

Clavícula, cantareira e saboneteira: as variedades do português formosense

Clavicle, Cantareira, and Saboneteira: the varieties of Formosan Portuguese

Karina de Jesus Araujo*

Universidade do Estado de Mato Grosso, Sinop, MT, Brasil

Manoel Mourivaldo Santiago-Almeida**

Universidade de São Paulo, São Paulo, SP, Brasil

Resumo: Esta pesquisa tem como objetivo analisar a forma “clavícula” a partir do(s) falar/falares dos moradores/migrantes de Formoso do Araguaia - TO, traçando representações sociolinguísticas no contexto semântico-lexical dos que nasceram ou vivem nessa comunidade. A questão central é compreender de que maneira essa lexia varia. Utiliza-se a abordagem qualitativa, fundamentada na Dialectologia Pluridimensional e Relacional de Thun (2010) e na técnica dos três passos: *perguntar, insistir e sugerir*, aplicando a questão 106 do QSL: *Como se chama o osso que vai do pescoço até o ombro?* A coleta foi dividida em dois pontos de inquérito: urbano e rural, com a participação de 48 informantes distribuídos em quatro grupos que representam diferentes variedades do português: maranhense (PM), gaúcho (PG), caipira (PC) e ribeirinho (PR) além de duas gerações, os mais jovens (GI) e os mais velhos (GII), abrangendo ambos os sexos. As análises se concentraram na ocorrência, frequência e divergência das formas utilizadas para referir-se à *clavícula*. Os resultados foram documentados em mapas polifórmicos e de status da forma. Assim, esta pesquisa contribui ao identificar as lexias predominantes na comunidade, além de oferecer *insights* relevantes para a sociedade em geral e para os estudos dialetais em todo o país.

Palavras-chave: Clavícula. Variante. Cantareira. Saboneteira.

Abstract: This research aims to analyze the term “clavícula” (clavicle) based on the speech of residents/migrants of Formoso do Araguaia - TO, tracing sociolinguistic representations in the semantic-lexical context of those born or living in this community. The central question is to understand how this lexeme varies. A qualitative approach is used, grounded in Thun’s (2010) Pluridimensional and Relational Dialectology, and employing the three-step technique: asking, insisting, and suggesting, applying question 106 from the QSL: *What is the name of the bone that goes from the neck to the shoulder?* Data collection was divided into two inquiry points: urban and rural, involving 48 informants distributed across four groups representing different varieties of Portuguese: Maranhense (PM), Gaúcho (PG), Caipira (PC), and Ribeirinho (PR), as well as two generations, the younger (GI) and the older (GII), encompassing both genders. Analyses focused on the occurrence, frequency, and divergence of the forms used to refer to the clavicle. Results were documented in polyformic and form status maps. Thus, this research contributes by identifying the predominant lexemes in the community, as well as offering relevant insights for society at large and for dialectal studies throughout

* Professora da Educação Básica, Secretaria de Educação do Estado de Mato Grosso, Doutoranda, Universidade de São Paulo - USP, São Paulo, SP, Brasil; kjaraujo@usp.br

** Professor Titular, Universidade de São Paulo - USP, Doutorado em Filologia e Língua Portuguesa - Letras, Universidade de São Paulo – USP, São Paulo, SP, Brasil; msantiago@usp.br

the country.

Keywords: Clavicle. Variant. Cantareira. Saboneteira.

1 INTRODUÇÃO

Este artigo exhibe um recorte do estudo de mestrado intitulado *Atlas Semântico-Lexical de Formoso do Araguaia - Tocantins: a Dialetologia Pluridimensional e Relacional na Amazônia Legal*. A pesquisa se baseia na abordagem da Dialetologia Pluridimensional e Relacional proposta por Thun (2010), bem como nos princípios labovianos da variação linguística (Labov, 2008). A análise engloba várias dimensões linguísticas, incluindo as diatópicas (geográficas), diastráticas (sociais), diagenéricas (relacionadas ao gênero), diageracionais (geracionais), diavarietais (variações regionais) e diarreferenciais (referentes a situações comunicativas). O principal objetivo deste estudo é contribuir para a pesquisa sobre as variedades da língua portuguesa brasileira, destacando a riqueza cultural local e investigando as variações linguísticas presentes no município de Formoso do Araguaia.

Nesse contexto, o foco central deste artigo é a variante *clavícula* e suas diferentes denominações documentadas nesse município, tendo como base a pergunta 106 do Questionário Semântico-Lexical do Atlas Linguístico do Brasil (QSL/ALiB): *Qual é o nome do osso que vai do pescoço até o ombro?*, pertencente à área semântica “corpo humano”. O município pesquisado é conhecido como a “Capital da Irrigação” e está localizado próximo à “Ilha do Bananal” e às comunidades indígenas do tronco Macro-jê. Além disso, é cortado por dois rios importantes: o rio Formoso, que desempenha um papel crucial no projeto de irrigação local e nas plantações de arroz e melancia, e o rio Javaés, que se forma a partir de um afluente do rio Araguaia e, é conhecido por sua riqueza em peixes, como surubim, pirarucu, piaçu, boto, entre outros. A região também abriga as comunidades indígenas nas proximidades do município, como: Canuanã, São João e Porto Piauí, representadas pelos povos indígenas Javaés, Avá-canoeiros e Krahô-canela, respectivamente.

Outro ponto de destaque é a comunidade de Canuanã, localizada na zona rural do município e nomeada em homenagem à comunidade indígena vizinha. Nesta região, destaca-se uma escola de educação básica da Fundação Bradesco, localizada na Fazenda Canuanã, desempenhando um papel fundamental na formação dos alunos da região, fica próxima à Ilha do Bananal e abriga aproximadamente dois mil moradores.

Dada a diversidade cultural presente, torna-se essencial examinar os aspectos semântico-lexicais dessa região. Portanto, é crucial investigar se as dimensões mencionadas anteriormente (diatópicas, diastráticas, diagenéricas, diageracionais, diavarietais e diarreferenciais, de acordo com Thun, (2000, 2005) exercem influência na linguagem cotidiana dos habitantes do município. Com o propósito de avaliar o impacto do processo migratório na fala local de Formoso do Araguaia, Tocantins, devido à diversidade linguística resultante da presença de migrantes maranhenses, gaúchos, caipiras, ribeirinhos e dos primeiros habitantes da região, este estudo concentrou-se em confirmar a hipótese de que o processo migratório influenciou e continua influenciando o português local.

