OPEN ACCESS

ESCOLA DE HUMANIDADES
PUCRS

LETRAS DE HOJE
Studies and debates in linguistics, literature and Portuguese language

Letras de hoje Porto Alegre, v. 61, n. 1, p. 1-14, jan.-dez. 2026
e-ISSN: 1984-7726 | ISSN-L: 0101-3335

ARTIGOS

Problemas geolinguísticos: cartografia de variantes lexicais e de respostas suspeitas

Geolinguistic problems: mapping lexical variants and suspicious responses

Problemas geolingüísticos: mapeo de variantes léxicas y respuestas sospechosas

Romário Duarte Sanches[1]
orcid.org/0000-0002-0571-303X
romario.duarte@unifap.br

Greize Alves da Silva[2]
orcid.org/0000-0002-2589-6750
greize_silva@uft.edu.br

Recebido em: 30 jun. 2025.
Aprovado em: 05 maio 2026.
Publicado em: 27 out. 2026.

Resumo: O presente artigo aborda de forma sistemática e pontual os desafios metodológicos enfrentados em estudos geolinguísticos, com foco no mapeamento linguístico e na validação de variantes lexicais coletadas *in loco*. A partir da constatação de que determinados atlas linguísticos apresentam respostas que não correspondem efetivamente a variantes linguísticas em uso, propomos aqui quatro critérios-base para garantir maior rigor na seleção e representação dessas variantes: (i) frequência lexical; (ii) estabilidade do signo linguístico e triangulação de dados; (iii) validação comunicativa (*members checking*); e (iv) revisão por informantes-chave (*external validation*). O artigo discute a importância da frequência de uso como indicador de variantes lexicais mais frequentes e menos frequentes, mas também alerta para os limites dessa métrica quando utilizada isoladamente. Já o processo de triangulação é realizado por meio da consulta a dicionários e atlas linguísticos já publicados, garantindo a verificação semântica e referencial dos itens coletados. A validação comunicativa, por sua vez, reconhece como parte integrante do processo de coleta a participação ativa dos informantes na validação dos dados, reforçando a confiabilidade dos resultados da pesquisa. Por fim, a validação externa é apresentada como uma instância complexa de se aplicar, mas não impossível, sendo fundamental para consolidar a credibilidade científica dos atlas linguísticos. A adoção desses critérios auxilia na elaboração de mapas linguísticos mais fidedignos à realidade dialetológica, evitando a inclusão de respostas suspeitas e não condizentes com o referente. Contudo, ressaltamos que esses quatro critérios arrolados não constituem um conjunto fechado, sendo possível a incorporação de novos ou a reformulação daqueles já existentes.

Palavras-chave: variantes lexicais; validação de dados; cartografia linguística.

Abstract: This article systematically and precisely addresses the methodological challenges faced in geolinguistic studies, focusing on linguistic mapping and the validation of lexical variants collected in situ. Based on the observation that certain linguistic atlases present responses that do not effectively correspond to actual linguistic variants in use, we propose four core criteria to ensure greater rigor in the selection and representation of such variants: (i) lexical frequency; (ii) stability of the linguistic sign and data triangulation; (iii) communicative validation (members checking); and (iv) review by key informants (external validation). The article discusses the relevance of usage frequency as an indicator to distinguish between more frequent and less frequent lexical variants, while also warning about the limitations of this metric when used in isolation. The triangulation process is carried out through consultation of dictionaries and previously published linguistic atlases, ensuring the semantic and referential verification of collected items. Communicative validation recognizes the active participation of informants as an integral part of the data collection process, strengthening the reliability of the research results. Lastly, external validation is presented as a complex, yet feasible instance, crucial for consolidating the scientific credibility of linguistic atlases. The adoption of these criteria supports the creation of linguistic maps that are more faithful to dialectological reality, avoiding the inclusion of suspicious or

Este é um artigo de acesso aberto distribuído sob a licença CC-BY 4.0, que permite a cópia e redistribuição do material em qualquer formato e para qualquer finalidade, desde que a autoria original e os créditos de publicação sejam mantidos.

[1] Universidade Federal do Amapá (Unifap), Santana, Amapá, Brasil.
[2] Universidade Federal do Tocantins (UFT), Porto Nacional, Amapá, Brasil.

non-representative responses. However, we emphasize that these four criteria do not form a closed set, allowing for the inclusion of new criteria or the reformulation of those already established.

Keywords: lexical variants; data validation; linguistic cartography.

Resumen: El presente artículo aborda de forma sistemática y puntual los desafíos metodológicos enfrentados en estudios geolingüísticos, con enfoque en el mapeo lingüístico y en la validación de variantes léxicas recolectadas in situ. A partir de la constatación de que ciertos atlas lingüísticos presentan respuestas que no corresponden efectivamente a variantes en uso, proponemos cuatro criterios base para garantizar mayor rigor en la selección y representación de dichas variantes: (i) frecuencia léxica; (ii) estabilidad del signo lingüístico y triangulación de datos; (iii) validación comunicativa (*members checking*); y (iv) revisión por informantes clave (*external validation*). El artículo discute la importancia de la frecuencia de uso como indicador para distinguir entre variantes léxicas más y menos frecuentes, pero también advierte sobre los límites de esta métrica cuando se utiliza de forma aislada. El proceso de triangulación se lleva a cabo mediante la consulta de diccionarios y atlas lingüísticos ya publicados, lo que garantiza la verificación semántica y referencial de los ítems recolectados. La validación comunicativa reconoce la participación activa de los informantes como parte integral del proceso de recolección, fortaleciendo la confiabilidad de los resultados. Finalmente, la validación externa se presenta como una instancia compleja de aplicar, pero no imposible, siendo fundamental para consolidar la credibilidad científica de los atlas lingüísticos. La adopción de estos criterios contribuye a la elaboración de mapas lingüísticos más fieles a la realidad dialectológica. No obstante, subrayamos que estos cuatro criterios no constituyen un conjunto cerrado, permitiendo la incorporación de nuevos o la reformulación de los ya existentes.

Palabras clave: variantes léxicas; validación de datos; cartografía lingüística.

1 Introdução

A Geolinguística, no contexto brasileiro, tem experimentado um crescimento considerável no que se refere à aplicação dos aspectos metodológicos e teóricos, impulsionado sobretudo pelo surgimento de diversos projetos de atlas linguísticos, desenvolvidos a partir de diferentes perspectivas e com distintos objetivos. Contudo, a feitura de uma obra dessa envergadura, independentemente do tamanho, envolve uma série de decisões técnicas que podem influenciar diretamente a qualidade e a representatividade dos dados coletados, impactando, assim, no produto deste trabalho, ou seja, nos mapas linguísticos de fenômenos diversos que compõem os atlas.

Dentre os desafios metodológicos da Geolinguística, destaca-se a seleção criteriosa das formas lexicais que deverão ser consideradas para compor o mapa linguístico. Esse processo suscita a seguinte questão: quais critérios os pesquisadores têm adotado para validar respostas dadas pelos informantes como uma variante lexical? Com o objetivo de responder a essa problemática, este estudo propõe um conjunto de critérios que pode orientar a seleção de formas lexicais para o processo de cartografia linguística. Cabe ressaltar que esses critérios foram desenvolvidos especificamente para a análise geolinguística de âmbito lexical e que, para outros níveis da língua, como fonético, morfossintático ou discursivo, parâmetros distintos ou similares devem ser elaborados, levando em consideração as propriedades e especificidades de cada domínio linguístico.

