https://doi.org/10.35520/diadorim.2018.v20n0a23279

ALiB: A DATABASE OF SPOKEN LANGUAGE FOR MAPPING LINGUISTIC VARIATION IN BRAZILIAN PORTUGUESE[1]

ALiB:UMA AMOSTRA DE LÍNGUA ORAL PARA MAPEAR A VARIAÇÃO LINGUÍSTICA NO PORTUGUÊS BRASILEIRO

Josane Moreira de Oliveira[2]
Jacyra Andrade Mota[3]
Suzana Alice Marcelino Cardoso[4†]

Resumo

O Atlas Linguístico do Brasil (ALiB) objetiva descrever e analisar o português brasileiro, documentando aspectos de diferentes níveis linguísticos. Usando uma metodologia multidimensional, o ALiB investiga a realidade linguística de 250 localidades, distribuídas por 26 Estados brasileiros. Aí estão incluídas todas as capitais – com exceção de Palmas (TO) e Brasília (DF), pois são cidades de fundação recente que ainda não têm falantes da segunda faixa etária contemplada no Projeto com pais nascidos na mesma cidade. Além da variável diatópica, o ALiB considera outras variáveis sociais, tais como sexo, idade e nível de escolaridade. O *corpus* foi constituído a partir de entrevistas com 1.100 informantes, estratificados por sexo, faixa etária e nível de escolaridade, totalizando, aproximadamente, 3.300 horas de gravação. Neste artigo, apresentamos nosso modelo de coleta de dados e análise, esta exemplificada pela variação fo-

1 An earlier version of this article was presented at the I International Symposium on Variation in Portuguese, University of Minho, Braga, April 28-30th, 2014.

2 Universidade Estadual de Feira de Santana. E-mail:josanemoreira@hotmail.com.
3 Universidade Federal da Bahia/CNPq 1.E-mail: jacymota@gmail.com.

4 †Participou deste texto, mas nos deixou em 2 de maio deste ano. Foi colega, amiga e parceira da homenageada na Dialetologia Brasileira. Temos a certeza de que concordaria, muito feliz, com a homenagem à Dinah.

Recebido em: 24/09/2018
Aceito em: 08/11/2018

nética entre [ti] / [di] (variantes estigmatizadas) ~ [tʃi] / [dʒi] (variantes *standard*) (*tia*, *parte*; *dia*, *desde*) no português brasileiro. O estudo considerou variáveis linguísticas e sociais que podem condicionar a realização dento-alveolar ou palatalizada de /ti/, /di/. Os dados foram analisados de acordo com o quadro teórico-metodológico da Sociolinguística Quantitativa, usando o pacote estatístico GoldVarb X. Nossos resultados mostram que a realização dento-alveolar ou palatalizada desses fonemas é condicionada geograficamente, além de aí atuarem também outras variáveis.

Palavras-chave: Atlas Linguístico do Brasil; Geolinguística Pluridimensional; Sociolinguística Quantitativa.

Abstract

The Linguistic Atlas of Brazil (ALiB) aims to describe and analyze Brazilian Portuguese, documenting aspects of its different linguistic levels. Using a multidimensional methodology, the ALiB investigates the linguistic reality of 250 locations, distributed across twenty-six Brazilian states. These include all the capital cities – with the exception of Palmas (capital of Tocantins State) and Brasília (capital of the Federal District) as they are new cities with no speakers of the second age group considered in this study, whose parents have been born in the same city. As well as the diatopic variable, the ALiB considers other social variables such as gender, age and level of education. The *corpus* was compiled by recording 1.100 respondents, totalling approximately 3.300 hours of recordings. In this article, we present our model of data collection and analysis, which is exemplified by means of the phonetic variation between a dento-palatal /t/ as [ti] / [di] (stigmatized variants) and a palatal fricative [tʃi] / [dʒi] (standard variants) (*tia*'aunt', *parte* 'part'; *dia* 'day', *desde* 'since') in Brazilian Portuguese. The data were analyzed according to the theoretical and methodological framework of Quantitative Sociolinguistics, using the GoldVarb X statistical package. Our results show that the dento-alveolar or palatal realizations of these phonemes are geographically conditioned, in conjunction with other social and structural variables.

Keywords: Linguistic Atlas of Brazil; Pluridimensional Geolinguistics; Quantitative Sociolinguistics.

Introduction

In the early nineteenth century, dialectology became established as a branch of linguistic studies, adopting a unidimensional methodology that gave precedende to diatopic variation and considered rural areas as a privileged *locus* of research, as exemplified in the *Atlas Linguistique de la France* (GILLIÉRON and EDMONT, 1902-1910). This methodological approach spread throughout Europe and even to Americas.

Although the importance of social variables such as sex, age and education was noted from the very beginning – Abbé Rousselot, in the late nineteenth century, expressed such concerns (cf. POP, 1950, p.43) –, the mapped results did not record this variation. It was the atlas by Hans Kurath *Linguistic Atlas of New England*, published between 1939-1943, that first incorporated the presence of variables other than diatopic variables in language maps, bringing

multidimensionality into play.

This methodological trend gradually took shape, especially since the 1960s, when theVariationist Sociolinguistics movement, led by William Labov(cf. WEINREICH; LABOV; HERZOG, 1968; LABOV, 1966, 1972, 1994, 2001, 2010) emphasized the relationship between language and social factors. This development led to the implementation of new guidelines in the methodology of dialect study, especially regarding the priority initially given to the geographic dimension in the analysis of speech. Henceforth, geolinguistic studies began not only to adopt some of the methodological assumptions of Sociolinguistics (social variables) but also felt the need to include different levels of dialogue, seeking to document register variation, as illustrated by the *Atlas Lingüístico Diatópico y Diastrático del Uruguay* (THUN; ELIZAINCÍN, 2000) and the Linguistic Atlas of Brazil (ALiB), in development.

Background

The idea of the Linguistic Atlas of Brazil (ALiB) was kindled in 1996, during the *Seminário Caminhos e Perspectivas para a Geolinguística no Brasil* (Seminar on Directions and Prospects for Geolinguistics in Brazil). At this seminar, three reasons for the development of the area were tabled: a) the absence of a general characterization of Brazilian Portuguese based on fieldwork data; b) the absence of data that would allow sketching of the dialectal divisions within the country; c) the need for knowledge of the country's linguistic multidimensionality in order to clarify and demarcate geolinguistic areas and thus better address the reality of each area and the teaching of the Portuguese language (CARDOSO, 1996).

In this way, the ALiB, a national project of institutional character, was born, with the following linguistic, political and social goals: a) to describe the reality of the Portuguese language in Brazil; b) to establish isoglosses in order to trace the dialectal divisions of Brazil; c) to examine the data collected interfacing with other branches of knowledge (history, sociology, anthropology etc.) in order to establish theoretical perspectives on the nature of the deployment and development of Brazilian Portuguese; d) to provide a database for lexicographers, grammarians, authors of textbooks, teachers and others interested in language studies; and e) to contribute to a wider understanding of the Portuguese language in Brazil as a social instrument of diversified communication, possessing various norms of use.

In this way, the - project was implemented in compliance - with Article 3 of Federal Decree 30.643, of March 20th, 1952, which states that the main purpose of the Philology Committee of the Casa de Rui Barbosa is to establish the Linguistic Atlas of Brazil[5].

At present, the entire corpus has been constituted and the first two volumes – *Volume I – Introdução* and *Volume II – Cartas linguísticas*–have been published (CARDOSO *et al*, 2014a;

5 Determination regulated by Federal Ordinance nº 536 of May 20th, 1952.

2014b). The Project benefited from the financial support of such bodies as the CNPq, CAPES[6], FAPESB[7]and the Fundação Araucária[8] as well as support from the home universities of the participating researchers.

