

Portuguese spoken in São Paulo State: Urban life lexies



Selmo Ribeiro Figueiredo Junior (Charles University, Prague)

ABSTRACT

This study unveils insights into the spatial and social distribution of Portuguese lexical forms, or lexies, pertaining to urban life in the interior of São Paulo State. These lexies were collected through oral responses from eighty participants, elicited by nine questions addressing variants for terms such as ‘traffic lights’, ‘speed bump’, ‘sidewalk’, ‘curb’, ‘traffic circle’, ‘plot of land’, ‘city bus’, ‘intercity bus’, and ‘a small grocery store where one can also drink liquor at a counter’. The analysis examines whether these lexies establish themselves as predominant norms, indicated by both consistent spatial distribution and a relative frequency exceeding 50% across the localities surveyed. Out of the nine cases studied, eight meet this criterion, with the following overall relative frequencies: ‘semáforo’ 53%; ‘lombada’ 96%; ‘calçada’ 94%; ‘guia’ 84%; ‘rotatória’ 74%; ‘terreno (baldio)’ 79%; ‘ônibus’ (in the sense of ‘city bus’) 71%; ‘ônibus’ (in the sense of ‘intercity bus’) 51%. The ninth variant, ‘bar,’ nearly meets this standard as well, with a relative frequency of 41% just falling short of the minimum threshold.

KEYWORDS

Caipira dialect, geolinguistics, lexicon, Pluridimensional Dialectology, São Paulo State Portuguese

DOI

<https://doi.org/10.14712/18059635.2024.1.2>

1 INTRODUCTION

This paper presents a study of the urban life lexies of Portuguese spoken in *Médio Tietê*, a region of São Paulo State. Such a variety of Portuguese is often referred to as the *Caipira* dialect, making primary reference to its countryside background (cf. Amaral, 1920). The framework adopted is Pluridimensional Dialectology (Radtke & Thun, 1996; Thun, 2000, 2005). The data were collected by means of a semantic-lexical questionnaire developed by the National Committee of the Linguistic Atlas of Brazil (Comitê Nacional do Projeto ALiB, 2001). Nine (Q194-202¹) of its questions were posed to volunteers in order to elicit lexies² as answers, respectively meaning ‘traffic lights’ (Q194), ‘speed bump’ (Q195), ‘sidewalk’ (Q196), ‘curb’ (Q197), ‘traffic circle’ (Q198), ‘plot of land’ (Q199), ‘city bus’ (Q200), ‘intercity bus’ (Q201), and ‘a small grocery store where one can also drink liquor at a counter’ (Q202). A total of 79 variants were gathered, ergo an average of 8.8 per dependent variable. These

1 “Q” stands for “question,” and “194 to 202” is the preserved, original numbering of the questions used in this study that are part of the questionnaire.

2 Pottier’s (1967, p. 17) classical terminology is adopted here, who defines lexie as a memorized lexical unit that can be, more specifically, simple (e.g., *house*), compound (e.g., *summerhouse*), or complex (e.g., *house of worship* ‘church’).



data were yielded by eighty volunteers living in ten localities representative of the region. The diatopic, diastratic, diagenational, and diasexual dimensions were the independent variables controlled in the study.³ Except for the latter, they are put into correlation with the data, and geolinguistic charts show the results, which are then examined.

Supported by the Cooperatio Program (Lingvistika) at Charles University, this article presents an original analysis of the data which I collected (cf. Figueiredo Jr., 2019) under the auspices of the São Paulo Research Foundation — FAPESP (process numbers 2015/14038-5 and 2011/51787-5), the German Academic Exchange Service — DAAD (funding program number 57214225), and the Coordination for the Improvement of Higher Education Personnel — CAPES (process number 0128-16-3/99999.000128/2016-03).⁴

2. LOCUS AND METHODOLOGY

2.1. LOCALITIES

Ten municipalities (P1–P10) in São Paulo State were visited for data collection in 2016 and 2017 and, as such, constitute the research network, as shown in Figure 1.

These localities belong to the *Médio Tietê* region, which is the cradle of the so-called *Caipira*⁵ culture in the state. The oldest and newest localities, Tietê and Piracicaba, were founded in 1570 and 1766, respectively. The region was one of the first in Brazil to be colonized by the Portuguese, a fact that makes its Portuguese one of the oldest varieties spoken in the country. Such historical relevance had been considered before the localities were defined as the research network. An important purpose is to capture the current synchronic state of the language, which may be compared to other synchronic states revealed by other studies.

2.2. INFORMANTS

Eighty research subjects were interviewed. They were evenly divided into groups according to sex (female vs. male), age (18–36 vs. ≥55), and education (low vs. high edu-

3 Flydal (1951) created the terms diatopic (T1) and diastratic (T2) (along with others) for stylistics and the study of loanwords (Rona, 1970, p. 201), and they have been adopted by many to deal with linguistic variation ever since (e.g., Coseriu, 1981), whereas the terms diagenational (T3) and diasexual (T4) (and more) have been invented by analogy by unclear authorship. T1, T2, T3, and T4 denote linguistic variation based on spatial distribution (T1), social factors in general or differences in social classes or in education in particular (T2), age groups (T3), and sexes (T4).

4 I acknowledge the support provided by the above-mentioned institutions and the valuable suggestions given by the anonymous reviewers of this journal.

5 Historically, the term *Caipira* denotes someone (or something) born, having originated, or simply living in rural areas of the Brazilian Southeast and Midwest regions, especially in São Paulo State. It is commonly associated with subsistence farming, itinerant travel, and sometimes landlessness. Details can be found in Candido (2010).

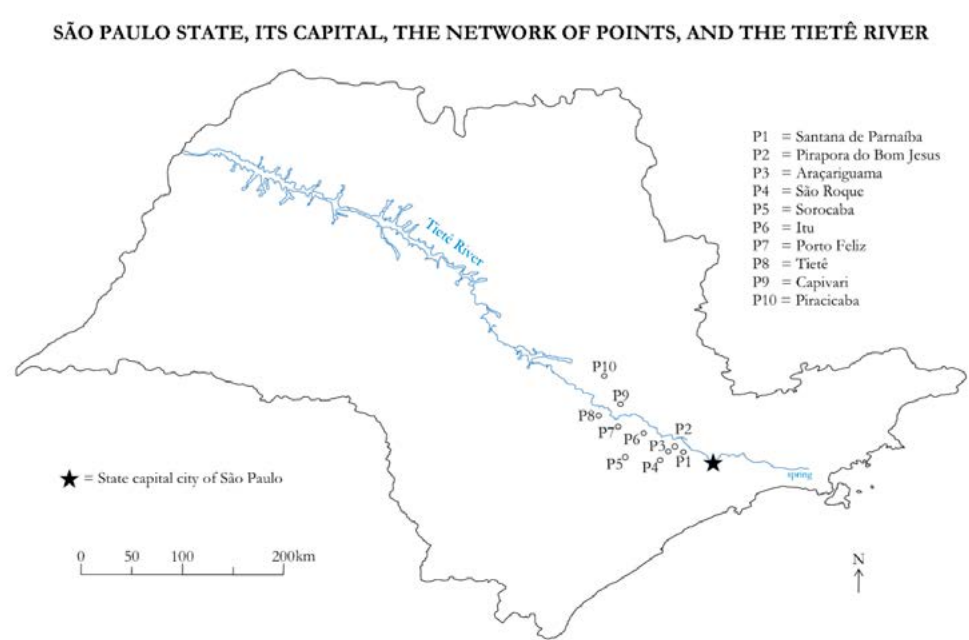


FIGURE 1: Network of points (Figueiredo Jr., 2019, p. 236) [my translation]

cation). Low and high education were deemed to range from illiteracy to incomplete high school and from higher education at least started, respectively.⁶

Furthermore, all informants by the time of the data collection had been living in their respective locality for at least three-quarters of their lifetime, with their last five years of residence being uninterruptedly spent there too. Each locality was represented by eight interviewees from different profiles: four CaGI, CaGII, CbGI, and CbGII women, as well as four CaGI, CaGII, CbGI, and CbGII men.

2.3. MATERIALS

The elicitation was conducted by means of an informant identification sheet, a semantic-lexical questionnaire, a laptop, a laser pointer, and an audio recorder. The identification sheet comprises fields of extralinguistic nature (age, sex, education, and address). The semantic-lexical questionnaire (best known among researchers through its Portuguese abbreviation “QSL”) has 170 questions and is a modified version derived from the 200-question original version made in 2001 by the National Committee of the Linguistic Atlas of Brazil Project (Comitê Nacional do Projeto ALiB, 2001⁷). Although

⁶ For simplicity, and in accordance with the terminology used by the framework adopted here, the young age group is abbreviated as GI, the old age group as GII, low education as Cb, and high education as Ca.

⁷ The Linguistic Atlas of Brazil Project, conducted by its National Committee, is aimed at the creation of a linguistic atlas of Brazil from a dialectological perspective. It involves



the modified version has fewer questions, it keeps the original numbering unaltered to avoid difficulties in the cross-referencing usually performed in comparative works dealing with the same questionnaire.

The laptop—utilized on the occasion to display slides in full screen illustrating referents (see Appendix) denoted by the questions as they were read out loud by the researcher—had a full HD 13.3” screen. Since the laptop was placed as close to the informants as possible, the researcher both moved the slides forward and pointed out specific parts of the images at a distance with the help of a laser pointer when necessary. The answers went into a digital voice recorder with a stereo microphone, an audio rate of 128 kbps, and 75–17,000 Hz.

