

PLURICENTRIC LANGUAGES WORLDWIDE
SERIES 2: RESEARCHING PLURICENTRIC LANGUAGES

**Rudolf Muhr, Eugênia L. Duarte,
Cilene Rodrigues, Juan Thomas (eds.)**

Pluricentric languages in the Americas



pcl
PRESS
Graz / Berlin

Pluricentric languages in the Americas

PLURICENTRIC LANGUAGES WORLDWIDE
SERIES 2: RESEARCHING PLURICENTRIC LANGUAGES

Edited by

**Rudolf Muhr, Juan Thomas, Eugenia L. Duarte,
Reglindis De Ridder, and Dawn Marley**



VOLUME 1

Bibliographic Information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at <http://dnb.d-nb.de>.

ISBN: On the back cover of the book

PCL-PRESS Graz 2022 (<http://www.pcl-press.org>)

All rights reserved.

PCL- Press is the publishing house of the International Working Group on Non-Dominant Varieties of Pluricentric Languages and their Non-Dominant Varieties (WGNDV) and a subdivision of the Austrian German Association (AGA).

Printed by epubli.de (Berlin) and obtainable as printed copy via <https://www.epubli.de/shop>
Downloadable as an ebook via <http://pcl-press.org/>

All papers of this publication have been peer-reviewed.

All parts of this publication are protected by copyright. Any utilisation outside the strict limits of the copyright law, without the permission of the publisher, is forbidden and liable to prosecution. This applies in particular to reproductions, translations, microfilming, and storage and processing in electronic retrieval systems.

PLURICENTRIC LANGUAGES WORLDWIDE
SERIES 2: RESEARCHING PLURICENTRIC LANGUAGES

**Rudolf Muhr, Eugênia L. Duarte,
Cilene Rodrigues, Juan Thomas (eds.)**

Pluricentric languages in the Americas



pcl
PRESS
Graz / Berlin
2022

Table of Contents

I. Indigenous Pluricentric Languages in the Americas	
1. Cilene RODRIGUES (Rio de Janeiro, Brazil): Migration within Tupi-Guarani Languages	9
2. Liliana SÁNCHEZ (Chicago, USA): Is Quechua II a pluricentric language family?	31
II. The Pluricentricity of Spanish in the Americas	
3. Sebastian GREUSSLICH (Bonn, Germany): The Pluricentricity of Spanish in the Americas – Current Perspectives on Variety Dominance, Standardization, and the Media	47
4. Benjamin MEISNITZER & Dennis SCHMECHEL (Leipzig, Germany): Pluricentricity and the Varieties of Spanish in Central America	75
5. Juan THOMAS (Utica NY, USA): What taboo Anglicisms say about U.S. Spanish	89
6. Soledad CHÁVEZ FAJARDO (Santiago, Chile): The (In)Visible (Hispanic) American in Hispanic Lexicography. 19th and 20th Centuries. Some Cases	105
7. Meredith CHURCH (Boulder, USA): Unstressed Vowel Devoicing in Andean Spanish	121
III. The Pluricentricity of Portuguese in the Americas and worldwide	
8. Augusto SOARES DA SILVA (Braga, Portugal): Portuguese, pluricentricity and Brazilian Portuguese: A case of a reverted asymmetry?	135
9. Maria Eugênia L. DUARTE / Christina A. GOMES / Maria da Conceição PAIVA (Rio de Janeiro, Brazil): Beyond the dichotomy Dominant and Non-Dominant varieties of Pluricentric Languages: the case of Brazilian Portuguese	157
10. Dinah CALLOU / Carolina SERRA / Aline FARIAS (Rio de Janeiro, Brazil): On R-deletion in final coda position: regional diversity in Brazilian Portuguese and syllable phonology	173
11. Christina GOMES / Marcelo L. MELO (Rio de Janeiro, Brazil): Linguistic and Social Patterns of Sound Variation in two Contemporary regional varieties of European and Brazilian Portuguese	189
12. Raquel Meister Ko. FREITAG (Sergipe, Brazil): Mobility and higher education in grammatical patterns of Brazilian Portuguese	201