The *corpus*

The team traveled throughout the country conducting surveys in 250 medium and large cities, of 26 states spread throughout Brazil's 8,515,767.049 km^2. The surveys included all capitals, except for Brasília (capital of the Federal District) and Palmas (Tocantins capital), because of the already mentioned recent foundation. The network extends from Oiapoque (point 001), in the far north, to Chuí (point 250), in the extreme south of the country, as shown on the map of the Points Network (*Rede de Pontos*–Figure 1).

Figure 1: CARTA V –BRASIL –Rede de pontos 'Map of network points' (CARDOSO *et al*, 2014b, p. 59)

The criteria for the definition of network points were: a) geographic distribution; b) history of the city; c) trajectory of development; d) demographic configuration; e) socio-cultural characteristics; and f) linguistic profile, where this was known. It also considered the rela-

6 Conselho Nacional de Pesquisa (National Council of Scientific Research) and Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (Coordination for Improvement of Higher Education Staff).

7 Fundação de Amparo à Pesquisa do Estado da Bahia (Bahia State Foundation for Research Support).

8 Fundação de Apoio ao Desenvolvimento Científico e Tecnológico do Paraná (Paraná State Foundation for Support of Scientific and Technological Research).

tionship between the total points and the population density of each state and each region. Other important input was provided by the proposal by Nascentes (1958), and previous studies for the establishment of networks of points of regional atlases and of interstate and international boundaries.

Recordings of a total of approximately 3,300 hours were made with 1,100 respondents, distributed by sex (male and female) and in two age groups (18-30 years and 50-65 years – covering two generations of speakers)[9]. In each capital, eight respondents were surveyed, stratified in two educational levels (up to 8 years of schooling and university level), because, owing to the government literacy programs, young people without basic education (4 years of schooling) are almost non-existent. In the other cities, because of the lack of individuals with higher education, informants have, at the most, 8 years of schooling. In each capital, eight informants were surveyed; in other cities, only four informants.

By controlling social and geographic variables, ALib provides a geo-sociolinguistic view of Brazilian Portuguese.

The interviews comprised a phonetic-phonological questionnaire (including prosodic issues), a semantic-lexical questionnaire (concerning various topic areas), and a morpho-syntactic questionnaire, in addition to questions of pragmatics, semi-directed speech, metalinguistic questions and a text for reading.

The informant record file contains social information (social level, socio-cultural habits, education) and provides information concerning the actual interview. There are 50 items in the file. Items 1 to 19 contain personal data of the informant; items 20 to 29 contain information on the degree of contact of the informant with the media, and their religious affiliation; items 30-37 concern the social life of the informant (entertainment and leisure); items 38-50, completed by the documenter after the audio recording, concern the documenter's impressions of this informant and the research environment (psychological characteristics of the informant, bystander interference etc.).

The questionnaires

In preparing the phonetic-phonological and morpho-syntactic questionnaires, facts of major interest were selected from the available data. The lexical-semantic questionnaire privileged items of greater magnitude that cover all or almost all of the national territory.

In this way, the three questionnaires were formulated: a) the phonetic-phonological questionnaire (QFF), with 159 questions, including prosody questions concerning the nature of in-

9 Although important for the study of changes in progress, the middle age-group (between 31 and 49 years) was discarded for operational reasons, because of the size of the area being investigated and the burden of the inclusion of a higher number of respondents.

terrogative sentences, statements and imperatives; b) the lexical-semantic questionnaire (QSL) with 202 questions covering 14 semantic areas; and c) the morpho-syntactic questionnaire (QMS), with 49 questions.

The Phonetic-phonological questionnaire (QFF)

The QFF sought to register phonetic variants in selected words. The questions devised in terms of the description of something so that the informant would pronounce the expected word. Some exemples of QFF are shown as follows:

- For the pronunciation of mid vowels in pretonic position

(1) T**E**L**E**VISÃO 'TELEVISION' (QFF 4)

... aquele aparelho onde se pode ver novela, jogo, programas...?

'… the appliance where one can watch soap-operas, games, programs…?'

(2) G**O**RDURA 'FAT' (QFF 22)

A carne de porco não é magra porque tem _____.

'Pork isn't lean because it has _____.'

- For the pronunciation of syllable-final /r/

(1) TO**R**NEIRA '(WATER) TAP' (QFF 12)

... aquilo que se abre quando se quer lavar as mãos numa pia?

'...the thing one opens when one wants to wash one's hands in the sink?'

(2) VARRE**R** 'SWEEP'(QFF 18)

Para limpar o chão, o que é que é preciso fazer (*mímica*)?

'To clean the floor, what is it necessary to do (*documenter mimics*)?'

- For the pronunciation of syllable-final /l/

(1) A**L**MOÇO 'LUNCH' (QFF 19)

... uma refeição que se faz, em geral, às 12 horas?

'...a meal generally eaten at 12 am?'

(1) AZU**L** 'BLUE'(QFF 89)

Que cor é esta? (*mostrar*)

'What color is this (*documenter shows*)?'

- For the pronunciation of syllable-final /S/

(1) ARRO**Z**‘RICE’ (QFF 21)

... o que se come no almoço, uns grãozinhos brancos que podem acompanhar o feijão, a carne?

‘...something that is eaten for lunch, small white grains that can go with beans and meat?’

(1) CA**S**CA ‘SKIN’ (QFF 31)

Para comer uma banana, o que é que se tira?

‘To eat a banana, what needs to be removed?’

(1) RA**S**GAR ‘TEAR’(QFF 88)

Fazer assim (*mímica*) em um papel é _____?

‘To do this with a piece of paper (*documenter mimics*) is to _____?’

- For prosody: interrogative sentences

(1) Eu vou sair hoje, doutor?(QP 6)

‘Am I going to get out today, doctor?’

Documenter says:

Uma pessoa está internada em um hospital e quer saber do médico se vai sair naquele dia. Como é que pergunta?

‘A person is interned in a hospital and asks the doctor if he/she will be checked out on that day. How does the person ask this?’

- For prosody: affirmative sentences

(1) Você vai sair hoje.(QP 1)

‘You will get out today.’

Documenter says:

E o médico, como é que responde?

‘And the doctor, what does the doctor say?’

- For prosody: exclammative sentences

(2) Oh, meu filho, saia da chuva!(QP 1)

‘Hey son, get out of the rain!’

Documenter says:

Como é que uma mãe diz ao filho para que ele saia da chuva?

‘How does a mother tell her son to get out of the rain?’

The semantic-lexical questionnaire (QSL)

The QSL sought to register lexical variants, characteristic of the locality, based on relevantsememes. Of the 14 semantic areas covered in the 202 questions, here are some examples:

- Atmospheric phenomena

(1) RELÂMPAGO 'LIGHTNING' (QSL 8)

... um clarão que surge no céu em dias de chuva?

'...a flash that happens in the sky on rainy days?'

- Stars and time

(1) ESTRELA CADENTE / ESTRELA FILANTE / METEORO / ZELAÇÃO 'SHOOTING STAR' (QSL 31)

De noite, muitas vezes pode-se observar uma estrela que se desloca no céu assim (*mímica*) e faz um risco de luz. Como chamam isso?

'At night, often a star can be seen moving through the sky like this (*documenter mimics*) and it makes a streak of light. What do you call it?'