2.4. METHODOLOGY

Once the researcher arrived in the localities, he looked for prominent people within the communities to help find volunteers who could fit into the predetermined profiles. As soon as they were found and accepted to participate, they were given information on the study and asked to give their informed consent in writing in accordance with the Nuremberg Code (1947), Universal Declaration of Human Rights (1948), Code of Ethics of the University of São Paulo (2001),⁸ Resolution No. 466 of the Brazilian Ministry of Health (2012), and Declaration of Helsinki (2013). The interviews were conducted as soon as the informants had time and at places of their choosing (commonly at their own houses).

Each interview kicked off with the identification sheet. It allowed me to collect extralinguistic information that I later used in correlation with linguistic data. After that, the questionnaire was orally conducted following the certitude-oriented technique (Figueiredo Jr., 2019), directly derived from Thun’s (2000) three-step technique.⁹ The latter consists of first posing the questions to collect spontaneous answers, secondly insisting (by restating the questions, regardless of answers possibly given in the previous step), and finally suggesting variants previously arranged so that the passive knowledge is also investigated. This technique is supposed neither to have its sequence altered nor to miss steps. The former departs from the latter by moving on to the next question after achieving a clear certitude of the informant in providing a beyond-doubt confident answer at any step. Such certitude comes either explicitly or contextually via inference and, as an answer, is named “relevance head” (RH) to be distinguished from other replies, relevance here being defined as the one given by certitude attitude. The charts ahead are built only upon RHs, unlike the tables, which allow for all responses. This technique increases speed significantly and favors contextually relevant responses (relevance here defined as the one given by certitude attitude), but it has a disadvantage in that it does not systematically elicit

different levels (phonetic-phonological, morphosyntactic, etc.) of spoken Portuguese nationwide. Its first official publication is Cardoso et al. (2014).

8 This research had no obligation to go through any ethics committees at the University of São Paulo. Nevertheless, it was conducted under its ethics code.

9 For both a review of the techniques mentioned and a presentation of a new one, see Figueiredo Jr. et al. (2021).

passive answers. However, the research aimed for active responses and to have these charted. Moreover, as mentioned before, a laptop and a laser pointer were utilized as elicitation aids.

2.5. QUESTIONS

This article selects data concerning all the nine questions within the last semantic area—called “urban life”—of the questionnaire, in which they are numbered from Q194 to Q202.

The questions in the questionnaire adopted here are divided into 14 “semantic areas” (which is the term used in the questionnaire) according to their referential proximity (or semantic affinity), as we see between “car” and “bus”, for instance, since both of them refer to types of vehicles. The areas are (1) landforms, (2) atmospheric phenomena, (3) celestial bodies and climate, (4) agropastoral activities, (5) fauna, (6) the human body, (7) life cycles, (8) social interaction and behavior, (9) religions and creeds, (10) children’s games and entertainment, (11) housing, (12) food and kitchen, (13) clothing and accessories, and (14) urban life. I have researched all of these areas across the same network of localities, and this paper—due to its spatial limit—randomly picks out only one of them, namely semantic area 14 with all of its questions, Q194–202. This numbering is kept here so that confusion is avoided and quick identification by QSL researchers is favored.

- Q194 *Na cidade, o que costuma ter em cruzamentos movimentados, com luz vermelha, amarela e verde?*
‘What is the object with red, yellow, and green lights that usually stands at busy traffic intersections?’
- Q195 *Como se chama o morrinho atravessado no asfalto para os carros diminuírem a velocidade?*
‘What is this narrow, raised area lying across a road to force vehicles to slow down?’
- Q196 *Como se chama este caminho, por onde as pessoas andam, ao lado da rua?*
‘What is the name of this path people walk on at the side of a street?’
- Q197 *Como se chama o que separa o/a [calçada] da rua?*
‘What does one call the sidewalk edge adjunct to a street?’
- Q198 *Como se chama o trecho da rua ou da estrada que é circular, que os carros têm que contornar para evitar o cruzamento direto?*
‘What name is given to a circular area where more than two roads join and vehicles must drive around?’
- Q199 *Como se chama a área que é preciso ter ou comprar para se fazer uma casa na cidade?*
‘What to call this plot on which a house is built?’
- Q200 *Como se chama a condução que leva mais ou menos quarenta passageiros e faz o percurso **dentro da cidade**?*
‘What is the name of a large vehicle that transports about 40 passengers **within a city**?’
- Q201 *E quando o percurso é **entre cidades**, como se chama?*
‘What about the one traveling **between cities**?’



Q202 *Como se chama um lugar pequeno com um balcão, onde se bebe [aguardente] e onde também se pode comprar alguma outra coisa?*
 ‘What is this small establishment with a counter at which customers drink liquor and may buy other things?’

3 RESULTS AND ANALYSIS

This section brings all the data collected through all the steps of the interview technique employed and analyzes them. It means that the total number of variants per locality varies. The answers to the questionnaire are presented both written in italics and individually enclosed by ‘<’ for clearer distinctiveness, especially in the running text, as opposed to mere mentions of linguistic forms in general, which are made distinct with italics only. No data are neglected, meaning that even one-occurrence variants—the so-called hapax legomena—are included, in accordance with the concerns and interests that characterize Pluridimensional Dialectology (Radtke & Thun, 1996; Thun 2000, 2005).

3.1. Q194 (‘WHAT IS THE OBJECT WITH RED, YELLOW, AND GREEN LIGHTS THAT USUALLY STANDS AT BUSY TRAFFIC INTERSECTIONS?’)

	Variants	Total occurrences distributed across the network of localities										Σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
Q194	<semáforo>	1	5	6	5	4	6	6	6	6	6	51
	<farol>	7	4	1	3	2	2	3	2	5	1	30
	<senal>	1	1	0	2	1	4	0	1	3	0	13
	<senal de trânsito>	0	0	0	0	0	1	1	1	0	0	3
	<sinaleiro>	0	0	0	0	1	3	2	0	4	2	12
	<sinaleira>	0	0	0	1	1	0	0	0	1	0	3
Σ	6											

TABLE 1: Answers to Question 194 (Figueiredo Jr., 2019, p. 1756) [my adaptation and translation]

Table 1 displays the responses to Question 194 in relation to locality. From this point of view, <semáforo> has the most occurrences. The second place goes to <farol>. The next two—<senal> and <senal de trânsito>—may be taken as a variant group by virtue of both sharing the noun phrase head *senal*. This can also be said for the last two—<sinaleiro> and <sinaleira>—because what distinguishes them is gender inflection (feminine -a vs. masculine -o), which is thus grammatical rather than lexical. All locales show variation, and the greatest one—composed of five lexies in competition—is observed in Sorocaba (P5), Itu (P6), and Capivari (P9).

Points 6 to 10 form, in great number, a continuous and quantitatively uniform route of <semáforo>. The furthest point from such a route within the network con-

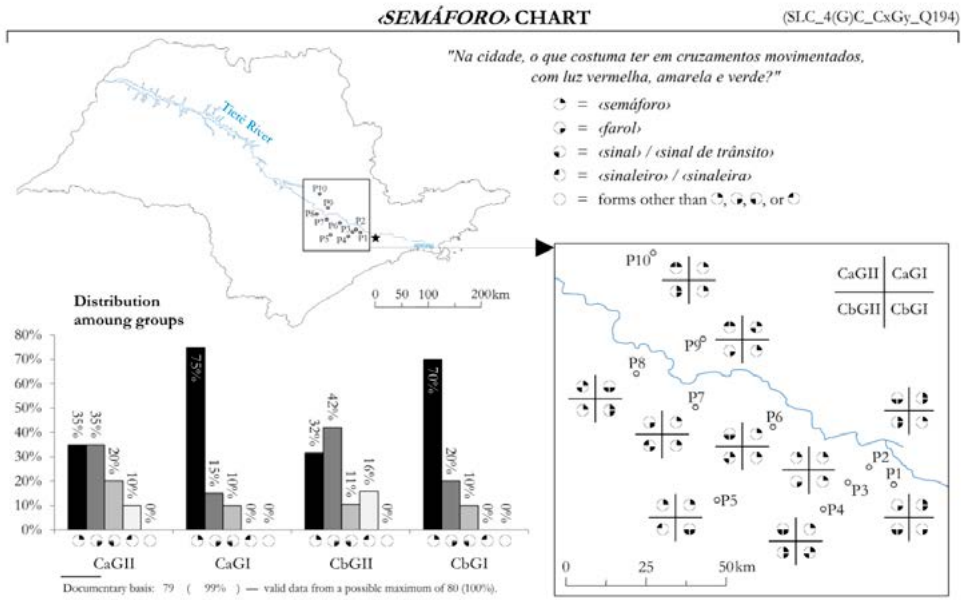


FIGURE 2: Chart of <semáforo> (Figueiredo Jr., 2019, p. 319) [my translation]

cerned is Santana de Parnaíba (P1), where <farol> is more frequent than <semáforo>. When both <senal> and <senal de trânsito> are considered, their spatial distribution does not show quantitative parallel, except in Araçariguama (P3), Tietê (P8), and Piracicaba (P10), where these variants have respectively zero, one, and zero occurrences each. As for <sinaleiro> and its feminine counterpart <sinaleira>, the former is preferred, especially in Capivari (P9).

Figure 2 has relevance heads (as defined earlier) geolinguistically charted. The four most frequent (groups of) variants from a non-correlational perspective are listed in descending order, the first of which gives its name to the chart. Further variants are indistinctly gathered under and represented by an empty circle symbol (all of them are explicitly presented in Table 1). The bar graph shows relative frequencies, and, to its right, one can see how the variants are diatopically distributed. Also note that the data are crossed with both diastratic (Cx) and diagenational variables (Gy), as introduced earlier in simpler terms.

