GROUPING TYPOLOGY AND DIFFUSION IN THE BRAZILIAN SOUTH – ESTABLISHING GUIDELINES FOR FUTURE RESEARCH

VIRGINIA MEIRELLES, ABDELHAK RAZKY

1. INTRODUCTION

The rhotic consonants in BP are identified by <r> and <rr> in written form and have historically changed from alveolar tap and trill to fricatives. Still the variants are numerous and may include [r, ɾ, ɹ, ɻ, ɽ, x, h, ɣ, Ø] depending on the position in the syllable (Cagliari 1982, 2007; Callou *et al*. 1996). In coda position, for example, Cagliari (1982, 2007) observes that the realizations can include [h, ɦ, χ, ɣ, ɹ, x, ɾ, r, ɻ, Ø] at the time that they are never deleted in onset.

Yet, fricative realizations are more common outside the southern states – Parana (PR), Santa Catarina (SC) and Rio Grande do Sul (RS). Indeed, according to Lima (2003), the realizations of 'r' have changed from apical alveolar to velar, but, in the southern states, alveolar sounds are still more frequent.

Callou and Brandão (2016: 118) argued that describing the distribution of the realizations of the 'r' in BP is complicated because of the many variants in external coda. For which reason, the authors recommended considering the hypothesis "that the varieties of Brazilian Portuguese may present different behaviors at the same stage of change". Indeed, the diachronic development of the variants in PB is often summarized as a multi-stage process that includes posteriorization, fricativization and elision: [r] > [x] > [h] > Ø (Callou *et al*. 1996), but most authors do not include 'retroflexion'. As retroflex -r is common in many localities in Brazil, Rennicke (2011, 2015) proposed that the 'r' in BP went through two different processes: fricativization, [r] > [ɾ] > [x] > [h] > Ø, and retroflexion, [r] > [ɾ] > [ɹ] > Ø.

As in the southern states the variable presents a different behavior if compared to the rest of the country, this investigation aims at explaining the diffusion of the rhotics in those states under the light of the grouping typology and to establish.

FD, XLII, Bucureşti, 2023, p. 51–72

2. DISTRIBUTION OF RHOTIC SOUNDS

2.1. Distribution according to the *ALiB*

ALiB (Cardoso *et al*. 2014) presents the data in five ranges: 1-25%; 26-50%; 51-75%; 76-99% and 100%. In the present investigation,only data from maos that show the results for the southern capitals – maps F04 C1, F04 C2, F04 C3 and F04 C4 – will be presented here. According to the cartographic results, the deletion of 'r' at the end of names represents 1-25% in Curitiba (PR) and Porto Alegre (RS) while in Florianópolis (SC) it represents 26-50% (ALiB map F04 C1). For verbs, Porto Alegre and Curitiba have 51-75% deletion and Florianópolis 76-99% deletion (ALiB map F04 C2).

In terms of realization of final 'r', according to map F04 C3 (names), in Curitiba 26-50% of the realizations are retroflex and 51-75% are taps; in Porto Alegre there is 1-25% of retroflex realizations and 76-99% of taps; in Florianópolis 51-75% of the realizations are of glottal fricatives and 26-50% of velar fricatives (Table 1 below).

In the case of verbs, map F04 C4 (verbs) shows that in Curitiba there are Â26-50% of retroflexes and 51-75%; in Florianopolis there are 1-25% of taps, 26-50% of velar fricatives and 26-50% of glottal fricatives; in Porto Alegre there are 1-25% glottal fricatives, 1-25% of velar fricatives, 1-25% of retroflexes and 76-99% taps. The following table summarizes the data (Table 1 below).

Table 1

Deletion and realization of 'r' in the three southern capitals

Deletion in names	**Deletion in verbs**	**Realization in names**	**Realization in verbs**	**State**
1-25%	51-75%	**h** - ∅ **x** - ∅ **ɾ** - 26-50% **ɾ-** 51-75%	**h** - ∅ **x** - ∅ **ɾ** - 26-50% **ɾ** -51-75%	Curitiba
26-50%	76-99%	**h** - 51-75% **x** - 26-50% **ɾ** - ∅ **ɾ-** ∅	**h** - 26-50% **x** - 26-50% **ɾ** - ∅ **ɾ** - 1-25%	Florianópolis
1-25%	51-75%	**h** - ∅ **x** - ∅ **ɾ** - 1-25% **ɾ** - 76-99%	**h** - 1-25% **x** - 1-25% **ɾ** - 1-25% **ɾ** - 76-99%	Porto Alegre

Thereby, in coda position in names and verbs, the retroflex and the tap are particular to Curitiba and Porto Alegre, while the fricatives are unique to Florianópolis. The deletion frequency demonstrates that Curitiba and Porto Alegre have similar patterns. Accordingly, there is a parallel between Porto Alegre and Curitiba and a differentiation with respect to Florianópolis.

2.1.1. Oliveira (2018)

Oliveira (2018) examined the realization of 'r' in final coda in six municipalities of the southern region: Santa Maria (RS), Caçapava do Sul (RS), Lages (SC), Criciúma (SC), Guarapuava (PR) and Campo Mourão (PR). It considered 24 interviews (four from each municipality) from the ALiB project (recorded from 2003 and 2009).

The study showed that, in the inland south, deletion of final 'r' can reach 92% in verbs and 11% in non-verbs. In Rio Grande do Sul, deletion in Santa Maria was 95% in verbs and 16% in non-verbs and, in Caçapava do Sul, 89% in verbs and 8% in non-verbs. In Santa Catarina, deletion in Criciúma was 97% in verbs and 22% in non-verbs and, in Lages, 87% in verbs and 6% in non-verbs. In Paraná, the percentage of deletion in Campo Mourão was 90% in verbs and 3% in non-verbs and, in Guarapuava, 94% in verbs and 11% in non-verbs. That is, while Criciúma (PR) and Santa Maria (RS) led the deletion process in both morphological classes, Campo Mourão (PR) and Caçapava do Sul (RS) had the lowest percentages of deletion.

The study also showed that in Santa Maria and Caçapava do Sul (RS) the most common realization is the tap. In Guarapuava and Campo Mourão (PR) the most common realization is the retroflex approximant. Santa Catarina presents different results in each town; in Criciúma, the retroflex is more frequent while in Lages it is the tap.

In summary, the results indicate a high frequency of deletion in verbs (between 97 and 87%) and a low frequency of deletion in non-verbs (between 3 and 22%). There is also evidence that the retroflex approximant and tap were the most used variants in both cases, while the trill occurs sporadically and the fricative variants are almost non-existent. Additionally, the distribution of the variables indicates that the retroflex dispersed from the north (PR) to the south (RS).

2.1.2. Comiotto and Margotti (2019)

The authors examined the realization of 'r' in twenty-seven bilingual communities in RS and SC using data from the *ALiB*. The investigation concluded that older men with less schooling are the ones who produce taps and that in municipalities where there is linguistic contact – such as Flores da Cunha (RS), Erechim (RS), Concórdia (SC) and São Miguel do Oeste (SC) – the trill and the tap are in variation, but the tap is the most frequent realization. In other cities – such as

Criciúma (PR) – only the velar fricative [x] is present, which is why the authors conclude that the typical mark of contact between BP and Italian dialects is absent there. They also report that there is no significant difference between the variants used by age groups and that the use of the tap is a linguistic and cultural trait of the descendants of Italian immigrants from the Veneto region.

