedos infantis de rua* – e três segmentos (fones) nos limites virtuais de sua distribuição diatópica” (Ribeiro *et al.* 1977, p. 18), constituindo uma obra de referência que ecoa até a atualidade dado o pioneirismo para a Geolinguística brasileira.

Este artigo objetiva apresentar a análise dialetométrica do EALMG a partir dos dados documentados nas cartas lexicais com vistas a discutir a divisão dialetal de Minas Gerais proposta por Zágari (2005). Para tal intento este trabalho se justifica por retomar a pesquisa do referido dialetólogo, apresentando os resultados do atlas a partir do uso de um método inovador para análise do conjunto do material publicado em um atlas de primeira geração (Cardoso, 2010) que só agora pode ser realizado, considerando os avanços tecnológicos e as tendências da moderna Geolinguística (Silva; Romano, 2022).

2 Os falares de Minas: apresentação e metodologia do EALMG

O volume publicado do EALMG apresenta um conjunto de 73 cartas linguísticas de caráter léxico-fonético, além da introdução em que constam as bases metodológicas do atlas, o plano da obra e um glossário ao final. O atlas traz o resultado da coleta de dados realizada em 116 localidades englobando “grandes cidades, as médias, as pequenas e, também, aqueles “grotões”, locais ermos e perdidos onde o único caminho a ser percorrido era o da volta” (Zágari, 2005, p. 52) e foi projetado para ser publicado em quatro volumes, conforme detalhamento a seguir:

- ♦ Volume I: metodologia, cartas lexicais e fonéticas dos campos semânticos tempo e folguedos infantis de rua, cartas de isófonas, cartas de isoléxicas e o glossário de ambos os campos.
- ♦ Volume II: cartas dos campos semânticos homem e animais, além de cartas de isófonas e isoléxicas e o glossário de ambos os campos.
- ♦ Volume III: cartas lexicais e fonéticas dos campos semânticos água e terra, além de cartas de isoléxicas e as abordagens sociolinguísticas que algumas localidades suscitaram com informação sobre agrupamentos de negros (Serro, Capela Nova e Soledade) e indígenas (Maxacalis). Desse volume também constaria o glossário dos respectivos campos mencionados.

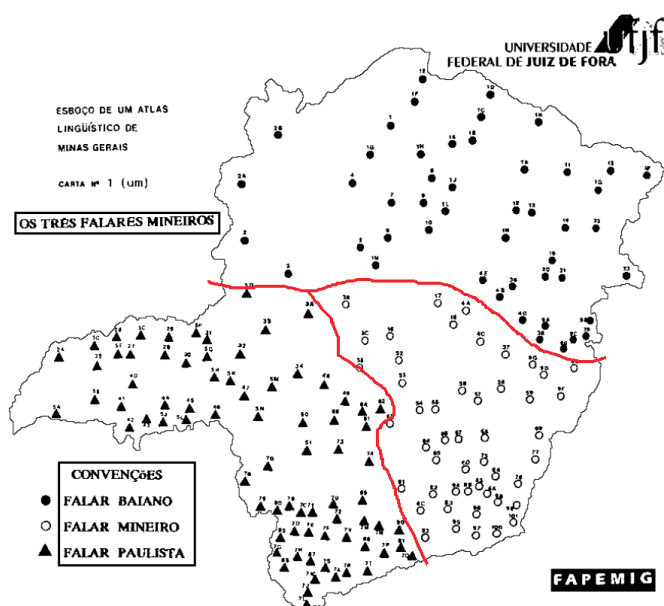
³ A equipe do EALMG foi composta por José Ribeiro, Mário Roberto Lobuglio Zágari, José Passini e Antonio Pereira Gaio, todos falecidos. O professor Zágari foi o que seguiu na linha da geolinguística e integrou o Comitê Nacional do Projeto Atlas Lingüístico do Brasil até 2010, ano de seu falecimento.

- ♦ Volume IV: dedicado para a sistematização e interpretação das formas, segmentos e construções obtidos. Neste volume, segundo Zágari (2005), estariam as análises e os mapas dialetométricos.

Para a seleção dos informantes, os autores do EALMG consideraram os seguintes requisitos: estar na faixa dos 30 e 50 anos de idade; ter baixa ou nenhuma escolarização, ser natural da localidade, não ter vivido em outro município, nem feito muitas viagens; não ter prestado serviço militar, ter boas condições de saúde e fonação e ser uma pessoa desinibida e capaz de representar o falar local (Ribeiro et al. 1977, p. 28). Portanto, o perfil dos informantes do EALMG é topoestático (Thun, 1998), atendendo aos pressupostos de atlas monodimensionais, visto que a variável controlada é a areal ou diatópica.

A partir das cartas do EALMG, Zágari (2005) faz a observação de que no território mineiro coexistem três falares, considerando aspectos sócio-históricos do processo de ocupação e povoamento: o falar baiano, o falar paulista e o falar mineiro, resumidamente apresentados a seguir:

Imagem 1 – Mapa Dialetal de Minas Gerais segundo Zágari (2005)



Fonte: Zágari (2005) – adaptado.

O falar baiano compreende toda a parte norte que vai de leste a oeste do território. Caracteriza-se pelo uso de vogais média-baixas em posição pretônica, como em c[o]ração. c[ɔ]ração, p[e]cado.p[ɛ]cado; o uso de africadas antes de vogal alta posterior, mui[t]o.mui[tʃ]o, pei[t]o.pei[tʃ]o; além da nasalização em sílabas pretônicas: b[a]nana.b[ã]nana. Sobre o léxico, o autor afirma: “Itens lexicais comuns ocorrem, mas veiculando significados ignorados⁴ nas outras regiões, a saber: neve (= cerração), chuva-de-flor (= granizo), zelação (= estrela

⁴ À época (década de 1970), não havia, praticamente, estudos geolinguísticos com os quais Zágari pudesse fazer a comparação. Muitas dessas denominações foram registradas no Atlas Linguístico do Paraná (ALPR) (Aguilera,

cadente), china (= bola-de-gude), queiro (= dente-de-siso), bituca (= toco de cigarro), ponga (= carona)” (Zágari, 2005, p.50).

Quanto ao falar paulista, engloba o sul de Minas e o Triângulo Mineiro, caracterizado tipicamente pelo /r/ retroflexo. Para o autor, “nesse falar, o ritmo de fala é mais veloz, contrastando com o ritmo mais arrastado do norte, verificando-se, lexicalmente, certas preferências como ramona (= grampo), rabicó (= animal sem rabo), cachopa ou caixote (= colmeia) e chuva-de-rosa (= granizo)” (Zágari, 2005, p. 51).

Já o falar mineiro, segundo o autor, está “preso entre essas duas áreas” (Zagari, 2005, p. 51) e não possui nenhuma das características dos outros dois. Caracteriza-se do ponto de vista fonético pela monotongação e ditongação, uma vez que “desfaz constantemente os ditongos [aj], [ej], [ow] quando não finais, caixa.“caxa”, peixe.“pexe”; ouro.“oro” e faz surgirem outros, quando finais e antecidos de sibilantes” (p. 51), como em arroz.‘arrois’, faz.“fais”, nós.“nóis”.. Compreende a parte centro-leste de Minas, englobando a Zona da Mata Mineira (região de Juiz de Fora) e a região metropolitana de Belo Horizonte.

Como síntese das constatações de Zágari (2005), cabem as palavras do próprio autor:

Ao estabelecer essas fronteiras, diga-se ser impossível demarcá-las como definitivas, quer por não se poder balizá-las sem inter cruzamentos (*grifo nosso*), quer porque aqui e ali elas se tocam desordenadamente, quer porque o tempo mostrará que elas se movem, quer porque o que existe são fenômenos fonéticos e lexicais cuja difusão, muitas vezes ou sempre, operam de forma independente. Isso não anula, contudo, a realidade que interpõe esses três falares num jogo contrastivo: um belo-rizantino, um januarioense e um uberlandense (*grifo nosso*) se sabem brasileiros e mineiros pela língua que falam, mas se sabem, também, participantes de uma variedade, de uma diferente norma de fala. Qualquer observador atento notará serem eles oriundos de espaços diferentes das Minas Gerais (Zágari, 2005, p. 51).

Considerando as análises geolinguísticas sob a metodologia da Dialetologia, ao modo clássico⁵, outros trabalhos discutem essa tripartição de Zágari (2005), dentre os quais, citam-se os trabalhos de Martins (2006), Mendonça e Romano (2020), Rocha e Ramos (2010), Rocha (2012), Rocha e Antunes (2014), Romano e Cruz (2020) e Romano e Seabra (2017), em diferentes perspectivas.

Esses estudos revelam que, embora alguns aspectos linguísticos evidenciem contradições sobre a tripartição do território, do ponto de vista da fonética, mostram-se “tendências de pronúncia que ainda continuam ocorrendo no estado” (Rocha e Antunes, 2010 p. 109-

1994) e no Atlas Linguístico-Etnográfico da Região Sul do Brasil (ALERS) (ALTENHOFEN et al., 2011): chuva de flor, bituca e ramona e no ALPR (chuva de rosa, dente queiro).

⁵ Por “modo clássico”, no âmbito da Geolinguística brasileira, entendem-se as análises de cunho quantitativo, segundo a estatística descritiva, e qualitativo, a partir das interpretações dos resultados em correlação com outros estudos, bem como com a história social das áreas investigadas. Esse é o modelo clássico, já tradicional, entre os geolinguistas e dialetólogos brasileiros para estudo dos atlas. Exceção a esse modelo, é a tese de Romano (2015) que avança em direção à estatística inferencial a partir de hipóteses com o uso do SPSS (Statistical Package for the Social Sciences). Outros dois estudos seguem essa mesma linha, também com dados geolinguísticos: Romano e Seabra (2014) e Romano e Seabra (2017). As abordagens dialetométricas tratadas nas seções seguintes também são um exemplo de reformulação do modo clássico.

110), porém, não se sustentam sobre os dados lexicais: Rocha (2012), Rocha e Ramos (2010), Romano e Cruz (2020) e Romano e Seabra (2017).

Mesmo assim, à luz da sociohistória das regiões, de certa forma, os trabalhos se coadunam às ideias de Zágari (2005), o que põe em relevo a perspicácia do dialetólogo mineiro na delimitação de áreas dialetais. A modéstia do autor revela-se ao afirmar que as áreas não são definitivas principalmente por “não se poder balizá-las sem inter cruzamentos” (Zágari, 2005, p. 51). Ora, neste momento, com os avanços da Dialetologia e recursos computacionais que se tem, é possível verificar, mesmo que parcialmente, os inter cruzamentos dessas áreas por meio de uma análise dialetométrica. Essa é a contribuição deste texto como uma forma de homenagem e reconhecimento ao legado do professor Zágari para a Geolinguística do Brasil.

3 Dialetometria nos atlas linguísticos brasileiros

A dialetometria é um método de análise matemático-estatística aplicada a *corpora* geolinguísticos (sobretudo atlas), cujo principal objetivo é medir distâncias e semelhanças linguísticas entre locais ou grupos de locais. Jean Séguéy (1973) foi o criador desse método em uma época em que não existiam ferramentas computacionais para correlacionar cálculos matemáticos com representação cartográfica. Ele e sua equipe de investigadores compararam as respostas de cada ponto de inquérito do *Atlas Linguístico da Gasconha* (1954) com as discordâncias de pontos vizinhos, construindo, desse modo, uma matriz de dissimilaridade. O número de diferenças entre os pontos foi reduzido a porcentagens e essa porcentagem tratada como uma pontuação de índices que evidenciam a distância linguística entre dois pontos.

Mas, foi com Hans Goebel (1974; 1976) que esse método se difundiu devido aos avanços não só do ponto de vista dos cálculos necessários para medir as distâncias, como também no que tange ao primeiro pacote computacional para tratar da dialetologia quantitativa, o programa desktop VDM (Visual DialectoMetry) (Goebel 2004, Haimenl 2006). A partir dos trabalhos de Goebel, outros o sucederam, inclusive, no que se refere às ferramentas computacionais que possibilitam as análises dialetométricas.