The graph lets us notice that <semáforo> reigns among the young, whereas it competes more closely with <farol> among the old. The dominance of <semáforo> reaches its zenith among young speakers with high education, where 60% separate <semáforo> from second-place <farol>. This percentage is 50% when it comes to young individuals with low education. The only group in which <farol> achieves preference to other variants is that of old people with low education. It obtains a relative frequency of 42%, and its rival <semáforo> 32%. However, they both tie (35% each) among old speakers with high education.



Both spatially and quantitatively, it must be said that <semáforo> establishes itself as an **absolute** norm, for it meets these two criteria: being present throughout the network—what is called regular distribution—and having an overall relative frequency greater than 50%. (One method for confirming this is to compute the arithmetic average of the four individual relative frequencies of <semáforo>, i.e., $(35 + 75 + 32 + 70) / 4 = 53\%$.) Additionally, it nearly becomes a **relative norm**—i.e., a norm defined as the same as the absolute norm except that it is related to a specific participant profile, not to all profiles (cf. Figueiredo Jr., 2019, p. 1955)—among young participants with low and high education. What prevents <semáforo> from becoming a relative norm is only that it does not occur in either the former group in Tietê (P8) or the latter in Santana de Parnaíba (P1) and Sorocaba (P5).

3.2. Q195 ('WHAT IS THIS NARROW, RAISED AREA LYING ACROSS A ROAD TO FORCE VEHICLES TO SLOW DOWN?')

	Variants	Total occurrences distributed across the network of localities										Σ	
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
Q195	<lombada>	8	7	7	8	8	8	8	8	8	8	8	78
	<quebra-molas>	2	1	1	0	0	1	0	1	0	0	0	6
	<guarda-deitado>	1	0	0	0	0	0	0	0	0	0	0	1
	<obstáculo>	0	0	1	0	0	0	0	0	0	0	0	1
	<prefeito-deitado>	0	0	0	1	0	0	0	0	0	0	0	1
	<tartaruga>	0	0	0	0	0	0	0	0	1	0	0	1
Σ	6												

TABLE 2: Answers to Question 195 (Figueiredo Jr., 2019, p. 1757) [my adaptation and translation]

Table 2 presents all the six responses to Question 195. Notice the great predominance of <lombada>, which is the exclusive form in all points, except Pirapora do Bom Jesus (P2) and Araçariçuama (P3). The second most frequent variant, <quebra-molas>, has only six occurrences in the network. The remaining ones appear only once each. Among them, there is <obstáculo>. Since it simply means 'obstacle' on its face value, it illustrates—and reminds us of—the fact that there is a varying degree of context-sensitivity associated with lexies in concrete usage that specifies meaning. This is to be borne in mind.

Moreover, notice that four localities present no variation, that is, no variant competition. They are: Sorocaba (P5), Porto Feliz (P7), Capivari (P9), and Piracicaba (P10). On the other hand, the greatest variation is displayed by three points: Santana de Parnaíba (P1), Araçariçuama (P3), and Tietê (P8). Each of them has three variants in competition under the leadership of <lombada>.

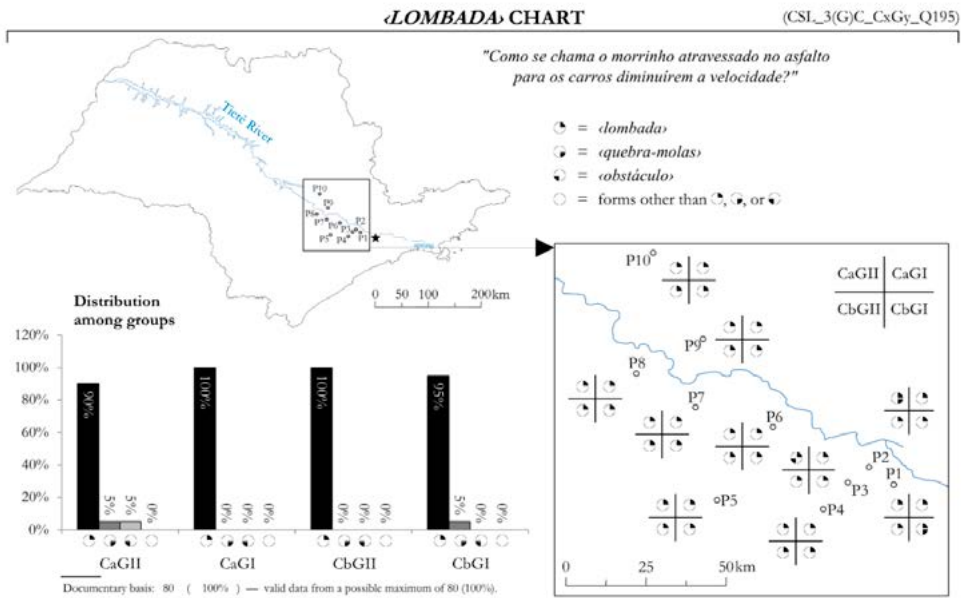


FIGURE 3: Chart of «lombada» (Figueiredo Jr., 2019, p. 543) [my translation]

Figure 3 shows that «lombada» is both an absolute and relative norm for all the participant profiles. The other variants stand no chance. The forms «quebra-molas» and «obstáculo» individually have only a 5% relative frequency among old speakers with high education, and «quebra-molas» has another 5% among young informants with low education.

As far as «lombada»'s absolute diatopic hegemony is concerned, all localities are involved, except Santana de Parnaíba (P1), Pirapora do Bom Jesus (P2), and Aracariçuama (P3). In the first and second of them, «quebra-molas» is also seen, and that on account of both the young with low education and old with high education groups' usage, respectively. As for P3, «obstáculo» is recorded as a hapax legomenon produced by old speakers with high education.

3.3. Q196 ('WHAT IS THE NAME OF THIS PATH PEOPLE WALK ON AT THE SIDE OF A STREET?')

Table 3 exhibits the unquestionable supremacy of «calçada». Its corresponding masculine augmentative, «calçadão», barely appears. Since they differ in neither lexical nor semantic but rather morphological terms, they may be deemed one variant group.¹⁰ The second position in frequency is held by «passeio». The remaining variants, when reckoned individually, have just one occurrence each. As a group, «passeio»,

¹⁰ The form «calçadão» may also mean something else, but here it has the same reference as «calçada».



	Variants	Total occurrences distributed across the network of localities										Σ		
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10			
Q196	⟨calçada⟩	8	7	7	8	7	8	6	8	8	8	75	76	
	⟨calçadão⟩	0	0	1	0	0	0	0	0	0	0	1		
	⟨passeio⟩	0	1	0	0	1	0	0	0	1	0	3		
	⟨passarela⟩	0	0	0	0	0	0	1	0	0	0	1		5
	⟨passatório⟩	1	0	0	0	0	0	0	0	0	0	1		
	⟨espaço para pedestre⟩	0	0	0	0	0	0	1	0	0	0	1		
	⟨pista de caminhada⟩	0	0	0	0	1	0	0	0	0	0	1		
Σ	7													

TABLE 3: Answers to Question 196 (Figueiredo Jr., 2019, p. 1757) [my adaptation and translation]

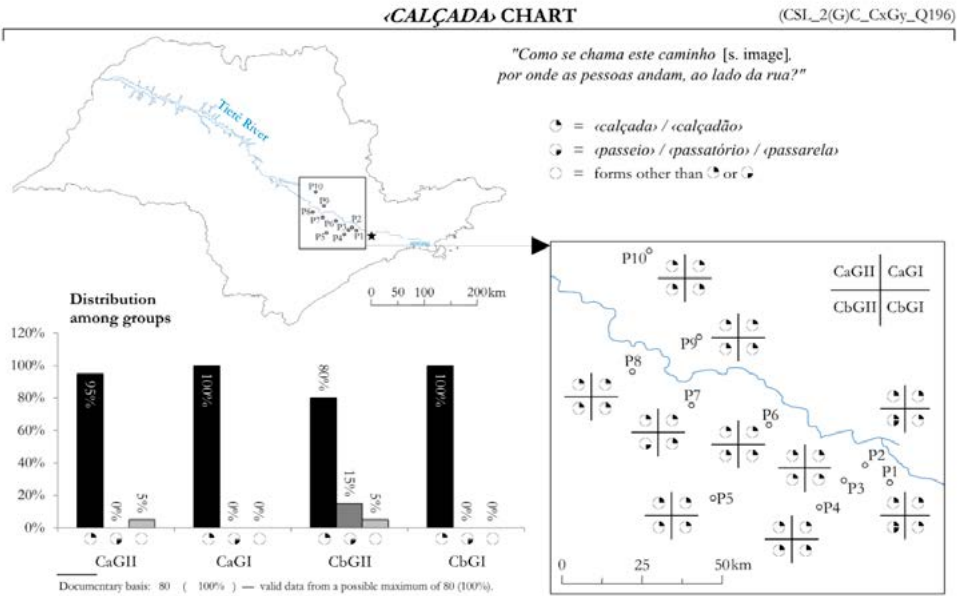


FIGURE 4: Chart of «calçada» (Figueiredo Jr., 2019, p. 656) [my modification and translation]

«passarela», and «passatório»—whose semantic components derived from their morphological suffixes are pragmatically neutralized for the question at issue—amount to five instances.