Com base nas especificidades da Geolinguística, propomos quatro critérios norteadores especialmente elaborados para o mapeamento do léxico em estudos dialetológicos. São eles: (i) a frequência lexical nas ocorrências coletadas; (ii) a estabilidade do signo linguístico aliada à triangulação com outros dados linguísticos; (iii) a validação comunicativa por meio do *members checking*; e (iv) a revisão por informantes-chave, visando à validação externa das formas selecionadas (*external validation*).

2 Frequência lexical

O primeiro critério a ser considerado para o mapeamento de variantes lexicais é a frequência quantitativa de respostas obtidas durante a coleta de dados. Pode-se afirmar que este critério parte dos princípios estatísticos de contagem de ocorrências, que diz respeito ao número de vezes que um determinado item ou estruturas frasais complexas (formada por mais de um vocábulo) aparecem em um *corpus* linguístico.

A formalização estatística da frequência lexical remonta à tradição matemática do século XVIII, mas sua incorporação sistemática na Linguística data, sobretudo, dos trabalhos de George K. Zipf (1935, 1949). Ao analisar textos, Zipf percebeu que a frequência das palavras seguia um padrão

específico, levando-o a desenvolver a ideia de que a linguagem humana possui uma estrutura subjacente que pode ser quantificada. Desde então, a Lei de Zipf tem sido aplicada em diversas áreas, incluindo Ciência da Computação, Sociologia e Biologia, demonstrando sua relevância em múltiplos contextos. Essa lei formula que a frequência de uma palavra é inversamente proporcional à sua posição no *ranking* de ocorrência, revelando padrões regulares mesmo em amostras linguísticas extensas[3].

Atualmente, o conceito de frequência lexical também é utilizado em outros campos da linguística, como: *Aquisição de linguagem:* palavras mais frequentes tendem a ser aprendidas primeiro por crianças e por aprendizes de segunda língua (Ellis, 2002); *Processamento linguístico:* palavras de alta frequência são reconhecidas e processadas mais rapidamente pelo cérebro (Brysbaert; New, 2009); *Lexicografia e dicionários:* a frequência ajuda a decidir quais palavras entram em um dicionário básico (Béjoint, 2010); *Ensino de línguas:* orienta a seleção de vocabulário para cursos de idiomas, priorizando palavras mais frequentes como forma de facilitar o aprendizado (Nation, 2001); e *Estudos de variação linguística:* a frequência de variantes pode indicar padrões de mudança linguística (Labov, 1972, 1994).

No contexto da Geolinguística, assume-se as mesmas premissas da Lexicografia e da Sociolinguística. Para o mapeamento linguístico, tal como para construção de dicionários, leva-se em consideração a frequência de uso das palavras que determina a representatividade de um item lexical, pois, quanto mais recorrente ele for, maior sua estabilidade e consolidação na variedade linguística estudada (Labov, 1972, 1994).

A frequência lexical pode ser classificada a partir da *frequência relativa*, isto é, o número de ocorrência de um item em relação ao tamanho total do *corpus*. Assim, os dados lexicais podem assumir uma frequência muito alta (75% a 100%), alta (50% a 75%), moderada (20% a 50%), baixa (5% a 20%) ou muito baixa (0% a 5%). A frequência relativa é calculada a partir da divisão entre a frequência absoluta de um determinado fenômeno e o total de ocorrências observadas, sendo expressa na forma decimal ou percentual.

Se considerarmos como exemplo o *corpus* do Atlas Linguístico do Brasil – ALiB, especificamente a questão de número 101 que trata das denominações para *fanhoso*, veremos que, na Região Sul, a variante lexical *fanho(a)* aparece com frequência muito alta (79%), *fanhoso(a)* com frequência baixa (9%) e *fonhém* com frequência muito baixa (4%). Essa porcentagem foi calculada em relação ao número total de 26 ocorrências fornecidas pelos informantes sulistas, sendo 19 para *fanho(a),* 02 para *fanhoso(a)*, 01 para *fonhém* e 02 para *não respostas* (Sanches; Razky, 2021, p. 22).

Aparentemente, pela frequência lexical, parece simples identificar quais respostas podem ser mapeadas como variantes lexicais. No entanto, é preciso lembrar que na pesquisa geolinguística, comumente, são utilizados como instrumentos de coleta questionários semiestruturados com perguntas abertas e fechadas. Tomando como exemplo o modelo metodológico do Projeto ALiB, tem-se o questionário fonético-fonológico, com perguntas fechadas; já no caso do questionário semântico-lexical, há as perguntas abertas. Para este último tipo de questionário, o informante pode fornecer mais de uma resposta. Por exemplo, a pergunta de número 01 do QSL é esta: "como se chama aquele rio pequeno de uns dois metros de largura?". Indagado pelo inquiridor, o informante poderá responder apenas *riacho* (se essa for a variante conhecida e utilizada por ele), mas, se ele achar necessário dar outras respostas, poderá também informar demais nomes que estão no seu repertório de conhecimento, como *córrego*, *igarapé*, *sanga* etc. (Sanches; Razky, 2016).

A partir dos dados coletados, é preciso iniciar a triagem do que será considerado como variante lexical ou não. Essa etapa é essencial para seleção de itens que irão compor o mapeamento lexical, pois, uma vez selecionadas respostas suspeitas,

---

[3] Posteriormente, John Sinclair (1991) aprimora o uso de frequência relativa para análise de *corpus* linguístico.

os mapas, em vez de representarem a realidade varietal de uma comunidade, poderão falsear essa realidade. Por isso, o primeiro critério a ser considerado deve ser a frequência lexical a partir das respostas obtidas; quanto menor a frequência de uma resposta, maior a probabilidade de não ser uma variante.

A falta de critérios como este tem gerado atlas linguísticos com dados problemáticos. É o caso do Atlas Linguístico do Estado do Ceará – ALECE: apesar do valor inestimável que a obra apresenta para o cenário dialetal do Nordeste brasileiro, é preciso questionar a validade de algumas respostas mapeadas, como as cartas 096 e 097 que tratam das denominações para *arco-íris*. Bessa *et al.* (2010) registraram nas cartas 88 e 89 variantes lexicais com maior frequência: *arco-íris*, *arco da velha*, *arco celeste* e *olho de boi*. Nas páginas seguintes (096 e 097) aparecem outras realizações com menor frequência: *as barras*, *capelo*, *cometa*, *nevoeros*, *os raios*, *os torrame*, *raio*, *torreame*, *torres*, *a neve*, *as nuve*, *as nuve escura*, *as torre*, *barra*, *barra de vento*, *barra do dia*, *carpejo*, *está barreando o tempo*, *os aico* e *raio do solo*. Diante desse conjunto de respostas com menor frequência, que denominaremos aqui de *respostas suspeitas*, o falante nativo de português brasileiro questionará se realmente *barras*, *nuvem*, *raio*, *torres* ou *neve*, por exemplo, podem assumir o *status* de variante lexical para denominar *arco-íris* no falar cearense.

