The Retroflex [ɽ] in Data of the Monitored Style of the ALiB: An Analysis of Reading in Seven Cities in the State of São Paulo

Suely Claudia Lobato MACIEL*
Edina de Fátima de ALMEIDA**
Dircel Aparecida KAILER**

* Master(2018) in Languages Studies in the State University of Londrina. Professor in the State of the Pará. E-mail: suely.claudia.lobato@uel.br
** Master(2017) in Languages Studies in the State University of Londrina. Professor in the State of the Pará. E-mail: edifatro@gmail.com
*** PhD in Linguistic and Portuguese Language in the Federal University of the Rio de Janeiro. Professor at the State University of Londrina. E-mail: ueldircel@hotmail.com

Abstract:
The rhotics, due to their different realizations, have been the object of study in many sociolinguistic and dialectological studies (Aguilera (2008), Botassini (2009), Callou, Moraes e Leite (2013), Almeida e Kailer (2016), Maciel (2018), Maciel e Kailer (2019, among others). Nevertheless, there seems to be a lot to be investigated regarding the use of /R/ variants. For this reason, this article analyses, in the light of the theoretical assumptions of Variationist Sociolinguistics (Labov, 2008 [1972]), the realization of /R/ in internal syllabic coda in data from the Linguistic Atlas of Brazil (ALiB). These data were collected from readings carried out by speakers from seven cities in the state of São Paulo1 and analyzed with the objective of verifying whether the retroflex variant [ɽ] is productive in formal contexts of language use. After analysis, we found that: a) young people are using the retroflex in a very significant way, which indicates the vitality of the variant; b) in cities in the state inlands, the use of [ɽ] is more frequent than in the capital; and c) the low vowel [a] and the open middle [ɛ] are the most favorable to the implementation of the variant under study.

Keywords:
Sociolinguistics. Rhotics. Retroflex.

1 The seven locations analyzed in this article were: Campinas (point 173), Bragança Paulista (point 174), Guaratinguetá (point 176), Sorocaba (point 178), São Paulo (point 179), Capão Bonito (point 182) and Registro (point 186).
The Retroflex \[\text{"[\text{t}]\text{"} \text{]}\text{ in Data of the Monitored Style of the ALIB: An Analisys of Reading in Seven Cities in the State of São Paulo

Suely Claudia Lobato Maciel; Edina de Fátima de Almeida; Dirce Aparecida Kailer

INTRODUCTION

Labov (2008 [1972]) states that variation is essential to the nature of human language and that it is inseparable from the (external) social factors that act on it. The research he carried out in New York on the use of /R/ and on the island of Martha’s Vineyard on the centralization of diphthongs [ay] and [aw] are considered consolidators of the systematized study of linguistic heterogeneity.

From these works, the study of linguistic variation gained strength, and it has been investigated all over the world. Brazil followed this tradition, and authors such as Tarallo (1987), Callou (1987), Oliveira, Silva and Scherre (1996), Naro and Scherre (1999), Bisol (1999), among others, initially embraced Labovian sociolinguistic studies and brought reflections that have contributed, for example, to understanding the variations at different linguistic levels that occur in Brazilian Portuguese, such as the phonetic-phonological level, for example, which is the case of the multiple realizations for the phoneme /R/ in coda syllabic, object of analysis of this work.

Our general objective is to investigate, based on data collected by the Atlas Linguístico do Brasil (ALiB) project, the occurrences of variants of /R/, especially the retroflex \[\text{"[\text{t}]\text{"}\], in internal syllabic coda, in the reading of a text by the informants of seven locations in the state of São Paulo: Campinas, Bragança Paulista, Guaratinguetá, Sorocaba, São Paulo, Capão Bonito and Registro, respectively points: 173, 174, 176, 178, 179, 182 and 186 of the ALiB.

As specific objectives, we have:

a) analyzing the linguistic and extralinguistic contexts that influence the use of the retroflex variant;

b) verifying whether the retroflex is, in fact, gaining strength among younger Brazilian Portuguese speakers, as showed by Botassini’s (2009), Aguilera and Silva’s (2011), Oushiro and Mendes’s (2013) and Almeida’s (2017) studies.

