ISSN 1983-9979
prolíngua

FROM FAVELA TO THE BAIXADA: THE VARIATION OF CODA (S) AND THE CONTINUUM OF URBAN NORM IN RIO DE JANEIRO

DA FAVELA À BAIXADA: A VARIAÇÃO DA CODA (S) E O CONTÍNUO DE NORMA URBANA NO RIO DE JANEIRO*

Marcelo Melo[1]
Gabriel Magalhães da Silva Gregório[2]
Kamila Leite Alvarez da Silva[3]
Karen Tavares dos Santos Lima[4]

ABSTRACT: This study investigates coda (s) variation in the Rio de Janeiro speech community, based on two new speech samples: FavRio, composed of residents from urban favelas, and Baixada-RJ, formed by speakers from the Baixada Fluminense region. The aim is to examine both linguistic and social constraints on the realization of alveopalatal, back fricative, alveolar, and absence variants, as well as to analyze how these communities are positioned within the urban norm continuum. Grounded in Variationist Sociolinguistics (WEINREICH; LABOV; HERZOG, 1968) and Usage-based Models (BYBEE, 2016; PIERREHUMBERT, 2003, 2016; CRISTÓFARO-SILVA; GOMES, 2017, 2020), the analysis shows that structural factors – such as following context, stress, and coda position – alongside social variables, such as sex and schooling, significantly influence the production of the back fricative variant. Findings indicate that, although from peripheral territories, FavRio and Baixada-RJ speakers do not align with the patterns of socially excluded adolescents (EJLA sample) but rather converge with middle-class speakers (Census 2000 sample) and socially integrated adolescents (Fiocruz sample). Moreover, the study reveals that the centrality of the alveopalatal or back fricative variant in the representation of specific lexical items varies according to speakers' degree of social insertion, reflecting different ways of organizing linguistic knowledge. These results support the hypothesis of an urban norm continuum in the Rio de Janeiro speech community, in which social insertion and access to prestigious institutions influence both production and lexical representation.

KEYWORDS: Coda (s). Rio de Janeiro speech community. Urban continuum. Usage-based Models. Variationist Sociolinguistics.

RESUMO: Este estudo investiga a variação da coda (s) na comunidade de fala do Rio de Janeiro, a partir de duas novas amostras de fala: FavRio, composta por moradores de favelas da capital, e Baixada-RJ, formada por falantes oriundos da Baixada Fluminense. O objetivo é compreender os condicionamentos linguísticos e sociais que atuam sobre a realização das variantes alveopalatal, posterior, alveolar e de ausência, bem como analisar o posicionamento dessas comunidades no contínuo de norma urbana. A análise, fundamentada na Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 1968) e nos Modelos Baseados no Uso (BYBEE, 2016; PIERREHUMBERT, 2003, 2016; CRISTÓFARO-SILVA; GOMES, 2017, 2020), revelou que fatores estruturais, como contexto seguinte, tonicidade e posição da coda, além de variáveis sociais, como sexo e escolaridade, condicionam significativamente a realização da variante posterior. Os resultados indicam que, embora oriundos de territórios periféricos, os falantes das amostras FavRio e Baixada-RJ não se aproximam do padrão dos adolescentes socialmente excluídos (Amostra EJLA), mas convergem para o comportamento observado entre

* Submmited: 06.09.2025 – Accepted: 13.01.2026 | DOI: 10.22478/ufpb.1983-9979.2025v20n2.76169
[1] Professor Adjunto do Departamento de Linguística e Filologia da Universidade Federal do Rio de Janeiro – Contato: malmelo.lopes@letras.ufrj.br | Orcid: https://orcid.org/0000-0002-8025-0530
[2] Mestrando do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro (PPGLIN/UFRJ) – Contato: gabrielgregorio@letras.ufrj.br| Orcid: https://orcid.org/0009-0008-3973-7946
[3] Mestranda em Linguística pela Universidade Federal do Rio de Janeiro – Contato: kamilalvarezs@letras.ufrj.br
[4] Licenciada em Português-Inglês pela Universidade Federal do Rio de Janeiro. Contato: karentavares@letras.ufrj.br

falantes da classe média (Amostra Censo 2000) e adolescentes inseridos socialmente (AmostraFiocruz). Além disso, observou-se que a centralidade da variante alveopalatal ou posterior na representação de itens lexicais mais frequentes varia conforme o grau de inserção social dos falantes, refletindo diferentes modos de organização do conhecimento linguístico. Esses achados sustentam a hipótese de um contínuo de normas urbanas na comunidade de fala do Rio de Janeiro, no qual a inserção social e o acesso a instituições de prestígio influenciam tanto a produção quanto a representação lexical.

PALAVRAS-CHAVE: Coda (s). Comunidade de fala do Rio de Janeiro. Contínuo urbano. Modelos Baseados no Uso. Sociolinguística Variacionista.

1 Introduction

Different samples of the Rio de Janeiro speech community have primarily mapped speakers belonging to different sectors of the middle class. Based on these samples, different variables have been analyzed, among which the variation of coda (s) has always generated particular interest, since the alveopalatal realization of this variable is recognized as a characteristic of the carioca variety.

Thus, the present study aims to analyze the variation of coda (s) in word-internal (medial) and final positions, such as in me[ʒ]mo ~ me[ɦ]mo (*mesmo*), nó[ʃ] ~ nó[h] (*nós*), in the Rio de Janeiro speech community, based on two new samples under construction, consisting of individuals from the peripheries of the state of Rio de Janeiro: the FavRio Sample (PEUL/UFRJ) and the Baixada-RJ Sample (PEUL/UFRJ). Through these two samples, the intention is to observe the constraints on the realization of the variable and its impacts on the organization of individuals' linguistic knowledge, as well as to delineate a continuum of urban norms based on the distance of the groups of individuals from the norm of the city of Rio de Janeiro. This distance will be mapped based on physical distance (territories far from the urban center, through the Baixada-RJ Sample) and social distance (favelas in the urban center, through the FavRio Sample). In other words, the reference to peripheries in this project is made in two dimensions: geographical (physical) and social.

