

Estudo dialetométrico do Atlas Semântico-lexical da Região do Grande ABC (ASL_ ABC)

Adriana Cristina **CRISTIANINI***

*Doutora em Linguística pela Universidade de São Paulo (2007). Professora Doutora Associada III_UFU.
adriana.cristianini@gmail.com.

Resumo:

Por meio de estudos dialetológicos, geolinguísticos, sociolinguísticos, entre outros, pesquisadores direcionam seus olhares para diferentes aspectos da língua e tipos de variação. Esses estudos nos oportunizam conhecer com maior exatidão, descrever e mensurar a atuação dos fenômenos linguísticos relacionados a fatores sociais, culturais, históricos, políticos, regionais e outros. Entretanto, o volume de dados tratados em tais estudos é gigantesco e tem-se tornado cada vez maior. Muitas vezes, somente com o auxílio de recursos tecnológicos podemos desenvolver pesquisas que, num passado não muito distante, seriam impossíveis. Dada essa breve reflexão, partimos dos registros do Atlas Semântico-lexical da Região do Grande ABC e, seguindo os princípios teórico-metodológicos da Escola Dialetométrica de Salzburgo, objetivamos apresentar neste artigo um estudo dialetométrico, com auxílio do software DiaTech. Essa ferramenta permite abordar um grande número de dados de um atlas linguístico, para relacionar dois ou mais pontos da pesquisa, auxiliando na constituição de uma visão geral e de padrões de variação diatópica. Segundo Saramago (2020), a Dialetometria permite um maior aproveitamento do material existente nos atlas linguísticos e possibilita a descoberta de relações que ficariam ocultas com a aplicação do método qualitativo ou tradicional na análise de dados.

Palavras-chave:

Dialetologia brasileira. Dialetometria. Região do Grande ABC.

Signum: Estudos da Linguagem, Londrina, v.26, n.3, p.78-101, dezembro. 2023

Recebido em: 27/09/23

Aceito em: 14/02/24

Estudo dialetométrico do Atlas Semântico-lexical da Região do Grande ABC (ASL_ABC)

Adriana Cristina Cristianini

INTRODUÇÃO

A constituição social e histórica do Brasil é repleta de pluralidade. A Língua Portuguesa, falada nesse País e que é o seu principal elemento cultural, apresenta, portanto, muita diversidade e variação que vêm sendo estudadas por vários caminhos.

Por meio de estudos da Dialetologia, da Geolinguística, da Sociolinguística, da Sociogeolinguística e da Dialetometria, entre outros, torna-se possível conhecer com maior exatidão, descrever e mensurar a atuação dos fenômenos linguísticos relacionados a fatores sociais, culturais, históricos, políticos, regionais e outros. Os pesquisadores, assim, direcionam seus olhares para diferentes aspectos da língua e tipos de variação.

No que se refere à variação diatópica, objeto deste estudo, partimos de um trabalho que seguiu os preceitos da Geolinguística, considerada um recorte da Dialetologia, que registra cartograficamente as respostas obtidas por meio de entrevistas aplicadas a certo grupo de sujeitos de uma dada localização, buscando precisar a repetição topográfica dos fenômenos observados.

Neste estudo, pretendemos analisar dados da variação diatópica retratada no *Atlas semântico-lexical da região do Grande ABC – ASL_ABC* (Cristianini, 2007), por meio da Dialetometria. Buscaremos direcionar nosso olhar principalmente a dois pontos da rede constituída para a elaboração desse atlas, mais especificamente aos pontos que correspondem a São Caetano do Sul e ao distrito de Santo André, Paranapiacaba.

Para tanto, além desta introdução e da conclusão, dividimos este trabalho em três seções: Fundamentação; Metodologia; e Resultados.

FUNDAMENTAÇÃO

Segundo García Mouton (1999), apesar dos avanços marcantes, ao longo do século passado, no desenvolvimento de teorias linguísticas e nos estudos da variação, sabemos que os atlas, com sua enorme riqueza, geralmente mantêm-se nas mãos quase que exclusivamente dos dialetólogos e dos geolinguistas, que, com grande frequência, são os mesmos que coletaram e cartografaram os dados.

A Geolinguística desenvolve estudos de variação linguística no espaço e registra em cartogramas¹ os dados coletados por meio de entrevistas, em uma rede de pontos de uma determinada região. A compilação desses cartogramas resulta nos atlas linguísticos. Segundo Álvarez Blanco, Dubert García e Sousa Fernández (2006, p. 461), a Geolinguística

[...] dedica-se ao estudo da variação linguística no espaço e baseia-se principalmente nos dados contidos nos atlas linguísticos, que são uma fonte muito rica de informações para o conhecimento profundo de traços fonéticos, morfológicos, sintáticos e lexicais de um conjunto de lugares em um território. Além dos dados específicos de cada um dos pontos

¹ Apesar de, tradicionalmente, encontrarmos na terminologia relacionada à Geolinguística e à Dialetometria os usos dos termos mapa ou carta linguística, desde 2007, temos utilizado o termo cartograma, considerando que este é um tipo de representação que se preocupa mais com informações que serão objetos da distribuição espacial no interior do mapa, que com os limites exatos e precisos ou com coordenadas geográficas (Sanchez, 1973 *apud* Cristianini, 2007).

pesquisados, os atlas nos permitem analisar a regularidade com que as diferentes unidades linguísticas são distribuídas no espaço, informação esta que os dialetólogos consideram crucial, uma vez que a descoberta de regularidades permite reconhecer as áreas do território que partilham traços linguísticos (tradução nossa)

Para identificar essas localidades que compartilham os mesmos traços linguísticos, os geolinguistas delimitam isoglossas nos cartogramas, que normalmente são em um número elevado, e as sobrepõem. Ocorre que, ao analisar essa sobreposição, percebemos que raramente é possível determinar com precisão áreas que apresentam características diferentes linguisticamente, na medida em que as isoglossas tendem a não coincidir de forma próxima umas com as outras, sobretudo se se utiliza um número substancial de fenômenos linguísticos na representação cartográfica. Diante disso, tornou-se importante a busca por uma maneira de sintetizar os dados dos atlas de modo que seja possível identificar os pontos em que ocorrem maiores coincidências e maiores diferenças em relação a outros pontos linguísticos.

A Dialetoimetria surge, então, da necessidade de, a partir do volume bastante grande de dados coletados por meio de pesquisas geolinguísticas e cartografados para apresentar os fenômenos abordados, sintetizar em um cartograma único os resultados de múltiplos cartogramas que compõem um atlas linguístico. Logo, surge da necessidade de registrar e analisar a totalidade dos registros contidos em um atlas, permitindo uma visão geral da variação contida em determinada localidade, região ou, em sentido lato, sistema dialetal ou linguístico.

Pioneiro nos estudos dialetométricos, Jean Séguy foi quem primeiro cunhou o nome Dialetoimetria e quem primeiro deu à disciplina um corpo teórico robusto, fundamentado em resultados práticos significativos, com a dialetometrização do *Atlas Linguistique de la Gascogne – ALG*, em 1973. A partir de Séguy, Guiter e os integrantes da escola de Toulouse – Philips, Fossat – desenvolveram o método e o aplicaram no âmbito românico (García Mouton, 1999). Segundo Saramago (2020), inicialmente, esse método “[...] consistia no traçado de uma rede triangular, resultado da ligação entre os pontos geograficamente mais próximos. Nessas linhas eram colocados os valores encontrados na totalidade do atlas para o fenômeno estudado”.

No final da década de 1970, Hans Goebel desenvolve um estudo dialetométrico com os dados contidos no *Atlas Linguístico-Etnográfico da Itália e da Suíça Meridional – AIS* e no *Atlas Linguistique de la France – ALF*. Podemos afirmar que, a partir dos estudos de Goebel, a Dialetoimetria se consolida e passa a ser considerada um importante estudo quantitativo, partindo do registro de dados dialetais fornecidos pela Geolinguística. É tendo em conta esse percurso já de cerca de meio século que Brissos e Saramago (2019, p. 353) definem a Dialetoimetria como

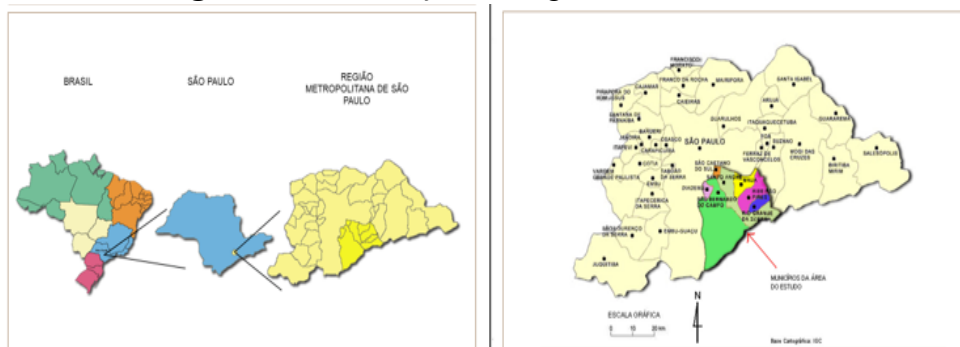
[...] uma abordagem quantitativa ao estudo dos dialetos com um enfoque na métrica, i.e., na mensuração dos fenômenos de variação dialetal por meio de procedimentos exatos e totalmente comparáveis, os quais importa da classificação numérica ou taxonômica. Aplica cálculos matemático-estatísticos elaborados à matriz de dados obtida a partir dos procedimentos referidos e representa cartograficamente (*espacializa*) os resultados desses cálculos, cabendo ao linguista, com a liberdade que a estatística confere, a tarefa final de interpretação do quadro geolinguístico que tem à frente.

A Dialetoimetria, portanto, permite abordar um grande número de dados, aproveitando os registros de um atlas linguístico, para relacionar dois ou mais pontos da pesquisa, auxiliando na constituição de uma visão geral e de padrões de variação diatópica. Segundo Saramago (2020), a Dialetoimetria permite um maior aproveitamento do material existente nos atlas linguísticos e possibilita a descoberta de relações que ficariam ocultas com a aplicação do método qualitativo ou tradicional na análise de dados.

A região do Grande ABC paulista é formada por sete municípios – Santo André, São Bernardo do Campo, São Caetano do Sul, Diadema, Mauá, Ribeirão Pires e Rio Grande da Serra – e tem uma importância considerável no contexto sócio-histórico-político-econômico-cultural do Brasil.

A Figura 1, a seguir, constituem-se em cartogramas de localização da região.

Figura 1 – Localização da Região do Grande ABC



Fonte: Elaborada pela autora.