2.1. Distribution according to the *ALERS* (Altenhofen and Klassmann 2011, Koch et al. 2011)

The results for specific target words are shown in Table 2 below (Koch et al. 2011, maps 055, 064, 070, 073, 085, 101, 144, 155, 167, 373) where the percentage of occurrences in each state is indicated.

Table 2

Occurrences by state (percentage).

ɾ	ɽ	r	x	R	State
a**r**vore 68 ve**r**de 22 po**r**ca 45 ga**r**ganta 49 pe**r**fume 78 joao-de-ba**rr**o 1 fe**rr**ão 11 te**rr**a 9 **r**ouge 8 beija-flo**r** 26 infnitive 10	a**r**vore 29 ve**r**de 70 po**r**ca 52 ga**r**ganta 51 pe**r**fume 22 - - - - - beija-flo**r** 59	- - porca 1 - - joao-de-ba**rr**o 59 fe**rr**ão 49 te**rr**a 62 **r**ouge 46 - -	- - - - - joao-de-ba**rr**o 38 fe**rr**ão 30 te**rr**a 29 **r**ouge 44 - -	- - - - - - - **r**ouge (ɹ) 2 - -	PR
a**r**vore 90 ve**r**de 95 po**r**ca 92 ga**r**ganta 97 pe**r**fume 100 joao-de-ba**rr**o 8 fe**rr**ão 12 te**rr**a 24 **r**ouge 5 beija-flo**r** 98 infinitive 8	- - - - - - - - - - - -	árvore 1 ve**r**de 1 - - - joao-de-ba**rr**o 49 fe**rr**ão 60 te**rr**a 60 **r**ouge 69 - -	- - - - - joao-de-ba**rr**o 10 fe**rr**ão 8 te**rr**a 15 **r**ouge 24 - -	- - - - - - te**rr**a 1 - **r**ouge (ɹ) 2 - -	RS
a**r**vore 65 ve**r**de 48 po**r**ca 38 ga**r**ganta 92	árvore 1 ve**r**de 5 po**r**ca 5 ga**r**ganta 8	árvore 19 ve**r**de 32 po**r**ca 28 -	- - po**r**ca 18 -	- - - -	SC

ɾ	ɽ	r	x	R	State
pe**r**fume 80 joao-de-ba**rr**o 5 fe**rr**ão 5 te**rr**a 14 **r**ouge 11 beija-flo**r** 51 infinitive 9	pe**r**fume 1 - - - - beija-flo**r** 22 -	- joao-de-ba**rr**o 78 fe**rr**ão 62 te**rr**a 64 **r**ouge 24 - -	pe**r**fume 8 joao-de-ba**rr**o 5 fe**rr**ão 21 te**rr**a 14 **r**ouge 61 - -	- - - - te**rr**a 8 **r**ouge (ɹ) 4 - -	

The data from the *ALERS* (Koch *et al.* 2011) indicate that retroflex realizations are non-existent in RS, infrequent in SC and common in PR – considering internal and external coda realizations. Also according to the *ALERS*, weak-r in inner coda is realized as [ɾ] in PR and RS, and as [ɾ, r, x] in SC. In the case of strong-r, [ɾ, r, x, ʁ] are possible in the three states, but the fricative is more common in PR.

2.2.1. *Margotti (2004)*

The work examined bilingual cities – BP and Italian – in RS and SC, separating the samples collected with bilinguals and monolinguals. The study considers the results of surveys for the *ALERS* and VARSUL. It also collects data in eight municipalities in each state – with similar sociocultural characteristics. The results indicate that there are different levels and degrees of bilingualism, but that in all the cases there is a tendency to replace the heritage language with BP.

The results show that the realizations of 'r' differ from place to place, but four variants of 'r' were recorded: the trill, the tap, the velar fricative and the alveolar approximant. The study states that since "speakers from the contact areas of BP and Italian have difficulty reproducing the opposition that exists in Portuguese" (Margotti 2004: 156), an intermediate sound – the alveolar approximant – indicates that speakers perceive the difference between tap and trill, but cannot reproduce it. Furthermore, the study states that, the substitution of the trill by the tap is a phenomenon that also happens in regions of Azorean and German immigration.

Therefore, since the studies described above indicate that the retroflex is more common in PR and less common in RS, it may be concluded that the diffusion in the southern states started from PR. Second, the studies indicate that SC has a different behavior from the other southern states, not only regarding realizations in the capital but also regarding those in inland towns. As such, fricative realizations in general are more common in that state. Lastly, the investigations demonstrate that the tap distinguishes places that have Italian influence. That is, even though there are coincidences among the three states, RS has its particularities, which will be investigated in the rest of the article.

3. METHODOLOGY

The data, used to generate distribution maps, was taken from the *ALERS* and corresponds to 92 towns in RS. The evaluated words were *arvore, verde, carta, corda, fervendo, porca, gordura, garganta, perfume* (examples of internal coda), *calor, beija-flor, revolver* (examples of final coda), *genro, caro, clara* (examples of internal onset), *joao-de-barro, terra, carro* (examples of 'strong-r'), *revolver* (word onset) – taken from the maps 055, 064, 070, 073, 101, 144, 167, 373 (Altenhofen and Klassmann 2011) and maps 44, 45, 46, 47, 48, 49, 50, 51, 52, 53 (Koch et al. 2011).

Gabmap, a web application that visualizes language variation, was used to generate the maps and analyze the data (see Nerbonne and Kleiweg 2005; Snoek 2014; Leinonen *et al*. 2016). The dialect data was prepared in two spreadsheets, one indicating the pronunciation of the word in the International Phonetic Alphabet (IPA) and one in which only the pronunciation of the target variable was represented using the IPA. In order to create the distribution maps, the data files were accompanied by a map file with the geographical coordinates of the data sites (created in Google Earth). Those maps were used to have a rough overview of the distribution of the dialects. The clusters that emerged were examined using the Multidimensional scaling (MDS), a technique for representing distances between objects in which the locations are represented as circles of matching color in a plot. The beam maps, where the strength of association between the two locations is encoded in colored lines, were used to examine the linguistic distance between neighboring towns. Lastly, to interpret the distribution, the beam maps and the distribution maps were weighted against the results of several studies carried out in RS.

4. DISCUSSION

4.1. The maps

Since the objective was to understand the distribution of 'r' in RS and to investigate the evolution of the realizations, the first step was to generate a map showing the distribution of the variable where only the target sound was phonetically transcribed – i.e. *carro* (car) could be ca[r]o, ca[ʁ]o, ca[x]o, ca[ɾ]o – no other differences were registered.

Figure 1 below shows that there are 4 clusters in the state, identified by the different colors. Noticeably, the clusters fall flat in terms of geographic distances because "the fundamental dialectological postulate that nearby varieties are normally more similar linguistically that distant ones" (Nerbonne 2010: 8) does not hold.