Para Brissos, Gillier e Saramago (2019, p.560):

[...] a dialetometria surgiu da necessidade de *digerir* cientificamente *corpora* dialetais de grande dimensão, como atlas linguísticos, que só podem ser analisados eficientemente a partir de uma abordagem matemático-estatística. O ponto de viragem foram os anos 70 do século passado, quando, por um lado, se tinha acumulado um número significativo de atlas linguísticos que era preciso interpretar de forma sistemática, e, por outro lado, a matemática começava a ser uma ferramenta de uso comum no estudo da variação linguística.

Com o incremento da informática, programas computacionais surgiram para auxiliar os pesquisadores no tratamento quantitativo do material geolinguístico. Wieling e Nerbonne (2015) assumem que durante muito tempo duas ferramentas foram utilizadas para análises dialetométricas: uma desenvolvida na Universidade de Salzburgo, a já referida VDM, mais ligada aos trabalhos de geolinguística românica e dependente dos princípios da conhecida Escola Dialetométrica de Salzburgo; e a outra desenvolvida na Universidade de Groningen,

por Peter Kleiweg, o RUG/LO4⁶, que em sua versão *on-line*, resultou na aplicação GapMap (Nerbonne et al., 2011), amplamente utilizada nas línguas germânicas e outras ligadas à Escola Dialetométrica de Groningen.

Outro programa de análise dialetométrica é a ferramenta *on-line* DiaTech (Aurrekoetxea et al., 2013; Aurrekoetxea et al., 2016), criada na Universidade do País Basco, com a finalidade de resolver duas questões importantes para a moderna dialetologia: (i) as respostas múltiplas em um único ponto de inquérito e (ii) a comparabilidade dos resultados da análise estatística de dados linguísticos de diferente natureza (Aurrekoetxea et al. 2013; Aurrekoetxea et al. 2016). Os autores afirmam que com a aplicação DiaTech há uma melhoria na medição das distâncias ou semelhanças linguísticas em relação às ferramentas da dialetologia tradicional e que, particularmente no ponto (i), é fornecida já uma solução eficiente. Os vários trabalhos da ainda jovem Escola Dialetométrica de Lisboa, que tem utilizado os princípios de Salzburgo com algumas modificações, incluindo o *software*, comprovam-no (Brissos 2016; Brissos, Gillier e Saramago 2017; Brissos 2020; Brissos 2021).

A aplicação DiaTech é gratuita e apresenta uma interface intuitiva e amigável para o dialetólogo, pois permite o *upload* de dados e desenhos de mapas correspondentes a partir do Google Maps. Há também a possibilidade de gerir bases de dados e utilizar muitos parâmetros dialetométricos, tais como mapas de similaridade, mapas de zonas de transição, análises de correlação, mapas de feixes, análises de *clusters*, entre outros, com diferentes algoritmos de análises estatísticas.

Para este trabalho, optou-se pelo uso da DiaTech, seguindo os procedimentos bem estabelecidos para dados de língua portuguesa pela referida Escola Dialetométrica de Lisboa a exemplo de pesquisas como as de Brissos (2015) e Brissos (2016) sobre o português europeu; Brissos, Gillier e Saramago (2017) e Brissos, Gillier e Saramago (2016), sobre os arquipélagos dos Açores e da Madeira, respectivamente e, sobre o Português Brasileiro, Brissos e Saramago (2019), Brissos (2021) e Cristianini (2023).

No Brasil, o trabalho pioneiro em dialetometria foi a tese de doutorado de Altino (2007) que, com base nos dados do *Atlas Linguístico do Paraná* (ALPR) (Aguilera 1994) e nos dados inéditos do volume II do ALPR, organizado para sua tese, apresenta duas cartas dialetométricas com Índice Relativo de Distância (IRD) e Índice Relativo de Identidade (IRI) entre os pontos linguísticos, seguindo o quadro metodológico de Salzburgo. O segundo trabalho sobre dialetometria em atlas linguístico brasileiro é o de Saramago e Cardoso (2010), que, consoante ao trabalho de Altino (2007), também apresentam uma análise de mapas dialetométricos dos dados do *Atlas Linguístico de Sergipe* (Ferreira, 1987) e do *Atlas Linguístico de Sergipe II* (Cardoso, 2005).

Ressalte-se que ambos os trabalhos, Altino (2007) e Saramago e Cardoso (2010), constituíram a matriz de dados e aplicaram os métodos estatísticos sem o uso de um *software* específico para os cálculos e representação cartográfica de forma automatizada. Contudo, essas pesquisas são de excepcional importância ao apontar para as especificidades no uso da dialetometria no Brasil, uma vez que é necessário considerar particularidades metodológicas que os atlas linguísticos brasileiros encerram: (i) densidade da rede de pontos e sua distribuição pelo território, (ii) redes de comunicação e intercomunicação, (iii) vários informantes por ponto linguístico, (iv) perfil dos informantes, (v) respostas múltiplas não mononímicas etc.

⁶ <http://www.let.rug.nl/kleiweg/lo4>

Os estudos dialetométricos na Geolinguística brasileira vêm se expandindo ligados à Escola de Lisboa, a exemplo das referências já citadas e de projetos de pós-doutoramento desenvolvidos no Centro de Linguística da Universidade de Lisboa cujos resultados se encontram ainda inéditos (citam-se Augusto 2018, Takano 2018 e Soares 2019, por exemplo). A presente pesquisa, que avança para a dialetometrização de um atlas e um estado ainda não explorados, fornece mais um contributo no sentido da dinamização dos estudos dialetométricos brasileiros.

Outro ponto salutar para o florescimento do método dialetométrico em solo brasileiro é a quantidade de atlas linguísticos de grandes e pequenos territórios já divulgados, conforme observa Romano (2020), ao indicar que, à época, se contavam 63 atlas linguísticos de pequeno domínio e 14 atlas estaduais concluídos no Brasil. A esse quantitativo acrescentam-se mais dezenas de trabalhos desenvolvidos e em desenvolvimento de 2020 a 2025, o que revela o quão profícua a Geolinguística brasileira se apresenta ao mundo, com dados suficientes para análises dialetométricas de atlas já publicados/divulgados. A utilização desse método de análise nos atlas brasileiros parece ser o passo seguinte para os dialetólogos e é nesse sentido que o presente trabalho avança. Um quadro atual dos estudos dialetométricos no Brasil estará disponível brevemente em Romano e Brissos (2025).

4 Metodologia da pesquisa

De posse do atlas, o estudo dialetométrico requer do geolinguista o trabalho inverso à cartografia, ou seja, a partir do material documentado nas cartas linguísticas, as respostas são transferidas para uma planilha de dados. Neste caso específico, seguimos os procedimentos formais exigidos pelo DiaTech. Assim, foram construídas planilhas, em ambiente LibreOffice®, com detalhamentos das respostas por ponto linguístico. Para cada conceito, carta linguística, foram documentadas as respostas correspondentes na rede de pontos, conforme exemplo:

Imagem 2 – Modelo de planilha de dados

	A	B	C	D	E	F
1		Januária	Paracatu	João Pinheiro	São Romão	Pirapora
2	Arco-íris	Arco-da-velha	Arco-íris	Arco-íris	arco-da-velha	Arco-íris
3	Orvalho	sereno	neve	sereno	sereno	orvalho
4	Mormaço	mormaço	mormaço	mormaço	mormaço quente	mormaço
5	Garoa	garoa	inverno	garoa	garoa	garoa
6	Neblina	neve	neblina	neblina	neve	neblina
7	Chuva de pedra	sem resposta	Chuva-de-pedra	Chuva-de-pedra	Chuva-de-flor	Chuva-de-flor
8	tempestade	temporal	tempestade	tempestade	tempestade	temporal
9	trovão	sem resposta	trovão	trovão	trovão	trovoada
10	relâmpago	relâmpago	raio	relâmpago	relâmpago	relâmpago
11	Tromba-d'água	Bomba-d'água	Bomba-d'água	sem resposta	Tromba-d'água	Tromba-d'água
12	Estrela cadente	zelação	Mãe-do-ouro	satélite	sem resposta	sem resposta
13	Anteontem	Ontem-ontem	Antes-de-ontem	Antes-de-ontem	sem resposta	Antes-de-ontem

Fonte: dados da pesquisa.

Na primeira coluna (A), encontram-se os conceitos, nominalizados pelas respectivas cartas do atlas. As colunas subsequentes apresentam as localidades, 118 ao todo, pois são 116 localidades de Minas Gerais às quais foram adicionados dois pontos artificiais, denominados

PPB (Português Padrão Brasileiro) e PPE (Português Padrão Europeu)⁷, seguindo a mesma metodologia da pesquisa de Brissos e Saramago (2019), com a finalidade verificar a integração das formas aos dois padrões de variação linguística mencionados, condizentes a cada uma das duas variedades da língua portuguesa.

Em cada linha, encontram-se as respostas. Quando não se tinha representação de variante para um ponto linguístico da carta, foi feita a anotação “sem resposta”. As respostas múltiplas em cada ponto de inquérito foram separadas por vírgula.

Para a análise dialetométrica do EALMG, foram utilizadas 22 cartas linguísticas que apresentam um total de 3090 dados, contabilizando deste universo 92% de respostas válidas (2.843 dados) no *corpus* analisado e 8% de localidades sem variante lexical (247 dados são de “sem resposta”). Ressalta-se que as 23 cartas fonéticas do EALMG não são objeto deste artigo, uma vez que se propõe um estudo dialetométrico da variação lexical.⁸ Da mesma forma, não se utiliza as 28 cartas que são de isófonas ou de isoléxicas, uma vez que já pressupõem certo grau de análise, abstração e espacialização das variantes

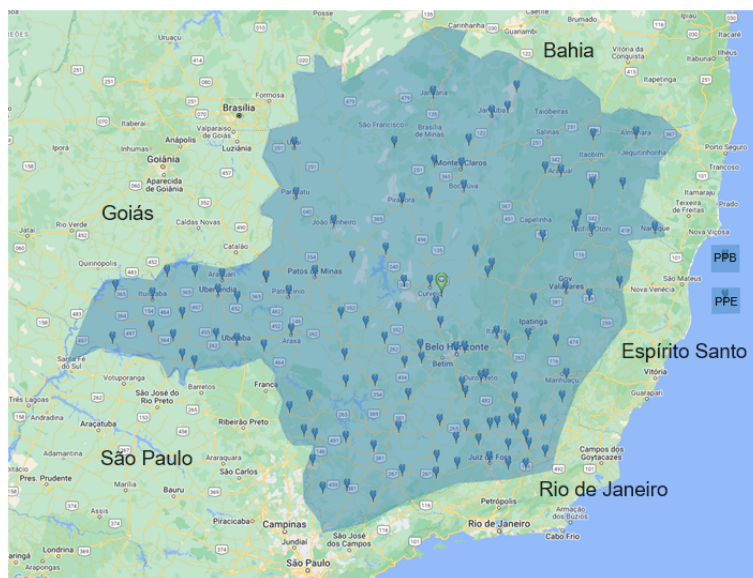
As cartas selecionadas são as que recobrem os designativos para os seguintes conceitos: *arco-íris*; *orvalho*; *mormaço*; *garoa*; *neblina*; *chuva de pedra*; *tempestade*; *trovão*; *relâmpago*; *tromba-d'água*; *estrela cadente*; *anteontem*; *cambalhota*; *bolinha de gude*; *búrica*; *pique*; *chicotinho-queimado*; *cabra-cega*; *porrinha*; *pular carniça*; *papagaio*; *estilingue*.