There are four points without variation: São Roque (P4), Itu (P6), Tietê (P8), and Piracicaba (P10). These have only «calçada». On the other hand, the places where one observes the greatest variation—made up of three lexies—are Sorocaba (P5) and Porto Feliz (P7). In addition to «calçada» in both cases, the former has «passeio» and «pista de caminhada», while the latter has «passarela» and «espaço para pedestre».



As Figure 4 displays, the variant group constituted by <calçada> and <calçada>—not losing sight of the fact that the latter is just a hapax legomenon, as shown in Table 3—is present all over the network and has a relative frequency way over 50%. Hence, it is an absolute norm. Moreover, it is also a relative norm within three specific profiles: old speakers with high education as well as young speakers with high and low education. Almost so is such a variant group within the remaining group, the old one with low education, where it is taken away from achieving the status of relative norm by the fact that the variant group under consideration does not show up in one locality, Porto Feliz (P7). As for the *pass-* variant group—made up of <passaio>, <passatório>, and <passarela>—it is spoken only by old participants with low education in the sample.

Diatopically, the absolute norm in question establishes its full dominance at all sites, except Santana de Parnaíba (P1) and Porto Feliz (P7). These two points have their old inhabitants with low education saying <passaio>, <passatório>, or <passarela> in addition to the absolute norm.

3.4. Q197 ('WHAT DOES ONE CALL THE SIDEWALK EDGE ADJUNCT TO A STREET?')

	Variants	Total occurrences distributed across the network of localities										Σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
Q197	<guia>	6	7	8	7	6	6	7	7	6	8	68
	<meio-fio>	3	1	0	2	2	3	0	0	1	1	13
	<sarjeta>	1	0	0	1	2	0	0	1	2	2	9
Σ	3											

TABLE 4: Answers to Question 197 (Figueiredo Jr., 2019, p. 1758) [my adaptation and translation]

Table 4 shows three variants. By far, <guia> is the most frequent, followed by <meio-fio> and then <sarjeta>. The first is all over the network. So is the second, except in Araçariguama (P3), Porto Feliz (P7), and Tietê (P8), and the absence of the third happens only in Pirapora do Bom Jesus (P2), Araçariguama (P3), Itu (P6), and Porto Feliz (P7). Two places do not see variation: Araçariguama (P3) and Porto Feliz (P7); where <guia> is all that there is from the data sample. Conversely, five points open the way to the maximum variation possible within the limits of this case—which involves three variants—and they are Santana de Parnaíba (P1), São Roque (P4), Sorocaba (P5), Capivari (P9), and Piracicaba (P10).

With the aid of Figure 5, one can see that <guia> emerges as an absolute norm due to both its regular distribution and a relative frequency greater than 50%. Simultaneously, it is a relative norm for all four participant groups. The form <guia> may well be a lexical norm but has no complete dominance. Its rivals <sarjeta> and <meio-fio> have some presence too, although to a fairly low extent. The former holds the position of the second most frequent variant, and its higher value in frequency, 12%, is found among old informants with low education. Its second higher value, 10%, is once again related to the old, this time of high education.

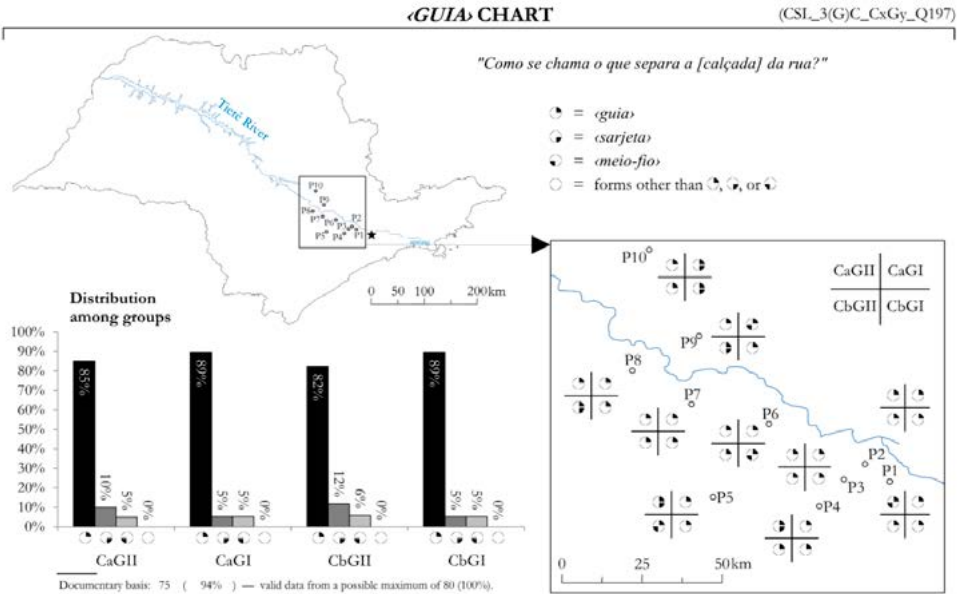


FIGURE 5: Chart of «guia» (Figueiredo Jr., 2019, p. 544) [my translation]

In terms of spatial distribution, «guia» exerts full dominance in some locales: Pirapora do Bom Jesus (P2), Araçariguama (P3), and Porto Feliz (P7). It should also be pointed out that the pairing of «guia» and «sarjeta» happens almost always on the left side of the crosses seen in Figure 5, that is, among old speakers. The only exception in this regard concerns Piracicaba (P10), where such a pairing is located on the young interviewees’ side. By contrast, the other pairing—of «guia» and «meio-fio»—varies as to the profiles it is associated with. It is found in Santana de Parnaíba (P1) among the old informants with high education, in Sorocaba (P5) among the old with low education, and in Itu (P6) among the young with low education.

3.5. Q198 (‘WHAT NAME IS GIVEN TO A CIRCULAR AREA WHERE MORE THAN TWO ROADS JOIN AND VEHICLES MUST DRIVE AROUND?’)

Question 198 received nine different responses, as shown in Table 5. Among them, «rotatória» is the most frequent by a large margin. It is extensively present over the network of points. The second position is occupied by «balão». Unlike «rotatória», «balão» is not omnipresent. Its absence is recorded in Araçariguama (P3), São Roque (P4), Tietê (P8), and Piracicaba (P10). The next two variants are «retorno» and «contorno». Both do obtain almost the same overall quantitative value, but they exhibit a partial divergence in spatial distribution. Individual values ranging from 0 to 2 are mirrored in four locations, from Santana de Parnaíba (P1) to São Roque (P4), and vary in the remaining points within the same quantitative range.

The fifth and sixth variants—having a total of three occurrences each—seem to behave in the sample as if they were in complementary distribution in terms of



	Variants	Total occurrences distributed across the network of localities										Σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
Q198	⟨rotatória⟩	6	5	8	5	6	6	8	7	7	7	65
	⟨balão⟩	2	2	0	0	1	2	1	0	2	0	10
	⟨retorno⟩	0	1	0	2	2	1	1	1	0	0	8
	⟨contorno⟩	0	1	0	2	1	0	0	0	2	1	7
	⟨rótula⟩	1	0	0	0	0	1	1	0	0	0	3
	⟨trevo⟩	0	1	0	1	1	0	0	0	0	0	3
	⟨círculo⟩	1	0	0	0	0	0	0	0	0	0	1
	⟨redonda⟩	0	0	0	0	0	0	0	0	1	0	1
	⟨retunda⟩	1	0	0	0	0	0	0	0	0	0	1
Σ	9											

TABLE 5: Answers to Question 198 (Figueiredo Jr., 2019, p. 1758) [my adaptation and translation]

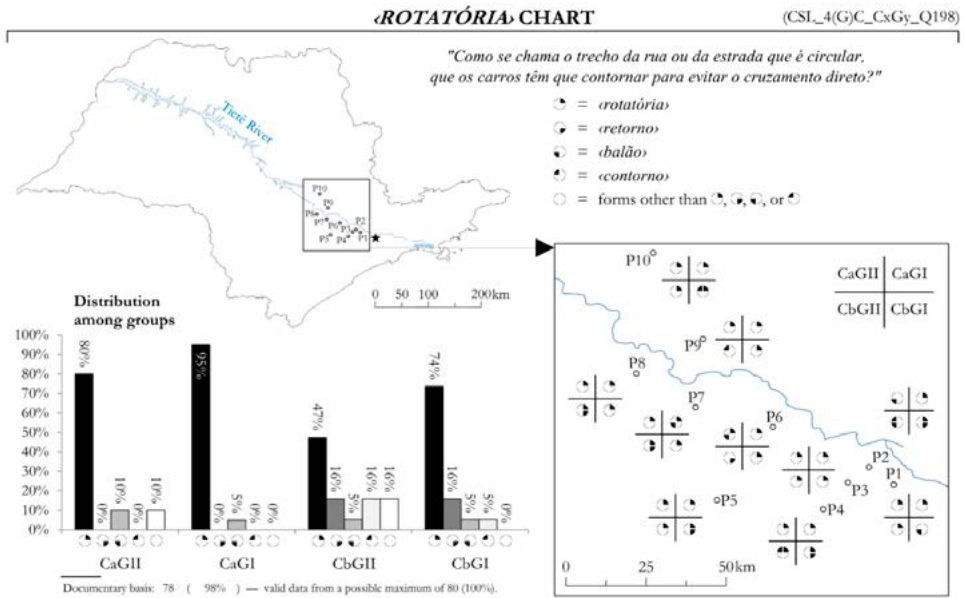


FIGURE 6: Chart of ⟨rotatória⟩ (Figueiredo Jr., 2019, p. 320) [my translation]

pragmatic usage, not language structure. Anyway, this is a hypothesis to be tested in future studies of the network in question, ruling in or out the case of a possible coincidence. In other words, as if they could not simultaneously occur in the same places. Such is what can be seen from Santana de Parnaíba (P1) to Porto Feliz (P7). There are no occurrences of ⟨rótula⟩ and ⟨trevo⟩ in the remaining points. Lastly, Table 5 shows that ⟨círculo⟩, ⟨redonda⟩, and ⟨retunda⟩ each have one occurrence, the first and third lexies in Santana de Parnaíba (P1), and the second in the far point of Capivari (P9).