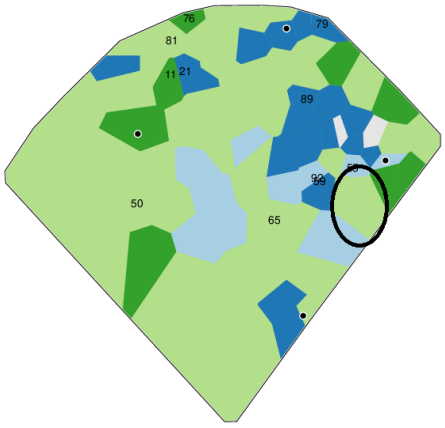

Figure 1. *Distribution of the realizations of 'r'.*

As an example it may be argued that some municipalities to the north of the state (dark blue) are more similar to distant places than to neighboring towns (light green). That is, villages and towns may belong to the same cluster even when far away. In turn, neighboring towns may not belong to the same cluster. Furthermore, it is clear that places that are at regular distances from Porto Alegre (circled on the map) – which is considered the reference site since it is the capital city – do not show similar results in terms of diffusion.

Since the distribution above was considered flawed, a second analysis was undertaken from the perspective of the aggregate analysis (see Nerbonne *et al.* 2008; Nerbonne 2010). The complete phonetic transcription of the same eighteen words was evaluated this time – car could be [karɷ], [karɹo], [karɷ], [kaʁɷ], [karo], [kaxɷ], etc. Figure 2 below shows the distribution of the clusters.

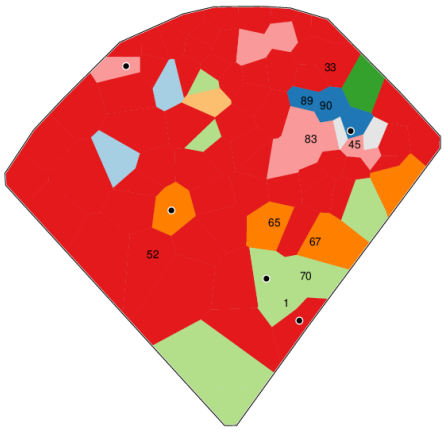

Figure 2. *Distribution by aggregate perspective*.

There are 8 clusters, but their distribution does not seem correct still. According to the new configuration, neighboring towns continue to be very different among each other, at the time that they are similar to geographically distant ones. This configuration indicates, for example, a cluster (in pink) composed of locations that are proximate but that is also composed of the areas which are more than 200 km away and not even proximate to each other.

That is, the two possible distributions (Figures 1 and 2) seem problematic since the distribution of variation does not represent a function of geographical distance. Since Nerbonne *et al*. (2008: 648) argue that "the research community is convinced that the linguistic varieties are hierarchically organized," the case in question would demonstrate that the data are inaccurate. For which reason, a further study contemplating more data from the atlases will be carried out. Yet, a second possibility is that the analysis needs to be supplemented by taking into account the grouping typology as proposed by Razky (2010a, 2010b).

4.2. The groupings as proposed by Razky (2010a, 2010b)

Traditional Dialectology methodologies trace isoglosses lines drawn following the unidimensional aspect as implemented by Gilliéron and Edmont when preparing *L'Atlas Linguistique de la France* (ALF) in the second half of the 19th century and corroborated by recent studies that have proven that geography is determinant in delineating varieties. Huisman *et al.* (2019: 14), for example, determined that it is important to consider the geographical configuration of a language area to explain language variation and change. Nerbonne and Heeringa (2007: 8) demonstrated that varieties that are closer (or larger) should resemble one another (2006: 12) and that "synchronic differences should reflect historical dynamics". However, Nerbonne (2010: 3827–3828) discovered that "linguistic contact operationalized through geography can account for about one-quarter of the aggregate linguistic variation we find in large collections such as dialect atlases". That is, the studies indicate that geographical configuration is more relevant to linguistic variation than language contact.

Yet, the use of models of diffusion like those that postulate that a given linguistic change is spread in waves from a given center (Schmidt 1872) or that innovations spread from a large population center directly to another similarly sized one (Trudgill 1974), may only be partially applied in the Brazilian context. The pluridimensional aspect, as stipulated by Radtke and Thun (1996) and Thun (1998, 2000a, 2000b, 2005) at the end of the 20th century, may be used to better represent dialectal distribution in Brazil. That methodology was, of late, enhanced by the notion of grouping as theorized by Razky (2010a, 2010b; Razky and Guedes 2015) and has been found to be the best approach to examine cartographic results in Brazil.

According to Razky (2010a, 2010b), the geographic space is conceived as part of the dialect continuum by analyzing geographic spaces beyond geopolitical

boundaries. For that reason, instead of tracing isoglosses, intersecting lines are drawn to show phonetic or lexical contours with more complex diatopic or diastratic configurations. This approach was found to be necessary, for example, to better map the complexity of the lexical continuum when Razky and Guedes (2015) examined the lexical distibution in Pará, Brazil.

The methodology is relevant in Brazil because the use of isoglosses cannot depict the great societal changes that have taken place since 1950 – which include great population movements, due to the development of the transportation sector and the decentralization of job opportunities. For that reason, a significant linguistic variability may be seen in the same geographic space in terms of grouping, a complex situation that could not be explained by using the traditional approaches (Razky 2010a, 2010b; Razky *et al*. 2019: 4; Rasky and Santos 2020). Additionally, the typology allows the examination of data belonging to a non-homogeneous lexical continuum and to reflect the historical characteristics of diachronic isoglosses that came into contact with variants that spread both, horizontally (geographically) and vertically (socially) (Sanches and Razky 2021: 14).

The grouping seeks to fill the theoretical-methodological gap with regard to social changes that directly reflect on complex diatopic configurations and that require a typology capable of responding to the complexity of variation in Brazil. The typology comprises macrogroupings, microgroupings, nanogroupings and supragroupings (Razky *et al*. 2019: 4). A macrogrouping, refers to a linguistic variant present in a set of locations belonging to more than one microregion in a state, at the time that a variant may occur only in a set of localities within a microregion; a nanogrouping is defined by the presence of a variant in a few localities belonging to a microregion. A supra-grouping, in turn, is a manipulation of the set of variants of a locality by selecting only the most productive variant.

It was considered that the grouping typology would be an important ally when explaining the distribution of the variants in RS. First, because the linguistic results call for a methodology appropriate to the present linguistic situation. Second, because the socio-history of the state shows that the occupation took place differently in different locations.

According to Albuquerque (1954: 44-47) the distance from southern RS to Rio de Janeiro made the influence from Rio de la Plata greater in the RS-Uruguayan border until the mid-1950s. This also meant that, until the mid-1950s, the center of influence for the border towns was Porto Alegre (Albuquerque 1954: 44-47). Consequently, equidistant areas from Porto Alegre could differ linguistically depending on being to the south or north of RS.

Another reason that justifies the use of the typology proposed by Razky (2010a, 2010b) is that not only the geographic configuration, but also settlement characteristics resulted in places that may be geographically close, but that did not have frequent contact with each other. As mentioned in the previous paragraph, the southern and western areas were under Spanish influence while the highlands were

influenced by European immigrants (Italians, Germans, and Polish). European immigration began on the eve of independence and, after 1888, immigration to the South increased remarkably. For example, between 1884 and 1894 Italian immigration reached its peak, with around 60 thousand immigrants. Today it is estimated that around 27% of the population of RS has Italian ancestry.

The settlement of the immigrants happened in 'colonies' – 'colony' is the name the descendants give to the communities. The colonies, which were self-sufficient, were founded by the different ethnic groups and remained isolated for a long time.

