

Variação lexical para “amanhecer” no estado do Amapá: uma análise geolinguística

Lexical variation for “amanhecer” in the state of Amapá: A geolinguistic analysis

Matheus Gomes dos Santos¹

Romário Duarte Sanches²

Resumo: Este trabalho tem como objetivo descrever as variantes lexicais referentes à carta L79-Amanhecer, pertencente ao banco de dados do Atlas Linguístico do Amapá (ALAP). A pesquisa está fundamentada nos princípios da Geolinguística Pluridimensional, que busca evidenciar as variantes, considerando o aspecto espacial e sociocultural de seus falantes (Cardoso, 2010). O *corpus* metodológico utilizado faz parte do Projeto ALAP, que contou com 40 informantes distribuídos em dez localidades do estado do Amapá (1-Macapá, 2-Santana, 3-Mazagão, 4-Laranjal do Jari, 5-Pedra Branca do Amapari, 6-Porto Grande, 7-Tartarugalzinho, 8-Amapá, 9-Calçoene e 10-Oiapoque). Em cada ponto de inquérito, foram entrevistados um homem e uma mulher da primeira geração (18 a 30 anos) e um homem e uma mulher da segunda geração (50 a 75 anos). Em relação à cartografia linguística, a carta com as denominações para *amanhecer* foi elaborada pelos autores desta pesquisa por meio do software de design gráfico *Inkscape*. Os resultados obtidos foram representados com base na distribuição percentual das ocorrências de cada variante. Assim, observou-se a seguinte frequência: *amanhecer* (54,36%), *raiar o dia* (13,04%), *dia clareando* (13,04%), *manhã* (6,52%), *nascer do sol* (4,35%), *aurora* (2,17%), *dia* (2,17%) e *sem resposta* (4,35%).

Palavras-chave: Amanhecer. Amapá. Geolinguística. Léxico.

Abstract: This study aims to describe the lexical variants associated with map L79-Amanhecer, belonging to the database of the Atlas Linguístico do Amapá (ALAP). The research is grounded in the principles of Pluridimensional Geolinguistics, which seeks to highlight linguistic variation by considering the spatial and sociocultural dimensions of its speakers (Cardoso, 2010). The methodological *corpus* used in this study is drawn from the project Atlas Linguístico do Amapá (ALAP), which comprises 40 informants distributed across ten localities in the state of Amapá (1-Macapá, 2-Santana, 3-Mazagão, 4-Laranjal do Jari, 5-Pedra Branca do Amapari, 6-Porto Grande, 7-Tartarugalzinho, 8-Amapá, 9-Calçoene e 10-Oiapoque), Brazil. In each survey location, one male and one female informant from the first generation (18–30 years old) and one male and one female informant from the second generation (50–75 years old) were interviewed. With regard to linguistic cartography, the map displaying the designations for *amanhecer* was produced by the authors of this study using the graphic design software *Inkscape*. The results were presented based on the percentage distribution of the occurrences of each variant. Thus, the following frequencies were observed: *amanhecer* (54,36%), *raiar o dia* (13,04%), *dia clareando* (13,04%), *manhã* (6,52%), *nascer do sol* (4,35%), *aurora* (2,17%), *dia* (2,17%), and *sem respostas* (4,35%).

Keywords: Amanhecer. Amapá. Geolinguistics. Lexicon.

¹ Graduando na Universidade do Estado do Amapá, Macapá, AP, Brasil. Endereço eletrônico: matheusgo23571@gmail.com.

² Docente na Universidade Federal do Amapá, Macapá, AP, Brasil. Endereço eletrônico: romario.duarte@unifap.br.

Introdução

No passado, o estudo do léxico era tradicionalmente difundido atrelado à linguística histórica, por meio da análise de documentação escrita (González, 2002). Ou seja, a lexicologia tinha como característica principal uma visão diacrônica, visto que havia como foco compreender o surgimento de novas palavras ou, quando já extintas, identificar o momento em que decaiu o seu uso (González, 2002). Pouco se buscou compreender o dinamismo por trás das ocorrências, tais como sua expansão relacionada ao espaço geográfico e aos fatores socioculturais nos quais o léxico das línguas em uso esteve inserido. Outra linha de estudo que muitas vezes se entrelaça com a lexicologia, ao investigar a origem e a história das palavras, é a etimologia. De acordo com Viaro (2014), a etimologia tem como objetivo compreender o surgimento e o sentido histórico das palavras. Desse modo, ambos os estudos são pertinentes para caracterizar o léxico e, de modo geral, a língua.

A língua é um sistema vivo que se propaga e se modifica em um ritmo indeterminado, haja vista que toda comunidade possui o seu próprio tempo de instabilidade linguística. É dessa forma que os estudos geolinguísticos se fazem necessários nesse contexto, pois registram o léxico em uso e viabilizam os resultados de forma espacializada, i.e. em formato de mapas ou cartas linguísticas. Por definição, de acordo com Cardoso (2010), a geolinguística pluridimensional é o método por excelência da dialetologia, que registra as variantes encontradas, valoriza os aspectos geográficos das localidades e as características extralinguísticas dos informantes.

Entre os conceitos lexicais que a geolinguística brasileira se propõe a investigar, destaca-se a questão 22 (*amanhecer*)³ do Questionário Semântico-Lexical (QSL), desenvolvido pelo Comitê do Projeto Atlas Linguístico do Brasil (2001). Podemos dizer que para os estudos descritivos, o tempo tem uma carga lexical de alto valor, visto que está integrado quotidianamente à vida das pessoas, portanto, há grande possibilidade de surgirem variantes distintas para o mesmo fenômeno em questão. Na visão de Elias (1988), a língua portuguesa dispõe de diferentes palavras para caracterizar o tempo na fala. Dessa

³ Um dos conceitos contidos no *questionário semântico-lexical* (Comitê Nacional, 2001, p. 23) do *Projeto Atlas Linguístico do Brasil* (Cardoso et al., 2014) que foi utilizado pelos inquiridores do *Projeto Atlas Linguístico do Amapá* (Razky; Ribeiro; Sanches, 2017).

forma, cabe ao falante realizar a escolha lexical referente ao contexto social em que sua comunidade está inserida, visto que o tempo cumpre a função de orientá-lo diante do mundo em que vive (Freitas; Margotti, 2012). De modo geral, o fenômeno do dia denominado *amanhecer* é perfeitamente compreensível para os falantes de língua portuguesa, mas há, certamente, outras denominações genéricas relacionadas diretamente ao dia ou ao sol no início da manhã, tais como: *raiar do dia/sol*, *nascer do dia/sol*, *surgir o dia/sol* etc. Outra variante bastante comum é a palavra *aurora*, que, na língua portuguesa, tem vários valores semânticos quando empregada em diferentes sentenças, além de ser normalizada nos principais dicionários de língua portuguesa e ser sinônimo de *amanhecer*.

