Mapping the variation of /R/ in internal syllabic coda in the Southeast Region: contributions of QGIS

Helen Cristina da **SILVA***.

* PhD in Language Studies from Universidade Estadual de Londrina (2016). Associate Professor II_UNIPAMPA (Campus Bagé). helensilva@unipampa.edu.br

Abstract

In this article, we present an excerpt from the author's doctoral thesis entitled *Pelas veredas do /r/ retroflexo* (Through the pathways of the retroflex /r/), defended at the Universidade Estadual de Londrina. On this occasion, we present the mapping of the variation of /r/ in internal syllabic coda, based on unpublished data collected by the *Atlas Linguístico do Brasil* (ALiB) in the Southeast region. Grounded in the assumptions of Geolinguistics, the research focuses on the diatopic distribution of the retroflex /r/, conceived in this study as one of the phonetic elements that may contribute to the dialectal division of Brazil (Aguilera 2020a; 2020b). The cartography of the results was conducted using the QGIS software, a Geographic Information System (GIS), which situates this study as pioneering in Brazil, by employing this tool for the mapping of linguistic data. Throughout the analysis, our aim is to present the vitality of the retroflex /r/, contradicting the prediction of its disappearance, as suggested by Amaral (1982 [1920]), and at the same time, to provide evidence confirming the connection of the retroflex /r/ to the historical movement of the Bandeiras Paulistas. To achieve this, we present the maps elaborated with QGIS, which can assist in representing the results, understanding, and consequently, analyzing the data of this research and others of linguistic nature.

Keywords:
Variation of /r/ in syllabic coda. Retroflex /r/. QGIS

Mapping the variation of /r/ in internal syllabic coda in the Southeast region: contributions of QGIS

Helen Cristina da Silva

Introduction

In this article, we present a segment of the research titled "Pelas veredas do /r/ retroflexo" (Through the pathways of the retroflex /r) (Silva, 2016), still unpublished, developed for the author's sandwich doctoral thesis, defended by the Universidade Estadual de Londrina and the Universidade de Santiago de Compostela. Based on Pluridimensional Dialectology (Thun, 1998), the general objective of the research was to map the variation of /r/ in syllabic coda in the Southeast region, focusing on the retroflex /r/. This phoneme, alongside other phonetic phenomena such as the realization of pretonic vowels or the voiceless and voiced fricative consonants [s] and [z] in syllabic coda, may characterize a proposal for dialectal division of Brazil (Aguilera; Silva, 2011; Silva, 2012, 2016).

The analyzed corpus consists of 336 interviews collected by the Linguistic Atlas of Brazil (ALiB) in 80 locations in the Southeast Region. The choice of this region, among other reasons, stems from the fact that São Paulo is believed to be the birthplace of the /r/ *caipira* and thus provides evidence both to confirm the hypothesis of the maintenance and diffusion of this phoneme and to demonstrate its importance for dialectal delimitation in Brazil. Additionally, the complex socio-history of the Southeast, coupled with its position in the national and international economy, endows it with undeniable importance for understanding the constitution of Brazilian Portuguese.

Therefore, in this article, we present some maps[1] developed for the thesis, along with their respective analyses. It is important to emphasize that this is a pioneering mapping in Brazil[2] in linguistic works, using the georeferencing software QGIS (Quantum Geographic Information System), whose "visualization techniques simplify the work of researchers and also help to discover patterns of behavior and relationships in the data that would otherwise remain hidden" (Moutinho *et al.*, 2019, p. 256). This technology is widely used in the cartography of dialectological studies developed at the Institute of Galician Language, Spain, where the author acquired the basic commands to operationalize it[3].

As is known, since its establishment in the early 20th century, Linguistic Geography, now known as Geolinguistics, has been adding numerous theoretical and methodological contributions to dialect studies. Among the advances, the improvement of digital tools in the elaboration of linguistic maps stands out, which, especially since the advent of Sociolinguistics, underwent a process of resignification in order to incorporate other factors beyond the diatopic (Cardoso, 2010). In Brazil, the great impetus of the field came with the development.

---

[1] Due to physical space limitations, as well as the scope of this article, it was not possible to present all the linguistic features developed for the thesis. Additionally, other social factors, beyond the diatopic ones, were taken into account for data analysis. Therefore, I invite the reader to consult the complete work at: Silva (2016).

[2] As far as we know, this is the first research to utilize the QGIS tool for this purpose, followed by the work of Dias (2017), which employed the tool for the development of maps in the *Atlas Geossociolinguístico Quilombola do Nordeste do Pará (AGQUINPA.*

[3] I reiterate here my sincere thanks to the entire research group at the *Instituto de Lingua Galega (ILG)*, especially to the computer scientist Cesar Osorio Peláez, Technical Specialist in Systems Informatics, for teaching me, step by step, how to use the basic commands of QGIS. I also extend my gratitude to Professor Xulio Sousa Fernández and Professor Rosário Álvarez Blanco, my co-supervisor, for their careful analyses and suggestions, which greatly contributed to the improvement of the maps and the study as a whole.

Currently, in addition to ALiB and state and small domain atlases, several studies focused on the variation and description of Brazilian Portuguese rely on cartography due to its proven efficiency and practicality in data analysis and, above all, in the photographic visualization of linguistic reality that only this form of representation seems capable of providing.

Among the facilitating programs for this task is QGIS, an open-source Geographic Information System (GIS) available for free, which enables, among numerous features, the generation of maps, including linguistic ones. According to Teles and Ribeiro (2006, p.213), GIS, since the 1980s, has revolutionized the concepts and techniques of research, analysis, and data manipulation, going beyond mapping, seeking to "answer the questions asked by the user, allowing the execution of complex combinations of data, as well as their analysis, with the possibility of automating calculations to evaluate real-world phenomena." Among the advantages of generating maps in QGIS, we highlight: graphic quality, time-saving in map composition, the possibility of exporting the map in various formats and supports, and the creation of customized maps (Dias, 2017).

1. Methodological Aspects

The corpus of this research consists of unpublished data from the *Atlas Linguístico do Brasil* (ALiB) archive (Cardoso *et al.*, 2014), covering the Southeast region, including both capital cities and inland towns. In total, we analyzed material from 80 linguistic points distributed as follows: 38 in the state of São Paulo, 23 in Minas Gerais, 14 in Rio de Janeiro, and five (5) in Espírito Santo, totaling 336 informants stratified by sex, age group (Group I – 18 to 30 years old, Group II – 50 to 65 years old), place of origin (native of the region), and educational level (complete or incomplete primary education and complete higher education, in the case of capital cities). From this material, the analyzed segment consists of questions from the Phonetic-Phonological Questionnaire (QFF) of the ALiB 2001 Questionnaires (National Committee…, 2001), which include responses containing /r/ in internal syllabic coda, namely: "torneira" (tap/faucet); "gordura" (fat); "fervendo" (boiling); "árvore" (tree); "borboleta" (butterfly); "tarde" (afternoon); "catorze" (fourteen); "pernambucano" (from Pernambuco); "certo" (certain); "perdão" (forgiveness); "perfume" (perfume); "dormindo" (sleeping); "perdida" (lost); and "esquerdo" (left).