Surely, sociolinguistic studies demonstrate that ethnicity is relevant when describing realizations in RS. For instance, Bisol (1981) showed there was a difference among the realizations of vowels in Porto Alegre, border towns, places of Italian origin, and of German origin.

Accordingly, since the gathered data reveals a complex dialectal continuum and the socio-history describes an intricate migratory situation, it was considered necessary to use dialectal typology to better understand the dynamics of variation in RS. It is expected that in this way each grouping may be correctly identified in the cartography and dynamics of the variants better understood.

As a matter of fact, if the data are separated according to the environment in which 'r' occurs (internal coda, onset, etc) and are then examined considering the socio-history of the locality, the interpretation of the occurrence of variants becomes more likely.

For example, a typical pronunciation of the word carro by speakers who have Italian as L2 is ca[ɾ]o a variant that does not happen among L1 speakers of BP. Figure 3 below shows the distribution of the variant in the state. As expected, the distribution of ca[ɾ]o corresponds to the places where Italian is L2.

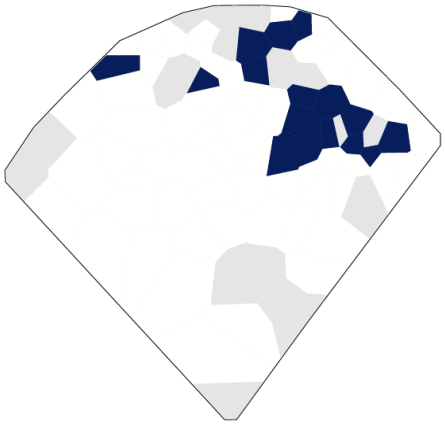

Figure 3. *Distribution of ca[ɾ]o.*

Another characteristic of Italian as L2 is the realization of [ɾ] in word onset. Figure 4 below shows the distribution of the word [ɾ]evolver (gun) in RS. Again, the distribution matches the disposition of Italian settlements.

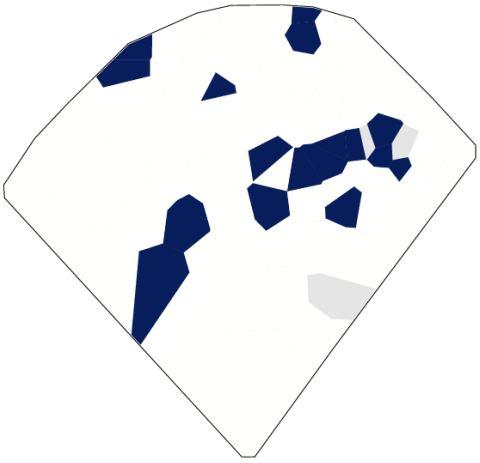

Figure 4. *Distribution of [ɾ]evolver*.

Even when Figures 3 and 4 are not identical, both represent places where the Italians have left marks in the most diverse domains. At the same time, it becomes evident that the distribution of the variants in isoglosses will present a limited representation of the distribution of variants in the state. Since the representation of variants in isoglosses seems flawed, the grouping typology as proposed by Razky *et al*. (2019) and Razky and Coimbra (2020) will be used to interpret the data.

The initial concept of grouping (Razky 1998), which was revised by Guedes *et al*. (2018) and, later by Razky *et al*. (2019) and Razky and Coimbra (2020), describes five linguistic groupings as follows:

(1) Nanogrouping: a variant that ranges from 1 to 10% of the investigated area;
(2) Microgrouping: variants that cover between 11% and 25% of the total space;
(3) Mesogrouping: variants that cover between 26% and 50% of the total space;
(4) Marcogrouping: variants that are present in 51% to 100% of the investigated territory;
(5) Supra-grouping: stems from an intervention in the statistical result. For example, only variants registered in a certain percentage in one or several diatopic spaces are considered.

After considering the results of the works works in section 1 and the distribution of the groupings in Figures 1, 2, 3 and 4, it was possible to conjecture about the existence of the following groupings in RS:

(1) Two macrogroupings comprising: the bilingual region; monolingual region;
(2) Four mesogroupings comprising: internal monolingual; border monolingual; bilingual rural communities; bilingual urban communities;
(3) Microgroupings: locations that show other kinds of influence (for example, from the drovers).

As such, there would be two macrogroupings: the bilingual regions (of European settlement) and the monolingual regions. Those macrogroupings would be composed of two other groupings each. The monolingual macrogrouping would include a mesogrouping of internal areas and one of border areas. The bilingual macrogrouping would be divided into rural and urban communities.

Dividing RS into bilingual and monolingual communities seems straightforward. For instance, opposite to what Camara Jr. (2004) and Cagliari (1982, 2007) point out for BP, i.e. that the tap never occurs in word-initial position, Monaretto (1992, 1997) reports that, in RS, bilinguals substitute [r] for [ɾ]. Also, the existence of two monolingual regions is justified, as there are several works that suggest that speakers in border areas understand and are able to interact with Spanish speakers (Carvalho 2004, 2007, 2022; Meirelles 2006, 2011) and that the movement of inhabitants on either side of the border implies language contact, but no stable bilingualism (cf. Carvalho 2022). Lastly, the microgroupings would correspond to places where the variant [ɽ] suggests the influence of the drovers (who transported goods and cattle from RS to São Paulo).

The configuration described above predicts that towns belonging to the same mesogrouping would present similar linguistic behavior, and that neighboring towns would have different realizations when they belong to different mesogroupings. At the same time, it predicts that bilingual towns would behave similarly even when they are distant. Indeed, when the distribution of the groupings in Figure 2 (section 4.1) is compared to the map that shows the distribution of bilingual informants (Koch *et al.* 2011), there is a close match. The comparison reveals that the larger area corresponds mostly to BP realizations and that the smaller areas correspond to places where the informants were bilingual.

However, the distribution of the groupings in Figures 1 and 2 is still not optimal because it fails when it presents towns that are geographically near in separate groupings even when they have had the same linguistic influences. For those reasons and to better test the validity of each grouping, beam maps that show the linguistic distances to the neighboring towns were generated using Gabmap.

4.3. The correlations between geographic and linguistic distance

Figure 6 below represents the linguistic distances when considering only the realization of 'r'.

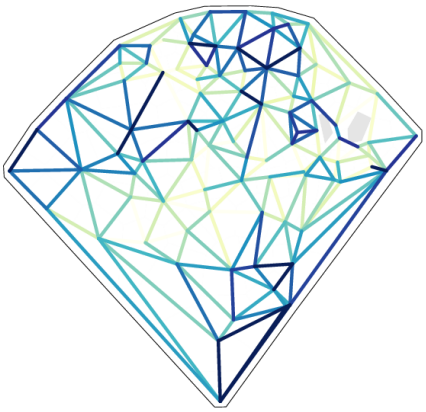

Figure 6. *Beam map for the 'r' variable.*

As already expected and contrary to what Nerbonne and Kleiweg (2005: 10) proposed, the "geographically proximate varieties" in RS do not regularly "tend to be more similar than distant ones". Additionally, Local Incoherence was 4.04, which means that the data represents poor measurements or that the correlation between linguistic and geographic distances does not hold locally. Indeed, the Figure above shows that some areas, mainly those to the northwest, are proximate linguistically even though they have different origins. An interesting fact to consider at this point is that all the municipalities were part of the trials followed by drovers on their way from São Paulo to Colonia del Sacramento (Uruguay). In other words, it seems that the influence of the drovers was initially under-estimated.