Diante desse fato, esta pesquisa integra os estudos geolinguísticos, com o propósito de proporcionar uma descrição lexical do português falado em contexto amazônico. De modo mais claro, visa esclarecer as denominações utilizadas pelos moradores do estado do Amapá, localizado na Região Norte do Brasil, que integram o quadro de informantes do *Atlas Linguístico do Amapá* (ALAP) (Razky; Ribeiro; Sanches, 2017). A questão-alvo está inserida no campo semântico-lexical *Astros e Tempo*, mais precisamente, o conceito *amanhecer* e suas eventuais variantes registradas. Além disso, também apresentamos um breve estudo histórico no que diz respeito às variantes padrão provenientes desse conceito, i.e. aquelas que estão dicionarizadas e, portanto, são reconhecidas em contextos gerais da língua portuguesa. Para esta análise, recorreremos às reflexões teóricas da lexicologia e da etimologia.

Dessa forma, nas próximas páginas apresentamos: (I) uma seção sobre lexicologia e etimologia; (II) uma seção sobre as variantes padrão (*amanhecer*, *raiar* e *aurora*), separadas, cada uma, por subseções; (III) a metodologia desta pesquisa; (IV) os resultados; (V) e a conclusão do estudo.

Léxico e etimologia: Uma relação com os estudos geolinguísticos

O estudo do léxico constitui um campo de investigação fundamental para a compreensão das línguas naturais, pois envolve a análise das unidades lexicais que os falantes utilizam para nomear, organizar e interpretar a realidade que os cerca (Abbade, 2011). Nesse contexto, duas áreas assumem papel central: a lexicologia e a etimologia. A

primeira dedica-se ao estudo científico do léxico, investigando sua estrutura, organização e funcionamento dentro do sistema linguístico, bem como suas relações semânticas e socioculturais (Abbade, 2011). Já a etimologia concentra-se na origem e na história das palavras, examinando os processos de formação, transformação fonológica e mudança de sentido que ocorrem ao longo do tempo (Isidoro de Sevilla, 2004; Baglioni, 2016; Viaro, 2011). Embora possuam objetivos específicos, ambas as áreas se articulam de forma complementar, uma vez que o conhecimento da origem das palavras contribui para a compreensão de sua organização no léxico e de sua difusão nas diferentes variedades linguísticas. Dessa maneira, o diálogo entre lexicologia e etimologia permite ampliar a compreensão acerca dos processos históricos e culturais que influenciam a formação e a evolução do vocabulário de uma língua, além de contribuírem e serem áreas requisitadas para os estudos em linguística, como a geolinguística, por exemplo.

Lexicologia

A lexicologia é o ramo da Linguística responsável pelo estudo científico do léxico, compreendido como o conjunto de unidades lexicais que compõem uma língua. Apesar de atualmente ocupar um espaço significativo nas pesquisas linguísticas, o estudo do léxico nem sempre recebeu a mesma atenção por parte dos estudiosos da linguagem. Conforme observa Abbade (2011), durante muito tempo os estudos lexicais foram considerados secundários, sendo frequentemente negligenciados pelos linguistas, que priorizavam a análise de estruturas consideradas mais sistemáticas da língua. De acordo com Abbade (2011), até o século XIX os linguistas concentravam seus esforços principalmente nas dimensões fonéticas, morfológicas e sintáticas das línguas, pois representavam estruturas internas mais relevantes para a descrição linguística. Nesse cenário, o léxico era frequentemente tratado apenas como um inventário de palavras, sem que fossem exploradas de maneira mais ampla suas relações semânticas, históricas e socioculturais que o constituíam.

Entretanto, com o desenvolvimento da linguística histórica no final do século XIX, o interesse pelo estudo do léxico começou a se intensificar. Segundo González (2002), foi nesse período que diversas pesquisas passaram a integrar a análise lexical à investigação

de documentos escritos, buscando compreender as transformações que as palavras sofrem ao longo do tempo. Paralelamente, outras disciplinas surgiram para se dedicar ao estudo das unidades lexicais, como a onomasiologia e a geolinguística, ambas interessadas nos processos de nomeação e na distribuição das variantes lexicais nas diferentes regiões e comunidades.

No âmbito da lexicologia, torna-se importante distinguir alguns conceitos fundamentais relacionados às unidades que compõem o léxico. Conforme explica Abbade (2011), é necessário diferenciar *palavra*, *vocabulo* e *lexia*. A *palavra* constitui um termo genérico que integra o *vocabulário* de uma língua, representando o conjunto de formas linguísticas utilizadas por uma determinada comunidade de falantes. O *vocabulário*, por sua vez, refere-se ao repertório lexical efetivamente empregado por esses falantes em contextos comunicativos específicos. A *lexia*, entretanto, apresenta um alcance conceitual mais amplo. Diferentemente da *palavra*, que pode ser compreendida como significação morfemática, i.e. gramatical, a *lexia* corresponde a uma unidade significativa do léxico que envolve também dimensões extralinguísticas. Nas palavras de Abbade (2011, p. 1334): “a *lexia* [...] é a unidade significativa do léxico de uma língua, ou seja, é uma palavra que tenha significado social”. Isso significa que, além de nomear objetos, ações ou conceitos, a *lexia* também carrega valores identitários, culturais e sociais associados aos contextos em que é utilizada. Assim, o estudo das *lexias* permite compreender não apenas os significados linguísticos das palavras, mas também as experiências culturais e históricas que influenciam seu uso em diferentes localidades.

Nesse sentido, como observa Biderman (1988, p. 88): “a atividade de nomear, isto é, a utilização de palavras para denominar os referentes extra-lingüísticos é específica da espécie humana”. Ou seja, por meio da linguagem, os indivíduos organizam e interpretam a realidade que os cerca, estabelecendo categorias que refletem suas experiências socioculturais. Dessa forma, o léxico de uma língua pode ser entendido como um reflexo da visão de mundo de seus falantes, uma vez que as palavras utilizadas para designar determinados objetos ou fenômenos estão diretamente relacionadas às práticas culturais e às necessidades comunicativas das comunidades linguísticas. Em outras palavras, língua, sociedade e cultura constituem uma tríade indissociável, i.e. não se explicam isoladamente, mas se constroem e se significam mutuamente em cada prática social.

Uma das abordagens teóricas que contribuem para compreender a organização interna do léxico é a teoria dos campos lexicais, desenvolvida por Coseriu (1977). Segundo essa perspectiva, as palavras de uma língua não existem de forma isolada, mas se organizam em conjuntos ou campos semânticos nos quais mantêm relações de interdependência. Isso denota que o significado de uma palavra é parcialmente definido pelas relações que estabelece com outras unidades lexicais pertencentes ao mesmo domínio semântico.

Os campos lexicais podem ser compreendidos, portanto, como agrupamentos de palavras que compartilham um mesmo universo de significado (Coseriu, 1977). Um exemplo simples é o campo lexical relacionado aos *fenômenos atmosféricos* (i.e. meteorológicos), que inclui palavras como chuva, neve, tempestade, granizo, entre outros. Dentro desse campo, cada palavra possui um significado específico, mas todas se relacionam semanticamente por fazerem parte de um mesmo domínio conceitual.

