

Dialectometrical study of the Semantic and Lexical of the Great ABC Region (ASL-ABC)

Adriana Cristina **CRISTIANINI***

*PhD in Linguistics from the University of São Paulo (2007). Associate Professor III_UFU.
adriana.cristianini@gmail.com.

Abstract:

Through dialectological, geolinguistic, sociolinguistic studies, among others, researchers direct their attention to different aspects of the language and types of variation. These studies provide us with the opportunity to understand more accurately, describe and measure the performance of linguistic phenomena related to social, cultural, historical, political, regional and other factors. However, the volume of data processed in such studies is gigantic and has become increasingly larger. Often, only with the help of technological resources can we carry out research that, in the not-too-distant past, would have been impossible. Given this brief reflection, we started from the records of the Semantic-lexical Atlas of the Great ABC Region and, following the theoretical-methodological principles of the Salzburg Dialectometric School, we aimed to present in this article a dialectometric study, with the help of the DiaTech software. This tool allows you to approach a large number of data from a linguistic atlas, to relate two or more points of research, helping to create an overview and patterns of diatopic variation. According to Saramago (2020), Dialectometry allows greater use of existing material in linguistic atlases and enables the discovery of relationships that would be hidden with the application of the qualitative or traditional method in data analysis.

Key words:

Brazilian dialectology. Dialectometry. Great ABC Region, Brazil.

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

Received on: 27/09/23

Accepted on: 14/02/24

Dialectometrical study of the Semantic and Lexical of the Great ABC Region (ASL_ABC)

Adriana Cristina Cristianini

INTRODUCTION

The social and historical constitution of Brazil is full of plurality. The Portuguese language, spoken in this country and which is its main cultural element, therefore presents a lot of diversity and variation that has been studied in several ways.

Through studies of Dialectology, Geolinguistics, Sociolinguistics, Sociogeolinguistics and Dialectometry, among others, it becomes possible to know with greater accuracy, describe and measure the performance of linguistic phenomena related to social, cultural, historical, political, regional and others. Researchers thus direct their attention to different aspects of the language and types of variation.

With regard to diatopic variation, the object of this study, we start from work that followed the precepts of Geolinguistics, considered a section of Dialectology, which cartographically records the answers obtained through interviews applied to a certain group of subjects from a given location, seeking to specify the topographic repetition of the observed phenomena.

In this study, we intend to analyze data on the diatopic variation portrayed in the semantic-lexical Atlas of the Great ABC region – ASL_ABC (Cristianini, 2007), through Dialectometry. We will seek to direct our attention mainly to two points in the network created to create this atlas, more specifically to the points that correspond to São Caetano do Sul and the district of Santo André, Paranapiacaba.

To this end, in addition to this introduction and conclusion, we divide this work into three sections: Literature review; Methodology; and Results.

LITERATURE REVIEW

According to García Mouton (1999), despite the remarkable advances, over the last century, in the development of linguistic theories and studies of variation, we know that atlases, with their enormous richness, generally remain in the hands almost exclusively of dialectologists. and geolinguists, who, very often, are the same ones who collected and mapped the data.

Geolinguistics develops studies of linguistic variation in space and records the data collected through interviews in cartograms, in a network of points in a given region. The compilation of these cartograms¹ results in linguistic atlases. According to Álvarez Blanco, Dubert García and Sousa Fernández (2006, p. 461), Geolinguistics

[...] is dedicated to the study of linguistic variation in space and is mainly based on data contained in linguistic atlases, which are a very rich source of information for in-depth knowledge of phonetic, morphological, syntactic and lexical features of a language set of places in a territory. In addition to

¹ Although, traditionally, we find in the terminology related to Geolinguistics and Dialectometry the use of the terms map or linguistic chart, since 2007, we have been using the term cartogram, considering that this is a type of representation that is more concerned with information that will be objects of spatial distribution within the map, with exact and precise limits or geographic coordinates (SANCHEZ, 1973 apud CRISTIANINI, 2007).

the specific data for each of the points researched, the atlases allow us to analyze the regularity with which the different linguistic units are distributed in space, information that dialectologists consider crucial, since the discovery of regularities allows us to recognize areas of the territory that share linguistic features (our translation).

To identify these locations that share the same linguistic features, geolinguists delimit isoglosses on the cartograms, which are usually high in number, and superimpose them. It turns out that, when analyzing this overlap, we realize that it is rarely possible to accurately determine areas that present linguistically different characteristics, as isoglosses tend not to coincide closely with each other, especially if a substantial number of phenomena are used. linguistic aspects in cartographic representation. Given this, it has become important to search for a way to synthesize atlas data so that it is possible to identify the points where the greatest coincidences and greatest differences occur in relation to other linguistic points.

Dialectometry arises, then, from the need, based on the rather large volume of data collected through geolinguistic research and mapped to present the phenomena addressed, to synthesize into a single cartogram the results of multiple cartograms that make up a linguistic atlas. Therefore, it arises from the need to record and analyze all the records contained in an atlas, allowing an overview of the variation contained in a given locality, region or, in the broadest sense, dialect or linguistic system.

A pioneer in dialectometric studies, Jean Séguy was the one who first coined the name Dialectometry and who first gave the discipline a robust theoretical body, based on significant practical results, with the dialectometricization of the Atlas Linguistique de la Gascogne – ALG, in 1973. From Séguy, Guiter and members of the Toulouse school – Philips, Fossat – developed the method and applied it in the Romanesque context (García Mouton, 1999). According to Saramago (2020), initially, this method “[...] consisted of tracing a triangular network, the result of the connection between the geographically closest points. In these lines, the values found in the entire atlas for the studied phenomenon were placed”.

At the end of the 1970s, Hans Goebel developed a dialectometric study with data contained in the Linguistic-Ethnographic Atlas of Italy and Southern Switzerland – AIS and the Atlas Linguistique de la France – ALF. We can say that, based on Goebel's studies, Dialectometry is consolidated and is now considered an important quantitative study, based on the recording of dialectal data provided by Geolinguistics. It is taking into account this journey of around half a century that Brissos and Saramago (2019, p. 353) define Dialectometry as

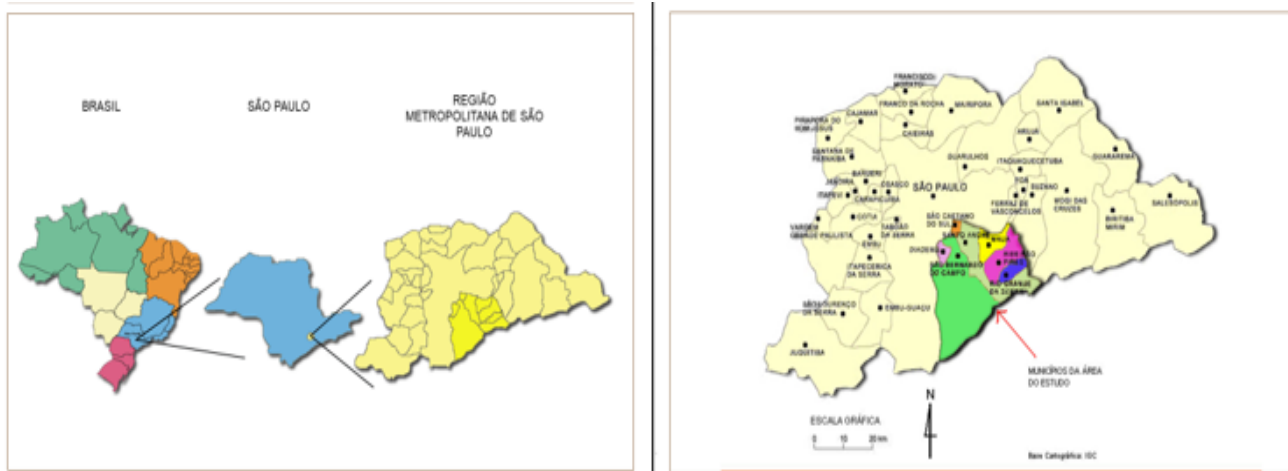
[...] a quantitative approach to the study of dialects with a focus on metrics, i.e., on measuring the phenomena of dialectal variation through exact and fully comparable procedures, which matter from numerical or taxonomic classification. It applies elaborate mathematical-statistical calculations to the data matrix obtained from the aforementioned procedures and cartographically represents (spatializes) the results of these calculations, with the linguist, with the freedom that statistics confers, having the final task of interpreting the geolinguistic framework he has at hand front.

Dialectometry, therefore, makes it possible to approach a large number of data, taking advantage of records from a linguistic atlas, to relate two or more points of research, helping to create an overview and patterns of diatopic variation. According to Saramago (2020), Dialectometry allows greater use of existing material in linguistic atlases and enables the discovery of relationships that would be hidden with the application of the qualitative or traditional method in data analysis.

The Great ABC region of São Paulo is made up of seven municipalities – Santo André, São Bernardo do Campo, São Caetano do Sul, Diadema, Mauá, Ribeirão Pires and Rio Grande da Serra – and has considerable importance in the socio-historical-political context. economic-cultural of Brazil.

Figure 1, below, constitutes cartograms of the region’s location.

Figure 1 – Location of the Great ABC Region



Source: Prepared by the author.