Outros exemplos de atlas que considerou para o mapeamento lexical itens que não são variantes e sim respostas dadas pelos informantes é o Atlas Linguístico do Amapá – ALAP (Razky; Ribeiro; Sanches, 2017). No caso do ALAP, os autores criaram o seguinte protocolo para o mapeamento: foram consideradas até cinco variantes e cada uma representaria uma cor no mapa, indo da cor vermelha ao rosa. Para perguntas que tivessem mais de cinco respostas, os autores criaram a categoria "outras variantes", marcada pela cor cinza. E é nesta categoria que está o problema de mapeamento, pois não houve seleção do que seria apenas respostas do que corresponderia de fato a uma variante lexical. Como exemplo, tem-se a carta L12 que trata das denominações para "macaxeira" (como chama aquela raiz branca por dentro, coberta por uma casca marrom, que se cozinha para comer?), em que foram consideradas como variantes macaxeira e mandioca, e na categoria "outras variantes" aparece uma única vez o termo "tajoba", mencionado por uma falante da segunda geração da cidade de Porto Grande.

No entanto, se levarmos em consideração os dicionários gerais de língua portuguesa, tajoba ou taioba refere-se a uma planta herbácea (*Xanthosoma violaceum*), da família das aráceas, nativa das regiões tropicais das Américas, de folhas sagitadas, grandes, com pecíolos longos, e rizomas tuberosos (Houaiss; Villar, 2009). Com isso, percebe-se que houve agrupamento de respostas dadas e não avaliadas se seriam de fato uma variante. Neste exemplo, fica claro que a informante se equivocou com resposta dada para descrição de mandioca/macaxeira.

Essa problemática nos mostra que apenas o critério de frequência lexical não é suficiente para validar respostas com baixa incidência. No entanto, é fato que avaliar a frequência de uso auxilia na distinção entre formas periféricas e centrais, permitindo que o pesquisador estabeleça critérios objetivos para a seleção de variantes nos mapas linguísticos. Para casos atípicos como o que foi apresentado no ALECE, é preciso incorporar o próximo critério para validação de respostas *suspeitas*. Estamos falando da consistência do signo linguístico e da triangulação de dados.

3 Estabilidade do signo linguístico e triangulação de dados

O estudo do significado e sua relação com o referente constitui um eixo central dos Estudos da Linguagem, especialmente relevante para a Geolinguística, à medida que subsidia a validação das formas lexicais proferidas pelos informantes na pesquisa de campo. Como indicado no tópico anterior, há muitos exemplos de formas cartografadas como variantes lexicais, mas que não mantêm correspondência semântica com os referentes especificados no instrumento de coleta. Este tópico centra-se justamente na importância

de analisar se a forma lexical declarada pelo informante, no momento elocucional, estabelece de fato uma relação direta com o referente, ou seja, se há estabilidade do signo linguístico, ou se se trata apenas de uma criação neológica momentânea. Também faz parte desse processo a triangulação dos dados por meio de obras lexicográficas e dos atlas linguísticos.

O significado, como conteúdo semântico do signo linguístico, é compreendido por Ullmann (1972) dentro de uma abordagem referencial, estruturada na tríade: *nome* (*name*), *sentido* (*sense*) e *coisa* (*thing*). O *nome* corresponde à forma da palavra, seus traços acústicos, sua configuração fonética. O *sentido* refere-se à informação que o nome comunica ao ouvinte, ou seja, à relação recíproca e reversível entre o som e o conceito associado. Tal relação é considerada reversível porque pode ser ativada em ambas as direções do processo comunicativo: o ouvinte acessa o conceito ao ouvir a palavra, e o falante enuncia a palavra ao ativar o conceito. Já a *coisa*, o referente propriamente dito, é o objeto, evento ou entidade no mundo real. A associação entre significado e referente é particularmente importante no contexto teórico e prático a que se propõe este texto, pois se evidencia que as línguas são construções históricas, sociais e culturais, moldadas por seu tempo e marcadas por fatores diatópicos, diageracionais e muitos elementos de ordem extralinguística.

O ponto de vista referencial é enriquecido por uma perspectiva funcionalista, como a proposta por Coseriu (1979), que introduz os planos de *sistema*, *norma* e *fala*, fundamentais para compreender a variação lexical documentada nos atlas linguísticos. O *sistema*, para esse autor, é o conhecimento linguístico abstrato e coletivo partilhado por todos os falantes de uma comunidade, correspondendo ao conjunto de possibilidades estruturais de uma língua (fonemas, morfemas, construções, léxico etc.). A *norma* diz respeito à forma socialmente consagrada de realizar as possibilidades do *sistema*, enquanto a *fala* é a atividade do falante, individual e elocucional; nela o indivíduo mobiliza o *sistema* e segue (ou não) a *norma*, conforme sua intenção comunicativa.

O *sistema*, para o autor, refere-se ao conhecimento linguístico compartilhado por uma comunidade e no léxico ele estaria envolto em duas funções: representativa e associativa. A primeira considera a língua como modo de representar o mundo por meio dos signos; a segunda percebe a língua como associação entre os signos, isto é, quando ele é criado e quando é reutilizado (Coseriu, 1979). As variantes lexicais, neste caso, como detentoras de significados semelhantes, estariam embutidos na função representativa do léxico, pois se ligam ao referente, ou seja, no plano de conteúdo[4].

Ainda, o estudo da significação e nomeação pode ser traduzido pelos planos semasiológico e onomasiológico – este último, de especial importância nos questionários dialetais, pois analisa as múltiplas expressões acerca de determinado conceito, os vários nomes atribuídos a um referente. Já o plano semasiológico parte da palavra especificamente para buscar seus muitos significados (Geeraerts, 2010).

No caso específico da pergunta onomasiológica, seu intuito é obter os vários nomes para o conceito, cuja indagação parte de uma pergunta do tipo 'gatilho', iniciada por pronomes interrogativos ou expressões interrogativas (*Como se chama...?*; *Qual o nome...*?). Obtidas as formas, na exegese do material, o geolinguista depara-se com uma questão fundamental: o que pode ou deve ser reconhecido como sinônimo lexical e representado nos mapas linguísticos? A esse respeito, conforme a reflexão de Barbosa (1998), é importante dis-

[4] A distinção entre significado e referente, conforme discutida por Ullmann (1972), é central para a análise geolinguística, pois evidencia que a mesma entidade extralinguística (referente) pode ser nomeada por diferentes palavras (signos linguísticos), conforme a região, o grupo etário, o contexto social ou histórico; e que, inversamente, uma mesma palavra pode adquirir significados diferentes conforme o contexto cultural em que circula – ao passo que as formas linguísticas documentadas em campo não são apenas expressões individuais (*fala*), mas estão enraizadas em um sistema de possibilidades estruturais (*sistema*) e refletem, ao mesmo tempo, padrões de uso consagrados socialmente (*norma*). O léxico, nesse sentido, é o ponto de encontro entre o plano abstrato da língua e a realização concreta da fala: ele revela como cada comunidade classifica conceitualmente o mundo (função representativa) e como associa palavras a esses conceitos por meio do uso reiterado ou inovador (função associativa).

tinguir entre sinonímia e parassinonímia, sendo esta última mais apropriada aos propósitos da cartografia linguística. Isso porque, nas práticas onomasiológicas, é comum encontrar unidades lexicais que compartilham o mesmo plano de conteúdo, mas que se diferenciam quanto ao plano de expressão (influenciadas por fatores culturais, geográficos, sociais ou situacionais).