After review, tabulation, and analysis, we submitted the data to the Geographic Information System (QGIS) [4], an open-source software licensed under the GNU General Public License, which enabled the generation of computerized maps, allowing for immediate and clear visualization of the phenomena under study.

"The QGIS software is an application that integrates digital data of any nature with geographic information (referred to as rasters) and can be arranged into vector maps" (Dias, 2017, p.148). One of its conveniences is the use of geographic meshes, which can be downloaded directly from the website of the Brazilian Institute of Geography and Statistics (IBGE), ensuring the reliability of the material.

Each map generated in the QGIS software consists of the combination of several map layers (shapefiles). The layers that compose the maps elaborated for our thesis are, respectively: region; mesoregion; municipalities; and, finally, the layer of linguistic data, previously tabulated in an Excel spreadsheet. After joining these layers, we can choose numerous forms of representation, adapting them to the study's objective. In our case, we used two: i) pie charts, which present the data more objectively, offering the percentages of each variant; and ii) Voronoi diagrams[5], which

---

[4] For the development of the linguistic maps, we utilized version 2.18 of QGIS (https://qgis.org/pt_BR/site/index.html); meanwhile, the geographic meshes were downloaded from the IBGE website (http://downloads.ibge.gov.br/downloads_geociencias.htm)

[5] Given a set of points (generators) in a plane, a Voronoi Diagram is nothing more than a subdivision of this plane into regions formed by the places closest to each of the points." Available at: http://sweet.ua.pt/leslie/Geocomp/Slides/GC_09_10_7_Diagramas_Voronoi.pdf. Accessed: September 2023.

allow us to transform the map of locations into a thematic map of polygons where the perimeter of each polygon, equidistant from neighboring points, defines an area of influence. The choice of this type of presentation of isophonic areas resides in the fact that it is a form of representation and visualization that facilitates the understanding of results obtained from quantitative analyses, since the relationships between compared spatial elements become more evident (Moutinho *et al.*, 2019).

The finalization of the maps is done in the print composer area of the software, where it is possible to insert the basic elements useful for reading the map, such as: title, legend, scale, north arrow, compass rose, frames, and so on. In this last step, it is possible to save the file in PNG or PDF formats.

2. The Retroflex /r/ in Brazil: Some Considerations

The Portuguese language transplanted to Brazilian lands underwent various processes that, combined, endowed it with a new appearance, attributing it with unique characteristics until it solidified into Brazilian Portuguese, reflecting the miscegenation of its people and its territorial extension. Upon arriving in Brazilian lands, European Portuguese encountered a strong rival, the Tupi language, spoken by various indigenous groups and maintained for a long time by mestizos, a fact that instilled it with significant lexical influence. Subsequently, it acquired significant African tones and, later, with the arrival of large waves of immigrants, underwent profound transformations in all its aspects.

From this mixture, some peculiarities stand out and aid in demarcating the speech varieties that constitute Brazilian Portuguese, such as the variants of /r/ in syllabic coda, particularly the retroflex /r/, which, according to the hypotheses we rely on, resulted from the contact between Tupi and Lusitanian Portuguese. From the village of São Vicente, this phoneme, the only genuinely Brazilian one, accompanied the incursions of the bandeiras (exploratory expeditions) and was disseminated by the São Paulo pioneers throughout the state's interior, radiating outwards and extending its boundaries to the states of Minas Gerais, Goiás, Mato Grosso, Mato Grosso do Sul, Paraná, Santa Catarina, and Rio Grande do Sul.

In 1920, Amadeu Amaral, in his pioneering work in the field of Dialectology in Brazil, *O Dialeto Caipira* (The Caipira Dialect), described morphological, phonetic, lexical, and syntactic aspects of the dialect spoken in the former province of São Paulo until the end of the 19th century. He mentions one of the phonetic characteristics of this dialect as the retroflex /r/, which has since been dubbed the /r/ *caipira*. According to the author, this dialect had expanded in the region and was even part of the educated class, a fact that gave the *paulistas*[6] the reputation of corrupting the vernacular with many and ugly language vices." He further states that "caipirism" existed not only in language but in all manifestations of our provincial life. However, the profound social changes that Brazil underwent since the colonial period led the author to affirm that the *caipira* dialect, "cornered in small localities that did not closely follow the general movement of progress […], is condemned to disappear in a more or less short time" (Amaral, p. 41-42).

Despite some research revealing possible linguistic prejudice directed at /r/ *caipira* (Silva, 2012, 2015; Botassini, 2009; Leite, 2004, 2010; Guioti, 2002; Head, 1987), dialectal studies (Aguilera, 2008, 2012, 2020a; AGuilera; Silva, 2011; Almeida, 2018; Brandão, 1991, 1995, 1997, 2007, 2017; Castro, 2006, 2009, 2011, 2013; Cuba, 2015; Silva, 2012; 2016; Callou *et al.*,1997, among others) tend to contradict Amaral's (1920) prediction, especially regarding the vitality of the retroflex /r/, one of the most authentic marks of the caipira dialect, demonstrating that this phoneme is alive and perhaps expanding in Brazilian Portuguese. According to Aguilera (2020, p.267):

---

[6]   People who live in São Paulo state.

Amaral (1976) surely never imagined, at the beginning of the 20th century, that such intense advancements in all sectors of human activities, especially in the fields of science and technology, would so radically transform Brazilian society in the 21st century, which at the time had a population of around 31 million inhabitants and the state of São Paulo with less than five million. We also believe that he never would have imagined that the influence of the caipira dialect would be so prominent in our urbanized society, whether in music, food, speech, particularly in the use of the retroflex /r/. It would have been very difficult to predict that, a hundred years later, the retroflex /r/, the most significant mark of the caipira dialect, would have the vitality it currently possesses[7]. (Our translation).
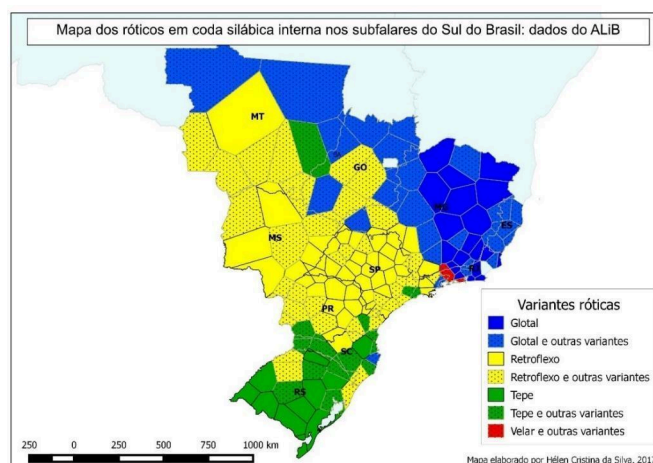
That said, after a century since the publication of *O Dialeto Caipira*, this phoneme, nationally known as the */r/ caipira*, has become one of the most representative features of the speech of people from the countryside, although it is not currently restricted solely to this reality. This incites researchers from all over the country to study its origin, its expansion, the social variables influencing its use, and also the linguistic attitudes it arouses. Adding to this fact, its well-marked distribution may characterize, alongside other phonetic aspects, a proposal for a dialectal division of Brazil, serving as a valuable contribution to discussions about the origins of Brazilian Portuguese.