Essa perspectiva orienta a elaboração de questionários em pesquisas geolinguísticas, como no caso do Atlas Linguístico do Brasil (ALiB) (Cardoso *et al.*, 2014). Conforme descrevem os autores, as questões de natureza lexical presentes no questionário do projeto são organizadas em campos semânticos, o que permite investigar diferentes formas de denominação associadas a um mesmo conceito, conforme propõe Coseriu (1977). Essa organização facilita a comparação entre as respostas fornecidas pelos informantes em diferentes localidades, contribuindo para a identificação de padrões de variação lexical no território brasileiro.

Etimologia

A etimologia foi alvo de estudos desde a Grécia Antiga e ganhou destaque na Idade Média com os esforços de Isidoro de Sevilla no século VII. A partir da visão medieval, conforme conceitua Isidoro de Sevilla (2004, p. 310): “*etymologia est origo vocabulorum, cum vis verbi vel nominis per interpretationem colligitur*”⁴. Na visão moderna de Baglioni (2016), esse conceito está correto, mas requer uma ampliação aos moldes científicos atuais, acrescentando a importância da relação histórica das palavras para desvendar suas

⁴ Tradução: “Etimologia é a origem das palavras, quando o sentido do verbo ou do nome é compreendido por meio da interpretação”.

origens ao longo do tempo, i.e. a etimologia constitui um campo de investigação dedicado ao estudo da origem e da história das palavras. Seu objetivo principal, na visão de Viaro (2011) é compreender de que maneira as unidades lexicais se formam, se transformam e se difundem ao longo do tempo, considerando tanto aspectos linguísticos quanto fatores históricos e culturais. Para Viaro (2011), a etimologia investiga os processos de formação das palavras, bem como as mudanças fonológicas e semânticas que ocorrem no decorrer da evolução das línguas.

Nesse sentido, a análise etimológica permite identificar as relações históricas entre diferentes línguas e compreender os caminhos percorridos pelas palavras até sua incorporação em determinado sistema linguístico. No caso do português brasileiro, por exemplo, o léxico apresenta contribuições provenientes de diversas tradições linguísticas, incluindo o latim, o árabe, as línguas indígenas, línguas africanas, entre outras. Essas influências refletem os processos históricos de contato cultural que marcaram a formação da língua portuguesa no Brasil ao longo dos séculos.

Nos estudos geolinguísticos, a etimologia desempenha um papel relevante, pois auxilia na interpretação das variantes lexicais registradas em diferentes regiões. De maneira sucinta, é possível conceituar a geolinguística pluridimensional como a área da linguística que descreve a língua por meio de mapas ou cartas linguísticas, além de considerar simultaneamente fatores geográficos e sociais (Cardoso, 2010). Estudar a origem e história das palavras também faz parte dos moldes metodológicos de descrição da geolinguística, pois contribui na compreensão dos processos de difusão e transformação do léxico em uso.

Nesse contexto, os pesquisadores frequentemente recorrem aos dicionários de orientação etimológica como instrumentos de apoio para a análise dos dados coletados em campo. Esses recursos permitem verificar a origem histórica das palavras registradas, bem como identificar possíveis relações entre diferentes variantes lexicais. Assim, a etimologia contribui para ampliar a compreensão das dinâmicas históricas e socioculturais que influenciam a formação do léxico.

De acordo com Nascentes (1953), um dos pioneiros da geolinguística no Brasil, a variação linguística está profundamente relacionada às condições geográficas e sociais que caracterizam cada comunidade de fala. Dessa forma, determinadas palavras podem

permanecer em uso em algumas regiões enquanto desaparecem ou são substituídas em outras, refletindo os processos históricos e culturais que moldam o desenvolvimento das línguas.

Um exemplo que ilustra a importância da análise etimológica em estudos que descrevem a língua a partir da abordagem geolinguística, pode ser observado na pesquisa realizada por Santos *et al.* (2024) acerca da palavra *papagaio*. Nesse estudo, os autores investigam a origem da palavra para explicar sua associação com o brinquedo conhecido como *papagaio de papel*. A partir da consulta ao Dicionário Houaiss (2009), os pesquisadores identificaram que o termo possui origem árabe, na forma *babaghá* (ببغاء). Além disso, foi constatado que a modificação do fonema /b/ para /p/ ocorreu possivelmente por influência das línguas europeias, uma vez que o árabe não apresenta a consoante oclusiva bilabial surda em seu inventário fonológico. Do ponto de vista histórico, a palavra apresenta registros sucessivos nas formas *papagay*, *papagayo* e *papagaio*, indicando um processo gradual de adaptação fonológica e ortográfica ao longo do tempo. Segundo Santos *et al.* (2024), esse percurso histórico sugere que o termo foi incorporado ao português no contexto das intensas trocas comerciais e culturais que ocorreram entre os séculos XIII e XV. Ademais, a presença moura na Península Ibérica pode ter desempenhado um papel importante na introdução dessa palavra no português europeu, favorecendo sua posterior difusão nas diferentes variedades da língua portuguesa. Esse exemplo evidencia como a etimologia contribui para compreender não apenas a origem das palavras, mas também os processos históricos e socioculturais que influenciam a constituição do léxico.

O surgimento do sol no início do dia

A língua exerce papel fundamental na apreensão e na organização da realidade. Assim, as múltiplas formas de nomear fenômenos atmosféricos, a exemplo do *amanhecer*, revelam não apenas estruturas lexicais distintas, mas também valores culturais subjacentes. Embora o surgimento do sol no início do dia seja um fenômeno astronômico universal, é designado por diferentes denominações nas variedades do português em uso nas regiões do Brasil, as quais se distinguem quanto à formação morfológica, ao conteúdo semântico e às funções discursivas que desempenham.

No português, coexistem diferentes expressões para designar o início do dia, entre as quais se destacam *amanhecer*, *raiar do sol* e *aurora*, cada uma portadora de nuances semânticas específicas que orientam os falantes no contexto sociocultural em que se inserem (Freitas; Margotti, 2012). Dessa forma, a palavra *amanhecer* apresenta uma carga semântica centrada no processo gradual de transição da noite para o dia. A denominação *raiar do sol*, por sua vez, caracteriza-se por apresentar uma estrutura composicional transparente (verbo + substantivo), remetendo de forma direta aos feixes luminosos emitidos pelo astro ao surgir no horizonte. Já o termo *aurora* configura-se como uma das designações mais antigas e estilisticamente marcadas para o nascer do sol, cujo emprego extrapola a função descritiva e assume valor simbólico, evocando simultaneamente o início do dia e expressões poéticas em textos literários (Houaiss, 2003).

Amanhecer

A denominação *amanhecer* ocupa posição de destaque no léxico da língua portuguesa, uma vez que articula dimensões etimológicas, culturais e simbólicas que a qualificam como objeto relevante de investigação linguística. Conforme registra o Dicionário Houaiss (2003), trata-se de um verbo impessoal e defectivo, cuja conjugação se restringe à terceira pessoa em todos os tempos e modos; todavia, em usos figurados, admite flexão em quaisquer pessoas gramaticais.