Regarding the variation of coda (s) in the Rio de Janeiro speech community, previous studies have shown three important points: strong phonetic conditioning, the possibility of lexical conditioning, and variant evaluation. In relation to phonetic conditioning, it is possible to observe in previous studies on coda (s) in the carioca variety that the realization of the back fricative variant (velar/glottal) is strongly favored when the following context is a voiced consonant and disfavored before a voiceless consonant, vowel, and pause (SCHERRE; MACEDO, 2000; MELO, 2012, 2017). Furthermore, some studies indicate that there may be lexical conditioning for the realization of the back fricative variant (AULER, 1992; SCHERRE; MACEDO, 2000; MELO, 2012, 2017), given that some items tend to be realized with this variant more often than others. Concerning the evaluation of variants, it is observed that the back fricative variant is stigmatized among more educated speakers, while the alveopalatal variant is perceived only as a characteristic of the carioca variety (MELO, 2012, 2017).

Based on the data from the new samples and the results of previous studies, the theoretical support for the present analysis combines the theoretical assumptions of Variationist Sociolinguistics and Usage-Based Models. Regarding Variationist Sociolinguistics, it is understood that linguistic knowledge is endowed with structured heterogeneity (WEINREICH; LABOV; HERZOG, 1968), which is why the variability observed in linguistic use is conditioned by linguistic, social, and cognitive factors (LABOV, 1994, 2001, 2010). Concerning Usage-Based Models (henceforth UBM), it is understood that there is a relationship between abstracted knowledge and usage, granting representational status to the variability

observed in use (BYBEE, 2016; PIERREHUMBERT, 2003, 2016; CRISTÓFARO-SILVA; GOMES, 2017, 2020).

Thus, the present study aims to contribute to the debate on issues that are central to Variationist Sociolinguistics: (i) whether speakers from different social groups within the same speech community share the same structural patterns of a given linguistic variable and, if not, to what extent these patterns diverge; (ii) whether favoring and disfavoring conditioning factors for the realization of a given variant operate in the same way across all social groups within a speech community; and (iii) to what extent different social experiences affect the linguistic knowledge of speakers belonging to the same speech community. To this end, the study is organized as follows: the next section presents the theoretical assumptions that underpin the present research; the following section revisits previous studies on the linguistic variable under analysis; the third section outlines the methodology adopted for data collection; the fourth section presents the data analysis; and, finally, some concluding remarks are offered.

2 Variation and sound change

Weinreich, Labov, and Herzog (1968) (henceforth WLH) broke with the tradition of linguistic studies consolidated in the first part of the 19th century, according to which language constitutes a homogeneous and abstract system, unrelated to historical and social issues. By breaking with this tradition, WLH conceived the linguistic system as being endowed with orderly heterogeneity, which is why the variation observed in speakers' usage is related to the heterogeneity of the system and to – or conditioned by – structural (linguistic), social, and cognitive issues (LABOV, 1994, 2001, 2010). Also, according to the authors, only a model of knowledge organization that incorporates variation into the system would be capable of explaining language change, since "the key to a rational conception of language change – indeed, of language itself – is the possibility of describing orderly differentiation in a language serving a community"(WEINREICH; LABOV; HERZOG, 1968, p.101). Certainly, the assumptions put forward by WLH had a major impact on studies of linguistic variation and change, especially for those who focused on the relationship between language and society.

Despite introducing variation into the linguistic system through the postulation of the concept of orderly heterogeneity, WLH adopted the formal grammar model prevalent at the time, and based on it, postulated the concept of a variable rule to explain how the different uses of language were not only produced but also systematized. Thus, the grammar model traditionally assumed by variationist studies was a formal model, with variation explained through rules: a rule is applied to an underlying invariant form, and different forms are generated. In other words, despite conceiving that change is intimately linked to variation and that variation is inherent to the linguistic system, variationist studies have adopted, over the last few decades, a model of linguistic knowledge organization whose core is invariant and from which variable forms are produced. This conception of the linguistic system raises an important question for the treatment of variation in sociolinguistic studies: how would a grammar whose core is invariant be able to accommodate variation (CRISTÓFARO-SILVA; GOMES, 2004, p.32)?

Usage-Based Models are presented as a grammar model that considers the speaker's real linguistic experience and that is built – or emerges – from language use. Pierrehumbert (2003) argues that linguistic variability, as it constitutes a fundamental element of the speaker's linguistic knowledge, cannot be reduced to the result of a process. Thus, UBM postulate variation as an intrinsic component of the speaker's linguistic knowledge, granting representational status to linguistic variation and not treating it as the result of a process. In other words, unlike formal models, mental representations are multiple and include different ways in which lexical items are produced and perceived. In this way, UBM conceive that

abstract representations contain all phonetic possibilities for the realization of lexical items (CONNINE; RANBOM; PATTERSON, 2008).

In addition to postulating that representations are redundant and include phonetic detail, UBM make no distinction between lexicon and grammar. To this end, Bybee (2016) and Pierrehumbert (2003) argue that linguistic items are not stored randomly and in an unstructured list, but that the phonetic and semantic similarities observed between the items are part of the storage structure, thus building a network of connections from the items (CRISTÓFARO-SILVA; GOMES, 2004, p.37). These networks are reinforced according to the experiences lived by each speaker individually, which is why the mapping of these networks is continuous, as it is always updated according to individual experience (BYBEE, 2016). This lexical organization in networks, in turn, is also dynamic, since the same item participates in connections with other lexical items based on different shared characteristics.

UBM applied in Phonology, based on Exemplar Theory, allow for the representation of both the "abstract information that makes up the lexical item, and empirical, or phonetically detailed, information, containing gradualness, that is, the phonetic detail present in speech" (CRISTÓFARO-SILVA; GOMES, 2017, p.13, our translation), since this information is part of word representations. It is from these detailed representations that abstractions emerge. Therefore, all concrete variants of language use are perceived and stored in exemplar clouds that encompass all the phonetic possibilities of the lexical items. This, according to Bybee (2002), allows for the capture of ongoing sound change processes, since multiple representations can mirror different organizations of lexical items according to the age and other social categories of speakers in the same community. In other words, as representations mirror aspects of the production and perception of what is produced by members of a given community, "the speakers of this speech community are also exposed to the variability resulting from the distribution of the stage of change according to the age and socioeconomic class of the speakers of that community" (CRISTÓFARO-SILVA; GOMES, 2020, p. 21, our translation).

