

Notes on the lexical continuum in Amazonian Portuguese: a preliminary discussion

Greize Alves da SILVA*
Romário Duarte SANCHES**

* PhD in Language Studies from the State University of Londrina (2018). Associate Professor I - UFT. Contact: greize_silva@uft.edu.br.

** PhD in Letters (Linguistic Studies) from the Federal University of Pará (2020). Adjunct Professor II - UNIFAP. Contact: romario.duarte@unifap.br

Abstract:

Based on the concepts of dialect continuum (Berruto, 2010; Trudgill; Hernández Campoy, 2007) and urbanization continuum (Bortoni-Ricardo, 2004), this article proposes the concept of lexical continuum based on data from the Portuguese spoken in the states of Amapá and Tocantins, both located in the Brazilian Amazon region. This is an initial proposal based mainly on the IBGE's new classification of urban, rural and nature areas. Based on this recategorization, which we consider more appropriate for delimiting Amazonian varieties, we identified lexical traits (continuous and discontinuous) based on the following studies: *Linguistic Atlas of Amapá* (Razky, Ribeiro, Sanches, 2017), *Lexical Mapping of the Portuguese spoken by the Wajãpi in the State of Amapá* (Rodrigues, 2017), *Linguistic Atlas of the Karipuna of Amapá* (Sanches, 2020) and *Topodynamic and Topostatic Linguistic Atlas of the State of Tocantins* (Silva, 2018). Specifically, we analyzed the designations for the marsupial mammal known as *gambá* (diffuse lexicon) or as *mucura* (focused lexicon).

Keywords: Lexical continuum; Amazonian Portuguese; Spaces of nature; Diffuse lexicon; Focused lexicon.

Resumo

Com base nas concepções de contínuo dialetal (Berruto, 2010; Trudgill; Hernández Campoy, 2007) e de contínuo de urbanização (Bortoni-Ricardo, 2004), propomos neste artigo o conceito de contínuo lexical a partir de dados do português falado nos estados do Amapá e Tocantins, ambos situados na região amazônica brasileira. Trata-se de uma proposta inicial que se fundamenta principalmente na nova classificação do IBGE entre áreas urbanas, rurais e de

natureza. A partir dessa recategorização, que consideramos mais adequada para delimitar as variedades amazônicas, identificamos traços lexicais (contínuos e descontínuos) com base nos seguintes estudos: *Atlas Linguístico do Amapá* (Razky; Ribeiro; Sanches, 2017), *Mapeamento Lexical do Português falado pelos Wajãpi no Estado do Amapá* (Rodrigues, 2017), *Atlas Linguístico dos Karipuna do Amapá* (Sanches, 2020) e *Atlas Linguístico Topodinâmico e Topoestático do Estado de Tocantins* (Silva, 2018). Especificamente, analisamos as designações para o mamífero marsupial conhecido como *gambá* (léxico difuso) ou como *mucura* (léxico focalizado).

Palavras-chave: Contínuo lexical; português amazônico; espaços da natureza; léxico difuso; léxico focalizado.

Signum: Estudos da Linguagem, Londrina, v.27, i.3, p.32-48, december. 2024
Received on: 28/06/2024
Accepted on: 03/12/2024

Notes on the lexical continuum in Amazonian Portuguese: a preliminary discussion

Greize Alves da Silva
Romário Duarte Sanches

INTRODUCTION

Amapá and Tocantins comprise, along with seven other federal units, the Brazilian Legal Amazon - a macro-region made up of Acre, Amapá, Amazonas, Pará, Rondônia, Roraima, Tocantins and part of Mato Grosso and Maranhão, with a population of approximately 30 million people, corresponding to an extensive 59% of the Brazilian territory (IBGE, 2013) -, sharing various socio-cultural and, therefore, linguistic characteristics with these states. An important aspect of these similarities is that the territorial breadth of the Legal Amazon has a low population density - only 5.5 inhabitants per square kilometer, which does not mean that linguistic research in the region is easy. The areas belonging to the Legal Amazon include most of the indigenous ethnic groups on Brazilian soil, a multitude of riverside communities and quilombolas, as well as bordering seven South American countries. In addition, historically, the Amazon region has always welcomed internal migratory flows, with waves of migrants coming from other Brazilian regions, usually linked to some specific economic activity (rubber, mining) and government incentives for populating and exploiting the Amazon.

Even more recently, there has been a new trend of transnational displacement, mainly Venezuelan, Colombian and Haitian. In fact, linguistic description work, for example, in these locations, encompasses different aspects of the languages in contact: Portuguese and indigenous languages, Afro-Brazilian languages, immigration languages (Haitians, Venezuelans), co-official languages (Nheengatu, Baniwa and Tukano, Akwê Xerente), border linguistic contacts (with neighboring countries), inter-varietal contacts of Portuguese (between speakers of regional varieties of Portuguese) (Altenhofen, 2013). Historical and ethnic mixtures translate into a complex reality from an analytical-descriptive point of view, when the researcher observes linguistic dynamics from different perspectives, be they phonetic, lexical, morphosyntactic, pragmatic or discursive.

In the case of the more recently constituted states - such as Amapá and Tocantins - these issues need further investigation. Tocantins is marked by cultural heterogeneity, the result of colonization, its geographical location and the intense migrations that have taken place over the centuries. The construction of its linguistic and cultural identity has its roots in the building and invention of a territory (Brito, 2016). At different times, it was circumscribed by the presence of indigenous people, Africans and miners in search of gold, factors which describe Tocantins' culture as hybrid, "the result of ethnic associations and cultural universes produced historically in different economic contexts and political conjunctures" (Cavalcante; Kimura, 2008, p. 97). Tocantins' sociocultural composition is a factor that organizes the state's linguistic characteristics in the context of the Legal Amazon, just as it does in others, but resulting in unique traits

In Amapá, the process of territory formation shares similarities with other states in the Legal Amazon, especially regarding interactions between various indigenous ethnic groups and traditional communities and migrants. In addition, the significant influence of allochthonous

elements is evident, with economic activities driving migration, mainly from Pará and Maranhão. For Nunes Filho (2009), the invasions and disputes over the territory of Amapá are common features of the roots of Amazonian formation and modulate the linguistic aspects of the region. In short, the linguistic features of Amapá and Tocantins are part of a macro context (of the macro-region to which they belong), both historically and culturally; at the same time, their particularities demarcate the linguistic composition of each. These are similarities and differences that characterize Brazilian dialect formation and are therefore relevant to study in order to trace the characteristics of Brazilian Portuguese. In view of the above, this article sets out to investigate the names given to the marsupial mammal of the *Didelphidae* family, called the 'gambá' or 'mucura', found mainly in the Americas and known for releasing a strong-smelling liquid when it feels threatened. Studying the names of natural elements, such as animals from the local fauna (like the opossum), allows us to understand the basis on which the linguistic characteristics of a given territory are organized, which is why the forementioned nomenclature was chosen.