Despite the total number of variants, there is no variation in Araçariçuama (P₃), where <rotatória> occurs alone. On the other hand, the greatest variation—consisting of five lexies—is found in Santana de Parnaíba (P₁), Pirapora do Bom Jesus (P₂), and Sorocaba (P₅).

Figure 6 shows that <rotatória> is an absolute norm. The graph lets us realize that there is the following correlation: the younger and more educated people are, the more use they make of <rotatória>. After all, its highest and lowest values, 95% and 47%, are found among the young with high education and the old with low education, respectively. Moreover, it is also a relative norm within the former group. As for the remainder of the speaker groups, two of them nearly have <rotatória> as a relative norm too. However, Santana de Parnaíba (P₁) and Pirapora do Bom Jesus (P₂) prevent it from happening in that the variant at issue does not occur among the young with low education and the old with high education, respectively.

All four profiles do display some degree of variation, but the one concerning old participants with low education comes across the most intense lexical competition. These individuals first employ <rotatória> and then get themselves evenly divided as to opting for the second most used variant in that <retorno> and <contorno> share the same rate of occurrence, 16%. It should also be noted that <retorno>¹¹ is spoken only by those with low education in the sample, and this happens quantitatively equally between its two generations. As far as pairings are concerned, the one of <rotatória> and <retorno> is noteworthy: it occurs five times in different locales—Pirapora do Bom Jesus (P₂), São Roque (P₄), Sorocaba (P₅), Porto Feliz (P₇), and Tietê (P₈)—always as a low-education product.

3.6. Q199 ('WHAT TO CALL THIS PLOT ON WHICH A HOUSE IS BUILT?')

	Variants	Total occurrences distributed across the network of localities										Σ	
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
Q199	<terreno>	7	7	7	8	6	6	6	3	8	7	65	68
	<terreno baldio>	0	0	0	0	0	0	1	2	0	0	3	
	<lote>	1	0	4	1	3	2	3	5	3	2	24	25
	<lote de terreno>	0	0	0	0	0	1	0	0	0	0	1	
	<bloco>	0	1	0	0	0	0	0	0	0	0	1	
Σ	5												

TABLE 6: Answers to Question 199 (Figueiredo Jr., 2019, p. 1759) [my adaptation and translation]

As displayed by Table 6, five variants were collected as answers to Question 199: <terreno>, <terreno baldio>, <lote>, <lote de terreno>, and <bloco>. Because the first and second variants, as well as the third and fourth ones, share the noun phrase heads *terreno* and *lote*, respectively, they may be taken as variant groups (this device is especially use-

¹¹ The form <retorno> may also mean something else, but here it has the same reference as <rotatória>.

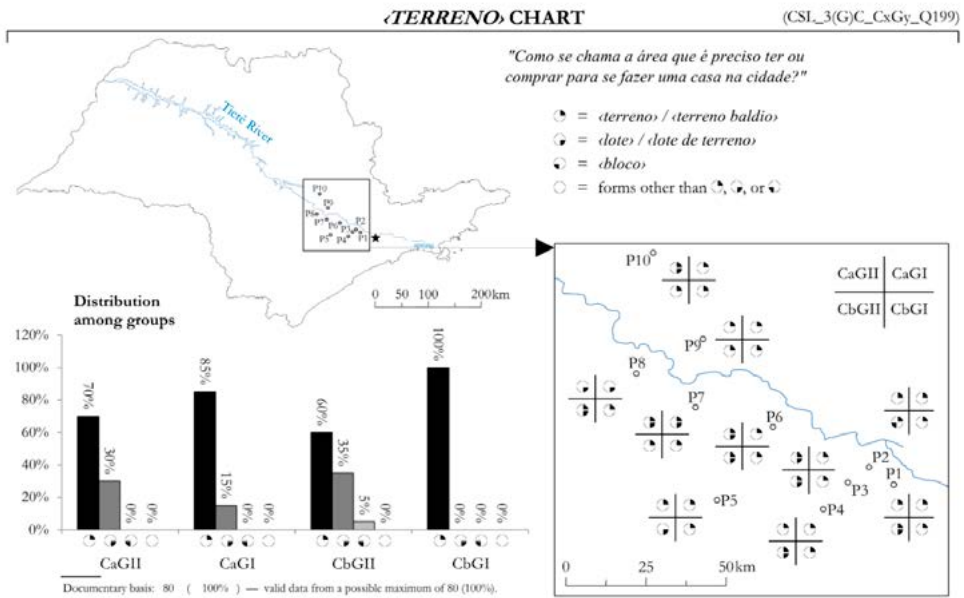


FIGURE 7: Chart of «terreno» (Figueiredo Jr., 2019, p. 321) [my modification and translation]

ful with regard to the chart in Figure 7). The form «terreno» reigns over the network, except in Tietê (P8), where the informants prefer «lote» the most. Speaking of which, this option holds the second position in frequency. It most certainly loses strength in Santana de Parnaíba (P1) and São Roque (P4) and is absent in Pirapora do Bom Jesus (P2). The remainder is formed by the hapax legomena «bloco» and «lote de terreno», when considered individually. The greatest variation is observed in three adjacent locales: Itu (P6), Porto Feliz (P7), and Tietê (P8). As for lack of variation, there are no points with only one variant.

Quite obviously, the variant group formed by «terreno» and «terreno baldio»—or simply “«terreno (baldio)»” for easier reference—reveals itself as an absolute norm, as Figure 7 indicates. More than that, it also made itself a relative norm among young speakers with low education. Regarding both the profiles with high education on one hand and the old one with low education on the other hand, «terreno (baldio)» falls short of becoming a relative norm solely due to one locality in each case, Tietê (P8) and Sorocaba (P5), where «terreno (baldio)» does not occur.

Leaving aside young informants with low education, among whom there is no variation owing to the absolute reign of «terreno (baldio)», lexical competition is observed among the other three profiles. Such variation is forged by virtually only two forms, «terreno (baldio)» and «lote (de terreno)», and is fiercer among old speakers with low education, with the former having a relative frequency of 60% and the latter having a relative frequency of 35%. Only one locality is a stage for the absolute dominance of «terreno (baldio)»: Capivari (P9). The remainder does see variation. The highlights in this respect are São Roque (P4), Itu (P6), and Porto Feliz (P7). These points each have



at least two participant profiles with two variants charted. A fact that connects P4, P6, and P7 is the feature that they, as neighbors to one another, lie in the network's central area.

3.7. Q200 ('WHAT IS THE NAME OF A LARGE VEHICLE THAT TRANSPORTS ABOUT 40 PASSENGERS **WITHIN A CITY?**')

	Variants	Total occurrences distributed across the network of localities										Σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
Q200	<ônibus>	6	7	7	8	7	6	6	5	3	6	61
	<circular>	2	1	0	0	0	2	0	1	6	0	12
	<ônibus circular>	0	0	0	0	1	1	2	2	0	1	7
	<busão>	0	1	0	0	1	1	1	1	1	1	7
	<jardineira>	0	0	0	0	0	0	2	1	2	0	5
	<coletivo>	1	0	1	0	0	0	0	0	0	0	2
	<ônibus coletivo>	0	0	1	0	0	0	0	0	0	0	1
	<buzu>	0	0	0	0	0	1	0	0	0	0	1
	<condução>	1	0	0	0	0	0	0	0	0	0	1
	<ônibus de linha>	0	0	0	0	0	0	0	1	0	0	1
	<ônibus municipal>	0	0	0	0	0	0	0	0	0	1	1
	<transporte>	1	0	0	0	0	0	0	0	0	0	1
Σ	12											

TABLE 7: Answers to Question 200 (Figueiredo Jr., 2019, p. 1759) [my adaptation and translation]

Note that grouping all the variants in Table 7 that either expressly or elliptically contain the form <ônibus> or some derivative of it within their respective noun phrases would be pointless, since almost all of them have it (exceptions are <jardineira>, <condução>, and <transporte>). So, it is preferable that a grouping criterion be something more plausible. A possible one involves modifiers applied to <ônibus>. From this perspective, we can group <circular> and <ônibus circular>, for the former is originally a modifier, an adjective that has been substantivized and is actually functioning as a noun itself. The same process is seen with <coletivo> and <ônibus coletivo>.

Table 7 exhibits 12 variants, led far and away by <ônibus>. Then come <circular> and <ônibus circular>, both reckoned as one variant group by virtue of the former obviously having the noun phrase head *ônibus* tacitly understood, which is enforced by two facts. First, *circular* is primarily an adjective. Second, the sample does not show another noun to form the noun phrase at issue other than *ônibus*. The next variant is <busão>, morphologically comprehended either as a result of both apheresis and augmentative suffixation applied to *ônibus* or as a result of just augmentative suffixation applied to the English word *bus* (which is short for *omnibus*; cf. Onions, 1982 [1966], p. 129).