3 Studies on the variable

The previously cited studies on the variation of coda (s) in the Rio de Janeiro speech community point to the predominance of the alveopalatal variant, with other realizations – alveolar variant, 'aspirated' or back fricative variant (glottal or velar), or coda absence – occurring in smaller percentages. Among speakers with higher education in the city of Rio de Janeiro, Callou and Brandão (2009) argue that the palatalization of coda (s) is generalized and independent of education level, gender, and age group, with the realization of the back fricative occurring in only 1% of the data. Among speakers with Elementary and Middle School (Census 1980 Sample), Scherre and Macedo (2000) also observed that the most frequent variant in the carioca variety would be the alveopalatal fricative (61%), with the realization of the glottal fricative observed in only 7% of the data. Still in relation to speakers with Elementary and Middle School (a subgroup of speakers from the Census 2000 Sample), Melo (2012) observed a variant distribution very similar to Scherre and Macedo (2000): predominance of the alveopalatal variant (74%) and much lower percentages for the other variants (19% for the alveolar, 3% for coda absence, and 6% for the back fricative variant).

Melo (2017) also observed the behavior of speakers in the Rio de Janeiro speech community who belonged to a social group markedly distinct from those previously observed: adolescents living in favelas in the city of Rio de Janeiro with different degrees of social insertion. For the group of socially integrated adolescents (Fiocruz Sample), the distribution of variants mirrors that observed in previous studies: 77% for the alveopalatal variant, 15% for the alveolar variant, 2% for coda absence, and 6% for the back fricative variant. On the other hand,

for the group of socially excluded adolescents, a very large difference was observed in the distribution of variants, especially regarding the back fricative variant: 53% for the alveopalatal variant, 12% for the alveolar variant, 4% for coda absence, and 30% for the back fricative variant. Although similar behavior was expected for these two groups of adolescents (Fiocruz and EJLA), the observed differences were attributed to the degree of social insertion of the speakers in both groups, with the behavior of adolescents with some degree of social insertion (Fiocruz) being very similar to that of middle-class speakers (Census 2000) and very different from socially excluded adolescents (EJLA).

Most studies already conducted on coda (s) variation in Brazilian Portuguese (henceforth BP), including the carioca variety, revolve around three important – and recurrent– issues: strong phonetic conditioning for variant realization, the possibility of lexical conditioning, and social evaluation of the variants, especially regarding the back fricative. Thus, regarding strong phonetic conditioning for coda (s) realization, various studies on the variable mention the importance of the following context as a conditioning factor of the variants (GRYNER; MACEDO, 2000; CARVALHO, 2000; SCHERRE; MACEDO, 2000; BRESCANCINI, 2006; SANTOS, 2009; MELO, 2012, 2017). Regarding the Rio de Janeiro speech community, Scherre and Macedo (2000) maintain that the realization of coda (s) is directly related to the sonority scale, i.e., "the lower the degree of the following segment on this scale, the less the aspirated variant (0.37 before a voiceless stop); the higher the degree, the more the aspirated variant (0.94 and 0.87 before a voiced lateral and before voiced nasals)" (p.55, our translation). Melo (2012, 2017) argues that the following context consisting of a voiced consonant favors the realization of the glottal variant, while the following context consisting of a voiceless consonant, vowel, or pause disfavors the realization of the same variant.

Melo (2017) points out that although there are no systematic studies on lexical conditioning for the realization of coda (s) in BP, it is possible to find mention of some lexical effects in different works. Gryner and Macedo (2000) allude to "a factor of morphological nature" for the realization of coda (s) in the Cordeiro region (RJ), and Carvalho (2000) observed a relationship between certain word classes and the realization of some variables. Santos (2009) observes a higher occurrence of the back fricative variant in certain items, which led the author to argue in favor of a process of lexical diffusion, considering that, according to the author, a process of glottalization (realization of the back fricative variant) would have started in specific items and spread to other items in the language. Auler (1992), regarding the Rio de Janeiro speech community, also defended that items would be ahead of a probable process of lexical diffusion, pointing only to the item *mesmo* as being predominantly realized with the back fricative variant. Still concerning the Rio de Janeiro speech community, Scherre and Macedo (2000) observed that some items were realized more frequently with the aspirated variant (*nós*, *mais*, and *mesmo*), which led the authors to postulate "the need for future analyses that take into account, for example, the frequency and formality of the items" (p.62, our translation). Melo (2012, 2017), by utilizing the theoretical assumptions of UBM and a statistical model capable of accommodating fixed and random effects variables, claimed that, among adolescents in the Rio de Janeiro speech community who are socially excluded, high-frequency items with a phonetic context favorable to the realization of the back fricative variant were predominantly realized with this variant. According to the author, this would indicate that the lexical item is a constraint on the realization of the back fricative variant.

Finally, regarding the evaluation of variants, some studies observe a degree of stigma attributed to the back fricative variant in different varieties of BP. These studies are based on production data, and assumptions are made based on the characteristics of speakers who use the aspirated or back fricative variant the most, i.e., less educated speakers belonging to lower

socioeconomic classes. Thus, as the back fricative variant is produced more by speakers from lower socioeconomic classes (lower class), different production studies argue that the back fricative is the stigmatized variant of coda (s).

Unlike other studies, Melo (2017, 2022), through a sociolinguistic evaluation experiment using the matched guise technique, sought to access the social evaluation of participants from four social groups in the Rio de Janeiro speech community: UFRJ (middle-middle and lower-middle class speakers, with higher education), Fiocruz (lower-class speakers, with social insertion and regular schooling), EPSJV (lower-class speakers, with a certain degree of social insertion and delayed schooling), and EJLA (lower-class speakers, socially excluded, with irregular schooling). The results indicated that the back fricative variant was more associated with a less socially prestigious professional profile by participants from the UFRJ, Fiocruz, and EPSJV groups, i.e., speakers with some degree of social insertion and schooling. In contrast, the EJLA group (socially excluded adolescents) did not distinguish between the variants in their evaluation, showing that, for these speakers, there was no negative or positive evaluation of the back fricative variant. Regarding the alveopalatal variant, this variant was not consistently perceived as prestigious by any of the groups, suggesting that it is seen as the expected or characteristic realization of the Rio de Janeiro speech community, with no associated positive value or stigma. Although speakers from the Fiocruz, EPSJV, and EJLA groups have the same social origin (favela residents), speakers from the Fiocruz and EPSJV groups show evaluation patterns closer to the university group, considerably distancing themselves from the EJLA group. Thus, Melo (2017) concludes that the degree of social insertion of the subjects (influenced by schooling, family ties, and access to prestigious institutions) is crucial to explaining the different evaluations, which is why the author argued that speakers may be arranged on a continuum of social insertion, with different degrees of access to socially prestigious values, which impacts their linguistic knowledge and the evaluations they perform.