Sob a perspectiva etimológica, *amanhecer* resulta da combinação do prefixo *a-* (acepção de aproximação), do substantivo *manhã* e do sufixo inceptivo *-ecer* (incoativo de manifestação com antecedência), aspectos linguísticos que são responsáveis por indicar mudança de estado ou passagem temporal (Houaiss, 2003). O Dicionário Houaiss (2003) registra, ainda, diferentes registros gráficos da palavra em sua transformação histórica, como: *manhecer*, *ameesçer* e *amaesçer*. Nascentes (1966) apresenta a palavra derivada do latim tardio (*admanescere*), forma reconstruída para justificar o valor inceptivo do sufixo *-ecer*. Para Houaiss (2003), trata-se de uma formação regular no próprio português, a partir de *manhã* (latim: *mane*) e do sufixo inceptivo *-ecer* (latim tardio: *-escere*).

No plano geolinguístico, o Comitê Nacional do Projeto *Atlas Linguístico do Brasil* reconhece *amanhecer* como a variante padrão para a designação de “a parte do dia

quando começa a clarear” (Comitê Nacional do Projeto ALiB, 2001, p. 23), sendo igualmente a forma mais difundida e socialmente reconhecida nas diversas regiões do país.

Raiar do sol

Raiar do sol caracteriza-se como expressão sinônima de *amanhecer*, sendo amplamente conhecida e utilizada em diferentes regiões do Brasil. Em determinados contextos socioculturais, a marcação do tempo pode estar associada à observação de fenômenos da natureza, em detrimento de horários convencionais, o que favorece o uso de expressões como “ao raiar do sol”, “depois que o sol raiar” e “esperar o sol raiar” no discurso cotidiano. Conforme discute Elias (1988), em diferentes sociedades a organização do tempo esteve historicamente vinculada à observação de fenômenos naturais, e.g. como o nascer e o pôr do sol, que funcionavam como referências temporais compartilhadas.

Acerca da perspectiva etimológica, o Dicionário Houaiss (2003) registra que o verbo *raiar* tem origem no substantivo latino *radius* (“raio, haste, vara”), ao qual se acrescenta o sufixo verbal *-ar*, responsável por indicar a ação de emitir ou projetar raios de luz. A associação do verbo *raiar* ao substantivo *sol*, mediada pela contração *do* (de + o), resulta em uma expressão de forte valor imagético, que remete diretamente à projeção dos primeiros feixes luminosos no início da manhã.

Nesse sentido, *raiar do sol* configura-se como uma denominação recorrente em comunidades tradicionais e em contextos nos quais a nomeação dos períodos do dia desempenha função orientadora das práticas sociais. Estudos sobre variação lexical no português brasileiro indicam que denominações desse tipo refletem não apenas escolhas linguísticas, mas também modos de apreensão da temporalidade associados à experiência social dos falantes (Freitas; Margotti, 2012). Além disso, a expressão remete ao momento em que os primeiros raios solares se projetam no horizonte, marcando simbolicamente o início de um novo dia.

Aurora

No português contemporâneo, *aurora* é empregada tanto para designar o alvorecer ou o amanhecer quanto em contextos de natureza literária. Conforme registra o Dicionário Houaiss (2003), o termo apresenta usos que extrapolam a descrição temporal do início do dia, assumindo valores simbólicos e poéticos. Na literatura de língua portuguesa, *aurora* constitui uma imagem recorrente, especialmente na tradição lírica, associada a ideias de pureza, juventude e renovação. No uso cotidiano, embora menos frequente do que denominações como *amanhecer* ou *raiar do sol*, a palavra *aurora* mantém-se em expressões cristalizadas, como “romper da aurora” e “antes da aurora”, amplamente registradas em textos literários, canções e provérbios, o que evidencia seu caráter estilisticamente marcado e sua permanência no léxico do português brasileiro.

No que tange à etimologia da palavra *aurora*, de acordo com Houaiss (2003), trata-se de um termo derivado do latim (*aurōra*) e sinônimo de *amanhecer*. Em congruência com Ernout e Meillet (1951), há uma antiga derivação da palavra *ab aurō* (proveniente do ouro), haja vista que Isidoro de Sevilla, importante estudioso e arcebispo visigodo do século VII, que desenvolveu a obra *Etymologiae*, i.e. enciclopédia com todo o conhecimento etimológico da época, expôs a seguinte frase, que reflete a elegância poética dessa denominação: *aurora dicitur ante solis ortum, ab eo quod ab igni solis tum aureo aer aurescit*⁵ (Ernout e Meillet, 1951). Segundo o dicionário Houaiss (2003), a palavra *aurora* também corresponde à deusa Eos, i.e. deusa da manhã na mitologia grega, que personificava a renovação e o ciclo de passagem da noite para o dia.

No contexto social, *aurora* tende a ganhar espaço em contextos muito específicos, como em regiões onde a tradição oral, a religiosidade popular e a literatura regional reforçam o uso de vocabulário poético para denominar fenômenos naturais. Haja vista que, ao observar em pesquisas de natureza geolinguística (atlas linguísticos) em diferentes regiões do Brasil, sua pouca ocorrência comparada à variante *amanhecer* é evidente.

⁵ Tradução: “A aurora é a chamada antes do nascer do sol, porque pelo fogo dourado do sol, o ar então se torna dourado”.

Metodologia do Atlas Linguístico do Amapá

A presente pesquisa utiliza como base metodológica os dados do Atlas Linguístico do Amapá (ALAP) (Razky, Ribeiro; Sanches, 2017). Tal abordagem fundamenta-se na tradição da geolinguística pluridimensional, conhecida pela incorporação de variáveis sociais, permitindo a análise da variação linguística não apenas em sua dimensão espacial, mas também em sua dimensão sociocultural (Cardoso, 2010).

Por sua vez, o Projeto ALAP fundamentou-se nos parâmetros metodológicos estabelecidos pelo Atlas Linguístico do Brasil (ALiB) (Cardoso et al., 2014), obra que se consolidou como referência na elaboração de atlas linguísticos no país. A partir desse modelo, o projeto passou a contemplar diferentes dimensões da variação linguística, a saber: a variação diatópica (variação geográfica dos pontos de inquéritos), a diassexual (variação por sexo biológico informado pelos colaboradores) e a diageracional (corresponde à idade dos colaboradores que é informada no formulário), assegurando uma abordagem mais abrangente e sistemática da diversidade linguística no estado do Amapá.