The urgency of rescuing the semantic-lexical variation of the language spoken in Great ABC, as a way of recording the linguistic memory of the community in this locality, was already perceived, even before the creation of the atlas, since, among other characteristics of this region, there are great mobility of the population.

The history of the founding and development of the municipalities of Great ABC shows us that original peoples, migration and immigration can be responsible for the great linguistic variation and change in the speech of the region's inhabitants. It is indisputable that linguistic uses in the Great ABC region are the result of influences from different peoples, in addition to the original ones, who already inhabited this land before the Portuguese came to Brazil.

When disembarking in São Vicente, the future Portuguese colonizers were faced with major obstacles when coming from the coast to reach the plateau. Two major challenges stand out: the unknown, steep and dangerous paths to cross the Serra do Mar; and the existence of great rivalry between the original peoples. On the plateau, we can highlight the Guaranis (Carijós) groups and the Jês groups, such as the Guaianás, Maromomis and Kaingangues. In addition to these, there are records of bloody wars, in the 16th and 17th centuries, between the Tupiniquins and the Tupinambás, bitter enemies. It is in this context that the name of João Ramalho stands out, who became leader of the Tupiniquin tribes in the São Paulo plateau region when he married Bartira (Tree Flower”, in Tupi), daughter of the chief Tibiriça (“Watchman of the land” , in Tupi). It was through João Ramalho that the Portuguese colonizers noticed that, above the coast and after the mountains, there were more populated, rich and fertile lands. With this brief preamble, we can delve into the history of the Great ABC region.

In fact, to tell the origin of the region that we now call Great ABC, history needs to be divided into two periods. The first began in 1553, when the Vila de Santo André da Borda do Campo was inaugurated, with the authorization of the then Governor General of Brazil, Tomé de Souza. A few years later, however, there was, in 1560, the extinction of Vila de Santo André, which ceased to exist as an administrative unit and became a neighborhood of São Paulo de Piratininga,

which had been created in 1554 (due to the Jesuits' interest in transferring their school, located in São Vicente, for the Piratininga fields region, given the large evasion of people from the coast to the interior). The lands of Vila de Santo André, in 1561, were granted as sesmaria to Amador de Medeiros, ombudsman of the Captaincy of São Vicente. This region, after that, went through a period of stagnation, with a restricted subsistence and pasture culture developing for the animals of drovers who came from the coast towards São Paulo and the interior. After 70 years, in 1631 and 1637, most of the territory was donated to the Order of São Bento and the São Caetano and São Bernardo farms were formed, respectively. These farms remained under the ownership of the Benedictines until 1870 (shortly after the beginning of the region's second period of development), when they were purchased by the State with a view to creating immigrant colonies).

The second period, which is a relevant milestone in the history of the Great ABC region, began in 1860, with the construction of the São Paulo Railway, connecting Santos to Jundiaí. There was then a major recovery and rapid development of the place, which saw its profile change considerably. The project, installed near the Tamanduateí River, was intended to improve the flow of agricultural products to the Port of Santos, especially coffee, whose large-scale production had already begun in the Province of São Paulo. The transport facilities brought to the region attracted industries (generally linked to chemical, textile and furniture production) which, relying mainly on immigrants as labor and with tax incentives, initially settled in areas close to the railway and the river. In this way, urban centers were formed that became villages and, from the 1940s onwards, they became independent and, thus, forming the municipalities that today make up the region.

From the 1950s onwards, transformations began that were felt due to the change in the types of industries that came to the region. Foreign and state investments boosted the growth of industries in the metallurgical, mechanical, automobile and electrical material sectors. As a result, the profile of a large part of the workforce also changed, attracting workers from different regions of the country. In the decades that followed, these industries, going through great expansion and subsequent decline, were decisive in the historical-social delineation of the region to the present day.

We can state, in view of the above, that the presence of drovers, migrants and immigrants stood out for the sociocultural formation of the region.

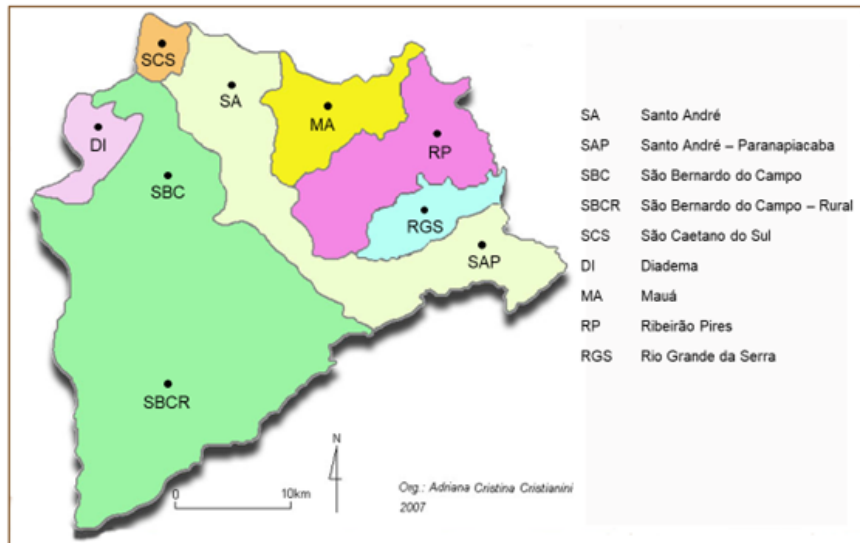
This brief explanation, even if superficial, is enough for us to conclude that the transfiguration that took place in relationships between people caused changes in the reality of the region's communities and their linguistic uses.

In order to study and record the linguistic landscape of the region, based on the postulates of contemporary Geolinguistics, a semantic-lexical questionnaire, with 202 questions, was applied in nine points in the region. To determine the network of points established for the preparation of the ASL_ABC, the following criteria were considered, according to Cristianini (2007): a) each municipality necessarily had at least one research point; b) the demographic density of the seven municipalities in the region, in which it was observed that two of the seven municipalities contained more than half of the region's population, namely, Santo André, which is home to 27.6% of the entire region's population, and São Bernardo do Campo, with 29.8%; the two municipalities together, therefore, are home to 57.4% of the entire population of Great ABC; these two municipalities were awarded one more point each; c) the municipalities of Santo André and São Bernardo do Campo, in addition to having greater demographic density, are also much larger in geographical area, which also justified the determination of two points in each of these municipalities; d) São Bernardo do Campo presents a natural geographic division that highlighted

the division of the municipality between urban and rural areas², which was considered relevant to establish the need for a point in the urban area and another in the rural area; e) Santo André contains a region considered a historical area, Paranapiacaba, which is somewhat isolated from the urban area of the municipality and this was decisive for the establishment of the two points in this municipality.

The delimitation of the ASL_ABC point network, therefore, is: Santo André; Santo André – Paranapiacaba; São Bernardo do Campo; São Bernardo do Campo – Rural area; São Caetano do Sul; Diadem; Mauá; Ribeirão Pires; and Rio Grande da Serra. Figure 2 shows the nine research points in the Great ABC region.

Figure 2– Research points



Source: Prepared by the author.

At each point, four subjects were interviewed, totaling 36 subjects subdivided into two age groups – 18 to 30 and 50 to 65 years of age – and two genders – female and male.

The determination of the subjects' profile mainly took into account the geographical issue. Therefore, only subjects born locally or living in the region for at least two-thirds of their lives participated in the research. The education variable was also observed: an education level of up to the eighth grade of elementary school was determined for the research, considered low education.

With the answers obtained in the interviews, a semantic-lexical database was created and, later, this data was treated and documented cartographically, so that it was possible to create a picture of the semantic-lexical diversity of speech in the Great ABC region.

In addition to the theoretical-methodological framework of Geolinguistics and Lexicology, based on works by Pottier (1978) and Barbosa (1978, 1981, 1989), the Norma approach carried out by Coseriu (1954, 1973) was used, and complemented it. with notions of Lexical Statistics, proposed by Muller (1968), especially regarding frequency.

The atlas, made up of a set of 202 linguistic cartograms, recorded, albeit partially, the linguistic heritage that consists of the semantic-lexical norm of the Great ABC region of São Paulo.

METHODOLOGY

² According to the 2000 Census from the Brazilian Institute of Geography and Statistics – IBGE. In this work, we maintained the demographic data used in ASL_ABC by Cristianini (2007).

This dialectometric study follows the theoretical-methodological principles of the Salzburg Dialectometric School, respecting the procedures, interpretation and analysis of data recommended by dialectometric studies of dialectal areas of the Portuguese language based on linguistic atlases at the University of Lisbon, such as: Brissos (2016), Brissos, Gillier and Saramago (2016, 2017) and Brissos and Saramago (2019).

To carry out this work, we revisited ASL_ABC to analyze and delimit the corpus used. To this end, we analyzed more than six thousand lexical items, resulting from the responses of 36 interviewed subjects, belonging to the seven municipalities that covered the studied region, with nine research points. In addition, we listened to around 70 hours of audio recordings relating to the interviews. We also studied 202 cartograms and graphs that demonstrate the semantic-lexical variation in the region, as well as considerations related to the genders and age groups of the subjects interviewed.