Já se percebia, desde antes da elaboração do atlas, a urgência de se resgatar a variação semântico-lexical da língua falada no Grande ABC, como forma de se registrar a memória linguística da comunidade dessa localidade, pois, entre outras características dessa região, há uma grande mobilidade da população.

A história da fundação e do desenvolvimento dos municípios do Grande ABC mostra-nos que os povos originários, a migração e a imigração podem ser responsáveis pela grande variação e mudança linguísticas na fala dos habitantes da região. É indiscutível que os usos linguísticos na região do Grande ABC é fruto de influências de diversos povos, além dos originários, que já habitavam essa terra antes da vinda dos portugueses para o Brasil.

Ao desembarcarem em São Vicente, os futuros colonizadores portugueses se deparavam com grandes obstáculos para, vindos do litoral, chegarem ao planalto. Destacam-se dois grandes desafios: os caminhos desconhecidos, íngremes e perigosos para transpor a Serra do Mar; e a existência de grande rivalidade entre os povos originários. No planalto, podemos destacar os grupos Guaranis (Carijós) e os grupos jês, como os Guaianás, Maromomis e Kaingangues. Além desses, há registro de guerras sangrentas, nos séculos XVI e XVII, entre os Tupiniquins e o Tupinambás, inimigos ferrenhos. É nesse contexto que se evidencia o nome de João Ramalho, o qual se tornara líder das tribos tupiniquins na região do planalto paulista ao se casar com Bartira (Flor de árvore”, em tupi), filha do cacique Tibiriça (“Vigilante da terra”, em tupi). Foi, por intermédio de João Ramalho que os colonizadores portugueses notaram que, acima do litoral e após a serra, havia terras mais povoadas, ricas e férteis. Com esse breve preâmbulo, podemos nos embrenhar na história da região do Grande ABC.

Na verdade, para relatar a origem da região a qual hoje chamamos Grande ABC, a história precisa ser dividida em dois períodos. O primeiro teve início em 1553, quando ocorreu a inauguração da Vila de Santo André da Borda do Campo, com a autorização do então Governador Geral do Brasil, Tomé de Souza. Poucos anos depois, contudo, houve, em 1560, a extinção da Vila de Santo André, que deixa de existir enquanto unidade administrativa e passa a ser bairro de São Paulo de Piratininga, que nascera em 1554 (devido ao interesse dos jesuítas em transferir seu colégio, instalado em São Vicente, para a região dos campos de Piratininga, dada a grande evasão de pessoas do litoral para o interior). As terras da Vila de Santo André, em 1561, foram concedidas como sesmaria a Amador de Medeiros, ouvidor da Capitania de São Vicente. Essa região, depois disso, passou por um período de estagnação, desenvolvendo-se restritamente uma cultura de subsistência e pastagem para os animais de tropeiros que vinham do litoral rumo a São Paulo e ao

interior. Passados 70 anos, em 1631 e em 1637, a maior parte do território foi doada à Ordem de São Bento e formaram-se, respectivamente, as fazendas de São Caetano e São Bernardo. Essas fazendas ficaram sob propriedade dos beneditinos até 1870 (pouco tempo depois do início do segundo período de desenvolvimento da região), quando foram compradas pelo Estado com vistas à criação de colônias de imigrantes).

O segundo período, que é marco relevante na história da região do Grande ABC, iniciou-se em 1860, com a construção da Estrada de Ferro São Paulo Railway, ligando Santos a Jundiaí. Ocorreu, então, uma grande retomada e um rápido desenvolvimento do local, que teve seu perfil consideravelmente mudado. O empreendimento, instalado nas proximidades do rio Tamanduateí, destinava-se à melhoria do escoamento de produtos agrícolas até o Porto de Santos, especialmente o café, cuja produção em larga escala já se iniciava na Província de São Paulo. As facilidades de transportes trazidas para a região atraíram indústrias (em geral ligadas à produção química, têxtil e moveleira) que, contando principalmente com os imigrantes como mão-de-obra e com incentivos fiscais, se instalaram, inicialmente, nas áreas próximas à ferrovia e ao rio. Formaram-se, desse modo, núcleos urbanos que se constituíram em povoados e, a partir da década de 1940, foram se tornando independentes e, assim, formando os municípios que hoje compõem a região.

A partir da década de 1950, iniciaram-se transformações que foram sentidas por conta da mudança dos tipos de indústrias que vieram para a região. Investimentos estrangeiros e estatais impulsionaram o crescimento de indústrias no setor metalúrgico, mecânico, automobilístico e de material elétrico. Com isso, o perfil de grande parte da mão-de-obra também se modificou, atraindo trabalhadores de diversas regiões do País. Nas décadas que se seguiram, essas indústrias, passando por grande expansão e posterior declínio, foram determinantes no delineamento histórico-social da região até a atualidade.

Podemos afirmar, diante do exposto, que para a formação sociocultural da região, destacou-se a presença de tropeiros, migrantes e imigrantes.

Essa breve explanação, mesmo que superficial, é suficiente para concluirmos que a transfiguração que se fez presente nas relações entre as pessoas provocou transformações na realidade das comunidades da região e de seus usos linguísticos.

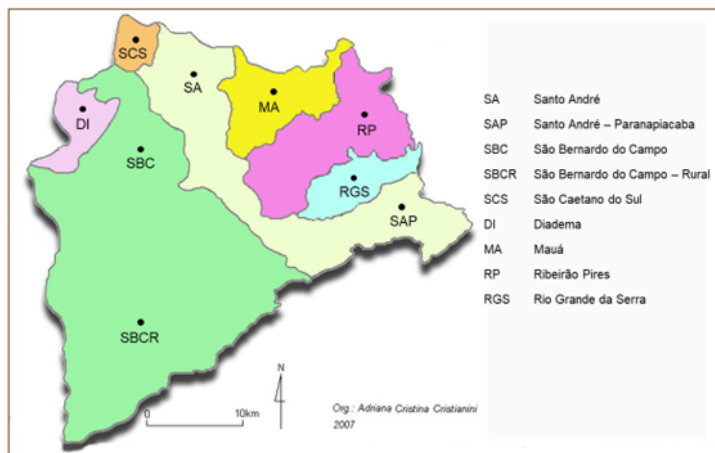
De modo a estudar e registrar a paisagem linguística da região, com base nos postulados da Geolinguística contemporânea, um questionário semântico-lexical, com 202 questões, foi aplicado em nove pontos da região. Para a determinação da rede de pontos estabelecida para a elaboração do *ASL_ABC*, foram considerados os seguintes critérios, segundo Cristianini (2007): a) cada município necessariamente teve, pelo menos, um ponto de pesquisa; b) a densidade demográfica dos sete municípios da região, na qual observou-se que dois dos sete municípios comportavam mais da metade da população da região, quais sejam, Santo André que abriga 27,6% da população de toda a região e, São Bernardo do Campo, com 29,8%; os dois municípios juntos, portanto, abrigam 57,4% de toda população do Grande ABC; esses dois municípios foram contemplados com mais um ponto cada; c) os municípios de Santo André e São Bernardo do Campo, além de comportar maior densidade demográfica, são igualmente muito maiores em área geográfica, o que também justificou a determinação de dois pontos em cada um desses municípios; d) São Bernardo do Campo apresenta uma divisão geográfica natural que evidenciou a divisão do município entre zona urbana e zona rural², o que foi considerado relevante para que se estabelecesse a necessidade de um ponto na área urbana e outro na área rural; e) Santo André contém uma região considerada área histórica, Paranapiacaba, que tem um certo isolamento da área urbana do município e isso foi determinante para a fixação dos dois pontos desse município.

A delimitação da rede de pontos do *ASL_ABC*, portanto, constitui-se em: Santo André; Santo André – Paranapiacaba; São Bernardo do Campo; São Bernardo do Campo – Zona rural; São

² Segundo o Censo de 2000 do Instituto Brasileiro de Geografia e Estatística – IBGE. Neste trabalho, mantivemos os dados demográficos utilizados no *ASL_ABC* por Cristianini (2007).

Caetano do Sul; Diadema; Mauá; Ribeirão Pires; e Rio Grande da Serra. A Figura 2 apresenta os nove pontos da pesquisa na região do Grande ABC.

Figura 2– Pontos da pesquisa



Fonte: Elaborada pela autora.

Em cada ponto, foram entrevistados quatro sujeitos, somando-se um total de 36 sujeitos subdivididos em duas faixas etárias – 18 a 30 e 50 a 65 anos de idade – e nos dois gêneros – feminino e masculino.

A determinação do perfil dos sujeitos atendeu principalmente à questão geográfica. Então, participaram da pesquisa apenas sujeitos nascidos no local ou residentes na região por, ao menos, dois terços da vida. Também foi observada a variável escolaridade: determinou-se um nível de escolaridade de até a oitava série do Ensino Fundamental para a pesquisa, considerada baixa escolaridade.

Com as respostas obtidas nas entrevistas, elaborou-se um banco de dados semântico-lexical e, posteriormente, esses dados foram tratados e documentados cartograficamente, a fim de que se pudesse elaborar o quadro da diversidade semântico-lexical do falar da região do Grande ABC.

Além do referencial teórico-metodológico da Geolinguística e da Lexicologia, a partir de obras de Pottier (1978) e Barbosa (1978, 1981, 1989), utilizou-se a abordagem de Norma efetuada por Coseriu ([1954] 1991, 1973), e complementou-se com noções de Estatística Lexical, propostas por Muller (1968), sobretudo referentes a frequência.

O atlas, constituído por um conjunto de 202 cartogramas linguísticos, registrou, ainda que parcial, o cabedal linguístico que consiste na norma semântico-lexical da região do Grande ABC paulista.

METODOLOGIA

O presente estudo dialetométrico segue os princípios teórico-metodológicos da Escola Dialetométrica de Salzburgo, respeitando os procedimentos, a interpretação e a análise dos dados preconizados pelos estudos dialetométricos de áreas dialetais da língua portuguesa a partir de atlas linguísticos, na Universidade de Lisboa, tais como: Brissos (2016); Brissos, Gillier e Saramago (2016, 2017); e Brissos e Saramago (2019).

Para a realização do presente trabalho, revisitamos o *ASL_ABC* para análise e delimitação do *corpus* utilizado. Para tanto, analisamos mais de seis mil itens lexicais, fruto das respostas de 36 sujeitos entrevistados, pertencentes aos sete municípios que abrangeram a região estudada, com nove pontos de pesquisa. Além disso, ouvimos em torno de 70 horas de gravação de áudio, referentes às entrevistas. Também estudamos 202 cartogramas e os gráficos que demonstram a

variação semântico-lexical na região, bem como as considerações relacionadas aos gêneros e às faixas etárias dos sujeitos entrevistados.