On the other hand, the lines that connect the towns to the southeast demonstrate that all those towns are linguistically proximate. In this case, the macrogrouping proves accurate because all those towns have a Portuguese origin and are monolingual. Yet, one of the places is a border town, which means it should behave differently.

Therefore, since the correlation between linguistic distance and geographic distance seemed far from optimal when considering only the realizations of 'r', the same correlation was explored from the aggregate perspective. Figure 7 shows that there is less correspondence between geographic and linguistic distance now than when considering only the variants of 'r'. In fact, the beams indicate that only a few neighboring towns are linguistically similar. That is, the distribution seems to be flawed.

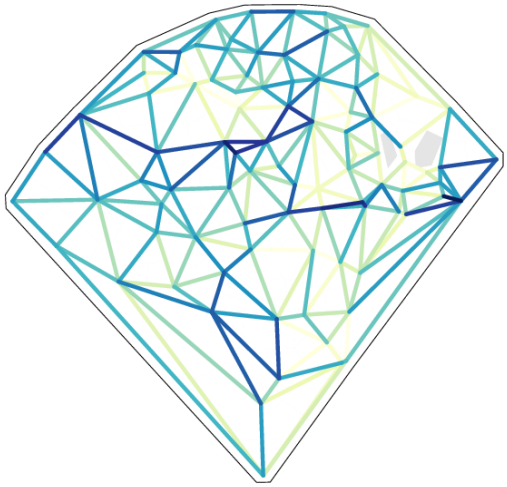

Figure 7. *Beam map of aggregate data.*

Based on the typology advanced by Razky (2010a, 2010b), Guedes (2020: 104) proposed the inclusion of another factor in the list of extralinguistic variables mapped in pluridimensional linguistic atlases, which he called the diaethnic variable. When examining data from communities where different ethnic groups were in contact, the author realized that the data could not be explained by applying common variables – such as: diatopic or dialinguistic. He was only able to explain variation by considering the speaker's ethnicity (Guedes 2020: 105). The author explains that the diaethnic variable differs from the dialinguistic variable because it does not represent the variation that occurs in the case of language contact, but the one that represents "the influences of ethnicity on linguistic variation, both within and outside situations of language and dialect contact" (Guedes 2020: 105).

There are many reasons why the adoption of a diaethnic variable in the description of the occurrences in Rio Grande do Sul seems pertinent. Societal bilingualism denotes the situation in a particular speech community in which more than one language is used. Nevertheless, in the Brazilian South the situation is not as straightforward as the definition may suggest. When investigating the varieties of 'r' in SC and RS, Margotti (2004) recognized that there are different levels and degrees of bilingualism. For example, today, in major urban centers, the immigrants' languages have already given way to BP, in smaller urban centers their language is losing ground while in rural communities, the European dialects are still spoken (Margotti 2004).

As mentioned before, the original settlement may be described as 'concentration in colonies'. Furthermore, in 1938, the Brazilian government implemented a series of restrictions that were part of its nationalization campaign, one of which banned foreign languages in schools. The situation became dreadful

in 1942, when German descendants were specially targeted and the government considered that maintaining the 'colonies' was problematic. Immigrants and their descendants who did not speak Portuguese could be taken as prisoners and part of the immigrants' memory was destroyed. Nonetheless, as in most cases of discrimination, those actions strengthened the bonds within the 'colonies'.

In fact, Mombach (2012) concludes that German culture, for example, was not erased while Fritzen and Ewald (2013) mention that during that period the individuals learned how to read and write in German with members of the family, in the Evangelical Church of the Lutheran Confession, or with other community members. More importantly, the authors claim that even though school was no longer the literacy agent, other strategies were adopted to continue to use the language (Fritzen and Ewald 2013: 253).

Currently, due to socio-political influences, many individuals are abandoning the use of their L1 and adopting BP even at home – which indicates a language shift (Horst *et al.*, 2019) or the foreign language is being maintained only in rural areas (Fraga, 2009; Battisti and Martins, 2011; Corrêa, 2017; Comiotto and Margotti, 2019; Horst *et al*., 2019). Yet, Ortale *et al*. (2015) suggest that in many cases, the two languages have clearly defined contexts of use: at home and in exchanges with Brazilians.

Other reasons belong to the linguistic domain. Brescancini and Monaretto (2008) suggest that the different realizations or 'r' are characteristic of the different ethnicities present in the southern states. Battisti and Martins (2011) conclude the use of the tap instead of the trill in the BP spoken in Flores da Cunha represents what they call an "unexpected use of Portuguese, but common to communities where linguistic contact was or has been common" (Battisti and Martins 2011: 147–148).

Velho (2017) investigated the realizations of 'r' in three bilingual municipalities in RS to verify whether it is possible to delimit a speech community in terms of the production of 'r'. The study concluded that despite the "shared Italianness", "different characteristics that correspond to the boundaries of the villages need to be addressed in the analysis of this variable" (Velho 2017: 318). Indeed, Corrêa (2017) examined 'r' in Antônio Prado and concluded that the stabilization between [r] and [ɾ] is a consequence of "an effort by residents to maintain or rescue dialectal speech and transmit it to children, ensuring identification with the Italian ethnicity" (Corrêa 2017: 102).

Fraga (2006, 2008) reports that bilingual speakers (Dutch/BP) perceive the rhotic as a retroflex approximant, an alveolar approximant or an alveolar retroflex approximant in Carambeí (PR). Fraga (2009) observed that there is a connection between identity and realization of 'r' because individuals who identify themselves as Dutch prefer the tap, while those who identify themselves as Brazilians prefer fricatives.

It is cogitated that one of the reasons for the lack of correspondence is that even though there is a coincidence within the bilingual and non-bilingual regions when it corresponds to the most frequent realizations – for example, in the latter 'r' in onset may be [r] or a fricative while in the former the most frequent realization is the tap – if different cities of Italian influence (or mesogroupings) are investigated in depth, the results demonstrate that the distribution may be different from one location to another (Brescancini and Monaretto 2008; Battisti and Martins 2011; Corrêa 2017; Velho 2017; Comiotto and Margotti 2019). For that reason the groupings were reexamined and a new distribution proposed as follows:

(1) Two macrogroupings comprising: the bilingual regions and the monolingual regions;
(2) Four mesogrouping comprising:internal monolingual; border monolingual; bilingual rural communities; bilingual urban communities;
(3) Two microgroupings comrpising: bilingual urban communities (-BP); bilingual urban communities (+BP).

The microgroupings would now correspond to the bilingual urban communities characterized by their ethnic and cultural practices. For example, Antônio Prado would be part of the bilingual urban community that is more distant form BP (-BP) while Flores da Cunha would belong to the bilingual urban community that is more akin to BP (+BP) (cf. Corrêa 2017; Velho 2017).

That way, the dynamics would reflect local identities with regard to ethnic origin and would allow us to better understand the formation of BP in the state. In the main, by carefully characterizing the groupings, the studies could yield a high-level understanding of the pervasiveness of variable patterns for both monolingual and bilingual communities.