- Fauna

(1) GALINHA-D'ANGOLA / GUINÉ / COCAR 'GUINEA FOWL' (QSL 67)

... a ave de criação parecida com a galinha, de penas pretas com pintinhas brancas?

'… a breeding bird similar to a chicken, with black feathers with white spots?'

- Children's games and amusements

(1) CAMBALHOTA 'SOMERSAULT' (QSL 155)

... a brincadeira em que se gira o corpo sobre a cabeça e acaba sentado? (*mímica*)

'... the game where you turn your body over your head and end up sitting? (*documenter mimics*)'

The morpho-syntactic questionnaire (QMS)

The QMS sought to register morpho-syntactic variants in questions focusing on the linguistic phenomenon or variants that could occur during the description of a picture or in an account given by the informant. The following are exemples of some of the questions:

- Definite article used with proper noun

(1) Tem filhos / irmãos? Como se chamam? O que eles fazem? (QMS 1)

'Do you have children / brothers? What are their names? What is their occupation?'

- Placement of NÃO in negative responses

(1) Você / o(a) senhor(a) sabe se tem vida em outro planeta / na lua? (QMS 47)

'Do you know if there is life on another planet / on the moon?'

(2) Você / o(a) senhor(a) já viu disco voador? (QMS 48)

'Have you ever seen a flying saucer?'

- TU / VOCÊ 'YOU' subject

(1) Quando se vê um amigo com uma mala e se quer saber para onde ele vai, como é que se pergunta? (QMS 24)

'When you see a friend with a suitcase and you want to know where they are going, what do you ask them?'

Pragmatic questions

The four pragmatic questions seek to ascertain how a young boy, finding an object on the street, addresses another youngster, male and female, or an elderly individual, also of both sexes, in order to tell them that the object fell from their pocket.

Semi-directed discourse

In the topics for semi-directed discourse, the informant must relate a personal experience, comment on a television program, describe his/her professional activity and tell a fact heard from another person of their acquaintance (reported speech).

Metalinguistic questions

The six metalinguistic questions ask about the local way of speaking, about people who, according to the informant, *speak differently* in the locality and in other Brazilian areas, and

also about the speech of earlier times. The questions solicit exemples of these variants. The following are two exemples of these questions:

(1) Como chama a língua que você / o(a) senhor(a) fala? (QM 1)

'What is the language you speak called?'

(2) No passado, falavam diferente aqui? (QM 6)

'In the past, did they speak differently here?'

The reading text

The text for the reading is the *Parable of the seven wicker baskets* used by Armando de Lacerda and Göran Hammarström (LACERDA; HAMMAARSTRÖM, 1953, p. 27-28), with the necessary adaptations.

Potential

In addition to using geolinguistics methodology, the ALiB is also a sociolinguistic atlas because of the social variables it monitors, presented earlier. It is important to highlight its innovation also with regard to the surveys, because in addition to phonetic issues (including prosody) and lexical (typical of a dialectology research), the ALiB also employed a morphosyntatic questionnaire, incorporated pragmatic and metalinguistic questions, recorded semi-directed discourse, and a reading. This variety of documentation furnishes material for analysis at various grammatical levels, permitting undoubtedly an overall profile of the reality of Brazilian Portuguese.

In order to make the ALiB corpus available for consultation and use by the scientific community and others interested in the area, the Project is building a database that permits access to the audio recordings and the linguistic maps of the collected material.

Many analyses based on the ALiB corpus have been presented in the form of conference papers, articles, Master's dissertations, PhD theses, books, book chapters, and others, including national and international publications, at present totalling more than a hundred research publications[10].

By way of illustration, we present the analysis of the variable realization of /t/, /d/ (dento-alveolar or palatalized) before the vowel /i/ (phonological) or [i] (derived[11]) in Brazilian capitals, following the theoretical and methodological framework of Labovian Sociolinguistics.

10 Further information concerning the ALiB is available at the site http://www.ALiB.ufba.br.
11 The term derived vowel refers to the realization of [i] of the archiphonemes /I/, /E/.

In addition to the geographical variable (cities: Aracaju, Belém, Belo Horizonte, Boa Vista, Campo Grande, Cuiabá, Curitiba, Florianópolis, Fortaleza, Goiânia, João Pessoa, Macapá, Maceió, Manaus, Natal, Porto Alegre, Porto Velho, Recife, Rio Branco, Rio de Janeiro, Salvador, São Luís, São Paulo, Teresina and Vitória), this analysis considered the linguistic and social variables, listed in the table below:

Linguistic variables	**Social variables**
Voicing: voiceless /t/ (*tio* 'uncle') or voiced /d/ (*dia* 'day')	Education: basic level or university level
Position in the syllable: initial (*tiara* 'tiara', *diabo*'devil'), medial (*prostituta* 'prostitute', *mandioca* 'cassava') or final (*bonitinho* 'cute', *papudinho* 'chubby cheeks'[1])	Sex: masculine or feminine
Tonic stress of the syllable: tonic (*adotivo* 'adopted', *tardinha* 'evening') or atonic (*estiar* 'to stop raining', *mandioca*'cassava')	Age-group: 18 - 30 years (Group 1) or 50 - 65 years (Group 2)
Preceding vowel: [a, ã] (*pratinho* 'little plate', *mandioca* 'cassava'), [E, e, ẽ] (*coletivo* 'bus', *dente* 'tooth', *pédianteiro* 'forefoot'), [i, ĩ] (*cabritinho* 'kid', *liquidificador* 'blender'), [ɔ, o] (*adotiva* 'adopted', *ondinha* 'ripple'), [u, ũ] (*sutiã* 'brassiere', *papudinho* 'chubby cheeks', *conjuntivite* 'conjunctivitis', *carrinho de mão*[2] 'wheelbarrow'), front semivowel [j] (*feitiço* 'spell') or back semivowel [w] (*São Tiago* 'St. James', *maldita*[3] 'accursed')	
Preceding consonant: alveolar constrictive [s, z] (*estilingue* 'slingshot', *patas dianteiras* 'front hooves'), palatal constrictive [ʃ, ʒ] (*prostituta* 'prostitute', *patas dianteiras* 'front hooves') or laryngeal or velar constrictives, or rhotic approximant or tap [h, ɦ, x, ɣ, r, ɾ, ɽ](*cortiça* 'cork', *jardineira* 'jardiniere')	
Nasality of vowel: oral (*estiar* 'to stop raining', *perdida* 'lost') or nasal (*cabritinha* 'kid', *ondinha* 'ripple')	
Type of questionnaire: more monitored (QFF, QSL and reading) or less monitored (other parts of the survey)	

Table 1: Independent variables

Realization of /t, d/ before phonological vowel /i/

For the analysis of the variable realization of /t, d/ before /i/ (phonological vowel) in words such as *tio*'uncle', *mentira* 'lie', *estilingue* 'slingshot', *dia* 'day', *perdida* 'lost', *diadema* 'tiara', data were considered from 14 capitals, from four Brazilian regions, namely: a) Northern region: Boa Vista (RR), Macapá (AP), Manaus (AM) and Porto Velho (RO); b) Northeast region: Aracaju (SE), João Pessoa (PB), Maceió (AL), Natal (RN), Recife (PE) and Teresina (PI); c) Midwest region: Campo Grande (MS) and Cuiabá (MT); d) Southern region: Curitiba (PR) and Florianópolis (SC).