Na sequência, transcrevemos todos os dados e os organizamos em planilhas e contabilizamos os itens lexicais para compor o banco de dados e posterior análise dialetométrica. Além da elaboração da planilha com todos os dados linguísticos do atlas, também elaboramos planilhas para respostas dos sujeitos de cada faixa-etária e de cada gênero.

Para o presente trabalho, optamos por fazer um recorte, apoiando-nos em parâmetros já utilizados em estudos dialetométricos anteriores, e utilizar as respostas dos sujeitos da segunda faixa etária³ e, para a seleção dos cartogramas dos conceitos que comporiam a pesquisa, seguimos tais critérios:

- haver respostas de todos os sujeitos da segunda faixa-etária, em todos os pontos da pesquisa;
- não ser um cartograma com respostas mononímicas, isto é, não apresentar a mesma resposta em todos os pontos e por todos os sujeitos.

Respeitando esses critérios, utilizamos os fenômenos linguísticos correspondentes a 74 conceitos distribuídos em 14 áreas semânticas⁴, constituindo um corpus com 1332 itens lexicais para a análise dialetométrica.

A dialetometrização dos dados e a cartografia dos resultados foram efetivadas por meio do *software DiaTech* (Aurrekoetxea *et al.*, 2013) e, para isso, utilizamos dois parâmetros de análise dos resultados: Análise de *cluster*⁵ e Mapa sinóptico.

Iniciamos pelos métodos de organização hierárquico-aglomerativa, ou seja, de análise de *cluster*, em que todos os pontos da pesquisa começam o processo de análise separados e vão sendo agrupados um a um, até que se obtenha um único *cluster* (ou agrupamento), contendo todos os anteriores. No *software*, para análise de *cluster* (ou análise de agrupamentos), de acordo com o aspecto linguístico semântico-lexical do nosso trabalho, selecionamos: Índice de similaridade – IRI⁶; e Algoritmo de análise estatística – “*Ward*”⁷. Na sequência, passamos para a análise com o parâmetro de Mapa sinóptico. Nesse parâmetro, empregamos dois tipos de operação estatística: a distribuição de similaridade e a distribuição de assimetria (ou *skewness*).

³ Cabe esclarecer que a opção por utilizar respostas de sujeitos da segunda faixa etária deve-se ao fato de constatarmos que os sujeitos de idade maior geraram respostas mais diversificadas e produtivas para a elaboração do *ASL_ABC*.

⁴ Conceitos considerados para a análise dialetométrica, distribuídos em 14 áreas semânticas: a) Acidentes geográficos – córrego, pinguela, onda de mar; b) Fenômenos atmosféricos – temporal, chuva de pedra, garoa, orvalho, nevoeiro; c) Astros e tempo – nascer (do sol), pôr (do sol); d) Atividades agropastoris – camomila, banana dupla, coração, espiga, sabugo, vagem de feijão, carrinho de mão, haste do carrinho de mão, jacá, picada; e) Fauna – urubu, João-de-barro, papagaio, crina do pescoço, lombo, úbere, rabo, mosca varejeira, bicho de fruta; f) Corpo humano – desdentado, meleca, nuca, axila, canhoto, seios, útero; g) Ciclos da vida – parteira, dar à luz, madrastra, finado; h) Convívio e comportamento social – pessoa sovina, mau pagador, bêbado, cigarro de palha, toco de cigarro; i) Religião e crenças – diabo, fantasma, feitiço, benzedeira; j) Jogos e diversões infantis – bolinha de gude, papagaio de papel, esconde-esconde, balanço; k) Habitação – tramela, vaso sanitário, lanterna, interruptor de luz; l) Alimentação e cozinha – café da manhã, curau, aguardente, empanturrado, glutão, pão francês, pão bengala; m) Vestuário e acessórios – sutiã, cueca, calcinha, grampo (com pressão), diadema; n) Vida urbana – sinaleiro, lombada, calçada, ônibus urbano, bodega.

⁵ Análise de *cluster* é o nome dado a um conjunto de técnicas utilizadas na identificação de padrões em bancos de dados por meio de formação de grupos tentativamente homogêneos. Segundo Brissos, Gillier e Saramago (2017, p. 17), “A técnica dialetométrica por excelência que nos permite identificar ou decompor uma dada área em grupos dialetais é a *análise cluster* ou *dendrográfica*, que segmenta, num diagrama em árvore, os núcleos dialetais num nível crescente de homogeneidade a partir do tronco da árvore” (grifo dos autores).

⁶ Os cálculos estatísticos ocorrem a partir das diferenças encontradas, observando-se os Índices Relativos de Diferenças (IRD) ou os Índices Ponderados de Diferenças (IPD); e, a partir das similaridades, por meio dos Índices Relativos de Identidade (IRI) ou dos Índices Ponderados de Identidade (IPI). O IRI, principal índice dialetométrico da Escola Dialetométrica de Salzburgo (Goebel, 2010), pode ser considerado como um índice padrão para a medição de similaridade linguística ou, em nosso estudo, lexical. O IRI constitui um quociente que, em última análise, mede a percentagem de identidades linguísticas entre dois pontos do atlas. O *software DiaTech* estabelece relação dialetal entre um ponto de referência e os outros pontos, demonstrando a aproximação ou o distanciamento linguístico, considerando o IRI. Para uma descrição mais detalhada acerca do IRI, vale consultar Goebel (1981, p. 357-360; 1984, p. 74-77).

⁷ *Ward* – Método estatístico para formação de grupos em que se dá a maximização da homogeneidade dentro dos grupos. Pode-se encontrar mais informações em Goebel e Smečka (2014, p. 460).

Com a análise de distribuição de similaridade, podemos verificar a hierarquia de semelhança/diferença do ponto específico relacionando-o com cada um dos pontos restantes.

Segundo Brissos (2020), ao analisarmos o grau de similaridade em relação a um determinado ponto, há necessidade de observarmos que os valores acima da média no histograma gerado são representados com cores quentes e à direita. Os pontos com valores de similaridade abaixo da média, por sua vez, são representados com cores frias e à esquerda no histograma. Isso significa que quanto mais quente a cor, mais alta é a similaridade do ponto analisado (que sempre é representado com a cor branca) e, quanto mais fria, mais baixa a similaridade. Tal quer dizer que, no histograma, a cor vermelha indica um nível de similaridade máxima, enquanto a cor azul, um nível mínimo.

Para análise da distribuição de similaridade, rodamos o *software DiaTech* com as seguintes especificações: Algoritmo estatístico – MinMwMax⁸; Unidade de distância – IRI; e Localização – um ponto diferente da pesquisa por vez. Com os resultados, elaboramos os cartogramas de similaridade.

A análise de distribuição de assimetria, por sua vez, faz a síntese das cartas de similaridade de todos os pontos e nos permite observar quais são os pontos mais e menos integrados no conjunto constituído pelo *corpus* linguístico utilizado no estudo. Segundo Brissos, Gillier e Saramago (2017, p. 19), “Esse parâmetro parte do facto de que nem todos os locais têm o mesmo grau de semelhança com os restantes locais do conjunto analisado e apresenta, a partir do coeficiente de assimetria de Fisher⁹ (CAF), uma síntese relacional dos locais mais e menos semelhantes com o conjunto”.

Para gerarmos os resultados da distribuição de assimetria no *DiaTech*, selecionamos as mesmas especificações que foram utilizadas para a análise da similaridade. A localização aqui deixa de ser solicitada, visto que a análise é do conjunto.

Passamos, a seguir, a discorrer sobre os resultados fornecidos pelo *DiaTech*.

RESULTADOS

Dividimos esta seção em duas partes para contemplarmos os dois parâmetros empregados na análise: análise de cluster e mapa sinóptico.

Iniciamos esta análise discutindo os cartogramas elaborados a partir dos resultados obtidos pelo *software DiaTech* com o corpus de 74 fenômenos linguísticos.

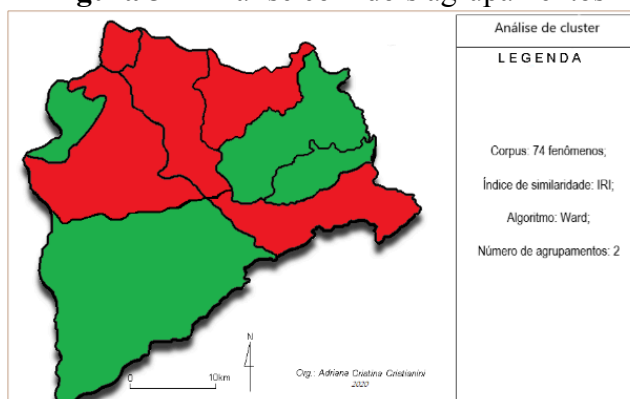
Percebemos que, ao analisarmos dois agrupamentos, observamos que são constituídos da seguinte maneira: 1) Rio Grande da Serra, Diadema, Ribeirão Pires e São Bernardo do Campo – Zona rural (verde); e 2) São Bernardo do Campo, Mauá, Santo André, Santo André – Paranapiacaba e São Caetano do Sul (vermelho). A Figura 3 apresenta o cartograma de análise de cluster com dois agrupamentos¹⁰:

⁸ MinMwMax é um algoritmo, geralmente utilizado pela Escola de Dialeto-metria de Salzburgo, que cria intervalos simétricos (n intervalos, cujo número deve ajustar-se ao conjunto de dados e ao número de pontos presentes no estudo) em ambos os lados da média aritmética (Mw), para baixo (valor mínimo – Min) e para cima (valor máximo – Max). Com o MinMwMax o intervalo entre a Mw e o Min e o intervalo entre o Max e Mw são divididos por n/2 a fim de obter os limites numéricos dos n intervalos. Para mais informações a respeito, consulte Brissos, Gillier e Saramago (2017).

⁹ Coeficiente de **assimetria de Fisher** é um dos coeficientes que, na estatística, se usa para calcular a assimetria de uma distribuição, permitindo-nos saber se uma distribuição de probabilidade é distorcida positivamente, distorcida negativamente ou simétrica. Para maiores informações, consulte Academia Balderix (2022).

¹⁰ Para este trabalho, elaboramos os cartogramas com os resultados em uma representação dos municípios, apesar de os mapas gerados pelo *DiaTech* apresentarem a poligonização de Voronoi (mosaico poligonal), considerando que esta assume uma dimensão espacial idêntica para cada ponto, o que, numa rede reduzida de pontos, como ocorre neste estudo, é especialmente falível. Para maiores informações, veja Saramago (2020).