Assim, na exegese dos dados, a partir do levantamento realizado pelo pesquisador das respostas fornecidas *in loco*, após aplicado o primeiro critério de validação (o da frequência), passa-se à segunda etapa do processo de validação e triangulação: a verificação dos parassinônimos com o apoio de i) dicionários, vocabulários e glossários e dos ii) atlas linguísticos.

Os dicionários e vocabulários, como produtos lexicográficos, constituem importante acervo da língua, ao passo que documentam em dado momento histórico o produto linguístico de seu tempo, as normas e marcas de uso circunscritas a um período (Isquerdo, 2007), e todos são importantes ao trabalho geolinguístico. Os dicionários de caráter mais geral da língua apresentam a norma lexical de um dado território, o vocabulário comum; os etimológicos apresentam marcas diacrônicas da palavra, sua possível origem e evolução semântica; já os dicionários regionalistas apresentam marcas culturais de uso restrito a determinadas comunidades de fala (Isquerdo, 2007)[5]. Quanto mais ampla for a pesquisa realizada em obras lexicográficas, maior será a segurança do geolinguista em considerar determinada forma como variante lexical e compor o mapa dialetal.

A relevância da consulta aos dicionários é ilustrada no exemplo de *cofo*, coletado para a questão 038 do ALiTTETO (Silva, 2018, Carta L3, Tomo II), como designativo para os objetos de vime, de taquara, de cipó trançado, para levar batatas, mandioca, no lombo animal. De um universo de 118 respostas, *cofo* obteve nove (09) alusões (7,7%). Para a validação da forma como variante, recorreu-se primeiramente ao dicionário geral da língua portuguesa Houaiss (Houaiss; Villar, 2009), que o indica como regionalismo da Guiné-Bissau, "cesto de vime, de tamanho menor do que o de um balaio, usado em casas rurais". Ainda nesse dicionário, encontram-se importantes marcações etimológicas: de *cofinho*, latim *cophinus* 'cesto' e, este, do grego *kóphinos* 'cestinha de flores'; no português medieval, provavelmente reduzido a *cofo*, que, em verdade, seria forma divergente de *cofre*. Já na obra etimológica de Aulete (1986), *cofo* é apresentado como objeto utilizado por pescadores e sua etimologia remete a *cofinho*, proveniente do Minho (Portugal) com datação do século XIV.

A presença de *cofo* em ambos os dicionários esclarece ao pesquisador que a forma deve ser considerada variante para o referente solicitado, pois é designativo para objetos tecidos com fibras naturais e serve ao propósito de 'carregar coisas', além de apresentar marcas diacrônicas importantes: ou se trata de variante procedente de Portugal, como indicado por Aulete (1986), ou de regionalismo bissau-guineense, segundo Houaiss e Villar (2009)[6].

Ainda no rol de obras lexicográficas, são também importantes os dicionários especializados em determinada área do conhecimento: botânica, jurídica, fauna, símbolos, folclore, na perspectiva de validação da forma presente em determinado campo léxico. Notou-se a importância dessas obras especializadas na confirmação de *corvo* como variante para *urubu*, a partir do *caput*: *Como se chama a ave preta que come animal morto, podre?* (Comitê Nacional do Projeto ALiB, 2001), cujos dicionários aludem a ambas as formas como sinônimas e fornecem a possível trajetória social e histórica da variante no Brasil.

Em uma rubrica ornitológica, fornecida por Houaiss e Villar (2009), a ave *corvo* é a designação comum a diversas grandes aves da família dos corvídeos, de plumagem preta. Ainda segundo esse dicionário, a ave é encontrada em todos

---

[5] Dicionários dialetais. Ver Américo: https://gruponemesis.ufba.br/sites/gruponemesis.ufba.br/files/projeto_ddb.pdf.

[6] É importante deixar claro que, neste momento, o objetivo do levantamento em obras lexicográficas é o de validar ou não uma forma lexical para compor a cartografia, não se pretendendo confirmar ou refutar os indicativos etimológicos fornecidos pelos dicionários, temática que poderá ser tratada em trabalhos futuros ao atlas.

os continentes, *com exceção da América do Sul*. No dicionário etimológico de Cunha (1999), por exemplo, há uma sutil comparação entre *corvo* e *urubu*, no que se refere ao tamanho e aos hábitos alimentares de ambas as aves. Aulete (1986) fornece *corvo* com brasileirismo dado 'erradamente' ao *urubu*, enquanto os especializados, como o *Dicionário de Símbolos* (2008) e o de *Folclore* de Cascudo (2012), indicam que o traço de semelhança entre o *urubu* e o *corvo* está relacionado ao comportamento alimentar de ambas as aves, o que as associa ao mau agouro e ao azar. Todas essas obras levam à compreensão de que o nome *corvo* fora introduzido pelo colonizador europeu ao se deparar com a ave brasileira, fornecendo-lhe o nome de uma ave mais próxima de sua realidade europeia[7].

A consulta maciça a obras lexicográficas foi de suma importância a Chofard (2023), em seu *Atlas Linguístico da Rota dos Tropeiros*, sobretudo na validação das 308 respostas coletadas para o questionamento sobre as designações para o veículo puxado por animais para transportar cargas ou pessoas, das quais puderam ser interpretadas como variantes 27 delas, com o auxílio dos dicionários gerais e especializados. No entanto, algumas unidades lexicais auferidas para o citado referente não se encontravam dicionarizadas, sendo necessário ampliar a pesquisa para as obras geolinguísticas. Um exemplo de forma não lexicalizada mas que fora registrado em atlas linguístico é o de *gaiota*, computada no *corpus* com 13 respostas, somatizando pouco mais de 4,0% no cômputo geral. Entretanto, quando observado o *Atlas Linguístico Etnográfico da Região Sul* (ALERS) (Altenhofen; Klassmann, 2011), a forma aparece de forma produtiva como *galhota* e marca neste atlas regional "[...] áreas de deslocamento gaúcho, como no oeste catarinense e paranaense, além de áreas fronteiriças com São Paulo com possível influência do vaivém tropeiro" (Chofard, 2023, p. 170).

Essa perspectiva realizada pela autora remete ao outro passo no processo de triangulação abordado neste tópico: a consulta aos diversos atlas linguísticos já confeccionados. Recomenda-se que, neste processo, o pesquisador inicie sua busca pelos atlas linguísticos relacionados à região em estudo. Por exemplo, se a obra em processamento tratar de um atlas de pequeno domínio, como o de uma comunidade ribeirinha do Amazonas, é aconselhável consultar inicialmente os demais atlas linguísticos produzidos sobre esse estado; em seguida, pode-se estender a busca aos demais atlas linguísticos das unidades federativas contíguas; se possível, averiguar a cartografia do item em atlas estrangeiros também, sobretudo nos atlas sul-americanos e nos europeus[8].

Em síntese, a análise do significado e sua relação com o referente é essencial para a pesquisa geolinguística, na perspectiva de distinguir entre formas lexicalmente legítimas ou criações momentâneas decorrentes do desconhecimento do informante. A estabilidade do signo linguístico é um princípio teórico que orienta essa distinção e compreender essas categorias auxilia o pesquisador a identificar os limites e as possibilidades do léxico, observando como os falantes, em contextos culturais e regionais diversos, nomeiam um mesmo referente de diferentes formas. Ainda, nesse processo, a triangulação dos dados com o auxílio de obras lexicográficas e com outros atlas linguísticos torna-se uma etapa importante na validação das formas obtidas em campo. A identificação de sinônimos ou parassinônimos garante maior precisão na representação cartográfica dos dados, como forma de assegurar que a forma registrada não seja um desvio referencial, mas uma expressão linguística compartilhada e culturalmente enraizada na comunidade de fala.