We start from the hypothesis that, in the northern, Amazonian localities due to their history of formation and population dispersion (within a vast territory), Tocantins and Amapá present characteristics of a Portuguese with a more rural tendency - that is, one that combines both variants with a more rural propensity and dialectal forms more typical of the urban area, tending towards linguistic hybridization. Here, we will work on the notion of continuum, that is, linguistic gradience, more specifically lexical gradience, whose variants oscillate gradually, going from the more gradual forms (common traits) to the other end of the spectrum, the discontinuous, local traits. Studies like this offer an initial discussion on the trends identified based on the recent IBGE proposal, which includes, in addition to the urban and rural categories, the proposition of natural spaces, which is one of the justifications for this undertaking. We would also like to point out that the choice of Tocantins and Amapá to make up the analytical *corpus* is based on their historicity, since these are states considered to be young, recently founded. Therefore, in the next topic, we will list some of the historical elements that surround the Legal Amazon, Tocantins and Amapá in terms of migration, which increases linguistic contact in the region - a crucial aspect for delineating inter and intravarietal variation.

1 MIGRATION AND LANGUAGE CONTACT IN THE AMAZON

In general terms, in Brazil, there is a kind of 'migratory culture', resulting from the perception that social inflexibility is so great that in order to achieve a better and more dignified life, one must migrate, or better still, it was only possible through migration, leaving one's place of origin and trying life in another socio-spatial context (Brito, 2000, p. 19). It's the projection of leaving absolute nothingness to at least seek a possibility of improving one's life, even if it's minimal. In addition, migratory flows are part of Brazilian history and culture in different time periods and are motivated by different objectives on the part of those who move, whether social, economic, cultural, political or environmental, all of them linked to a common denominator: opportunities not offered in the original space. In this sense, migrations transcend exclusively geographical aspects and are not limited to individual choices; they are fundamentally social movements in which groups move to a new reality, carrying with them their beliefs, values, culture and language. Specifically in the Amazon context, the focus of this study, until the 1980s, the region went through three major phases of demographic and economic occupation (Brasil, 1997).

The first consisted of the entry into the territory by the Portuguese who, with a view to defending it from foreigners, established a partial and dispersed settlement along the region's many rivers (Brasil, 1997). The second phase (Brazil's imperial period) was more extractive in nature, with the region serving as an economic base for collecting species of fauna and flora, the so-called "drugs of the sertão" cycle (Brasil, 1997, p. 61). Also, during this phase, between 1879 and 1920, the first Rubber Cycle began, followed by other extractive activities (such as chestnuts), bringing

many northeasterners from the poorest regions of the country in search of better living conditions. It was during this second period that cities like Manaus (Amazonas) and Belém (Pará) achieved great economic growth (Brasil, 1997). The third phase concerns the Brazilian government's intervention in the settlement and development of the Amazon, starting in the 1960s, with various programs, such as the "National Integration Program"¹ (Brasil, 1997); since then, the demographic scenario has changed drastically, with a vertiginous increase in the population from "1.9 million in 1960 to almost 6 million inhabitants in 1980" (Brasil, 1997, p. 62).

The migratory fronts were predominantly northeasterners and southerners, with the former accounting for a significant 35.4% of the overall total in the 1970s, with the state of Pará being the largest recipient of northeastern flows (55%), followed by Mato Grosso and Rondônia, states covered by the public policy of encouraging the use of these territorial spaces (Jakob, 2015, p. 251). Of the states belonging to the Northeast region that sent the most migrants, Maranhão (58%), Ceará (18%) and Bahia (11%) stand out (Brasil, 1997, p. 67).² Recently, the Amazon region has experienced a new stage of displacement with the entry of cross-border migrants, who have left their countries of origin as a result of economic, political and humanitarian crises. The northern states have been the gateway for these migrants, through twin cities, such as the entry of Venezuelans through Pacaraima (RR) or Colombians through Tabatinga (AM). It is important to note that the North is home to nine twin cities with neighboring countries, involving at least five states and with different identities, whose border can be dry, i.e. without a water barrier, or is demarcated by a river barrier with a bridge³

Specifically about the two federal units selected here for analysis - Amapá and Tocantins - only the former shares borders with other countries, Suriname and French Guiana - the latter borders Oiapoque (AP), considered to be the northernmost city in the country and there is a dynamic of cooperation and interdependence between the two cities, sharing resources, infrastructure and economic relations. With regard to Tocantins, its location in the center of Brazil hinders the entry or settlement of cross-border migrants, however, it has historically facilitated (and still favors) the marked presence of interstate migratory fronts, especially from neighboring states: Maranhão, Goiás and Pará, and population movements within its own territory, mainly along the BR-153 highway that cuts through the state in a south/north direction (Silva; Borges, 2019). Created on October 5, 1988, it is considered the most recent state in the Brazilian Federation. Its dismemberment from Goiás was mainly due to political forces, articulated by separatist movements in the territory.

The distances and spreads typical of the sertão meant that the people of Tocantins developed diverse cultural characteristics, especially in relation to religious festivities and the worship of patron saints, since in the "vicinal coexistence that was limited to the cowboys of the same area, what prevailed was the isolation of the sertanejo nuclei, each structured autarchically and turned in on itself, in the immensity of the sertões" (Ribeiro, 1995, p. 343). Brasil (1997) points out that in the 1970s and 1980s, almost 70% of the migratory flow to Tocantins was carried out by northeasterners, mainly people from Maranhão.⁴

¹ This was a government initiative whose main objective was to promote the economic and social development of the country's most remote and least developed regions, especially in the Legal Amazon and the Northeast.

² Also in this decade (1970s), the Center-West region was also responsible for a significant flow of migrants to the Amazon, 21%, especially those coming from Goiás. The Southeast accounted for 14% of the migration rate. In the following decade, the 1980s, the number of migrants entering the Amazon went from 850,000 to 1.3 million people, with Pará still the state that received the most migrants (55%). However, the states of Roraima (29%) and Tocantins (18%) also came into this scenario as attractors of these nuclei, Tocantins being the last federal unit created, after the 1988 Constitution, from the dismemberment of the territory of Goiás (Midwest region), arousing migratory interest, in search of land or public jobs in the newly created state.

³ Jacob (2015) points out that in data obtained from the 2005-2010 censuses, Bolivia was the country that sent the most migrants to the Legal Amazon (16%), followed by Peru (15.4%). However, when looking at the international migratory flow in recent decades, specifically between 2011 and 2020, the largest movements came from Venezuela, Haiti, Bolivia and Colombia and the number of refugees recognized annually in the country increased from 86 in 2011 to 26,500 in 2020 (Número [...], 2021).