A planilha de dados foi importada no DiaTech, a partir do qual se verificou cada um dos pontos linguísticos inserindo individualmente suas coordenadas no Google Maps, que está vinculado à aplicação. Depois de feita a conferência de todos os pontos, passou-se ao contorno do mapa (feito manualmente com o *mouse* seguindo o mais perto possível a fronteira política do território) conforme se observa na geolocalização de cada município. O PPB e PPE foram geolocalizados na margem direita do território, em cima do oceano (Imagem 3) como dois retângulos:

⁷ Sobre o PPB e o PPE, para definição da forma ‘padrão’ de cada variedade lexical, foram considerados os itens de uso mais difundido a partir conhecimento dos autores sobre o léxico do Português Brasileiro e o Português Europeu e também considerando a dicionarização em duas obras lexicográficas, o Dicionário Houaiss da Língua Portuguesa (Houaiss, Villar e Franco, 2001), para o PPB; e o Dicionário da língua portuguesa, da Porto Editora(2011), para o PPE..Exemplos: para o PPB, foi considerada a variante estrela cadente em vez de mãe-de-ouro, arco-íris em vez de arco-da-velha entre outras, uma vez que a primeira não revela marcas regionais, já a segunda apresenta traços de marcação diatópica. Para o PPE, foi considerada variante berlinde para bolinha-de-gude, fisga para estilingue e assim sucessivamente.

⁸ Uma etapa seguinte para esta pesquisa será elaborar a análise dialetométrica das cartas fonéticas.

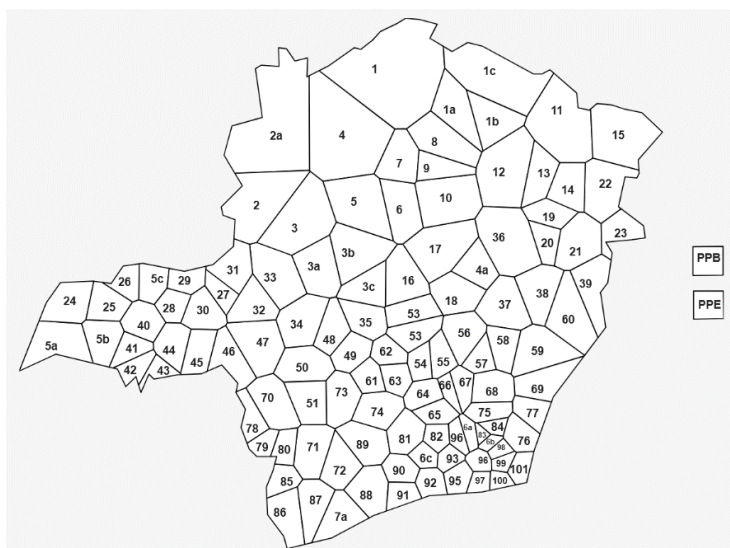
Imagem 3 – Rede de pontos do EALMG georreferenciada



Fonte: dados da pesquisa.

No DiaTech obteve-se o mapa base das cartas dialetométricas no qual o *software* inseriu automaticamente a poligonação de Voronoi (Imagem 4), e foi feita em ferramenta de designer a numeração dos polígonos conforme a rede de pontos (Quadro 1).

Imagem 4 – Rede de pontos do EALMG (Poligonação de Voronoi)



Fonte: dados da pesquisa.

A Imagem 4 permite verificar os limites entre determinado ponto linguístico em relação aos seus vizinhos, formando polígonos de um ponto em intersecção com outro, partindo-se do conceito interpontual dos primeiros trabalhos dialetométricos.

A rede de pontos linguísticos do EALMG é numerada de 1 a 101, aos quais os autores do atlas acrescentaram 15 pontos complementares: 1A, 1B, 1C, 2A, 3A, 3B, 3C, 4A, 5A, 5B, 5C, 6A, 6B, 6C e, por fim, 7A; detalhados no Quadro 1.

Quadro 1 – Relação da rede de pontos do EALMG (número/localidade)

1	Januária	26	Centralina	65	Conselheiro Lafaiete
1A	Janaúba	27	Monte Alegre	66	Ouro Preto
1B	Porteirinha	28	Uberlândia	67	Mariana
1C	Mato Verde	29	Araguari	68	Ponte Nova
2	Paracatu	30	Indianópolis	69	Manhumirim
2A	Unaí	31	Monte Carmelo	70	Passos
3	João Pinheiro	32	Patrocínio	71	Alfenas
3A	São Gonçalo do Abaeté	33	Patos de Minas	72	Varginha
3B	Barreiro Grande	34	São Gotardo	73	Formiga
3C	Felixlândia	35	Pompéu	74	Oliveira
4	São Romão	36	Água Boa	75	Viçosa
4A	Santo Antônio do Itambé	37	Açucena	76	Muriáe
5	Pirapora	38	Governador Valadares	77	Carangola
5A	Iturama	39	Mantena	78	São Sebastião do Paraíso
5B	Campina Verde	40	Prata	79	Guaxupé
5C	Tupaciguara	41	Comendador Gomes	80	Muzambinho
6	Jequitaí	42	Frutal	81	São João Del Rei
6A	Silverânea	43	Planura	82	Barbacena
6B	Rodeiro	44	Veríssimo	83	Ubá
6C	Ibertioga	45	Uberaba	84	Visconde do Rio Branco
7	Coração de Jesus	46	Sacramento	85	Poços de Caldas
7A	Itajubá	47	Araxá	86	Ouro Fino
8	Capitão Enéias	48	Dores do Indaiá	87	Pouso Alegre
9	Montes Claros	49	Bom Despacho	88	Caxambu
10	Bocaiúva	50	Bambuí	89	Lavras
11	Medina	51	Piui	90	Andrelândia
12	Virgem da Lapa	52	Cordisburgo	91	Liberdade
13	Araçuaí	53	Sete Lagoas	92	Olaria
14	Padre Paraíso	54	Belo Horizonte	93	Santo Dumont
15	Almenara	55	Sabará	94	Mercês
16	Curvelo	56	Itabira	95	Juiz de Fora
17	Diamantina	57	Nova Era	96	São João Nepomuceno
18	Serro	58	Timóteo	97	Mar de Espanha
19	Ladainha	59	Caratinga	98	Cataguases

20	Poté	60	Galiléia	99	Leopoldina
21	Teófilo Ottoni	61	Divinópolis	100	Além Paraíba
22	Águas Formosas	62	Pará de Minas	101	Pirapetinga
23	Nanuque	63	Itaúna		
24	Santa Vitória	64	Congonhas do Campo		
25	Ituiutaba	67	Mariana		

Fonte: EALMG (1977).

Quanto aos parâmetros utilizados, foram selecionados a análise de *clusters* e os mapas sinóticos. Dentre as análises de *clusters*, inicialmente foi utilizada uma linha de corte de dois e três grupos, para verificar a tripartição de Zágari (2005) e, por fim, de sete grupos, um nível ainda inteligível para interpretação.

Quanto aos mapas sinóticos, foram selecionados dois tipos de análise: (i) distribuição de similaridade, tomando como pontos de comparação, além do PPB e do PPE, três cidades que representam cada um dos três falares de Minas Gerais, Belo Horizonte (falar mineiro), Uberlândia (falar paulista) e Januária (falar baiano); (ii) a distribuição de assimetria entre todas as localidades. Em ambos os parâmetros dialetométricos foi utilizado como ferramenta de computação das semelhanças e diferenças entre pontos o Índice Relativo de Identidade (IRI) e como sistema de visualização o algoritmo MinMwMax.

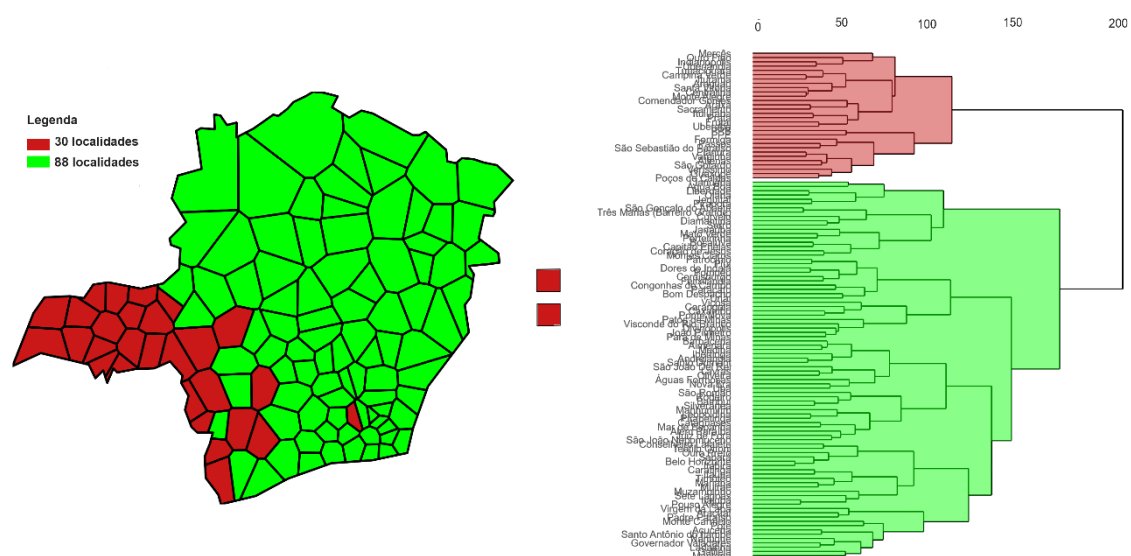
5 Resultados

Os resultados estão apresentados em duas partes. A primeira considera a análise de *clusters* e a segunda os mapas sinóticos de similaridade e assimetria.

5.1 Análise de *clusters*

Neste parâmetro dialetométrico, a partir de dois agrupamentos – ou seja, *clusters* –, encontram-se duas grandes áreas linguísticas representativas em Minas Gerais: (i) uma localizada no Triângulo Mineiro que se expande para o sul do estado e (ii) outra mais estendida que compreende o restante do território do estado, conforme a Figura 5.

Imagem 5 – Resultados de dois *clusters*

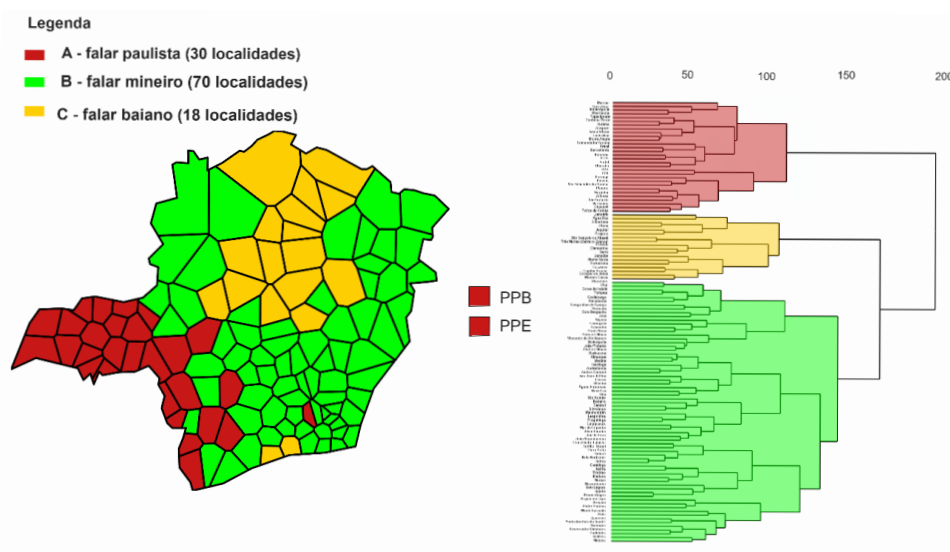


Fonte: dados da pesquisa.

Das 118 localidades, 30 estão representadas no grupo homogêneo (cor vermelha) concentrado na porção ocidental com espalhamento para o sul do estado; enquanto grande parte do território está compreendida em outro grupo, com 88 localidades, diferenciando-se dialetalmente da porção ocidental-sul de Minas Gerais. O dendrograma à direita do mapa apresenta a hierarquização das localidades. Devido à rede de pontos densa, observa-se a dificuldade de separar os agrupamentos mínimos da hierarquização (os *best friends*, i.e. os pares de localidades do final do dendrograma).

A partir desses resultados gerais, parece ser interessante apresentar o recorte do dendrograma com três grupos, com vistas a adentrar às análises detalhadas dos *clusters* e verificar se há coincidência das áreas com a tripartição do território proposta por Zágari (2005).

Imagem 6 – Resultados de três *clusters*



Fonte: dados da pesquisa.