The form <jardineira> is up next from Table 7 and occurs in quantitative similarity to <busão>, which may be perceived as a slang. The lexies <coletivo> and <ônibus coletivo>

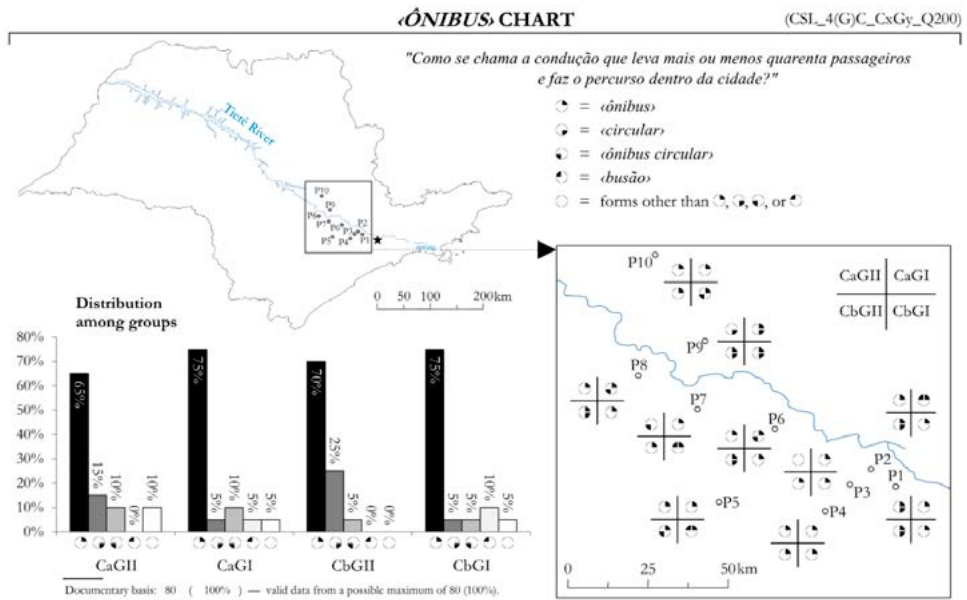


FIGURE 8: Chart of <ônibus> [Q200] (Figueiredo Jr., 2019, p. 322) [my translation]

come as a variant group. Again, the noun *ônibus* is implicitly understood in the former alternative, along with the explicit lexie *coletivo*, which primarily belongs to adjective category. The remainder is made up of five hapax legomena. The first of them, <buzu>, merits a brief note. It is a slang word often associated with metropolitan-suburban low-educated youth, and its spatial distribution in Brazilian territory is yet to be diachronologically revealed and backed up by large data.¹²

Diatopically, <ônibus> is omnipresent across the network of localities. The variant group constituted by <circular> and <ônibus circular> fails to occur in Araçariçuama (P3) and São Roque (P4). The form <busão> manifests itself in a one-occurrence series throughout the network, except at three sites: Santana de Parnaíba (P1), Araçariçuama (P3), and São Roque (P4). By contrast, <jardineira> is recorded only in three points: Porto Feliz (P7), Tietê (P8), and Capivari (P9). The variant group consisting of <coletivo> and <ônibus coletivo> occurs only in Santana de Parnaíba (P1) and Araçariçuama (P3). The remainder of the forms have one occurrence each in different locations, except <condução> and <transporte>, which are both seen in Santana de Parnaíba (P1). São Roque (P4) has no variation due to the occurrence of just <ônibus>. On the other hand, the greatest variation, made up of six lexies, is recorded in Tietê (P8).

This time, the first three variants (as listed in Figure 8) are not grouped. The intent is to observe how the presumed underlying full expression *ônibus circular* comes into utterance and then capture its distribution. This measure thus presupposes that

¹² The Linguistic Atlas of Brazil Project does provide for such a purpose, but its publication (cf. Cardoso et al., 2014) does not contain Q199 results.



both *⟨ônibus⟩* and *⟨circular⟩* are concrete realizations of *ônibus circular* licensed by the principle of context-sensitivity. With this considered, Figure 8 shows that *⟨ônibus⟩* is an absolute norm, and its power finds its pinnacle in São Roque (P₄), where it reigns in its fullest expression, as all profiles use it and there are no competitors.

The form *⟨circular⟩* holds the second most frequent variant position, and the old speakers are the ones to utilize it in a quantitatively significant way, with over 15%, whereas the young use it at a rate of just 5%. The second position among the young is divided by two lexies, *⟨ônibus circular⟩* and *⟨busão⟩*, the former being preferred by individuals with high education and the latter by those with low education.

With the aid of the enlarged map in Figure 8, one sees that the pairing of *⟨ônibus⟩* and *⟨circular⟩* is concentrated within the group with low education, especially among its old individuals. The pairing of *⟨ônibus⟩* and *⟨ônibus circular⟩* happens more frequently among young informants. Lastly, the pairing of *⟨ônibus⟩* and *⟨busão⟩* is found among young speakers. As for the places in the network where variation is more intense in general, they are Santana de Pirapora do Bom Jesus (P₂), Araçariguama (P₃), Sorocaba (P₅), Itu (P₆), Porto Feliz (P₇), and Tietê (P₈).

3.8. Q201 ('WHAT IS THE NAME OF A LARGE VEHICLE THAT TRANSPORTS ABOUT 40 PASSENGERS **BETWEEN CITIES?**')

Question 201 received the greatest number of variants among the questions dealt with in this article. As Table 8 shows, 17 answers were collected. Since many of them contain *ônibus*, this noun is not taken as a criterion for grouping variants. With this considered, the most frequent response is *⟨ônibus⟩*, present in all localities, though less strongly in Santana de Parnaíba (P₁), Araçariguama (P₃), Capivari (P₉), and Piracicaba (P₁₀). Then comes *⟨ônibus intermunicipal⟩*, absent at four sites: Santana de Parnaíba (P₁), Pirapora do Bom Jesus (P₂), São Roque (P₄), and Itu (P₆).

The variant group formed by *⟨circular⟩* and *⟨ônibus circular⟩* holds the third position in frequency. It appears in half of the network: Santana de Parnaíba (P₁), Araçariguama (P₃), Tietê (P₈), Capivari (P₉) and Piracicaba (P₁₀). The forms *⟨busão⟩* and *⟨ônibus de viagem⟩* have five occurrences each and occur in the same locales, except Araçariguama (P₃), where *⟨ônibus de viagem⟩* is recorded, but not *⟨busão⟩*. The complex lexie *⟨ônibus interurbano⟩* has a total of three occurrences that are distributed among Itu (P₆), Capivari (P₉), and Piracicaba (P₁₀).

The next three—*⟨ônibus municipal⟩*, *⟨ônibus rodoviário⟩*, and *⟨ônibus urbano⟩*—occur twice each over the network. The remaining variants are seven hapax legomena that roughly cluster in the first half of the points (P₁–P₅). More precisely, the places are Santana de Parnaíba (P₁), Pirapora do Bom Jesus (P₂), Araçariguama (P₃), and Sorocaba (P₅). Excluding P₁ and P₅, these variants do not co-occur. Note that none of them are documented in no less than six localities (those not mentioned). Moreover, there is no locale without variation. By contrast, two points—Santana de Parnaíba (P₁) and Capivari (P₉)—exhibit the greatest variation in the case, which comprises seven variants. One of them merits a brief note. *⟨Cometa[®]⟩*—as opposed to *cometa* ‘comet’—is so written to denote that it primarily refers to a transport company whose popularity seems to lend its name to serve sometimes as another synonym of the items listed in Table 8.

	Variants	Total occurrences distributed across the network of localities										Σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
Q201	⟨ônibus⟩	2	6	2	6	4	7	6	5	1	2	41
	⟨ônibus intermunicipal⟩	0	0	2	0	1	0	1	1	4	4	13
	⟨circular⟩	1	0	1	0	0	0	0	1	2	1	6
	⟨ônibus circular⟩	0	0	0	0	0	0	0	1	0	0	1
	⟨busão⟩	1	0	0	0	0	0	0	2	1	1	5
	⟨ônibus de viagem⟩	1	0	1	0	0	0	0	1	1	1	5
	⟨ônibus interurbano⟩	0	0	0	0	0	1	0	0	1	1	3
	⟨ônibus municipal⟩	0	0	0	1	1	0	0	0	0	0	2
	⟨ônibus rodoviário⟩	0	0	1	1	0	0	0	0	0	0	2
	⟨ônibus urbano⟩	0	0	0	0	0	0	1	0	1	0	2
	⟨coletivo⟩	1	0	0	0	0	0	0	0	0	0	1
	⟨Cometa [®] ⟩	0	0	1	0	0	0	0	0	0	0	1
	⟨condução⟩	0	1	0	0	0	0	0	0	0	0	1
	⟨executivo⟩	1	0	0	0	0	0	0	0	0	0	1
	⟨micro-ônibus⟩	0	0	0	0	1	0	0	0	0	0	1
	⟨ônibus de carreteira⟩	0	0	0	0	1	0	0	0	0	0	1
	⟨ônibus de luxo⟩	1	0	0	0	0	0	0	0	0	0	1
Σ	17											

TABLE 8: Answers to Question 201 (Figueiredo Jr., 2019, p. 176o) [my adaptation and translation]

What has been said about Figure 8 in connection with the decision not to group the variants applies here to Figure 9. With this considered, ⟨ônibus⟩ emerges as an absolute norm. It does not come close to being a relative norm, except among young speakers with high education, inasmuch as ⟨ônibus⟩ fails to be used by them in only one locality, Piracicaba (P10). The profile making the most use of ⟨ônibus⟩ is that of young individuals with high education. In the matter of ⟨ônibus intermunicipal⟩, it appears in notable numbers within all informant groups, excluding young individuals with low education, among whom ⟨ônibus intermunicipal⟩ does not occur.