Além disso, o objetivo geral deste trabalho foi analisar como a variação e a mudança linguística contribuem para a formação do léxico dos habitantes do município, explorando representações Dialetológicas Pluridimensionais e Relacionais,

com foco no aspecto semântico-lexical resultante da interação entre os falantes migrantes, indígenas e nativos da comunidade. Para atingir esse objetivo, seguimos as dimensões estabelecidas por Thun (2000, 2005) na Dialetologia Pluridimensional e Relacional, que incluem as dimensões supracitadas. Dessa forma, o estudo busca compreender como o processo migratório e as diversas dimensões da variação linguística moldaram o léxico e a fala na comunidade de Formoso do Araguaia, Tocantins.

2 A DIALETOLOGIA PLURIDIMENSIONAL E RELACIONAL

A Dialetologia Pluridimensional e Relacional de Thun (1998) emerge da fusão da “Dialetologia Areal e da Sociolinguística”, que antes eram tratadas de forma independente. Essa abordagem se integra a um novo conceito de Geolinguística com o aparecimento da Geolinguística Moderna, conhecida como: “Dialetologia Pluridimensional”. Nesse contexto, ela é reconhecida como uma parte essencial da ciência da variação linguística, considerando as variantes e variáveis identificadas pelos falantes.

Desse modo, a Dialetologia Pluridimensional e Relacional amplia o campo de estudo Geolinguístico para uma perspectiva tridimensional. No entanto, isso não implica na rejeição da Dialetologia monodimensional/tradicional, da variação diatópica ou da Sociolinguística, pois cada uma dessas dimensões continua sendo relevante. Thun (1998, 2009) destaca que, ao abordá-las de forma separada, essas disciplinas podem apresentar limitações. Porém, ele também enfatiza que seu campo de estudo preferido reside na superfície geográfica e nas interdependências que dão origem aos fenômenos linguísticos. Nesse contexto, é essencial ressaltar algumas características fundamentais que permeiam a abordagem da Dialetologia Pluridimensional e Relacional de Thun (2010):

Primeiramente, observa-se uma ampliação considerável no escopo da Dialetologia monodimensional/tradicional e da análise da variação linguística (dimensão diatópica). Concomitantemente, a Geolinguística transcende a consideração exclusiva dos aspectos monodimensionais/tradicionais e passa a contemplar a importância de uma Geolinguística Pluridimensional e Relacional. Isso implica que a pesquisa se dedique à análise das manifestações linguísticas dos falantes, suas interações sociais e as variáveis extralinguísticas, tais como o estilo e a dimensão diafásica.

Uma característica fundamental dessa abordagem é a inclusão de múltiplos informantes, o que aumenta a representatividade do estudo, tanto em aspectos quantitativos quanto qualitativos. A presença de diversos informantes fomenta um diálogo produtivo entre o pesquisador e os informantes.

É importante destacar, em segundo lugar que os comentários metalinguísticos dos informantes não se limitam a enunciados relacionados ao mundo extralinguístico; eles também englobam considerações sobre as formas linguísticas encontradas, criando assim o que Thun (2010) denominou de “dimensão diarreferencial”. Em terceiro lugar, a pesquisa não se circunscreve mais ao estudo do dialeto, *standart* permitindo, em vez disso, a investigação das variedades *substandart* e das variações linguísticas em situações de contato, por meio da dimensão dialinguística.

Para desenvolver a metodologia da Dialetoologia Pluridimensional e Relacional, o pesquisador pode oferecer sugestões ou recomendações aos informantes durante as entrevistas. Após o registro das respostas espontâneas, o pesquisador apresenta ao entrevistado sugestões de expressões linguísticas. Além disso, incorporam-se perspectivas como a nanocronologia, a microcronologia, a mesocronologia e a macrocronologia, que estão relacionadas à análise dos dados e ao estudo da cronologia (tempo) nos atlas pluridimensionais. Também são adotadas mudanças no processo de mapeamento e tratamento dos dados, utilizando um procedimento que pode ser comparado à técnica cinematográfica de corte, chamada cronofotografia.

Para Thun (2017), a Dialetoologia Pluridimensional e Relacional emerge da consolidação de conceitos de cunho humanista da linguística, aliada à preservação de conceitos de menor expressão na Dialetoologia, conforme destacado por Abbé Rousselot (1888) e Tomás Navarro Tomás (1966).

A sistematização da Dialetoologia Pluridimensional e Relacional ganha destaque por meio do *Atlas Lingüístico Diatópico y Diastrático del Uruguay* (ADDU, 1989). Thun (2000) esclarece que as pesquisas que se baseiam na pluridimensionalidade têm a capacidade de responder a duas questões que, até então, permaneciam sem solução. A primeira delas se refere à extensão de um fenômeno linguístico observado entre os falantes de uma determinada região. Esses falantes compartilham o mesmo nível social, grupo geracional, estilo de expressão, entre outros? Enquanto a segunda busca compreender o grau de disseminação de um fenômeno identificado em uma ou várias camadas, investigando se ele pode coexistir em uma mesma região, envolvendo diversas gerações, estilos linguísticos, e outros fatores.

Assim, para a elaboração desta pesquisa, adota-se os parâmetros estabelecidos por Thun (2000, p. 189), levando em consideração as diversas dimensões da pluridimensionalidade linguística:

Quadro 1 - Parâmetros/dimensões da Dialetoologia Pluridimensional e Relacional.

Dimensões	Parâmetros
Diatópica	Rede de pontos: zona urbana e zona rural
Diageracional	(GI) Geração de Jovens - 18 a 40 anos
	(GII) Geração de Velhos - acima de 50 anos
Diassexual	Homem – masculino
	Mulher - feminino
Diastrática	Estratos sociais/escolaridade
	Classe baixa (Cb) - até ensino médio completo
Diarreferencial	Fala objetiva
	Fala metalinguística
Diafásica	Ferramenta dos três passos: perguntar, insistir e sugerir/sugêrência.

Fonte: elaborado pelos autores a partir de Altenhofen (2016, p. 375).

De acordo com Radkte e Thun (1999), a dimensão diatópica é caracterizada por ser “estrática”, pois concentra seus estudos em um espaço ou região específicos. No entanto, também é considerada “monoestrática”. Os autores argumentam que os fenômenos linguísticos têm sua origem em um espaço delimitado, mas os usos linguísticos estão intrinsecamente ligados às comunidades de falantes que estão diretamente associadas a esse espaço ou região. Portanto, não é possível dissociá-los.

A dimensão diastrática contempla diversos elementos que influenciam os comportamentos linguísticos dos falantes em uma comunidade. Para Radtke e Thun (1999, p. 36) a ampliação mais frequente ocorre com a combinação dos parâmetros diatópico, diastrático e diageracional. Portanto, esta pesquisa, adota as dimensões labovianas: diatópico, diastrático, diageracional, diassexual, diafásico e diarreferencial, embasadas nos princípios da Sociolinguística.