IV. The Pluricentricity of French and Dutch in the Americas

13. Karine GAUVIN (Moncton, New Brunswick, Canada): On the matter of an endogenous norm in Acadia (New Brunswick, Canada)	219
14. Eric MIJTS (Aruba, Aruba): Pluricentric languages in the Americas: the case of Dutch in the Dutch Caribbean	231

Preface

This is the first of two volumes that present part of the outcome 9th World Conference on Pluricentric Languages that took place from August 26-28 2021. The volume contains fourteen contributions. The conference was organised by the “Working Group on Non-Dominant Varieties of Pluricentric Languages” (WGNDV), this time hosted by the Austrian German Research Centre in Graz. Due to the pandemic and the complications for travelling and health connected with it, the conference was held online via ZOOM much to the satisfaction of all participants. The conference had three main topics: (1) Pluricentric languages in the Americas; (2) The localisation of global audiovisual and print media in pluricentric language areas. (3) Pluricentric languages worldwide. The papers of this volume refer to topic (1) while the papers of theme (3) are published in volume (2).

All articles of this volume deal with pluricentric languages that exist in the Americas. It is the first volume of this kind that has been ever published on this topic. Section (1) contains two papers that concern the indigenous language Quechua and the Guarani-Tupi language family. Cilene Rodriguez deals in her paper with the Guarani languages, which might be classified as pluricentric languages. She presents an overview of some phonemic and grammatical aspects of these languages. Liliana Sanchez deals with Quechua that is present in three South-American countries. She provides a brief introduction to the complexity of the language family and an overview of current legislation and policies in the three Andean countries with the largest Quechua-speaking populations.

Section (2) contains five papers about the pluricentricity of Spanish in the Americas. Sebastian Greusslich gives an overview about Spanish in the Americas and shows that there is a dominance hierarchy among these varieties. And there are specific social conditions as well as political aims relevant for standardization efforts in different Hispanic Latin American countries. The author points to the impact of mass media on these processes. Benjamin Meisnitzer and Dennis Schmechel discuss whether or not there is a Central American standard variety of Spanish – or several – based on the analysis of a feature catalogue for Central American phonetics and morphosyntax that was compiled and investigated by Quesada Pacheco (2010/2013). Juan Thomas describes fourteen Anglicisms that are homonyms with a taboo word in Spanish or that give a taboo meaning from English to another word present in U.S. Spanish. His study wishes to shed light on how and why these taboo expressions can enter Spanish if they truly are unique to the Spanish of the U.S. Soledad Chávez Fajardo discusses various ideologising instances present in Spanish-American lexicography such as some aspects related to the question of Eurocen-

trism, which results in a certain Spanish lexicographical ignorance of the Latin American reality. Meredith Church's research concerns the phonetic phenomenon of unstressed vowel devoicing in Andean Spanish assessing the patterning of this phenomenon as a characteristic of a non-dominant variety of Spanish.

In section three about Portuguese in the Americas and worldwide there are five papers. Augusto Soares da Silva explores the status of the Brazilian variety of Portuguese within the pluricentricity of Portuguese and argues that the great influence of Brazilian culture, resulting from the proliferation of Brazilian audiovisual cultural products, is gradually leading to a reverted asymmetry. Four papers deal with specific linguistic features of Brazilian Portuguese often compared to other varieties of Portuguese. The paper of Eugenia Duarte et al. compares four different varieties of Portuguese that in relation to EP have developed convergent linguistic patterns; but differ from BP in relation to the other features like referential pronominal subjects. Dinah Callou et al. deal with the phonetic development of Brazilian Portuguese in respect to R-deletion in final coda position which shows considerable regional diversity in Brazil. Christina Gomes and Marcelo L. Melo researched the social patterns of sound variation in two regional varieties of European and Brazilian Portuguese which turned out to be related to the adoption of an exogenous pattern that was introduced by the royal family at the beginning of the 19th century. Raquel Meister Ko. Freitag explores the hypothesis that a trigger for the further development of Brazilian Portuguese is the effect of the changes in higher education on linguistic patterns which is due to a strong enlargement of the Brazilian educational system that has changed the profile of students. In the last section of the volume Karine Gauvin explores the question why an endogenous norm has not developed of French in Acadia (Canada). Eric Mijts gives an overview about the complex language situation of Dutch in Aruba, where a non-dominant variety of Dutch has developed, which is strongly influenced by the iterative migration of large groups of Dutch Caribbean citizens to the Netherlands.