- (i) A partir desse detalhamento (Figura 6), os agrupamentos podem ser divididos conforme segue:
 - (ii) Área A (cor vermelha), correspondente ao Triângulo Mineiro com extensão à parte do sul de Minas Gerais (30 localidades);
 - (iii) Área B (verde), que compreende a maior parte do território (70 localidades);
- Área C (amarelo), na porção centro-norte do estado (18 localidades).

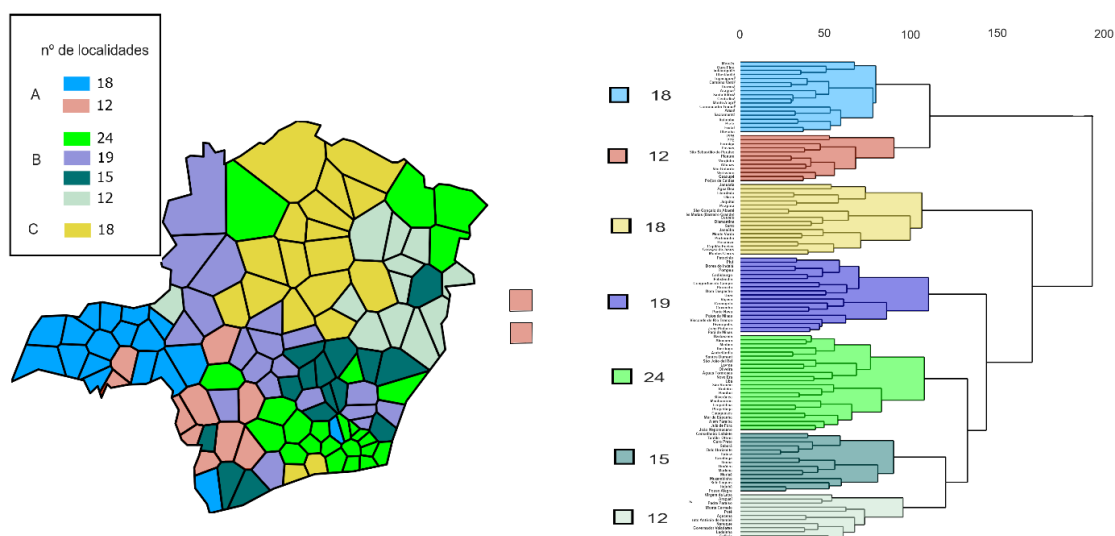
Observe-se que esses dados indicam os possíveis três falares de Zágari (2005), podendo a Área A ser identificada como *falar paulista*, igualmente distribuída pelo Triângulo Mineiro e sul de Minas Gerais; a Área B, como *falar mineiro* que nesta análise dialetométrica expande-se para o noroeste e nordeste de Minas Gerais e também a localidades ao sul; por fim, (iii) uma faixa do centro-norte, analogamente denominada como *falar baiano* (C), neste caso em específico, com ressalvas, pois há neste grupo duas localidades na Zona da Mata: Liberdade e Olaria⁹.

Considerando os três falares (A, B, C), passa-se a verificar detalhadamente cada um dos grupos. Para isso, faz sentido interpretar os resultados a partir de subgrupos:

- A) A - *falar paulista* em dois subgrupos (a¹ e a²);
- B) B - *falar mineiro* em quatro subgrupos (b¹, b², b³ e b⁴);
- C) C - o *falar baiano*, único grupo (C), pois, na análise dendrográfica de sete *clusters*, continua homogêneo, conforme a Imagem 7. Um detalhamento acima de sete grupos torna os dados de difícil compreensão para os demais grupos de falares, ou seja, os dados ficam demasiadamente fragmentados.

⁹ Para se referir às regiões de Minas Gerais, neste texto, opta-se pela distribuição conforme os pontos cardeais: Norte, Sul, Leste, Oeste etc. Em alguns momentos, as áreas linguísticas são mencionadas pelas Mesorregiões do território, conforme a classificação do Instituto Brasileiro de Geografia e Estatística. O estado de Minas Gerais compõe-se de 12 mesorregiões: Noroeste de Minas, Norte de Minas, Jequitinhonha, Vale do Mucuri, Triângulo Mineiro e Alto Paranaíba, Central Mineira, Metropolitana de Belo Horizonte, Vale do Rio Doce, Oeste de Minas, Sul e Sudoeste de Minas, Campo das Vertentes e Zona da Mata. Ressalte-se, contudo, que, para descrição e análise dos resultados, não se ateve estritamente a essa distribuição. Confira: <https://www.ibge.gov.br/cidades-e-estados/mg.html>

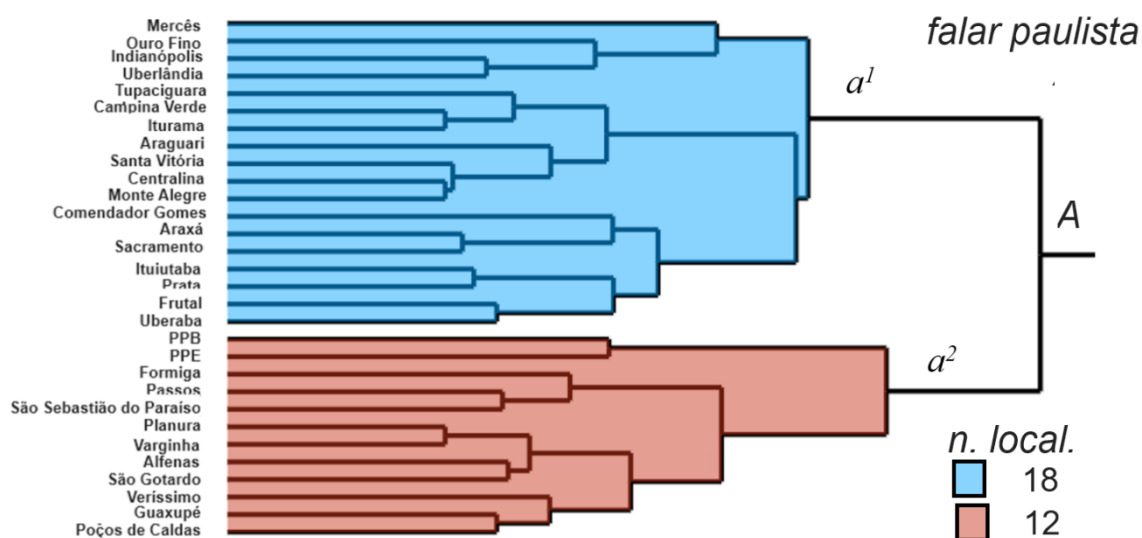
Imagem 7 – Resultado de sete *clusters*



Fonte: dados da pesquisa.

Entre as localidades do *falar paulista* (30 no total), observam-se, basicamente, dois grupos: o primeiro (a^1), que compreende 18 localidades e o segundo (a^2), com 12 localidades (Figura 8). Embora pertençam ao mesmo falar no território, caracterizado pela influência linguística e proximidade geográfica do estado de São Paulo, observa-se que a^1 e a^2 também se dividem conforme a distribuição geográfica das localidades. Para melhor visualização dessa bipartição do *falar paulista*, na Imagem 8, é possível observar cada par de *best friends*.

Imagem 8 – Dendrograma do *falar paulista*



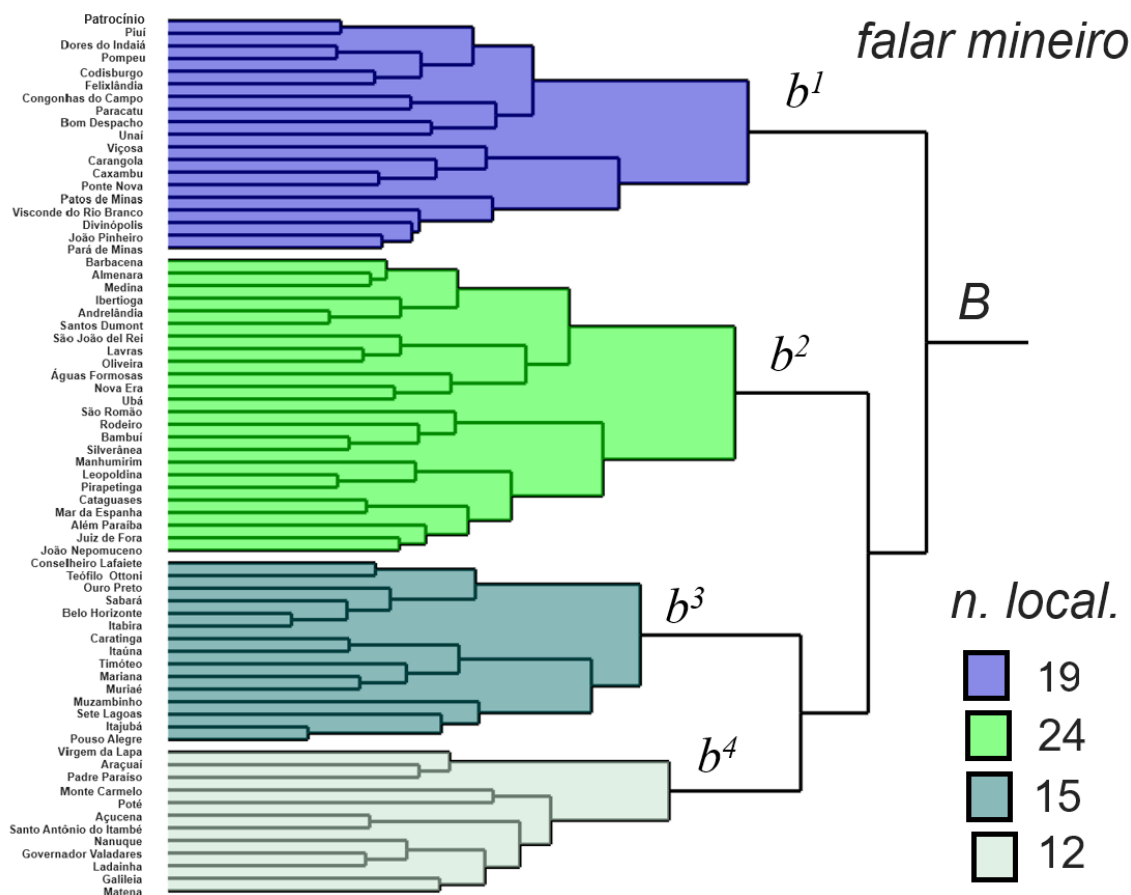
Fonte: dados da pesquisa.

O subgrupo a^1 é composto por Mercês, Ouro Fino, Indianópolis e Uberlândia e outro conjunto de localidades que se inicia em Tupaciguara e se encerra em Uberaba, todas localidades do Triângulo Mineiro. Ressalte-se que tanto Uberlândia quanto Uberaba são as maio-

res cidades da região (aproximadamente 100 km de distância em linha reta pela estrada BR-050), porém, linguisticamente, não formam *best friends*. A segunda une-se a Frutal, ao passo que Uberlândia divide o *cluster* com Indianópolis.

Já na porção do sul de Minas Gerais (α^2), encontram-se o PPB e o PPE integrados entre si como *best friends*, consoante aos resultados da pesquisa de Brissos (2021), que dialetometrizou o *Atlas Linguístico do Amazonas* (ALAM) (Cruz, 2004), e, em menor medida, aos resultados de Brissos e Saramago (2019), que dialetometrizam o ALERS (Altenhofen; Klassmann, 2011). No seu todo, o grupo α^2 compreende, para além desses dois pontos, dez localidades entre Formiga e Poços de Caldas, todas no sul de Minas Gerais. Dessa porção, as principais cidades são Varginha, *best friend* de Planura, e Poços de Caldas, que se agrupa com Guaxupé, e estas duas integradas a Veríssimo. Vale notar, contudo, que, nos dados analisados, algumas localidades do sul de Minas não ficaram agrupadas no *falar paulista*, a saber: Itajubá, Pouso Alegre, Muzambinho e Bambuí. Estas, por sua vez, estão no bojo das localidades recobertas pelo falar mineiro (B), conforme a Imagem 9.

Imagem 9 – Dendrograma do *falar mineiro*



Fonte: dados da pesquisa.