The remaining 11 capitals were removed from the analysis because they display categorical realization of the palatalized variant: a) Northern region: Belém(PA) and Rio Branco (AC); b) Northeastern region: Fortaleza (CE), Salvador (BA) and São Luís (MA); c) Midwest region: Goiânia (GO); d) Southeastern region: Belo Horizonte (MG), Rio de Janeiro (RJ), São Paulo (SP) and Vitória (ES); e) Southern region: Porto Alegre (RS).

The distribution of the 2.327 data items are presented in Table2, as follows:

Variants	Ocurrences	Frequency
Palatalized	1575	68%
Alveolar	752	32%
Total	2327	100%

Table 2: Overall result for /t, d/ before /i/ (phonological vowel)

According to the results, the palatalized variant [tʃ], [dʒ] is predominant in Brazil – 68% against 32% of the alveolar variant [t], [d] before the phonological vowel /i/.

In order to verify the context favoring the palatalized variant, the data were submitted to the GoldVarb Program[12], using this variant as the application value[13]. Motivation from this option comes from the fact that the palatalized variant is considered more innovative, as it was the dento-alveolar realization that arrived in Brazil, and which has been maintained in Portugal until today.

The analysis revealed an initial *input* of 0,680 and a final *input* of 0,940, which shows that the palatalized variant is undergoing a process of actuation in Brazilian Portuguese. The *log-likelihood* of -639,921 and the level of significance of 0,000, attest to the statistical reliability of

12 SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. GoldVarb X – a multivariate analysis application. 2005. Toronto: Department of Linguistics; Ottawa: Department of Mathematics. Available in: http://individual.utoronto.ca/tagliamonte/ Goldvarb/GV_index.htm#ref. Acessed on 20th October, 2011.
13 The value that is the outcome of the variable rule.

the results.

The program selected as statistically relevant, in sequential order, the following variables: a) city; b) preceding consonant; c) syllable stress; and d) age-group. This selection shows that the variable phenomenon under study illustrates a particularly diatopic variation, which facilitates the definition of geographical areas in Brazil.

Next, we present the results for the variable *city* only, in the following section.

Realization of /t, d/ before /i/ (phonological vowel) and the variable *City*

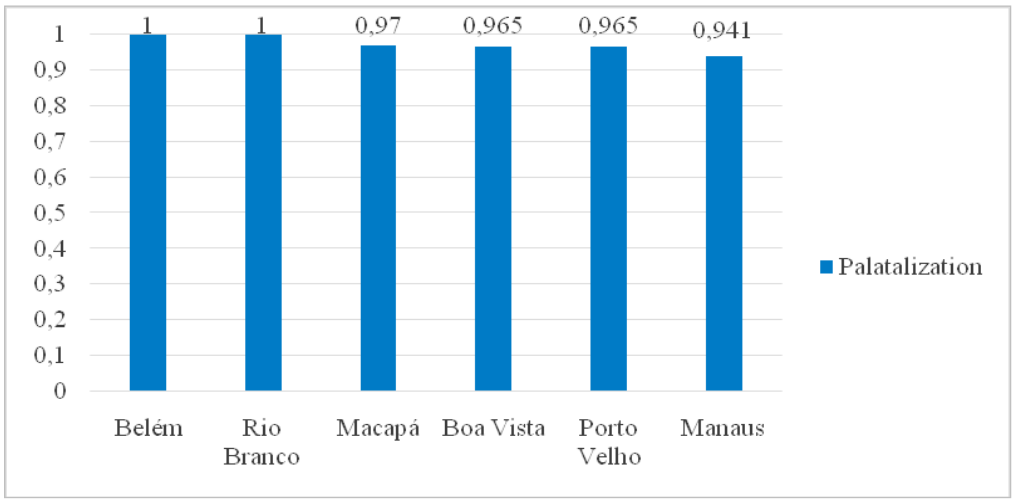
The results for the 14 cities displaying the dento-alveolar or palatalized realization of /t, d/ before /i/ (phonological vowel) are shown in Table 3 below:

City	Ocurrences/Total	Frequency	Relative weight
Macapá	187/188	99%	**0,970**
Boa Vista	182/183	99%	**0,965**
Porto Velho	171/172	99%	**0,965**
Manaus	206/208	99%	**0,941**
Teresina	137/140	97%	**0,893**
Campo Grande	200/204	98%	**0,885**
Curitiba	133/143	93%	**0,696**
Florianópolis	122/191	63%	0,210
Cuiabá	90/176	51%	0,132
Natal	38/126	30%	0,045
João Pessoa	23/125	18%	0,026
Recife	37/174	21%	0,025
Maceió	25/148	16%	0,020
Aracaju	24/149	16%	0,018

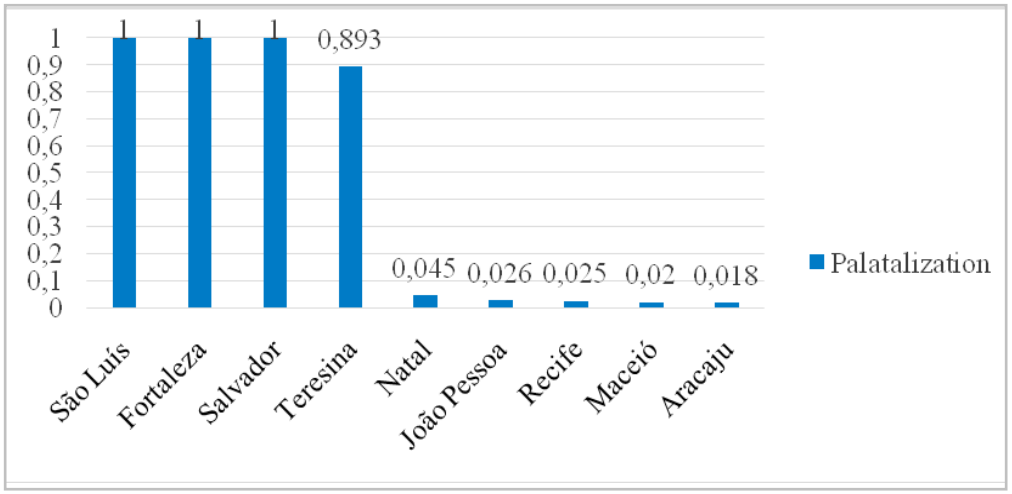
Table 3: Application of the palatalized realization of /t, d/ before /i/ (phonological vowel) by *City*

The results indicate that the rule of palatalization of /t, d/ before /i/ (phonological vowel) is favored in Macapá, Boa Vista, Porto Velho, Manaus, Teresina, Campo Grande and Curitiba. With the exception of this last city, in which the palatalization affects 93% of the data, in the other cities mentioned, it is a semi-categorical rule. The remaining capitals – Florianópolis, Cuiabá, Natal, João Pessoa, Recife, Maceió and Aracaju –inhibit the realization of the palatalized variant.