The *Atlas Linguístico do Brasil* – AliB (Cardoso *et al.*, 2014) presents six phonetic maps on the variation of /r/ in coda position. Among these, map F04 C6 shows that the retroflex /r/ occurs most frequently in the capitals of Mato Grosso, Mato Grosso do Sul, and Goiás, followed by São Paulo and Paraná, and is present, to a lesser extent, in Santa Catarina and Rondônia. Although we do not yet have the mapping of data from the interior in the other regions, it is already possible to outline an isophone of this variant, which is concentrated in the Midwest, part of the Southeast, and South of Brazil, to be completed by the next volumes of the atlas.

The diatopic distribution of the rhotics, in internal syllable coda, combined with other linguistic phenomena uncovered by studies of the ALiB, led Aguilera (2009, 2020a, 2020b) to advocate the need for a new interpretation of the pioneering division proposed by Nascentes (1953) of the sub-dialects that make up the Southern Speech, namely: the *Baiano*, the *Fluminense*, the *Mineiro*, and the *Sulista*. Thus, the dialectologist argues (Aguilera, 2020b, p. 265), based on the presence and expansion of the retroflex /r/, for the insertion of a fourth sub-dialect into this group: "the Paulista, which would cover the states of Mato Grosso, Mato Grosso do Sul, São Paulo, Paraná, part of Goiás, Minas Gerais, Santa Catarina, and Rio Grande do Sul." To illustrate and support such a proposal, Aguilera (2020b) presents a linguistic map (Figure 1), developed in the QGIS software by the author of this article, with unpublished data from the ALiB.

Figure 1 – Map of rhotics in internal syllable coda in the sub-dialects of Southern Brazil: ALiB data

---

[7] (Original text): Amaral (1976), com certeza, jamais teria imaginado, no início do século XX, que avanços tão intensos em todos os setores das atividades humanas, sobretudo na área da ciência e da tecnologia, iriam transformar tão radicalmente a sociedade brasileira do século XXI que, na época, contava com uma população em torno de 31 milhões habitantes e o Estado de São Paulo com menos de cinco milhões. Também acreditamos que jamais teria imaginado que a influência do caipira seria tão marcante em nossa sociedade urbanizada, seja na música, na alimentação, na fala, em particular no uso do /r/ retroflexo. Seria muito difícil prognosticar que, cem anos depois, o /r/ retroflexo, a marca mais significativa do dialeto caipira, teria a vitalidade que tem atualmente.

Mapa dos róticos em coda silábica interna nos subfalares do Sul do Brasil: dados do ALiB

Variantes róticas
- Glotal
- Glotal e outras variantes
- Retroflexo
- Retroflexo e outras variantes
- Tepe
- Tepe e outras variantes
- Velar e outras variantes

Mapa elaborado por Hélen Cristina da Silva, 2017.
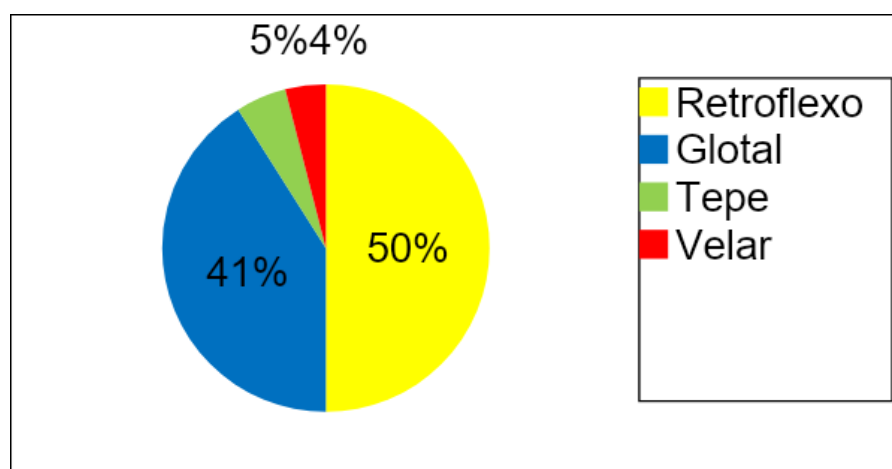
Source: AGUILERA (2020b, p. 266).

The linguistic map represented by Figure 1 demonstrates that, as expected when dealing with language, there are zones of rhotic intersections, but there is a spatially well-defined coexistence of three variants: the glottal, the retroflex, and the tap, distributed as follows: in the northern portion, the glottal covers the states of Espírito Santo, Rio de Janeiro, parts of Minas Gerais, Goiás, and Mato Grosso; in the central part, the retroflex, in São Paulo, Paraná, Mato Grosso do Sul, and parts of Goiás and Mato Grosso; and in the south, the tap.

In summary, this linguistic snapshot provides support to reinforce the proposal of a "Paulista" speech in Brazil, based on the distribution, especially, of the retroflex /r/. It also allows us to affirm that the center of irradiation of this variant is the Southeast region, mainly the state of São Paulo, from where the yellow color seems to spread to all adjacent areas. With that said, we present in the following section the detailed distribution of the retroflex /r/ in this region and how it can aid in the understanding of this rhotic.

2.2 The Retroflex /r/ in the Southeast Region: Mapping and Analysis

Based on the 80 locations comprising the corpus of this research, we recorded 4.829 occurrences. As shown in Graph 1, the realizations of /r/ in coda position are distributed as follows: 50% (2.396) retroflex; 41% (2.002) glottal; 5% (245) tap; and 4% velar.
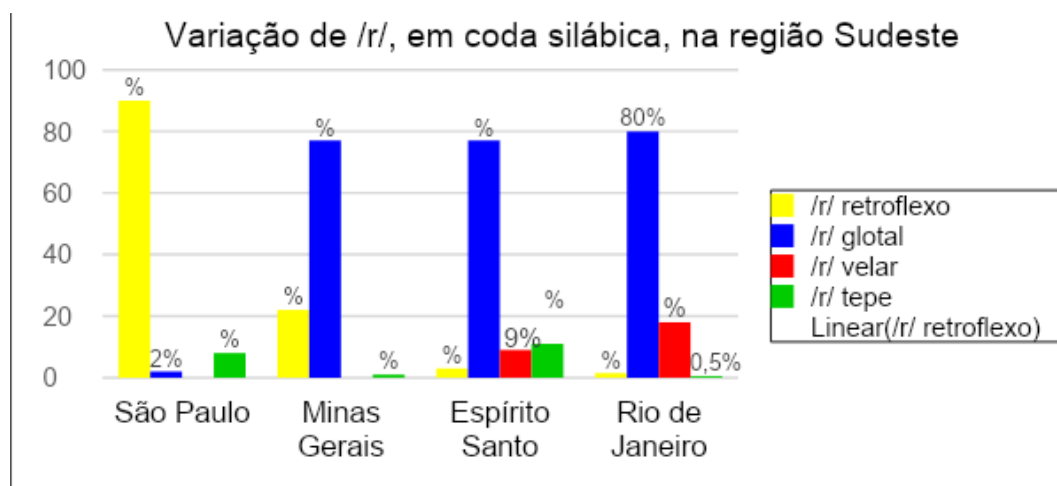
Graph 1 – Overall occurrence of /r/ variants in internal coda in the Southeast Region

Source: compiled by the author with unpublished data from the ALiB

The numbers indicate the existence of four variants for /r/ in coda position in the Southeast Region, with the retroflex (50%) being the most used by speakers, followed by the glottal (41%); whereas the other variants, namely tap and velar, present much lower percentages, 5% and 4% respectively, in the overall count. Upon initial analysis, the results reveal that the characteristic rhotic of the Southeast Region is the /r/ *caipira*. However, these data need to be contextualized, taking into account the distribution of records by state, which is depicted in Graph 2.