The editors would like to thank the authors for their contributions and would also like thank those colleagues who – in addition to the editors – acted as reviewers. Our sincere thanks also go to Dawn Marley and the student proof readers at the University of Surrey who corrected the texts. The government of the federal state of Styria is thanked for its financial support of the publication.

Rudolf Muhr, Eugênia L. Duarte, Cilene Rodrigues, Juan Thomas
Graz, Rio de Janeiro, and Utica in June 2022

I. Indigenous Pluricentric Languages in the Americas

Cilene RODRIGUES

(PUC-Rio - Pontifical Catholic University of Rio de Janeiro, Brazil)
crodrigues@puc-rio.br

Migration within Tupi-Guarani Languages¹

Abstract

Around 2,000 years ago, the Guarani languages, a subgroup of the Tupi-Guarani family (Tupi Stock), migrated from the Amazon to the lowlands of South America forming the modern Guarani languages, which might be classified as pluricentric languages. The present paper aims at presenting an overview of some phonemic and grammatical aspects of these languages, focusing on possible effects of their expansion southwards. First, it is shown that Guarani languages display a founder effect, a reduction in phonemic variability. Second, it is hypothesized that the possible migratory routes explored by Xetá and Aché might have resulted in language contact and depopulation, causing internal grammatical changes. Third, intensive language contact with positive language attitude and practices might have favored Paraguayan Guarani, which has the biggest number of speakers and the largest phonemic inventory among Guarani languages.

1. Introduction

Formation of pluricentric languages can involve long distance migrations. A group of individuals from population A spreads across the territory carrying A's language, but reaching unconnected distances with respect to A. Also, speakers can cross international borders, thus being subject to different political regulations that can affect their language. With that in mind, in this paper, we assess some historical migrations within Tupi-Guarani (Tupi stock), and their grammatical consequences, focusing on varieties of Guarani, which are spoken in the lowlands of South America. The following topics will be addressed:

- (a) Consequences of migrations to languages
- (b) How language contact and depopulation can affect founder languages
- (c) Influences of language attitudes and practices on the grammar of founder languages

¹ The study was financially supported by CNPq - Brazilian National Council for Scientific and Technological Development (Grant # 311093/2019-1).

An often-ignored treasure of South America is linguistic diversity, and little is known about the history of South American native languages. Thus, most of them are out of the scope of current research on grammar and its relationship with historical issues, such as migrations, depopulation and language contact. In the present paper, however, we emphasize that studies on these languages can bring important insights to our understanding of language as a biosocial phenomenon.

South America houses one fourth of the world's languages (Campbell & Grondona 2012), the majority of which are spoken by minorities, communities with less than 1 million speakers. According to Lewis (*Ethnologue: languages of the world* 2009), there are currently 624 languages in South America, but 179 of them are dying. Brazil is the country with the largest number of native languages within South America, but Brazilian native languages are very small (Rodrigues 2013), and Brazil is listed as the third country in terms of language loss (Lewis 2009). Therefore, we need to draw attention to the importance of these languages.

The paper is organized as follows: section 2 presents the history of Tupi, focusing on the phonemic consequences of migration within Guarani. In section 3, we discuss how certain grammatical features of two modern Guarani languages, Aché and Xetá, might reflect issues related to language contact and depopulation during migration and settlement in the south. Section 4 is devoted to Paraguayan Guarani, a discussion on how, under intense contact with Spanish, Paraguayan Guarani's grammatical system was preserved, arguably due to positive local social attitudes and language practices. Our main conclusions are in section 5.