Somado a isso, a dimensão diageracional enfoca o perfil dos informantes em relação à idade. Desse modo, Marques (2018) conduziu um estudo comparativo entre os usos linguísticos dos falantes da geração mais velha (acima de 50 anos) e os da geração mais jovem (18 a 40 anos) no município de Cólíder localizado no estado de Mato Grosso em sua pesquisa de mestrado intitulada: *Atlas Semântico-Lexical de Cólíder – Mato Grosso*. Cardoso (2010) ressalta que a especificação da idade dos informantes já era considerada importante pelos pesquisadores desde o século XIX. Já a dimensão diassexual refere-se à diferenciação dos usos linguísticos entre os falantes do sexo masculino e feminino. Cardoso (2010) destaca que o perfil dos informantes, incluindo idade e gênero, passou a ser relevante para os estudos dialetais, visando ao registro documental.

Ademais, a dimensão diafásica aborda aspectos estilísticos, como leitura, respostas as perguntas do questionário, conversas livres e semidirigidas. Neste estudo, a dimensão diafásica é analisada durante as aplicações do Questionário Semântico-Lexical (QSL/ALiB) nas entrevistas gravadas e na aplicação dos três passos. Por fim, a dimensão diarreferencial foca nos comentários metalinguísticos dos informantes e em suas atitudes linguísticas durante as entrevistas. O inquiridor procura estimular as formulações, opiniões e percepções sociais dos falantes sobre a variedade linguística durante o inquérito. Borella (2014) destaca a importância de adotar a técnica dos três tempos, que envolve a pluralidade de informantes e leva os entrevistados a tecerem comentários metalinguísticos que podem revelar outras variantes presentes no repertório passivo ou ativo da comunidade pesquisada.

Nesse contexto, as dimensões labovianas contribuem para a evolução da Dialetoлогия, uma vez que voltada para a integração com a Sociolinguística, tem o intuito de equiparar-se à mesma em termos de documentação da variação linguística. Assim, a Sociolinguística se destaca por seus métodos modernos, que conseguem relacionar as dimensões sociais com as mudanças na língua.

Dessa forma, os dialetólogos buscam uma abordagem que permita a documentação adequada do fenômeno linguístico. O professor Harald Thun introduz o princípio do sistema em cruz, empregando a ferramenta dos três passos. Isso agrega várias dimensões da Sociolinguística, incluindo estratos sociais como níveis de escolaridade, faixas etárias, gênero e outros, ampliando as perspectivas da pesquisa em relação à Geolinguística que, historicamente, estava mais focada em informantes masculinos.

Essa abordagem metodológica resulta em uma produção significativamente maior de informações, como mencionado por Krug e Horst (2022). Thun destaca que, em cada grupo, os informantes podem fornecer dados diferentes, resultando em múltiplas variações do mesmo fenômeno. Isso torna a cartografia dos dados em um momento desafiador, já que até oito resultados diferentes podem ser registrados em um mesmo ponto geográfico. Para simplificar a cartografia, Thun propõe a criação de “mapas fenotípicos”, que representam apenas uma primeira aproximação do fenômeno, reduzindo a variação entre quatro a oito pessoas a um único símbolo (Krug; Horst, 2022).

É importante notar que a Geolinguística não se limita apenas à documentação da língua falada, mas também busca registrar os estágios iniciais dos processos de inovação linguística que começam na oralidade. No entanto, nem todas as inovações apresentadas pela Sociolinguística são realmente novas, uma vez que alguns aspectos já haviam sido pesquisados, testados ou questionados pela Geolinguística.

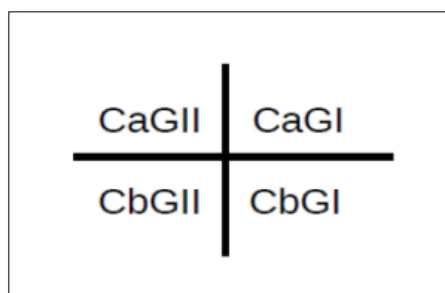
Tomás Navarro Tomás, por exemplo, questionava a questão da classe social antes mesmo de William Labov. De acordo com Thun (2000) as lacunas na pesquisa geolinguística levaram a um período de estagnação e esquecimento nas décadas de 70 a 90, com muitos projetos de atlas geolinguísticos interrompidos, sendo esse período referido como “cemitérios geolinguísticos” por Thun (1998).

Durante esse período, a Sociolinguística ganhou destaque com a teoria laboviana, contribuindo para a diminuição do interesse pela Geolinguística. Harald Thun não foi o primeiro a propor a pluridimensionalidade na Geolinguística. Além disso, a combinação da Dialetoлогия tradicional com a teoria laboviana permite a representação de um aspecto importante da metodologia: o contraste entre diferentes gerações. Isso é essencial para tornar visível a dimensão temporal durante a pesquisa, envolvendo entrevistas com pelo menos duas gerações.

A cartografia linguística utiliza o modelo em cruz para distinguir os quatro grupos de maior representatividade, chamados de grupos *standard*, que variam socioculturalmente de acordo com o grau de escolaridade. A linha horizontal divide a cruz e diferencia dois grupos, enquanto a linha vertical separa as gerações, com a geração mais velha à esquerda e os jovens à direita (Krug; Horst, 2022).

Assim, a Dialetoлогия Pluridimensional e Relacional condensa quatro mapas em uma única representação, sintetizando assim quatro grupos em um só mapa. É relevante destacar que esta abordagem metodológica é capaz de espelhar a situação linguística de uma comunidade. Portanto, se em um ponto específico, todos os grupos apresentam o mesmo resultado em uma cruz, sendo representados pelo mesmo símbolo, isso denota unidade. Essa unidade linguística é característica da comunidade em questão. Além disso, os dados podem convergir para o mesmo desfecho “[...] tanto nas gerações velhas, quanto nas gerações novas, mas diferentes da Classe alta da geração mais velha (CaGII), isso nos indica o rumo da mudança linguística no lugar” (Krug; Horst, 2022, p. 12).

FLP 25(2)



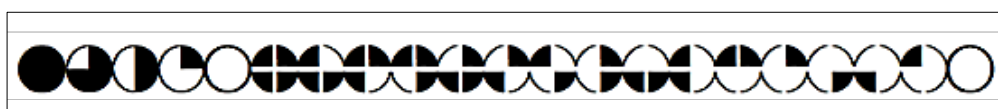
Fonte: Thun (2010, p. 509).

Figura 1 – Sistema em cruz.