Given the results presented for coda (s) variation, especially concerning the Rio de Janeiro speech community, it is expected that the production of speakers from the FavRio and Baixada-RJ Samples will exhibit behavior that mirrors the variant distribution patterns closer to speakers from different sectors of the middle class, considering what was observed by Melo (2017) for speakers from the Fiocruz Sample (favela residents with a certain degree of social insertion). On the other hand, it is also expected that there may be productions closer to those of speakers from the EJLA Sample (socially excluded), especially concerning some lexical items. In other words, it is expected that speakers from the FavRio and Baixada-RJ Samples, in an urban norm continuum of the Rio de Janeiro speech community, will occupy an intermediate position, between middle-class speakers and speakers with a lower degree of social insertion. These different positions in the proposed continuum may, in turn, point to different centralities of the coda (s) variants for specific lexical items, depending on the speakers' degree of social insertion.

4 Methodology

In recent decades, speech samples from the Rio de Janeiro speech community have been composed predominantly of individuals from different sectors of the middle class who, consequently, have access to the institutions responsible for shaping linguistic values in a given society, such as the Census 1980 and Census 2000 Samples, from the *Programa de Estudos sobre Usos da Língua* (PEUL/UFRJ);the NURC Sample, from the *Projeto Norma Culta Urbana* (UFRJ); and *Discurso e Gramática* (UFRJ). The MOBRAL, APERJ, Fiocruz (PEUL/UFRJ), and EJLA (PEUL/UFRJ) samples are a few examples of speech samples with individuals of a social origin different from the other samples. Thus, the FavRio Sample and

the Baixada-RJ Sample emerged from the need to observe the dynamics of variation and language change in the Rio de Janeiro speech community more broadly, encompassing individuals with a social profile that is still little studied: speakers residing in the peripheries of the capital.

Thus, the data for this research were collected from two spontaneous speech samples, composed of residents from peripheral regions of the state of Rio de Janeiro: the FavRio sample, consisting of residents of carioca favelas, and the Baixada-RJ Sample, consisting of individuals born and residing in the Baixada Fluminense. The samples were constituted considering the methodology of Variationist Sociolinguistics (LABOV, 2008) for obtaining spontaneous speech and are part of the PEUL/UFRJ collection, representing lower-class social groups and residents of the peripheries of the state of Rio de Janeiro. It is important to note that the interviews were duly approved by the *Comitê de Ética em Pesquisa* (CEP) of UFRJ[5]. The interviews for both samples were conducted using a digital recorder, and data collection follows the standards established by Variationist Sociolinguistics, ensuring the anonymity of the interviewees.

Regarding the FavRio Sample, 44 interviews have been conducted to date, all lasting between approximately 30 and 60 minutes. The favelas of Rio de Janeiro and the number of interviewees – per favela – are distributed, to date, as follows: *Arará* (1 interview); *Carobinha* (1); *Manguinhos* (2); *Maré* (1); *Parada de Lucas* (3); *Santo André* (10); *Vila Aliança* (3); *Vila Cruzeiro* (2); *Vila Kennedy* (18); *Vila Vintém* (2); and *Pavãozinho* (1). Due to the peculiarity of the sample, the interviewees were people known by the interviewers or indicated by close contacts, which did not alter the random nature of the sample. Regarding the Baixada-RJ Sample, 17 interviews, lasting between 30 and 60 minutes, have been conducted, mapping the municipalities of *Duque de Caxias*, *Mesquita*, *Nilópolis*, *Nova Iguaçu*, and *Magé*.

For comparison purposes, the stratification of the FavRio and Baixada-RJ samples follows the stratification patterns of the Census 1980 and 2000 samples, which mapped speakers from the Rio de Janeiro speech community: (a) three age groups (18 to 29 years; 30 to 49 years; above 50 years); (b) three levels of schooling (Elementary/MiddleSchool, High School, and Higher Education); and (c) sex (male and female). In addition, the results will be compared to the results obtained for the EJLA and Fiocruz Samples, both consisting of adolescents living in Rio de Janeiro favelas, with different degrees of social insertion. Although the Census 1980 and 2000 samples do not include speakers with a higher education level, it is understood that, due to the increase in social profiles that entered university in recent decades, the inclusion of this schooling level in these two samples is pertinent for a comparative analysis.

To date, data have been collected from a total of eight speakers, with four speakers drawn from each sample: four speakers from the FavRio Sample, aged between 18 and 29, including 01 man and 01 woman with Elementary/MiddleSchool and 01 man and 01 woman with High School; four speakers from the Baixada-RJ Sample, including 01 man and 01 woman aged between 18 and 29 with High School, as well as 01 man and 01 woman aged between 30 and 49, with Higher Education. The following independent variables were tested for both samples: following context, word size, stress of the syllable where the coda (s) occurs, morphological status of the coda (s), position of the coda in the word, in addition to the speakers' sex, schooling, and age group. The obtained data were subjected to the logistic regression model on the RStudio platform and, subsequently, to a likelihood-ratio test, which compares the full model with reduced versions, each with one variable removed.