Following Tarallo’ (1997) assumption that, in order to understand how language is actually used and what it is made of, it is necessary to make a correlation between linguistic and extralinguistic factors, in this work we investigate: (i) among the linguistic contexts, the preceding and following segments, as well as the morphological word class; and (ii) among the extralinguistic contexts, gender, age group and place of origin of the informants.

---

2 “The ALiB project has a national character, started in 1996 and is still under development, with the objective of producing a general atlas of Brazil regarding the Portuguese language. […] the ALiB team seeks to document different types of linguistic variation that make it possible to establish, in addition to dialect areas, the predominance of the use of a more colloquial or more formal variant in Brazilian speech” (AGUILERA; KAILER, 2015, p. 2).
These contexts were coded, and then the data were submitted to binomial analysis using the Goldvarb X Application\(^1\), which provided the results in percentage and relative weight for analysis, pointing out the most relevant contexts for the use of the retroflex variant [ɾ].

By analyzing the contexts that favor or not the occurrence of the retroflex variant, considered a linguistic mark of what Amaral (1920) called caipira speech (country accent), in more monitored speeches of language use (reading), when the informant is expected to make use of the variant which they consider most prestigious, we believe that this work will bring significant contributions to Brazilian phonetic-phonological and sociolinguistic studies.

**Sociolinguistics**

Sociolinguistics investigates language in its concrete use, aiming to understand its varieties and value all the heterogeneity present in its system (COELHO, 2010). It arises, as already mentioned in the introduction to this article, in the 1960s, from studies carried out by William Labov, in New York. It conceives, therefore, a heterogeneously structured language, with categorical rules (which do not vary within the speech community) occurring in parallel with variable rules (which may change from speaker to speaker). The different realizations of /R/ in syllabic coda are part of the variable rules and may even vary in the speech of the same person, as can be seen in data from some informants of this research.

Within the scope of variationist studies, different perspectives of analysis are presented (ALKMIM, 2001), such as the following variations: a) Diatopic: related to geographic space; b) Diastratic: related to social factors; and c) Diaphasic or Stylistic: related to the linguistic variants used by the speaker in different contexts and with different objectives. This last type of variation will be the basis for the work presented here and, for that reason, we will talk more about it later.

**Stylistic Variation**

Hora (2011, p. 148) states that the word *Style* can assume different concepts in the Portuguese language. However, when relating style to language use, he conceptualizes it as “a restriction that may or may not favor the choice of one among the different variants that constitute a variable”, corroborating the Botassini’s assumption (2009, p. 101) that “Depending on the formality of the interview (more informal and less structured, more formal and more controlled), the informants change the variant used”.

For Labov (2008 [1972]), this choice is related to the attention the speaker pays to speech; this conception of Style that we adopt in this article. From this perception, the author presents a monitoring continuum that goes from the least monitored to the most monitored speech. In a sociolinguistic interview, this continuum could be represented as follows:

---

\(^1\) Goldvarb X is a quantitative analysis program for computers, which provides results in percentage and relative weight, selecting the most favorable contexts for the use of a given variable rule.
The analysis carried out for this article is based on data obtained in situations of more monitored use, since it involves reading a text, a constitutive part of the data collection carried out by the ALiB project team. In this reading, we observe the realization of rhotics in internal coda, seeking to understand the linguistic and extralinguistic contexts that most favor the use of the retroflex variant.

Before dealing specifically with rhotics, it should be said that the Linguistic Atlas of Brazil Project, in a very general way, aims to provide empirically collected data so that we know Brazilian Portuguese (BP). In this sense, it maps the linguistic reality of BP at its different levels; establishes isoglosses to demonstrate the dialectal differences of each region; offers these data to scholars so that, within their theoretical-methodological perspectives, they can update, improve, expand and deepen their knowledge regarding the linguistic variety of the Portuguese language spoken in Brazil (COMITÊ NACIONAL, 2001, p. 16). Let us, then, talk a little more about rhotics, especially the retroflex [t].