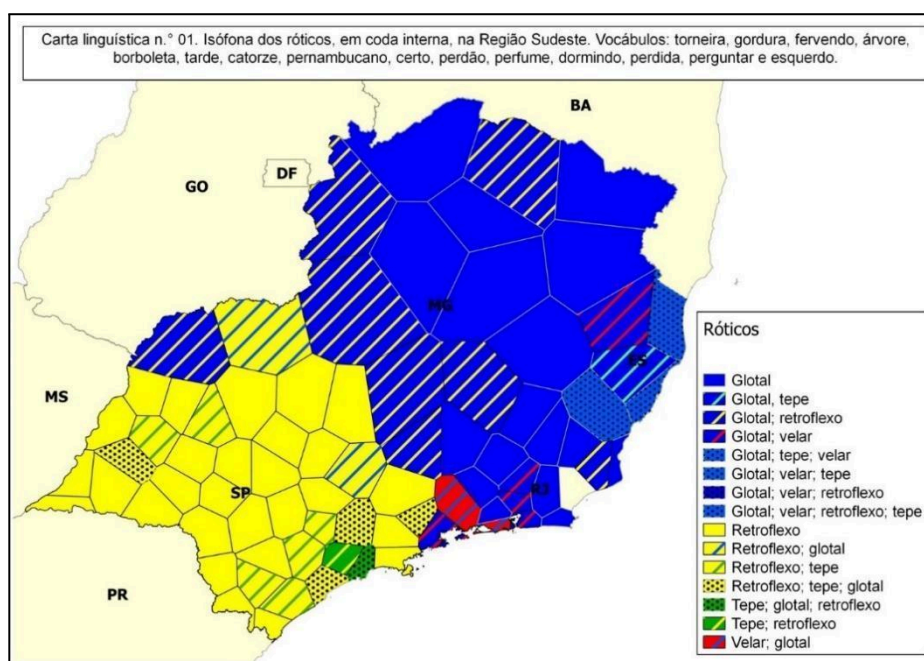**Graph 2** – Occurrence of /r/ variants in internal coda in the states of the Southeast Region

According to Graph 2, in the state of São Paulo, 2.300 rhotics were recorded, among which 90% (2.059) correspond to the retroflex, 8% (196) to the tepe, and 2% (45) to the glottal. In Minas Gerais, the total amounts to 1,395 rhotics, with 77% (1.068) being the glottal /r/, 22% (314) the retroflex /r/, and only 1% (12) the tepe /r/, with a single occurrence of the velar /r/ (0%). In Espírito Santo, we obtained 332 records, with 77% (258) glottal, 11% (36) velar, 9% (32) tepe, and 3% (11) retroflex. Finally, in Rio de Janeiro, we recorded 793 rhotics, with 80% glottal (636), 18% (140) velar, 1.5% (12) retroflex, and 0.5% (5) tepe.

To comprehend these results in the context of the diatopic dimension and to fulfill one of the objectives proposed in this study, we present the results through Figure 2, which depicts isophonic areas of /r/ in coda syllable in the Southeast Region.

Figure 2. - Isophonic areas of rhotics in internal coda in the Southeast Region

Carta linguística n.° 01. Isófona dos róticos, em coda interna, na Região Sudeste. Vocábulos: torneira, gordura, fervendo, árvore, borboleta, tarde, catorze, pernambucano, certo, perdão, perfume, dormindo, perdida, perguntar e esquerdo.

Source: Silva (2016, p. 122).

The geographical distribution of the variants, as depicted in Figure 2, reveals that the retroflex /r/, represented by the yellow color, and the glottal /r/, by the blue color, form isophonic areas responsible for representing the speech of the Southeast Region, configuring them as important and indispensable phonetic elements for the linguistic delimitation of this area. Alongside these rhotics, albeit with a more modest productivity, are the tepe and the velar.

Regarding the retroflex /r/ specifically, this study confirms its predominance throughout the state of São Paulo, covering the bordering areas of the states of Paraná, Mato Grosso do Sul, and Goiás. It also reveals that this variant is virtually categorical in parts of southern Minas Gerais and the Triângulo Mineiro and presents, albeit sporadically, occurrences in the north and northeast of MG, the south of Espírito Santo, and the northern part of Rio de Janeiro. Thus, it is possible to confirm that, to a greater or lesser extent, the retroflex /r/ is present throughout the Southeast Region.

The significant presence of the retroflex /r/ (variant, starting from São Paulo and entering the state of Minas Gerais, confirms, according to our analysis, the hypothesis that this rhotic would have originated among the mamelucos[8] who inhabited the region of Piratininga and spread through the interior of São Paulo through the incursions of the bandeirantes and later, with the discovery of gold, would have been taken to Minas Gerais.

The results also lead to endorsing Santiago-Almeida's statement (2012, p. 282) that considering the socio-historical-cultural aspects since the first linguistic contacts occurred where São Paulo city is today, in the 16th century when the region was not yet the capital, it is possible to perceive that "the caipira dialect or variety had its origin in a place distinct from what is now known as the interior." In other words, the birthplace of this variety, particularly the retroflex /r/, is not linked to inland areas but rather to the coast of São Paulo, from where the bandeirantes departed, simultaneously expanding territories and leaving marks of their speech.
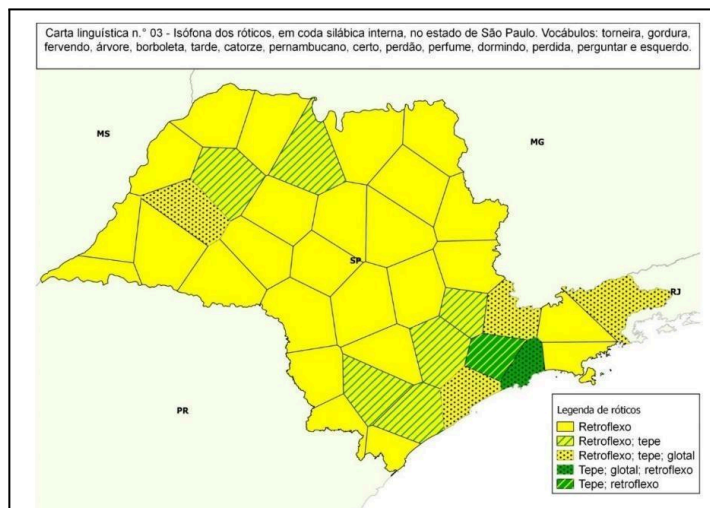
The timid presence of the retroflex /r/ in the states of Rio de Janeiro and Espírito Santo may be linked to migratory movements, of which miners and Paulistas were a part.