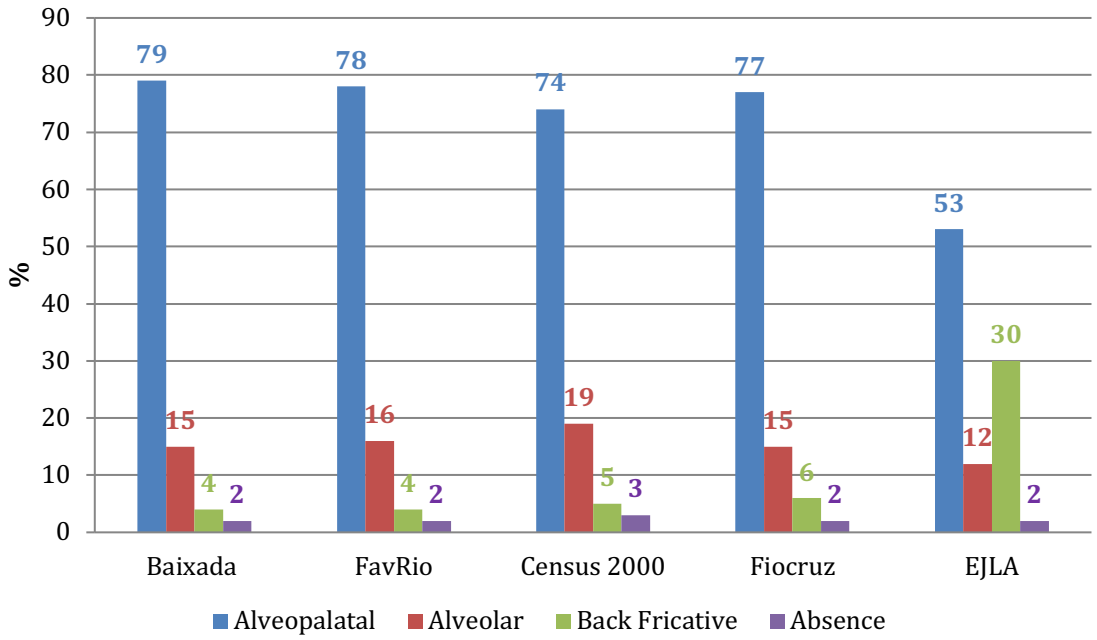
[5] (a) approval of the FavRio Sample: *Certificado de Apresentação de Apreciação Ética* (CAAE) number 19063219.0.0000.5286 (2020); (b) approval of the Baixada-RJ Sample: *Certificado de Apresentação de Apreciação Ética* (CAAE) number 78895424.4.0000.5582 (2024)

ISSN 1983-9979
prolíngua

5 Results

This section presents the preliminary results obtained from the FavRio Sample and the Baixada-RJ Sample for the analysis of coda (s) variants. The general distribution of variants for the two samples can be observed in Figure 01 below, along with the percentages for the other samples from the Rio de Janeiro speech Community:

Figure 01 — Distribution of coda (s) variants



Source: Authors' own elaboration

Comparing the data in Figure 01, it is o bserved that speakers from the FavRio and Baixada-RJ Samples exhibit behavior very similar to that of speakers from the Census 2000 Samples (middle-middle and lower-middle class speakers) and Fiocruz (socially integrated favela resident adolescents), especially regarding the percentages of the back fricative variant. Consequently, speakers from the FavRio and Baixada-RJ Samples distance themselves from speakers in the EJLA Sample (socially excluded adolescents). These results follow the same pattern observed by Melo (2017) for speakers in the Fiocruz Sample. As with Melo (2017), it was expected that speakers from the FavRio Sample, due to their social origin (favela residents), would exhibit behavior closer to that of speakers from the EJLA Sample (favela resident adolescents with low social insertion). However, this was not observed: like the speakers from the Fiocruz Sample, speakers from the FavRio Sample approach the behavior of middle-middle class speakers (Census 2000 Sample) and move away from speakers with a low degree of social insertion (EJLA Sample).

Regarding the variables tested, after submitting the data to logistic regression in R, the individual contribution of each explanatory variable to the model was evaluated using the likelihood-ratio test (Likelihood Ratio Test), through the drop1() function with test = "Chisq". The adjusted model had the dependent variable (DV) and included the variables following context to the coda (FOLLOWING), position of the coda in the word (POSITION), size of the word with the coda (SIZE), stress where the coda occurs (STRESS), morphological status of the coda (MORPH.STATUS), in addition to the speakers' sex (SEX) and schooling (SCHOOLING). A binary logistic regression was adopted, considering the realization of the back fricative variant as the event of interest.

ISSN 1983-9979
prolíngua

Table 01 shows the variables that showed a significant effect on the model in the Baixada-RJ Sample:

Table 01 — Variables with significant effect - coda (s): Baixada-RJ Sample

Variable	GL	Deviance	χ^2 (LRT)	p-value	Significance
FOLLOWING	3	325.19	126.03	< 2.2e-16	***
POSITION	1	199.40	0.24	0.6243	
SIZE	3	202.11	2.96	0.3985	
STRESS	2	208.23	9.08	0.0107	*
MORPH.STATUS	1	199.17	0.02	0.8978	
SEX	1	227.54	28.38	9.54e-08	***
SCHOOLING	1	214.46	15.30	1.54e-05	***

Source: Authors' own elaboration

As observed in Table 01, the results revealed that the variables FOLLOWING (χ^2 = 126.03; $p < 0.001$), STRESS (χ^2 = 9.08; $p = 0.0107$), SEX (χ^2 = 28.38; $p < 0.001$), and SCHOOLING (χ^2 = 15.30; $p < 0.001$) contributed significantly to the model. The variables POSITION, SIZE, and MORPH. STATUS did not show a statistically significant contribution ($p > 0.05$).

Table 02 shows the variables that showed a significant effect on the model in the FavRio Sample:

Table 02 — Variables with significant effect - coda (s): FavRio Sample

Variable	GL	Deviance	χ^2 (LRT)	p-value	Significance
FOLLOWING	3	405.80	84.34	< 2e-16	***
POSITION	1	327.49	6.03	0.01405	*
SIZE	3	325.54	4.08	0.25320	
STRESS	2	330.05	8.59	0.01365	*
MORPH.STATUS	1	321.49	0.03	0.86690	
SEX	1	413.10	91.64	< 2e-16	***
SCHOOLING	1	420.84	99.38	< 2e-16	***

Source: Authors' own elaboration

As observed in Table 02, the results revealed that the variables FOLLOWING (χ^2 = 84.34; $p < 0.001$), POSITION ($\chi^2 = 6.03$; $p = 0.014$), STRESS ($\chi^2 = 8.59$; $p = 0.0137$), SEX ($\chi^2 = 91.64$; $p < 0.001$), and SCHOOLING ($\chi^2 = 99.38$; $p < 0.001$) showed a significant effect. The variables SIZE and MORPH. STATUS did not show a statistically significant contribution ($p > 0.05$).