O Atlas Linguístico do Amapá surgiu em 2010 e teve início por meio de *workshops* desenvolvidos na Universidade Federal do Amapá (UNIFAP), em parceria com o professor Abdelhak Razky, da Universidade Federal do Pará (UFPA), que se responsabilizou por ser um dos coordenadores do projeto. Essa etapa se caracterizou como o início da propagação teórica da geossociolinguística, i.e. termo cunhado por Abdelhak Razky para designar a abordagem geolinguística articulada às variáveis sociolinguísticas, evidenciando a integração entre a dimensão espacial da língua e os fatores sociais que condicionam sua variação. Dessa forma, em 2011, iniciaram-se inquéritos experimentais com o objetivo de proporcionar um treinamento prático aos acadêmicos que se tornariam os futuros inquiridores da pesquisa. Após esse período, a coleta dos dados válidos para a composição do *corpus* foi realizada entre 2012 a 2014 pelos professores e acadêmicos da Universidade Federal do Amapá (Razky, Ribeiro e Sanches, 2017). Dessa forma, de acordo com Razky, Ribeiro e Sanches (2017), o ALAP seguiu três etapas:

1ª etapa: formação e treinamento dos membros do grupo;

2ª etapa: realização dos inquéritos experimentais, treinamento para transcrição fonética, aplicação da pesquisa *in loco* (questionários fonético-fonológico e semântico-lexical), junto a 40 informantes distribuídos pelos dez pontos de inquéritos;

3ª etapa: transcrição dos dados, revisão das transcrições fonéticas, confecção das cartas e mapeamento dos dados registrados, tendo em vista a sistematização, organização e publicação dos resultados (Razky; Ribeiro; Sanches, 2017, p. 36).

Assim como informado pelos autores, o Atlas Linguístico do Amapá contou com uma rede de pontos formado por dez municípios. É correto afirmar que o estado do Amapá é composto por 16 localidades (Amapá, Calçoene, Cutias, Ferreira Gomes, Itaubal, Laranjal do Jari, Macapá, Mazagão, Oiapoque, Pedra Branca do Amapari, Porto Grande, Pracuúba, Santana, Serra do Navio, Tartarugalzinho, Vitória do Jari), no entanto, foram selecionadas apenas dez pontos para a pesquisa (Macapá, Santana, Mazagão, Laranjal do Jari, Pedra Branca do Amapari, Porto Grande, Tartarugalzinho, Amapá, Calçoene e Oiapoque) por motivos de enquadramento metodológico. De acordo com Razky, Ribeiro e Sanches (2017):

[...] foram considerados critérios diversificados na seleção da rede de pontos de inquéritos, tais como a densidade demográfica; a distribuição espacial das localidades, a fim de que um ponto não ficasse muito próximo de outro; aspectos históricos, ligados ao tempo de origem de cada município; a importância econômica e sociocultural da localidade (Razky; Ribeiro; Sanches, 2017, p. 37).

Diante desse fato, os municípios de Vitória do Jari, Serra do Navio, Ferreira Gomes, Cutias, Pracuúba e Itaubal não foram incluídos na rede de pontos de inquérito por não satisfazerem os critérios adotados no estudo.

Para a seleção dos informantes, os inquiridores do Atlas Linguístico do Amapá desenvolveram uma ficha catalográfica para registrar o perfil dos participantes e verificar se estavam aptos a colaborar com a pesquisa. De acordo com Sanches (2025), os seguintes critérios foram adotados:

a) ter nascido no município; b) ser filho de pais nascidos na região; c) não ter morado em outro Estado ou Região por mais de um ano; d) ter nível de instrução escolar variando de analfabeto ao Ensino Fundamental completo; e) possuir boas condições de saúde e de fonação; e f) ter disponibilidade para a entrevista (Sanches, 2025, p. 67).

Desse modo, foram selecionados 40 informantes distribuídos entre as dez localidades, conforme faixa etária e sexo, para manter o equilíbrio e a organização metodológica. Em cada ponto de inquérito houve quatro informantes por localidade: dois do sexo masculino e dois do sexo feminino, sendo um homem e uma mulher na faixa etária de 18 a 30 anos e um homem e uma mulher na faixa etária de 50 a 75 anos.

Após a seleção das localidades e dos informantes, os dados foram coletados com o auxílio dos seguintes questionários desenvolvidos pelo Comitê do Atlas Linguístico do Brasil (2001): fonético-fonológico (QFF) que é composto por 159 questões fechadas; e o questionário semântico-lexical (QSL) com 202 perguntas abertas, distribuídas em 14 campos semânticos. De maneira geral, as entrevistas tiveram duração média de duas horas, com os dados registrados a partir de um gravador de voz, em formato MP3, e organizados no *Soft Cool Edit Pro 2.1* (Sanches, 2025). Para a análise dos dados, foram consideradas as respostas múltiplas fornecidas pelos informantes durante a aplicação do Questionário Semântico-Lexical, i.e. os casos em que um mesmo colaborador apresentou mais de uma denominação para o conceito investigado. Dessa forma, todas as variantes foram contabilizadas e incluídas na análise quantitativa, seguindo os procedimentos adotados na cartografia lexical do Projeto ALAP. Em razão disso, o número total de ocorrências pode ser superior ao número de informantes entrevistados, conforme veremos na carta referente ao conceito *amanhecer*.

O último período de análise e cartografia ocorreu entre os anos de 2023 e 2024, por meio do projeto de pesquisa denominado *Atlas Linguístico do Amapá: Fase III*, do qual participaram acadêmicos da Universidade Federal do Amapá (UNIFAP) e da Universidade do Estado do Amapá (UEAP), sob a coordenação dos professores Romário Duarte Sanches e Celeste Maria da Rocha Ribeiro, ambos da Universidade Federal do Amapá. Os integrantes cartografaram um total de 86 conceitos lexicais – que já haviam sido coletados pela primeira equipe do ALAP, formada em 2010 – distribuídos pelos campos semânticos apresentados no Quadro 1.

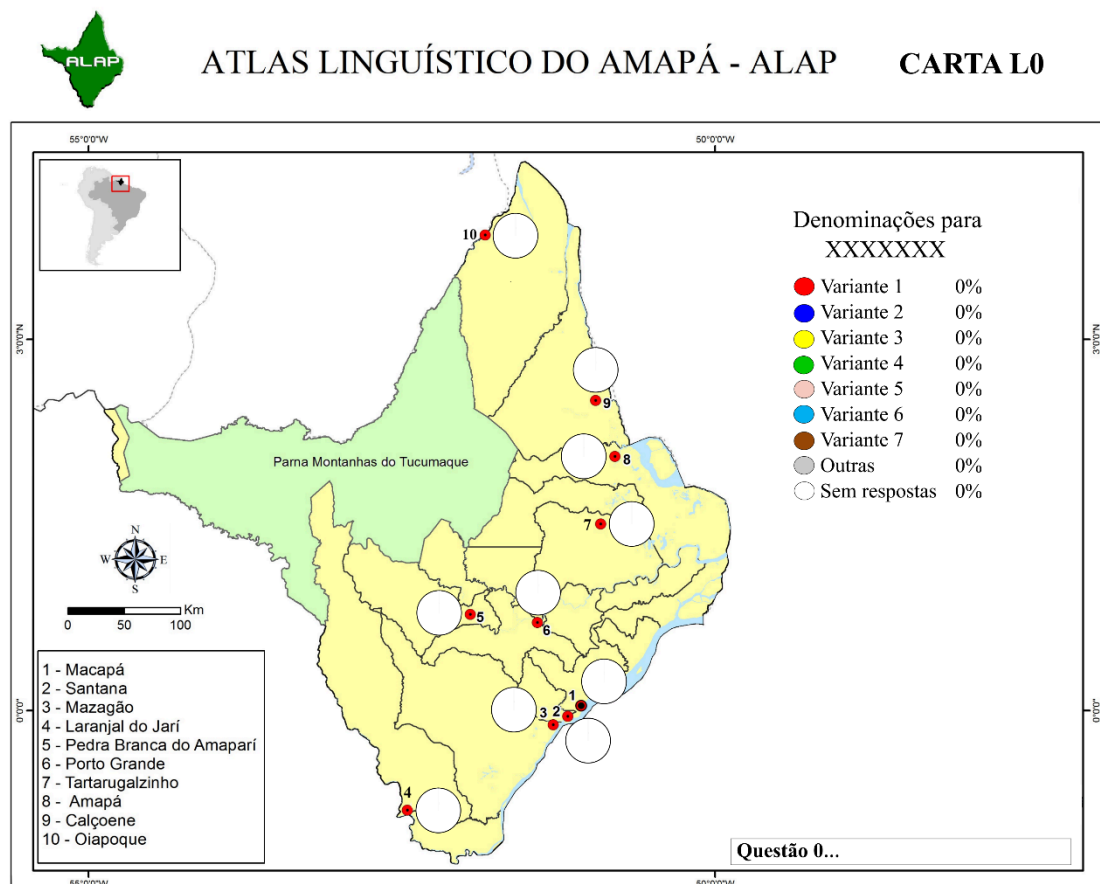
Quadro 1. Lista dos campos semântico-lexicais cartografados pelo Projeto ALAP Fase III