Although numbers are essential for understanding the reality intended to be presented in this work, it is undeniable that the linguistic photograph provided by the map presents the data in such

---

[8]   In the Brazilian historical context, "mameluco" typically referred to the children of Portuguese or Spanish settlers with indigenous women

a way that makes the visualization of the phenomena undoubtedly clearer, didactic, and even palpable. To reinforce these assertions, we present two more linguistic maps: the first one (Figure 3) shows isophonic areas of /r/ in coda in the state of São Paulo and a second (Figure 4), in the state of Minas Gerais.

Figure 03 – Isophonic areas of rhotics in internal coda in the state of São Paulo
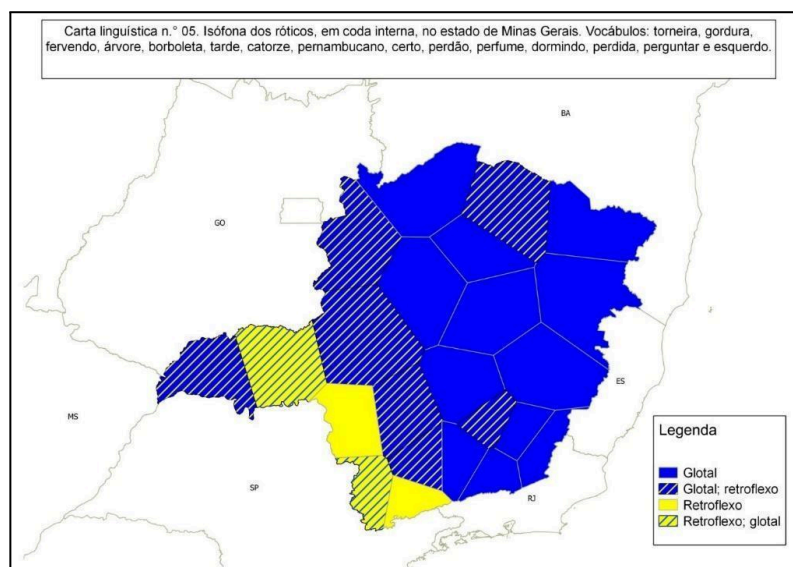

Source: Silva (2016, p. 140).

The base background color of the map represents the predominance of the rhotic in the area; meanwhile, the hatching indicates areas where the two predominant variants, namely the tepe and the retroflex, co-occur, with the base color prevailing. The black dots show the areas where, in addition to these rhotics, the glottal /r/ was also recorded.

With that said, the map reveals that the retroflex /r/ is practically categorical throughout the state of São Paulo, constituting the norm of São Paulo speech, except in the capital and the city of Santos, presenting itself as a rule throughout the bordering strip with the states of Paraná, Mato Grosso do Sul, and Minas Gerais, as well as throughout the São Paulo interior. The tepe, on the other hand, predominates only in the capital and the city of Santos, reflecting occurrences in adjacent points to these two locations and in the western region. The glottal /r/ is more vigorous only in Santos, occurring sporadically in other points.

The map also allows verification of two areas defined by rhotic polymorphism. The first, and more pronounced, is concentrated in points closest to the capital. We attribute this variational pattern to the linguistic influence of the capital, which, as is known, is a cosmopolitan, multifaceted city forged by unquestionable social complexity both nationally and internationally. The second area covers the western region of the state, and the most acceptable reason for the observed variation lies in its historical context of occupation, marked, on one hand, by migrants, mostly from the Northeast, and on the other hand, by immigrants, mainly Italians. These possibly left as a linguistic-phonetic heritage the presence of the tepe variant; those the glottal /r/. Results regarding such influences can also be confirmed in the *Atlas Linguístico Topodinâmico do Estado de São Paulo* (Santos, 2014).

The next linguistic map (Figure 4) presents the diatopic distribution of /r/ variants recorded in the state of Minas Gerais.

Figure 4 – Isophonic areas of rhotics in internal coda in the state of Minas Gerais - data from the ALiB

Carta linguística n.° 05. Isófona dos róticos, em coda interna, no estado de Minas Gerais. Vocábulos: torneira, gordura, fervendo, árvore, borboleta, tarde, catorze, pernambucano, certo, perdão, perfume, dormindo, perdida, perguntar e esquerdo.

Legenda
- Glotal
- Glotal; retroflexo
- Retroflexo
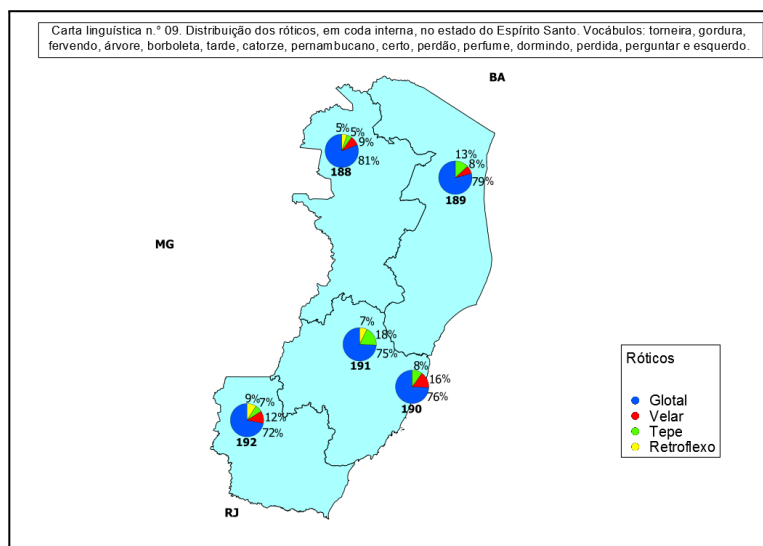- Retroflexo; glotal

Source: Silva (2016, p. 146).

Taking into account the same reading guidelines as the previous map, Figure 4 shows that the region of predominance of the retroflex /r/ in coda syllable is concentrated in the Triângulo Mineiro and southern Minas Gerais, areas bordering the state of São Paulo, losing density (in terms of geographical distribution and in terms of alternation with other variants) as it advances further into the interior of the state. These data confirm results from previous research, such as those obtained by Silva (2012) in his master's thesis entitled " *O /r/ caipira no Triângulo Mineiro: um estudo dialetológico e de atitudes linguísticas,*" and especially those published in the six maps dealing with this aspect, from the *Esboço de um Atlas Linguístico de Minas Gerais* - EALMG - (Ribeiro *et al.*, 1977), as well as the proposed division of MG dialects presented by Zágari (2005, p.50-51) for whom: there is a Bahian dialect that extends from the North to the East-West of Minas, with notable linguistic characteristics such as the predominance of low pretonic vowels, the presence of the affricate, and nasality occurring outside the tonic syllable; a Paulista dialect that encompasses the entire Triângulo Mineiro and the South of the state, characterized by the prominent retroflex /r/; and finally, the Minas Gerais dialect, influenced by colonization during the gold and diamond cycles, "which constantly breaks diphthongs [aj] [ej] and [ow] when not final and creates others when final and preceded by a sibilant: [a'xoys], ['fays] ['nɔys]". Additionally, the diatopic distribution of [ɻ] corroborates the hypothesis of crediting the irradiation of this variant to the bandeirantes, as previously explained.