According to the results presented in Tables 01 and 02, it is possible to observe that following context, stress, speaker's sex, and schooling were relevant for the realization of the back fricative variant of coda (s) in both samples, in addition to the position of the coda in the word for the FavRio Sample. In general, these results replicate the findings of previous studies, especially regarding the following context (strong effect of phonetic conditioning).

Regarding the following context, the results follow the same conditioning factors as prior studies:

Table 03 — Coda (s) Results - Baixada-RJ Sample: FOLLOWING CONTEXT

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Voiced	219/276	79,35%	46/276	16,67%	1/276	0,36%	10/276	3,62%
Voiceless	543/552	98,36%	2/552	0,36%	3/552	0,72%	4/552	0,72%
Vowel	18/215	8,37%	0/215	0	187/215	86,98%	10/215	4,65%
Pause	202/207	97,58%	0/207	0	0/207	0	5/207	2,42%

Source: Authors' own elaboration

Table 04 — Coda (s) Results - FavRio Sample: FOLLOWING CONTEXT

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Voiced	328/406	80,79%	59/406	14,53%	02/406	0,49%	17/406	4,19%
Voiceless	810/845	95,86%	10/845	1,19%	04/845	0,47%	21/845	2,49%
Vowel	03/341	0,88%	4/341	1,17%	323/341	94,72%	11/341	3,22%
Pause	117/118	99,15%	0/118	0%	0/118	0%	11/118	0,85%

Source: Authors' own elaboration

The results for the FOLLOWING variable corroborate what was pointed out by Scherre and Macedo (2000) and Melo (2012, 2017), that is, the following context consisting of a voiced consonant favors the realization of the back fricative variant. As observed in Table 03, among the 48 items realized with the back fricative variant in the Baixada-RJ Sample, 46 were realized when the following context consisted of a voiced consonant. Something very similar happens in the FavRio Sample: of the 73 items realized with the back fricative variant, 59 had the following context to coda (s) consisting of a voiced consonant. Regarding this conditioning, Melo (2017) attributed the fact that the realization of the back fricative variant is favored when the coda is followed by a sonorant consonant or voiced obstruent is probably due to the same vocal cord approximation configuration during the production of the glottal fricative and sonorant and voiced consonants.

The results for the STRESS variable, detailed in Tables 05 and 06 below, follow the same tendency observed by Gryner and Macedo (2000) and Scherre and Macedo (2000):

Table 05 — Coda (s) Results – Baixada-RJ Sample: STRESS

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Post-tonic	385/498	77,31%	08/498	8%	78/498	15,67%	27/498	5,42%
Pre-tonic	234/244	95,90%	02/244	2%	07/244	2,87%	01/244	0,41%
Tonic	363/508	71,46%	38/508	38%	106/508	71,46%	01/508	0,20%

Source: Authors' own elaboration

Table 06 — Coda (s) Results – FavRio Sample: STRESS

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Post-tonic	372/529	70,32%	15/529	2,84%	107/529	20,23%	35/529	6,62%
Pre-tonic	317/320	99,06%	02/320	0,63%	01/320	0,31%	0/320	0
Tonic	569/861	66,09%	56/861	6,50%	221/861	25,67%	15/861	1,74%

Source: Authors' own elaboration

As can be inferred from the results for STRESS, the realization of the back fricative variant tends to occur in tonic (stressed) syllables: of the 48 occurrences of the back fricative variant in the Baixada-RJ Sample, 38 happened in a tonic syllable; of the 73 occurrences of the back fricative variant in the FavRio Sample, 56 happened in a tonic syllable. This effect for this variable run counter to weakening processes, which would primarily affect non-tonic syllables, since segments tend to be preserved in tonic syllables. Thus, these results raise a question for the hypothesis according to which the aspiration of coda (s) would be a stage of a coda weakening process toward the absence of the segment: if it is indeed a stage of a lenition process, how can we explain that the back fricative variant occurs predominantly in tonic syllables?

Regarding the variable position of the coda (s) in the word (POSITION), a variable relevant only for the FavRio Sample data, Table 07 presents the distribution of variants:

Table 07 — Coda (s) Results - FavRio Sample: POSITION

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Final	794/1232	64,45%	64/1232	5,19%	327/1232	26,54%	47/1232	3,82%
Internal	464/478	97,07%	09/478	1,88%	02/478	0,42%	03/478	0,63%

Source: Authors' own elaboration

Like the other structural variables analyzed, the results for the POSITION variable follow the results obtained by Melo (2012, 2017) for the EJLA and Fiocruz samples: there are more occurrences of the back fricative variant when the coda (s) is realized at the end of the word. Benayon (2010, p.88) verified that the most frequent context for the fricative in internal coda is that consisting of a voiceless fricative (*pasta*, *casca*, *aspa*), while the context consisting of a voiced fricative (*asma*, *asno*, *rasga*) is less frequent in BP. Thus, the fact that the internal context is more frequently constituted by a consonant that disfavors the realization of the back fricative would explain the lower realization of this variant word-internally.

Tables 08 and 09 present the results for the SEX variable in the analyzed samples:

Table 08 — Coda (s) Results - Baixada-RJ Sample: SEX

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Final	446/611	73,00%	25/611	7,04%	97/611	15,87%	25/611	4,09%
Internal	536/639	83,88%	05/639	0,78%	94/639	14,71%	04/639	0,63%

Source: Authors' own elaboration

Table 09 — Coda (s) Results - FavRio Sample: SEX

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Final	619/819	69,47%	66/819	7,40%	170/819	19,08%	36/819	4,04%
Internal	639/891	78,02%	07/891	0,85%	159/891	19,41%	14/891	1,71%

Source: Authors' own elaboration

Even in distinct varieties of BP, Gryner and Macedo (2000), Carvalho (2000), Santos (2009), and Autho (2012) observe that the back fricative variant of coda (s) is more frequently realized by men than by women. The results for the SEX variable indicate that the alveopalatal variant is the one that predominates for both sexes. However, it is possible to note that male individuals, in both samples, produce the back fricative variant more than women, which proved to be an important conditioning factor for the realization of this variant. If we understand the

variation of coda (s) in the Rio de Janeiro speech community as a process of stable variation, the literature that observes the effects of the sex variable in variationist studies points out that men tend to use more non-standard forms. In the case of the coda (s) variants, the back fricative is the stigmatized variant among more educated and/or socially inserted speakers, as indicated by evaluation experiments conducted by Melo (2017, 2022). Thus, the results for this variable follow the tendency observed in the literature.