Na análise dialetométrica, a área do *falar mineiro* está expandida, conforme os mapas das Figuras 5 e 6, compreendendo localidades do nordeste e noroeste mineiro, além da região central do estado e alguns pontos do sul do território. O *falar mineiro*, segundo os *clusters*, pode ser dividido em quatro grupos, a saber:

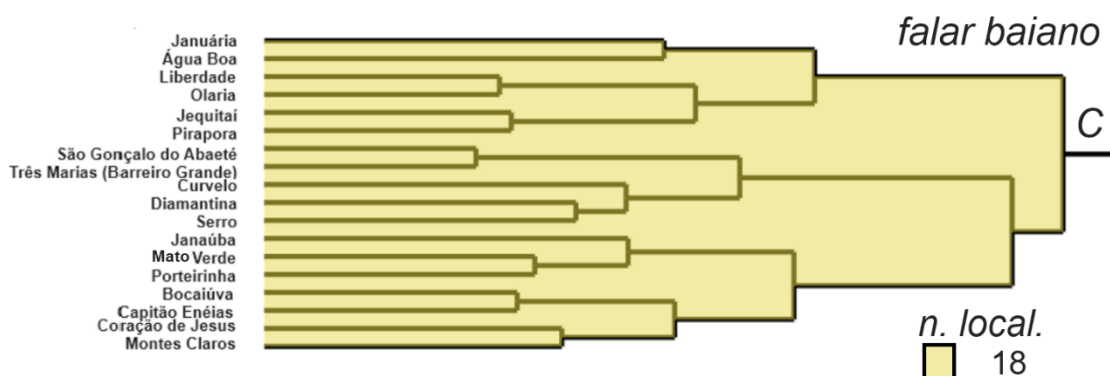
O primeiro subgrupo (b^1) engloba 19 localidades (em roxo), com os pontos linguísticos em diferentes regiões, a citar: Patrocínio, Paracatu, a noroeste; ao sul, Caxambu e na região central, Bom Despacho, por exemplo. Portanto, tem uma distribuição esparsa no território.

O segundo subgrupo (b^2) encerra 24 localidades, entre as quais Barbacena (Campo das Vertentes), Juiz de Fora (Zona da Mata) e Águas Formosas (Vale do Mucuri), o que mostra que também está disperso pelo território. O terceiro *cluster* (b^3) compreende 15 localidades como, por exemplo, Conselheiro Lafaiete (no Campo das Vertentes), Belo Horizonte (região metropolitana de BH), e cidades do sul do estado, Pouso Alegre, Itajubá, Muzambinho, no qual também se observa que, pela distribuição, não apresenta um nexo geográfico contínuo entre as localidades.

Por fim, o menor subgrupo do *falar mineiro* (b^4), com 12 localidades, que está numa região com pontos mais integrados entre si, no nordeste do estado (Vale do Jequitinhonha e do Mucuri). Em síntese, consoante os resultados dialetométricos do EALMG, observa-se que, na análise dos *clusters*, o *falar mineiro* está amplamente difundido pelo território, não se circunscrevendo a uma região em específico.

Ainda sobre análise de *clusters*, passa-se à área C, denominada *falar baiano*, conforme Imagem 10.

Imagem 10 – Dendrograma do *falar baiano*



Fonte: dados da pesquisa.

Uma visualização geral dos dados permite observar que a área do *falar baiano* se concentra, principalmente, em localidades do Norte de Minas, como Pirapora, Januária, Janaúba, Montes Claros, entre outras, mostrando-se mais homogêneo geograficamente que os outros dois falares. Uma única ressalva que se pode notar nessa integração é que Olaria e Liberdade (localidades do sul do estado), como *best friends*, estão também neste *cluster*. Os outros *best friends* são grupos de localidades próximas: Januária e Água Boa, ao norte, são *best friends*, seguindo-se Jequitai com Pirapora, São Gonçalo do Abaeté e Três Marias, Curvelo e Diamantina. Ressalte-se, contudo, que, apesar de haver uma relativa homogeneidade, a área do falar baiano, nas análises de *clusters*, não compreende todo o norte do estado, conforme Zágari (2005), pois não engloba localidades ao noroeste e nordeste de Minas Gerais.

Após ter detalhado cada um dos *clusters* a partir da tripartição de Zágari (2005), focalizam-se nas próximas seções as análises de similaridade inicialmente em relação a três cidades que representam cada um dos três falares (Uberlândia, Belo Horizonte e Januária)

e, depois, em relação ao PPB e ao PPE, seguindo-se a análise da distribuição de assimetria de todas as localidades. Cabe, neste momento, retomar a afirmação de Zágari (2005, p.51) sobre o fato de que

um belorizontino, um januarenses e um uberlandense (*grifo nosso*) se sabem brasileiros e mineiros pela língua que falam, mas se sabem, também, participantes de uma variedade, de uma diferente norma de fala. Qualquer observador atento notará serem eles oriundos de espaços diferentes das Minas Gerais.

É nessa direção que as próximas seções seguem.

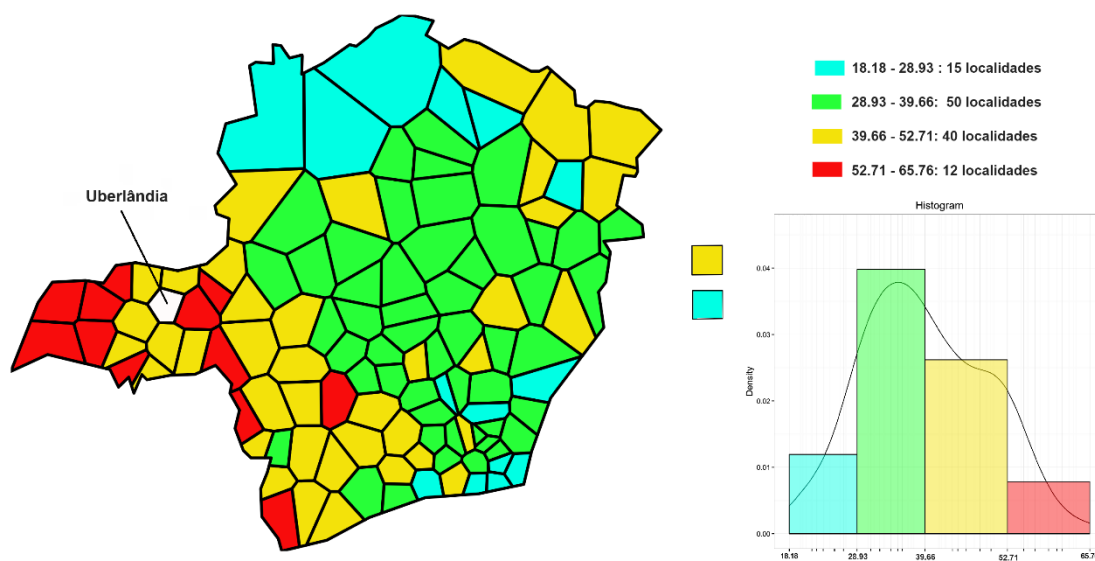
5.2 Mapas Sinópticos

Para manter a coerência da análise, a visualização desses índices nos mapas é sempre feita em quatro escalões, portanto, quatro cores. As cores quentes (vermelho e amarelo) representam as áreas com um grau de similaridade acima da média em relação ao ponto selecionado e, simetricamente, as cores frias representam (verde e azul) os domínios com similaridade abaixo da média em relação ao mesmo ponto. Após a análise da distribuição de similaridade das três cidades representativas de cada falar, é feita a análise das distribuições do PPB e do PPE.

5.2.1 Distribuição de similaridade de Uberlândia (Falar Paulista)

A localidade escolhida para representação do falar paulista é Uberlândia. Essa cidade está localizada no Triângulo Mineiro e é de grande importância econômica e social para a região, constituindo um dos principais núcleos urbanos do estado, com ligação direta às cidades paulistas e goianas, por exemplo.

Imagem 11 – Carta de similaridade de Uberlândia (falar paulista)



Fonte: dados da pesquisa.

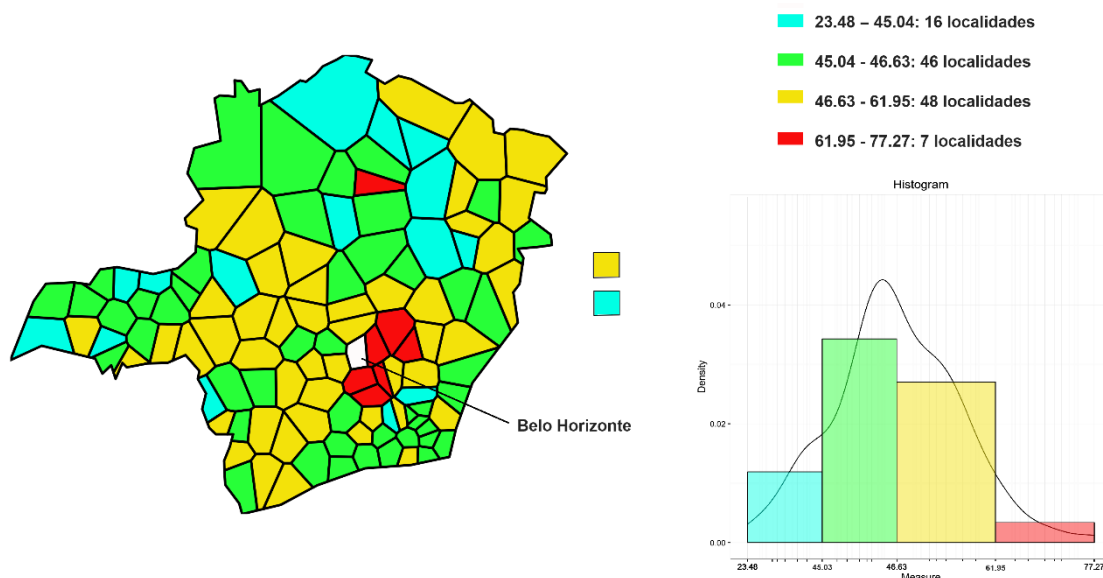
A carta de similaridade revela que as localidades com cores quentes estão mais próximas linguisticamente de Uberlândia. Em vermelho, encontram-se 12 localidades mais similares a Uberlândia e distribuídas, majoritariamente, pelo Triângulo Mineiro e por três localidades no sul de Minas Gerais: São Sebastião do Paraíso, Pouso Alegre e Formiga.

No outro escalão acima da média, em amarelo, encontram-se 40 localidades, situadas, predominantemente, no sul de Minas Gerais e no Triângulo Mineiro, mas com ocorrência também no norte do estado, incluindo neste grupo o Português Padrão Brasileiro. Por fim, a maioria dos pontos linguísticos do território mineiro são mais distantes linguisticamente de Uberlândia, representadas pelas cores verde e azul, com destaque para os pontos mais díspares localizados na fronteira com os estados do Rio de Janeiro e Espírito Santo, bem como no extremo norte na fronteira política com o sudoeste baiano, incluindo, neste grupo, o PPE.

5.2.2 Distribuição de similaridade de Belo Horizonte (falar mineiro)

A capital mineira localiza-se no centro do estado e concentra mais de 2 milhões de habitantes. É a terceira capital mais populosa da região Sudeste do Brasil e a sexta mais populosa de todo o país, tendo grande importância social, histórica e econômica tanto a nível regional como nacional. Nas imediações de Belo Horizonte, há cidades históricas, como Ouro Preto e Mariana, erigidas durante o período colonial. A Figura 12 traz a distribuição de similaridade entre as localidades do EALMG e a capital.

Imagem 12 – Carta de similaridade de Belo Horizonte (falar mineiro)



Fonte: dados da pesquisa.