The state of Espírito Santo, although presenting the glottal variant as the norm throughout its territory, with rates ranging from 72% to 81%, circumscribes itself as the most polymorphic in the Southeast region, as shown in Graph 2 and Figure 5. The data confirm that, in addition to the glottal /r/, the state, after São Paulo, is the one that uses the tepe variant the most, and that this co-occurs with the velar variant; furthermore, the *caipira* /r/, although less prominent, is more frequent than in Rio de Janeiro. The pronounced rhotic variation may, as everything indicates, be linked to the historical migration of miners, Paulistas, and Fluminenses, which would justify the presence of [ɻ], of [h] and of [x]; whereas, the [R] may be attributed to the influence of Italian immigrants, one of the largest ethnic groups that settled in the state.

The simultaneous occurrence of at least three rhotics in each point makes it difficult to establish isophonic areas. Therefore, we present the linguistic map with pie charts (Figure 5), which, in addition to the diatopic distribution, provides detailed data for each investigated locality.

Figure 5 - Distribution of rhotics in internal coda in the state of Espírito Santo

Carta linguística n.º 09. Distribuição dos róticos, em coda interna, no estado do Espírito Santo. Vocábulos: torneira, gordura, fervendo, árvore, borboleta, tarde, catorze, pernambucano, certo, perdão, perfume, dormindo, perdida, perguntar e esquerdo.
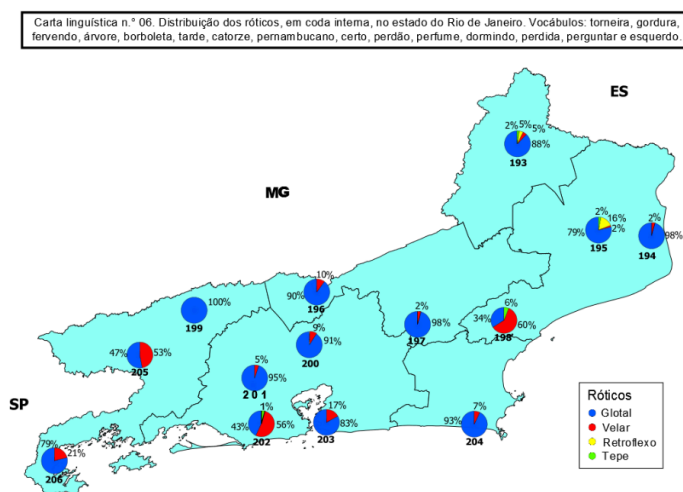
Source: Silva (2016, p. 146).

The retroflex /r/, despite being less productive, was recorded in three out of five points in the state, namely: Alegre (Southern mesoregion), Santa Teresa (Central region), and Barra de São Francisco (Northwest), totaling 11 registrations with percentages of 9%, 7%, and 5%, respectively. These results indicate that the area of occurrence of the retroflex /r/ in Espírito Santo extends from the interior of the state towards Minas Gerais. However, the analyzed data are numerically insufficient to prove whether this is indeed the determinant governing the presence of this variant in Espírito Santo. It is also important to consider that the informant from Santa Teresa lived in the state of Paraná for two years, which may have influenced the use of this rhotic.

In the state of Rio de Janeiro, the rhotic background is also drawn by the glottal /r/, and due mainly to the speech of the capital, it presents the highest usage rate of the velar /r/ in the Southeast region and reveals, albeit sparsely, the occurrence of the tepe and retroflex variants, as illustrated in the linguistic map (Figure 6).

Figure 6 - Distribution of rhotics in internal coda in the state of Rio de Janeiro

Carta linguística n.º 06. Distribuição dos róticos, em coda interna, no estado do Rio de Janeiro. Vocábulos: torneira, gordura, fervendo, árvore, borboleta, tarde, catorze, pernambucano, certo, perdão, perfume, dormindo, perdida, perguntar e esquerdo.

Source: Silva (2016, p. 166).

Figure 6 shows that the *caipira* /r/ is realized only in Itaperuna (point 193), a city located in the Northwest of the state, and in Campos dos Goytacazes (point 195), situated in the Northern region, reaching percentages of 5% and 16%, respectively. The diatopic distribution of this variant

in the state is corroborated by Brandão (2017), who, analyzing 1.845 occurrences of /r/ in medial context in 13 communities in the North and Northwest regions of RJ, recorded, among other rhotics, 13% of retroflex /r/. According to the author, sociogeographic factors such as older age groups and rural traits influence the realization of the [+ant] variants.

Given the appropriate proportions between the corpora, our results also resemble those recorded by the *Atlas Fonético do Entorno da Baía de Guanabara – AFeBG* (Lima, 2006), which provides 30 data points with /r/ in internal coda, considering that in this research we obtained 77% of the glottal variant, 22% of the velar, 1.5% of the retroflex, and 0.5% of the tepe (cf. Graph 2); whereas in the AFeBG data (2006), the percentages are respectively 77%, 18%, 0.3%, and 3%.

Regarding specifically the retroflex /r/, the AFeBG (2006) records two occurrences in Magé, a city located in the metropolitan region of Rio de Janeiro. In the *Micro AFERJ* (2008), on the other hand, the realization of [ɻ] is concentrated in Resende, a city in the southern region of Rio de Janeiro. Nascimento (2009), analyzing 5,972 rhotics collected in the cities of Petrópolis, Itaperuna, and Parati, records only four cases in the latter location. As can be seen, the results regarding the retroflex /r/ available to us from RJ are not sufficient to trace areas of concentration of this variant, as its presence is sparse and occurs either in the North and Northwest or in the metropolitan and Southern regions.

CONCLUSÃO

The main question raised in this research revolves around the importance of rhotic variation in internal coda, especially the retroflex variant, in delineating the varieties that constitute Brazilian Portuguese, and how linguistic cartography, especially facilitated by the QGIS software, can aid in visualizing and better understanding the data.

Thus, through linguistic maps, some elaborated using Voronoi Polygons methodology and others with pie chart graphics, we demonstrated, based on the analysis of 4.829 data points collected in 80 locations, that the retroflex /r/ and the glottal /r/ form isophonic areas responsible for representing the rhotic reality of speech in the Southeast Region. These elements constitute important and indispensable phonetic elements for delineating the varieties that make up BP, contributing even to the revision (Aguilera 2009, 2020a, 2020b, Romano, 2015) of Nascentes' (1953) proposal regarding the dialectal division of the country.

The results regarding the retroflex /r/, the focus of this study, demonstrated that this rhotic:
i) is present, to a greater or lesser extent, throughout the Southeast Region;
ii) is predominant in the state of São Paulo, except in the capital, covering the border areas with the states of Paraná, Mato Grosso do Sul, and Goiás;
iii) is practically categorical in parts of Southern Minas Gerais and the Triângulo Mineiro;
iv) occurs, albeit sporadically, in the north and northeast of MG; in the south, northwest, and central regions of Espírito Santo; and in the northwest and north of Rio de Janeiro.