O sistema em cruz representa as divisões dos informantes por nível de escolaridade: a classe alta/maior escolaridade (Ca) e a classe baixa/menor escolaridade (Cb) pela linha horizontal. Enquanto, a linha vertical, representa as divisões geracionais: (GI) mais novos e (GII) mais velhos. No entanto, este método impõe a restrição de utilizar apenas um símbolo em cada divisão da cruz, não permitindo a presença de dois ou mais símbolos. Isso se deve ao objetivo de simplificar a interpretação das informações cartografadas, considerando que nossa percepção visual é tridimensional, e qualquer informação adicional não é registrada pelo nosso cérebro. Portanto, é crucial evitar a inclusão de mais de um símbolo em cada célula do sistema da cruz.

Historicamente, a cartografia linguística representa um campo relativamente recente, uma vez que todos os componentes que a constituem foram desenvolvidos no passado e estão distantes da disciplina linguística tradicional. Além disso, é comum encontrar modelos de mapas mais antigos e anteriores à cartografia linguística, nos quais todo o aparato utilizado é notavelmente diferente. Por exemplo, em mapas que datam dos períodos de descobrimento da América, é comum encontrar símbolos que representam lugares com igrejas, que podem ter sido fundadas tanto por portugueses quanto por espanhóis. Nesses mapas, as igrejas são frequentemente representadas como pequenas estruturas simplificadas, muitas vezes reduzidas a uma forma abstrata com uma cruz. Essa uniformização na simbolização era o modelo predominante para esse tipo de mapa.

No campo da linguística, a abordagem adotada é a do “ponto-símbolo”, seguindo os parâmetros estabelecidos por Thun (2010). Essa abordagem pode ser comparada a um dicionário, pois é capaz de retratar uma quantidade significativa de informações em uma única página. O princípio subjacente a essa técnica parte do ponto central e se estende até a zona circundante, sendo representado por um círculo que vai do ponto cheio ao ponto vazio, conforme demonstrado a seguir:



Fonte: Krug; Horst (2022, p. 12).

Figura 2 - Símbolos *Kiel*.

A perspectiva apresentada por Thun (2010) comunica a ideia da existência de zonas linguísticas, como exemplificado com a palavra *caçula* na região Norte do Uruguai, onde os símbolos são representados em preto, criando assim uma zona distinta. Por outro lado, na região Sul dessa fronteira uruguaia, os símbolos são representados em branco, criando, por conseguinte, outra zona (Krug; Horst, 2022).

No entanto, a abundância de dados gerados pela abordagem Pluridimensional e Relacional também apresenta restrições no que diz respeito à cartografia, uma vez que ela só é capaz de registrar até quatro formas predominantes no mapa polifônico. Somente no início do século XIX, esse desafio de representar tudo em um único mapa se consolidou com o formato cartográfico em vigor. Mesmo assim, os cartógrafos da época reconheceram que, para que essa metodologia produzisse resultados satisfatórios, o mapa precisava ser legível. A Dialectologia Pluridimensional e Relacional então, sugere que “[...] a solução foi separar as informações em mapas sucessivos e unidos tematicamente” (Krug; Horst, 2022, p. 12).

A abordagem Pluridimensional e Relacional permitiu o desenvolvimento de mapas simples e claros. No entanto, é importante destacar que esses resultados foram fruto de intensas discussões entre colegas que também contribuíram para a elaboração de atlas linguísticos. Além disso, esses avanços realizaram-se através de pesquisas que tinham como objetivo promover o crescimento da cartografia embasados em experiências concretas e experimentos (Krug; Horst, 2022).

Thun (2022) explicou também a relação entre a metáfora dos cavalos e a cartografia linguística pluridimensional. Ele ressaltou que essa metáfora foi fundamentada na cronofotografia para justificar o desenvolvimento da cartografia linguística na época. Quando o francês Jules Gilliéron e seus colegas publicaram o *Atlas Linguistique de France* (ALF, 1902-1910), eles utilizaram a terminologia “cronofotografia” relacionada à fotografia daquela época. No entanto, os elementos essenciais para a fotografia moderna só foram criados nos Estados Unidos por volta dos anos 80.

Metaforicamente, o francês explicou que um atlas linguístico poderia ser comparado a uma “foto instantânea de um momento da história e da língua, sem pose” (Krug; Horst, 2022, p. 14), sem retoques. Uma vez que, com os avanços tecnológicos e os investimentos da Kodak na criação de fotos instantâneas, muitos pesquisadores introduziram a fotografia para desenvolver e controlar experimentos científicos, tornando a cronofotografia um grande avanço nessa época.

Krug e Horst (2022, p. 14) acrescentam que “o inglês Muybridge, que trabalhou com cavalos, desejava estudar e representar as fases do movimento do cavalo nos diferentes momentos de seu passo, como do passo até o galope”. A técnica de combinar fotos sucessivas em uma única imagem permitiu a observação do movimento progressivo do cavalo em uma única foto, e essa abordagem de fotos sucessivas ficou conhecida como cronofotografia.

É por isso que o pesquisador Thun usa a metáfora do cavalo, representando várias fotos do fenômeno linguístico. Ele não busca apenas capturar uma única imagem, mas, ao contrário, deseja registrar várias fotografias desse fenômeno linguístico em constante mudança e em diferentes contextos, ou seja, em movimento. A imagem ideal para a Geolinguística tradicional é aquela que encontramos nos atlas mais antigos, como o *Atlas Linguistique de France* (ALF, 1902-1910), que representa todas as variações das formas linguísticas. Sabe-se que essas variações existem, mas

FLP 25(2)

não se conhece o status delas, como se são mais usadas por homens ou mulheres, pelas classes sociais mais altas ou mais baixas. Isso representa uma limitação dos dados da Geolinguística tradicional.

No entanto, para capturar várias imagens desse fenômeno linguístico, é necessário estabelecer previamente dimensões e parâmetros. Dessa forma, podem-se aplicar ferramentas que tornam as imagens mais nítidas, permitindo um zoom maior. As ferramentas que Thun utiliza incluem diferentes tipos de questionários, objetivos, conversa livre e de leitura. Além disso, ele desenvolveu a ferramenta dos três passos - *perguntar*, *insistir* e *sugerir* - que possibilita um zoom aprimorado no fenômeno linguístico e, conseqüentemente, o registro de várias imagens desse fenômeno.

3 METODOLOGIA

3.1 Tratamento dos dados

Este artigo descreve a metodologia utilizada em um estudo qualitativo baseando-se na obra *Pesquisa Qualitativa com texto, imagem e som* de Martin e Gaskell (2008). O objetivo desta pesquisa é analisar os aspectos semântico-lexicais das palavras em diferentes áreas semânticas, considerando sua ocorrência, frequência e variações. A pesquisa se concentra na comunidade linguística de Formoso do Araguaia, localizada em Tocantins, reconhecida por sua rica diversidade lexical.