⁴ In the following decades, between 2005 and 2010, the Midwest was the region that sent the most migrants to Tocantins (17.6%), especially those from Goiás (12.1%). The Brazilian Northeast accounted for 13% of the migratory flow in the Censuses, especially

Thinking about the second producer of the analytical *corpus* of this research, Amapá was one of the last Brazilian states that had its legal incorporation late; created in 1943, it was considered a territory of Pará until 1988, when it achieved the status of an autonomous federal state. According to IBGE (2023) data, the state has 774,268 inhabitants, spread over 16 municipalities. Amapá's territory is 142,828.521 km², bounded by the state of Pará to the west and south; by French Guiana to the north; by the Atlantic Ocean to the east; and by Suriname to the northwest. In the colonial period, because of invasions and land disputes, Portugal began (in the 18th century) to build forts and fortresses with the aim of occupying the region at various points where the state of Amapá is located today. Nunes Filho (2009) adds that in order to start the process of settlement and fortification in the lands of the Northern Cape, men were needed to work and occupy the territory.

This is how the then Governor Mendonça Furtado brought some families (settlers) from the Azores islands to the region, with the aim of starting a small settlement and building huts to house the soldiers who were guarding the area. After the colonial period, another factor that boosted the recognition of Amapá was its legitimization as a federal territory, as a result of territorial and diplomatic disputes (especially the gold dispute in the region) between France and Brazil. Over the years, Amapá underwent another transformation: in 1988, through the Brazilian Constitution, promulgated at the time, it was elevated to statehood. According to Andrade (2005), this transformation made it possible for new job opportunities to be offered, mainly in construction, which influenced the migration process to the state. The author also points out that in 1990 the migratory dynamic was consolidated in a significant way, as Amapá received around 42,000 people from other units of the federation, with 58% (31,009) coming from the state of Pará and 13.98% (5,973) from Maranhão. Prior to this period, due to the Amapá government's incentives for farmers who had already settled there, hundreds of families from other states arrived in Amapá, increasing the migratory flow.⁵

By observing the dynamics of migration and language in the Legal Amazon, it is possible to identify a significant parallel between these phenomena. Migration, motivated by the search for better opportunities and living conditions, reflects the need to leave an environment to find new possibilities in different socio-spatial contexts. As well as transforming the regions of destination, this movement also brings with it cultural and linguistic elements, enriching local diversity. In the Amazonian context, for example, migration has contributed to economic and social development, while simultaneously introducing new forms of expression and communication, shaping a plural and dynamic linguistic landscape.

2 THE DIALECTAL AND URBANIZATION CONTINUUM IN VARIETIES OF BRAZILIAN PORTUGUESE

In dialectal studies, the varietal landscape is often seen as a continuum of linguistic varieties that overlap, sometimes in a constant way, sometimes in a fragmented, discontinuous way. From this perspective, Berruto (2010) points out that there are few distinct or discontinuous linguistic traits among varieties from nearby localities belonging to the same linguistic area. On the other hand, as these locations become more distant, there is a greater varietal difference recorded at the ends of the continuum. The concept of dialectal continuum has been understood based on two dimensions: geographical and social (Trudgill; Hernández Compoy, 2007). The first refers to the common situation in which geographically close dialects - especially traditional rural

the state of Maranhão (7.3%). Intraregional migration, i.e. from the same region, in the case of the North, accounted for 12.1%, with Pará (8.7%) as the center of migration. The Southeast and South contributed just 2.1 and 1.2%, respectively (Ramalho *et al.*, 2016, p.74). As you can see, the biggest states sending migrants to Tocantins are the border states: Goiás, Maranhão and Pará.

⁵ This large influx was due to various factors, such as "the installation of ICOMI (Indústria e Comércio de Minérios S/A), [...] gold mining [...], the creation of ALCMS (Macapá and Santana Free Trade Area) and the actions of the federal government, which promoted social and economic infrastructure works (Andrade, 2005, p. 94).

ones - are minimally distinguishable from each other, although their differences are cumulative, i.e. they increase according to the geographical distance between them. All dialects are intelligible to speakers of nearby dialects, so the greater the distance between them, the greater the difficulties of mutual understanding. If a geographical area is large enough, it is possible that the dialects situated at its extremes will not be mutually intelligible. However, they will be connected by a chain of mutual intelligibility through intermediate varieties (Trudgill; Hernández, Campoy, 2007).

The authors point out that in the chain of the dialectal continuum there is no complete cut that makes the contiguous dialects mutually incomprehensible, since geographical (regional) dialects are not discrete entities, and dialectal boundaries do not exist as defined areas, except as transitional areas within a linguistic continuum, like the colors of the rainbow that gradually change from one to the other. Just as there are geographical dialectal continuums, there are also social dialectal continuums, in which sociolects or dialects vary according to their location, higher or lower on the social scale. In this way, despite being from the same geographical region and sharing common linguistic characteristics, the language of a speaker or a group of speakers can vary from that of other groups within the same community. These linguistic differences or similarities between social groups can be conditioned by socio-demographic factors such as socio-economic class, gender, age, ethnicity, religion, level of literacy, etc. However, they are always gradual, cumulative and never discrete (Bagno, 2017, p. 63).

In the Brazilian context, sociolinguistic research carried out since the second half of the 20th century has shown that the main differentiating factor between sociolects is the level of formal education, i.e. access to schooling and literate culture. Confirming this argument, Bortoni-Ricardo (2004) presents the concept of the urbanization continuum, where Brazilian Portuguese speakers can be situated according to their geographical location and their literacy practices. Thus, at one end of the spectrum would be rural speakers and at the other urban speakers, and between the two ends would be the traits known as rural Portuguese. At one end of the continuum are the rural varieties used by the most geographically isolated communities. At the opposite pole are the urban varieties that have received the greatest influence from language standardization processes such as the press, literary works and, above all, schools. Within these institutions, monitored styles of language are preferably used in both written and oral forms. In the space between the two poles lies a rural area.

According to Bortoni-Ricardo (2004), rural groups are formed within communities, usually in the countryside, where there is a mixture of rural and urban ways of life, culture and economy. Commonly, these speakers preserve much of their rural cultural background, especially in their linguistic repertoire, but they also make use of more gradual forms, submitted through urban influence, either through the media or the absorption of agricultural technology. The author states that if we take the urbanization continuum as a methodology for analysis, we can place any Brazilian Portuguese speaker at a certain point on this continuum, considering the region where they were born and live. For her, the urbanization continuum shows that there are no rigid boundaries separating rural, rural or urban speakers. The boundaries are fluid and there is a lot of overlap between these types of speech - which is why they talk about a continuum. Speech in isolated rural communities can be considered as a highly focused form of rural vernacular, which contains virtually the entire set of non-standard features that define a sharp distinction between rural and urban varieties, as well as a high incidence of diffuse features of popular Portuguese

To the extent that speakers of rural vernaculars come into direct or indirect contact with the standard language, written or oral, their dialect tends to become more diffuse, i.e. the occurrence of typical rural traits decreases and some non-standard rules in their repertoire, previously almost categorical, tend to become variable rules. Diffusion is not necessarily seen as assimilation into standard Portuguese, but rather as a move away from stigmatized rural dialects. This may be due to the speaker's need to be accepted in their new daily reality and, as a result, they tend to adopt speech patterns closer to those used in their new environment.