Next, we transcribed all the data and organized it into spreadsheets and counted the lexical items to compose the database for subsequent dialectometric analysis. In addition to preparing the spreadsheet with all the linguistic data from the atlas, we also created spreadsheets for responses from subjects of each age group and each gender.

For the present work, we chose to make a selection, relying on parameters already used in previous dialectometric studies, and use the responses of subjects from the second age group³ and, to select the cartograms of the concepts that would make up the research, we followed these criteria:

- there are responses from all subjects in the second age group, at all points in the survey;
- not be a cartogram with mononymous answers, that is, not present the same answer at all points and by all subjects.

Respecting these criteria, we used linguistic phenomena corresponding to 74 concepts distributed in 14 semantic areas⁴, constituting a corpus with 1332 lexical items for dialectometric analysis.

The dialectometricization of data and mapping of results were carried out using the DiaTech software (Aurrekoetxea *et al.*, 2013) and, for this, we used two parameters for analyzing the results: Cluster analysis⁵ and Synoptic map.

We started with hierarchical-agglomerative organization methods, that is, cluster analysis, in which all research points begin the analysis process separately and are grouped one by one, until a single cluster (or grouping) is obtained. containing all of the above. In the software, for cluster analysis (or grouping analysis), according to the semantic-lexical linguistic aspect of our work, we

³ It is worth clarifying that the option to use responses from subjects in the second age group is due to the fact that we found that older subjects generated more diverse and productive responses for the elaboration of the ASL_ABC.

⁴ Concepts considered for the dialectometric analysis, distributed in 14 semantic areas: a) Geographic features – stream, pinguela, sea wave; b) Atmospheric phenomena – storm, hail, drizzle, dew, fog; c) Stars and time – sunrise, sunset; d) Agropastoral activities – chamomile, double banana, heart, cob, cob, bean pod, wheelbarrow, wheelbarrow stem, jacá, chopped; e) Fauna – vulture, vulture, parrot, neck mane, loin, udder, tail, blowfly, fruit bug; f) Human body – toothless, booger, neck, armpit, left-hander, breasts, uterus; g) Life cycles – midwife, giving birth, stepmother, death; h) Socializing and social behavior – stingy person, bad payer, drunk, straw cigarette, cigarette butt; i) Religion and beliefs – devil, ghost, spell, faith healer; j) Children's games and entertainment – marbles, kite flying, hide and seek, swing; k) Housing – frame, toilet, lantern, light switch; l) Food and cuisine – breakfast, curau, brandy, gorged, gluttonous, French bread, Bengal bread; m) Clothing and accessories – bra, underwear, panties, clip (with pressure), diadem; n) Urban life – traffic lights, speed bumps, sidewalks, urban buses, bodega.

⁵ Cluster analysis is the name given to a set of techniques used to identify patterns in databases through the formation of tentatively homogeneous groups. According to Brissos, Gillier and Saramago (2017, p. 17), “The dialectometric technique par excellence that allows us to identify or decompose a given area into dialectal groups is cluster or dendrographic analysis, which segments, in a tree diagram, the dialectal nuclei at an increasing level of homogeneity from the trunk of the tree” (emphasis added).

select: Similarity index – IRI⁶; and Statistical analysis algorithm – “Ward”⁷. Next, we move on to the analysis with the Synoptic Map parameter. In this parameter, we employ two types of statistical operations: the similarity distribution and the asymmetry (or skewness) distribution.

With similarity distribution analysis, we can check the similarity/difference hierarchy of the specific point by relating it to each of the remaining points.

According to Brissos (2020), when analyzing the degree of similarity in relation to a given point, it is necessary to observe that values above the average in the generated histogram are represented with warm colors and to the right. Points with similarity values below the average, in turn, are represented with cold colors and to the left in the histogram. This means that the warmer the color, the higher the similarity of the analyzed point (which is always represented with white) and the colder it is, the lower the similarity. This means that, in the histogram, the red color indicates a maximum level of similarity, while the blue color indicates a minimum level.

To analyze the similarity distribution, we ran the DiaTech software with the following specifications: Statistical algorithm – MinMwMax⁸; Distance unit – IRI; and Location – one different point of the search at a time. With the results, we created similarity cartograms.

The asymmetry distribution analysis, in turn, summarizes the similarity charts of all points and allows us to observe which are the most and least integrated points in the set constituted by the linguistic corpus used in the study. According to Brissos, Gillier and Saramago (2017, p.19), “This parameter starts from the fact that not all locations have the same degree of similarity with the remaining locations in the analyzed set and presents, based on Fisher's asymmetry coefficient⁹ (CAF), a relational synthesis of the places most and least similar to the set”.

To generate the results of the asymmetry distribution in DiaTech, we selected the same specifications that were used for the similarity analysis. The location is no longer requested here, as the analysis is of the whole.

We now discuss the results provided by DiaTech.

RESULTS

We divided this section into two parts to cover the two parameters used in the analysis: cluster analysis and synoptic map.

We begin this analysis by discussing the cartograms created from the results obtained by the DiaTech software with the corpus of 74 linguistic phenomena.

We noticed that, when analyzing two groups, we observed that they are constituted as follows: 1) Rio Grande da Serra, Diadema, Ribeirão Pires and São Bernardo do Campo – Rural area

⁶ Statistical calculations occur based on the differences found, observing the Relative Difference Indices (IRD) or the Weighted Difference Indices (IPD); and, based on similarities, through Relative Identity Indices (IRI) or Weighted Identity Indices (IPI). The IRI, the main dialectometric index of the Salzburg Dialectometric School (GOEBL, 2010), can be considered as a standard index for measuring linguistic or, in our study, lexical similarity. The IRI constitutes a quotient that, ultimately, measures the percentage of linguistic identities between two points in the atlas. The DiaTech software establishes a dialectal relationship between a reference point and other points, demonstrating linguistic approximation or distance, considering the IRI. For a more detailed description of IRI, it is worth consulting Goebel (1981, p. 357-360 and 1984, p. 74-77).

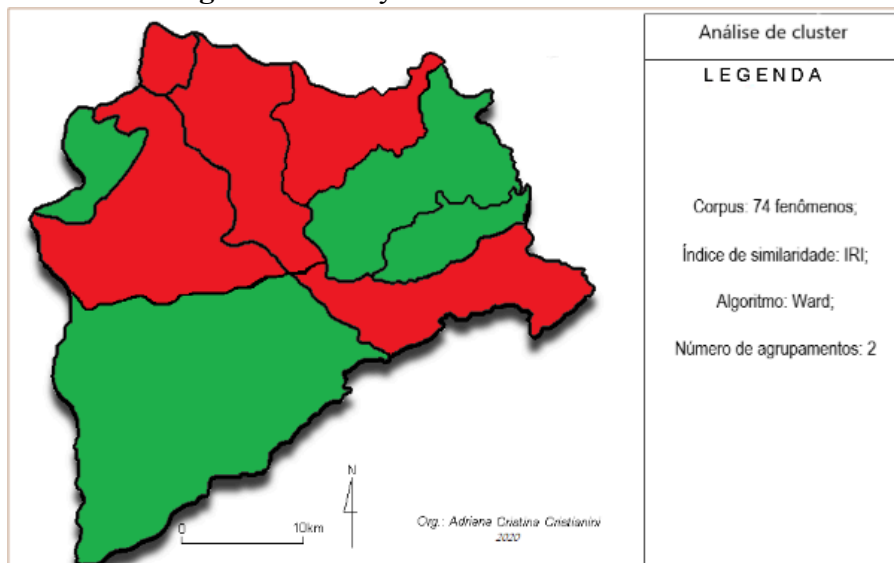
⁷ Ward – Statistical method for forming groups in which homogeneity within groups is maximized. More information can be found in Goebel and Smečka (2014, p. 460).

⁸ MinMwMax is an algorithm, generally used by the Salzburg School of Dialectometry, that creates symmetric intervals (n intervals, the number of which must fit the data set and the number of points present in the study) on both sides of the arithmetic mean (Mw), downwards (minimum value – Min) and upwards (maximum value – Max). With MinMwMax the interval between Mw and Min and the interval between Max and Mw are divided by n/2 in order to obtain the numerical limits of the n intervals. For more information on this, see BRISSOS, GILLIER and SARAMAGO (2017).

⁹ Fisher's asymmetry coefficient is one of the coefficients that, in statistics, is used to calculate the asymmetry of a distribution, allowing us to know whether a probability distribution is positively skewed, negatively skewed or symmetric. For more information, see Balderix Academy (2023).

(green); and 2) São Bernardo do Campo, Mauá, Santo André, Santo André – Paranapiacaba and São Caetano do Sul (red). Figure 3 presents the cluster analysis cartogram with two groupings¹⁰:

Figure 3 – Analysis with two clusters

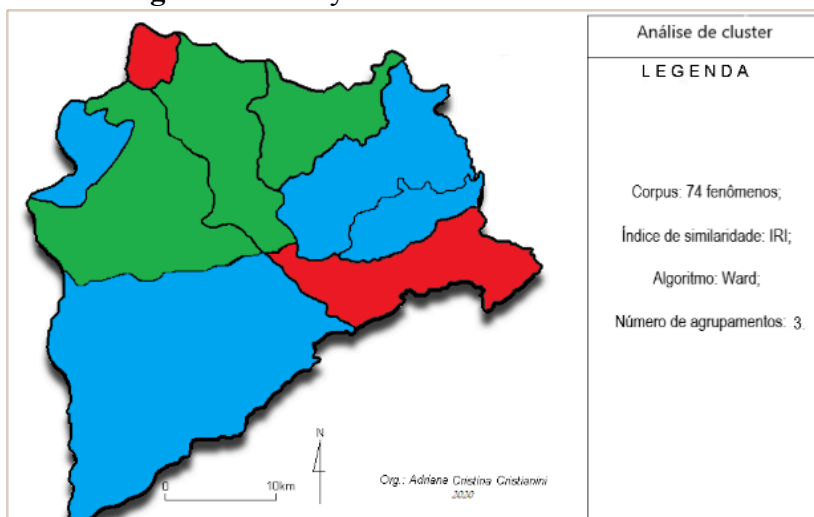


Source: Prepared by the author.