Rhotics

The great diversity of realizations that rhotics can assume in speech, especially in syllabic coda, is, without a doubt, one of the main factors in the investigation of this archiphoneme.

Callou and Leite (1996, p. 465) state that the /R/ presents “in the syllabic coda position, a high degree of polymorphism, lending itself exemplarily to the characterization of variation in Brazilian Portuguese”. Bisol (1999) declares that there are multiple variants for the postvocalic /R/ in Brazilian Portuguese. According to Aguilera (2008, p. 1), “the /r/ in syllabic coda is the phoneme with the greatest possible number of variants in Brazilian Portuguese, especially when diatopic-regional dimensions are considered”. For Callou, Moraes and Leite (2013), the /R/ can be performed orally, in syllabic coda, as alveolar vibrating [ɾ], tap [ɾ], uvular fricative [ʁ], retroflex [t], velar fricative [x], glottal [h] or zero Ø.

Almeida and Kailer (2016), in turn, declare, based on Lindau (1985), that these realizations are socially evaluated in different ways, some being stigmatized ([ɾ], [ɻ]) and others prestigious ([ɦ], [h], [ɤ], [x], [ɾ]). The authors also present a distribution of the realizations of the rhotic in Brazil, as shown in Figure 01 below.

Figure 01 – Distribution of the achievements of the rhotic – Predominance

As can be seen in the map in Figure 1, the realization of the rhotic in the national territory shows an area further north where the glottal variant predominates, much of the Brazilian coast where there is a
record of the velar/uvular – which is recorded from Sergipe to Rio de Janeiro, passing through Goiânia, in the country inlands – and the predominance of the vibrant alveolar, retroflex and approximating variants further south of the country. From these different rhotic variants, we focus, in this article, on the studies on the retroflex [ɾ] that we are going to describe.

**Retroflex [ɾ]**

In this section we will cover the retroflex variant. It is carried out, according to Silva (2003, p. 34), by “lifting and curving the tip of the tongue towards the hard palate. They occur in the “caipira” dialect and in the accent of North Americans speaking Portuguese, as in the words mar, carta.”

For Silva (2016), the retroflex originated from the contact between Tupi and the Portuguese language where, today, is the state of São Paulo, and spread across the South, Southeast and Midwest regions from the incursion of the Bandeiras through Brazil. In 1920, Amadeu Amaral described the use of the retroflex, presenting it in a stigmatized way, as a mark of caipira speech. This is a brand that the author associated with the paulistas in general “The caipirismo was not only part of the language, but also of the way of life of the paulistas” (AMARAL, 1982 [1920], p. 41).

The stigma attributed by the author to the retroflex seems to be present even today in the evaluation that Brazilian Portuguese speakers make about the use of this variant. Almeida and Kailer (2017, p. 356), when analyzing the use of /R/ in syllabic coda in data from Goiás inlands, concluded that “the only style that disfavors the retroflex variant is reading (0.395)”. Also according to the authors, this occurs because, in reading, due to the degree of monitoring, the speaker tends to use the variant that he considers most prestigious, which may demonstrate a possible stigma suffered by the retroflex in this region. This result echoes the postulate by Labov (2008 [1972]) about the different degrees of stylistic monitoring.

In the same direction, regarding linguistic prejudice linked to retroflex realization, Botassini (2009), when researching linguistic beliefs and attitudes of speakers from the city of Maringá-PR, found the occurrence of linguistic disloyalty of most informants, since they stigmatize their own variant. However, when analyzing the speech of these informants in less and more monitored situations, the author observed that the younger ones prefer the use of the retroflex, presenting a more characteristic speech of Maringá’s dialect, although they affirm that “the Maringaenses do not speak very well, that they pull the /r/ a lot, which they don’t like the way Maringaenses speak” (BOTASSINI, 2009, p. 93).