Campo Semântico	Número de Itens Cartografados
Acidentes geográficos	01
Fenômenos atmosféricos	04
Astros e tempo	12
Atividades agropastoris	16
Fauna	08
Corpo humano	09
Ciclos da vida	08
Religião e crenças	04
Jogos e diversões infantis	06
Habitação	03
Alimentação e cozinha	05
Vestuário e acessórios	03
Vida urbana	07
Total Geral	86 conceitos

Fonte: Elaborado pelos autores.

A cartografia realizada pela equipe formada por graduandos da UNIFAP e da UEAP teve duração de 12 meses, com término em 2024, e, no momento, segue em processo de triagem pelos coordenadores do projeto para que as cartas linguísticas sejam publicadas no Atlas Linguístico do Amapá vol. 2. Diferentemente do primeiro volume, no qual a cartografia dos conceitos foi elaborada por meio do *software CorelDRAW X5*, na fase III do ALAP optou-se pela utilização do editor de gráficos vetoriais *Inkscape*. Essa mudança metodológica evidencia não apenas uma atualização nos procedimentos técnicos adotados pelo projeto, mas também a incorporação de uma ferramenta de código aberto, amplamente reconhecida por sua versatilidade e precisão na produção cartográfica. Além disso, o uso do *Inkscape* possibilita maior flexibilidade na edição, padronização visual das cartas linguísticas e potencial de reprodutibilidade dos dados, contribuindo, assim, para o aprimoramento da qualidade gráfica e científica do atlas. Na figura 1, apresenta-se a carta-base do ALAP, atrelada das informações acrescidas no editor de gráficos vetoriais *Inkscape*.

Figura 1. Carta-base do Atlas Linguístico do Amapá



Fonte: Elaborado pelos autores.

Segundo Sanches (2025), a equipe do projeto ALAP, em sua fase III, ampliou o número de variantes lexicais consideradas no processo de cartografia linguística. Diferentemente do primeiro volume, que priorizou apenas as cinco variantes mais frequentes, essa nova etapa passou a incluir até sete variantes, incorporando também formas menos recorrentes. Entre essas, destacam-se aquelas registradas uma única vez, mas que apresentam indícios de estarem em processo de desuso em determinadas comunidades de fala (Sanches, 2025).

Outra alteração relevante a ser destacada refere-se às modificações na tabela de descrição das variantes, que deixam de ser representadas por meio de símbolos fonéticos (i.e. utilização do Alfabeto Fonético Internacional – IPA) e passam a ser registradas sob a forma de transcrição lexicográfica. Tal mudança visa tornar a leitura mais acessível, não

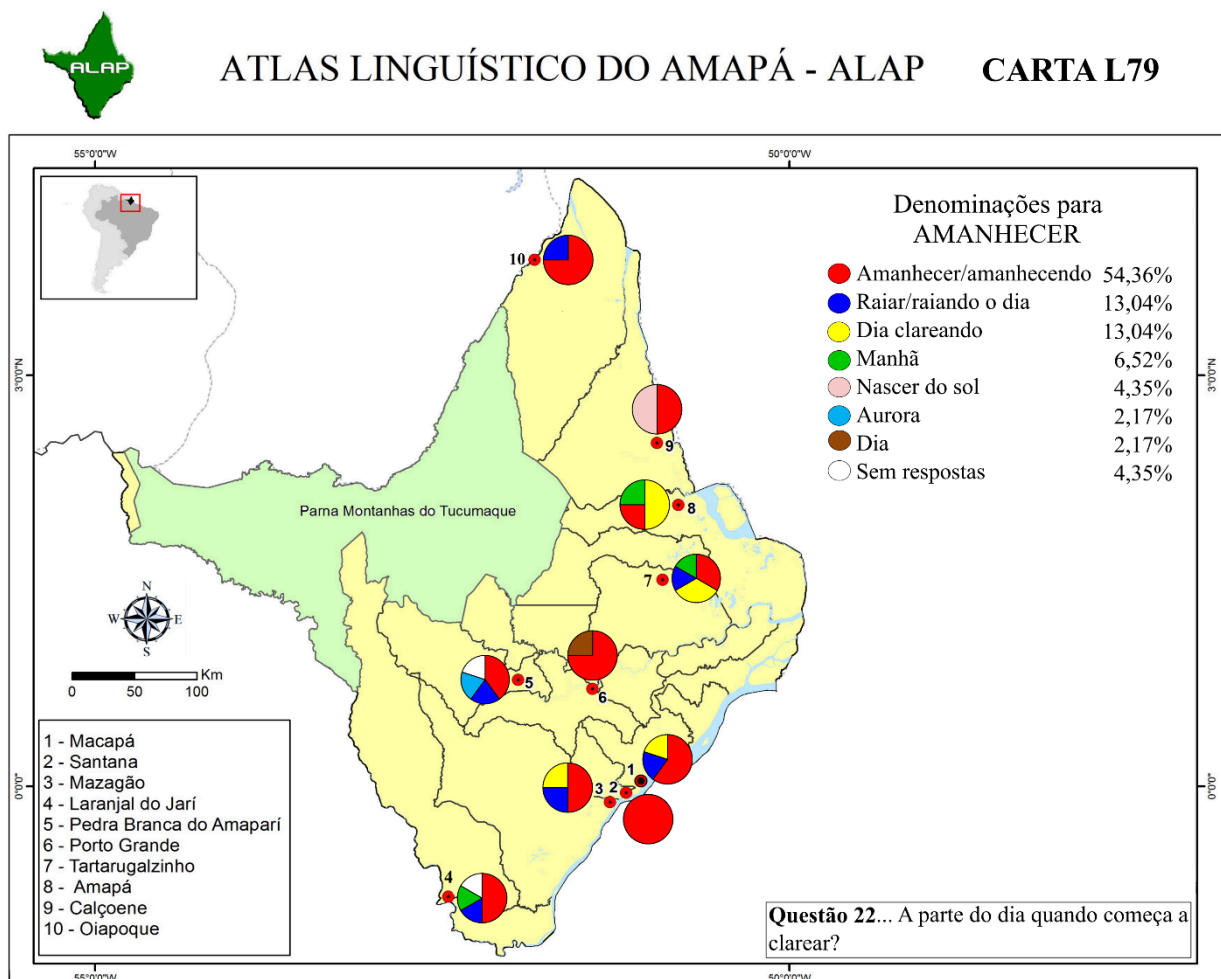
apenas para especialistas, mas também para estudiosos e interessados na variedade amapaense do português, ampliando, assim, o alcance e a compreensão dos dados apresentados.