The analysis with three groups, or more, highlights additional information that already allows us to make some inferences, as it makes it possible to associate the results with socio-historical-cultural and geographic aspects of the region.

In Figure 4, below, we observe the groupings of locations as follows: 1) Rio Grande da Serra, Diadema, Ribeirão Pires and São Bernardo do Campo – Rural area (blue); 2) São Bernardo do Campo, Mauá and Santo André (green); and (iii) Santo André – Paranapiacaba and São Caetano do Sul (red).

Figure 4 – Analysis with three clusters



Source: Prepared by the author.

¹⁰ For this work, we created the cartograms with the results in a representation of the municipalities, although the maps generated by DiaTech present the Voronoi polygonation (polygonal mosaic), considering that this assumes an identical spatial dimension for each point, which, in a network reduction of points, as occurs in this study, is especially fallible. For more information, see Saramago (2020).

Historically, Santo André – Paranapiacaba and São Caetano do Sul are two points that received considerable influence from immigrant cultures. São Caetano do Sul, throughout its history, was a farm destined to receive European immigrants, more specifically, it became a colony for Italian immigrants mainly from Treviso. These lands, at the time, were practically abandoned and, then, the government made it an object of agreement with the Italians who, initially, received lots, housing and food as long as they produced in the region. This fact is an important factor for the identity of the municipality. Geographically, São Caetano do Sul, an exclusively urban municipality, does not have conditions for expansion and, therefore, compared to other municipalities in the region, it hosted a smaller number of emigrants among those who came to the region during the installation and expansion of industries (Consórcio..., 2020).

The history of Paranapiacaba, in turn, is also marked by the presence of European immigrants, generally English and, later, Portuguese. Paranapiacaba emerged in the 1860s, as a camp for workers (many of them of English origin) from the company The São Paulo Railway Company Ltd. – SPR, responsible for the construction of the Serra do Mar section of the railway in the region (Funicular System). Even with the inauguration of the railway, workers were kept on site to operate the services and maintain the works.

In the 1940s, the place underwent some changes, as the railway was incorporated into the Union Heritage Site and the presence of the English in the region decreased. Transformations occurred with the marked presence of Portuguese culture. During the time of the English, according to Santo André (2020), the village of Paranapiacaba maintained British characteristics, with a romantic air,

[...] with wooden houses, backyards separated by hedges and calm streets, lined with pine trees, in contrast to the Upper Part, which received an urban occupation marked by Portuguese heritage, with narrow streets and houses with small fronts built next to to alignment. Connecting the Upper Part to the Lower Part there is a metal bridge intended exclusively for pedestrians and bicycles, which remains today after some renovations (Santo André, 2020).

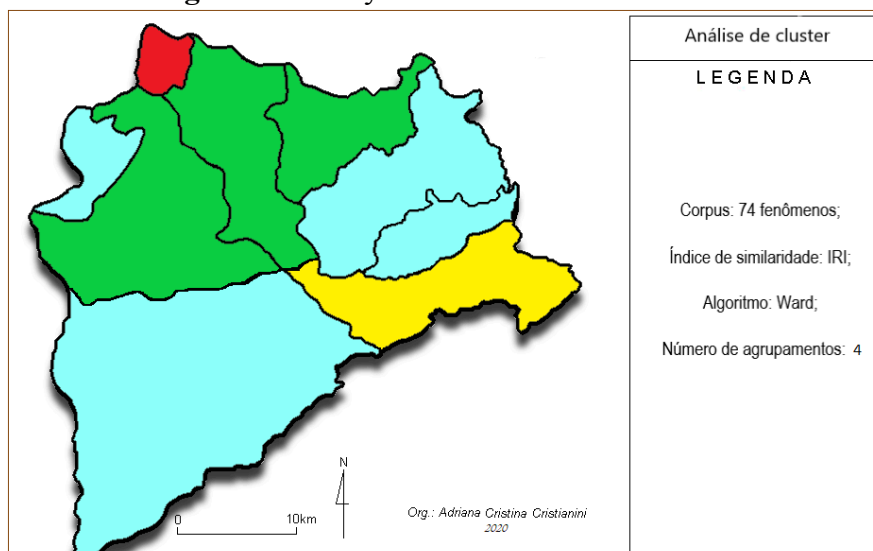
In the 1980s, with the end of the Funicular System's operation, an era ended and a struggle for the historical and environmental preservation of the site began. A movement for Paranapiacaba to become a tourist hub was promoted and, later, its historical listing was published, legalizing the place as being of public interest. This listing covers the urban area, railway equipment and the natural area around it.

The village was incorporated into the municipality of Santo André in 2001 and, in 2003, the Nascentes de Paranapiacaba Municipal Natural Park was created, a green area of Atlantic Forest surrounding the village. Furthermore, also in 2003, Paranapiacaba became one of the centers of the Biosphere Reserve program of the United Nations Educational, Scientific and Cultural Organization – UNESCO¹¹.

It is worth noting, however, that, when analyzing the data with four groupings, in Figure 5, São Caetano do Sul and Santo André – Paranapiacaba do not remain in the grouping. We can observe that two clusters stand out with the highest number of points and two other clusters constitute unitary groups, namely: 1) Santo André, São Bernardo do Campo and Mauá (green); 2) São Bernardo do Campo – Rural area, Diadema, Ribeirão Pires and Rio Grande da Serra (blue); 3) São Caetano do Sul (red); and 4) Santo André – Paranapiacaba (yellow).

¹¹ According to UNESCO (2020), Biosphere Reserves seek to reconcile “[...] the conservation of biological and cultural diversity with economic and social development, through partnerships between people and nature”. It is worth noting that Paranapiacaba is surrounded by three important Conservation Units (protected areas): Parque Nascentes; the Alto da Serra de Paranapiacaba Biological Reserve; and Serra do Mar State Park.

Figure 5 – Analysis with four clusters

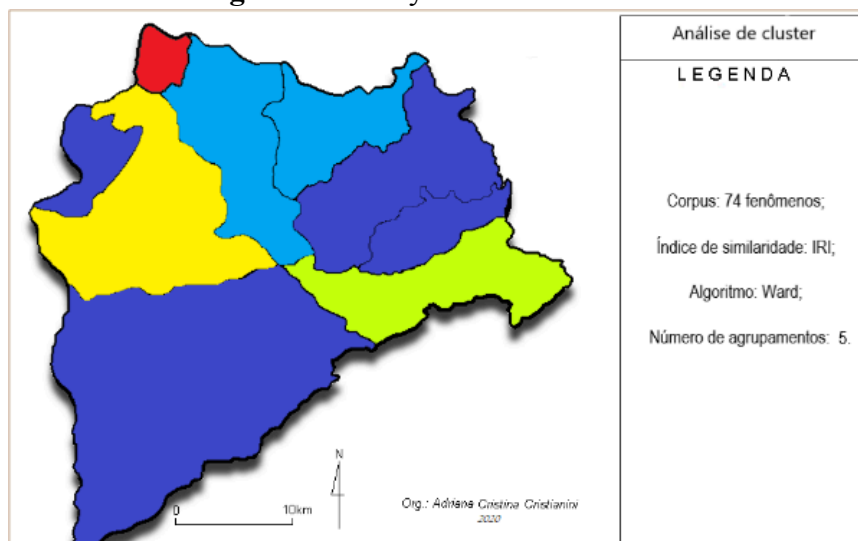


Source: Prepared by the author.

This observation allows us to infer that, although these two points, when there are fewer clusters, come closer together in the same group, when we analyze a larger number of clusters, they move further apart. The geographic distance between the two points may justify this result. It must also be considered that Santo André – Paranapiacaba is a point that remains considerably geographically isolated from the others as it is located close to the top of Serra do Mar.

In an analysis with five groupings, as shown in Figure 6, we have: 1) São Bernardo do Campo – Rural area, Diadema, Ribeirão Pires and Rio Grande da Serra (dark blue); 2) Santo André and Mauá (light blue); 3) São Bernardo do Campo (yellow); 4) São Caetano do Sul (red); and 5) Santo André – Paranapiacaba (green).