A abordagem escolhida para este estudo é a Dialectologia Pluridimensional e Relacional, que visa documentar diferentes formas de expressão dentro da comunidade. Para alcançar esse objetivo, utiliza-se várias ferramentas, incluindo o sistema em cruz de Thun (2010), as dimensões da teoria Sociolinguística de Labov (2008) e o Questionário Semântico-Lexical (QSL) composto por 202 perguntas e divididas em 14 áreas semânticas distintas. No entanto, neste artigo, opta-se por focar na área semântica “corpo humano” e, mais especificamente, na pergunta 106: *Como se chama o osso que vai do pescoço até o ombro?*

Além disso, destaca-se a importância das características históricas, culturais e linguísticas de Formoso do Araguaia para este estudo. Dessa forma, a pesquisa recebeu aprovação do Comitê de Ética em Pesquisa da Universidade Estadual de Mato Grosso e seguiu as diretrizes estabelecidas pelas resoluções 466/2012 e 510/2016 da CONEP. Durante a coleta de dados, foram adotadas medidas de segurança devido à pandemia de Covid-19, como o uso de máscaras, álcool em gel e o cumprimento do distanciamento social. Portanto, este artigo, aborda detalhadamente a metodologia adotada na condução da pesquisa.

3.2 Rede de pontos

No contexto apresentado torna-se evidente a importância de desenvolver um Atlas Semântico-Lexical de Formoso do Araguaia (ASeLFO) com base no Atlas Linguístico do Brasil. Isso requer a definição de locais de pesquisa com base em critérios como relevância histórica, densidade populacional e distribuição geográfica da comunidade. O estudo abrange tanto a zona urbana quanto a rural do município, considerando o aumento da população urbana em detrimento da rural e as conseqüentes mudanças linguísticas.

A decisão de incluir a área rural, especificamente a Fazenda de Canuanã, destaca-se como fundamental para a análise das disparidades linguísticas entre a zona urbana e a rural. A pesquisa engloba diversos pontos de investigação em várias variantes do português e divide o município em setores socioculturais, examinando tanto bairros antigos como novos, além de considerar as diferenças educacionais.

Portanto, podemos observar que a delimitação da área de pesquisa dialetal é influenciada por múltiplos fatores, como localização geográfica, contexto histórico, economia e demografia da região. Nesse contexto, a inclusão da comunidade rural como área de pesquisa é de suma importância para a análise e comparação do léxico utilizado em Formoso do Araguaia.

Quadro 2 - Pontos de Inquérito – variedades e números de habitantes.

Área Urbana e Rural	Ponto	Grupos por Variedades	N.º de hab.
Zona Urbana	1	Português Maranhense	13.333
		Português Gaúcho	
		Português Caipira	
Área Rural	2	Fazenda de Canuanã/Fundação Bradesco	1.626

Fonte: Elaboração própria.

3.3 Perfil dos informantes

As concepções metodológicas desta pesquisa concentram-se na análise da variação linguística em Formoso do Araguaia, localizado no estado do Tocantins. Para isso, consideram-se duas gerações de informantes: a Geração Um (GI), constituída por indivíduos mais jovens, com idades entre 18 e 40 anos, e a Geração Dois (GII), composta por pessoas mais velhas, com 50 anos de idade ou mais, que residem em Formoso do Araguaia há pelo menos trinta anos.

Nesse contexto, as respostas dos informantes foram analisadas com base em quatorze áreas semânticas do Questionário Semântico Lexical (QSL), parte integrante do Atlas Linguístico do Brasil (ALiB). Para garantir a precisão dos resultados, seguiu-se rigorosamente os critérios estabelecidos pelo ALiB na seleção dos informantes, evitando a inclusão de membros da mesma família, embora tenhamos enfrentado desafios ao encontrar participantes que se encaixassem no perfil desejado.

As entrevistas foram conduzidas individualmente, com a aplicação do QSL. O grupo de informantes foi composto por um total de 32 pessoas, distribuídas igualmente entre homens e mulheres, abrangendo diferentes níveis de escolaridade. Além disso, foram considerados quatro grupos varietais, representando distintas variedades do português (maranhense, gaúcho, caipira e ribeirinho). Para assegurar a representatividade equitativa dos grupos varietais, entrevistamos mais 16 informantes adicionais durante a aplicação do terceiro passo da pesquisa. Ao final, o conjunto de informantes totalizou 48 indivíduos, com uma distribuição equitativa em termos de gênero, idade e escolaridade, com o objetivo principal de analisar a variação linguística em Formoso do Araguaia.

3.4 Processo de coleta dos dados

Os métodos empregados na coleta de dados para este estudo linguístico resultaram na formação do seu corpus através de uma combinação de encontros presenciais gravados, questionários impressos com anotações, transcrições e medidas rigorosas de biossegurança para mitigar qualquer risco de contaminação durante a pandemia de Covid-19. A fim de facilitar o contato com os informantes, intermediários foram utilizados como mediadores antes das visitas aos locais de pesquisa.

Todas as entrevistas foram registradas, tanto em encontros presenciais quanto por meio de plataformas digitais, e as respostas referentes a um total de 202 perguntas foram minuciosamente transcritas. Além disso, enfatiza-se a importância do pesquisador dominar as técnicas de pesquisa e conduzir testes preliminares para garantir a obtenção de dados confiáveis. O corpus desse estudo foi formado por variantes que representam as variações linguísticas formosenses.

4 RESULTADOS E DISCUSSÕES

4.1 Descrição e análise

Os resultados obtidos neste estudo teve como eixo norteador o Questionário Semântico-Lexical (QSL) do Atlas Linguístico do Brasil (ALiB), juntamente com a aplicação da ferramenta dos “três passos” desenvolvida por Thun (2010) durante a pesquisa de campo. A coleta de dados foi realizada através do QSL, combinada com entrevistas utilizando a técnica dos “três passos”. Estes três passos consistiram em questionar inicialmente os entrevistados e aguardar suas respostas espontâneas, em seguida, insistir para estimular a evocação de memórias linguísticas mais profundas. E, por fim, sugerir variantes linguísticas que talvez não tivessem sido mencionadas inicialmente.

A adoção dessa abordagem permitiu que os entrevistados recordassem outras variantes linguísticas conhecidas, sendo de grande relevância para a Geolinguística contemporânea. Durante a fase de insistência, a pesquisadora encorajou os entrevistados a refletirem sobre suas memórias linguísticas, incluindo influências de gerações passadas, como as de seus avós e bisavós. Isso proporcionou uma compreensão mais profunda da maneira como a língua se desenvolve ao longo das gerações.