Figura 3 – Análise com dois agrupamentos

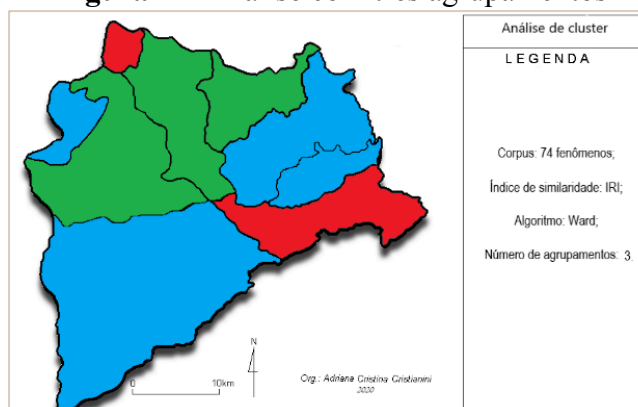


Fonte: Elaborada pela autora.

A análise com três agrupamentos, ou mais, evidencia informações adicionais que já nos permitem algumas inferências, pois possibilita associar os resultados com aspectos sócio-histórico-culturais e geográficos da região.

Na Figura 4, a seguir, observamos os agrupamentos das localidades da seguinte forma: 1) Rio Grande da Serra, Diadema, Ribeirão Pires e São Bernardo do Campo – Zona rural (azul); 2) São Bernardo do Campo, Mauá e Santo André (verde); e (iii) Santo André – Paranapiacaba e São Caetano do Sul (vermelho).

Figura 4 – Análise com três agrupamentos



Fonte: Elaborada pela autora.

Historicamente, Santo André – Paranapiacaba e São Caetano do Sul são dois pontos que receberam considerável influência de culturas de imigrantes. São Caetano do Sul, no decorrer de sua história, foi uma fazenda destinada ao recebimento de imigrantes europeus, mais especificamente, tornou-se colônia para imigrantes italianos principalmente provenientes de Treviso. Essas terras, na ocasião, estavam praticamente abandonadas e, então, o governo a tornou um objeto de acordo com os italianos que, inicialmente, recebiam lotes, casa e alimentos desde que produzissem na região. Esse fato é um fator importante para a identidade do município. Geograficamente, São Caetano do Sul, município exclusivamente urbano, não possui condições de expansão e, portanto, comparado aos outros municípios da região, abrigou um menor número de

emigrantes dentre os que vieram para a região por ocasião da instalação e da expansão das indústrias (Consórcio..., 2020).

A história de Paranapiacaba, por sua vez, também é marcada pela presença de imigrantes europeus, em geral, ingleses e, depois, portugueses. Paranapiacaba surgiu na década de 1860, como acampamento para trabalhadores (muitos deles de origem inglesa) da empresa The São Paulo Railway Company Ltd. – SPR, responsável pela construção do trecho da Serra do Mar da estrada de ferro na região (Sistema Funicular). Mesmo com a inauguração da ferrovia, os operários foram mantidos no local para operacionalização dos serviços e manutenção das obras.

Na década de 1940, o local passou por algumas mudanças, pois a estrada de ferro foi incorporada ao Patrimônio da União e a presença dos ingleses na região foi diminuindo. Transformações ocorreram com a marcante presença da cultura portuguesa. No tempo dos ingleses, segundo Santo André (2013), a vila de Paranapiacaba mantinha características britânicas, com ar romântico,

[...] com casas de madeira, quintais separados por cercas vivas e ruas calmas, ladeadas de pinheiros, em contraste com a Parte Alta, que recebeu uma ocupação urbana marcada pela herança portuguesa, com ruas estreitas e casas de pequenas frentes edificadas junto ao alinhamento. Unindo a Parte Alta à Parte Baixa há uma ponte metálica destinada exclusivamente aos pedestres e bicicletas, que se mantém até hoje após algumas reformas. (Santo André, 2013).

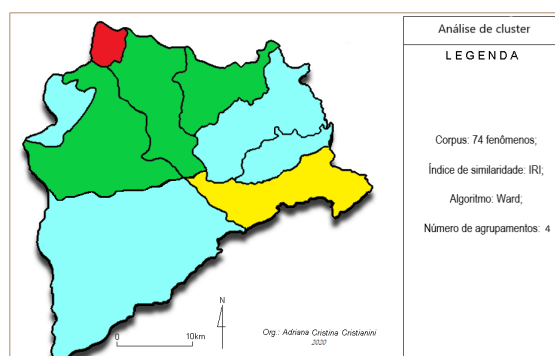
Na década de 1980, com o fim do funcionamento do Sistema Funicular, finalizou-se uma era e iniciou-se uma luta pela preservação histórica e ambiental do local. Um movimento para que Paranapiacaba se tornasse um polo turístico foi promovido e, posteriormente, foi publicado seu tombamento histórico legalizando o local como de interesse público. Esse tombamento abrange a área urbana, os equipamentos ferroviários e a área natural ao seu redor.

A vila foi incorporada ao município de Santo André em 2001 e, em 2003, foi criado o Parque Natural Municipal Nascentes de Paranapiacaba, área verde de Mata Atlântica no entorno da Vila. Além disso, também no ano de 2003, Paranapiacaba tornou-se um dos núcleos do programa da Reserva da Biosfera da Organização das Nações Unidas para a Educação, a Ciência e a Cultura – UNESCO¹¹.

Cabe ressaltar, entretanto, que, ao analisarmos os dados com quatro agrupamentos, na Figura 5, São Caetano do Sul e Santo André – Paranapiacaba não se mantêm no agrupamento. Podemos observar que dois agrupamentos se destacam com maior número de pontos e dois outros agrupamentos constituem-se em grupos unitários, quais sejam: 1) Santo André, São Bernardo do Campo e Mauá (verde); 2) São Bernardo do Campo – Zona rural, Diadema, Ribeirão Pires e Rio Grande da Serra (azul); 3) São Caetano do Sul (vermelho); e 4) Santo André – Paranapiacaba (amarelo).

Figura 5 – Análise com quatro agrupamentos

¹¹ Segundo a UNESCO (2020), as Reservas da Biosfera buscam conciliar “[...] a conservação da diversidade biológica e cultural com o desenvolvimento econômico e social, por meio de parcerias entre pessoas e a natureza”. Cabe destacar que Paranapiacaba é cercada por três importantes Unidades de Conservação (áreas protegidas): o Parque Nascentes; a Reserva Biológica do Alto da Serra de Paranapiacaba; e o Parque Estadual da Serra do Mar.

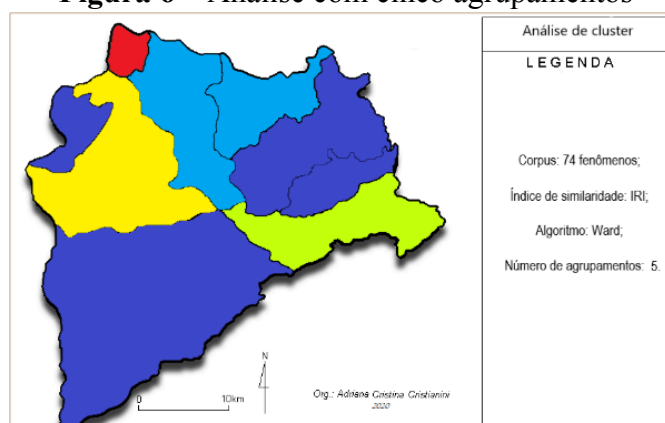


Fonte: Elaborada pela autora.

Essa observação nos permite inferir que, apesar de esses dois pontos, quando há menos agrupamentos, se aproximarem em um mesmo grupo, ao analisarmos um número maior de agrupamentos, eles se distanciam. A distância geográfica que existe entre os dois pontos pode justificar esse resultado. Há de se considerar, também, que Santo André – Paranapiacaba constitui-se em um ponto que se mantém consideravelmente isolado geograficamente dos demais por estar localizado próximo ao topo da Serra do Mar.

Em uma análise com cinco agrupamentos, conforme a Figura 6, temos: 1) São Bernardo do Campo – Zona rural , Diadema, Ribeirão Pires e Rio Grande da Serra (azul escuro); 2) Santo André e Mauá (azul claro); 3) São Bernardo do Campo (amarelo); 4) São Caetano do Sul (vermelho); e 5) Santo André – Paranapiacaba (verde).

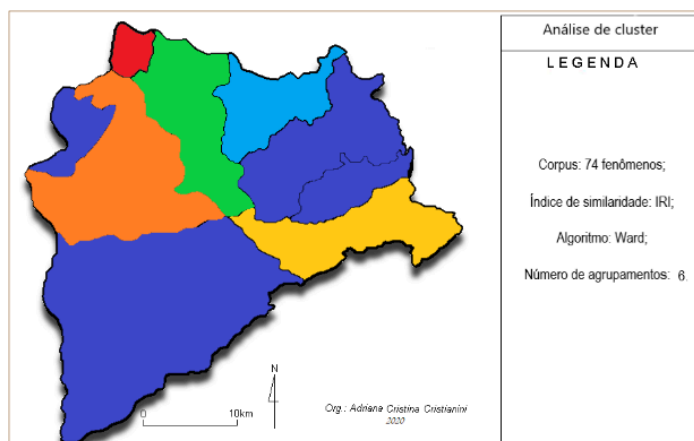
Figura 6 – Análise com cinco agrupamentos



Fonte: Elaborada pela autora.

Na Figura 7, temos os seis agrupamentos: 1) São Bernardo do Campo – Zona rural , Diadema, Ribeirão Pires e Rio Grande da Serra; 2) Santo André; 3) Mauá; 4) São Bernardo do Campo; 5) São Caetano do Sul ; e 6) Santo André – Paranapiacaba. Podemos observar, aqui, que São Bernardo do Campo – Zona rural , Diadema, Ribeirão Pires e Rio Grande da Serra (azul escuro) mantêm-se agrupados, enquanto os pontos Santo André (verde), Mauá (azul claro), São Bernardo do Campo (laranja), São Caetano do Sul (vermelho) e Santo André – Paranapiacaba (amarelo) constituem-se em grupos unitários.

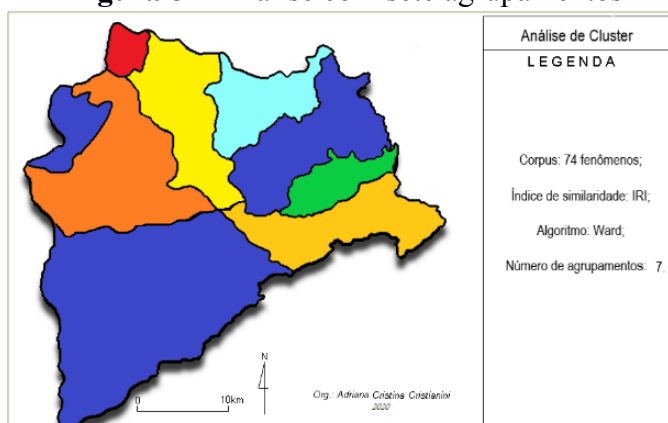
Figura 7 – Análise com seis agrupamentos



Fonte: Elaborada pela autora.