Figure 6 – Analysis with five clusters

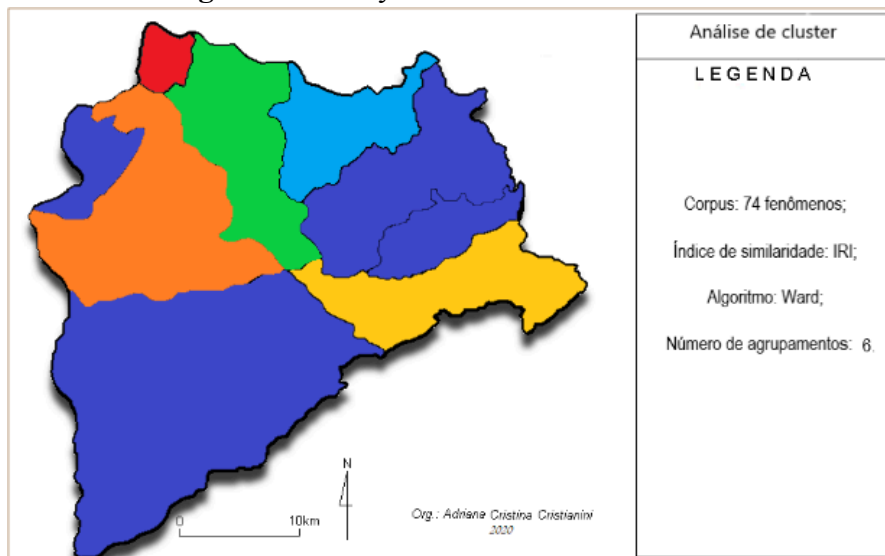


Source: Prepared by the author.

In Figure 7, we have the six groupings: 1) São Bernardo do Campo – Rural area, Diadema, Ribeirão Pires and Rio Grande da Serra; 2) Santo André; 3) Mauá; 4) São Bernardo do Campo; 5) São Caetano do Sul; and 6) Santo André – Paranapiacaba. We can observe, here, that São Bernardo

do Campo – Rural area, Diadema, Ribeirão Pires and Rio Grande da Serra (dark blue) remain grouped, while the points Santo André (green), Mauá (light blue), São Bernardo do Campo (orange), São Caetano do Sul (red) and Santo André – Paranapiacaba (yellow) constitute unitary groups.

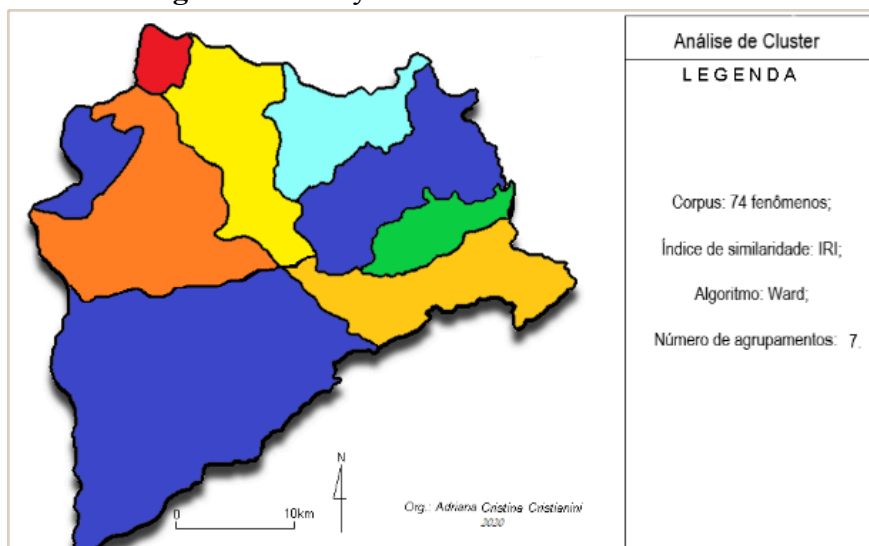
Figure 7 – Analysis with six clusters



Source: Prepared by the author.

The trend continues when we observe a section with seven clusters, in Figure 8. To the unitary groups, we add the point Rio Grande da Serra (green) which, with fewer clusters, was part of the larger group. We thus have: 1) São Bernardo do Campo – Rural area, Diadema, Ribeirão Pires (dark blue); 2) Santo André (yellow); 3) Mauá (light blue); 4) Rio Grande da Serra (green); 5) São Bernardo do Campo (orange); 6) São Caetano do Sul (red); and 7) Santo André – Paranapiacaba (ochre).

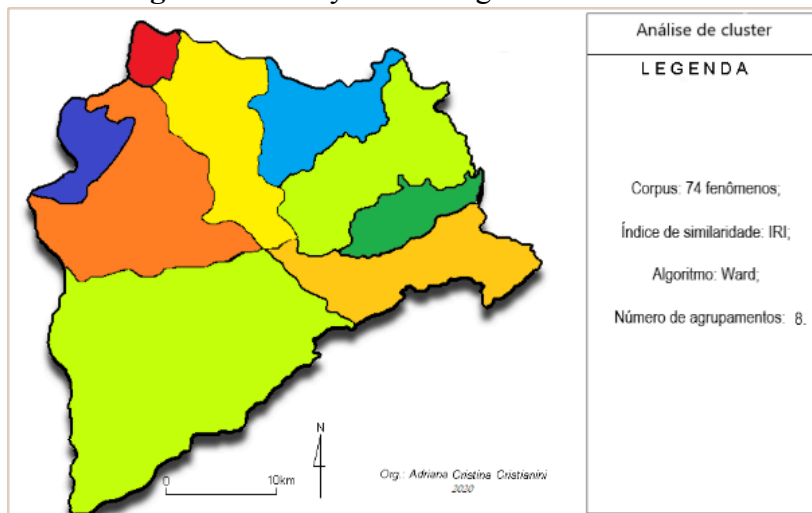
Figure 8 – Analysis with seven clusters



Source: Prepared by the author.

With the maximum possibility of groupings (eight) for the amount of points in our research, in Figure 9, two points remain grouped, São Bernardo do Campo – Rural zone and Ribeirão Pires and all the others constitute unitary groups, namely: 1) São Bernardo do Campo – Rural area and Ribeirão Pires (light green); 2) Santo André (yellow); 3) Mauá (light blue); 4) Rio Grande da Serra; 5) São Bernardo do Campo; 6) Diadem (dark blue); 7) São Caetano do Sul (red); and 8) Santo André – Paranapiacaba (ochre).

Figure 9 – Analysis with eight clusters



Source: Prepared by the author.

We can observe, through the different numbers of clusters, that there is a clear lack of geographic cohesion in the clusters, possibly justified by the geographical composition, the historical process by which the region was formed, and the socioeconomic reality of its current population. The greater ABC region, in fact, is not considered a cohesive region in many aspects.

Geographically, it is important to highlight Serra do Mar and the various paths that connect the plateau to the coast. The Portuguese quickly realized that there was a trail that followed the Mogi River. The route of this trail, in the 19th century, was used by English engineers to build the Santos-Jundiaí Railway. When the Jesuits arrived on the coast and decided to found a nucleus on the plateau, they first used the Mogi River trail, but the proximity to hostile original groups made them consider opening a new trail, away from that river. It was up to Father José de Anchieta to undertake the opening of the new trail that became known as “Caminho do Padre José”. This route, however, at the time was considered very bad, with records that it was treated as one of the “worst in the world”, as it only allowed traffic in single files and, clinging to the roots of trees, loads were carried on shoulders and the sick, when necessary, carried in hammocks. This trail, although it was improved in the second half of the 17th century, with the construction of small bridges, cuts and embankments, retaining walls and, on the plateau, a log road, was insufficient and impractical, especially during the rainy season. The Padre José Path is currently demarcated by researchers with gaps resulting from landslides that occur in the mountains. Although rudimentary in form, the Caminho do Padre José was practically the only connection between the coast and the plateau until 1792, when the Calçada do Lorena was built. This road, unlike the previous one, even by European standards, was considered a very advanced construction for the time and was a historical and commercial landmark, as it boosted national and international trade. In 1840, another route began to be built, the Estrada da Maioridade, which was later adapted for car traffic in the 20th century and is currently called Caminho do Mar. It was also built in the 19th century, a path to the plateau, after climbing the mountain, the Estrada do Vergueiro. In 1947, the Anchieta Highway was built, marking the beginning of a phase of accelerated transformation and great economic, social and

urban development. Another extremely important route, which connects Santos to São Paulo, crossing the municipalities of São Bernardo do Campo and Diadema, is the Imigrantes Highway. Roughly speaking, these are the main paths¹² along which, in addition to wealth, products and merchandise, families, enslaved men and women, migrants, immigrants, tourists, ideas and words circulated.

It is also worth noting that, in 1925, the Bilings¹³ dam was created in the Great ABC region, whose hydrographic basin extends through the municipalities of Ribeirão Pires, Diadema, Rio Grande da Serra, São Bernardo do Campo, Santo André and São Paulo. Geographically, the Bilings dam constitutes a division of the region into two parts, considering that five, of the seven municipalities in the region, received the largest body of water in South America. We infer that this geographic barrier has contributed to the lack of cohesion dialect to which we refer.