---

[7] *Corvo* foi auferido nas capitais estaduais de São Paulo, Mato Grosso, Paraná, Santa Catarina e Rio Grande do Sul, formando um percurso diatópico por onde passavam as antigas rotas comerciais dos tropeiros paulistas em direção ao Rio Grande do Sul (Silva; Altino, 2023).

[8] Até 1996, ano de criação do Projeto ALiB, não havia um questionário dialetal de caráter mais geral a ser adotado pelos diferentes autores para a elaboração de seus atlas linguísticos, o que facilitaria uma análise comparativa entre cartas lexicais coincidentes. Entretanto, o ALiB como projeto interinstitucional sedimentou as bases da geolinguística brasileira, e seu questionário, embora elaborado para a finalidade de um atlas nacional, vem sendo utilizado em diferentes projetos geolinguísticos, tanto nos atlas estaduais como nos de pequeno domínio, permitindo a comparação cruzada entre muitos dessas obras.

4 Validação comunicativa (*members checking*)

Outro processo de sumária importância na validação das formas lexicais diz respeito ao que a bibliografia em Epistemologia da Pesquisa ou Metodologia Científica intitula de *member checking*, ou validação pelo participante, um tipo de técnica usada para conferir maior confiabilidade a pesquisas de caráter qualitativo. A prática consiste em incluir, de forma mais presente, a validação dos dados por parte de quem os proferiu, "[...] garantindo que os significados e perspectivas dos próprios participantes estejam representados e não sejam suprimidos pela agenda e conhecimento do pesquisador"[9] (Tong *et al.*, 2007, p. 356, tradução nossa).

Segundo Birt *et al.* (2016), o *member checking* pode incluir uma série de abordagens, tais como: i) devolução da transcrição da entrevista ao participante; ii) entrevista de verificação com base na transcrição ou nos dados interpretados; iii) grupos focais para verificação; e iv) devolução de dados analisados e sintetizados – cada uma delas, com pontos positivos e negativos. Especificamente para a pesquisa dialetológica, interessam os itens ii e iii, pelo seu caráter interpretativista e pela possibilidade de adição de novos dados à pesquisa, sobretudo os metalinguísticos e epilinguísticos.

O item ii, entrevista de verificação, consiste em realizar uma primeira coleta com o participante e, posteriormente, retornar a esse contato para validar, inserir ou retirar alguma informação anteriormente fornecida. Esse procedimento é muito importante, principalmente quando a pesquisa geolinguística assume uma vertente etnolinguística; no entanto, infelizmente, o retorno à localidade envolve custos e demanda tempo para que seja desenvolvido.

Esse procedimento foi adotado por Coelho (2024) em seu trabalho *Diversidade lexical em comunidades afroamapaenses*, no item 033 – *constelação*. Em inquéritos realizados entre os anos de 2018 e 2019, os informantes de Mazagão Velho não mencionaram a variante *caminho de Santiago*, marcadamente de origem portuguesa, o que causou estranhamento à pesquisadora, uma vez que a referida localidade abriga a festa religiosa de São Tiago, trazida pelas famílias portuguesas à região. Diante disso, Coelho (2024, p. 134-136) retornou a duas localidades (Mazagão Velho e Kulumbu do Patuazinho) e aplicou novamente o questionamento, recebendo, então, a devolutiva com a forma lexical portuguesa.

No procedimento empreendida por Coelho (2024) também houve a inclusão do processo identificado por Thun (1999) como 'Técnica em três tempos' (*Dreischritt-Methode*), que consiste em perguntar (resposta espontânea), insistir e sugerir[10]. Ela favorece a identificação, por parte do informante, de dados de natureza referencial, já que frequentemente emergem comentários sobre as variantes investigadas, indicando se fazem parte de seu uso cotidiano ou se estão associadas a grupos de outras localidades. Tais relatos são de fundamental relevância para a elaboração de atlas linguísticos pluridimensionais[11]. Entretanto, ao uso dessa técnica cabem considerações: por vezes, entre os geolinguistas, discute-se a espontaneidade das variantes lexicais decorrentes desse processo e se elas devem ou não figurar na carta linguística contígua a outras formas fornecidas espontaneamente. Cada pesquisador deve adotar o que melhor lhe convier, mas sugere-se que a variante obtida por meio de técnica de Thun (1999) seja marcada com algum tipo de sinal diferenciativo na superfície do mapa (um diacrítico * por exemplo), para que o leitor tenha acesso a essa informação.

Ainda em relação à entrevista de verificação,

---

[9] No original: *"[...] ensuring that the participants' own meanings and perspectives are represented and not curtailed by the researchers' own agenda and knowledge"*.

[10] Exemplifica-se da seguinte forma: no momento do inquérito, formula-se a pergunta base do questionário e aguarda-se a resposta. Na ausência de resposta, insiste-se geralmente por meio de uma reformulação da pergunta; caso o informante negue conhecer o referente, o inquiridor sugere algumas formas variantes para verificar se o interlocutor, de fato, reconhece ou não o item solicitado. A sugerência, última etapa do processo, parte do plano semasiológico, ou seja, da variante lexical para os seus outros significados.

[11] Coelho (2024, p. 137-138), ao fazer uso da estratégia, captou informações relacionadas ao caráter diarreligioso da forma lexical *caminho de Santiago*, como evidenciado no relato do informante: "[...] ouvi falar sim, é o caminho de São Tiago. Fica muito mais visível no mês de julho, que é quando acontece a festa dele, é como se fosse o nosso santo querendo anunciar que o dia dele já vai chegar".

nos projetos geolinguísticos brasileiros, a técnica não é comumente aplicada como descrito por Birt *et al.* (2016), em que se realiza uma primeira entrevista e, posteriormente, após os dados transcritos, retorna-se ao mesmo informante para confirmar ou refutar os dados, pois uma nova recolha traria elevados custos, sobretudo em projetos de médio e grande domínio. Em um macroprojeto como o ALiB, por exemplo, cuja coleta foi operada em 250 localidades, seria praticamente impossível retornar aos 1.100 informantes.

A terceira abordagem elencada por Birt *et al.* (2016) para validação dos dados é chamada de iii) grupos focais para verificação, e consiste em expor, discutir e validar os dados em coletivos compostos por mais de três pessoas. Como pontos positivos, os grupos focais podem suscitar muitos comentários epilinguísticos e valorativos sobre a língua, mas, como pontos negativos, Birt *et al.* (2016, p. 1804, tradução nossa) indicam: "A coerção do grupo pode dificultar a manifestação de uma voz dissonante individual; pode haver coerção, já que pode ser difícil discordar do pesquisador se ele estiver conduzindo o grupo focal"[12].

Inspirado em conceito similar, Radtke e Thun (1996) descrevem a técnica em geolinguística como 'pluralidade simultânea dos informantes' e consiste em entrevistar mais de um participante ao mesmo tempo. Para os autores, eles se sentiriam mais livres para expressar concordância ou discordância entre as formas lexicais, além de fornecerem uma série de dados contextuais, por vezes não coletados em entrevistas tradicionais realizadas entre um participante e o inquiridor apenas; ademais, essa também é uma forma de neutralizar a presença e o monopólio do entrevistador[13].