The complex lexie ⟨ônibus de viagem⟩ is not spoken by the old, and most of the young who speak it have low education. As for ⟨circular⟩, it is not spoken by individuals with high education, and those who have low education who speak it most are the old. Another fact indicated by Figure 9 that deserves attention is the relatively high rate of forms other than the four explicitly cited on the chart within the four participant profiles. The highest profile percentage, 32%, is found among old participants with low education. It is so high among them that it comes quite near the absolute norm's relative frequency, 37%.

When diatopically compared to its three charted competitors, ⟨ônibus⟩ exerts complete dominance in two points: São Roque (P4) and Itu (P6). The pairing of ⟨ônibus⟩

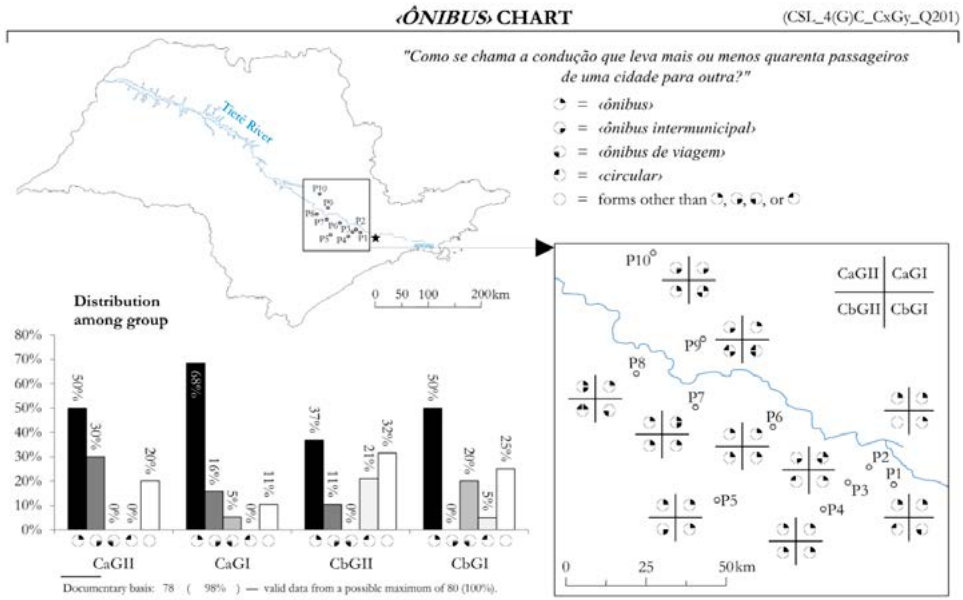


FIGURE 9: Chart of <ônibus> [Q201] (Figueiredo Jr., 2019, p. 323) [my translation]

and <ônibus intermunicipal> occurs in two adjacent locales, Porto Feliz (P7) and Tietê (P8), within their group with high education. The pairing <ônibus> and <ônibus de viagem> is recorded at two sites distant from each other, Araçariguama (P3) and Piracicaba (P10), spoken by the young. The other pairings only happen once each.

3.9. Q202 ('WHAT IS THIS SMALL ESTABLISHMENT WITH A COUNTER AT WHICH CUSTOMERS DRINK LIQUOR AND MAY BUY OTHER THINGS?')

By looking at Table 9, one can tell that the top two items are in closer competition for the first position in frequency than any other, being <bar> somewhat preferred to <boteco>. Then comes <armazém>, which, unlike the previous ones, is not present throughout the network of points. It fails to occur in Araçariguama (P3) and Piracicaba (P10). Up next is the variant group formed by <venda> and <vendinha>. The latter is a mere diminutive of the former, hence not being a different lexical item. As a group, they are not represented in Pirapora do Bom Jesus (P2), Araçariguama (P3), Itu (P6), and Porto Feliz (P7).

The alternative <mercearia> comes next and is not seen only in P2 and P7. The simple lexie <balcão> is half-present and half-absent over the network. The forms <bodega>, <botequim>, <adega>, and <empório> each have two to three appearances in different localities. The last three variants—<mercadinho>, <mercado>, and <quitanda>—share the fact that they have only one occurrence each in the network.

As Figure 10 allows us to infer, <bar> is close to being an absolute norm. It does have regular distribution in its favor, but its overall relative frequency (41%) falls below the minimum rate required (>50%). As a logical consequence, neither is it a relative



	Variants	Total occurrences distributed across the network of localities										Σ	
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
Q202	<bar>	2	4	6	6	4	5	3	5	4	6	45	
	<boteco>	2	4	2	6	3	3	3	1	6	6	36	
	<armazém>	2	2	0	1	1	1	5	1	4	0	17	
	<venda>	1	0	0	1	1	0	0	2	3	3	11	
	<vendinha>	0	0	0	0	0	0	0	1	0	0		1
	<mercearia>	2	0	1	1	2	1	0	1	1	1	10	
	<balcão>	1	1	0	0	1	1	0	0	0	1	5	
	<bodega>	0	0	0	0	0	1	0	0	1	2	4	
	<botequim>	1	0	0	2	0	0	0	0	1	0	4	
	<adega>	0	0	1	0	0	0	0	2	0	0	3	
	<empório>	0	0	0	1	0	0	1	0	0	1	3	
	<mercadinho>	0	0	1	0	0	0	0	0	0	0	1	2
	<mercado>	0	0	0	1	0	0	0	0	0	0	1	
<quitanda>	0	1	0	0	0	0	0	0	0	0	1		
Σ	14												

TABLE 9: Answers to Question 202 (Figueiredo Jr., 2019, p. 1761) [my adaptation and translation]

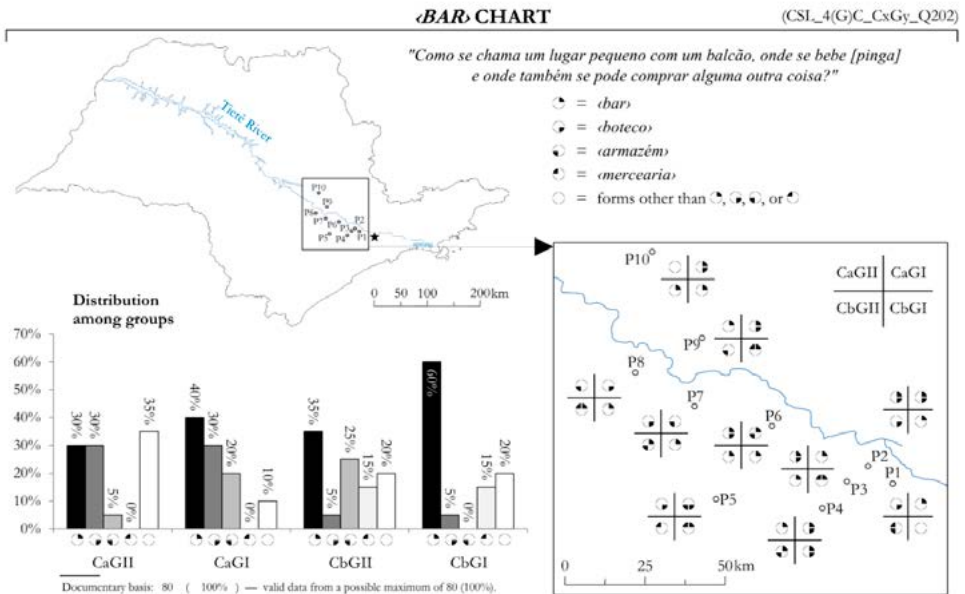


FIGURE 10: Chart of <bar> (Figueiredo Jr., 2019, p. 324) [my translation]



norm. Be that as it may, <bar> has a significant 60% among young informants with low education and, with lower values, is nevertheless the most preferred individual variant by both the old with low education and the young with high education. As for the old with high education, they evenly alternate between <bar> and <boteco>.

Speaking of <boteco>, it is used virtually only by informants with high education. The form <armazém> is spoken virtually only by two interviewee groups, the young with high education and the old with low education. The simple lexie <mercearia> is exclusively spoken by speakers with low education in the sample. For the rest, note that the other variants represented by an empty circle achieve important rates, meaning that variation as to Question 202 is very intense. Such other variants, when combined, even outnumber every variant of the four charted lexies among old informants with high education.

No localities are under the dominance of just one variant. Intense variation is hence observed in terms not only of frequency but also of spatial distribution. There are some copious pairings. That of <bar> and <boteco> comes about in Pirapora do Bom Jesus (P2), Araçariçuama (P3), São Roque (P4), Itu (P6), Capivari (P9), and Piracicaba (P10) and is very prevalent among those with high education. The pairing of <bar> and <mercearia> occurs in Araçariçuama (P3), Sorocaba (P5), Tietê (P8), and Capivari (P9), and is exclusive to the speakers with low education, with prevalence among the young. Another noteworthy pairing is that of <bar> and <armazém>. Its occurrences are located south of the Tietê River; to wit, in São Roque (P4), Itu (P6), and Porto Feliz (P7). The participant profile more inclined to yield it seems to be that of the old with low education.

4 CONCLUSIONS

This section summarizes the main findings laid out in this article. All variants are brought up. However, consistently with previous sections, the analytical focus is on the cited ones in the charts. These are the most frequent ‘relevance heads’ (as defined in the methodology section). With such a caveat in mind, let us begin with the lexies as answers to Question 194 (‘What is the object with red, yellow, and green lights that usually stands at busy traffic intersections?’). Six of them were collected: <semáforo>, <farol>, <senal>, <senal de trânsito>, <sinaleiro>, and <sinaleira>. Analytically, the third and fourth ones were deemed to be a variant group, and so were the last two. This decision was based on their sharing the same stem as the noun phrase head. The form <semáforo> is most often spoken by the young, while the old prefer <farol>. Be that as it may, <semáforo> is an absolute norm, i.e., a form with both regular distribution (occurrence throughout the network of localities investigated) and a relative frequency greater than 50%.