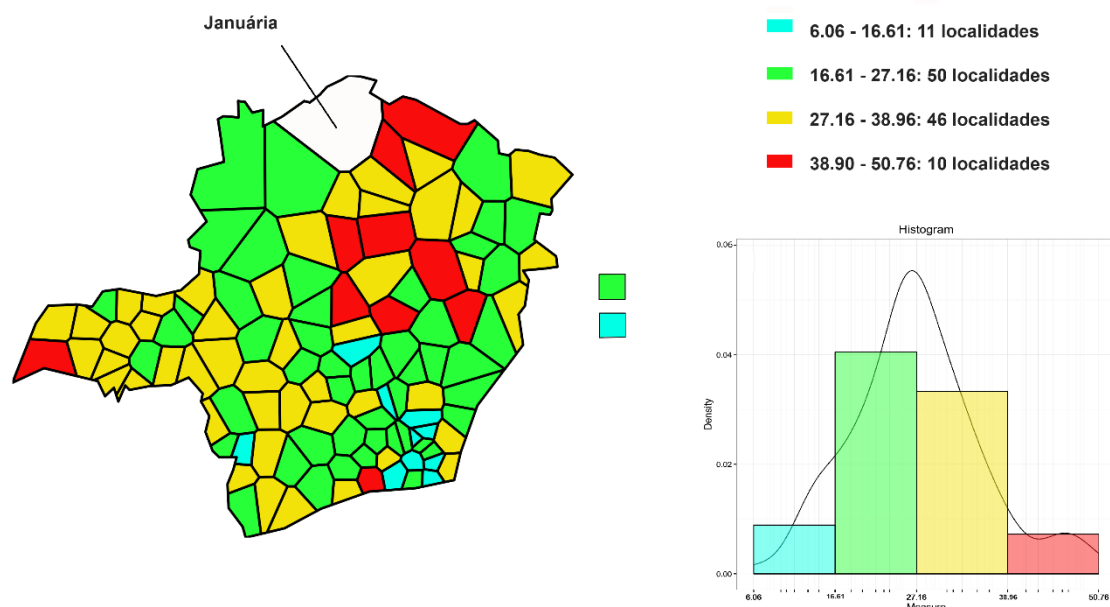
O mapa evidencia sete cidades no escalão mais alto de semelhança linguística com a capital mineira (vermelho), sendo seis delas localizadas em áreas vizinhas: Conselheiro Lafaiete, Congonhas, Sabará, Itabira, Ouro Preto e Nova Era, além de uma localidade no norte do estado (Montes Claros). Em amarelo, ainda similares a Belo Horizonte, seguem 48 pontos

linguísticos, localizados no centro-sul do estado, mas também no nordeste mineiro, incluindo o PPB. Quanto aos pontos menos similares, encontra-se a distribuição das cores mais frias por todo o estado, com diferenças no Triângulo Mineiro, centro-norte do estado e região da Zona da Mata. Observa-se no mapa que os pontos na cor azul estão espalhados, principalmente, no Triângulo Mineiro e Noroeste de Minas, incluindo nesse grupo o PPE.

5.2.3 Distribuição de similaridade de Januária (falar baiano)

Às margens do Rio São Francisco, Januária, cidade de médio porte de aproximadamente 65 mil habitantes, segundo o IBGE, está localizada no Norte do Estado, a quase 600 km de distância da capital mineira, e enquadra-se na área do falar baiano.

Imagem 13 – Carta de similaridade de Januária (falar baiano)



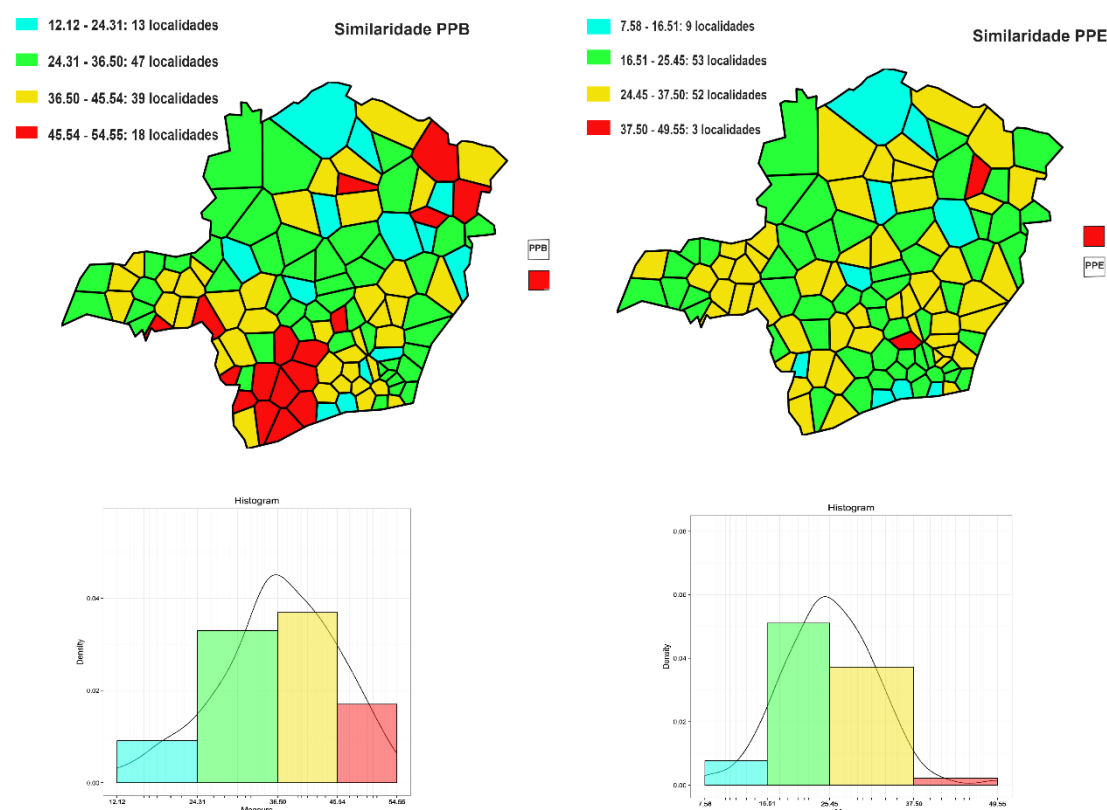
Fonte: dados da pesquisa.

No mapa, observam-se 10 localidades que mais se aproximam linguisticamente deste ponto (em vermelho), coincidentes e a distribuição diatópica mais concentrada no centro-norte de Minas Gerais (8 pontos), um ponto no extremo do Triângulo Mineiro (Iturama) e um ponto na região de fronteira com o estado do Rio de Janeiro (Olaria). Há ainda pontos mais semelhantes a Januária (em amarelo) por todo o Triângulo Mineiro, em algumas localidades do sul do estado e em localidades no norte de Minas Gerais. Quanto às regiões menos similares, portanto, mais distantes linguisticamente (escalão azul), encontram-se 11 localidades: Juiz de Fora, Além Paraíba, Rodeiro, Leopoldina, Visconde do Rio Branco, Viçosa, Ouro Preto, Sete Lagoas e Muzambinho e o PPE. Em verde, está a maioria das localidades menos semelhantes a Januária, nas diferentes regiões do estado.

5.2.4 Distribuição de similaridade de PPB e PPE

Com a finalidade de verificar o grau de relação de todos os pontos linguísticos de Minas Gerais com o Português Padrão Brasileiro e o Português Padrão Europeu¹⁰, a exemplo dos trabalhos de Brissos e Saramago (2019) e Brissos (2021), apresenta-se a Imagem14.

Imagem 14 – Cartas de similaridade - PPB e PPE



Fonte: dados da pesquisa.

No mapa da esquerda, observa-se que o PPB tem pontos mais próximos linguisticamente, sobretudo, no Sul de Minas Gerais e no Triângulo Mineiro, além de localidades no norte do estado, a saber: Montes Claros, Medina, Águas Formosas e Ladainha, inclusive o PPE tem uma grande similaridade com o PPB. Os pontos menos similares com o PPB, por sua vez, apresentam-se espalhados por todo o estado (azul), de forma descontínua e em menor número de localidades.

Por outro lado, ao observar os dados da carta à direita, verifica-se que há apenas duas localidades com alto grau de similaridade com o PPE: Araçuaí ao norte e Conselheiro Lafaiete, além do próprio PPB, o que confirma que tanto PPB quanto PPE são próximos linguisticamente, embora a rede de pontos apresente localidades mais próximas apenas do PPB. Esses resultados são análogos ao que se observa nas pesquisas de Brissos (2021) e de Brissos e

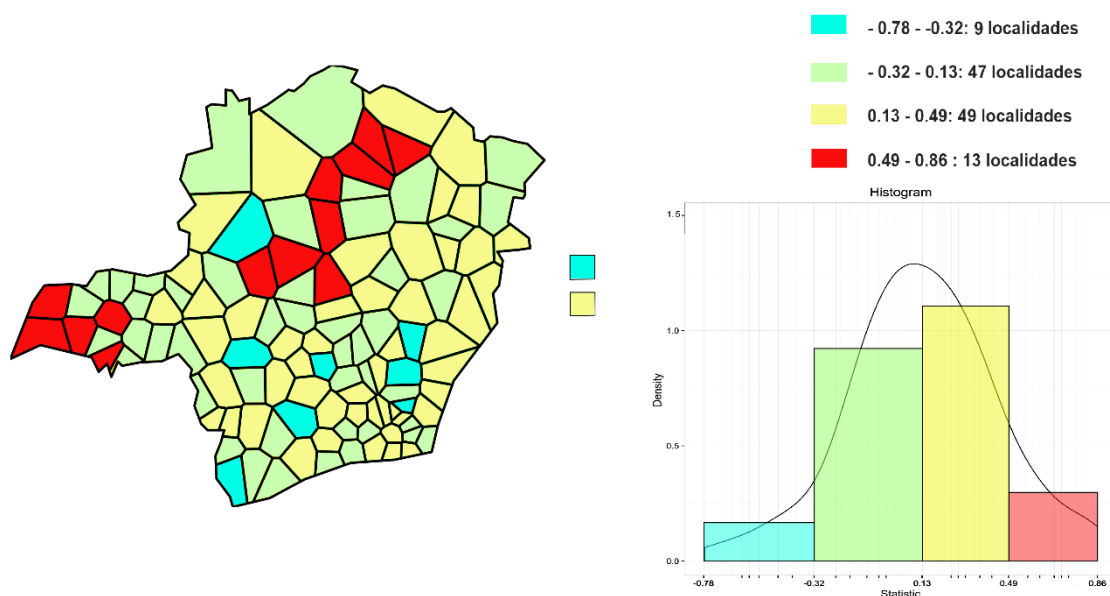
¹⁰ Por PPB e PPE, entende-se formas mais gerais conforme dois dicionários Houaiss (2001) e Porto (2011), ou seja, formas que não apresentam marcas de usos regionais, por exemplo.

Saramago (2019) em que constatarem uma elevada semelhança linguística entre PPE e PPB, por um lado, e, por outro – como seria previsível –, uma maior integração linguística do PPB no conjunto dos dados do que o PPE.

5.2.5 Distribuição de Assimetria (coeficiente de assimetria de Fisher)

A distribuição de assimetria pelo coeficiente de Fisher revela o grau de integração de cada ponto de inquérito ao conjunto dos dados, conforme a Figura 15.

Imagem 15 – Distribuição de Assimetria (Coeficiente de Fisher)

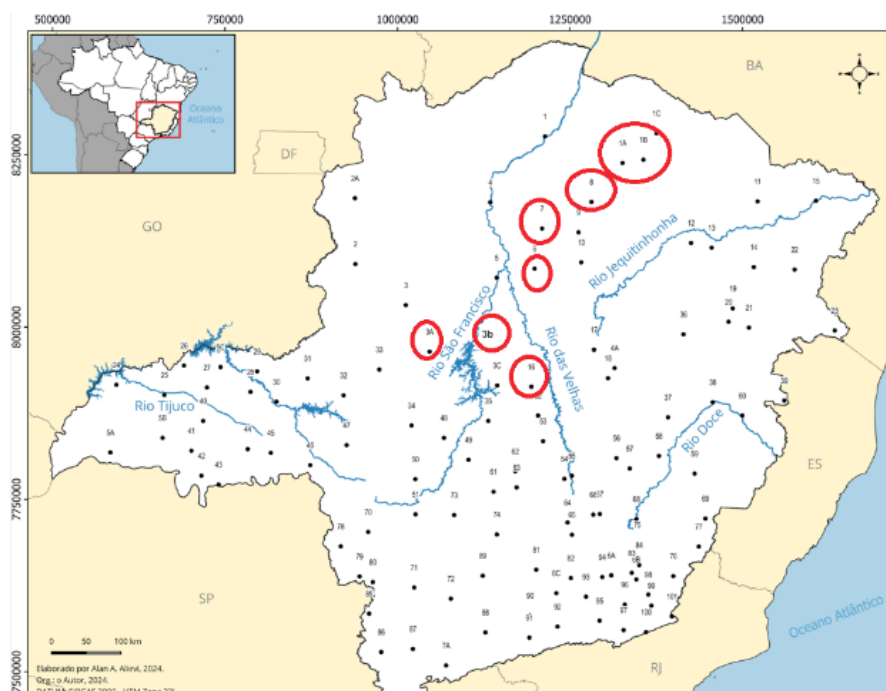


Fonte: dados da pesquisa.