Although in a preliminary way, in studies such as that of Botassini (2009), the retroflex variant in syllabic coda is increasingly present in the speech of young people. Aguilera and Silva (2011) arrived at similar results when investigating the community of Lavras-MG, in which the informants in the age group 01 (18 to 30 years) present a percentage of 55% of retroflex to 45% of glottal, while those of the second age group (50 to 65 years), conversely, present 62% glottal and 38% retroflex. Oushiro and Mendes (2013) also found a higher incidence of retroflex pronunciation among young people from São Paulo (0.550). Although the relative weight of the variable was close to neutrality, it is considerably higher than that related to younger people (0.440).

Let us now proceed to the analysis of the data used to carry out this work.

---

1 People who live in São Paulo state.
2 The way caipira of speaking
3 People who live in Maringá.
The CORPUS

Next, we present the constitution and analysis of the corpus. First, we specify the linguistic and extralinguistic variables analyzed and those most relevant to the application of the variable rule of use of the retroflex variant in a more monitored style of speech in seven locations in São Paulo.

Constitution

For the elaboration of this article, we analyzed ALiB data collected in seven cities in São Paulo regarding the use of rhotics in internal syllabic coda. The state of São Paulo was chosen for this research because it is the birthplace of the retroflex variant (AMARAL, 1982[1920]). The target locations were selected from the Map presented by Silva (2016, p. 118) and reproduced below as Figure 02.

In the lower part of the map, there is a demarcated area in which there is variation in the use of the rhotic in internal coda with the retroflex, tap and glottal variants. Six cities of the seven chosen for this work are within this area: Campinas (point 173), Bragança Paulista (point 174), Sorocaba (point 178), São Paulo (point 179), Capão Bonito (point 182) and Registro (point 186). Only Guaratinguetá (point 176), outside it, was included in the analysis for presenting a behavior similar to the other six.

It is worth mentioning that Silva’s (2016) thesis uses data from the ALiB Phonetic-Phonological Questionnaire (QFF), that is, it considers a less monitored style than that of reading within the Labovian continuum of stylistic variation (LABOV, 2008 [1972]). In this article, we work with the extreme of the
continuum, analyzing the same phenomenon in data collected from the reading carried out by the same informants investigated by her.

We also point out that, for each location, the ALiB has four informants: two men and two women in age group 01 (18 to 30 years old) and two men and two women in age group 02 (50 to 65 years old), with, preferably, incomplete basic education.

In the text that ALIB selected for reading (Comitê Nacional, 2001), the Parable of the seven wickers, “a text already used by Lacerda and Hammarström (1953) to analyze phonetic facts in Portuguese from Portugal” (CARDOSO et al., 2014, p. 90), there are 14 words with /R/ in internal coda (perguntou, parti, parte, tornou, partir, parte-o, ordenou, esforçar, partir, força, estiverdes, paterno, fortes and irmãos). However, it is necessary to explain that not all interviewees produced the 14 possibilities of the rhotic realization. Some informants read only a part of the text, others enunciated words different from those that were written and that did not present the rhotic in syllabic coda, and there were two cases of deletion, in forte ([ˈfɔØtʃɪ]), pronounced by a man in age group 02 from São Paulo City; and, in perguntou ([peØɡũˈto]), produced by a man also in age group 02 from Capão Bonito, both occurrences were excluded because deletion is not the object of analysis in this study. For these reasons, our corpus consisted of 317 words.

ANALYSIS

After identifying all the possibilities collected and excluding the two cases of deletion mentioned, we detected the occurrence of five variants of rhotic in internal syllabic coda in the researched locations: retroflex [ɾ], tap [ɾ], multiple vibrator [r], glottal [h] and velar [x], as shown in graph 1 below.

Graph 01 – Realization of /R/ in Internal Coda

It is possible to verify, through the reading of Graph 1, that the retroflex variant is quite productive in the investigated localities, having been categorical in Capão Bonito, which, because it does not present variation in the use of rhotics in syllabic coda in the clipping under study, was excluded from the binomial analysis. As our main objective is to observe the behavior of the retroflex in the analyzed data, we made a new coding merging the other variables (glottal, velar, tap and vibrating) into a single code to contrast with the retroflex. We then submitted the data to the Goldvarb X Application, which provided the results in percentage and relative weight for analysis, selecting the most favorable contexts for the use of [ɾ].
The first factor selected by the application was the age group of the informants, as shown in Table 01.