2. Founder effects: a consequence of migration

Within genetics, a *founder effect* is defined as a case of genetic drift related to territorial expansions (Mayr 1954, 1963), when a small subgroup of individuals (the founder) gets territorially disconnected from its origins. Carrying just a small fraction of the total genetic variants of the original population, it may experience a founder effect, a reduction in its genetic pool, with a cline in heterozygosity and allelic richness.

Examples of founder effects are abundant. For example, Marques & Renesto (2017) report that, due to alterations in their habitat (introduction of dams), *Hemiodus orthonops*, a small fish endemic from the Paraná-Paraguay basin river, migrated to the upper part of the Paraná River, and genetic analyses of the

genotype of two founder populations indicate that they display a low mean of alleles per locus and an excess of homozygotes. Another example is that the rate of blood-type “O” among native Americans is 98%, but these natives descend from Asians, among whom a balanced 50% rate of “O” type is attested. Thus, Cavalli-Sforza (2001) hypothesizes that the founders who crossed over the Bering strait towards North America were exclusive carriers of the “O” allele.

Arguably, the most stunning and impactful human geographic expansion was the exodus from Africa in the lower Paleolithic period (~ 60,000 to 100,000 B.P.). This territorial expansion gradually resulted in the occupation of all continents, with Oceania and the Americas being the last ones to be occupied. This expansion went hand in hand with a subsequent serial founder effect: the further a modern human population is from Africa the poorer it is in genetic diversity. Sub-Saharan native African populations have an exceptional number of genetic variants, whereas native populations of the Americas and Oceania display a significant reduction (Cavalli-Sforza 2001, Ramanchandran et al. 2005), as shown in figure 1.

Africa	16.6	Europe
	20.6	Asia
	22.6	America
	24.7	Oceania

Fig. 1: Genetic distance between Africa and other continents, based on Cavalli-Sforza (2001:52)

Another impact of this great migration is on language. The phonemic inventories of non-African modern languages display a founder effect.

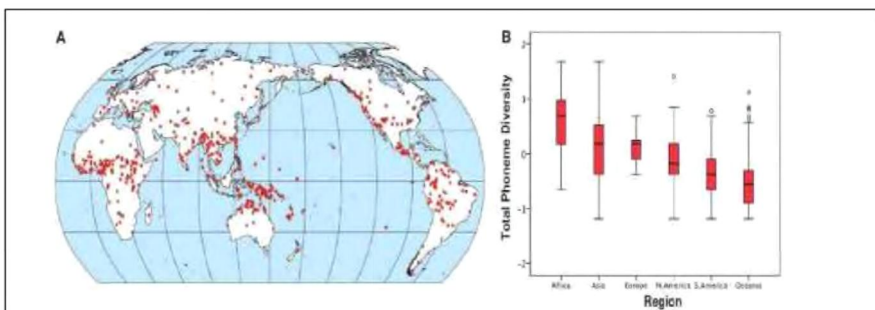


Fig. 2: A: location of analyzed languages, B: Phonemic diversity plotted against region (Atkinson 2011: 333).

Based on data from WALS - *The world Atlas of Languages Structures* (Dryer and Haspelmath 2005), Atkinson (2011) conducted a statistical analysis on a

sample of 504 modern languages from different families and continents. A correlation between phonemic reduction and physical distance from Africa was found, the further a language is from eastern Africa, the smaller its phoneme inventory is. (See Fig. 2).