Na carta de assimetria, observa-se que as cores quentes revelam menor integração ao conjunto, ou seja, são áreas que mais se diferenciam, se destacam linguisticamente – que têm maiores níveis de assimetria. Neste caso, compreende em vermelho as 13 localidades de menor integração às outras regiões do estado.

A interpretação dessa distribuição pode ser correlacionada a fatores que ratificam a tripartição de Zágari (2005) no que se refere ao falar baiano e ao falar paulista, mesmo que de forma parcial. Ao observar o mapa com os principais rios do estado (Figura 16), pode-se constatar que sete das localidades mais assimétricas ao conjunto acompanham o caminho do Rio São Francisco em direção à Bahia (São Gonçalo do Abaeté, Três Marias, Curvelo, Jequiá, Coração de Jesus, Capitão Eneias, Janaúba e Porteirinha). Essa região assimétrica se diferencia do conjunto, então, por se caracterizar como pontos pertencentes ao falar baiano: note-se que todas elas pertencem ao *cluster* C da Imagem 6, que representa precisamente esse falar.

Imagem 16 – Rede de pontos do EALMG e a hidrografia



Fonte: dados da pesquisa.

Por outro lado, o extremo do Triângulo Mineiro apresenta mais cinco localidades que se destacam do conjunto, portanto, são assimétricas ao restante da rede de pontos: Frutal, Campina Verde, Iturama, Santa Vitória e Prata. Neste caso, o destaque é para o falar paulista. Vale lembrar que os limites e as áreas de ocorrência de ambos os falares são fluidos e não estáveis, pois há grande número de localidades representadas também pela cor amarela que revelam certo grau de não integração ao conjunto dos dados.

Ressalte-se ainda que as demais localidades mineiras, no conjunto, não apresentam tantas diferenças, portanto, estão mais integradas. Do ponto de vista da tripartição do território, pode-se afirmar que parcialmente a proposta de Zágari (2005) pode ser validada dialetometricamente, em relação ao falar baiano que se projeta no território de Minas Gerais acompanhando o caminho do Rio São Francisco até Curvelo (no centro do estado), e o falar paulista que se destaca ainda mais no Triângulo Mineiro em relação ao Sul de Minas Gerais. O restante das localidades, por fim, englobaria o falar mineiro numa área mais expandida em relação à classificação de Zágari.

6 Considerações Finais

A dialetometria, enquanto método de análise de *corpora* geolinguísticos, permite com exatidão quantificar dados que ultrapassam a interpretação por vezes impressionista na delimitação de isoglossas e definição de áreas dialetais. O método, ainda com relativamente pouca tradição na generalidade das línguas, começa a criar um corpo significativo de estudos no português, tanto europeu quanto brasileiro.

Considerando a profícua atividade da Geolinguística brasileira no início do século XXI, Romano (2020) arrola uma centena de estudos que apresentam a diversidade diatópica de um país-continente com dados suficientes para análises quantitativas ainda a serem feitas. O presente trabalho contribui para o desenvolvimento da dialetometria no Brasil e está inserido em programa maior sobre dialetometrização de atlas linguísticos brasileiros. Trata-se de um estudo que traz resultados parciais do projeto de pós-doutoramento no Centro de Linguística da Universidade de Lisboa, vinculado à equipe do Atlas Linguístico do Brasil (ALiB) e à equipe do Atlas Linguístico-Etnográfico Português (ALEPor).

O trabalho dá continuidade às pesquisas pioneiras de Saramago (1986), no caso português, e de Altino (2007), no caso brasileiro, bem como outros estudos que só recentemente começam a despontar, como o de Brissos e Saramago (2019) e Brissos (2021), principalmente porque, a partir de 2011, surgiram ferramentas computacionais (Aurrekoetxea et al. 2013) que permitem o tratamento de dados de atlas brasileiros, considerando-se as suas especificidades, pois são pluridimensionais por natureza (Thun, 1998). O *software* DiaTech (Aurrekoetxea et al. 2013) foi a ferramenta utilizada para este estudo, amplamente testada já em ambiente de língua portuguesa. O estudo trouxe resultados a partir de três parâmetros dialetométricos para verificar as áreas lexicais no território de Minas Gerais, com base em cartas linguísticas do *Esboço de Atlas Linguístico de Minas Gerais* (Ribeiro et al. 1977).

Considerando a importância do EALMG para a Geolinguística brasileira, parte-se da proposta de tripartição do território feita por Zágari (1998, 2005): o falar paulista, o falar mineiro e o falar baiano. Com base na análise de *clusters* e nos mapas sinópticos de similaridade e assimetria, pode-se confirmar a existência dessas áreas cujos limites são fluidos.

Pela distribuição de assimetria, observa-se que grande parte da rede de pontos está integrada ao conjunto, na região central de Minas Gerais, avançando para o noroeste e nordeste do estado. Essa distribuição pode ser considerada a área correspondente ao falar mineiro que, na análise dialetométrica, compreende uma área maior do que a proposta por Zágari (2005). Ainda sobre a assimetria, duas áreas estão menos integradas ao conjunto: uma que segue o curso do rio São Francisco, a partir de Curvelo (ponto 16) em direção a Janaúba (norte de Minas) e fronteira com a Bahia. Esta pode ser considerada a área do falar baiano, cuja área de ocorrência é menor em relação à que Zágari propõe; mas ainda faz certo sentido. Acrescenta-se que o extremo do Triângulo Mineiro também está menos integrado ao conjunto, caracterizado como falar paulista.

Conforme se observou neste estudo, os limites dos falares de Zágari (2005) não seguem estritamente os limites apresentados nesta pesquisa, uma vez que a área do falar mineiro aparenta estar mais ampla do que a delimitada pelo dialetólogo. Cabe aqui, porém, a ressalva de que os resultados são sobre dados de natureza lexical, que, por si mesmos, são desafiadores, uma vez que o léxico é um nível de análise da língua particularmente difícil de sintetizar em áreas dialetais bem definidas (Ribeiro, 2012; Romano, 2015).

Quanto às cartas de similaridade, tomando-se três pontos de inquérito que se localizam em cada um dos falares: Uberlândia (paulista), Januária (baiano) e Belo Horizonte (mineiro), novamente se observa a coerência de Zágari (2005), porém com ressalvas, pois há pontos que não estão necessariamente na mesma região, ou seja, os limites não são estáticos, mas sim fluidos, ora avançando ora retraindo de um falar para outro.

Em relação ao PPE e ao PPB, observou-se que a rede de pontos é mais próxima do segundo do que do primeiro, como já se esperava. Por fim, a análise de *clusters* permitiu

concluir que as localidades do falar baiano estão mais homogêneas em relação às localidades do falar paulista e mineiro que se subdividem em subgrupos formando *best friends*, não necessariamente vizinhos geograficamente.

Com este texto foi possível terminar uma tarefa planejada por um dos primeiros geolinguistas brasileiros, o professor Mário Roberto Lobuglio Zágari, como forma de homenagem e reconhecimento pelo seu árduo trabalho no estudo e descrição da língua falada em Minas Gerais, ao menos em relação ao léxico. Este artigo, portanto, realça a importância e o legado que o EALMG deixou para os estudiosos da língua portuguesa, em especial para os que se ocupam da diversidade linguística mineira, e revela que as sementes plantadas na década de 1970 produziram frutos.

A dialetometria na Geolinguística brasileira é o próximo passo para a nova geração de dialetólogos como forma de gerir o grande volume de dados já coletados e divulgados nos atlas. O estudo conjugou o tradicional e o moderno, o velho e o novo na pesquisa geolinguística dando indícios para a necessidade de um novo atlas linguístico do estado de Minas Gerais, cujos dados poderão ratificar ou retificar os resultados que a análise dialetométrica do EALMG sintetizou.

Declaração de autoria

Valter Pereira Romano organizou a base de dados e escreveu fundamentação teórica-metodológica e descrição dos resultados. Fernando Brissos auxiliou na interação com a ferramenta DiaTech, na interpretação dos resultados e fez a revisão final do manuscrito.

Agradecimentos

Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, pois esta pesquisa fez parte do projeto de pós-doutorado do primeiro autor do artigo e foi financiada com bolsa de Professor Visitante Júnior (PVJ) CAPES/Print (Programa Institucional de Internacionalização, Edital n. 41/2017, Processo: 88887.936562/2024-00), e desenvolvida no Centro de Linguística da Universidade de Lisboa, no período de 01/05/2024 a 31/10/2024. Por ter sido levado a cabo no quadro das atividades de investigação do Centro de Linguística da Universidade de Lisboa, este artigo também foi financiado por fundos portugueses através da FCT – Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UIDB/00214/2020.

Referências

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

ALTENHOFEN, C. V.; KLASSMANN, M. S. (Orgs.). *Atlas Linguístico-Etnográfico da Região Sul do Brasil: cartas semântico-lexicais*. Porto Alegre: Editora UFRGS; Florianópolis: Editora UFSC, 2011.

ALTINO, F. C. *Atlas Linguístico do Paraná II*. 2007. 2 v. 183 f. Tese (Doutorado em Estudos da Linguagem) — Centro de Letras e Ciências Humanas, Universidade Estadual de Londrina, Londrina, 2007.

AUGUSTO, V. *Variação lexical goiana*: Estudo dialetométrico do Atlas Semântico-Lexical do Estado de Goiás — Brasil. Relatório inédito de pós-doutoramento, Universidade de Lisboa / Universidade de São Paulo, 2018.

AURREKOETXEA, G. SANTANDER, G. USOBIAGA, I.; IGLESIAS, A. Diatech: tool for making dialectometry easier. *Dialectologia*, Barcelona, v.1, n.17 p. 1-22, 2016. Disponível em: <https://raco.cat/index.php/Dialectologia/article/view/312056>. Acesso: 18 nov. 2025.

AURREKOETXEA, G.; FERNANDEZ-AGUIRRE, K.; RUBIO, J.; RUIZ, B.; SANCHEZ, J. DiaTech: A new tool for dialectology. *Literary and Linguistic Computing*, New York, v. 28, n.1, p. 23-30, 2013. DOI: <https://doi.org/10.1093/lc/fqs049>.

BRISSOS, F. Dialectos portugueses do centro-sul: corpus de fenómenos e revisão do problema da (des)unidade. *Zeitschrift für romanische Philologie*, Berlim, v. 131, n.4, p. 999-1041, 2015. DOI: <https://doi.org/10.1515/zrp-2015-0071>

BRISSOS, F. Análise dialetométrica do Atlas Linguístico do Amazonas: variação lexical. *Revista Internacional de Linguística Iberoamericana*, Madrid, v. 19 n.37p.167-207, 2021. Disponível em: <https://www.jstor.org/stable/48637006> . Acesso em: 18 nov. 2025.

BRISSOS, F. Portugal: a cidade e o interior. I — Centro-sul. *Limite*: Revista de Estudios Portugueses y de la Lusofonía, Cáceres, v. 10. n.1 p. 85–107, 2016. Disponível em: <https://revista-limite.unex.es/index.php/limite/article/view/1556/1519>. Acesso: 18 de nov. de 2025.