The diatopic distribution provided by the linguistic maps, with the prominent presence of the retroflex /r/ in the state of São Paulo, covering the South and the Triângulo Mineiro, may support the conjecture of the origin of this rhotic among the mamelucos, inhabitants of the region of the Village of São de Piratininga, who, through the movement of the *bandeiras Paulistas*, spread it throughout the interior of São Paulo, and later, with the discovery of gold, throughout the state of Minas Gerais.

Furthermore, the individual analysis of the data from each state reveals particularities linked, as could not be otherwise, to socio-historical processes. In the state of São Paulo, for example, we observe the existence of two well-marked areas characterized by rhotic variation: the capital and its metropolitan region, and the western part of the state. In the first case, we attribute this variation to the large number of migrants and immigrants attracted to the capital and its

surroundings due to its diverse nature and economic prominence. In the western region of São Paulo, we link rhotic polymorphism to its historical context of occupation, marked on one hand by migrants, mostly from the Northeast, and on the other hand by immigrants, especially Italians.

In Minas Gerais, the results confirm, among other issues, the maintenance of the retroflex /r/, as the comparison between the data collected by ALiB and those of EALMG (1977) reveals that, even 36 years after the publication of the atlas, this rhotic remains predominant in the Triângulo Mineiro and in the southern part of the state, contrary to the pessimistic prediction of Amaral (1920). The geographical distribution of this rhotic in the state of Minas Gerais also confirms that this allophone would have been spread by the bandeirantes, given its proximity to the state of São Paulo. In Rio de Janeiro, where the norm is the glottal variant, the *caipira* /r/ is present in a less significant manner, in the northwest and north regions, starting from the center towards the state of Minas Gerais, as previously documented by Brandão (2017) and by the maps of AFeBG (Lima, 2006).

In the state of Espírito Santo, the presence of the retroflex /r/, although modest, is greater than that recorded in RJ. However, it is not concentrated, as it occurs in the northwest, south, and central regions. The results of these last two states, regarding the retroflex variant, call for more comprehensive research aimed at explaining its scattered distribution, as well as the factors that influence it.

Although further complementary and in-depth studies of other areas are needed, given the extent and socio-historical complexity of the regions that make up Brazil, we believe we have achieved the objectives set forth in this study. Furthermore, we believe that its results can contribute to solving important issues for all those who seek, through Dialectology and linguistic cartography, to research, understand, and - why not - teach the Portuguese language on more solid foundations that truly reflect Brazilian Portuguese.

References

AGUILERA, Vanderci de Andrade. A distribuição dos róticos em coda silábica nos dados do Atlas linguístico do Brasil PR: um estudo geo-sociolinguístico. *In:* ARAGÃO, Maria do Socorro Silva de. (org.). *Estudos em Fonética e Fonologia no Brasil.* Goiânia: GT de Fonética e Fonologia da ANPOLL, 2008. p. 1-14.

AGUILERA, Vanderci de Andrade. O /r/ caipira está ganhando status? O que dizem os dados do Atlas Linguístico do Brasil coletados no Paraná. *Papéis*, v. 16, p. 13-26, 2012.

AGUILERA, Vanderci de Andrade. Léxico e áreas dialetais: o que podem demonstrar os dados do ALIB. *In:* HORA, Dermeval da (org.). *ANAIS - VI Congresso Internacional da Abralin.* v. 2. João Pessoa: Ideia, 2009. p. 4219-4233.

AGUILERA, Vanderci de Andrade. Aspectos fonéticos do português brasileiro e as diferentes possibilidades de delimitação de áreas dialetais: os róticos em coda silábica. *In:* MOTA, Jacyra Andrade *et al.* (org.). *Contribuições de estudos geolinguísticos para o português brasileiro*: uma homenagem a Suzana Cardoso. Salvador: EDUFBA, 2020b. p. 253-276.

AGUILERA, Vanderci de Andrade. Da divisão dialetal de Nascentes ao Atlas Linguístico do Brasil: os róticos em coda silábica interna nos subfalares do Sul do Brasil. *In:* RAZKY, Abdelhak; SFAR, Inès; SOUTET, Olivier; MEJRI, Salah (org.). *A variação nas línguas*: universais compartilhados e idiomaticidade dinâmica/ De la variation dans les langues: universaux partagés et idiomaticité dynamique. v. 1. Araraquara: Letraria, 2020a. p. 89-110.

AGUILERA, Vanderci de Andrade; SILVA, Hélen Cristina da. Dois momentos do /r/ retroflexo em Lavras – MG: no Atlas linguístico de Minas Gerais e nos dados do projeto do Atlas Linguístico do Brasil. *Revista Diadorim* – Revista de Estudos Linguísticos e Literários, n. 8, p. 125-142, 2011.

ALMEIDA. Fabiana da Silva Campos. *Micro Atlas fonético do Estado do Rio de Janeiro (Micro AFERJ):* uma contribuição para o conhecimento dos falares fluminenses. 2008. v. 1. Tese (Doutorado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008. 2 v.

ALMEIDA. Fabiana da Silva Campos. *Estudo dos róticos na Região Centro-Oeste*. 2018. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2018.

AMARAL, Amadeu. *O dialeto caipira.* 3. ed. São Paulo: HUCITEC; Secretaria da Cultura, Ciência e Tecnologia do Estado de São Paulo, [1920] 1982.

BAJUELOS, Antonio L. Problemas de Proximidade: Diagramas de Voronoi. *In:* BAJUELOS, Antonio L. *Geometria Computacional 2007/2008*. Aveiro: Department of Mathematics, University of Aveiro, 14 dez. 2007. Disponível em: http://sweet.ua.pt/leslie/Geocomp/Slides/GC_09_10_7_Diagramas_Voronoi.pdf. Acesso em: 24 set. de 2023.

BOTASSINI, Jacqueline Ortelan Maia. Crenças e atitudes linguísticas quanto ao uso de róticos. *Signum:* Estudos da Linguagem, v. 12, n. 1, p. 85-102, 2009.

BRANDÃO, Silvia Figueiredo. *A geografia lingüística no Brasil.* São Paulo: Ática, 1991.

BRANDÃO, Silvia Figueiredo. O /R/ implosivo no Norte do Estado do Rio de Janeiro. *In:* PEREIRA, Cilene C.; PEREIRA, Paulo R. D. (org.). *Miscelânea de estudos lingüísticos e literários in memoriam Celso Cunha*. Rio de Janeiro: Nova Fronteira, 1995. p. 49-58.

BRANDÃO, Silvia Figueiredo. Aspectos sociolingüísticos de um dialeto rural. *In:* HORA, Dermeval da (Org). *Diversidade lingüística no Brasil.* João Pessoa: Idéia, 1997. p. 61-69.

BRANDÃO, Silvia Figueiredo. Nas trilhas do –R retroflexo. *Signum: Estudos da Linguagem,* v. 10, n. 2, p. 265-283, 2007.

BRANDÃO, Silvia Figueiredo. Variação em coda silábica na fala popular fluminense. *Revista da Abralin*, v. 7, n. 1, p. 177-189, 2017.