Historically, it is important to highlight the influences of immigration and migration, together with the presence of original peoples, especially the Tupi and Guarani groups, in our culture. The Brazilian Portuguese language is constituted by that brought by the Portuguese in the times of domination and exploration, with evident influence of the languages spoken by the original peoples and, also, of the other languages and dialects brought with the migratory and immigration flows, over the centuries. One cannot omit the presence of clerics, always present in the region, who contributed to the lexical constitution of the region and are responsible, among other things, for naming the three main municipalities in the Great ABC region, named after three saints: André, Bernardo and Caetano. It is known about the intense arrival of black Africans, which lasted until the 19th century. The arrival of immigrants from many countries was intensified at the end of the 19th century and at the beginning of the 20th century, and continues to this day¹⁴. The state of São Paulo was the main region for attracting immigrants in Brazil (with almost sixty percent of all foreigners entering the country). Until a few decades ago, most of these immigrants entered our territory through the Port of Santos and then went up the mountains to the Great ABC region and the city of São Paulo. Those who had a work contract continued their journey to other locations, but many settled in the region. And so, the municipalities and paths of Great ABC became increasingly important in the regional and national context. The immigration of Italians, Portuguese, Spanish, Germans, Jews, Arabs, Japanese, among others, stands out. It is worth drawing attention to the English, who arrived to build the Santos-Jundiaí Railway from 1860 onwards and settled, first in Paranapiacaba and, later, close to the route of the railway which, in the Great ABC region, passes through Paranapiacaba, Rio Grande da Serra, Ribeirão Pires, Mauá, Santo André and São Caetano do Sul.

In addition to immigrants, the Great ABC region has, in its history, migratory flows from different parts of the country. According to the 2010 census¹⁵, released by IBGE (Brazilian Institute of Geography and Statistics (IBGE)), a quarter of the population of the Great ABC region is made up of migrants. Internal mobility in the Great ABC region is also constant, since, for various reasons, families move within the seven cities in the region.

The processes, whether migration or immigration, existing throughout the history of the Great ABC region, intensified in some periods, inevitably generated and continue to generate very important impacts on socioeconomic and demographic development and, in particular, on customs and culture of the people. As language is the main cultural element of a people and the means of

¹² For more information, you can consult, among others, São Paulo (2014), and Nascimento (2021).

¹³ For more information about the Bilings Dam, see Caastilho (1997).

¹⁴ According to IBGE (2024), it is estimated that 4 million Africans were forcibly brought to Brazil between the 16th and mid-19th centuries. More information about Brazilian immigration can be found on the IBGE page "Brazil 500 years": <https://brasil500anos.ibge.gov.br/territorio-brasileiro-e-povoamento.html>.

¹⁵ Information about the 2010 Census is available at: <https://censo2010.ibge.gov.br/sobre-censo.html>.

transmitting that culture, LPB evidently reflects all the influences of people from different locations in Brazil and the world.

Another aspect to consider when dealing with dialectal non-cohesion in the Great ABC region is related to socioeconomic issues in the region. It must be considered that the Great ABC Region, as Silva (2013, p. 175) points out, must be understood as “[...] a universe of practices experienced by the different human groups that are part of it; which encompass relief, personal relationships, family memory, working conditions, sexuality, association, among others”. The delimitation of its municipalities, therefore, is not established by territorial and political limits, but by who lives in the locality. Often, the territorial limits are even unknown to the population or to those who pass through the region, since, in some points, it is even unclear which municipality you are in. This occurs, for example, because there are places where, on a street, on one side is the municipality of Santo André and, on the other, São Caetano do Sul.

As for the segments into which the region's productive activities are divided, today there is a well-defined demarcation: Santo André and Mauá, are dedicated to petrochemical and plastics industrial sectors; São Bernardo do Campo is known for the presence of the automobile and auto parts industry; São Caetano do Sul and Diadema, despite the presence of industry, stand out in the commercial and services sector; Ribeirão Pires and Rio Grande da Serra, as they are in watershed regions, have a shortage of industries.

Social inequality is also discrepant as there is a contrast between: the three richest municipalities, Santo André, São Bernardo do Campo and São Caetano do Sul; and the four poorest municipalities, Diadema, Mauá, Ribeirão Pires and Rio Grande da Serra. When we refer to financial issues, the discrepancy is huge. To give a concise example, we can observe that the largest municipal Gross Domestic Product (GDP) in the region is São Bernardo do Campo, which, in fact, is the fourth largest in the state of São Paulo, the state with the highest GDP in Brazil; however, when analyzing GDP per capita, São Caetano do Sul has practically twice as much compared to the aforementioned municipality. With regard to the Human Development Index (HDI) in the region, the discrepancy between municipalities is also striking. São Caetano do Sul has the highest HDI in Brazil, being considered the best municipality to live in our country. Santo André and São Bernardo do Campo occupy good positions 15 and 28, respectively, in the ranking. The other municipalities in the region, in turn, do not have such a good HDI, being in the following positions: Ribeirão Pires, 100; Diadem, 420; and Rio Grande da Serra, 565¹⁶.

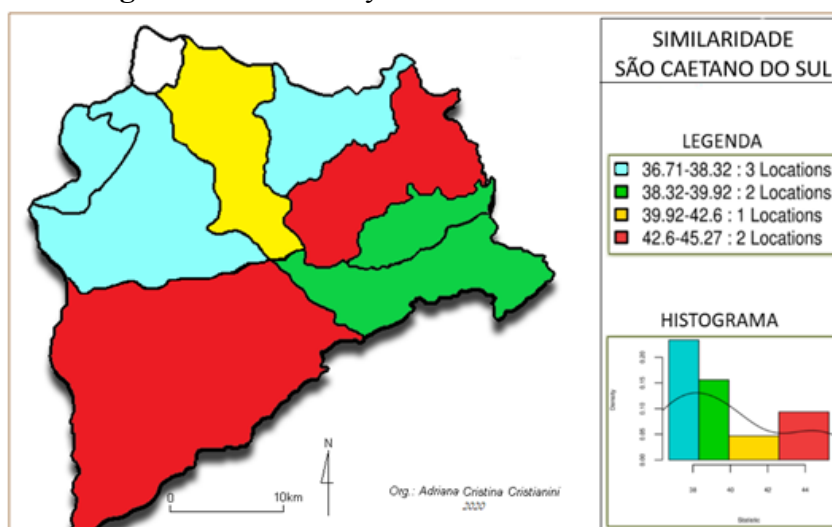
We seek to demonstrate, through the brief observations listed about the municipalities in the Great ABC region, that heterogeneity is very large, in several aspects, in the municipalities of the region. Given this, we can conclude that it is unquestionable that there are many reasons why this heterogeneity is also evident in the linguistic uses of the population.

We now move on to the analysis using synoptic maps.

When observing the results, we realized that, statistically, São Caetano do Sul, compared to the other points, presents a low similarity (or high difference), as most points are below the average, as we can see in Figure 10.

¹⁶ Data from IBGE and administrative records, as specified in the metadata available at: <http://atlasbrasil.org.br/acervo/biblioteca>.

Figure 10 – Similarity – São Caetano do Sul

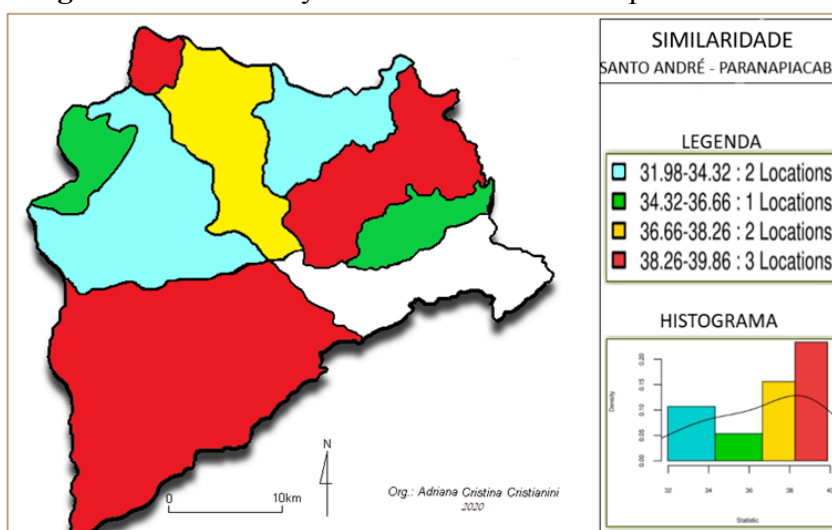


Source: Prepared by the author.

We noticed that, in relation to São Caetano do Sul, the histogram shows a curve markedly to the left with a predominance of cold colors, that is, the analysis indicates that there is a low level of similarity when compared to the points in the region as a whole.

The same does not occur with Santo André – Paranapiacaba, because, as we can see in Figure 11, there is a tendency, in the histogram, for values above the average, which means that there is a high degree of similarity with most other points in the region.

Figure 11 – Similarity – Santo André – Paranapiacaba



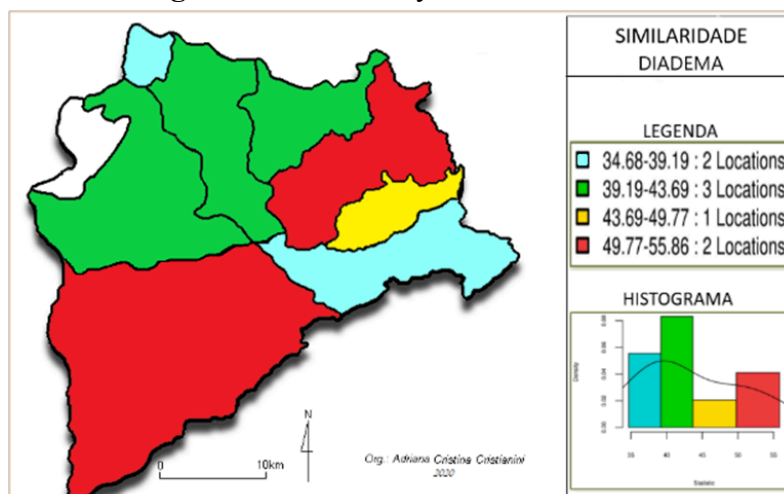
Source: Prepared by the author.