Além disso, a pesquisa faz menção aos comentários metalinguísticos feitos pelos informantes durante as entrevistas, nos quais eles relacionavam certas variantes linguísticas a regiões geográficas ou a gerações anteriores, como quando afirmavam “isso é típico dos gaúchos” ou “meus avós costumavam falar assim”. Esses comentários revelaram-se de suma importância para a pesquisa, proporcionando informações valiosas sobre a evolução linguística.

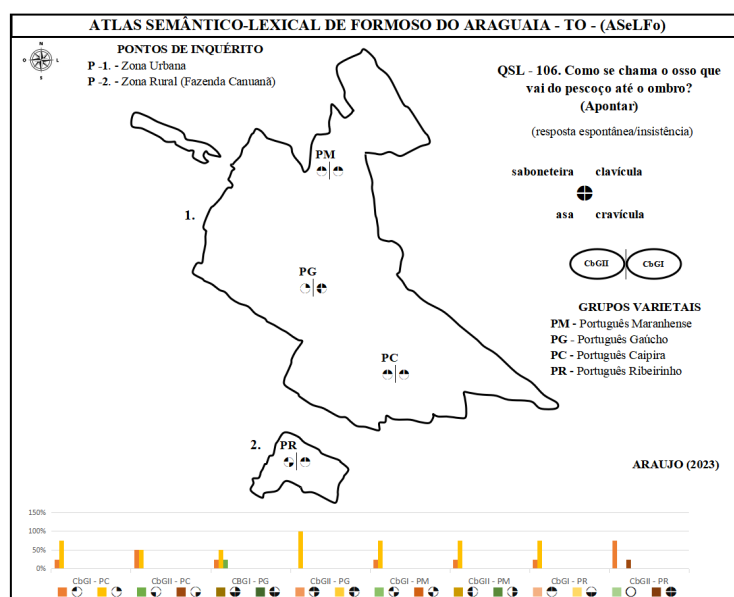
É crucial ressaltar a importância de levar em consideração a dimensão diarreferencial ao explorar a diversidade de informantes, uma vez que cada entrevistado pode representar uma rede complexa de influências linguísticas que inclui avós, pais, vizinhos e outros membros de sua comunidade.

Portanto, a aplicação da técnica dos “três passos” de Thun (2010) e a atenção dada aos comentários metalinguísticos emergentes são ferramentas indispensáveis na

pesquisa linguística que busca compreender a evolução da língua ao longo do tempo e em diferentes contextos sociais e geográficos. A seguir, serão apresentados os mapas resultantes da análise dos dados coletados da questão 106 (QSL/ALiB).

4.2 Mapas polifórmicos e de status das formas

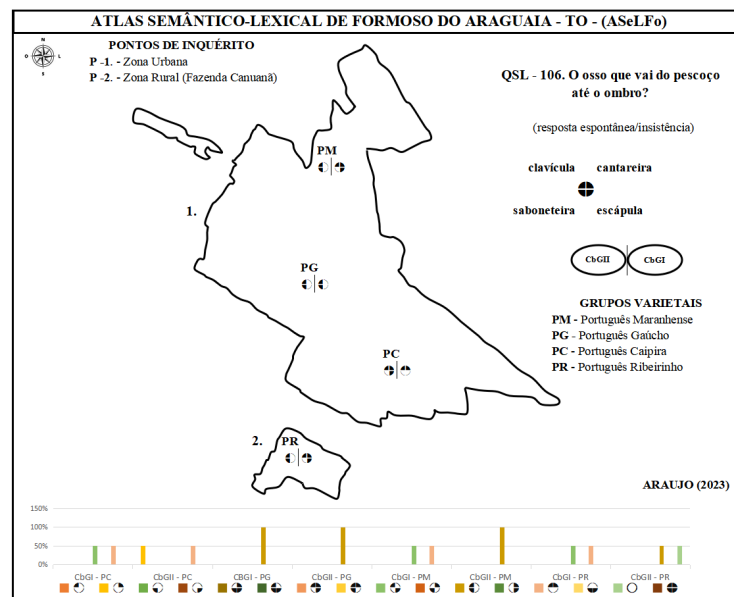
O Mapa polifórmico *clavícula* registra as formas das lexias mais comuns nas respostas espontâneas à pergunta 106: *Qual é o nome do osso que se estende do pescoço até o ombro?*



Fonte: Araujo (2023, p. 347).

Figura 3 - Mapa polifórmico - *Clavícula*.

Com base nas informações fornecidas pelos participantes dos quatro grupos de variedades investigados, registram-se as seguintes variações linguísticas: *saboneteira*, *clavícula*, *asa* e *cravícula*. De acordo com Machado (1977, p. 165), a variante *clavícula* é descrita da seguinte maneira: “Do latim *clāvīcula*, que significa *chavinha*, *rollha*, *batoque*, certa espécie de fortificação, *gavinha*”. Nesse contexto, os grupos de informantes das variedades do português (PM) e (PC), nas duas gerações de informantes (GII) e (GI), apresentam as formas *saboneteira* e *clavícula*, evidenciando um conhecimento espontâneo de 50% das variações mapeadas. Por outro lado, o grupo do português gaúcho (PG) registra, para a (CbGII), apenas a forma *clavícula*, correspondendo a um percentual de 25% das variações documentadas. Enquanto isso, a (CbGI) documenta de forma espontânea as formas *saboneteira*, *clavícula* e *asa*, demonstrando um conhecimento de 75% das variações mapeadas. Adicionalmente, o grupo (PR) apresenta, para a (CbGII), as formas *saboneteira* e *cravícula*, totalizando um conhecimento de 50% das variações documentadas. Por sua vez, a (CbGI) exibe as formas *saboneteira* e *clavícula*, somando um conhecimento de 50% das variações mapeadas.



Fonte: Araujo (2023, p. 348).

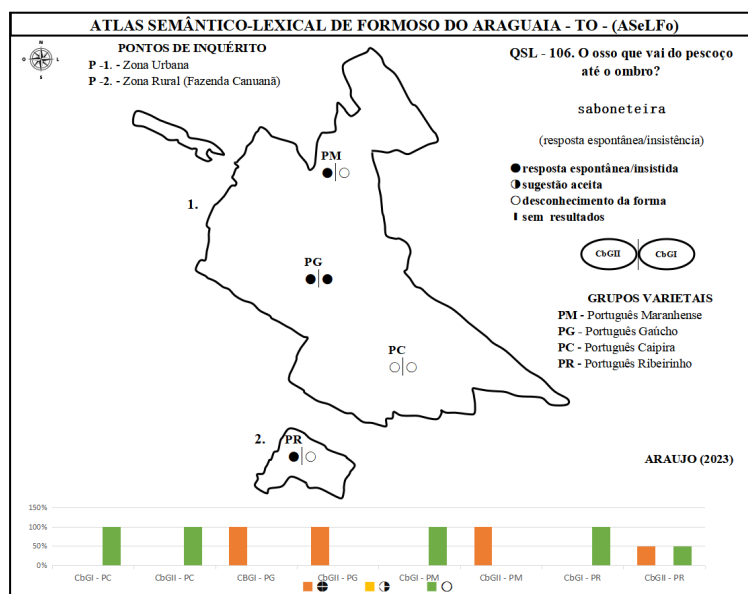
Figura 4 - Mapa polifórmico - *Clavícula* – 3.º passo.