Dentre todos os pontos discutidos neste tópico, uma observação é fundamental: o êxito do *members checking* depende primordialmente da atuação e do comprometimento do inquiridor, responsável por conduzir a entrevista, cabendo a ele conhecer detidamente os objetivos gerais da pesquisa, bem como cada item presente no questionário a ser aplicado, além de um bom relacionamento interpessoal, paciência e cordialidade[14]. Neste sentido, entre o ideal e o possível, caso o pesquisador identifique a impossibilidade financeira de retornar à localidade de inquérito, é necessário considerar diversos critérios para manter a confiabilidade dos dados. No momento da coleta, o pesquisador deve realizar todos os testes de verificação e percepção possíveis em relação às formas proferidas, solicitando ao informante que descreva o item. Sempre que possível, recomenda-se o uso de fotos e objetos que facilitem a compreensão do entrevistado, além de possibilitar a obtenção de relatos com valor diatópico e diastrático.

5 Revisão por informantes-chave (*external validation*)

Seguindo as mesmas premissas da abordagem *members checking*, o último critério a ser aplicado deve ser a *external validation* ou revisão por informantes-chave: trata-se de um procedimento em que uma pessoa qualificada, mas que não participou da coleta ou análise de dados, revisa o processo de pesquisa para avaliar se os dados foram coletados de maneira sistemática, se a análise foi lógica e consistente e se as interpretações estão fundamentadas nos dados.

Esse tipo de proposição não é novo e tem sido bastante utilizado nas Ciências Humanas, sobretudo em pesquisas etnográficas. Guba e Lincoln (1985), por exemplo, defendem as auditorias externas como uma forma de assegurar consistência e neutralidade na pesquisa. O foco

---

[12] No original: *"Group coercion can make it difficultor single disconfirming voice; Coercion as it might be hard to disagree with the researcher if they are leading the focus group"*.

[13] O último item diz respeito à iv) devolução de dados analisados e sintetizados, e consiste em discutir com os participantes o conjunto total dos dados para ver se os resultados ressoam de fato com a experiência vivenciada em uma comunidade. Para isso, faz-se necessário que a apresentação desses dados seja acessível ao informante, devendo o pesquisador responsável fazer uso, preferencialmente, de materiais mais didáticos para facilitar a compreensão e a responsividade do participante (Bird *et al.*, 2016).

[14] É sumariamente importante que, antes de iniciar o trabalho de campo, cada membro deve realizar inquéritos experimentais ou entrevistas-teste, expondo-se a situações adversas. Isso permite que o pesquisador se familiarize com diversos desafios e desenvolva um bom senso crítico para conduzir as entrevistas (Silva; Borges, 2022).

é observar o comportamento e os fenômenos de forma natural, sem manipulação. Já Yin (2009) defende a importância da auditoria externa para dar credibilidade a estudos de caso, garantindo que outros pesquisadores possam seguir os mesmos procedimentos.

No âmbito dos Estudos da Linguagem, cita-se a obra de Gumperz (1982), *Discourse Strategies*, que mostra como as interações sociais e culturais influenciam a comunicação e como os informantes desempenham papel importante na validação e interpretação dos dados linguísticos. Embora Gumperz não se refira diretamente à "revisão por informantes-chave" como um termo específico, ele utiliza o método etnográfico no qual os informantes ajudam a interpretar e contextualizar os dados coletados, especialmente em suas discussões sobre variabilidade linguística e contextos de interação verbal.

A seleção do revisor ou informante-chave pode ser um membro da comunidade que participou ou não participou diretamente da pesquisa, mas conhece a dinâmica social, cultural e política da comunidade. Ele pode ser um ancião, cacique, professor, pesquisador, presidente da comunidade etc. Esse olhar externo sobre os dados é fundamental para que o trabalho final ganhe confiabilidade, trazendo uma visão de fora do processo, demonstrando rigor metodológico para o público acadêmico ou para órgãos de financiamento e garantindo transparência no processo científico.

Como exemplo de aplicação deste critério, cita-se o *Microatlas linguístico (português-kheuól) da área indígena dos Karipuna do Amapá* (Sanches, 2020), que trata do contato linguístico entre o português e o kheuól, línguas faladas pelo povo indígena Karipuna, residente no extremo norte do Amapá, na fronteira com a Guiana Francesa. A pesquisa adotou uma abordagem geolinguística, de perspectiva pluridimensional, analisando a variação lexical em nove aldeias: Manga, Santa Izabel, Espírito Santo, Açaizal, Curipi, Kariá, Ahumã, Ariramba e Kunanã. Foram entrevistados 36 indígenas bilíngues, estratificados por faixa etária e sexo, utilizando o questionário semântico-lexical para mapear o repertório lexical em português e em kheuól.

Para o mapeamento dos dados coletados, Sanches (2020) realizou a validação das respostas de duas maneiras: a triangulação de dados e a revisão por informantes-chave. A primeira diz respeito à triangulação de dados linguísticos (conforme mencionado na seção 2), contando com o auxílio de duas gramáticas, Tobler (1983) e Spires (1984); dois dicionários, Tobler (1987) e Montejo (1988); e um vocabulário experimental, Green e Green (2004). As gramáticas sobre o kheuól foram fundamentais para entender o funcionamento da estrutura da língua e de seus aspectos fonéticos e fonológicos. Já os dicionários e o vocabulário auxiliaram no reconhecimento de itens mencionados na pesquisa e na escrita ortográfica de variantes lexicais em kheuól.

Além da consulta aos materiais citados, a validação das variantes também contou com a revisão por informantes-chave. Para isso, o autor teve o apoio de dois professores indígenas que ensinam a variedade kheuól em comunidades indígenas: o professor Odilton Felipe da Paixão, que atua na escola Jorge Iaparra, localizada na aldeia Manga, e a professora Edilena dos Santos, educadora modular nas aldeias Karipuna. Esses colaboradores receberam uma versão preliminar de todas as cartas lexicais produzidas para o trabalho final. Juntamente com o pesquisador, revisaram o material impresso, apontando possíveis desvios ortográficos e variantes lexicais em kheuól que necessitavam de correção. Após essa etapa de ajustes, o conteúdo foi analisado e, em seguida, publicado.

6 Respostas não validadas: o que fazer?

As respostas não validadas em pesquisas geolinguísticas são informações ou dados coletados junto aos informantes que não foram confirmados ou verificados por meio de procedimentos aqui descritos. Tais respostas, muitas vezes registradas durante a coleta de dados, não se consolidam como variantes lexicais legítimas por apresentarem baixa frequência de uso, ausência de estabilidade semântica, falta de reconhecimento

em fontes lexicográficas e ausência em registros anteriores de atlas linguísticos. Diante desse cenário, impõe-se a necessidade de estabelecer critérios objetivos para decidir o que fazer com tais ocorrências.

Em primeiro lugar, é fundamental compreender que a rejeição automática de uma forma lexical de baixa ocorrência pode ser tão problemática quanto sua aceitação acrítica. As respostas não validadas podem, em determinadas situações, refletir tendências emergentes da língua, regionalismos localizados ou mesmo formas em processo de obsolescência; portanto, seu tratamento metodológico deve ser cauteloso.