When analyzing each of the other points (Diadema, Mauá, Ribeirão Pires, Rio Grande da Serra, Santo André, São Bernardo do Campo and São Bernardo do Campo – Rural Zone) with regard to the degree of similarity, one consideration was constant: the linguistic uses of each of these points invariably have a low degree of similarity with São Caetano do Sul and Santo André – Paranapiacaba.

Analyzing the similarity of other points, related to the set, we observed that the Diadema point, in Figure 12, through analysis, presents a histogram that represents a low degree of similarity,

with a curve to the left that means a level of similarity below the average. As previously mentioned, São Caetano do Sul and Santo André – Paranapiacaba are points with a color that indicates low similarity.

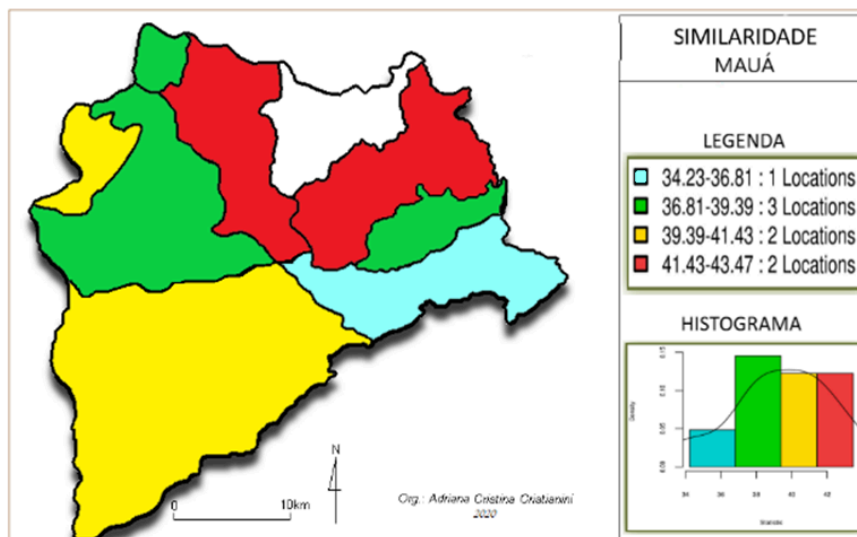
Figure 12 – Similarity – Diadem



Source: Prepared by the author.

Mauá, in turn, has a higher degree of similarity, mainly with the points that constitute neighbors in a geographic continuity, as seen in Figure 13. However, in relation to the points São Caetano do Sul and Santo André – Paranapiacaba (in addition to São Bernardo do Campo and Ribeirão Pires), we have, again, a low degree of similarity.

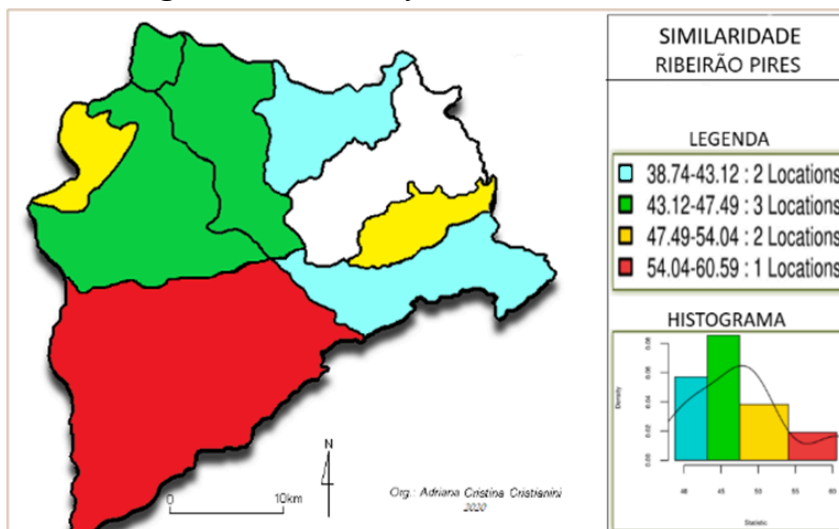
Figure 13 – Similarity – Mauá



Source: Prepared by the author.

Ribeirão Pires, as shown in Figure 14, presents a level of similarity below average, with a curve in the histogram to the left. The point São Bernardo do Campo – Rural zone is the one that is closest to Ribeirão Pires in terms of linguistic issues of a semantic-lexical aspect. The same low degree of similarity occurs in relation to São Caetano do Sul and Santo André – Paranapiacaba.

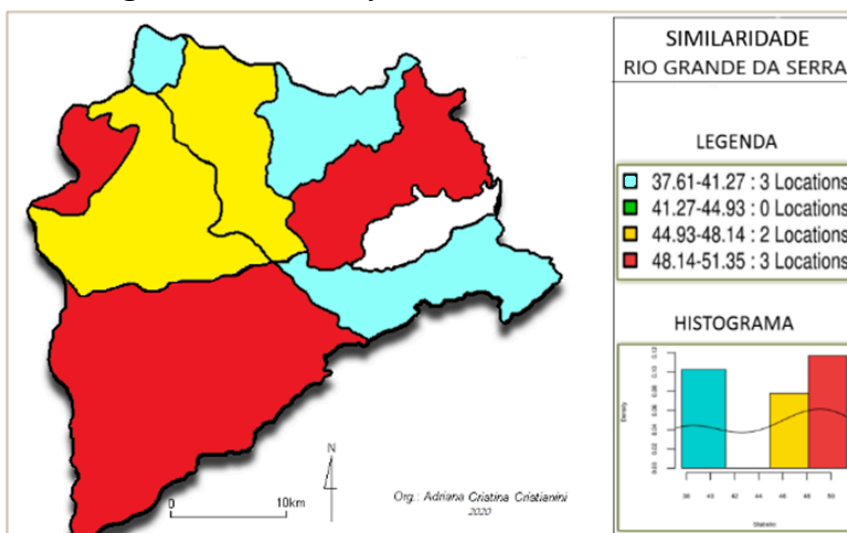
Figure 14 – Similarity – Ribeirão Pires



Source: Prepared by the author.

Rio Grande da Serra, as shown in Figure 15, shows similarity with a tendency to the right in the histogram, above the average. Ribeirão Pires, São Bernardo do Campo – Rural zone and Diadema are the points with the highest levels of similarity when related to Rio Grande da Serra. Once again, we observed that, relating Rio Grande da Serra to São Caetano do Sul and Santo André – Paranapiacaba, there is low similarity.

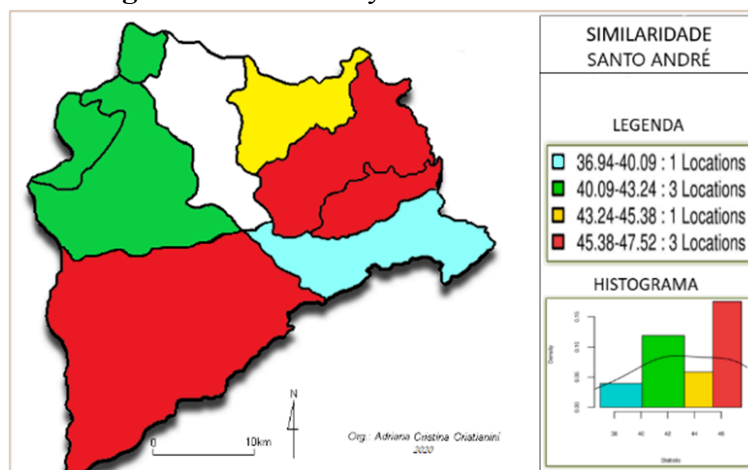
Figure 15 – Similarity – Rio Grande da Serra



Source: Prepared by the author.

Figure 16 shows that the Santo André point has greater similarity with São Bernardo do Campo – Rural zone, Ribeirão Pires, Rio Grande da Serra and Mauá, but the level of similarity is low when related to São Bernardo do Campo, São Caetano do Sul, Diadema and Santo André – Paranapiacaba. Therefore, São Caetano do Sul and Santo André – Paranapiacaba also have low similarity in relation to Santo André. The statistical curve suggests that the similarity is slightly above average.