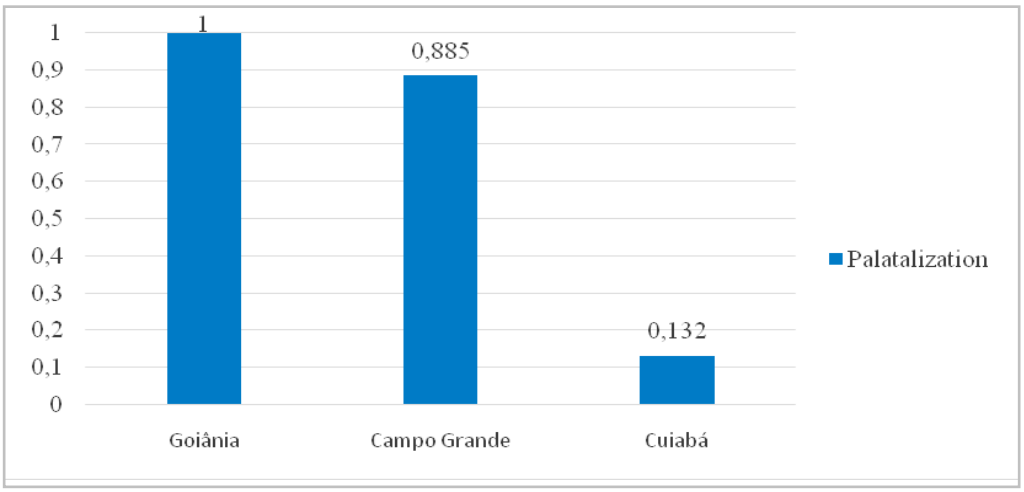
The following Graphs illustrate the distribution in Brazil by region, considering all the capitals:



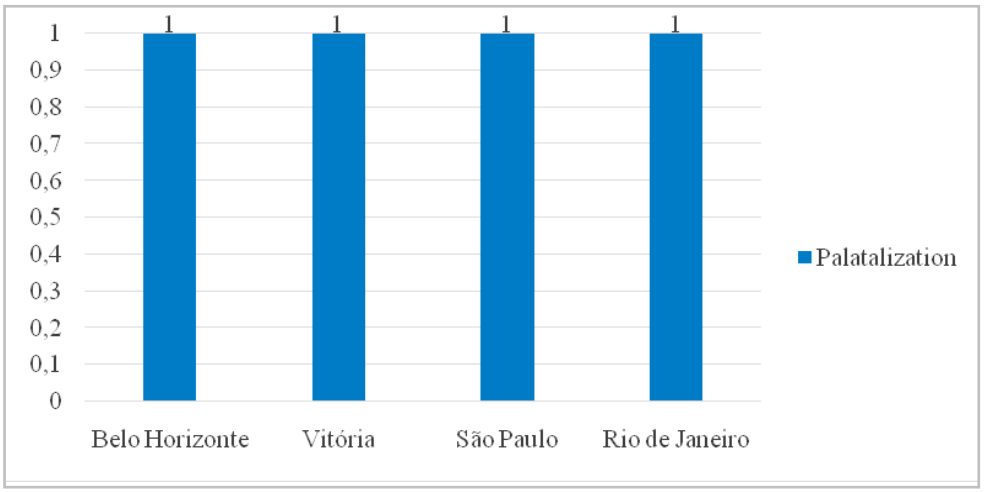
Graph 1: Palatalization of /t, d/ before /i/ (phonological vowel) in the Northern region (RW)



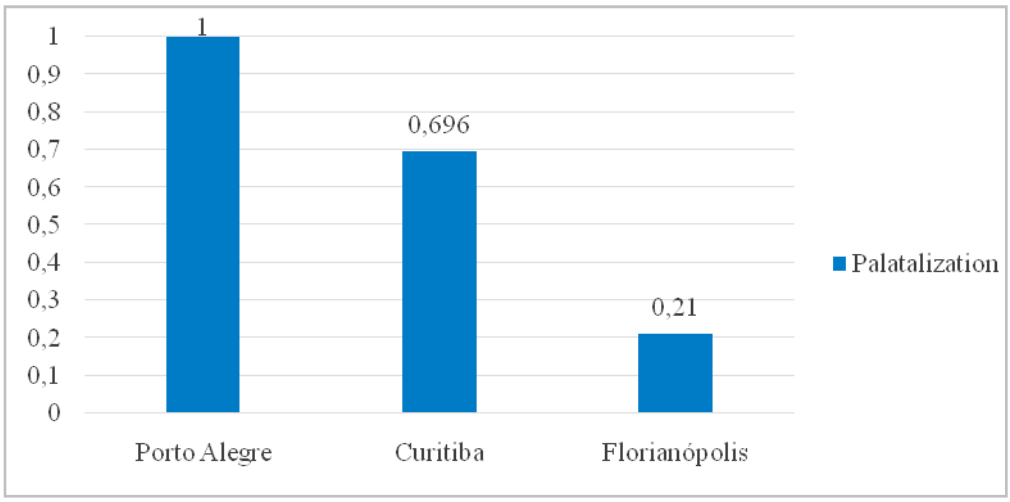
Graph 2: Palatalization of /t, d/ before /i/ (phonological vowel) in the Northeast region (RW)



Graph 3: Palatalization of /t, d/ before /i/ (phonological vowel) in the Midwest region (RW)



Graph 4: Palatalization of /t, d/ before /i/ (phonological vowel) in the Southeast region (RW)



Graph 5: Palatalization of /t, d/ before /i/ (phonological vowel) in the Southern region (RW)

Based on these findings, we can state that the following constitute areas of palatal realization: the entire North and Southeast of Brazil, part of the Northeast (except Natal, João Pessoa, Recife, Maceió and Aracaju) and of the Midwest (with the exception of Cuiabá) and part of the South (except Florianópolis). Therefore, the dento-alveolar realization persists in only 7 of the 25 capitals: Natal, Recife, João Pessoa, Maceió, Aracaju, Cuiabá and Florianópolis.

Realization of /t, d/ before the derived vowel [i]

For the analysis of the variable realization /t, d/ before [i] (derived vowel) in words such as *tesoura* 'scissor', *prateleira* 'shelf', *tomate*'tomato', *desmaio*'fainting', *redemoinho*'whirlwind', *tarde* 'afternoon', we considered data from 15 capitals, from four regions of Brazil, namely:a) the Northern region – Boa Vista (RR), Macapá (AP), Manaus (AM) and Porto Velho (RO); b) the Northeastern region: Aracaju (SE), João Pessoa (PB), Maceió (AL), Natal (RN), Recife (PE) and Teresina (PI); c) the Midwest region: Campo Grande (MS), Cuiabá (MT) and Goiânia (GO); d) the Southern region: Curitiba (PR) and Florianópolis (SC).

The remaining 10 capitals were excluded from the analysis because they display categorical realization of the palatalized variant: a) the Northern region: Belém (PA) and Rio Branco (AC); b) the Northeastern region: Fortaleza (CE), Salvador (BA) and São Luís (MA); c) the Southeastern region: Belo Horizonte (MG), Rio de Janeiro (RJ), São Paulo (SP) and Vitória (ES); d) the Southern region: Porto Alegre (RS).

The distribution of the 5.639 data items are presented in Table4, as follows:

Variants	Ocurrences	Frequency
Palatalizada	3721	**66%**
Alveolar	1918	34%
Total	5639	100%

Table 4: Overall result for /t, d/ before [i] (derived vowel)

According to the results, there is a predominance of the palatalized variant [tʃ, dʒ] in Brazil –occurring with a frequency of 66% as opposed to 34% for the alveolar variant [t, d],before the derived vowel [i] – a situation similar to that of the realization of /t, d/ before the phonological vowel /i/, presented earlier.

The data were then submitted to the GoldVarb program, using as the application value the palatalized variant, in order to ascertain the contexts that favor it. The initial input was 0,660 and the final input was 0,750, which indicates that the palatalized variant is being implemented in Brazilian Portuguese also in the context of the derived vowel [i], although to a slightly lesser degree than in the context of the phonological vowel. The log likelihood was -1,986.654 and the significance level was 0,014, which attests to the reliability of the statistical results also for this run.

The program selected as statistically relevant, in sequential order, the following variables: a) city; b) age-group; c) preceding vowel; d) preceding consonant; e) sex; f) education; and g) position in the syllable. This selection shows that the variable phenomenon under study, also in the context of the derived vowel [i], also indicates the relevance of diatopic variation, which permits the definition of geographical areas in Brazil.