Aktinson's findings correspond to some other important facts about phoneme inventory size. First, languages with too many neighbors present less phonemic variability (Creanza et al. 2015), but languages spoken by large populations have more phonemes than languages spoken by small populations (Trudgill 2005, Hay & Bauer 2007). Also, languages spoken by large populations increase their phonemic inventory over time at a rate between 0.26 and 0.38 per 1,000-year period (Perreault & Mathew 2012). Importantly, however, sonorant consonants and monophthongs seem to be less affected by this type of fluctuation. Spearman correlations between the means of language family, population and phonemic inventory size, were not significant for sonorant consonants and monophthongs (Hay & Bauer 2007:394)

2.1. Tupian languages: birthplace and migrations

Evidence from archeology, linguistics and genetics indicates that Proto-Tupi emerged around 5,000 years ago in the central-western part of the Amazon, in the Madeira-Guaporé Region (MGR), an area bounded by the rivers Amazon, Tocantins, Madeira and Guaporé, arguably where the current Brazilian state of Rondônia is located (Rodrigues 1964, Brochado 1984, Noelli 1998).

Around 3,000 years ago, Tupi started branching into different languages (Rodrigues 1964, Urban 1992, 1996) which corresponds to 10 modern language families (Rodrigues 1986, Rodrigues & Dietrich 1997). Currently, five of these language families are still located in MGR: *Arikém*, *Mondé*, *Ramaráma*, *Tuparí*, *Puruborá*, while others are located either to the north or to the east of MGR (*Awetí*, *Juruna*, *Mauwé*, *Mundurukú*), as shown in figure 3. Tupi-Guarani is an exception.

Fig. 3: Approximate location of Tupi-Languages: **I** *Arikém* (a. Karitiana, b. †Arikén), **II** *Aweté* (Awetí), **III** *Juruna* (a. Juruna, b. Xipaya, c. † Minitawá), **IV** *Mawé* (Mawé), **V** *Mondé* (a. Mondé, Aruá, c. Gavião, d. Suruí, e. Cinta-Larga, f. Zoró), **VI** *Mundurukú* (a. Mundurukú, b. Kuruáya), **VII** *Puruborá* (Puruborá), **VIII** *Ramaráma* (a. Karo, b. Urumi), **IX** *Tuparí* (a. Tuparí, b. Wayoró, c. Menkéns, d. Makuráp, e. Sakirabiát, f. Kepkiriwát). (Rodrigues, 2007: 108)