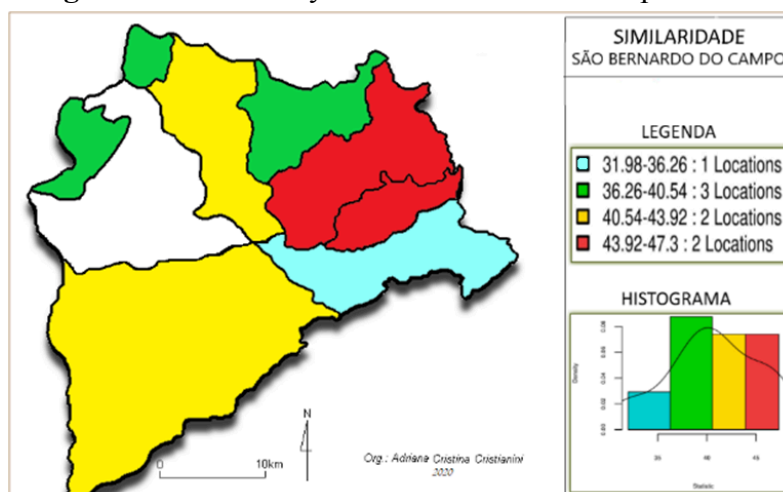
Figure 16 – Similarity – Santo André



Source: Prepared by the author.

It can be said that São Bernardo do Campo, based on Figure 17, has greater similarity with Ribeirão Pires and Rio Grande da Serra; still with a high degree of similarity with São Bernardo do Campo – Rural zone and Santo André. The level of similarity is low when related to the points Diadema, Mauá, São Caetano do Sul, and, mainly, Santo André – Paranapiacaba.

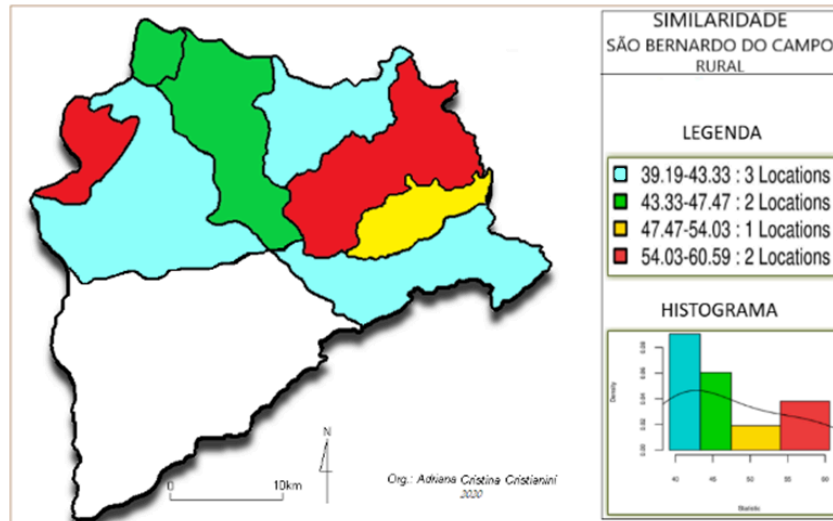
Figure 17 – Similarity – São Bernardo do Campo



Source: Prepared by the author.

Figure 18 shows the similarity of the point São Bernardo do Campo – Rural zone related to the other points. We observed that the similarity of this point is below the average, represented in the histogram with a curve to the left. The points Diadema and Ribeirão Pires have more linguistic similarity with São Bernardo do Campo – Rural zone, followed by Rio Grande da Serra. The points with a low level of similarity when related to São Bernardo do Campo – Rural area are São Caetano do Sul, Santo André and, even less similar, São Bernardo do Campo, Mauá and Santo André – Paranapiacaba. Once again, we find São Caetano do Sul and Santo André – Paranapiacaba in the groups whose similarity is considered low.

Figure 18 – Similarity – São Bernardo do Campo – Rural area



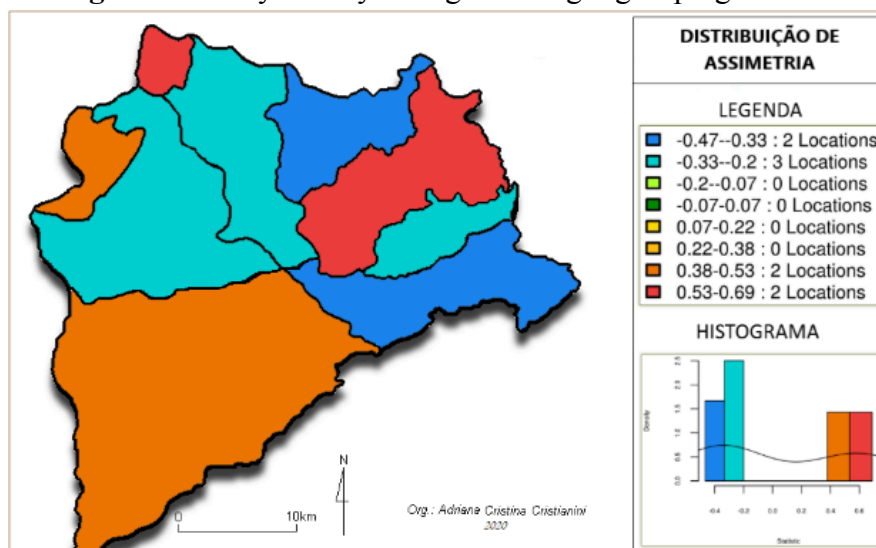
Source: Prepared by the author.

In view of the above, we can observe that regardless of which point is selected as a reference for the similarity analysis related to the remaining points, São Caetano do Sul and Santo André – Paranapiacaba always present a level of similarity below the average. Only when the reference point is Santo André – Paranapiacaba itself, do we have São Caetano do Sul with an above average grade.

However, even though Santo André – Paranapiacaba and São Caetano do Sul present similar results when another point is a reference in the similarity analysis, we can verify that the integration of these two points in the set is divergent when we calculate the asymmetry distribution.

As we can see in Figure 19, five points have high integration (São Bernardo do Campo, Santo André, Mauá, Rio Grande da Serra and Santo André – Paranapiacaba) and four have low integration in the region (Ribeirão Pires, São Bernardo do Campo – Rural area, Diadema and São Caetano do Sul).

Figure 19 – Asymmetry cartogram – eight groupings



Source: Prepared by the author.

The asymmetry cartogram shows us that, in addition to the points with low integration being fewer (four) than those with high integration (five), no point in the group with low integration in the set is contiguous to each other. The same does not occur in the group with points with high integration.

As our view, this study, is more focused on the points São Caetano do Sul and Santo André – Paranapiacaba, it is important to highlight that, with regard to integration as a whole, these two points have opposite results, as São Caetano do Sul has a degree minimum level of integration, in red, and Santo André – Paranapiacaba, the maximum degree, in blue.

CONCLUSION

In this work, we set out to highlight the global analysis of linguistic uses based on ASL_ABC, especially in the areas of São Caetano do Sul and Santo André – Paranapiacaba.

We found that São Caetano do Sul and Santo André – Paranapiacaba, possibly due to their socio-historical-geographical constitution, have their dialects in distinct groups from the others if we segment the corpus into more than three clusters. Furthermore, when we reflect through the study of similarity starting from the other points with São Caetano do Sul and Santo André – Paranapiacaba, we realize that there is always a degree of similarity below the average, except when the reference point is Santo André – Paranapiacaba, with which São Caetano do Sul has a high degree of similarity.

This last consideration is evident when observing the similarity cartograms of the points Diadema, Mauá, Ribeirão Pires, Santo André, São Bernardo do Campo and São Bernardo do Campo – Rural zone with the complex. In them, São Caetano do Sul appears in the green echelon in the similarity cartograms of Mauá, Santo André, Ribeirão Pires, São Bernardo do Campo and São Bernardo do Campo – Rural area and in the blue echelon in the Diadema and Rio Grande da Serra cartograms, always indicating below average degree of similarity. Santo André – Paranapiacaba, in turn, appears in the blue category in all these cartograms, also indicating a low degree of similarity.

Starting from the similarity cartograms of São Caetano do Sul and Santo André – Paranapiacaba, relating them to the set, we realize that Santo André – Paranapiacaba occurs in the green category, presenting a degree of similarity below average also in the São Caetano do Sul cartogram. However, in the Santo André – Paranapiacaba cartogram, São Caetano do Sul is among the points with the highest level of similarity, that is, in red.

With the asymmetry analysis, we noticed that the points São Caetano do Sul and Santo André – Paranapiacaba diverge, as the latter has maximum integration in the data set, while the former has minimum integration. The present work is an example of how Dialectometry allows us, starting from data from a linguistic atlas, through quantitative analysis, to carry out a global analysis of linguistic phenomena in a given region. With the linguistic atlas, we have a landscape of the linguistic norm used by speakers with regard to the use of each lexical item that was part of the list of concepts selected for the research. A linguistic set view, however, is only possible through mathematical methods, in our case, through dialectometricization.

REFERENCES

ACADEMIA BALDERIX. Coeficiente de asimetría 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 dialectométrica 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 Dialectometria ajudam-nos a perceber. *In: I CIPAL – Ciclo de Palestras em Dialectologia, 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 Dialectal 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 dialetal madeirense: estudo dialectométrico 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 dialectométrico do Atlas Lingüístico – Etnográfico dos Açores. *In: GRANJA, M. A.; AGRELO, A. B.; SEOANE, E. G. (ed.). Aproximações á variación lexical no domínio galego-portugués. Revista Galega de Filoxía: Monografia*. 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 dialectométrica do Atlas Lingüístico-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.consortioabc.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 estudios 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. Introducción 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.