Based on the postulates of Bortoni-Ricardo (2004), we seek to complement the proposal of a continuum of urbanization by presenting a new proposition of what the author calls rural-urban-urban. To do this, we used the current IBGE (2023) proposal, which classifies Brazil's territorial spaces as rural, urban and natural. In this sense, the category 'nature' has been included to take account of natural spaces (forests, mangroves, rivers) that could not be relegated to the rural universe, thus improving the ability to make a detailed reading of the territory in the task of delimiting rural and urban areas in Brazil (IBGE, 2023, p. 63).

About 'urban space', the Institute used, above all, the classification proposed by Sposito (2006), in which three criteria must be observed in order to interpret what is urban: i) demographic dynamics, ii) social differentiation and iii) the spatial unity of urbanization. Thus, urban is understood as: "Areas with high population densities, buildings and streets where the landscape is intensely altered. Predominantly, people are occupied in secondary and tertiary activities" (IBGE, 2023, p. 64). To define the 'rural space', criteria such as population dispersion in the territory, the influence of the natural environment and occupations were considered, arriving at this proposition: "areas of low population density characterized, in general, by the alteration of the landscape due to anthropic activities linked to agricultural production [...]. In addition, its rural fabric may contain small urbanized nuclei and/or natural fragments" (IBGE, 2023, p. 72). Finally, the third spatial category proposed, 'nature spaces', encompasses areas where natural processes prevail; they embrace human action, if this action is surrounded by and in line with the space.

Thus, nature's space is understood to be: "Areas where natural processes that shape the landscape prevail. When there is human action, it occurs in a dispersed way, according to the meanings attributed by the actors involved, ensuring relative stability of biotic and abiotic factors" (IBGE, 2023, p. 79). Based on this third category, it is understood that analyzing the varieties of Portuguese spoken in the Amazon only through the lens of urban or rural is unfeasible. In the Amazon, there are traditional community languages, such as indigenous ethnic groups, located in areas of nature and, therefore, it would be inappropriate to include these languages in the 'rural' category. Our proposal focuses on the classification between speakers who are inserted in nature-rural-urban-urban spaces. With this, it can be observed that within the urbanization continuum there can be, at one end, the traits of indigenous varieties located in the space of nature, and at the other end there can be the traits of standardized urban varieties, located in the urban space. Between these two ends are the traces of rural and rural-urban varieties located in the rural and rural-urban space.

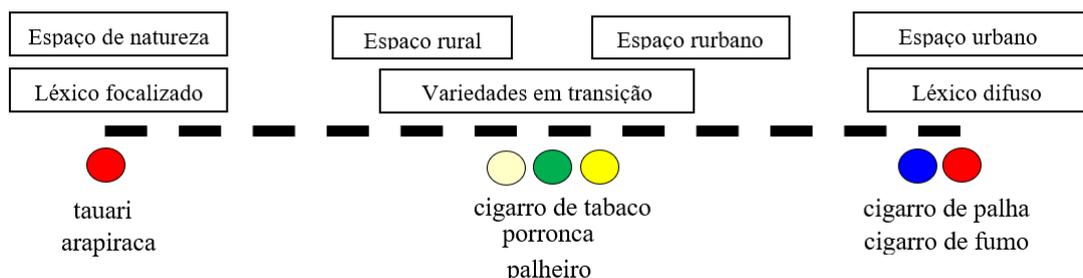
Thus, taking up the concept of the dialectal continuum, we developed the concept of the lexical continuum - which refers to the set of diffuse or focused lexical traits that spread across geographical space under the influence of historical, social, political, migratory, ethnic, media and ecosystemic factors (the interactions between man and the elements of nature). The lexical continuum is categorized with a view to forming spectra of variation that can change according to the phonetic, morphosyntactic or discursive categories of the language. At one end of this continuum we can find diffuse lexical features, i.e. lexical variants that systematically cover a geographical area, coexisting in more than one speech community or social groups with distinct linguistic varieties. At the other end of the spectrum are focused lexical features, i.e. lexical variants with more restricted linguistic use compared to diffuse lexical variants. Typically, the focused lexicon is identified as a non-standard, non-dictionary variant or as a variant entry in lexicographic manuals, while the diffuse lexicon is understood as a standard variant, i.e. a socially prestigious lexical item that is widely taught and disseminated in the institutions that perpetuate literate culture.

In general, the focused lexicon has been found in traditional communities or in areas considered rural and with difficult access to urban areas, while the diffuse lexicon has been found frequently in urban cities and less frequently in rural or nature spaces, because of the schooling process and access to different digital media. An example of this continuous model is the data on

the names of 'cigarro de palha' in Brazilian capitals, contained in the *Linguistic Atlas of Brazil Project* (ALiB). The mapping shows the names given by Brazilians to the cigarette that people used to make by hand. Letter L16 shows the use of diffuse lexical variants, i.e. there is a certain geographical spread of lexemes competing in other regions of Brazil, such as the lexical variants 'cigarro de palha' and 'cigarro de fumo', with a greater distribution in the Northeast, Midwest, Southeast and South and a lower concentration in the North. There are also the variants 'porronca' and 'palheiro', the first of which stands out in the North and the second in the Midwest and South. The data collected in the Brazilian capitals by the ALiB team shows that there was no mention of the '*tauari*' and '*arapiraca*' variants, which are now being referred to as focalized. The former was recorded in the Portuguese spoken by the Karipuna indigenous people of Amapá (Sanches, 2020) and the latter in the Portuguese spoken by the Wajãpi indigenous people of Amapari (Rodrigues, 2017), both ethnic groups located in the state of Amapá. The use of '*tauari*' and '*arapiraca*' by indigenous people from different ethnic groups is conditioned by the constant interactions they have with the ecosystem in which they live. Sanches (2021), for example, comments that *tauari* is a species of tree native to the region from which the Karipuna indigenous people usually remove the periderm to roll tobacco, hence the name '*tauari* cigarette'.

Thus, we can summarize the concept of lexical continuum, for this item in particular, based on the following diagram, Figure 1:

Figure 1 - Lexical continuum for straw cigarette



Source: Prepared by the authors (2024).

At the left end of the continuum are the focused lexicons ('*tauari*' and '*arapiraca*'), found in indigenous communities, hardly found in Portuguese dictionaries and used in urban spaces. At the other end, on the right, are the lexicons belonging to standard Portuguese, usually found in Brazil's urban areas, dictionarized and disseminated as a standard variant.⁶ Finally, 'cigarro de tabaco', 'porronca' and 'palheiro' are in a process of transition: they are in the intermediate zone, between rural and rurano spaces. These variants are partially widespread and can be identified in lexicographic manuals, treated as variant entries, linked to the standard variant 'cigarro de palha'. There are other lexical items that also show the same lexical behavior, such as the names for 'riacho'/córrego' (Razky; Sanches, 2016) and 'fanhoso' (Sanches; Razky, 2021). This evidence allows us to hypothesize that the cities of the Northern region (capitals and inland cities) have, in their lexical norm, traces of varieties in transition (rural and rurano spaces) and, at the same time, a diffuse lexicon, as a result of numerous factors, especially the process of industrialization, migration and late development, which took place after the second half of the 20th century.