However, there is still a factor to be discussed: the influence of the drovers. In the revised version above the influence of the drovers, needs to be contemplated as a possibility because some localities that apparently do not have the same linguistic background, show linguistic proximity (Figures 1, 2, 3 and 4).

From the end of the 17th century to the beginning of the 20th century, there was a wide network of trails and locations that served as a landing place for the drovers who were responsible not only for transporting goods, but were also agents of territorial integration. Domingues (1999) holds that the drovers were "directly related to the Brazilian settlement" and to the consolidation of the borderlines between Brazil and the Spanish territories. In fact, Koch, Klassmann and Altenhofen (2011: 22) claim that the Caminho dos Conventos (1728) and Caminho das Tropas (1731) guaranteed "the linking of the southern lands to the Portuguese colonial heritage".

The first route was the Caminho dos Conventos, but over time other routes were opened. The main and most traveled route was Caminho do Viamão, which

started in Uruguaiana, and continued through Alegrete to Vacaria (see Oliveira 2009). Figure 8 below shows the trail that connected Minas Gerais and São Paulo to the Brazilian south. The trail was approximately 4000 kilometers long and it was normally traveled in stretches of about 30 kilometers a day.

The drovers needed to rest and to fatten up cattle along the route, for periods that varied depending on the weather conditions and the conditions of the cattle – sometimes reaching up to 6 months. Those resting places gradually become villages, towns, and cities that were influenced by the customs of the travelers (cf. Silva Neto 2018: 31-40). After 1890 and as a result of the arrival of the train, many cattlemen from RS continued to travel with "the so-called 'loose troops'" (Domingues 1999: 282). Those travelers spoke a variety of BP called hillbilly BP.

In the first years of the 20th century, many studies started to describe what the so-called hillbilly BP. Amaral was one of the first to describe the inter- and post-vocalic 'r' produced by the countrymen or hicks as having "a peculiar value" because "instead of projecting the tip against the alveolar ridge, as it is produced in Portuguese, the tongue […] turns upwards, without touching the palate" (1955: 47). Melo described that in the South of Minas Gerais, northern São Paulo, Pernambuco, Alagoas, and Goias there was a voiceless 'r' similar to the Northamerican 'r'. Rodrigues (1974) registered [ɾ, ɽ, ɦ] as realizations of /ɾ/ in Piracicaba (São Paulo) in the 1970s.

The cartographic results from the *ALiB* reveal that Curitiba and Porto Alegre have [ɽ]. Also, the data from the *ALiB* analyzed by Oliveira (2018) showed that the two inland towns in RS produced mostly the tap, that the two in PR produced mostly the retroflex and that one of the towns in SC produced the retroflex and the other produced mostly the tap. Lastly, the examined data from the *ALERS* demonstrates that retroflex realizations are almost non-existent in RS, infrequent in SC and common in PR. That is, it is possible to deduce that the diffusion of the retroflex was from north to the south. Indeed, when contemplating the distribution of the towns along the trail, it becomes evident that there is a compatibility between the sites and the present distribution of retroflex-r.

Accordingly, the grouping for RS needs to be reconsidered, weighting that the variant ranges from 1 to 10% of the investigated area. Then, it will not conform to a microgrouping, but to a nanogrouping. In summary, the resulting grouping for RS, in terms of the distribution of 'r', will be:

(1) Two macrogroupings comprising: the bilingual regions and the monolingual regions;
(2) Four mesogroupings comprising: internal monolingual; border monolingual; bilingual rural communities; bilingual urban communities;
(3) Two microgroupings comprising: bilingual urban communities (-BP); Bilingual urban communities (+BP);
(4) One nanogrouping: places along the trails traveled by drovers.

To elaborate, there are 2 macrogroupings in RS that correspond to bilingual and monolingual regions. Each macrogrouping is composed of 2 mesogroupings. The monolingual macrogrouping is composed of the internal and border mesogroupings while the bilingual macrogrouping is made up by the urban and the rural mesogroupings. There are two microgroupings: one of bilingual urban communities that are more distant from BP (-BP) and another one of communities that show realizations to BP (+BP). Lastly, there is a nanogrouping composed by localities that show influence of the drovers.

One of the objectives of the study was to establish guidelines for research that investigates the situation from a perspective different from the generalizing approaches that are commonly applied when examining linguistic variables in RS. As such, it proposes that future research should consider the groupings in RS not only for analyzing the data but also for collecting data that truly represents the investigated localities.

5. FINAL WORDS

The present study postulates that the many variants present in RS, at least, are due to two phenomena: (1) there are two different paths for the evolution of ‘r’ and (2) the geographic distance from one location to another may not be the only determinant factor. That means that even when two locations are proximate, they may perform differently and that neighboring towns may present different variants.

Based on the argumentation presented, this study suggests that when examining the groups in RS the geographic distance a should be considered to ponder the connections the different towns have with one another, but that the grouping typology is necessary to validate the different groupings and understand the relation that exists among them.

In summary, this study intended to understand how the diffusion of ‘r’ took place in RS and to indicate the possibilities for future research. This work demonstrated that Gabmap can help identify the different varieties of ‘r’ used in RS, demonstrate which neighboring places have similar realizations, and delimit the extent to which the influence of a given variety extends. At the same time, it is useful to identify varieties that, although geographically distant, are linguistically close.

The future calls for a deepening in the distinction of varieties with more data for the examined locations. For that reason, a more detailed study is needed, using more data from the atlas to categorically validate the grouping as proposed in this article.

REFERENCES

Aguilera and Romano 1998 = Vanderci de Andrade Aguilera, Valter Pereira Romano, *A Geolinguística no Brasil: caminhos e perspectivas*, Londrina, Eduel.

Albuquerque 1955 = Arcy Tenório Albuquerque, *Gauchismos- a linguagem do Rio Grande do Sul*, Porto Alegre, Livraria Sulina.

Altenhofen and Klassmann 2011 = Cleo Altenhofen, Mario Klassmann, *ALERS: Atlas Lingüístico-Etnográfico da Região Sul do Brasil*. Porto Alegre, Florianopolis, Ed. UFRGS, Ed. UFSC.

Amaral 1920 = Amadeus Amaral, *O Dialeto Caipira*. Sao Paulo, O livro.

Battisti and Martins 2011 = Elisa Battisti, Luisa Martins, *A realização variável de vibrante simples em lugar de múltipla no português falado em Flores da Cunha (RS): Mudanças Sociais e Linguísticas*, in "Cadernos do IL" 42, p. 146–158. https://doi.org/10.22456/2236-6385.26018

Bisol 1981 = Leda Bisol, *Harmonia Vocálica: Uma regra variável*, Doctoral dissertation. UFRJ.

Brescancini and Monaretto 2008 = Claudia Brescancini, Valeira Monaretto, *Os róticos no sul do Brasil: panorama e generalizações*, "SIGNUM: Estudos Linguísticos", p. 51–66. https://doi.org/10.5433/2237-4876.2008v11n2p51

Cagliari 2007 = Luiz Carlos Cagliari, *Elementos de Fonética do Português Brasileiro*, Sao Paulo, Paulistana.

Cagliari 1982 = Luiz Carlos Cagliari, *Elementos de Fonética do Português Brasileiro*. Thesis (Licentiate's degree), Unicamp.