Question 195 (‘What is this narrow, raised area lying across a road to force vehicles to slow down?’) also got six variants as responses: <lombada>, <quebra-molas>, <guarda-deitado>, <obstáculo>, <prefeito-deitado>, and <tartaruga>. The first two attained some significant values, whereas the others had only one occurrence each, a fact that then makes them hapax legomena. The form <lombada> is so strong that it makes it not only



an absolute norm but also a relative one for all four participant profiles, meaning that this variant is present within every profile throughout the network. As for the compound lexie <quebra-molas>, it appears only timidly, solely among the old with high education and the young with low education.

Seven variants were elicited by Question 196 ('What is the name of this path people walk on at the side of a street?'): <calçada>, <calçadão>, <passείο>, <passarela>, <passatório>, <espaço para pedestre>, and <pista de caminhada>. The first two constitute a variant group (G₁), and so do the three that follow (which can be referred to as *pass*-group, or G₂). G₁ is an absolute norm (with virtual hegemony across the network) in general and a relative norm for the old with high education and the young with high and low education in particular. G₂ is used only by old participants with low education.

A total of three variants were used to reply to Question 197 ('What does one call the sidewalk edge adjunct to a street?'): <guia>, <sarjeta>, and <meio-fio>. The first manifests itself as being both an absolute and relative norm for all four participant profiles. The other two lexies occur in <guia>'s shadow in terms of relative frequency. The form <sarjeta> is mostly found among the old individuals, and <meio-fio> seems to be indifferent to participant profiles.

Question 198 ('What name is given to a circular area where more than two roads join and vehicles must drive around?') received nine variants as responses: <rotatória>, <balão>, <retorno>, <contorno>, <rótula>, <trevo>, <círculo>, <redonda>, and <retunda>. The first is an absolute norm, and its usage appears to be in the following relationship: the younger and more educated, the more speakers utter the norm mentioned. Additionally, <rotatória> is a relative norm among young interviewees with high education. On the other hand, the opposite profile—i.e., the old one with low education—displays the most intense variation, which is forged by the charted variants <rotatória>, <retorno>, and <contorno>. The form <retorno> is spoken only by individuals with low education in the sample.

As for Question 199 ('What to call this plot on which a house is built?'), it is the case of five variants: <terreno>, <terreno baldio>, <lote>, <lote de terreno>, and <bloco>. Analytically, they correspond to three, as in <terreno (baldio)> (G₁), <lote (de terreno)> (G₂), and <bloco>. G₁ is an absolute norm and, among young informants with low education, is a relative one that gives other variants no room. Among the other participant profiles, variation arises through the competition mainly between G₁ and G₂. Spatially speaking, Capivari (P₉) is a stage for the complete dominance of G₁, while especially São Roque (P₄), Itu (P₆), Porto Feliz (P₇) are a stage for noteworthy variation. (P₄, P₆, and P₇ are neighbors to each other and lie in the network's central area, and this fact has its importance, too.)

In turn, Question 200 ('What is the name of a large vehicle that transports about 40 passengers **within a city**?') got twelve variants as answers: <ônibus>, <circular>, <ônibus circular>, <busão>, <jardineira>, <coletivo>, <ônibus coletivo>, <buzu>, <condução>, <ônibus de linha>, <ônibus municipal>, and <transporte>. The lexies from <buzu> to <transporte> are hapax legomena. As for <ônibus>, it is an absolute norm, and its strength is particularly noticeable in São Roque (P₄), where it reigns without competitors among all the participant profiles. The form <circular> holds the second most frequent



one, chiefly owing to old speakers. Putting aside the absolute norm, the young prefer <ônibus circular> and <busão> in alternation instead of <circular>. Among such participants, the alternative <ônibus circular> is favored among the ones with high education, and <busão> among the ones with low education. Lastly, let it be said that variation as to Question 200 is more intense in Santana de Pirapora do Bom Jesus (P2), Araçariçuama (P3), Sorocaba (P5), Itu (P6), Porto Feliz (P7), and Tietê (P8).

By receiving seventeen lexies as individual replies, Question 201 ('What about the one traveling **between cities**?') is the one to display the greatest number of variants in the sample. They are: <ônibus>, <ônibus intermunicipal>, <circular>, <ônibus circular>, <busão>, <ônibus de viagem>, <ônibus interurbano>, <ônibus municipal>, <ônibus rodoviário>, <ônibus urbano>, <coletivo>, <Cometa®>, <condução>, <executivo>, <micro-ônibus>, <ônibus de carreteira>, <ônibus de luxo>. What has been said about variant grouping in connection with the previous question (involving the lexie <ônibus>) also applies here. The simple lexie <ônibus> is an absolute norm, which is mostly used by young people with high education and exerts full dominance in São Roque (P4) and Itu (P6). As for the complex lexie <ônibus intermunicipal>, it is quite well spread among all the informant groups, except the young one with low education, where such a variant does not occur. The complex lexie <ônibus de viagem> is not spoken by the old, and the young who use it most have low education. The simple lexie <circular> is not employed by informants with high education, and the speakers with low education who employ it most are the old.

Lastly, answers to Question 202 ('What is this small establishment with a counter at which customers drink liquor and may buy other things?') are presented. Fourteen simple lexies were recorded: <bar>, <boteco>, <armazém>, <venda>, <vendinga>, <mercearia>, <balcão>, <bodega>, <botequim>, <adega>, <empório>, <mercadinho>, <mercado>, and <quitanda>. Such a great number of variants, along with their not inconsiderable relative frequencies and wide spatial distribution, constitute a factor not that conducive to the emergence of an absolute norm. Nonetheless, <bar>—one of the four more closely examined in accordance with the previously exposed criteria—was nearly there, with the young as its primary speakers. The form <boteco> is utilized virtually only by informants with high education, and <armazém> by both the young with high education and the old with low education, whereas <mercearia> is uttered exclusively by individuals with low education.

REFERENCES

- Amaral, A. (1920). *O dialeto caipira*. São Paulo: Anhembi.
- Candido, A. (2010). *Os parceiros do rio Bonito. Estudo sobre o caipira paulista e a transformação dos seus meios de vida*. 11. ed. Rio de Janeiro: Ouro sobre Azul.
- Cardoso, S. A. M. S. et al. (2014). *Atlas Linguístico do Brasil*. Vol. 2. Londrina: Eduel.
- Comitê Nacional do Projeto ALiB (2001). *Atlas Linguístico do Brasil: questionários 2001*. Londrina: Eduel.
- Coseriu, E. (1981). Los conceptos de 'dialeto', 'nivel' y 'estilo de lengua' y el sentido propio de la dialectología. *Linguística española actual*, 3, 1–32. Madrid: Ediciones Cultura Hispánica del Centro Iberoamericano de Cooperación.

- Figueiredo Jr., S. R. (2019). *Atlas linguístico pluridimensional do português paulista: níveis semântico-lexical e fonético-fonológico do vernáculo da região do Médio Tietê* [Doctoral dissertation, University of São Paulo]. <https://www.teses.usp.br/teses/disponiveis/8/8142/tde-12072019-155328/en.php>
- Figueiredo Jr., S. R., Margotti, F. W., Santiago-Almeida, M. M., & Nascimento, J. F. (2021). Metodologia geolinguística: agentes em geral e técnicas de inquérito semântico-lexical em específico. *Cadernos de Estudos Linguísticos*, 63(0), 1-16 (e021010).
- Flydal, L. (1951). Remarques sur certains rapports entre le style et l'état de langue. *Norsk Tidsskrift for Sprogvidenskap*, 16, 240-257.
- Onions, C. T. (Ed.). (1982 [1966]). *The Oxford dictionary of English etymology*. Oxford / New York: Clarendon Press / Oxford University Press.
- Pottier, B. (1967). *Présentation de la linguistique. Fondements d'une théorie*. Paris: Klincksieck.
- Radtke, E., & Thun, H. (1996). Neue Wege der romanischen Geolinguistik: eine Bilanz. In E. Radtke, & H. Thun (Eds.), *Akten des Symposiums zur empirischen Dialektologie* (pp. 1-24). Kiel: Westensee-Verl.
- Rona, J. P. (1970). A structural view of sociolinguistics. In P. L. Garvin (Ed.), *Method and Theory in Linguistics* (pp. 199-211). The Hague / Paris: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9783110872521.199>.
- Thun, H. (2005). A dialetologia pluridimensional no Rio da Prata. In A. M. S. Zilles (Ed.), *Estudos de variação linguística no Brasil e no Cone Sul* (pp. 63-92). Porto Alegre: Editora da UFRGS.
- Thun, H. (Ed.). (2000). *Atlas lingüístico diatópico y diastrático del Uruguay*. Vol. 1. Kiel: Westensee-Verl.



Selmo Ribeiro Figueiredo Junior

Faculty of Arts, Charles University
 nám. J. Palacha 2, 110 00 Prague 1, CZ
 ORCID-ID: 0000-0001-8367-0306
 selmojunior@gmail.com



APPENDIX

Images to Questions 194–202 (Figueiredo Jr., 2019, p. 2017–2020).



Image to Q194



Image to Q195



Image to Q196



Image to Q197



Image to Q198



Image to Q199



Image to Q200



Image to Q201



Image to Q202