⁶ The standard variants are, "broadly speaking, those that belong to the varieties prescribed in the standard norm manuals; the non-standard variants tend to deviate from these varieties" (Coelho *et al.*, 2015, p. 18).

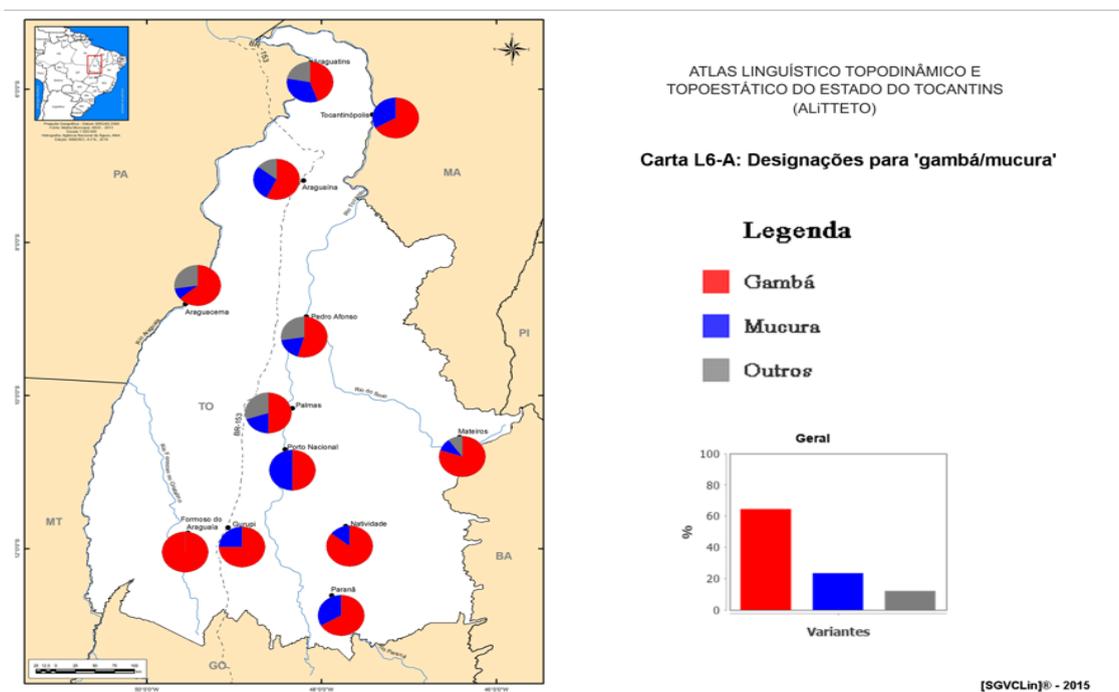
On the other hand, in spaces considered to be natural, where traditional communities are found, the lexical norm is made up of focused lexical traits and, at the same time, a diffuse lexicon. With these considerations in mind, the next section will deal a little more with the lexical continuum in Amazonian Portuguese, specifically the case of 'gambá' (or 'mucura') in the territories of Amapá and Tocantins, to structure the analytical proposal for linguistic variation in these regions and broaden the line of argument regarding the presence of a lexical continuum

3 EVIDENCE OF THE LEXICAL CONTINUUM IN AMAZONIAN PORTUGUESE: TOCANTINS AND AMAPÁ

To demonstrate how lexical variants are arranged in the Amazonian Portuguese continuum, sometimes as a focused lexicon (discontinuous), sometimes as a diffuse lexicon (continuous), this research was based on works that share the same research method, Geolinguistics, i.e. they present the public with lexical data that can be intra- and/or inter-compared. In this case, the lexical item that makes this process possible concerns the names for 'gambá', a common lexicon mapped and described in dialectal research in Tocantins and Amapá. In the case of Tocantins, the data in question relates to the collection carried out in 2015 for the doctoral thesis *Atlas Linguístico Topodinâmico e Topoestático do Tocantins (ALiTTeTO)* (Silva, 2018). Ninety-six informants were interviewed, distributed among 12 localities and grouped according to the following variables: gender, age (two age groups - 18 to 30 years old and between 50 and 65 years old) and type of mobility (topostatic: born and established in the locality being researched; and topodynamic: migrant informants or those coming from displacement, but who have lived in the locality for more than ten years). As is common practice in dialectology, semi-structured questionnaires were administered to all the informants

We therefore worked with the answers obtained through question number 53: What do you call the animal that gives off a bad smell when it feels threatened? In all, 107 answers were obtained, with the following forms, in order of productivity: 'gambá' (64.5%), 'mucura' (23.4%) and other forms totaling 12.1%. Nine informants were unable to answer the question, all of them female. Synoptically, in the linguistic chart published, four of the most productive forms were mapped. However, for the purposes of this article, only the two most latent will be described: 'gambá', representing the diffuse lexicon, the standardized form of Brazilian Portuguese, and 'mucura', representing the focused lexicon, the regional form.

Figure 2 - Lexical mapping for the *gambá* item in Tocantins



Source: Silva (2018, p. 363).

As can be seen in Figure 2, the predominant form 'gambá', marked in red, was produced throughout Tocantins, with occurrences of over 40%; in the city of Formoso do Araguaia, for example, it was categorically obtained. The entry for 'gambá' is dictionarized as a Tupimism in the Houaiss dictionary (2009), but its origin is controversial; the dictionary cites that the description comes from *gã'ba* ('seio oco') or from the Tupi '*gua*' ('seio', 'ventre') + '*ambá*' = ('vazio', 'oco'). 'Gambá' is also predominant in the data from the Linguistic Atlas of Brazil, specifically in the Central-West region (Silva-Costa; Isquerdo, 2012), which borders Tocantins, with percentages above 88%⁷. However, when we look specifically at the more northerly areas, we see that 'gambá' competes with 'mucura', the second variant (23.4%), being recorded throughout Tocantins and, like 'gambá', its areal distribution, present on the L06 chart, was contiguous. Houaiss (2009) presents the term as a synonym for 'opossum' and as a regionalism of the Amazon, information also pointed out by Silva-Costa and Isquerdo (2012) in a study on the variants for the marsupial animal in the Midwest. According to the authors, 'mucura' was collected in Mato Grosso, in the city of Aripuanã, and in Goiás, in Porangatu, pointing "to confirmation of a possible isoglossa⁸ [...] since Aripuanã/MT is in the far north of Mato Grosso, close to the state of Amazonas, and Porangatu/GO borders the state of Tocantins" (Silva-Costa; Isquerdo, 2012, p. 785).