Finally, the SCHOOLING variable proved to be relevant to the model. As observed in Tables 10 and 11, similarly to the results of Gryner and Macedo (2000), Carvalho (2000), and Melo (2012, 2017), the back fricative variant is more frequently realized by less educated speakers in the FavRio Sample. On the other hand, individuals with a higher degree of schooling (High School and Higher Education) realize the back fricative variant in lower percentages than less educated speakers:

Table 10 — Coda (s) Results - Baixada-RJ Sample: SCHOOLING

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
High	379/490	77,35%	26/490	5,31%	64/490	13,06%	21/490	4,28%
Higher	603/760	79,34%	22/760	1,05%	127/760	16,71%	08/760	2,89%

Source: Authors' own elaboration

Table 11 — Coda (s) Results - FavRio Sample: SCHOOLING

	Alveopalatal		**Back Fricative**		**Alveolar**		**Absence**	
	Apl/N	%	Apl/N	%	Apl/N	%	Apl/N	%
Elementary/ Middle	524/759	69,04%	61/759	8,04%	148/759	19,50%	26/759	3,43%
High	734/951	77,18%	12/951	1,26%	181/951	19,03%	24/951	2,52%

Source: Authors' own elaboration

Despite differences between High School speakers in the two samples, what is verified is an effect of schooling on the realization of the variable: the less educated speakers in the FavRio and Baixada-RJ Samples are the ones who realize the back fricative variant more. Thus, more educated speakers tend to realize the back fricative variant less, likely due to the stigma attributed to this variant (MELO, 2012, 2017).

Regarding lexical items, as only 121 occurrences of items with the back fricative variant were observed in the two samples (73 occurrences in FavRio and 48 occurrences in Baixada-RJ), it was not possible to analyze lexical effects. However, in a preliminary survey, it is possible to observe that few items were produced three or more times with the back fricative variant: *mas* (19/289), *mais* (14/168), *eles* (07/78), *mesmo* (06/54), *dois* (16/44), depois (03/42), nós (05/32), *àsvezes* (02/29), *desde* (07/20), *atrás* (05/10). Although future analyses are necessary to observe the effect of lexical conditioning, it is possible to observe that, unlike what happens with frequent items in the EJLA Sample, none of the items is realized predominantly with the back fricative variant. Thus, based on the assumptions of UBM, it is possible to argue, for now, that the items in the two samples have the alveopalatal variant in the central representations of the items containing coda (s).

The results presented in this section demonstrate that, despite the speakers in the two samples belonging to distinct peripheral territories in the state of Rio de Janeiro, their linguistic productions concerning coda (s) converge toward results very similar to the results for both the Fiocruz Sample, with socially inserted favela resident adolescents, and the Census 2000 Sample, with middle-middle and lower-middle class speakers. Consequently, the behavior of

the speakers in the FavRio and Baixada-RJ Samples distances itself from the behavior observed for socially excluded adolescents (EJLA Sample). This fact led to the following question: is it possible to argue that there is a continuum between the different norms observed in the Rio de Janeiro speech community, or is it possible to think of a much more severe rupture that places socially inserted speakers on one side and socially excluded speakers on the other? In Melo's terms (2017, p.133, our translation), it is important that "the different degrees of social insertion of the subjects are also considered so that it is possible to capture and explain the dynamism of the speech community, since social origin and hierarchy alone are not capable of translating such dynamism."

6 Discussion

The preliminary results obtained in this study consistently demonstrate that the variation of coda (s) in the FavRio and Baixada-RJ samples is strongly conditioned by linguistic and social factors. This is because, among the structural constraints, the strong influence of the phonetic context following the coda stands out, with the realization of the back fricative variant being favored when followed by voiced consonants, a conditioning factor already identified by Scherre and Macedo (2000) and Melo (2012, 2017). The stress of the syllable and the position of the coda in the word also proved relevant, reinforcing results from previous studies on the variable with respect to the realization of the back fricative variant.

Regarding social constraints, the effects of the sex and schooling variables of the speakers reinforce the association of the back fricative variant with groups with lower social insertion and lower schooling, which contributes to explaining the social stigma that weighs on the back fricative variant, as already pointed out in previous studies. However, the data show that, despite being speakers from peripheral regions – whether social peripheries (favelas) or geographical peripheries (Baixada Fluminense) – the speakers in the FavRio and Baixada-RJ Samples do not approach the pattern observed among socially excluded speakers, such as those in the EJLA Sample, but rather replicate the behaviors found in samples like Fiocruz and Census 2000, indicating a convergence with the prestigious urban norm with regard to the variable under analysis.

Thus, the results obtained in this work support the hypothesis of an urban norm continuum in the Rio de Janeiro speech community. This continuum is established not only in geographical but also in social terms, reflecting the different degrees of access to institutions that legitimize linguistic standards, such as schools, media, and access to the labor market. The insertion – even if partial – of these speakers into broader social networks and participation in spaces where prestigious norms circulate enables their linguistic behavior to reflect values and expectations close to those of the urban middle classes, at least in relation to some variables, as seems to be the case with coda (s). Therefore, instead of a rupture between "favela" and "asphalt" or between "periphery" and "center," the data suggest a gradation, in which different degrees of social insertion and mobility (social and spatial) influence the linguistic behavior of speakers and can impact the organization of the linguistic knowledge of these same speakers. A greater rupture, as is the case with socially excluded adolescents (EJLA Sample), leads not only to different variant distributions but also to different centralities in the representation of very frequent lexical items, which are migratorily produced with the back fricative variant. On the other hand, as observed so far for speakers in the FavRio and Baixada-RJ samples, greater insertion and mobility lead these speakers to share patterns very similar to those of speakers from different sectors of the middle class, both in relation to the distribution of variants and in relation to the centrality of variants in the representations of items with coda (s).