A tendência mantém-se ao observarmos um corte com sete agrupamentos, na Figura 8. Aos grupos unitários, acrescenta-se o ponto Rio Grande da Serra (verde) que, com menos agrupamentos, integrava o grupo maior. Temos, assim: 1) São Bernardo do Campo – Zona rural , Diadema, Ribeirão Pires (azul escuro); 2) Santo André (amarelo); 3) Mauá (azul claro); 4) Rio Grande da Serra (verde); 5) São Bernardo do Campo (laranja); 6) São Caetano do Sul (vermelho); e 7) Santo André – Paranapiacaba (ocre).

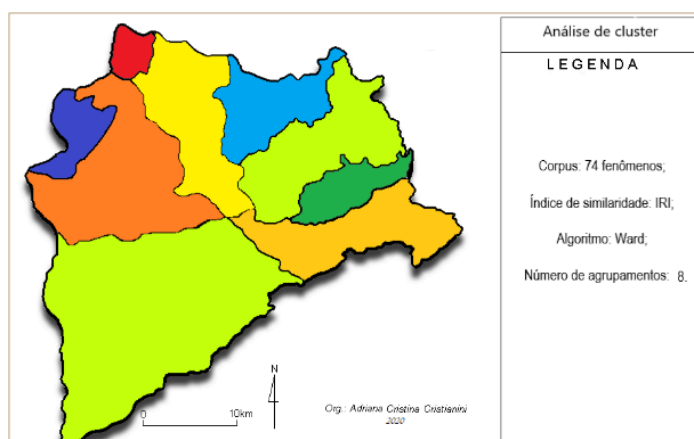
Figura 8 – Análise com sete agrupamentos



Fonte: Elaborada pela autora.

Com a possibilidade máxima de agrupamentos (oito) para a quantia de pontos de nossa pesquisa, na Figura 9, dois pontos mantêm-se agrupados, São Bernardo do Campo – Zona rural e Ribeirão Pires e todos os outros constituem-se em grupos unitários, quais sejam: 1) São Bernardo do Campo – Zona rural e Ribeirão Pires (verde claro); 2) Santo André (amarelo); 3) Mauá (azul claro); 4) Rio Grande da Serra; 5) São Bernardo do Campo; 6) Diadema (azul escuro); 7) São Caetano do Sul (vermelho); e 8) Santo André – Paranapiacaba (ocre).

Figura 9 – Análise com oito agrupamentos



Fonte: Elaborada pela autora.

Podemos observar, por meio dos diversos números de agrupamentos, que existe uma nítida falta de coesão geográfica nos agrupamentos, possivelmente justificada pela composição geográfica, pelo processo histórico pelo qual constituiu-se a região, e pela realidade socioeconômica de sua população atual. A região do grande ABC, aliás, não é considerada uma região coesa em muitos aspectos.

Geograficamente, é importante destacar a Serra do Mar e os vários caminhos que ligam o planalto ao litoral. Os portugueses rapidamente perceberam que havia trilha que acompanhava o Rio Mogi. O percurso dessa trilha, no século XIX, foi utilizado pelos engenheiros ingleses para a construção da Estrada de Ferro Santos-Jundiaí. Quando os jesuítas chegaram ao litoral e resolveram fundar um núcleo no planalto, primeiramente utilizavam a trilha do Rio Mogi, mas a proximidade com grupos originários hostis fez com que cogitassem a abertura de uma nova trilha, afastada daquele rio. Coube ao padre José de Anchieta empreender a abertura da nova trilha que ficou conhecida como “Caminho do Padre José”. Esse caminho, contudo, à época era considerado muito ruim, havendo registro que o tratam como um dos “piores do mundo”, pois só permitia o trânsito em filas indianas e agarrando-se às raízes das árvores, as cargas eram carregadas nos ombros e os doentes, quando necessário, carregados em redes. Essa trilha, ainda que tenha sido melhorada na segunda metade do século XVII, com construção de pequenas pontes, cortes e aterros, muros de arrimo e, no planalto, uma estrada de troncos, era insuficiente e impraticável, principalmente na época das chuvas. O Caminho do Padre José é atualmente demarcado por pesquisadores com lacunas decorrentes de deslizamentos que ocorrem na serra. Ainda que de forma rudimentar, o Caminho do Padre José foi praticamente a única ligação do litoral com o planalto até 1792, quando foi construída a Calçada do Lorena. Essa estrada, diferentemente da anterior, mesmo para padrões europeus, foi considerada uma construção muito avançada para a época e foi um marco histórico e comercial, visto que impulsionou o comércio nacional e internacional. Em 1840, mais uma rota começou a ser construída, a Estrada da Maioridade, que, mais tarde, foi adaptada para o tráfego de automóveis, já no século XX e, atualmente, recebe o nome de Caminho do Mar. Também foi construída, no século XIX, um caminho para percurso no planalto, após a subida da serra, a **Estrada do Vergueiro**. Em 1947, ocorre a construção da Rodovia Anchieta, que marca o início de uma fase de acelerada transformação e grande desenvolvimento econômico, social e urbano. Outro caminho de extrema importância, que liga Santos a São Paulo, cortando os municípios de São Bernardo do Campo e Diadema, é a Rodovia Imigrantes. Grosso modo, esses são os principais caminhos¹² por onde circulavam/circulam, além de riquezas, produtos e mercadorias, famílias, homens e mulheres escravizados, migrantes, imigrantes, turistas, ideias e palavras.

¹² Para mais informações, pode-se consultar, entre outros, São Paulo (2014), e Nascimento (2021).

Cabe destacar também que, em 1925, na região do Grande ABC foi criada a represa Bilings¹³, cuja bacia hidrográfica estende-se pelos municípios de Ribeirão Pires, Diadema, Rio Grande da Serra, São Bernardo do Campo, Santo André e São Paulo. Geograficamente, a represa Bilings constitui-se uma divisão da região em duas partes, considerando que cinco, dos sete municípios da região, receberam o maior espelho d'água da América do Sul. Inferimos que essa barreira geográfica tenha colaborado para a falta de coesão dialetal a que nos referimos.

Historicamente, é importante destacar as influências da imigração e da migração, juntamente com presença dos povos originários, em especial dos grupos Tupi e Guarani, em nossa cultura. A Língua Portuguesa Brasileira vai-se constituindo por aquela trazida pelos portugueses nos tempos de dominação e exploração, com evidente influência das línguas faladas pelos povos originários e, também, das outras línguas e dialetos trazidos com os fluxos migratórios e imigratórios, no decorrer dos séculos. Não se pode omitir a presença dos clérigos, sempre presentes na região, que contribuíram para a constituição lexical da região e são responsáveis, entre outras coisas, pela nomeação dos três principais municípios da região do Grande ABC, com nomes de três santos: André, Bernardo e Caetano. Sabe-se da intensa chegada de negros africanos, que perdurou até o século XIX. A vinda dos imigrantes de muitos países foi intensificada no final do século XIX e no início do século XX, perduram até os dias atuais¹⁴. O estado de São Paulo foi a principal região de atração de imigrantes no Brasil (com quase sessenta por cento do total dos estrangeiros entrados no país). Até há poucas décadas, a maioria desses imigrantes adentraram em nosso território pelo Porto de Santos e, então, subiam a serra para a região do Grande ABC e para a cidade de São Paulo. Aqueles que tinham contrato de trabalho continuaram seu percurso para outras localidades, mas muitos se estabeleceram na região. E, assim, os municípios e os caminhos do Grande ABC foram se tornando cada vez mais importantes no contexto regional e nacional. Destaca-se a imigração de italianos, portugueses, espanhóis, alemães, judeus, árabes, japoneses, entre outras. Cabe chamar a atenção para os ingleses, que chegaram para a construção da Estrada de Ferro Santos-Jundiaí a partir de 1860 e se fixaram, primeiramente em Paranapiacaba e, depois, próximos ao percurso da ferrovia que, na região do Grande ABC, passa por Paranapiacaba, Rio Grande da Serra, Ribeirão Pires, Mauá, Santo André e São Caetano do Sul.

Além dos imigrantes, a região do Grande ABC tem, em sua história, fluxos migratórios de diversas partes do país. Segundo o censo de 2010, divulgados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), um quarto da população da região do Grande ABC é constituída de migrantes. A mobilidade interna na região do Grande ABC também é constante, visto que, por vários motivos, as famílias se mudam dentro das sete cidades da região (IBGE, 2010).

Os processos, quer seja de migração, quer seja de imigração, existentes durante toda a história da região do Grande ABC, intensificados em alguns períodos, inevitavelmente geraram e geram impactos muito importantes no desenvolvimento socioeconômico, demográfico e, em especial, nos costumes e na cultura do povo. Como a língua é o principal elemento cultural de um povo e o meio de transmissão dessa cultura, evidentemente, a LPB reflete todas as influências de povos de diferentes localidades do Brasil e do Mundo.

Outro aspecto a se considerar ao se tratar da não coesão dialetal na região do Grande ABC está relacionado a questões socioeconômicas da região. Há de se considerar que a Região do Grande ABC, conforme pontua Silva (2013, p. 175), deve ser entendida como “[...] um universo de práticas vivenciadas pelos diversos grupos humanos que nela se inserem; que englobam o relevo, as relações pessoais, a memória familiar, as condições de trabalho, a sexualidade, a associação, entre outras”. A delimitação dos seus municípios, portanto, não é estabelecido pelos limites territoriais e políticos, mas por quem vive na localidade. Muitas vezes, os limites territoriais são até desconhecidos pela população ou por quem passa pela região, visto que, em alguns pontos, chega a ser impreciso em

¹³ Para mais informações acerca da Represa Bilings, veja Castilho (1997).

¹⁴ Segundo o IBGE, estima-se que 4 milhões de africanos foram trazidos à força para o Brasil entre os séculos XVI e meados do século XIX. Mais informações sobre a imigração brasileira podem ser verificadas na página “Brasil 500 anos”, do IBGE (2000).

que município se está. Isso ocorre, por exemplo, porque há locais onde, em uma rua, de um lado é o município de Santo André e, de outro, São Caetano do Sul.