In Guedes (2012), specifically with data from the state of Pará, 'mucura' is predominant and only competes with 'gambá' in two localities close to the state of Tocantins. Therefore, the presence of 'mucura' in our data, corroborated by Guedes (2012) and Silva-Costa and Isquerdo (2012), indicates that it is a northern variant that has possibly spread to other Brazilian regions, although the form 'gambá' remains consistently present. Regarding extralinguistic variables, all the informants interviewed maintained similar rates of variants, except for the second generation (50-65 years old), who cited 'mucura' twice as often as the young people. In addition, young people are the ones who most often recorded the unit 'gambá', understood here as the other end of the continuum, the diffuse trait, the form belonging to the variety of standard Portuguese. We will now present the analysis in the state of Amapá, based on lexical data extracted from three

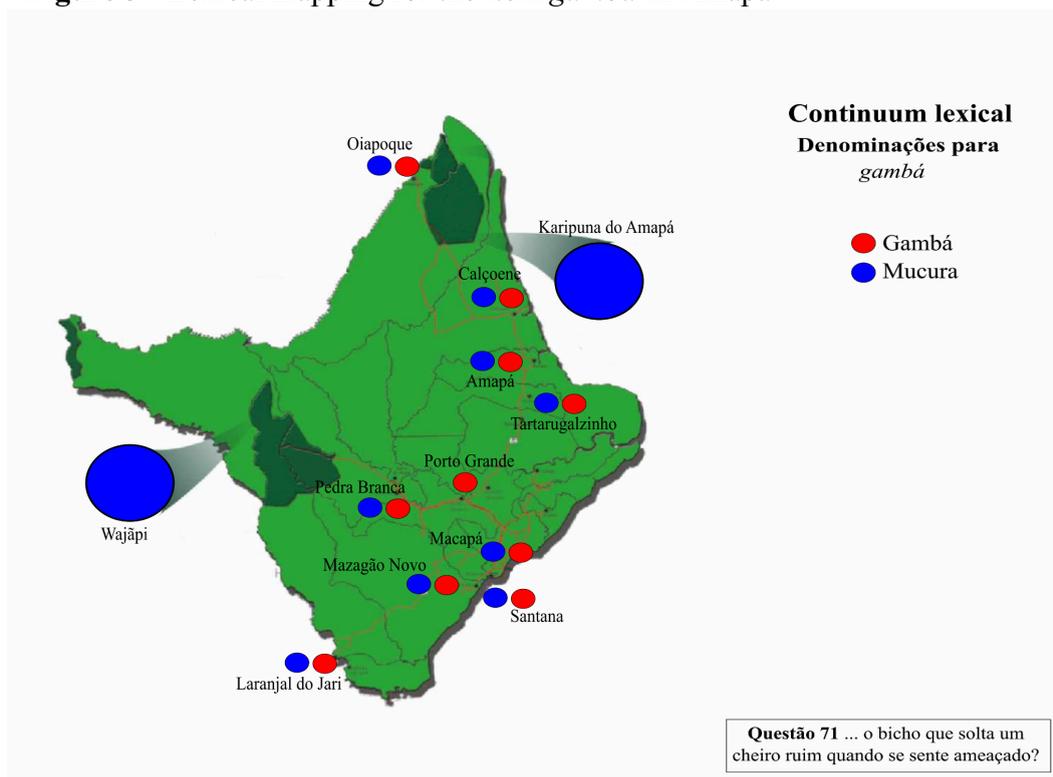
⁷ Silva-Costa and Isquerdo will present the cartography for the item in question for volume 4 of the *Linguistic Atlas of Brazil (ALiB)* containing data from the capitals.

⁸ Imaginary line drawn on a map demarcating geographical areas where different language variants are used.

geolinguistic works on the Portuguese spoken in the state of Amapá: the *Linguistic Atlas of Amapá* (Atlas Linguístico do Amapá - ALAP) (Razky; Ribeiro; Sanches, 2017), the *Lexical Mapping of the Portuguese spoken by the Wajãpi in the State of Amapá* (Rodrigues, 2017) and the *Linguistic Atlas of the Karipuna of Amapá* (Sanches, 2020).

ALAP presents 16 phonetic and 73 lexical maps of the Portuguese spoken in ten locations in the state (Macapá, Santana, Mazagão Velho, Laranjal do Jari, Porto Grande, Pedra Branca do Amapari, Tartarugalzinho, Amapá, Calçoene and Oiapoque). To compile the atlas, 40 Amapá men and women with incomplete primary or secondary education were interviewed between 2011 and 2013, divided into two age groups (A: 18-30 years old; B: 50-75 years old). Regarding the *lexical mapping of Portuguese spoken by the Wajãpi in the state of Amapá*, 20 interviews were carried out in 2016 with indigenous Wajãpi belonging to five villages (Aramirã, Pairakae, CTA, Mariry and Kurani'yty), men and women who were not literate or literate up to the 8th grade of elementary school, aged between 18-30 and 40-70. Finally, for the *Linguistic Atlas of the Karipuna of Amapá*, 36 Karipuna indigenous people (Portuguese and Kheuól speakers) from nine villages were interviewed between 2017 and 2018, with the following profile: men and women from the first generation (18-30 years old) and the second generation (over 50 years old). Based on the lexical data collected in the forementioned works, it was possible to group together information on the use of the variants 'gambá' and 'mucura' present in the Portuguese spoken in Amapá, as shown in Figure 3.

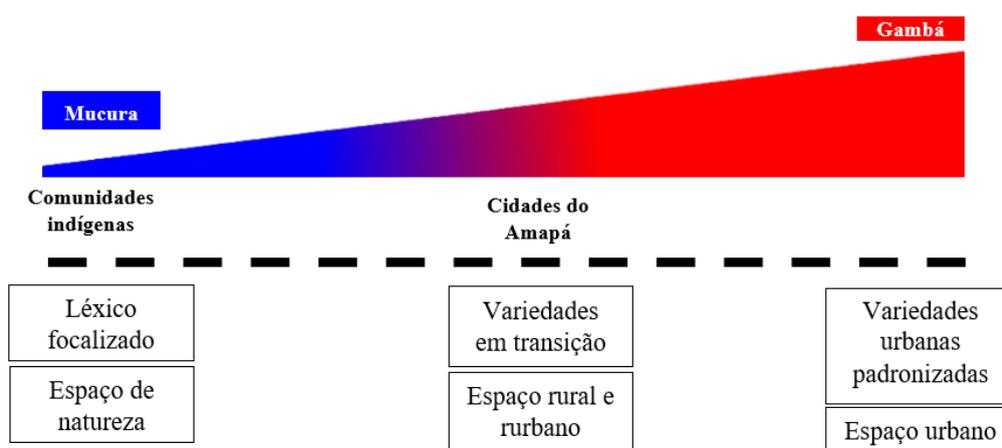
Figure 3 - Lexical mapping for the item *gambá* in Amapá



Source: Prepared by the authors.