Alguns dados inéditos, que integrarão o segundo volume do Atlas Linguístico do Amapá, já se encontram disponíveis em formato de artigo científico. Segundo Sanches (2025), há, até o momento, duas publicações que contemplam a descrição de campos semânticos distintos: *corpo humano* e *habitação*. O primeiro, referente ao campo semântico *corpo humano*, que foi descrito por Santos e Sanches (2024), em artigo publicado na revista *Muiraquitã*, no qual são apresentadas e analisadas nove cartas lexicais. No que diz respeito ao campo semântico *habitação*, a descrição linguística foi realizada por Santos, Carvalho e Sanches (2024), em trabalho publicado na revista *Acta Semiótica et Lingvistica*, abrangendo a análise de três cartas lexicais. Dessa maneira, a divulgação prévia desses dados em formato de artigo não apenas antecipa resultados do atlas em elaboração, mas também fortalece o diálogo acadêmico e contribui para a consolidação de uma base consistente, fundamental para a compreensão da variação lexical do português amapaense.

Resultados

Dentre os 86 conceitos lexicais cartografados pela equipe do ALAP Fase III, para este trabalho apresentamos as respostas referentes ao conceito *amanhecer*, que representa a questão 22 (“... a parte do dia quando começa a clarear?”) do *questionário semântico-lexical* (Comitê Nacional, 2001, p. 23). Para este conceito, foram obtidas as seguintes variantes: *amanhecer* com 25 ocorrências (54,36%), *raiar o dia* e a variante *dia clareando* com seis ocorrências (13,04%), *manhã* com três ocorrências (6,52%), *nascer do sol* com duas menções (4,35%), *aurora* e *dia* com uma ocorrência cada (2,17%), além de dois informantes que não conseguiram responder à questão, representando a legenda denominada *sem respostas* (4,35%), conforme apresentado na figura 2.

Figura 2. Carta L79-Amanhecer



Fonte: Elaborado pelos autores.

A partir da figura 2, foi possível observar o número significativo de lexias informadas pelos colaboradores, abrangendo denominações padronizadas e também formas mais genéricas. Dessa forma, a variante predominante, i.e. a variante informada em todas as localidades, foi *amanhecer*, que apresentou maior ocorrência na localidade 2-Santana e menor frequência na localidade 8-Amapá. No que diz respeito à variante *raiar o dia*, sua ocorrência esteve presente nos pontos 1-Macapá, 3-Mazagão, 4-Laranjal do Jari, 5-Pedra Branca do Amapari, 7-Tartarugalzinho e 10-Oiapoque. Quanto à variante *dia clareando*, embora tenha ocorrido em quatro pontos de inquérito (1-Macapá, 3-Mazagão, 7-Tartarugalzinho e 8-Amapá), evidenciou uma dispersão geográfica que não estabelece

relação do uso dessa variante entre os pontos de inquéritos. Acerca da variante *manhã*, houve menções nos pontos 7-Tartarugalzinho e 8-Amapá que são próximos geograficamente, mas também apresentou ocorrência dispersa na localidade 4-Laranjal do Jari, ponto de inquérito longínquo ao comparar com os dois primeiros. Para a variante *nascer do sol*, suas duas ocorrências foram registradas no ponto 9-Calçoene, enquanto a única ocorrência da variante *aurora* foi registrada na localidade 5-Pedra Branca do Amapari. Para a variante *dia*, a mais genérica entre todas as respostas registradas, houve registro único na localidade 6-Porto Grande. Quanto à ausência de respostas, somente os pontos 4-Laranjal do Jari e 5-Pedra Branca do Amapari tiveram informantes que não responderam à questão, com duas ausências. De maneira geral, os dados revelaram que, para este conceito, os colaboradores do ALAP reconhecem e utilizam a variante padronizada *amanhecer*, além de diferentes lexias relacionadas aos primeiros feixes luminosos do sol no início do dia.'

Em relação à variação diassexual e diageracional, dados que não estão evidentes na figura 1, foi elaborada uma tabela para auxiliar e facilitar a interpretação. Nesta tabela constam as variantes, o sexo e a faixa etária dos informantes, além das respectivas ocorrências. Para a ausência de resposta na fala do informante, foi adotado o símbolo (-), conforme consta na tabela 1.

Tabela 1. Variação diassexual e diageracional

Variantes	Mulheres (18-30 anos)	Mulheres (50-75 anos)	Homens (18-30 anos)	Homens (50-75 anos)
	Ocorrências	Ocorrências	Ocorrências	Ocorrências
Amanhecer	9	5	7	4
Raiar o dia	-	3	-	3
Dia clareando	-	2	2	2
Manhã	1	1	-	1
Nascer do sol	-	-	1	1
Aurora	-	1	-	-
Dia	-	1	-	-
Sem respostas	-	-	1	1

Fonte: Elaborado pelos autores.

De acordo com a tabela 1, a variante *amanhecer* ocorreu na fala dos homens e das mulheres das duas faixas etárias, com maior ocorrência entre os informantes de 18 a 30

anos (16 ocorrências) em comparação aos informantes de 50 a 75 anos (nove ocorrências). No que se refere à variante *raiar o dia*, não houve ocorrências registradas na fala das mulheres e dos homens pertencente a primeira faixa etária (18-35 anos); em contrapartida, houve três ocorrências na fala dos homens e três ocorrências na fala das mulheres, ambos os sexos pertencentes a segunda faixa etária (50-75 anos). Quanto à variante *dia clareando*, houve maior frequência na fala dos homens (quatro ocorrências) em comparação com as mulheres (duas menções). Para a variante *manhã*, sua menção se deu na fala de um colaborador do sexo masculino (uma ocorrência) e das mulheres (duas menções). No que tange à variante *nascer do sol*, houve ocorrência na fala dos homens (duas menções) e total ausência na fala das mulheres. Quanto às variantes *aurora* e *dia*, ambas ocorreram somente na fala das mulheres que correspondem à faixa etária de 50 a 75 anos, com uma ocorrência registrada para cada uma. Em relação à ausência de respostas, apenas os homens não conseguiram nomear o conceito, com duas omissões, uma na faixa etária mais jovem e outra na faixa etária mais velha.