Quanto aos seguimentos em que estão divididas as atividades produtivas da região, hoje há uma demarcação bem definida: Santo André e Mauá, dedicam-se a setores industriais petroquímicos e de plásticos; São Bernardo do Campo é conhecido pela presença da indústria automobilística e de autopeças; São Caetano do Sul e Diadema, apesar da presença da indústria, destacam-se no setor comercial e de serviços; Ribeirão Pires e Rio Grande da Serra, por estarem em região de mananciais, possuem uma escassez de indústrias.

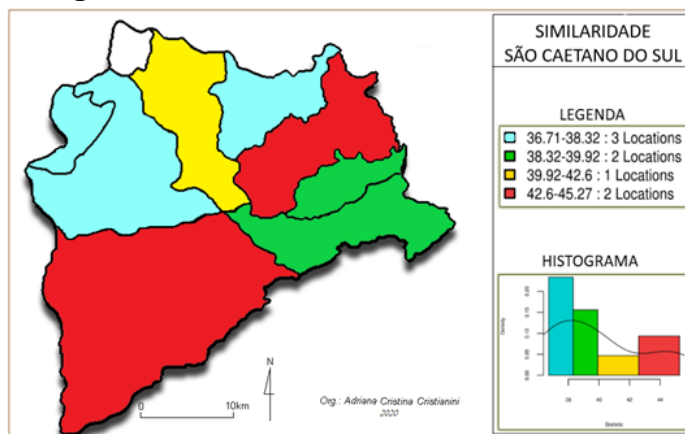
A desigualdade social também é discrepante pois há contraste entre: os três municípios mais ricos, Santo André, São Bernardo do Campo e São Caetano do Sul; e os quatro municípios mais pobres, Diadema, Mauá, Ribeirão Pires e Rio Grande da Serra. Quando nos referimos a questões financeiras, a discrepância é gigantesca. Para exemplificar de maneira concisa, podemos observar que o maior Produto Interno Bruto (PIB) municipal da região é de São Bernardo do Campo que, aliás, é o quarto maior do estado de São Paulo, estado com o maior PIB do Brasil; entretanto, ao se analisar o PIB *per capita*, São Caetano do Sul detém praticamente o dobro, se comparado ao referido município. No que se refere ao Índice de Desenvolvimento Humano (IDH) na região, a discrepância entre os municípios também é gritante. São Caetano do Sul tem o maior IDH do Brasil, sendo considerado o melhor município para se morar em nosso país. Santo André e São Bernardo do Campo ocupam as boas posições 15 e 28, respectivamente, no ranking. Os demais municípios da região, por sua vez, não possuem IDH tão bons, estando nas seguintes posições: Ribeirão Pires, 100; Diadema, 420; e Rio Grande da Serra, 565¹⁵.

Procuramos demonstrar, mediante às breves observações elencadas acerca dos municípios da região do Grande ABC, que a heterogeneidade é muito grande, em diversos aspectos, nos municípios da região. Diante disso, podemos concluir ser inquestionável que há muitos motivos para que essa heterogeneidade seja evidente também nos usos linguísticos da população.

Passamos, agora, à análise por meio de mapas sinópticos.

Ao observarmos os resultados, percebemos que, estatisticamente, São Caetano do Sul, comparado aos demais pontos, apresenta uma baixa similaridade (ou alta diferença), pois a maioria dos pontos estão abaixo da média, como podemos observar na Figura 10.

Figura 10 – Similaridade – São Caetano do Sul



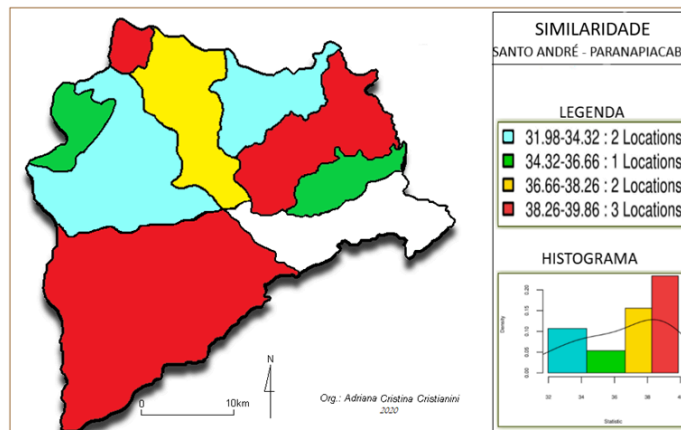
Fonte: Elaborada pela autora.

Percebemos que, com relação a São Caetano do Sul, o histograma mostra uma curva marcadamente à esquerda com predomínio de cores frias, ou seja, a análise indica que há um nível de similaridade baixo se confrontado com os pontos da região como um todo.

¹⁵ Dados do IBGE e de registros administrativos, conforme especificados nos metadados disponíveis em: <http://atlasbrasil.org.br/acervo/biblioteca>.

O mesmo não ocorre com Santo André – Paranapiacaba, pois, como observamos na Figura 11, há uma tendência, no histograma, de valores acima da média, o que significa que há um grau alto de similaridade com a maioria dos outros pontos da região.

Figura 11 – Similaridade – Santo André – Paranapiacaba

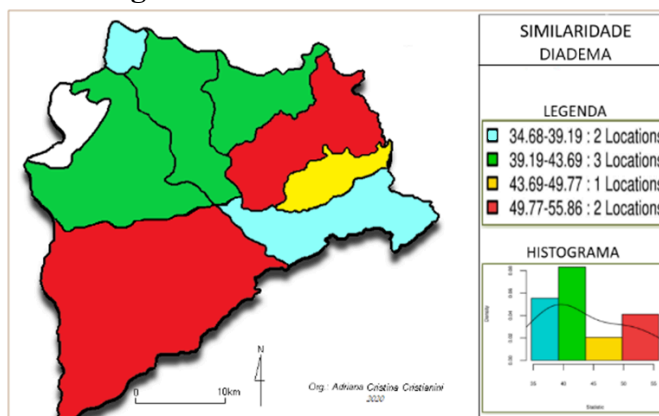


Fonte: Elaborada pela autora.

Ao analisarmos cada um dos outros pontos (Diadema, Mauá, Ribeirão Pires, Rio Grande da Serra, Santo André, São Bernardo do Campo e São Bernardo do Campo – Zona rural) no que se refere ao grau de similaridade, uma consideração foi constante: os usos linguísticos de cada um desses pontos, invariavelmente, possuem um baixo grau de similaridade com São Caetano do Sul e Santo André – Paranapiacaba.

Analisando a similaridade de outros pontos, relacionada ao conjunto, observamos que o ponto Diadema, na Figura 12, pela análise, apresenta um histograma que representa um baixo grau de similaridade, com uma curva à esquerda que significa um nível de similaridade abaixo da média. Como referido anteriormente, São Caetano do Sul e Santo André – Paranapiacaba são pontos com cor que indica baixa similaridade.

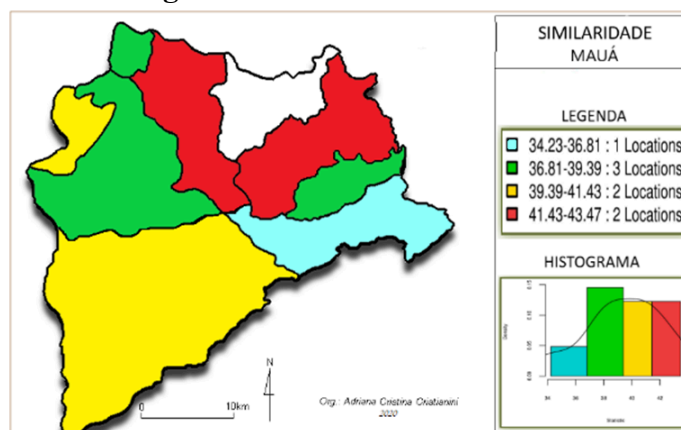
Figura 12 – Similaridade – Diadema



Fonte: Elaborada pela autora.

Mauá, por sua vez, possui um grau de similaridade mais alto, principalmente com os pontos que se constituem em vizinhos em uma continuidade geográfica, conforme observamos na Figura 13. Entretanto, em relação aos pontos São Caetano do Sul e Santo André – Paranapiacaba (além de São Bernardo do Campo e Ribeirão Pires), temos, novamente, um baixo grau de similaridade.

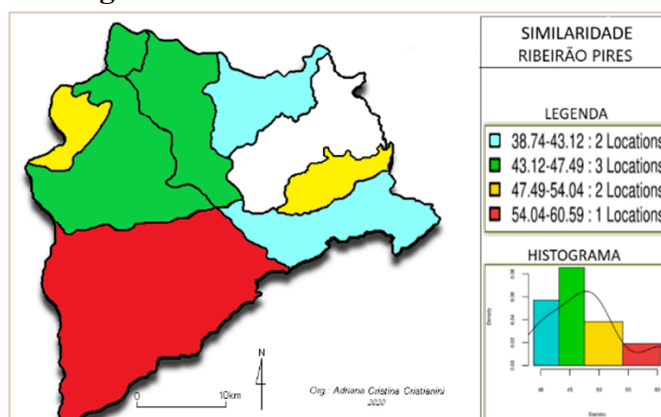
Figura 13 – Similaridade – Mauá



Fonte: Elaborada pela autora.

Ribeirão Pires, conforme Figura 14, apresenta um nível de similaridade abaixo da média, com uma curva no histograma à esquerda. O ponto São Bernardo do Campo – Zona rural é aquele que mais se aproxima a Ribeirão Pires nas questões linguísticas de aspecto semântico-lexical. O mesmo baixo grau de similaridade ocorre com relação a São Caetano do Sul e Santo André – Paranapiacaba.

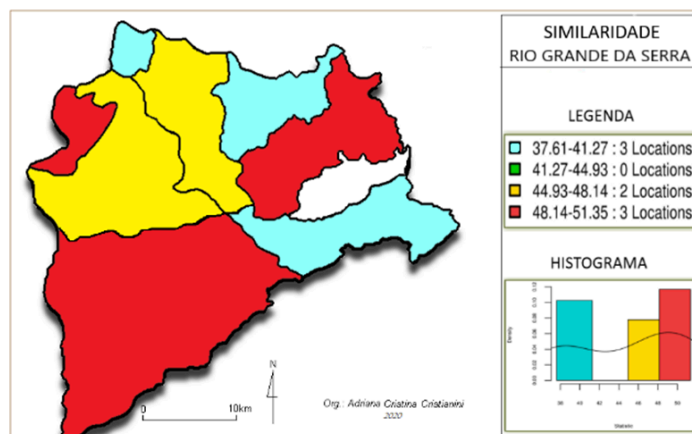
Figura 14 – Similaridade – Ribeirão Pires



Fonte: Elaborada pela autora.