Callou and Brandão 2016 = Dinah Callou, Silvia Brandão, *Caracterização de Áreas Dialetais no Português do Brasil: Análise de duas variáveis*. in, Lucrécio Araújo de Sá Júnior, Marco Antonio Martins (2016), p. 97–119. https://doi.org/10.5151/9788580391824-05

Callou *et al*. 1996 = Dinah Callou, Yonne Leite, Joao Moraes, *Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil*, in: Ingedore Koch (1996), *Gramática do português falado*, p. 465–493.

Cardoso *et al*. 2014 = Suzana Alice Marcelino da Silva Cardoso, Jacyra Andrade Mota, Vanderci de Andrade Aguilera, Maria do Socorro Silva de Aragão, Aparecida Negri Isquerdo, Abdelhak Razky, Felício Wessling Margotti, Cléo Vilson Altenhofen, *Atlas Linguístico do Brasil*. Londirna, EDUEL.

Comiotto and Margotti 2019 = Ariela Fátima Comiotto, Felício Wessling Margotti, *Uso dos róticos do português em contato com os dialetos italianos*, "Acta Sci. Lang. Cult." 41, e48857. https://doi.org/10.4025/actascilangcult.v41i2.48857

Corrêa 2017 = Rachel Corrêa, *Estigmatização em Antônio Prado (RS): ilustrando as sete consequências de atitudes linguísticas negativas de Grosjean (2001) através de trechos de entrevistas sociolinguísticas*, "Entrepalavras" 7, p. 91–104. https://doi.org/10.22168/2237-6321.7.7.1.91-104

Domingues 1999 = Julio Manoel Doningues, *Tropeirismo*. "Seminário Paulista de Estudos Tropeiros", p. 277–286. http://porangabasuahistoria.com/wp-content/themes/Porangaba/upload/downloads/tropeirismo.pdf

Fraga 2009 = Leticia Fraga, *A variedade de português falada pelos "holandeses" de Carambeí: identidade e r-forte*, "Revista brasileira de linguística aplicada" 9, p. 621–645. https://doi.org/10.1590/S1984-63982009000200013

Fraga 2008 = Leticia Fraga, *O português falado por descendentes de holandeses em Carambeí (PR) e os róticos em coda*, "Linguagem e Ensino" 11, p. 349–376. https://periodicos-old.ufpel.edu.br/ojs2/index.php/rle/article/view/15687

Fraga 2006 = Leticia Fraga, "O róticos no portugues de Carambeí/PR", "Estudos Linguísticos" 35, p. 1113–1122.

Fritzen and Ewald 2013 = Maristela Pereira Fritzen, Luana Ewald, *Aqui somos protegidos pelas nossas quatro paredes. Aqui nós falamos alemão: histórias de letramentos interculturais no Vale do Itajaí, SC*, "Trab. linguist. apl" 52, p. 239–258. https://doi.org/10.1590/S0103-18132013000200004

Gãrtner *et al*. Eberhard Gãrtner, Christine Hundt, Axel Schõnberger (eds.), *Estudos de geolingüística do português americano*, p. 185-227.

Guedes 2020 = Regis J. C. Guedes, *Variável diaétnica: repensando a variação geolinguística pluridimensional contatual*. "Revista Brasileira de Linguística Antropológica" 12(1), p. 101–116. https://doi.org/10.26512/rbla.v12i1.29735

Guedes *et al*. 2018. = Regis José da Cunha Guedes, Abdelhak Razky, Eliane Oliveira da Costa, *A Pesquisa Geolinguística em Áreas Indígenas Brasileiras: desafios e estratégias*, "Signum: Estudos da Linguagem" 21(1), 126–138. https://doi.org/10.5433/2237-4876.2018v21n1p126

Huisman *et al*. 2019 = John L. A. Huisman, Asifa Majid, Roeland van Hout, *The geographical configuration of a language area influences linguistic diversity*, "PLoS ONE" 14(6), e0217363. https://doi.org/10.1371/journal.pone.0217363

Koch *et al*. 2011 = Walter Koch, Cléo Vilson Altenhofen, Mário S. Klassmann, *Atlas Lingüístico-Etnográfico da Região Sul do Brasil (ALERS): Introdução, Cartas fonéticas e morfossintáticas*, Porto Alegre, Florianopolis, Ed. UFRGS, Ed. UFSC.

Koch 1996 = Ingedore Koch, *Gramática do português falado*. Sao Paulo, UNICAMP.

Leinonen *et al*. 2016 = Therese Leinonen, Çağrı Çöltekin, John Nerbonne, *Using Gabmap*, "Lingua" 178, p. 71–83. https://doi.org/10.1016/j.lingua.2015.02.004

Lima 2003 = Alcides Fernandes de Lima, *A variação do (r) Posvocálico em Cametá-PA: uma Abordagem Geo-sociolingüística*, Doctoral dissertation, UFPA.

Margotti 2004 = Felício W. Margott, *Difusão sócio-geográfica do português em contato com o italiano no sul do Brasil*, UFRGS. http://hdl.handle.net/10183/3911

Meirelles 2011 = Virginia Meirelles, *Elementos de fonética do português falado no Rio Grande do Sul*, Doctoral dissertation, University of Brasília. https://repositorio.unb.br/handle/10482/8701

Meirelles 2006 = Virginia Meirelles, *Phonetic-phonological aspects of contact between Portuguese and Spanish in the city of Sant'ana do Livramento-Rivera*, Master's thesis, University of Brasilia. https://repositorio.unb.br/handle/10482/3284

Melo 1971 = Gladstone Ch. Melo, *A Língua do Brasil*. Sao Paulo, Fundação Getulio Vargas.

Mombach 2012 = Clarissa Mombach, *O governo Vargas e suas implicações na produção literária teuto-brasileira*, "Literatura e Autoritarismo" 10, p. 31–44. https://doi.org/10.5902/1679849X75392

Monaretto 1997 = Valeria Monaretto, *Um reestudo da vibrante: análise variacionista e fonológica*. Doctoral dissertation, PUC-RS.

Monaretto 1992 = Valeria Monaretto, *A vibrante: representação e análise sociolinguística*. Master's thesis, UFRGS.

Nerbonne 2010 = John Nerbonne, *Measuring the diffusion of linguistic change, "*Phil. Trans. R. Soc." B365, p. 3821–3828. https://doi.org/10.1098/rstb.2010.0048

Nerbonne and Heeringa 2007 = John Nerbonne, Wilbert Heeringa, *Geographic Distributions of Linguistic Variation Reflect Dynamics of Differentiation*, "Studies in generative grammar", p. 267–298. https://doi.org/10.1515/9783110198621.267

Oliveira 2018 = Ingrid da Costa Oliveira, *Os Róticos em Coda Silábica e Externa: O Interior da Região Sul no Projeto ALiB*, Master's thesis, UFRJ.

Oliveira 2009 = Maria E. R. de Oliveira, *Os Tropeiros*. Parana, Secretaria de Estado de Educação. http://www.diaadiaeducacao.pr.gov.br/portals/cadernospde/pdebusca/producoes_pde/2009_uepg_historia_md_maria_elizabeth_ribeiro_de_oliveira.pdf

Ortale *et al*. 2015 = Fernanda Ortale, Giliola Maggio, Paola Baccin, *Identidade e bilinguismo em contexto de núcleo familiar de imigrantes italianos, "*Revista de Italianística" 30, p. 146–163. https://doi.org/10.11606/issn.2238-8281.v0i30p146-163

Paixão Côrtes 1999 = João Carlos D'Ávila Paixão Côrtes, *Danças Birivas do Tropeirismo Gaúcho*. Porto Alegre, CORAG.