Uma proposta possível para lidar com tais respostas consiste na criação de uma categoria intermediária, que as classifique como "formas não validadas" ou "respostas marginais". Essa nomenclatura indica que, embora os dados não cumpram integralmente os critérios de validação propostos (frequência, estabilidade semântica, triangulação e validação participativa ou externa), possuem potencial valor informativo e devem ser preservados no registro da pesquisa, ainda que não figurem na cartografia principal, podendo figurar como notas explicativas ou apêndices.

Outra possibilidade é a representação dessas respostas nos mapas linguísticos, com o uso de marcadores gráficos específicos, como asteriscos, sombreado diferenciado ou legenda explicativa, no intuito de informar ao usuário que se trata de uma forma com validade parcial, cuja representação deve ser interpretada com cautela. Dessa forma, o pesquisador mantém o compromisso com a transparência científica e com a documentação da variação, sem incorrer no erro de atribuir *status* de variante consolidada a formas isoladas ou episódicas.

No entanto, há situações em que a resposta não validada não pode ser aproveitada nem como forma marginal. É o caso de respostas que não mantêm qualquer relação semântica com o referente questionado, ou que se configuram como erros interpretativos, lapsos de memória, tentativas de adivinhação ou até mesmo respostas humorísticas por parte do informante. Tais ocorrências devem ser descartadas do processo de mapeamento, ainda que sejam registradas no diário de campo ou no banco de dados para fins de controle e documentação metodológica, como o exemplo de *a neve*, *as nuve*, *as nuve escura*, *barra de vento* para arco-íris.

A decisão sobre a exclusão ou a manutenção de uma forma deve sempre estar documentada no relatório metodológico da pesquisa e, para este fim, recomenda-se que os pesquisadores construam quadros ou tabelas metodológicas nas quais fiquem registrados os critérios de aceitação e rejeição, acompanhados de exemplos concretos retirados do *corpus* analisado.

Por fim, cabe mencionar que, em casos específicos, uma resposta inicialmente não validada pode vir a ser reconsiderada em etapas posteriores da pesquisa, a partir da consulta a novos dados, dicionários atualizados, outros atlas, ou mesmo após retorno ao campo para novos inquéritos. Dessa forma, é recomendável que os dados classificados como "não validados" sejam armazenados de forma organizada e acessível, permitindo sua reanálise à luz de novos aportes teóricos ou empíricos.

7 Considerações finais

Os critérios apresentados, quando articulados, conferem maior rigor metodológico aos dados dialetológicos, evitando que respostas suspeitas ou não validadas comprometam a confiabilidade dos dados representados nos atlas linguísticos. Assim, este trabalho contribui para a Geolinguística ao oferecer um protocolo sistematizado de validação de variantes, cuja aplicação amplia o alcance descritivo e interpretativo das pesquisas dialetológicas no Brasil.

O primeiro critério listado é o da frequência lexical, amplamente utilizado nos atlas já publicados. O pesquisador deve considerar que variantes altamente frequentes possuem maior relevância para o mapeamento, enquanto formas de baixa frequência podem indicar mudanças linguísticas em andamento ou tendências de desuso. São essas formas menos frequentes que têm se traduzido em desafios ao geolinguista.

Como segundo critério, propõe-se a verificação da estabilidade do signo linguístico, aliada à triangulação dos dados por meio de obras lexicográficas e atlas linguísticos, com o intuito de analisar se a unidade lexical enunciada pelo informante mantém, de fato, correspondência com o referente solicitado. A consulta a outros atlas linguísticos oferece respaldo metodológico ao geolinguista na validação de formas consideradas duvidosas, possibilitando, inclusive, identificar se a forma decorre de padrões diatópicos topodinâmicos, revelando o percurso migratório da variante.

O terceiro passo diz respeito à validação das formas lexicais por meio do *member checking* (entrevista de verificação), pluralidade simultânea de informantes ou dos grupos focais, pois permite ao pesquisador captar também dados metalinguísticos e epilinguísticos. Em virtude das limitações logísticas e financeiras que muitas vezes inviabilizam o retorno às localidades investigadas, é fundamental que, já no primeiro inquérito, o pesquisador realize os procedimentos para assegurar uma posterior análise dessas formas, tais como: solicitar a descrição do item pelo informante, o uso de imagens e objetos de apoio, bem como a aplicação das técnicas indicadas por Thun (1999) e pela Dialetologia Pluridimensional.

Por último, a revisão por informantes-chave (*external validation*) busca a validação do item por meio de pessoa qualificada e externa ao processo, que assegura rigor metodológico, neutralidade e fidelidade às dinâmicas sociais e culturais da comunidade inquirida. Após esses passos, às respostas não validadas, indica-se tratamento metodológico cuidadoso, que reconheça seu potencial informativo, sem necessariamente figurar nas cartas linguísticas. Deve-se indicar de forma clara as decisões de inclusão ou exclusão dessas respostas no relatório metodológico e a guarda desses dados permite futuras reavaliações por parte do pesquisador.

Por fim, acreditamos que esse conjunto de critérios pode orientar de forma sistemática as decisões do geolinguista na elaboração cartográfica de seu atlas. Contudo, ressaltamos que esses quatro critérios arrolados não constituem um conjunto fechado, sendo possível a incorporação de novos critérios ou a reformulação daqueles já existentes.

Referências

ALTENHOFEN, Cléo Vilson; KLASSMANN, Mário Silfredo (org.). *Atlas Linguístico-Etnográfico da Região Sul do Brasil*: cartas semântico-lexicais. Porto Alegre: Editora UFGRS; Florianópolis: Editora UFSC, 2011.

AULETE, Caldas. *Dicionário contemporâneo da língua portuguesa*. 5. ed. Rio de Janeiro: Delta, 1986.

BARBOSA, Maria Aparecida. Relações de significação nas unidades lexicais. *In*: CARVALHO, Nelly Medeiros; SILVA, Maria Emília Barcellos da (org.). 1º Encontro Nacional do GT de Lexicologia, Lexicografia e Terminologia da ANPOLL. *Anais [...]*. Recife: UFPE/CNPq, 1998.

BÉJOINT, Henri. *The Lexicography of English*: From Origins to Present. Oxford University Press, 2010.

BESSA, José Rogério *et al*. *Atlas linguístico do Estado do Ceará*. v. 2. Fortaleza: UFC, 2010.

BIRT, Linda *et al.* Member checking: a tool to enhance trustworthiness or merely a nod to validation? *Qualitative Health Research*, Thousand Oaks, v. 26, n. 13, p. 1802–1811, 2016. DOI: 10.1177/1049732316654870.

BLOOMFIELD, Leonard. *Language*. New York: Henry Holt and Company, 1933.

BRYSBAERT, Marc; NEW, Boris. Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, Windsor, v. 41, n. 4, p. 977–990, 2009.

CARDOSO, Suzana Alice Marcelino. *Geolinguística*: Conceitos e Aplicações. São Paulo: Parábola, 2010.

CASCUDO, Luís Câmara. *Dicionário do folclore brasileiro*. 12. ed. São Paulo: Global, 2012.

CHOFARD, Amanda. *Contatos intervarietais das variedades Sul-Rio-Grandense e Paulista nos dados do Projeto Atlas Linguístico da Rota dos Tropeiros*. 2023. 306 f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2023. Disponível em: https://repositorio.ufsc.br/handle/123456789/252290. Acesso em: 8 maio 2025.