**Table 01 – Performance of the Age Range Variable for the Implementation of [ɾ] in Internal Coda**

<table>
<thead>
<tr>
<th>AGE GROUP</th>
<th>APPLICATION/TOTAL</th>
<th>%</th>
<th>RELATIVE WEIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 (18-30 years)</td>
<td>154/170</td>
<td>90.6</td>
<td>0.731</td>
</tr>
<tr>
<td>02 (50-65 years)</td>
<td>99/147</td>
<td>67.3</td>
<td>0.240</td>
</tr>
</tbody>
</table>

Input: 0.876  Significance: 0.000

Source: Prepared by the authors based on ALiB data.

From the reading of Table 01, it can be said that people in age group 01 – young people – (0.731) are using the retroflex variant more significantly than people in age group 02 (0.240). This conclusion is in line with what has been presented by recent research, such as those by Botassini (2009), Aguilera and Silva (2011) and Oushiro and Mendes (2013), who claim the maintenance and expansion of the retroflex. This scenario makes us think that, in fact, Amaral’s proposition (1982 [1920], p. 41), that the so-called caipira dialect “[...] is condemned to disappear in a more or less short period of time”, was mistaken.

Our hypothesis for this behavior among young people is that of hidden prestige, whose notion, according to Roncarati (2008, p. 52), “is associated with social identity, linguistic pride, belonging to a given social class or speech community”, that is, the retroflex is used by young people as a regional identity mark. However, further studies are still needed to prove this hypothesis.

The second factor selected by the application as being more significant for the implementation of the retroflex variant was the place of origin of the informants, as shown in table 02.

**Table 02 – Performance of the Locality Variable for the Implementation of [ɾ] in Internal Coda**

<table>
<thead>
<tr>
<th>LOCATION</th>
<th>APPLICATION/TOTAL</th>
<th>%</th>
<th>RELATIVE WEIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guaratinguetá</td>
<td>37/38</td>
<td>97.4</td>
<td>0.926</td>
</tr>
<tr>
<td>Registro</td>
<td>46/52</td>
<td>88.5</td>
<td>0.720</td>
</tr>
<tr>
<td>Campinas</td>
<td>42/54</td>
<td>77.8</td>
<td>0.501</td>
</tr>
<tr>
<td>Bragança Paulista</td>
<td>28/38</td>
<td>73.7</td>
<td>0.332</td>
</tr>
<tr>
<td>Sorocaba</td>
<td>18/30</td>
<td>60</td>
<td>0.229</td>
</tr>
<tr>
<td>São Paulo</td>
<td>28/51</td>
<td>54.9</td>
<td>0.165</td>
</tr>
</tbody>
</table>

Input: 0.876  Significance: 0.000

Source: Prepared by the authors based on ALiB data.

The data presented in Table 02 show that Guaratinguetá (0.926) and Registro (0.720) are very favorable locations to the application of the variable rule of using the retroflex variant in internal coda, while São Paulo (0.165), Sorocaba (0.229) and Bragança Paulista (0.332) are the ones that least favor the application of this rule, while Campinas is neutral regarding the use of [ɾ] (0.501).

It is important to note that the six locations served as a route for the pioneers who entered Brazil in search of gold and that were originally inhabited by native speakers of the Tupi language, which corroborates the thinking of Silva (2016) about the origin of the retroflex to be linked to the contact between Tupi and the Portuguese language.

Another observation related to the data in Table 02 is the fact that the capital of São Paulo has the lowest relative weight for the application of the retroflex variant, which suggests that the stigma as a mark of rural speech still walks along with the retroflex variant. This hypothesis is reinforced by the fact that the two locations that most favored the use of the variant are the most geographically distant from the capital of São Paulo, respectively 176 and 185.9 km.
The third and last factor listed by the application as being relevant to the implementation of the retroflex in internal syllabic coda was the preceding context, as shown in Table 3.