Radtke and Thun 1996 = Edgar Radtke, Harald Thun, *Nuevos caminos de la geolinguística románica. Un balance*, in Radtke and Thun (1996), *Neue Wege der Romanischen Geolinguistik*, p. 25–49.

Radtke and Thun 1996 = Edgar Radtke, Harald Thun, *Neue Wege der Romanischen Geolinguistik*. Kiel, Westensee-Verlag.

Razky 1998 = Abdelhak Razky, *O Atlas Geo-sociolinguístico do Pará: abordagem metodológica,* in Aguilera and Romano (1998), *A Geolinguística no Brasil: caminhos e perspectivas*, p. 155–164.

Razky 2010a = Abdelhak Razky, *Uma perspectiva geo-sociolinguística para a análise do status da variável /s/ em contexto pós-vocálico no nordeste do estado do Pará,* "Estudos Linguísticos e Literários" 41, p. 170–188.

Razky 2010b = Abdelhak Razky, *Pour une approche géo-sociolinguistique de la variation phonétique*. "Lenguaje" 38(2): 313–330. https://doi.org.10.25100/lenguaje.v38i2.4918

Razky and Guedes 2015 = Abdelhak Razky, Regis Guedes, *Le continuum des regroupements lexicaux dans l'atlas géosociolinguistique du Pará.* "Géoliguistique" 15, 149–162. https://doi.org/10.4000/geolinguistique.645Razky 2010b = Abdelhak Razky, *Pour une approche géo-sociolinguistique de la variation phonétique*. "Lenguaje" 38(2): 313–330. https://doi.org.10.25100/lenguaje.v38i2.4918

Razky *et al*. 2019 = Abdelhak Razky, Iara M. Teles, Diego Coimbra, *Agrupamentos fonéticos da vogal média anterior /e/ em posição pretônica no Atlas Linguístico de Rondônia (AliRo), "*Cadernos de Estudos Linguísticos" 61, p. 1–19. https://doi.org/10.20396/cel.v61i0.8654429

Razky and Coimbra 2020 = Abdelhak Razky, Diego Coimbra, *Estudo comparativo da variação do /S/ em posição de coda silábica nos atlas linguísticos estaduais da Região Norte,* "Moara" 55, p. 316–338. http://dx.doi.org/10.18542/moara.v0i55.9072Rennicke 2015 = Iiris Rennicke, *Variation and change in the rhotics of Brazilian Portuguese,* Doctoral dissertation, University of Helsinki and Universidade Federal de Minas Gerais. https://repositorio.ufmg.br/bitstream/1843/MGSS-A8ZH8U/1/rennickerhotics.pdf

Rennicke 2011 = Iiris Rennicke, *The retroflex r of Brazilian Portuguese: theories of origin and a case study of language attitudes in Minas Gerais,* "Linguística: Revista de Estudos Linguísticos da Universidade do Porto" 6, p. 149–170. https://ojs.letras.up.pt/index.php/EL/article/view/2741

Rodrigues 1974 = Ada N. Rodrigues, *O dialeto caipira na região de Piracicaba,* Sao Paulo, Ática.

Ruffino 1998 = Giovanni Ruffino, *Dialettologia, geolinguistica, sociolinguistica*, Berlin, Boston, Max Niemeyer Verlag. https://doi.org/10.1515/9783110934038

Sá and Martins 2016 = Lucrécio A. de Sá Jr., Marco A. Martins, *Rumos da linguística brasileira no século XXI*, Sao Paulo, Blucher, https://doi.org/10.5151/9788580391824-05

Sanches and Razky 2021 = Romario D. Sanches, Abdelhak Razky, *Análise geossociolinguística das designações para fanhoso nas capitais brasileiras, "*DELTA. Documentação de Estudos em Linguística Teórica e Aplicada" 37, p. 1–22. https://doi.org/10.1590/1678-460X202143836

Schmidt 1872 = Johannes Schmidt, *Die Verwantschaftsverhältnisse der indogermanischen Sprachen*. Weimar, Böhlau.

Silva Neto 2018 = Casimiro P. da Silva Neto, *Desbravadores do Brasil*. Brasilia, Sesc-DF.

Snoek 2014 = Conor Snoek, *Review of Gabmap: Doing Dialect Analysis on the Web,* "Language Documentation and Conservation" 8, p. 192–208.

Thun 2005 = Harald Thun, *A dialetologia pluridimensional no Rio da Prata,* in, Zilles (2005), Estudos de variação linguística no Brasil e no Cone Sul, p. 63–92.

Thun 2000a = Harald Thun, *Atlas Lingüístico Diatópico y Diastrático del Uruguay-Norte (Addu-Norte): Parte Cartográfica. 1,* Kiel, Westensee-Verl.

Thun 2000b = Harald Thun, *O português americano fora do Brasil, in,* Gärtner *et al*. (2000), Estudos de geolingüística do português americano, p. 185–227.

Thun 1998 = Harald Thun, *La geolingüística como lingüística variacional general (con ejemplos del Atlas lingüístico Diatópico y Diastrático del Uruguay, in, Ruffino (1998),* Dialettologia, geolinguistica, sociolinguistica, p. 701–730. https://doi.org/10.1515/9783110934038.701

Trudgill 1974 = Peter Trudgill, *Linguistic Change and Diffusion: Description and Explanation in Sociolinguistic Dialect Geography*, "Language in Society" 3(2), p. 215–246. http://www.jstor.org/stable/4166764

Velho 2017 = Priscila S. A. Velho, *A variação da vibrante em início de sílaba no português falado na antiga região de colonização italiana do Rio Grande do Sul: em busca de um padrão regional*, "Letronica" 10, p. 313–324. https://doi.org/10.15448/1984-4301.2017.1.25075

Zilles 2005 = Ana M. S. Zilles, *Estudos de variação linguística no Brasil e no Cone Sul*, Porto Alegre, UFRGS.

GROUPING TYPOLOGY AND DIFFUSION IN THE BRAZILIAN SOUTH – ESTABLISHING GUIDELINES FOR FUTURE RESEARCH

Abstract

This investigation examines the realizations of 'r' in the variety of Brazilian Portuguese (BP) spoken in the three southern states, giving special attention to Rio Grande do Sul (RS). The objective is to understand the diffusion of rhotic sounds in the region and mainly to establish guidelines for future research to examine the situation in a different way to the generalizing approaches usually applied to investigating the linguistic phenomena in those states. The investigation begins by considering the data presented in the Atlas Linguístico do Brasil (ALiB) and the Atlas Linguístico-etnográfico da Região Sul do Brasil (ALERS) in terms of the distribution of the variables. Next, by using Gabmap, distributional maps are generated and the clusters are validated considering multidimensional scaling (MDS). Those clusters are examined against the grouping typology in the attempt to explain the chronology of the diffusion. Finally, the article suggests courses for future investigation.

Universidade de Brasilia