A figura 4 apresenta o mapa polifórmico da variante *clavícula*, aplicando o terceiro passo da metodologia de Thun (2010), *sugerir/sugerência*. As variantes documentadas pelos participantes deste estudo englobam os termos *clavícula*, *cantareira*, *saboneteira* e *escápula*. É importante notar que essas variantes já foram identificadas em pesquisas anteriores que exploraram as formas selecionadas, a saber: o *Atlas linguístico-etnográfico da Região Sul do Brasil* (ALERS, 2011), o *Atlas Linguístico Topodinâmico e Topoestático do Tocantins* (ALITTETO, 2018) e o *Atlas semântico-lexical do estado de Goiás* (ASELGO, 2013). Dessa forma, as sugestões adotadas para compor este estudo são as seguintes: *clavícula* e *paleta* conforme ALERS, 2011, *saboneteira* e *cantareira* conforme ALITTETO, 2018, *escápula* e *pá*, conforme ASELGO, 2013.

Ferreira (2004, p. 481) define a forma *clavícula* da seguinte maneira: “do latim *clavicula*, pequena chave. Cada um de dois ossos, situados ântero-superiormente em relação ao tórax, que se articulam, por um lado, com a omoplata e, pelo outro lado, com o esterno”. Nesse contexto, Romaguera Côrrea et al. (1964), Ferreira (2004) e Ortêncio (1983) registram a variante *paleta* com as acepções “1. é a omoplata ou a espádua do animal. 2. Como expressão anatômica, *paleta* é o termo castelhano que significa *pá*, nome vulgar para *espádua* ou *omoplata*” (Romaguera Corrêa et al., 1964, p. 333).

Ao analisar os dados coletados pelos grupos (PM) e (PR) observa-se que, para a (CbGII), as variantes *clavícula* e *saboneteira*, totalizam um conhecimento de 50% das formas registradas. Por outro lado, a (CbGI) apresentam as formas *clavícula*, *cantareira*, e *escápula* de forma espontânea, o que equivale a 75% das variantes mapeadas. Assim, o grupo do português gaúcho (PG) demonstra, em ambas as gerações (CbGII) e (CbGI), um conhecimento de 50% das formas cartografadas, incluindo as variantes

clavícula e *saboneteira*.



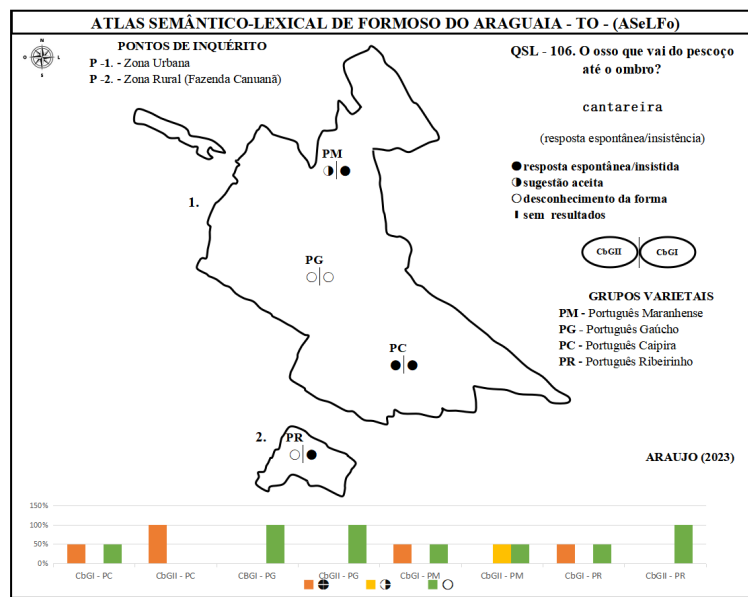
Fonte: Araújo (2023, p. 350).

Figura 5 - Mapa lexical - status da forma – *Saboneteira*.

Além disso, o grupo do português ribeirinho (PR) apresenta, para a geração (CbGII), as formas *clavícula* e *saboneteira*, representando um conhecimento de 50% das formas documentadas. Na geração (CbGI), as formas registradas são *clavícula*, *cantareira* e *escápula*, totalizando um conhecimento de 75% das variantes mapeadas. Diante do exposto, segue o status da forma *saboneteira*.

É importante destacar que o status da forma *saboneteira* reflete o conhecimento espontâneo da variante linguística por diferentes grupos de variedades do português. No grupo maranhense (PM), essa forma é registrada apenas para a (CbGII), enquanto os gaúchos (PG) a reconhecem em ambas as gerações (CbGII e CbGI), assim como os ribeirinhos (PR), que também aceitam essa variação para a (CbGII). Ferreira (2004, p. 1785) fornece as seguintes definições para *saboneteira*: “1. *Caixinha ou local específico para o sabonete; saboneira, saboeira, saboeiro, saboneiro*. 2. Cada uma das duas depressões que, em pessoas magras, destacam a saliência das *clavículas*”. Além disso, Ortêncio (1983, p. 395) documenta a forma *saboneteira* com a variante *sabonete de macaco*.

Em relação à (CbGII), os grupos varietais (PM) e (PR) não registraram conhecimento da forma *saboneteira* entre os informantes pesquisados. Da mesma forma, o grupo do português caipira (PC) demonstra desconhecimento da forma *saboneteira* em ambas as gerações, tanto para (CbGII) quanto para (CbGI). Dando continuidade a essa análise, é importante acrescentar o status da forma *cantareira*, conforme segue:



Fonte: Araújo (2023, p. 351).

Figura 6 - Mapa lexical - status da forma - *Cantareira*.

Ao analisar o status da variante denominada *cantareira*, é evidente que esta é considerada uma sugestão para a (CbGII) do grupo da variedade do português maranhense (PM) que foi confirmada em uso pelos seguintes grupos das variedades de português (GI/PM), (GI e GII/PC) e (GI/PR). É importante mencionar que os grupos (GII/PG) e a (GII/PR) demonstraram desconhecimento em relação à forma *cantareira* de acordo com os informantes. Uma vez que, Ferreira (2004, p. 390) também registrou a variante *cantareira* com a seguinte acepção: “[...] 2. Brasileirismo popular para designar a *clavícula*”.