Conclusão

A análise das variantes lexicais empregadas para designar o fenômeno *amanhecer* no estado do Amapá, a partir dos dados do Atlas Linguístico do Amapá (ALAP) (Razky; Ribeiro; Sanches, 2017), evidenciou a riqueza e a dinamicidade do léxico regional, reafirmando o papel da geolinguística pluridimensional como método privilegiado para o registro e a interpretação da variação linguística em sua dimensão espacial e social (Cardoso, 2010). Os resultados demonstraram que, embora *amanhecer* se configure como a variante predominante, outras formas, como *raiar o dia*, *dia clareando*, *manhã* e *aurora*, coexistem no uso, refletindo diferentes estratégias de nomeação associadas à experiência sociocultural dos falantes.

Essa diversidade lexical confirma a perspectiva de Abbade (2011), segundo a qual o léxico deve ser compreendido como um sistema em constante transformação, profundamente relacionado às práticas sociais e culturais das comunidades linguísticas. Do mesmo modo, os dados geolinguísticos desta pesquisa corroboram com o método proposto por Coseriu (1977) acerca da organização do léxico em campos semânticos, uma

vez que as variantes identificadas se articulam em torno de um mesmo domínio conceitual, i.e. o campo semântico-lexical *astros e tempo* proposto pelo Comitê do Atlas Linguístico do Brasil (2001), mantendo entre si relações de proximidade semântica e funcional.

A ocorrência de variantes menos frequentes, como *aurora*, especialmente entre informantes mais velhos, evidencia a permanência de formas lexicalmente marcadas, o que pode ser interpretado à luz das reflexões de Nascentes (1953), que relaciona a variação linguística às condições históricas, geográficas e sociais das comunidades de fala. Nesse sentido, tais formas não apenas designam um fenômeno natural, mas também preservam traços da tradição cultural e da memória coletiva, conforme também sugerido por Viaro (2011), ao destacar que o estudo etimológico permite compreender os percursos históricos e as transformações semânticas das palavras.

Dessa forma, o presente estudo contribui para ampliar a compreensão da variação lexical no português brasileiro, especialmente em contextos amazônicos, historicamente menos contemplados em investigações de maior abrangência. Por fim, ressalta-se a importância de iniciativas como o ALAP para a documentação sistemática das variedades linguísticas regionais, recomendando-se, para pesquisas futuras, a incorporação de outras variáveis socioculturais, e.g. como a dimensão diastrática, bem como o desenvolvimento de microatlas em comunidades específicas (i.e. comunidades ribeirinhas, indígenas e quilombolas), a fim de aprofundar a análise das relações entre léxico, cultura e território.

Referências

- ABBADE, Celina Márcia de Souza. A lexicologia e a teoria dos campos lexicais. **Cadernos do CNLF**. Rio de Janeiro. v. 15, n. 5, p 1332-1343, 2011.
- BAGLIONI, Daniele. **L'etimologia**. Roma: Carocci editore, 2016.
- BIDERMAN, Maria Tereza Camargo. Dimensões da palavra. **Filologia e Linguística Portuguesa**. São Paulo, v. 1, n. 2, p. 81-118, 1998.
- CARDOSO, Suzana Alice Marcelino. **Geolinguística: tradição e modernidade**. São Paulo: Parábola, 2010.
- CARDOSO, Suzana Alice Marcelino et al. **Atlas Linguístico do Brasil**. V.1. Londrina: Eduel, 2014.

COMITÊ NACIONAL DO PROJETO ALiB. **Atlas Linguístico do Brasil: questionário 2001 / Comitê Nacional do Projeto ALiB**. Londrina: Editora UEL, 2001.

COSERIU, Eugenio. **Teoría del lenguaje y lingüística general**: estudios de teoría e historia de la lingüística. Madrid: Gredos, 1977.

ELIAS, Norbet. **Sobre o tempo**. Rio de Janeiro: Jorge Zahar, 1988.

ERNOUT, Alain; MEILLET, Antoine. **Dictionnaire étymologique de la langue latine**. Paris: Klincksieck, 1951.

FREITAS, Paulo Garcia; MARGOTTI, Wessling Felício. c. **Caligrama**. Belo Horizonte, v. 17, n. 2, p. 209-227, 2012.

GONZÁLEZ, Manuel González. A dialectoloxía ó servicio do estudo da estratigrafia do léxico. In: ÁLVARES, Rosário et al. (Org.). **Dialectoloxía e léxico**. Santiago de Compostela: Consello de Cultura Galega: Instituto de Língua Galega, p. 29-40, 2002.

HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa**. Lisboa: Círculo Leitores, 2003.

HOUAISS, Antônio; VILLAR, Mauro de Salles; FRANCO, Francisco Manoel de Mello. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2009.

ISIDORO DE SEVILHA, San. **Etimologías**. Edición bilingüe. Texto latino, versión española y notas por José Oroz Reta e Manuel-Antonio Marcos Casquero; introducción general por Manuel Cecilio Díaz y Díaz. Madrid: Biblioteca de Autores Cristianos, 2004. Disponível em: <https://archive.org/details/etimologias-san-isidoro-de-sevilla/page/n7/mode/2up>. Acesso em: 19 mar. 2026.

NASCENTES, Antenor. **Geografia dos Dialectos Brasileiros**. Rio de Janeiro: Livraria Acadêmica, 1953.

NASCENTES, Antenor. **Dicionário etimológico resumido da língua portuguesa**. Rio de Janeiro: Instituto Nacional do Livro, 1966.

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

SANCHES, Romário Duarte. As fases do Atlas Linguístico do Amapá: uma breve historiografia. **Revista de Letras Norte@mentos**, v. 18, n. 54, p. 63-85, 2025.

SANTOS, Matheus Gomes dos et al. Denominações para papagaio de papel no Amapá. **SIGMA**. Macapá, v. 5, n. 6, p. 09-25, 2024.

SANTOS, Matheus Gomes dos; CARVALHO, Mônica dos Santos; SANCHES, Romário Duarte. Mapeamento lexical do português falado no estado do Amapá referente ao campo semântico *habitação*. **Revista Acta Semiótica et Lingvistica**, v. 30, n. 1, p. 61-77, 2024.

SANTOS, Matheus Gomes dos; SANCHES, Romário Duarte. Mapeamento lexical do português falado no Amapá: um estudo sobre o campo semântico “corpo humano”. **Muiraquitã - Revista de Letras e Humanidades**, v. 12, n. 1, p. 06-28, 2024.

VIARO, Mário Eduardo. **Etimologia**. 2. ed. São Paulo: Contexto, 2011.

Sobre os autores

Matheus Gomes dos Santos

Orcid: <https://orcid.org/0009-0004-2903-122X>

Graduado em Letras pela Universidade do Estado do Amapá (UEAP) e, atualmente, mestrando em Letras (Linguística) pela Universidade Federal do Amapá (UNIFAP).

Romário Duarte Sanches

Orcid: <https://orcid.org/0000-0002-0571-303X>

Doutor em Letras (Linguística) pela Universidade Federal do Pará (UFPA). Atualmente, é professor do Colegiado de Letras da Universidade Federal do Amapá (UNIFAP).