**Table 03 – Performance of the Precedent Context Variable for the Implementation of $\mathbf{\check{r}}$ in Internal Coda**

<table>
<thead>
<tr>
<th>PHONEME</th>
<th>APPLICATION/TOTAL</th>
<th>%</th>
<th>RELATIVE WEIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[a]</td>
<td>103/119</td>
<td>86.6</td>
<td>0.638</td>
</tr>
<tr>
<td>[ɛ]</td>
<td>38/44</td>
<td>86.4</td>
<td>0.619</td>
</tr>
<tr>
<td>[o]</td>
<td>57/67</td>
<td>85.1</td>
<td>0.570</td>
</tr>
<tr>
<td>[ɔ]</td>
<td>32/42</td>
<td>76.2</td>
<td>0.425</td>
</tr>
<tr>
<td>[i]</td>
<td>12/20</td>
<td>60</td>
<td>0.152</td>
</tr>
<tr>
<td>[e]</td>
<td>11/25</td>
<td>44</td>
<td>0.081</td>
</tr>
</tbody>
</table>

Input: 0.876  
Significance: 0.000

**Source:** Prepared by the authors based on ALiB data.

From the reading of the data presented in table 03, we can observe that the low vowel [a] (0.638) and the low mid front vowel [ɛ] (0.619) are the most favorable to the implementation of the retroflex; while the high mid back vowel [o] (0.570) and the low mid back vowel [ɔ] (0.425) are close to neutrality; and the high front vowel [i] (0.152) and the high mid front vowel [e] (0.081) disfavor the use of that variant. These data partially corroborate what was observed by Oushiro and Mendes (2013) in São Paulo city. The authors concluded that lower vowels are more favorable to retroflex (0.620), probably because this variant is also performed less loudly. It seems that a study with a greater number of words is necessary to see if there is, in fact, a relationship between the vowels preceding the rhotic and its realization as a retroflex.

One of the linguistic factors that we analyzed, which was not selected by the program, but which we consider important, refers to the grammatical class, which we divided into verbs and nouns, and which had a very close percentage of realization, as shown in graph 02.

**Graph 02 – Realization of $\mathbf{\check{r}}$ in Verbs and Nouns**

![Graph](image)

Source: Prepared by the authors based on ALiB data.

In relation to the data presented in Graph 02, we verified that the percentage of retroflexion in verbs and nouns is practically the same, which was probably decisive for their non-selection by the application. This result is very similar to that found by Oushiro and Mendes (2013) in São Paulo, where
nouns (0.510) and verbs (0.550) reached practically the same relative weight regarding the implementation of [t] in internal syllabic coda. This result gives evidence that the variable grammatical class has no influence on the use of one or another rhotic variant in internal syllabic coda.

**Final Considerations**

This article, whose objective was verifying whether the retroflex variant occurs in reading, a context of greater formality, in which, according to Labov (2002), one believes that the informant uses the variant they consider most prestigious, brings results that prove the vitality of the aforementioned variant, because, even being the context in which the informant pays more attention to their own speech, the retroflex prevailed, mainly, in the state of São Paulo inlands.

It also proves, in the seven locations under study, that the informants in age group 01 (18 – 30 years old) use the variant more significantly than those in age group 02 (50 – 65 years old), which leads us to believe that the retroflex [t], contrary to what had been stated by Amaral (1982 [1920]), does not tend to disappear quickly. It is getting stronger among informants in the state inlands and can be considered a regional identity mark. This result is echoed in Almeida’s (2017) studies on the variant retroflex in the speech of the Center-West region of Brazil.

Regarding the realization of the retroflex in the geographic space of São Paulo, we found that the capital is the location with the lowest percentage of use of this variant, which confirms studies already mentioned and that confirm the stigmatization of [t] as a mark of the “caipira” speech and, therefore, the little use among São Paulo speakers.

**References**