5 CONSIDERAÇÕES FINAIS

Em síntese, este estudo metódico de cartografia linguística revelou-se um instrumento valioso para a compreensão das múltiplas variantes faladas no município de Formoso do Araguaia, Tocantins. Além disso, destacou a importância do processo migratório na influência das formas linguísticas locais. Utilizando técnicas metódicas, como o registro das lexias, a transcrição das gravações e a elaboração de mapas pluridimensionais, conseguimos mapear com precisão as nuances das variedades linguísticas presentes na região.

Ao focar especificamente na classe social baixa e na sua interação com o contexto migratório, observa-se uma interessante dinâmica linguística, representada pela cruz que divide as gerações. Esta pesquisa não apenas alcançou seus objetivos previamente estabelecidos, mas também ofereceu valiosas contribuições para a comunidade acadêmica, para a comunidade local e suas instituições de ensino. A análise dos resultados confirmou que os maranhenses desempenham um papel fundamental na formação da variedade dialetal local, exercendo influência sobre os informantes ribeirinhos e outros grupos linguísticos. Portanto, esta pesquisa não apenas enriqueceu o campo da dialetologia, mas também promoveu o reconhecimento e a valorização das realidades linguísticas e culturais da região, contribuindo para uma compreensão mais profunda da língua portuguesa local.

Dessa forma, este estudo representa uma significativa contribuição para os estudos dialetológicos em Formoso do Araguaia, no estado do Tocantins, e em todo o país. Ele ressalta a importância de continuar investigando e documentando as variações linguísticas em nosso país, a fim de promover a preservação e a compreensão de nossa rica diversidade linguística e cultural.

REFERÊNCIAS

- Altenhofen CV, Thun H. As migrações e os contatos linguísticos na geografia linguística do sul do Brasil - bacia do Prata. In: Aguilera VA, Romano VP, organizadores. *A geolinguística no Brasil: caminhos percorridos, horizontes alcançados*. Londrina: Edue; 2016.
- Araujo KJ. Atlas semântico-lexical de Formoso do Araguaia - Tocantins: a dialetologia pluridimensional e relacional na Amazônia legal [dissertação]. Universidade do Estado de Mato Grosso, Sinop, Mato Grosso; 2023.
- Augusto VLDS. Atlas semântico-lexical do estado de Goiás (ASELGO) [tese]. São Paulo: Universidade de São Paulo; 2013.
- Borella SG. Tu dampém fala assim?: macroanálises pluridimensionais da variação de sonorização e desonorização das oclusivas do português de falantes bilíngues hunsriqueano – português [tese]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2014. [citado 23 set. 2023]. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/108953/000948325.pdf?sequence=1&isAllowed=y>.
- Cardoso SAM. Geolinguística: tradição e modernidade. São Paulo: Parábola Editorial; 2010.
- Ferreira ABH. Novo Aurélio século XXI: o dicionário da língua portuguesa. Curitiba: Positivo; 2004.
- Koch W, Altenhofen CV, Klassmann M, organizadores. Atlas Linguístico-Etnográfico da Região Sul do Brasil (ALERS): introdução, cartas fonéticas e morfosintáticas. 2.^a ed. Porto Alegre: Editora da UFRGS; Florianópolis: Editora da UFSC; 2011. [citado 23 set. 2023]. Disponível em: <https://lume.ufrgs.br/handle/10183/232185>.
- Krug MJ, Horst C. Dialetologia pluridimensional e relacional: entrevista com o professor Dr. Harald Thun. *Working Papers em Linguística*. 2022;23(1):08-16. [citado 23 set. 2023]. Disponível em: <https://doi.org/10.5007/1984-8420.2022.e78597>.
- Labov W. Padrões sociolinguísticos. São Paulo: Parábola Editorial; 2008.
- Machado JP. Dicionário etimológico da língua portuguesa. 3.^a ed. Lisboa: Livros Horizonte; 1977.
- Marques MJB. Atlas semântico - lexical de Colíder – MT 2018 [dissertação]. Sinop: Universidade do Estado do Mato Grosso; 2018.
- Martin WB, Gaskell G. Pesquisa qualitativa com texto: imagem e som: um manual prático. 7.^a ed. Petrópolis: Vozes; 2008.
- Ortêncio WB. Dicionário do Brasil central. São Paulo: Ática; 1983.
- Radtke E, Thun H. Novos caminhos da geolinguística românica: um balanço. In: *Cadernos de Tradução do Instituto de Letras*. 1999;5:31-51.
- Romaguera Córrea JC, et al. Vocabulário sul-rio-grandense. Rio de Janeiro, Porto Alegre, São Paulo: Globo; 1964.
- Silva GA. Atlas linguístico e topoestático do estado do Tocantins (ALITTETO) [tese]. Londrina: Universidade Estadual de Londrina; 2018.
- Thun H. La geolinguística como linguística variacional general (con ejemplos del Atlas linguístico Diatópico y Diastrático do Uruguay). In: *International Congress of Romance Linguistics and Philology. Atti del XXI Congresso Internazionale di Linguistica e Filologia romanza*, 1995. V. 5. Palermo: Giovanni Ruffino, Tübingen: Niemeyer; 1998. p. 701-729, 787-789.
- Thun H. O português americano fora do Brasil. In: Gärtner E, Hundt C, Schonberger A. *Estudos de geolinguística do português americano*. Frankfurt am Main: TFM; 2000. p. 185-227.

FLP 25(2)

Thun H. A dialetologia pluridimensional no rio da Prata. In: Zilles AMS, organizador. Estudos de variação linguística no Brasil e no Cone Sul. Porto Alegre: Ed. da UFRGS; 2005. p. 63-92.

Thun H. A geolinguística pluridimensional, a história social e a história das línguas. In: Aguilera VA, organizadora. Para a história do português brasileiro. Vol. VII: vozes, veredas, voragens. Londrina: EDUEL; 2009. p. 531- 558.

Thun H. Pluridimensional cartography. In: Lameli A, et al. Language and space: language mapping: an international handbook of linguistic variation. Berlin, New York: Walter Gruyter GmbH; 2010. p. 506-523.

Thun H. O velho e o novo na geolinguística. In: Altenhofen CV, Pavan CF, organizadores. Cadernos de Tradução: percursos teóricos e metodológicos da dialetologia. 2017;40:59-81. [citado 23 set. 2023]. Disponível em: <https://seer.ufrgs.br/index.php/cadernosdetraducao/issue/view/3444>.