CALLOU, Dinah *et al.* Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. *In:* KOCH, Ingedore Villaça (org.) *Gramática do Português falado:* desenvolvimentos. v. 6. Campinas: Unicamp, 1997. p. 465-493.

CARDOSO, Suzana Alice Marcelino da Silva. *Geolinguística*. Tradição e modernidade. São Paulo: Parábola Editorial, 2010.

CARDOSO, Suzana Alice Marcelino da Silva *et al. Atlas Linguístico do Brasil:* cartas Linguísticas. v. 2. Londrina: EDUEL, 2014.

CASTRO, Vandersí Sant' Ana. *A resistência de traços do dialeto caipira*: estudo com base em atlas Lingüísticos regionais brasileiros. 2006. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Campinas, Campinas, 2006.

CASTRO, Vandersí Sant' Ana. *O r "caipira" em São Paulo:* estudo com base em dados do Atlas linguístico do Brasil (ALiB). Comunicação oral apresentada no SILEL, Uberlândia, 17-19 nov. 2009.

CASTRO, Vandersí Sant' Ana. *O "r caipira" em localidades do norte do Estado de São Paulo*: estudo com base em dados do Atlas linguístico do Brasil (ALiB). Comunicação oral apresentada no GEL, 07-08 jul. 2011.

CASTRO, Vandersí Sant' Ana. O 'r caipira" em Mato Grosso do Sul – estudo baseado em dados do Atlas linguístico do Mato Grosso do Sul (ALMS). *Estudos linguísticos*, São Paulo, v. 42, n. 1, 2013a, p. 566 -575.

CASTRO, Vandersí Sant' Ana. O "r caipira" em localidades paulistas do Vale do Paraíba e do litoral - estudo com base em dados do Atlas linguístico do Brasil (ALiB). *In*: BORBA, Lilian do Rocio; LEITE, Cândida Mara Brito (org.). *Diálogos entre língua, cultura e sociedade*. Campinas: Mercado de Letras, 2013b. p. 99-115.

NATIONAL COMMITTEE OF ALiB PROJECT. COMITÊ NACIONAL DO PROJETO ALiB (Brasil). *Atlas lingüístico do Brasil:* questionários 2001. Londrina: EDUEL, 2001.

CUBA, Marigilda Antônio. *Atlas linguístico topodinâmico do Território Incaracterístico*. 2015. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2015.

DIAS, Marcelo Pires. *Atlas geossociolinguístico do Nordeste do Pará* (AGQUINPA). 2017. Tese (Doutorado em Letras) – Universidade Federal do Pará, Instituto de Letras e Comunicação, Belém, 2017.

GUIOTTI, Luciana. Prudente. *O Estudo da Variante Retroflexa na Comunidade de São José do Rio Preto*. 2002. Dissertação (Mestrado em Estudos Linguísticos) – Universidade do Estado de São Paulo, São José do Rio Preto, 2002.

HEAD, Franklin Brian. Propriedades fonéticas e generalidades de processos fonológicos: o caso do "R Caipira". *Cadernos de Estudos Lingüísticos*, n. 13, p. 5-39, 1987.

IBGE. Instituto Brasileiro de Geografia e Estatística. *Donwloads*. Rio de Janeiro: IBGE, 2023. Disponível em: http://downloads.ibge.gov.br/downloads_geociencias.htm. Acesso em: 24 set. 2023.

LEITE, Cândida Mara Brito. *Atitudes linguísticas:* a variante retroflexa em foco. 2004. Dissertação (Mestrado em Linguística) – Universidade Estadual de Campinas, Campinas, 2004.

LEITE, Cândida Mara Brito. *O /R/ em posição de coda silábica no falar campineiro*. 2010. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas. Instituto de Estudos da Linguagem, 2010.

LIMA, Luciana Gomes de. *Atlas Fonético do entorno da Baía de Guanabara-AFeBG*. 2006. Dissertação (Mestrado em Língua Portuguesa) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006. 2 v.

MOUTINHO, Lurdes de Castro *et al.* Análise dialectométrica e cartográfica da variação linguística. *In:* MOUTINHO, Lurdes de Castro (org.). *Estudos em variação linguística nas línguas românicas*. Aveiro: UA, 2019. p. 249-256.

NASCENTES, Antenor. *O linguajar carioca*. 2. ed. Rio de Janeiro: Simões, [1922] 1953.

NASCIMENTO, Tiana Andreza Melo do. *Os róticos na fala de três municípios fluminenses: Petrópolis, Itaperuna e Parati*. 2009. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

QGIS DEVELOPMENT TEAM. *QGIS Geographic Information System.* Open Source Geospatial Foundation Project. Chicago-IL, 2008. Disponível em: https://qgis.org/pt_BR/site/index.html. Acesso em: 24 set. 2023.

RIBEIRO, José *et al. Esboço de um Atlas Lingüístico de Minas Gerais.* Rio de Janeiro: Fundação Casa de Rui Barbosa, 1977.

ROMANO, Valter Pereira. do *Em busca de falares a partir de áreas lexicais no CentroSul Brasil.* 2015. 2 v. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2015.

SANTIAGO-ALMEIDA. Sobre o Falar Cuiabano e o Dialeto Caipira. *In:* ALTINO, Fabiane Cristina (org.). *Múltiplos olhares sobre a diversidade linguística*: uma homenagem à Vanderci de Andrade Aguilera. Londrina: Midiograf, 2012. p. 272-286.

SANTOS-IKEUCHI, Ariane dos. *Atlas Linguístico Topodinâmico do Estado de São Paulo.* 2014. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2014.

SILVA, Hélen Cristina da. *O /r/ caipira no Triângulo Mineiro*: um estudo dialetológico e de atitudes linguísticas. 2012. Dissertação (Mestrado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2012.

SILVA, Hélen Cristina da; AGUILERA, Vanderci de Andrade. Uma nova configuração do caipira: ecos do /r/ retroflexo. In.: *Revista da Abralin* – Associação Brasileira de Linguística, v. 14. n. 1, p 171-194, 2015.

SILVA, Helen Cristina. da. *Pelas veredas do /r/ retroflexo.* 2016. Tese (Doutorado em Estudos da Linguagem) – Universidade Estadual de Londrina, Londrina, 2016.

TELES, Ana Regina; RIBEIRO, Silvana Soares Costa. Apresentando a cartografia aos linguistas: o Projeto ALiB. *In:* MOTA, Jacyra Andrade; CARDOSO, Suzana Alice Marcelino da Silva. *Documentos 2:* Projeto Atlas Linguístico do Brasil. Salvador: Quarteto, 2006, p. 207-228.

THUN, Harald. La géographie linguistique romane à la fin du XX siècle. *In:* CONGRÈS INTERNATIONAL DE LINGUISTIQUE ET DE PHILOLOGIE ROMANES, 22., 23-29 juillet 1998, Bruxelles. *Actes* […]. Tübingen: Niemayer, 1998. p. 367-409.

ZÁGARI, Mário Roberto Lobúglio. Os Falares Mineiros: Esboço de um Atlas Lingüístico de Minas Gerais. *In:* AGUILERA, Vanderci de Andrade (org.). *A Geolingüística no Brasil* - trilhas seguidas, caminhos a percorrer. Londrina: EDUEL, 2005. p. 45-72.