Rio Grande da Serra, conforme Figura 15, apresenta similaridade com uma tendência à direita no histograma, acima da média. Ribeirão Pires, São Bernardo do Campo – Zona rural e Diadema são os pontos com níveis mais altos de similaridade quando relacionados a Rio Grande da Serra. Mais uma vez, observamos que, relacionando Rio Grande da Serra a São Caetano do Sul e Santo André – Paranapiacaba, existe baixa similaridade.

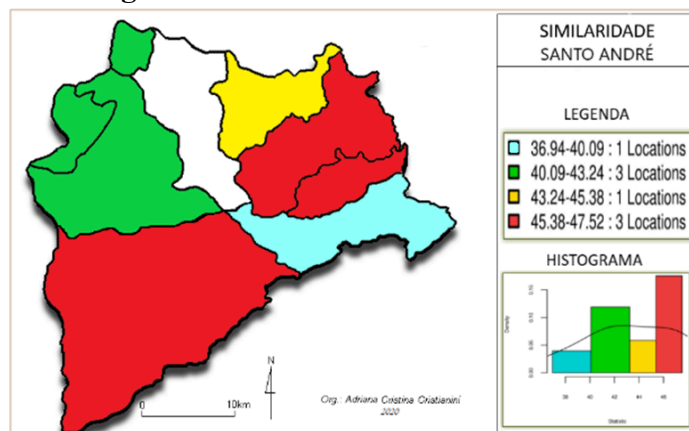
Figura 15 – Similaridade – Rio Grande da Serra



Fonte: Elaborada pela autora.

A Figura 16 mostra que o ponto Santo André possui maior similaridade com São Bernardo do Campo – Zona rural, Ribeirão Pires, Rio Grande da Serra e Mauá, mas o nível de similaridade é baixo se relacionado com São Bernardo do Campo, São Caetano do Sul, Diadema e Santo André – Paranapiacaba. Portanto, São Caetano do Sul e Santo André – Paranapiacaba apresentam baixa similaridade também em relação a Santo André. A curva estatística sugere que a similaridade está levemente acima da média.

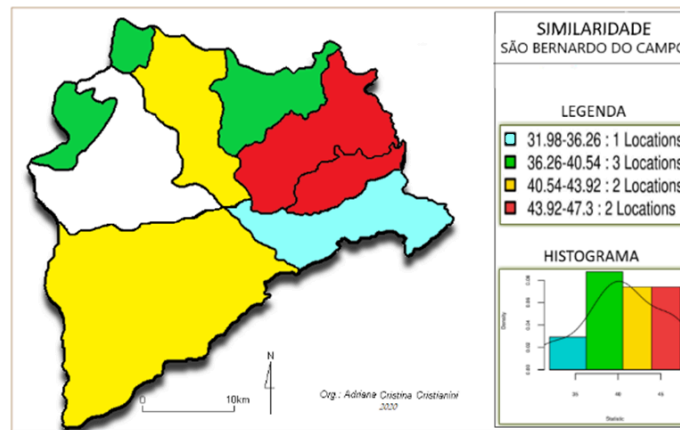
Figura 16 – Similaridade – Santo André



Fonte: Elaborada pela autora.

Pode-se dizer que São Bernardo do Campo, baseado na Figura 17, possui maior similaridade com Ribeirão Pires e Rio Grande da Serra; ainda com grau alto de similaridade com São Bernardo do Campo – Zona rural e Santo André. O nível de similaridade é baixo se relacionado com os pontos Diadema, Mauá, São Caetano do Sul, e, principalmente, Santo André – Paranapiacaba.

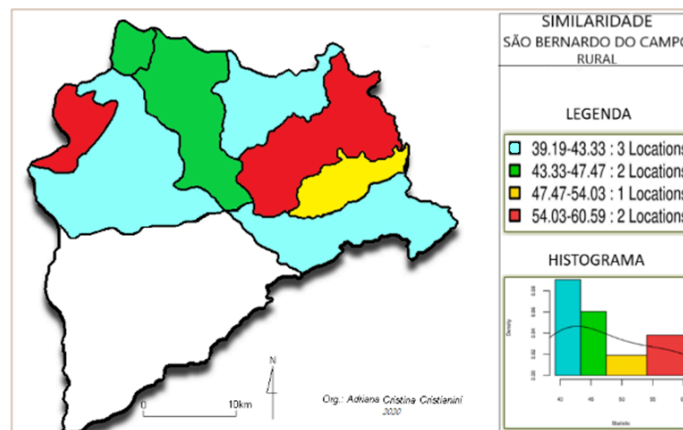
Figura 17 – Similaridade – São Bernardo do Campo



Fonte: Elaborada pela autora.

A Figura 18 apresenta a similaridade do ponto São Bernardo do Campo – Zona rural relacionada aos demais pontos. Observamos que a similaridade desse ponto se apresenta abaixo da média, representada no histograma com uma curva à esquerda. Os pontos Diadema e Ribeirão Pires possuem mais similaridade linguística com São Bernardo do Campo – Zona rural, seguido por Rio Grande da Serra. Os pontos com baixo nível de similaridade quando relacionado a São Bernardo do Campo – Zona rural são São Caetano do Sul, Santo André e, ainda menos similar, São Bernardo do Campo, Mauá e Santo André – Paranapiacaba. Mais uma vez, encontramos São Caetano do Sul e Santo André – Paranapiacaba nos grupos cuja similaridade é considerada baixa.

Figura 18 – Similaridade – São Bernardo do Campo – Zona rural



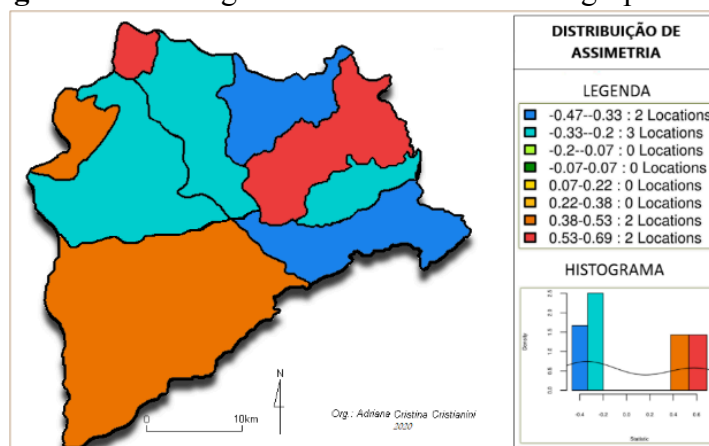
Fonte: Elaborada pela autora.

Diante do exposto, podemos observar que independente de qual seja o ponto selecionado como referência para a análise de similaridade relacionada aos pontos restantes, São Caetano do Sul e Santo André – Paranapiacaba sempre apresentam um grau de similaridade abaixo da média. Somente quando o ponto de referência é o próprio Santo André – Paranapiacaba, temos São Caetano do Sul com grau acima da média.

Entretanto, mesmo Santo André – Paranapiacaba e São Caetano do Sul apresentando resultados semelhantes quando um outro ponto é referência na análise de similaridade, podemos verificar que a integração desses dois pontos no conjunto é divergente quando calculamos a distribuição de assimetria.

Conforme podemos observar na Figura 19, cinco pontos possuem uma integração alta (São Bernardo do Campo, Santo André, Mauá, Rio Grande da Serra e Santo André – Paranapiacaba) e quatro, uma integração baixa na região (Ribeirão Pires, São Bernardo do Campo – Zona rural, Diadema e São Caetano do Sul).

Figura 19 – Cartograma de assimetria – oito agrupamentos



Fonte: Elaborada pela autora.

O cartograma de assimetria nos mostra que, além de os pontos que apresentam baixa integração serem em número menor (quatro) que aqueles com alta integração (cinco), nenhum ponto do grupo com baixa integração no conjunto é contíguo entre si. O mesmo não ocorre no grupo com os pontos com alta integração.

Como nosso olhar, este estudo, está mais direcionado aos pontos São Caetano do Sul e Santo André – Paranapiacaba, é importante ressaltarmos que, no que se refere à integração no conjunto, esses dois pontos possuem resultados opostos, pois São Caetano do Sul possui grau mínimo de integração, com a cor vermelha, e Santo André – Paranapiacaba, o grau máximo, na cor azul.

CONCLUSÃO

Neste trabalho, propusemo-nos a destacar a análise global dos usos linguísticos a partir do *ASL_ABC*, em especial dos pontos São Caetano do Sul e Santo André – Paranapiacaba.

Verificamos que São Caetano do Sul e Santo André – Paranapiacaba, possivelmente devido a sua constituição socio-histórico-geográfica, têm seus dialetos em grupos distintos dos demais se segmentarmos o corpus em mais de três *clusters*. Além disso, quando refletimos por meio do estudo da similaridade partindo dos demais pontos com São Caetano do Sul e Santo André – Paranapiacaba, percebemos que existe sempre um grau de similaridade abaixo da média, exceto quando o ponto de referência é Santo André – Paranapiacaba, com o qual São Caetano do Sul possui um grau alto de similaridade.

Essa última consideração evidencia-se ao observarmos os cartogramas de similaridade dos pontos Diadema, Mauá, Ribeirão Pires, Santo André, São Bernardo do Campo e São Bernardo do Campo – Zona rural com o conjunto. Neles, São Caetano do Sul aparece no escalão verde nos cartogramas de similaridade de Mauá, Santo André, Ribeirão Pires, São Bernardo do Campo e São Bernardo do Campo – Zona rural e no escalão azul nos cartogramas Diadema e Rio Grande da Serra, sempre indicando grau de similaridade abaixo da média. Santo André – Paranapiacaba, por

sua vez, surge no escalão azul em todos esses cartogramas, indicando, também, grau baixo de similaridade.

Partindo dos cartogramas de similaridade de São Caetano do Sul e de Santo André – Paranapiacaba, relacionando-os ao conjunto, percebemos que Santo André – Paranapiacaba ocorre no escalão verde, apresentando um grau de similaridade abaixo da média também no cartograma de São Caetano do Sul. Entretanto, no cartograma de Santo André – Paranapiacaba, São Caetano do Sul está entre os pontos com maior nível de similaridade, ou seja, em vermelho.

Com a análise de assimetria, notamos que os pontos São Caetano do Sul e Santo André – Paranapiacaba divergem, pois este tem integração máxima no conjunto dos dados, enquanto aquele, integração mínima.