BRISSOS, F. *Digestão dialetométrica de atlas linguísticos do português: problemas, soluções e resultados atualizados*. Conferência apresentada ao evento ABRALIN ao Vivo 2020. Associação Brasileira de Linguística. Disponível em: <https://aovivo.abralin.org/lives/fernando-brissos/>

BRISSOS, F. Estudos de dialetometria brasileira: problemas, soluções e estado da questão. In: Cristianini, A. et al. (orgs.). *Geolinguística brasileira numa abordagem dialetométrica*: Teoria e prática. Campinas: Pontes. [no prelo]

BRISSOS, F. ; GILLIER, R. ; SARAMAGO, J. As variedades açorianas no sistema dialetal português: síntese atualizada. In. VII COLOQUIO - O FAIAL, 2019, Horta..Atas... (Separata), Horta: Núcleo Cultural da Horta,.D. L 2019, p. 557-176.

BRISSOS, F.; GILLIER, R.; SARAMAGO, J. O problema da subdivisão da variedade dialetal madeirense: estudo dialetométrico da variação lexical. *Revista da Associação Portuguesa de Linguística*, Porto, v.1.. n.2, , p. 31–47, 2016. DOI: 10.26334/2183-9077/rapln2ano2016a2

BRISSOS, F.; GILLIER, R.; SARAMAGO, J. Variação lexical açoriana: estudo dialetométrico do Atlas Linguístico-Etnográfico dos Açores. *Revista Galega de Filoloxía: MonoGrafía 11* -.Aproximacións á variación lexical no dominio galego-portugués, A Coruña , p.11-28, 2017.

BRISSOS, F.; SARAMAGO, J. Análise dialetométrica do Atlas Linguístico Etnográfico da Região Sul do Brasil: variação lexical. In.: Carrilo, E.; Martins, A. M.; Pereira, S.; Silvestre, J. P. (eds.). *Estudos linguísticos e filológicos oferecidos a Ivo Castro*. Lisboa: Centro de Linguística da Universidade de Lisboa. 2019. p. 349-379. Disponível online em <http://hdl.handle.net/10451/39619>.

- CARDOSO, S. A. M. *Geolinguística: tradição e Modernidade*. São Paulo: Parábola, 2010.
- CARDOSO, S. A. M. *Atlas Linguístico de Sergipe II*. Salvador: EUFBA, 2005.
- CRISTIANINI, A. C. Atlas Semântico-lexical da região do Grande ABC. 2007. 3. v. 772 f. Tese (Doutorado em Linguística) — Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2007.
- CRISTIANINI, A. C. Estudo dialetométrico do Atlas Semântico-lexical da Região do Grande ABC. *Signum: Estudos da Linguagem*, Londrina, v.26, n.3, p.78-101, 2023. DOI: 10.5433/2237-4876.2023v26n3p78-101
- CRUZ, M. L.de C. *Atlas Linguístico do Amazonas – ALAM*, 2 v. 400 f. Tese (Doutorado em Letras Vernáculas) — Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.
- FERREIRA, C. *et al.* *Atlas linguístico de Sergipe*. Salvador: Universidade Federal da Bahia; Fundação de Cultura de Sergipe, 1987.
- GOEBL, H. Eléments d'analyse dialectométrique avec application à l'AIS, *Revue de Linguistique Romane*, Strasbourg, v.45, n.179, p. 349-420, 1981. DOI : <http://doi.org/10.5169/seals-399711>.
- GOEBL, H. La dialectométrie appliquée à l'ALF (Normandie). In: CONGRESSO INTERNAZIONALE DI LINGUISTICA E FILOGIA ROMANZA, 14., 1974, Nápoles. *Atti*. Nápoles; Amsterdã, v. II, p. 165-195, 1976.
- GOEBL, H. VDM—visual dialectometry. Vorstellung eines dialektometrischen Software-Pakets auf CD-ROM (mit Beispielen zu ALF und Dees 1980). *Romanistik und neue Medien*, ed.W Dahmen, GHoltus, J Kramer, M Metzeltin, W Schweickard, O Winkelmann, Tübingen, 2004., pp. 209–41.
- GOEBL, H. Introduction aux problèmes et méthodes de l'“École dialectométrique de Salzbourg” (avec des exemples gallo-, italo- et ibéro-romans). In: ALVAREZ PEREZ, X.; CARRILHO, E.; MAGRO, C. (Eds.). *Proceedings of the International Symposium on Limits and Areas in Dialectology (LimiAr)*, 2011. Lisboa: CLUL, 2012. p. 117-166.
- HAIMERL, E. Database design and technical solutions for the management, calculation, and visualization of dialect mass data. *Literary and Linguistic Computing*, Oxford, v.21 n.4, p. 437-444, 2006. DOI: <https://doi.org/10.1093/lc/fql037>
- HOUAISS, A.; VILLAR, M. de S.; FRANCO, F. M. de M. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001.
- MARTINS, E. F. Atlas lingüístico do Estado de Minas Gerais: o princípio da uniformidade da mudança lingüística nas características fonéticas do português mineiro. *Revista Virtual de Estudos da Linguagem*, v. 4, n. 7, p. 1-13, 2006. Disponível em: http://www.revel.inf.br/files/artigos/revel_7_atlas_linguistico_do_estado_de_minas_gerais.pdf. Acesso: 20 nov. 2025.
- MENDONCA, L. A. L. ; ROMANO, V. P. . Contribuições do Esboço de um Atlas Linguístico de Minas Gerais para o ensino de língua portuguesa na educação básica. *Sociodialeto*, v. 10, n.30, p. 180-200, 2020. Disponível em: <https://periodicosonline.uems.br/sociodialeto/article/view/8034> Acesso em: 20 nov. 2025.

NERBONNE, J., R. COLÉN, R. GOOSKENS, C.; KLEIWEG, P.; LEINONEN, T. Gabmap: a web: Aplicativo para dialetologia. *Dialectologia*, Barcelona, Special Issue, p.65-89, 2011. Disponível em: <file:///C:/Users/valte/Downloads/ub,+dialectol+Therese+LEINONEN.pdf>. Acesso: 20 nov. 2025.

PORTO EDITORA. *Dicionário da Língua Portuguesa*. Porto: Porto Editora, 2011.

RIBEIRO, J. et al. *Esboço de um atlas linguístico de Minas Gerais*. Rio de Janeiro: Casa de Rui Barbosa, 1977.

RIBEIRO, S. S. C. *Brinquedos e brincadeiras infantis na área do falar baiano*. 2012, 466 p. Tese (Doutorado em Língua e Cultura) — Instituto de Letras, Universidade Federal de Bahia, Salvador, 2012.

ROCHA, A. P. A. Notas sobre o léxico de brincadeiras infantis usado em Minas Gerais à Luz de dois trabalhos geolinguísticos: o ALEMIG (1977) e o Projeto ALiB. In: ALTINO, F. C. (Org.). *Múltiplos olhares sobre a diversidade linguística: uma homenagem à Vanderci de Andrade Aguilera*. Londrina: Midiograf, 2012. p. 79-92.

ROCHA, A. P. A.; ANTUNES, L. B. Divisão dialetal em Minas Gerais: notas sobre aspectos fonéticos. In: RAZKY, A.; LIMA, A. F. de.; OLIVEIRA, M. B.; COSTA, E. O. da. (orgs.). *Estudos sociodialetais do português brasileiro*. Campinas: Pontes, 2014. p. 97-111.

ROCHA, A. P. A.; RAMOS, J. M. Estudos dialetais em Minas Gerais. *Estudos linguísticos e literários*, Salvador, v.1, n. 41, p. 70-86, jan./jun. 2010. Disponível em: <https://periodicos.ufba.br/index.php/estudos/issue/view/1094/10>. Acesso: 20 nov. 2025.

ROMANO, V. P. Desdobramentos, desafios e perspectivas da Geolinguística Pluridimensional no Brasil. In: MOTA, J. A.; OLIVEIRA, J. M.; PAIM, M. M. T.; RIBEIRO, S. S. C. (Org.). *Contribuições de estudos geolinguísticos para o Português Brasileiro: uma homenagem a Suzana Cardoso*. 1ed. Salvador: EDUFBA, 2020, v. 1, p. 11-39. Disponível em: <https://edufba.ufba.br/livros-publicados/catalogo/contribuicoes-de-estudos-geolinguisticos-para-o-portugues-brasileiro-uma>. Acesso: 20 nov. 2025.

ROMANO, V. P. *Em busca de áreas lexicais no centro-sul do Brasil*. 2015. 2v. 367 f. Tese (Doutorado em Estudos da Linguagem) — Centro de Letras e Ciências Humanas, Universidade Estadual de Londrina, Londrina, 2015.

ROMANO, V. P.; BRISSOS, F. A dialetometria no Brasil. In: Romano, V. P.; Altino, F. C.; Aguilera, V. de A. (Org.). *A Geolinguística no Brasil: contatos, interfaces e entremeios*. São Carlos: Pedro & João Editores, 2025. p. 95-119.

ROMANO, V. P.; CRUZ, J. A. Entre raios e coriscos: estudo geolinguístico em Minas Gerais nos dados do ALiB e do EALMG. *Sociodialetos*, v. 10, n.30, p. 274-300, 2020. Disponível em: <https://periodicosonline.uems.br/index.php/sociodialetos/article/view/8003>. Acesso: 20 nov. 2025.

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

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

SARAMAGO, J. Differentiation lexicale (un essai dialectométrique appliqué aux matériaux portugais de l'A.L.E.)». *Géolinguistique* II, Grenoble, v.1, n.2, p. 1-31, 1986. Disponível em: https://www.persee.fr/doc/geol_0761-9081_1986_num_2_1_1058?utm. Acesso: 20 nov. 2025.

SARAMAGO, J.; CARDOSO, S. A. M. Atlas Linguístico de Sergipe: história, metodologia e análise dialetométrica. *Estudos Linguísticos e Literários*, Salvador, v.1, n. 41, p. 121-158, 2010. Disponível em: <https://periodicos.ufba.br/index.php/estudos/issue/view/1094/10>. Acesso: 20 nov. 2025.

SÉGUY, J. “La dialectométrie dans l’Atlas Linguistique de la Gascogne”, *Revue de Linguística Romane*, Zurique, v.37, n.1, p.1-24, 1973.

SOARES, R. de C. da S. *Atlas Semântico-Lexical da Região Norte do Alto Tietê/SP numa perspectiva dialetométrica*. Relatório inédito de pós-doutoramento, Universidade de Lisboa / Universidade de São Paulo, 2019.

SILVA, G. A. ; ROMANO, V. P. (Orgs.) . *Tendências da Geolinguística brasileira e a nova geração e atlas linguísticos*. São Carlos: Pedro & João Editores, 2022.

SILVA NETO, S. da. *Guia para estudos dialectológicos*. 2. ed. Belém: Instituto Nacional de Pesquisas da Amazônia, 1957.

TAKANO, Y. *Atlas Linguístico Semântico-Lexical do Falar Nipo-Brasileiro do Distrito Federal: Perspectiva dialetométrica*. Relatório inédito de pós-doutoramento, Universidade de Lisboa / Universidade de São Paulo, 2018.

THUN, H. La géographie linguistique romane à la fin du XX siècle. In.: XXII CONGRÈS INTERNATIONAL DE LINGUISTIQUE E PHILOGIE ROMANES. *Actes...* Bruxelles, Max Niemeyer Verlag, 1998, 367-409.

VITORINO, G. *Atlas linguístico do litoral português: flora e fauna*. 1987. Lisboa: Centro de Linguística da Universidade de Lisboa; Instituto Nacional de Investigação Científica 1987.

WIELING, M.; NERBONNE, J. Advances in dialectometry. *The Annual Review of Linguistics*, San Mateo, v.1, n.1, p.243-267, 2015. DOI: <https://doi.org/10.1146/annurev-linguist-030514-124930>.

ZÁGARI, J. R. L. Os falares mineiros: esboço de um atlas linguístico de Minas Gerais. In: AGUILERA, V. de A. (org.) . *A geolinguística no Brasil: trilhas seguidas, caminhos a percorrer*. Londrina: EDUEL, 2005 [1998]. p. 46-72.