COELHO, Helen Costa. *Diversidade lexical em comunidades afro-amapaenses*: contornos diatópicos, diastráticos e diarreligiosos. 2024. 417 f. Tese (Doutorado em Linguística) – Universidade Federal do Pará, Belém, 2024.

COMITÊ NACIONAL DO PROJETO ALiB. *Atlas Lingüístico do Brasil*: questionários. Londrina: Eduel, 2001.

COSERIU, Eugenio. *Linguística del texto y teoría de la competencia comunicativa*. Madrid: Gredos, 1979.

CUNHA, Antônio Geraldo da. *Dicionário etimológico Nova Fronteira da língua portuguesa*. 2. ed. Rio de Janeiro: Nova Fronteira, 1999.

DICIONÁRIO DE SÍMBOLOS. *Significado dos símbolos e simbologias*. Matosinhos: 7Graus, 2008. Disponível em: https://www.dicionariodesimbolos.com.br/. Acesso em: 6 maio 2025.

ELLIS, Nick. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, Cambridge, v. 24, n. 2, p. 143–188, 2002.

GEERAERTS, Dirk. *Theories of Lexical Semantics*. Oxford: Oxford University Press, 2010.

GREEN, Diana; GREEN, Harold. *Vocabulário Português – Paliku - Kheuol*. Edição preliminar e experimental. Belém: Sociedade Internacional de Lingüística, 2004.

GUBA, Egon G.; LINCOLN, Yvonna S. *Naturalistic Inquiry*. Beverly Hills: Sage, 1985.

GUMPERZ, John Joseph. *Discourse Strategies*. Cambridge: Cambridge University Press, 1982.

HOUAISS, Antônio; VILLAR, Mauro de S. *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Objetiva, 2009. CD-ROM.

ISQUERDO, Aparecida Negri (org.). *A propósito de dicionários de regionalismos do português do Brasil*. As ciências do léxico: lexicologia, lexicografia e terminologia. v. III. Campo Grande: Editora UFMS, 2007.

LABOV, William. *Principles of Linguistic Change*: Internal Factors. Oxford: Blackwell, 1994.

LABOV, William. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.

MONTEJO, Francisca Pincaço. *Dicionário Kheuól - Português e Português-Kheuól*. Belém: Mensageiro, 1988.

NATION, Paul. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press, 2001.

RADTKE, Edgar; THUN, Harald. Nuevos caminos de la geolinguística románica. Un balance. *In*: RADTKE, Edgar; THUN, Harald. *Neue Wege der Romanischen Geolinguistik*. Kiel: Westensee-Verlag, 1996. p. 25-49.

RAZKY, Abdelhak; RIBEIRO, Celeste Maria da Rocha; SANCHES, Romário Duarte. *Atlas linguístico do Amapá*. São Paulo: Labrador, 2017.

SANCHES, Romário. *Microatlas linguístico (português-kheuól) da área indígena dos Karipuna do Amapá*. 2020. 247 f. Tese (Doutorado em Letras) – Universidade Federal do Pará, Belém, 2020.

SANCHES, Romário; RAZKY, Abdelhak. Análise geossociolinguística das designações para fanhoso nas capitais brasileiras. *DELTA. Documentação de Estudos em Lingüística Teórica e Aplicada*, São Paulo,_v. 2, p. 1-22, 2021.

SANCHES, Romário; RAZKY, Abdelhak. Variação geossocial do item lexical riacho/córrego nas capitais brasileiras. *Revista Gragoatá*, Niterói, v. 21, p. 70-89, 2016.

SILVA, Greize Alves da. *Atlas Linguístico Topodinâmico e Topoestático do Estado do Tocantins (ALiTTETO)*. 2018. 2 v. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2018.

SILVA, Greize Alves da; ALTINO, Fabiane Cristina . Uma ave de mau agouro? o urubu e o corvo nas capitais brasileiras do ALiB. *In*: RAMOS, Conceição de Maria de Araujo; ALTINO, Fabiane Cristina; PAIM, Marcela Moura Torres (org.). *O ALiB e suas mulheres*: linhas de saudade e homenagem (Documentos 8). Cascavel: Edunioeste, 2023. v. 1, p. 309-319.

SILVA, Greize Alves da; BORGES, Patrícia Andréa. Relações de poder na pesquisa de campo sociodialetal: algumas reflexões necessárias. *Humanidades & Inovação*, Palmas, v. 9, n. 21, p. 11-23, 2022. Disponível em: https://revista.unitins.br/index.php/humanidadeseinovacao/article/view/7663. Acesso em: 28 jun. 2026.

SINCLAIR, John. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

SPIRES, Rebecca. *Gramática Kheuól*. Belém: CIMI, 1984.

THUN, Harald. O tratamento do material etnográfico no Atlas Lingüístico Diatópico y Diastrático del Uruguay (ADDU). *In*: ENCONTRO SOBRE CULTURA POPULAR, 1. 1999, Ponta Delgada. *Anais [...]*. Ponta Delgada: Universidade dos Açores, 1999. p. 481-499.

TOBLER, Joy. *The grammar of Karipuna Creole*. Brasília: SIL, 1983.

TOBLER, Alfred. *Dicionário Crioulo Karipuna - Português e Português - Crioulo Karipuna*. Brasília: Summer Institute of Linguistics, 1987.

TONG, Allison; SAINSBURY, Peter; CRAIG, Jonathan. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, [*s. l.*], v. 19, n. 6, p. 349–357, dez. 2007.

ULLMANN, Stephen. *Semântica*: uma introdução à ciência do significado. Lisboa: Fundação Calouste Gulbenkian, 1972.

YIN, Robert K. *Case Study Research*: Design and Methods. 4. ed. Thousand Oaks: Sage Publications, 2009.

ZIPF, George Kingsley. *The Psychobiology of Language*: An Introduction to Dynamic Philology. Boston: Houghton Mifflin, 1935.

ZIPF, George Kingsley. *Human Behavior and the Principle of Least Effort*: an introduction to human ecology. Cambridge: Addison-Wesley Press, 1949.

Romário Duarte Sanches

Doutor em Letras (Linguística) pela Universidade Federal do Pará (UFPA, Belém, PA, Brasil), professor da Universidade Federal do Amapá, Campus Santana, Brasil.

Greize Alves da Silva

Doutora em Estudos da Linguagem pela Universidade Estadual de Londrina (UEL, Londrina, PR, Brasil), professora da Universidade Federal do Tocantins, Campus Porto Nacional, Brasil.

Endereço para correspondência

ROMÁRIO DUARTE SANCHES

Universidade Federal do Amapá, Campus Santana, Brasil

Rodovia Duca Serra, 2191

Fonte Nova 68925-000

Santana, Amapá, Brasil

GREIZE ALVES DA SILVA

Universidade Federal do Tocantins, Campus Porto Nacional, Brasil.

Rua 03, Quadra 17, Lote 11, S/N,

Jardim dos Ipês, 77500-000

Porto Nacional, Tocantins, Brasil

Como citar este artigo

Sanches, R. D., & Alves da Silva, G. Problemas geolinguísticos: cartografia de variantes lexicais e de respostas suspeitas. Letras De Hoje, e48366. https://doi.org/10.15448/1984-7726.2026.1.48366

Os textos deste artigo foram revisados por Araceli Pimentel Godinho e submetidos para validação dos autores antes da publicação.