To read Figure 3, red was inserted to highlight the use of the 'gambá' variant and blue for the 'mucura' variant. The map also indicates, with smaller circles, the localities surveyed in the ALAP and, with larger circles, the traditional (indigenous) communities. The ALAP data shows that most Amapá residents in urban cities tend to use 'gambá' as the predominant variant and 'mucura' as the secondary variant, the former with 67% and the latter with 33%. In the indigenous communities (Wajãpi and Karipuna of Amapá), the categorical use of 'mucura' stands out, without the presence of the 'gambá' variant. Based on this, the configuration of a lexical continuum for the item *gambá* is as follows, as illustrated in Figure 4:

Figure 4 - Lexical continuum for *gambá*



Source: Prepared by the authors.

At the extreme left of the continuum, the focused lexicon 'mucura' stands out, categorically present in the indigenous communities of Amapá, marking this indigenous Portuguese with discontinuous traits, as there was no record of the 'gambá' variant. In the inland cities and the capital of Amapá, the two variants are present, with a greater focus on the use of the diffuse lexicon that is 'gambá', as a trait of varieties in transition. Our hypothesis consists of the assumption that the indigenous Wajãpi and Karipuna Portuguese-speaking communities of Amapá, because they are further away from urban areas and because there is no constant contact with the standard variety of urban Portuguese, tend to adopt a more focused lexicon, while the people of Amapá who live in rural and rurban areas have increasingly adopted a diffuse and standardized lexicon, especially as a result of the schooling process. This same scenario is repeated in the state of Tocantins, where the 'gambá' variant stands out in relation to the 'mucura' variant. Another important aspect worth highlighting concerns the social group that tends to use 'mucura', both in ALAP and ALiTTETO: this use predominates in the speech of second generation informants, over 50 years old.

This shows that the 'gambá' variant has been used constantly by the first generation (18-30 years old), and that this lexicon has been institutionalized and propagated as a standard variant by the analogue (radio, television, newspapers, magazines, etc.) and digital (smartphone, computer, websites, social networks, etc.) media, especially in school environments with the distribution of textbooks and dictionaries adopted to teach standard Portuguese in Brazilian schools. Based on the initial hypothesis that Amazonian/northern locations have the characteristics of a rurban Portuguese, some considerations are in order based on the elements mobilized here, especially the socio-demographic ones allied to economic activities, which are responsible for providing the tone for migratory movements. The Amazon region has always been marked by low population density, in an environment of large territorial extensions, covering 60% of Brazil's territory, causing its population centers to be marked by dispersion and a certain geospatial isolation. Thus, the more isolated a community is, the greater the possibility of maintaining a focused lexicon, with regional and rural characteristics, as demonstrated by the categorical occurrence of 'mucura' in indigenous communities.

On the other hand, still on the theme of economic activities as a factor driving urbanization, and the consequent diffuse changes in the lexicon, from the second half of the 20th century onwards, government programmes boosted the economy in the Legal Amazon, which attracted waves of migrants to this region in search of new opportunities, creating centers with urban characteristics. Allied to this fact, it is very common for populations to move internally in search of this 'progress', in other words, former residents of rural areas migrate to more developed centers in search of better living conditions. It is in this context that inter-varietal contacts take place, giving rise to a lexicon with a rurban tendency, i.e. a mixture of focused ('mucura') and diffuse ('gambá') traits, as observed in the ALAP and ALiTTETO state atlases.

FINAL CONSIDERATIONS

Based on the short description of the sociological and demographic veins in the context of the Legal Amazon, we started from the hypothesis that the region has the characteristics of a lexical continuum, with a rurban tendency, that is, a tendency towards linguistic hybridization, whose speakers know both ends of the spectrum. The analysis showed that at one end, there is the use of 'gambá', an educated variant, with a normative tendency in Brazil, a form used in textbooks and the media; and at the other end, there is 'mucura', a regional variant, with a strong tendency to be used in rural areas and traditional communities. Some explanations can be given, such as the disorganized growth, especially in the middle of the 20th century, based on government programs,

which meant that the region, which used to be home to less than two million people, tripled its population to six million, swelling the urban centers.

In addition, the mark of 'progress' is internal displacement, i.e. populations that used to live in rural areas moving to centers with higher rates of urbanization. Based on the postulates of Bortoni-Ricardo (2004), we sought to complement the proposal for a continuum of urbanization with the new IBGE (2023) classification, which distinguishes Brazil's territorial spaces into rural, urban and nature. The inclusion of the 'nature' category covers natural spaces such as forests, mangroves and rivers, which cannot simply be categorized as rural, represented here by indigenous communities. Thus, at the end of the continuum, the lexicon 'mucura' stands out, categorically present in the indigenous communities of Amapá, with no records of the variant 'gambá'. In the inland cities and the capital of Amapá, both variants are present, with a greater use of gambá as a diffuse lexicon, indicating varieties in transition. In Tocantins the situation is similar, since both forms are used, although with a predominance of gambá, especially by the younger generation (18 to 30 years old). This research will certainly be expanded with a focus on checking other lexical items to confirm our hypothesis.

Finally, we believe that in the Amazonian context, speakers tend to preserve the focused lexicon in isolated communities or in natural spaces. On the other hand, in the capitals and inland cities of the North, through inter-varietal contacts, speakers tend to preserve a rural lexicon that mixes focused and diffuse traits, as observed in the state atlases. This scenario supports the concept of a "lexical continuum" to understand the synchronic behavior of the lexicon in Amazonian Portuguese, highlighting the strong influence of geographical, historical and social factors. It is also worth mentioning that this model instigates the field of Sociolinguistics and Dialectology in Brazil to revisit pre-established concepts and theories, providing a new perspective for the recent categorization of areas in the Amazon published by the Brazilian Institute of Geography and Statistics

REFERENCES

ALTENHOFEN, Cléo Wilson. Migrações e contatos linguísticos na perspectiva da Geolinguística Pluridimensional e Contatual. *Revista de Letras Norte@mentos*, Sinop, MT, v. 6, n. 12, p. 31-52, jul./dez. 2013. DOI: <https://doi.org/10.30681/rln.v6i12.6876>

ANDRADE, Rosemary Ferreira. *Migração no Amapá: projeção espacial num contexto de crescimento populacional*. Belém: NAEA, 2005.

BAGNO, Marcos. *Dicionário crítico de sociolinguística*. São Paulo: Parábola Editorial, 2017.

BERRUTO, Gaetano. Identifying dimensions of a linguistic variation in a language space. In: AUER, Peter; SCHMIDT, Jürgen Erich (ed.). *Language and space: an international handbook of linguistic variation. Theories and methods*. Berlin: De Gruyter Mouton, 2010. DOI: <https://doi.org/10.1515/9783110220278>

BORTONI-RICARDO, Stella Maris. *Educação em língua materna: a sociolinguística na sala de aula*. São Paulo: Parábola, 2004.

BRASIL, Marília Carvalho. Os fluxos migratórios na região norte nas décadas de 70 e 80: uma análise exploratória. *Caderno de Estudos Sociais*, Recife, v. 13, p. 61-84, jan./jun. 1997.