For the derived vowel [i] context, we present only the results for the variable *city*, in the following section.

Realization of /t, d/ before [i](derived vowel) and the variable *City*

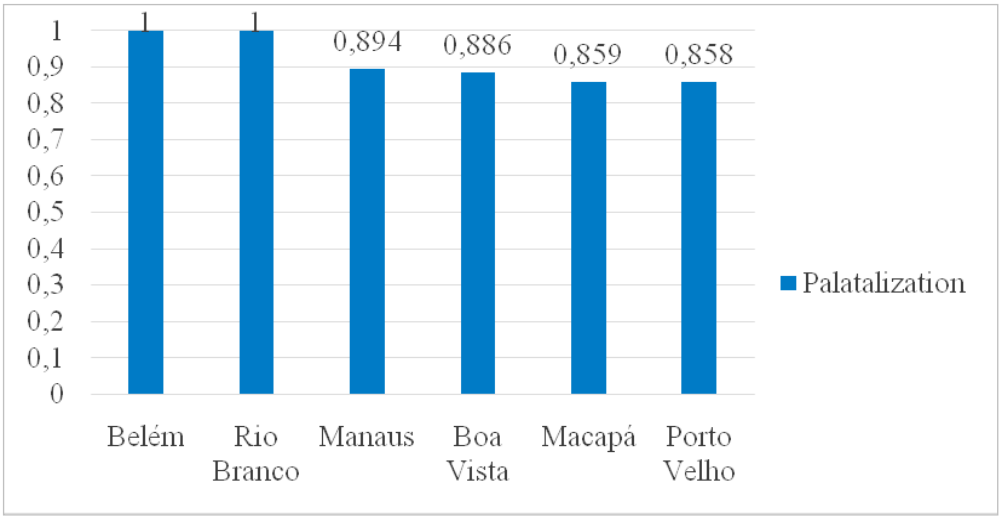
The results for the 15 cities displaying the dento-alveolar or palatalized realization of /t, d/ before [i] (derived vowel) are shown in Table 5 below:

City	Ocurrences/Total	Frequency	Relativeweight
Goiânia	433/446	97%	**0,907**
Manaus	579/601	96%	**0,894**
Boa Vista	300/312	96%	**0,886**
Macapá	446/468	95%	**0,859**
Porto Velho	415/436	95%	**0,858**
Campo Grande	420/443	94%	**0,845**
Teresina	293/320	91%	**0,763**
Florianópolis	259/445	58%	0,297
Cuiabá	130/372	34%	0,123
Curitiba[4]	96/283	33%	0,111
Natal	84/268	31%	0,102
Recife	92/341	26%	0,084
Aracaju	68/274	24%	0,073
João Pessoa	49/291	16%	0,047
Maceió	57/339	16%	0,047

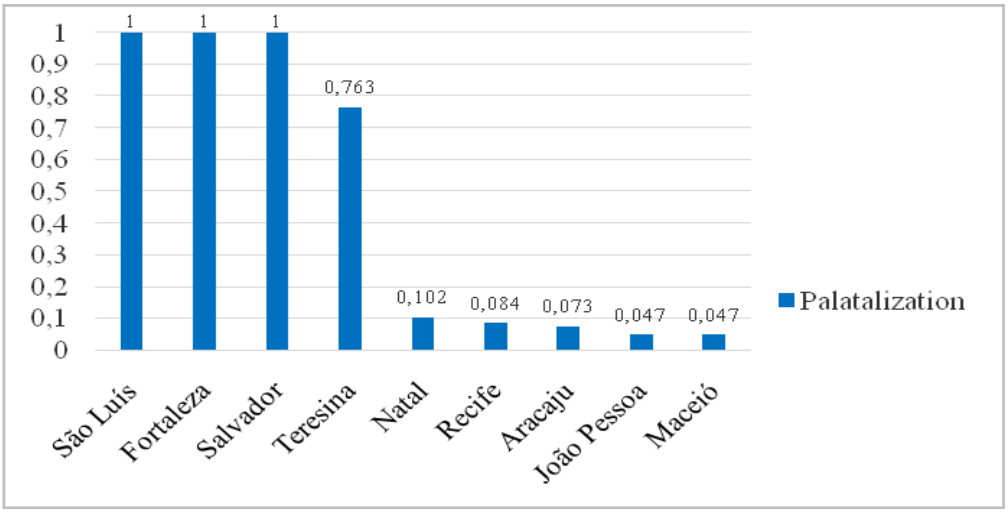
Table 5: Application of the palatalized realization of /t, d/ before [i] (derived vowel) by 'City'

The results indicate that the rule of palatalization of /t, d/ before [i] (derived vowel) is favored in Goiânia, Manaus, Boa Vista, Macapá, Porto Velho, Campo Grande and Teresina, constituting a semi-categorical rule in the first three cities. The palatalized realization is inhibited in Florianopolis capital, Cuiabá, Curitiba, Natal, Recife, Aracaju, João Pessoa and Maceió.

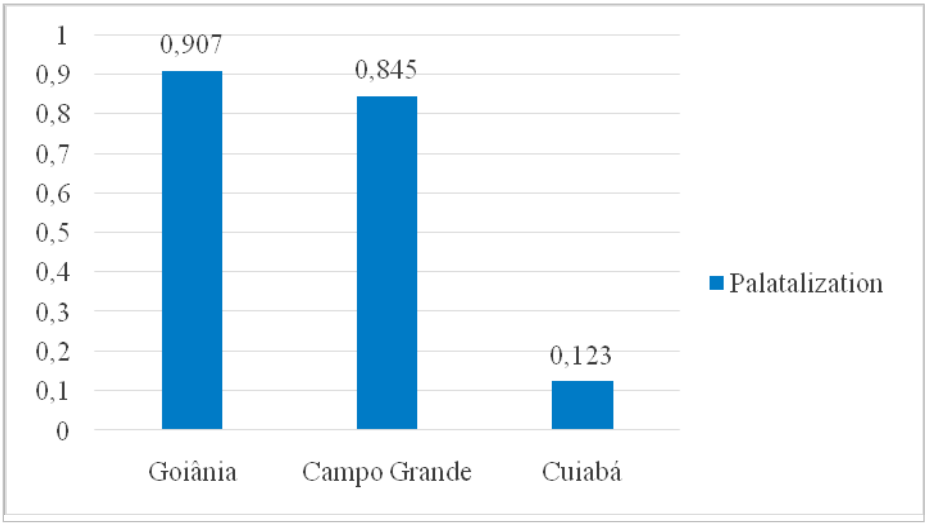
The following graphs illustrate this distribution throughout the country, by region, considering all the capitals:



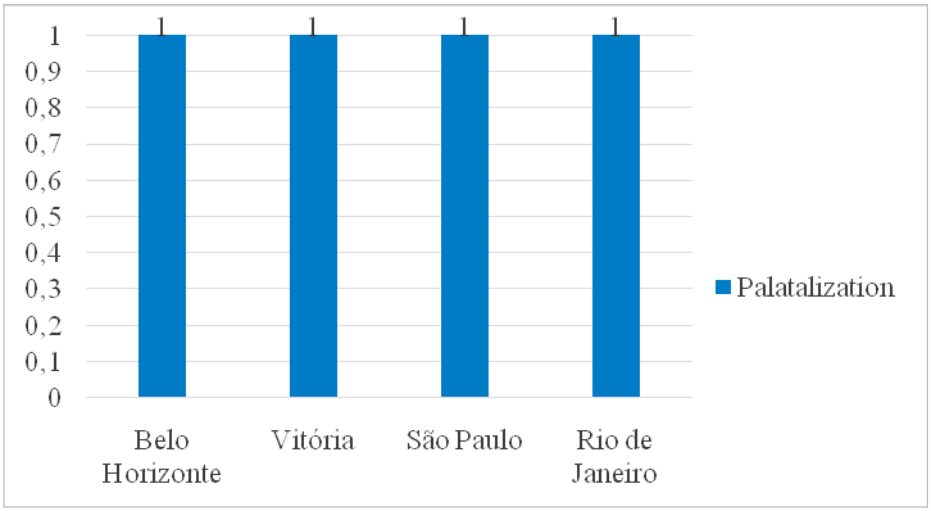
Graph 6: Palatalization of /t, d/before [i] (derived vowel) in the Northern region (RWs)



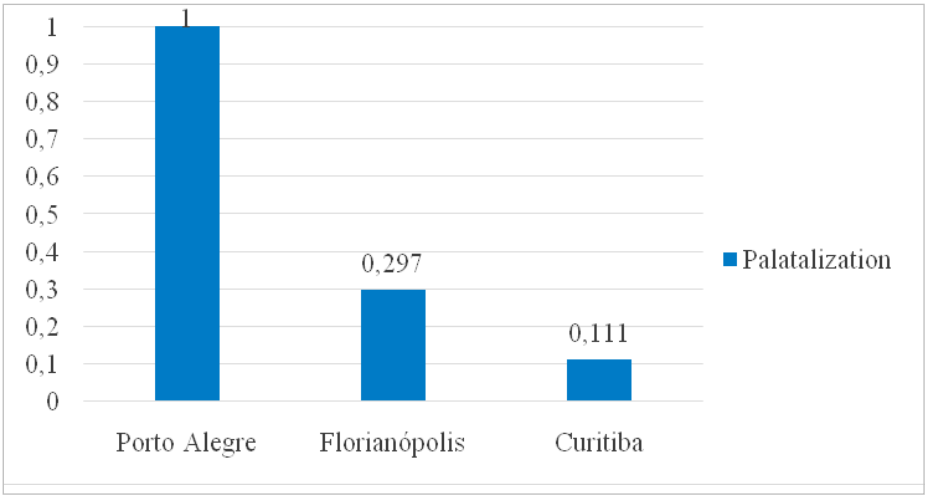
Graph 7: Palatalization of /t, d/before [i] (derived vowel) in the Northeast (RWs)



Graph 8: Palatalization of /t, d/before [i] (derived vowel) in the Midwest region (RWs)



Graph 9: Palatalization of /t, d/before[i] (derived vowel) in the Southeast (RWs)



Graph 10: Palatalization of /t, d/ before [i] (derived vowel) in the Southern region (RWs)

Based on the data presented, we can state that the area of palatal realization comprises the entire North and Southeast of Brazil, part of the Northeast (except Natal, Recife, Aracaju, João Pessoa and Maceió), part of the Midwest (except Cuiabá), and Porto Alegre, in the Southern region. Therefore, the alveolar realization persists in 8 of the 25 capitals: Natal, Recife, Aracaju, João Pessoa, Maceió, Cuiabá, Curitiba and Florianópolis.

Comparing the contexts

As can be seen in the previous sections, the results for the palatalization of /t, d/ are similar both in the context of the phonological vowel /i/ and in that of the derived vowel [i]. Table 6, below, showing the relative weights of all capitals, illustrates this fact:

City	Phonological vowel /i/	Derived vowel [i]
Northern Region		
Belém	**1**	**1**
Rio Branco	**1**	**1**
Macapá	**0,970**	**0,859**
Boa Vista	**0,965**	**0,886**
Porto Velho	**0,965**	**0,858**
Manaus	**0,941**	**0,894**
Northeastern Region		
São Luís	**1**	**1**
Fortaleza	**1**	**1**
Salvador	**1**	**1**
Teresina	**0,893**	**0,763**
Natal	0,045	0,102
João Pessoa	0,026	0,047
Recife	0,025	0,084
Maceió	0,020	0,047
Aracaju	0,018	0,073
Midwest Region		
Goiânia	**1**	**0,907**
Campo Grande	**0,885**	**0,845**
Cuiabá	0,132	0,123
Southeast Region		
Belo Horizonte	1	1
Vitória	1	1
São Paulo	1	1
Rio de Janeiro	1	1
Southern Region		
Porto Alegre	**1**	**1**
Curitiba	**0,696**	0,111
Florianópolis	0,210	0,297

Table 6: Application of palatalized realization of /t, d/ before /i/ (phonological vowel) and before [i] (derived vowel) by *City* (RWs)

Examining the results presented in Table 6, it is clear that the palatalization of /t, d/ begins in the context of a following phonological vowel /i/ in all capitals, with the exception of the cities of Natal, João Pessoa, Recife, Maceió and Aracaju – in the Northeast and Florianópolis – in the Southern region.

All these cities inhibit palatalization in both contexts, as in Cuiabá, in the Midwest region, which inhibits palatalization more frequently with the phonological vowel /i/. The behavior of Curitiba speakers is noteworthy, because they favor the palatalization of /t, d/ with the phonological /i/ vowel, but they inhibit it with the derived vowel [i]. This occurs because, in reality, they generally do not realize [i], retaining the pronunciation [e] in unstressed position, especially in the final syllable, as in *noi[te]* 'night' and *tar[de]* 'afternoon'.

Thus, with some generalization – since there are exceptions, as we have seen – we can say that the palatalization of /t, d/ before /i/ or [i] is being implemented in Brazilian Portuguese, and is even categorical in some capitals.

Diatopic representation

Linguistic map F06 C 1 of the *Atlas Linguístico do Brasil*–AliB (cf. CARDOSO *et al*, 2014b, p. 123), reproduced below, shows the area of the palatal realization of /t, d/ before phonological /i/ and derived [i], in Brazilian Portuguese, considering only the capitals, in percentage data, distributed in five groups.