O presente trabalho é um exemplo de que a Dialeto-metria nos permite, partindo dos dados de um atlas linguístico, por meio de uma análise quantitativa, fazer uma análise global dos fenômenos linguísticos de uma dada região. Com o atlas linguístico, temos uma paisagem da norma linguística utilizada pelos falantes no que se refere ao uso de cada item lexical que fez parte do rol de conceitos selecionados para a pesquisa. Uma visão de conjunto linguístico, entretanto, só é possível por meio de métodos matemáticos, em nosso caso, por meio da dialeto-metriação.

REFERÊNCIAS

ACADEMIA BALDERIX. Coeficiente de asimetria de Fisher. *ProbabilidadyEstadistica.net*, 31 maio 2022. Disponível em: <https://www.probabilidadyestadistica.net/coeficiente-de-asimetria-de-fisher/>. Acesso em: 10 jul. 2023.

ÁLVAREZ BLANCO, R.; DUBERT GARCÍA, F.; SOUSA FERNÁNDEZ, X. Aplicación da análise dialeto-métrica aos datos do Atlas Lingüístico Galego *In: ÁLVAREZ BLANCO, R.; GARCÍA, F. D.; FERNÁNDEZ, X. S. (ed.). Lingua e territorio*. Santiago de Compostela: Consello da Cultura Galega; Instituto da Lingua Galega, 2006. p. 461-493. Disponível em: https://www.researchgate.net/publication/257895687_Aplicacion_da_analise_dialectometrica_aos_datos_do_Atlas_Linguistico_Galego/link/5cc9ddbc92851c8d2213e99a/download. Acesso em: 10 out. 2020.

AURREKOETXEA, G. *et al. DiaTech: A new tool for dialectology*. Literary and Linguistic Computing, Leioa - Bizkaia, 2013. Disponível em: <https://eudia.ehu.eus/diatech/index/>. Acesso em: 02 jan. 2020.

BARBOSA, M. A. Da microestrutura dos vocabulários técnico-científicos. *In: ENCONTRO NACIONAL DA ANPOLL*, 4., jun. 1989, Recife. *Anais [...]*. Recife: ANPOLL, 1989. p. 567-578.

BARBOSA, M. A. Lexicologia: aspectos estruturais e semântico-sintáticos *In: PAIS, C. T. et al. Manual de Linguística*. São Paulo: Vozes, 1978. p. 81-125.

BARBOSA, M. A. *Léxico, produção e criatividade: processos do neologismo*. São Paulo: Global, 1981.

BRISSOS, F. O léxico são números ou são palavras? Os dialetos e a Dialeto-metria ajudam-nos a perceber. *In: I CIPAL – Ciclo de Palestras em Dialeto-logia, Geolinguística e Sociogeolinguística: teoria e prática*. São Paulo: USP; Uberlândia; UFU, 15 set. 2020. 1 vídeo (2h 09min 47seg).

Publicado pelo canal Adriana Cristianini. Disponível em:

<https://www.youtube.com/watch?v=lol4bQJg26Q>. Acesso em: 30 set. 2020.

BRISSOS, F. Portugal: a cidade e o interior. I – Centro-sul. *Limite: Revista de Estudos Portugueses y la Lusofonia*. Variação Dialetoal e História da Língua Portuguesa. Coordenação de Paulo Osório, v. 10, n. 1, p. 85-106, 2016. Disponível em:

<https://revista-limite.unex.es/index.php/limite/article/view/1556/1519>. Acesso em: 03 fev. 2019.

BRISSOS, F.; GILLIER, R.; SARAMAGO, J. O problema da subdivisão dialetoal madeirense: estudo dialetométrico da variação lexical. In: *XXXI Encontro Nacional da Associação Portuguesa de Linguística*. Porto: Faculdade de Letras da Universidade do Porto Associação Portuguesa de Linguística, 2016. p. 31-47. Disponível em:

<https://repositorio.ul.pt/bitstream/10451/32733/1/Brissos%20Gillier%20e%20Saramago%202016.pdf>. Disponível em: 10 nov. 2018.

BRISSOS, F.; GILLIER, R.; SARAMAGO, J. Variação lexical açoriana: estudo dialetométrico do Atlas Linguístico – Etnográfico dos Açores. In: GRANJA, M. A.; AGRELO, A. B.; SEOANE, E. G. (ed.). Aproximacións á variación lexical no domínio galego-portugués. *Revista Galega de Filoxía*: Monografía. v. 11, p. 11-17, 2017. Disponível em:

<https://repositorio.ul.pt/bitstream/10451/32734/1/Brissos%20Gillier%20e%20Saramago%202017.PDF>. Acesso em: 20 jan. 2019.

BRISSOS, F.; SARAMAGO, J. Análise dialetométrica do Atlas Linguístico-Etnográfico da Região Sul do Brasil: variação lexical. In: CARRILHO, E.; MARTINS, A. M.; PEREIRA, S.; SILVESTRE, J. P. (org.). *Estudos linguísticos e filológicos oferecidos a Ivo Castro*. Lisboa: Centro de Linguística da Universidade de Lisboa, 2019. p. 349-379. Disponível em:

<https://repositorio.ul.pt/handle/10451/39619>. Acesso em: 03 jan. 2020.

CASTILHO, J. C. *História do reservatório Billings e as bacias hidrográficas do ABC*. [s.l.: s.n.], 1997. Disponível em:

<https://acervo.socioambiental.org/sites/default/files/documents/24D00015.pdf>. Acesso em: 13 dez. 2023.

CONSÓRCIO INTERMUNICIPAL GRANDE ABC. *São Caetano do Sul*. São Caetano do Sul, 2020. Disponível em:

<https://www.consorcioabc.sp.gov.br/pagina/82/municipios-consorciados/sub-pagina/8/>. Acesso em: 02 dez. 2020.

COSERIU, E. La Geografía Lingüística In: COSERIU, E. *El hombre y su lenguaje*: Estudios de teoría y metodología lingüística. Madrid: Biblioteca Románica Hispánica; Editorial Gredos, [1954] 1991. p. 103-158.

COSERIU, E. Sistema, Norma y Habla In: COSERIU, E. *Teoría del Leguaje y Lingüística General*. 3. ed. Madrid: Biblioteca Românica Hispânica; Editorial Gredos, 1973. p. 11-113.

CRISTIANINI, A. C. *Atlas semântico-lexical da região do Grande ABC*. 2007. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo, 2007. Disponível em:

<https://www.teses.usp.br/teses/disponiveis/8/8139/tde-28012008-115533/pt-br.php>. Acesso em 04 nov. 2020.

SILVA, E. C. da. O Grande ABC Paulista: É Possível Pensar em Coesão Regional? *RDE - Revista de Desenvolvimento Econômico*, v. 1, n. 38, p. 173-187, 2013. Disponível em:

<https://revistas.unifacs.br/index.php/rde/article/view/2258/2138>. Acesso em 04 nov. 2020.

GARCÍA MOUTON, P. Dialectometría In: BLECUA, J. M.; CLAVERÍA, G.; SÁNCHEZ, C.; TORRUELLA, J. (ed.). *Filología e informática: Nuevas tecnologías em los estúdios filológicos*. Barcelona: Universidad Autónoma de Barcelona; Departamento de Filología Española, 1999. p. 3-24. Disponível em: http://adim.cchs.csic.es/sites/default/files/publicaciones/garcia_mouton_pilar_1999a.pdf. Acesso em 04 nov. 2020.

GOEBL, H. Eléments d'analyses dialectométrique (avec application à l'AIS). *Revue de linguistique romane* 45, p. 349-420, 1981.

GOEBL, H. *Dialektometrische Studien*. Anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF. Tübingen: Niemeyer, 1984. 3 v.

GOEBL, H. Introdução a los problemas y métodos según los principios de la Escuela Dialectométrica de Salzburgo (con ejemplos sacados del “Atlante Italo-Svizzero”, AIS) In: AURREKOETXEA, G.; ORMAETXEA, J. (ed.), *Tools for linguistic variation*. Leioa: Universidad del País Vasco, 2010. p. 3-39. Disponível em: <https://ojs.ehu.eus/index.php/ASJU/article/view/9887>. Acesso em: 20 nov. 2020.

GOEBL, H.; SMEČKA, P. L'analyse dialectométrique des cartes de la série B de l'ALF *Revue de Linguistique Romane*, v. 78, n. 311-312, p. 439-497, 2014.

IBGE. Instituto Brasileiro de Geografia e Estatística. Brasil 500 anos. *Território brasileiro e povoamento*. Rio de Janeiro, 2000, atualizado em 2007. Disponível em: <https://brasil500anos.ibge.gov.br/territorio-brasileiro-e-povoamento.html>. Acesso em: 03 jan. 2020.

IBGE. Instituto Brasileiro de Geografia e Estatística. Censo 2010. *Sobre*. Rio de Janeiro, 2010. Disponível em: <https://censo2010.ibge.gov.br/sobre-censo.html>. Acesso em: 03 jan. 2020.

MULLER, C. *Initiation à la statistique linguistique*. Paris: Librairie Larousse, 1968.

NASCIMENTO, L. D. do (org.). *Atlas socioambiental de São Bernardo do Campo*. São Bernardo do Campo: Prefeitura Municipal; Secretaria de Meio Ambiente e Proteção Animal, 2021. Disponível em: <https://www.saobernardo.sp.gov.br/web/sma/atlas>. Acesso em: 10 out. 2021.

POTTIER, B. *Linguística Geral: teoria e tradução*. Tradução de Walmiro Macedo. Rio de Janeiro: Presença; Universidade Santa Úrsula, 1978.

SANTO André. Vila de Paranapiacaba. Santo André, 29 abr. 2013. In: *Prefeitura de Santo André*. Disponível em: <https://www2.santoandre.sp.gov.br/index.php/paranapiacaba>. Acesso em: 02 dez. 2020.

SARAMAGO, J. O léxico são números ou são palavras? Os dialetos e a Dialectometria ajudam-nos a perceber In: I CIPAL – *Ciclo de Palestras em Dialectologia, Geolinguística e Sociogeolinguística: teoria e prática*. [S. l.: s.n], 2020. 1 vídeo (2h 09min 47seg). Publicado pelo canal Adriana Cristianini. Disponível em: <https://www.youtube.com/watch?v=lol4bQJg26Q>. Acesso em: 30 set. 2020.

SÃO PAULO. *Serra do Mar e Mosaicos da Mata Atlântica: Uma Experiência de Recuperação Socioambiental*. Coordenação editorial de Keila Prado Costa. São Paulo: KPMO Cultura e Arte, 2014.

UNESCO. *Reservas da Biosfera da UNESCO*. Lisboa: Unesco, 2020. Disponível em: unescoportugal.mne.gov.pt/pt/redes-unesco/reservas-da-biosfera-da-unesco/. Acesso em: 02 dez. 2020.