In these terms, in the future, the intention is to group the speakers in the FavRio and Baixada-RJ Samples based on criteria that reflect the level of social interaction of each speaker

with the place where they live, in order to measure the degree of integration into the territory, the identification of the interviewees with the favela, and the mobility of these speakers. What is observed in the peripheries is a very great heterogeneity, making it impossible to measure the degree of social insertion, as proposed by Melo (2017), without recognizing that this social insertion is intimately linked to access to spaces and institutions responsible for shaping socially prestigious values. In other words, it must be assumed that the behavior of individuals from lower classes is not uniform and that, in such an unequal society, the degree of mobility of individuals affects the linguistic behavior of speakers belonging to lower classes, impacting the sociolinguistic dynamics of the speech community.

References

AULER, M. A difusão lexical num fenômeno de aspiração em português. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 1, n. 1, p. 43-51, 1992.

BENAYON, A. **Aquisição das fricativas no Português Brasileiro:** Propriedades distribucionais e variação. 141 f. 2010. Tese (Doutorado em Linguística) — Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

BRESCANCINI, C. A fricativa em posição de coda no PB. In: RAMOS, J. (org.). **Estudos sociolinguísticos**: os quatro vértices do GT da ANPOLL. Belo Horizonte: Editora da FALE/UFMG, 2006. p. 06-20.

BYBEE, J. **Língua, Uso e Cognição**. Tradução: Maria Angélica Furtado da Cunha. São Paulo: Cortez, 2016.

BYBEE, J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. **Language Variation and Change**, *[s. l.]*, v. 14, n. 3, p. 261-290, 2002.

CALLOU, D.; BRANDÃO, S. Sobre o /S/ em coda silábica no Rio de Janeiro: falas culta e popular. In: SALGADO, A.; BARRETTO, M. (orgs.). **Sociolinguística no Brasil**: uma contribuição dos estudos sobre línguas em/de contato: homenagem ao Prof. Jürgen Heye. Rio de Janeiro: 7 Letras, p. 27-34, 2009.

CARVALHO, R. **Variação do /S/ pós-vocálico na fala de Belém**. 112 f. 2000. Dissertação (Mestrado em Linguística) — Universidade Federal do Pará, Belém, 2000.

CONNINE, C.; RANBOM, L.; PATTERSON, D. Processing variant forms in spoken word recognition: The role of variant frequency. **Perception & Psychophysics**, v.70, n. 3, p. 403–411, 2008.

CRISTÓFARO-SILVA T.; GOMES, C. Fonologia na perspectiva dos Modelos de Exemplares. In: GOMES, C. **Fonologia na perspectiva dos Modelos de Exemplares**: para além da dicotomia nature/nurture na ciência linguística. São Paulo: Contexto, 2020. p. 13-36.

CRISTÓFARO-SILVA T.; GOMES, C.Teoria de Exemplares. In: HORA, D. da; MATZENAUER, C. (orgs.). **Fonologia, fonologias**: uma introdução. São Paulo: Contexto, 2017, p. 157-168.

CRISTÓFARO-SILVA T.; GOMES, C.Variação linguística: antiga questão e novas perspectivas. **Linguagem**, v. 1, n. 2, p. 31-41, 2004.

GRYNER, H.; MACEDO, A. A pronúncia do –S pós-vocálico na região de Cordeiro – RJ. In: MOLLICA, M.; MARTELLOTA M. (orgs.). **Análises linguísticas**: a contribuição de Alzira Macedo. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, 2000. p. 26-51.

LABOV, W. **Principles of linguistic change**: Cognitive and Cultural Factors. Cambridge: Blackwell Publishers, 2010.

LABOV, W. **Padrões Sociolinguísticos**. Tradução: Marcos Bagno; Mª Marta Pereira Scherre; Caroline R. Cardoso. São Paulo: Parábola Editorial, 2008.

LABOV, W. **Principles of linguistic change**: Social Factors. Cambridge: Blackwell Publishers, 2001.

LABOV, W. **Principles of linguistic change**: Internal Factors. Cambridge: Blackwell Publishers, 1994.

MELO, M. **Desenvolvendo novos padrões na comunidade de fala**: um estudo sobre a fricativa em coda na comunidade de fala do Rio de Janeiro. 107 f. 2012. Dissertação (Mestradoe m Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

MELO, M. **Direcionalidade da mudança sonora**: o papel do item lexical e da avaliação social. 153 f. 2017. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.

MELO, M. Padrões de avaliação de duas variáveis sonoras na comunidade de fala do Rio de Janeiro: uniformidade ou diferentes tendências? **ORGANON**, v. 37, p. 102-124, 2022.

MELO, M. Fala, favela: a voz dos excluídos e a construção de identidades sociolinguísticas. In: Ana Pedro. (Org.). **Geografias (in)temporais**: identidade, memória e cultura. 1 ed. Aveiro: Humus, 2024. p. 183-203.

PIERREHUMBERT, J. Probabilistic Phonology: discrimination and robustness. In: BOD R.; HAY, J.; JANNEDY, S. (eds.).**Probabilistic Linguistics**. Cambridge: MIT Press, 2003. p. 177-228.

PIERREHUMBERT, J. Phonological representation: Beyond abstract versus episodic. **Annual Review of Linguistics**, v. 2, p. 33-52, 2016.

SANTOS, D. **A Variação do /S/ pós-vocálico da fala de Petrópolis, Itaperuna e Paraty**. 149 f. 2009. Dissertação (Mestrado em Linguística) — Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

SCHERRE, M.; MACEDO, A. Restrições fonético-fonológicas e lexicais: o -S pós-vocálico no Rio de Janeiro. In: MOLLICA, M.; MARTELLOTA M. (orgs.). **Análises linguísticas**: a contribuição de Alzira Macedo. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, 2000. p. 26-51.

WEINREICH, U.; LABOV, W.; HERZOG, M. **Fundamentos empíricos para uma teoria da mudança linguística**; Tradução: Marcos Bagno. São Paulo: Parábola Editorial, 2006 [1968].