- BRITO, Fausto. Brasil, final do século: a transição para um novo padrão migratório. *In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS*, 12., 2000, Caxambu, MG. *Anais [...]*. Caxambu: ABEP, 2000. v. 1, p. 1-44.
- CAVALCANTE, Ivana; KIMURA, Simone. Mapeamento do patrimônio cultural do estado do Tocantins. *In: CAVALCANTE, Ivana; KIMURA, Simone (org.). Vivências e sentidos: o patrimônio cultural do Tocantins*. Goiânia: Iphan, 2008. p. 90-159.
- COELHO, Izete L.; GÖRSKI, Edair M.; SOUZA, Cezarina Maria Nobre de; MAY, G. H. *Para conhecer Sociolinguística*. São Paulo: Contexto, 2015.
- GUEDES, Regis José da Cunha. *Estudo geossociolinguístico da variação lexical na zona rural do estado do Pará*. Dissertação (Mestrado em Letras) - Instituto de Letras e Comunicação, Universidade Federal do Pará, Belém, PA, 2012.
- HOUAISS, Antônio. *Dicionário Houaiss eletrônico*. Rio de Janeiro: Instituto Antônio Houaiss: Objetiva, 2009.
- IBGE. Coordenação de Geografia. *Proposta metodológica para classificação dos espaços do rural, do urbano e da natureza no Brasil*. Rio de Janeiro: IBGE, 2023. Disponível em: <https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=2102019>. Acesso em 27 jun. 2024.
- IBGE. Estados. *Tocantins*. Disponível em: <https://cidades.ibge.gov.br/brasil/to/panorama>. Acesso em: 20 out. 2017.
- IBGE. *Metodologia do censo demográfico de 2010*. 2. ed. Rio de Janeiro: IBGE, 2013. Acompanha 1 CD-ROM. (Série relatórios metodológicos, v. 41).
- JAKOB, Alberto Augusto Eichman. A migração internacional recente na Amazônia brasileira. *REMHU: Revista Interdisciplinar da Mobilidade Humana*, Brasília, DF, v. 23, n. 45, p. 249-271, jul./dec. 2015. Disponível em: <https://www.scielo.br/j/remhu/a/JVQM3sggXyBWrrf8jxX4CPh/?lang=pt>. Acesso em: 27 jun. 2024.
- NÚMERO de novos imigrantes cresce 24,4% no Brasil em dez anos. *Agência Brasil*, Brasília, 7 dez. 2021. Geral. Notícia. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2021-12/numero-de-novos-imigrantes-cresce-244-no-brasil-em-dez-anos>. Acesso em: 24 jun. 2024.
- NUNES FILHO, Edinaldo Pinheiro. Formação histórica, econômica, social, política e cultural do Amapá: descrição e análise do processo de formação histórica do Amapá. *In: OLIVEIRA, Augusto; RODRIGUES, Randolfe (org.). Amazônia, Amapá: escritos de história*. Belém: Paka-Tatu, 2009. p. 211-234.
- RAMALHO, Hilton Martins de Brito; FIGUEIREDO, Erik Alencar de; SILVA NETTO JÚNIOR, José Luís da. Determinantes das migrações interestaduais no Brasil: evidências a partir de um modelo gravitacional. *Pesquisa e planejamento econômico*, Rio de Janeiro, RJ, v. 46, n. 1, abr. 2016.
- RAZKY, Abdelhak; RIBEIRO, Celeste Maria da Rocha; SANCHES, Romário Duarte. *Atlas linguístico do Amapá*. São Paulo: Labrador, 2017.

- RAZKY, Abdelhak; SANCHES, Romário Duarte. Variação geossocial do item lexical riacho/córrego nas capitais brasileiras. *Gragoatá*, Niterói, RJ, n. 40, p. 70-89, 2016.
- RIBEIRO, Darcy. *O povo brasileiro: a formação e o sentido do Brasil*. São Paulo: Companhia das Letras, 1995.
- RODRIGUES, Maria Doraci Guedes. *Mapeamento lexical do português falado pelos Wajãpi no estado do Amapá: uma abordagem geossociolinguística*. Dissertação (Mestrado em Letras). Universidade Federal do Pará, 2017. Disponível em: <https://repositorio.ufpa.br/jspui/handle/2011/10235>. Acesso em: 22 jun. 2024.
- SANCHES, Romário Duarte; RAZKY, Abdelhak. Análise geossociolinguística das designações para fanhoso nas capitais brasileiras. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 37, n. 2, p. 1-22, 2021. Disponível em: <https://www.scielo.br/j/delta/a/4NH4NzT46FHNHLXXh8c68Kt/abstract/?lang=pt>. Acesso em: 22 jun. 2024.
- SANCHES, Romário Duarte. *Atlas linguísticos dos Karipuna do Amapá*. Rio Branco: NEPAN, 2020.
- SILVA, Greize Alves da. *Atlas linguístico topodinâmico e topoestático do Estado do Tocantins (ALITTETO)*. 2018. Tese (Doutorado em Estudos da Linguagem) - Universidade Estadual de Londrina, PR, Londrina, 2018. Disponível em: <http://www.bibliotecadigital.uel.br/document/?code=vtls000218332>. Acesso em: 22 jun. 2024.
- SILVA, Greize Alves da; BORGES, Patrícia Andréa. Presença vs ausência de traços de ruralidade no léxico tocantinense. *Revista do Instituto de Estudos Brasileiros*, São Paulo, n. 72, p. 83–105, 2019. Disponível em: <https://www.revistas.usp.br/rieb/article/view/157030>. Acesso em: 27 jun. 2024.
- SILVA, Leonardo Luiz Silveira; MARQUES, Renato de Oliveira. O papel das cidades-gêmeas da região norte brasileira na articulação do comércio exterior. *Caderno de Geografia*, Belo Horizonte, MG, v. 30, n. 60, 2020. Disponível em: <https://periodicos.pucminas.br/index.php/geografia/article/view/20743>. Acesso em: 15 maio 2024.
- SILVA-COSTA, Daniela de Souza; ISQUERDO, Aparecida Negri. Um estudo etnolinguístico de designativos para “gambá” no Brasil Central: contribuições do Projeto ALiB. *Estudos Linguísticos*, São Paulo, v. 41, n. 2, p. 779-792, maio/ago. 2012. Disponível em: <https://revistas.gel.org.br/estudos-linguisticos/article/view/1196>. Acesso em: 24 jun. 2024.
- SPOSITO, Maria Encarnação Beltrão. A questão cidade-campo: perspectivas a partir da cidade. *In: SPOSITO, Maria Encarnação Beltrão; WHITACKER, Arthur Magon (org.). Cidade e campo: relações e contradições entre urbano e rural*. São Paulo: Expressão Popular, 2006. v. 2, p. 111-130. (Geografia em movimento).
- TRUDGILL, Peter; HERNÁNDEZ COMPOY, Juan M. *Diccionario de sociolingüística*. Madrid: Gredos Editorial, 2007.