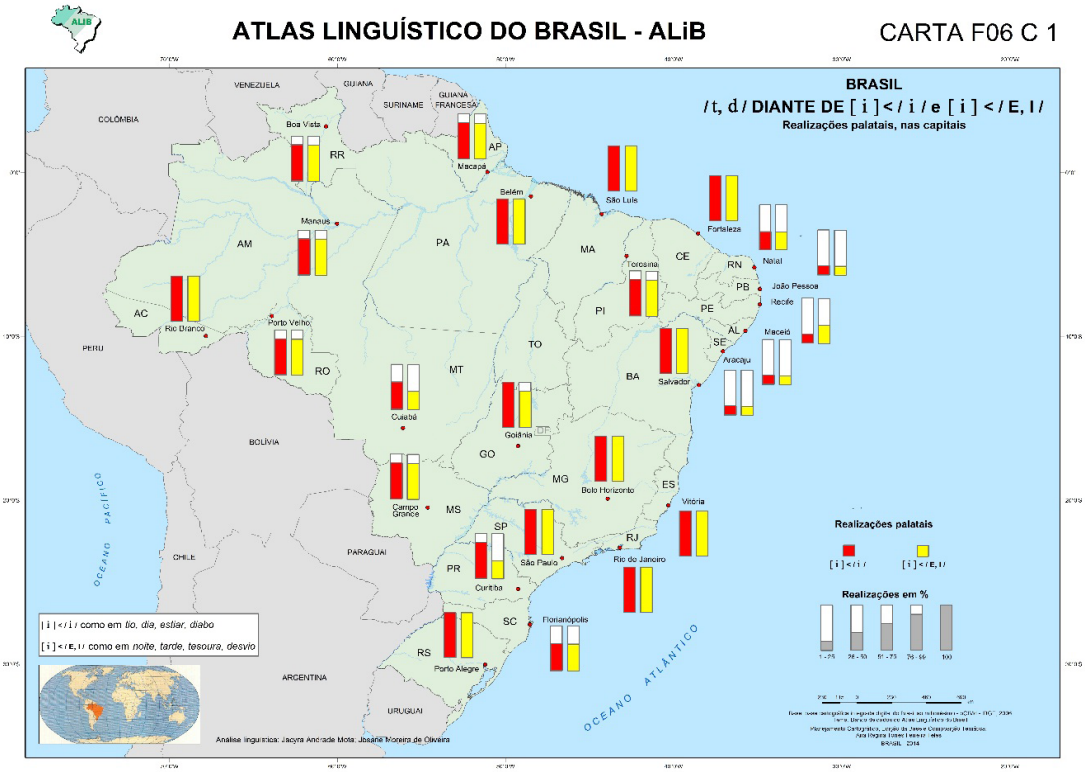


Figure 2: CARTA F 06 C 1 – BRASIL –/t, d/ before [i] < /I/ and [i]< /E, I/ – Palatal realizations in the capitals(CARDOSO et al, 2014b, p. 123)

As can be seen in map F06 C 1 ALiB, the area of alveolar realization, both with the phonological vowel /i/ and with the derived vowel [i], forms a continuum in the far east of the northeast coast.This realization is also found in Florianópolis, where it may be a remnant of the Azorean colonization in the history of the city, and is also found in Cuiabá and, with derived vowel [i], in Curitiba.

As far as Cuiabá is concerned, it must be said that the current area of Mato Grosso was only defined in the early eighteenth century, about 1719. The region attracted waves of individuals in search of gold, including people from Pernambuco, a region characterized by the realization of alveolar /t, d/ before [i]. Cuiabá was elevated to the status of capital only in 1825, and its development was delayed due to internal politics and its distance from the political and economic center of the country.

According to Palma (2005), in the speech of Cuiabá, the palatal realization of /t, d/ before [i] is prestigious only among young people, who mimic carioca (i.e. Rio de Janeiro) speech. The same author attests the realization of palatalized [tʃ, dʒ] also in words such as *cachorro* 'dog', *peixe* 'fish' and *jogo* 'game', which are quite stigmatized socially, a fact that could explain the maintenance of the alveolar realization before [i]. In Curitiba, as already noted, the preservation of the dento-alveolar variant is due to the mid-high realization of the archiphonemes /I, E/, in unstressed positions following /t, d/.

Conclusions

The ALiB project entails an innovative methodology and an open policy of distributing information in order to reach different layers of society.

The ALiB is an atlas that permits the description of the linguistic reality of Brazil based on field research, the systematic collection of oral data by means of a unified methodology. The data represent the Brazilian geolinguistic reality, facilitating the drawing of dividing lines or approximate areas identifying the map of dialectal regions of Brazilian Portuguese.

The knowledge of the diversity of uses in Brazilian Portuguese and of the variation registered from one region to another can contribute both to a better equation of the teaching and learning of Portuguese and also to an understanding of linguistic diversity without stigmatizing speakers.

Comprising a large corpus compiled by applying a set of questionnaires aimed at identifying facts at all grammatical levels, to be accessed from a multi-dimensional perspective–according to locality, social level, generation,style register,sex, and context – the ALiB Project is open to different lines of analysis.

In addition, that what was revealed when we analyzed the pronunciation of variable /t, d/ before /i/ and [i] in the Brazilian capitals. We demonstrated that the alveolar or palatal re-

alizations of these phonemes are geographically conditioned, in conjunction with other variables,which were mentioned but not discussed. This is merely one example of the potential that a project of this size presents.

May the ALiB *family* continue to make it possible for further studies on the Portuguese of Brazil grow and multiply!

References

CARDOSO, S. A. M. S. O Atlas Linguístico do Brasil: uma questão política. *Atas* do Seminário Nacional Caminhos e Perspectivas para a Geolinguística no Brasil. Salvador: UFBA, 1996, p. 87-96.

CARDOSO, S. A. M. S. *et al*.*Atlas linguístico do Brasil*, v.1 (Introdução). Londrina: EDUEL, 2014a.

CARDOSO, S. A. M. S.*et al*.*Atlas linguístico do Brasil*,v.2 (Cartas linguísticas 1). Londrina: EDUEL, 2014b.

COMITÊ NACIONAL do Projeto ALiB. *Atlas Linguístico do Brasil:*questionários 2001. Londrina: EDUEL, 2001.

GILLIÉRON, J.; EDMONT, E. *Atlas Linguistique de la France*. Paris: Honoré Champion, 1902-1910.

KURATH, H. *etalii*. *Linguistic Atlas of New England* (LANE). Providence: Brown University Press, 1939-1943.

LABOV, W. *The social stratification of English in New York City*. Washington: Center of Applied Linguistics, 1996.

LABOV, W. *Sociolinguistics patterns*. Philadelphia: University of Pennsylvania Press, 1972.

LABOV, W. *Principles of linguistic change* (internal factors). v.1. Oxford: Blackwell, 1994.

LABOV, W. *Principles of linguistic change* (social factors). v.2. Oxford: Blackwell, 2001.

LABOV, W. *Principles of linguistic change* (cognitive and cultural factors). v.3. Oxford: Willey-Blackwell, 2010.

LACERDA, A.; HAMMARSTRÖM, G. *Transcrição fonética do português normal*. Coimbra: [s/e], 1953.

NASCENTES, A. *Bases para a elaboração do Atlas Linguístico do Brasil*. v. 1. Rio de Janeiro:

MEC / Casa de Rui Barbosa, 1958.

PALMA, M. L. C. O falar cuiabano em Mato Grosso – estigma, status e atalhos.In: ALMEIDA, M. M. S; COX, M. I. P. (org.). *Vozes cuiabanas*: estudos linguísticos em Mato Grosso. Cuiabá: Cathedral, 2005.

POP, S. *La dialectologie*. Aperçu historique et méthodes d'enquêtes linguistiques, v. 1 and 2. Louvain: Chez l'Auteur / J. Duculot, 1950.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *GoldVarb X* – a multivariate analysis application. Toronto: Department of Linguistics; Ottawa: Department of Mathematics, 2005. Available at in http://individual.utoronto.ca/tagliamonte/Goldvarb/GV_index.htm#ref. Accessed on October 20th, 2011.

THUN, H.; ELIZAINCÍN, A. *Atlas Lingüístico Diatópico y Diastrático del Uruguay* (ADDU). Kiel: Westensee, 2000.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for a theory of language change. In: LEHMANN, W. P.; MALKIEL, Y. (Ed.) *Directions for historical linguistics*. Austin: University of Texas Press, 1968.

Sobre Dinah Callou — *Josane Moreira de Oliveira*[14]

Sobre Dinah Callou — *Jacyra Andrade Mota*[15]

14 Este texto é dedicado a Dinah Callou, minha orientadora de Doutorado, em reconhecimento e homenagem pelo seu trabalho na Linguística Brasileira.

15 Este texto, em homenagem à colega e amiga Dinah, lembra a sua presença no Projeto ALiB, que ela acompanha, desde o início, e tem ajudado a implementar.