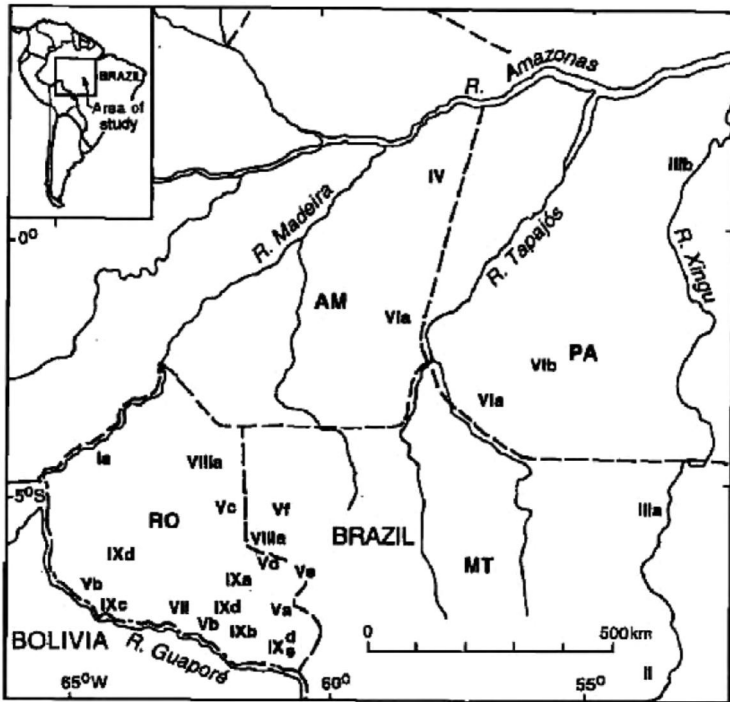


Fig. 3 Tupi languages with approximate locations

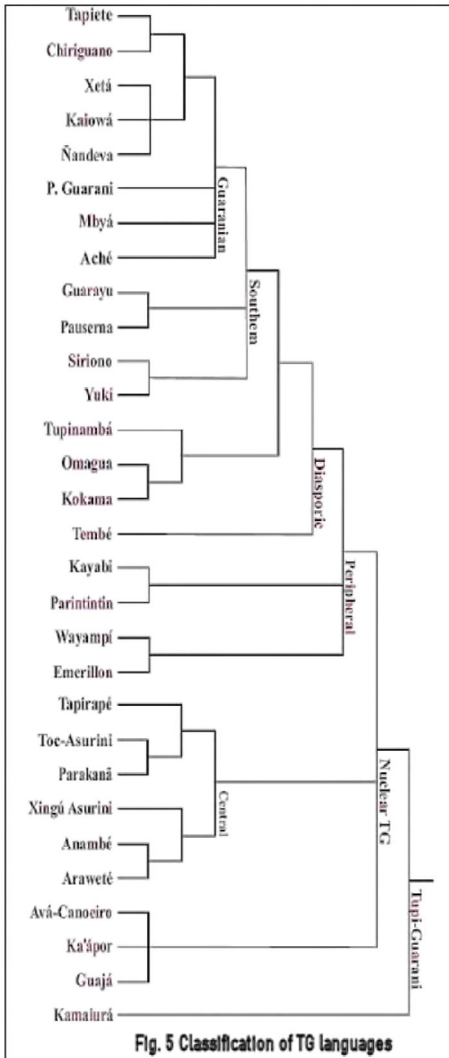
Tupi-Guarani (TG) is the biggest language family within the Tupi stock. It contains approximately 40 languages (Rodrigues 1985) which can be divided into 8 subgroups based on grammatical similarities (Dietrich 1990, Rodrigues & Cabral 2002, Michael et al. 2015). TG languages are spread out within the South American territory and are currently spoken in Brazil, Bolivia, Argentina, Paraguay, Peru and French Guiana. It is hypothesized that right after the formation of Proto-TG, its speakers fanned out in a radial fashion across South America (figure 4). The reason for this dispersal is unknown (see Noeli (2008) for a hypothesis based on demographic growth), but it was fast and took TG languages far from MGR. They expanded northward (Emerillon- Wayampi) and southward through routes in the east (Tupinambá) and in the west (Guarani). Around 2,000 B.P., some TG people (pre-Guaranis) were already in the lowlands of South America, occupying fertile fluvial zones in the Paraná-Paraguay River

basin, on lands that are now part of five countries, Bolivia, Brazil, Argentina, Paraguay and Uruguay (Mello 2000, Rodrigues 2007).



Fig. 4: Location of modern TG languages (Urban 1996:66)²

² Abbreviations for Tupi: Stock Languages: A=Arara; Au=Aruá; Ka-Kepkiriwat; Ma=Makurap; Mo=Mondé, P=Puruboná, S=Surui; T=Tupari; Abbreviations for Tupi-Guaraní Family/Languages/dialects: Am=Amanajé; An=Anambé; Ar=Araweté; AsT=Assuriní do Tocantins; AsX=Assuriní do Xingú; Em=Einérillon; Gj=Guajá; Gjj=Gujajára ST=Suruí do Tocantins. Not listed on map: Horá=southernmost dialect of Siriono and adjacent to it. Kokamiya=almost identical to Kokdma and adjacent to it on map.



Guarani languages correspond to Group I in Rodrigues' (1985) classification, according to whom they are variants, dialects of one single language. Rodrigues' grouping is confirmed by recent studies such as Michael et al. (2015) (see figure 5), where the Guaraní branch is a monophyletic group within southern TG.³

The term pluricentric languages (Clyne 1992) may apply to Guaraní languages. There are plenty of linguistic interactions among their speakers and the term Guaraní provides a sense of belonging and identity to all Guaranis⁴ while to its speakers, each Guaraní language represents a separated, distinct language, thus playing an important role in community cohesion and identity. Particularly, Guaraní languages might be classified as migrant PLC languages, type 10 in Muhr's (2016) types of PLC languages. But a proviso is in order: there is no dominant relationship or shared standardized form among Guaraní varieties.⁵ Also, there is no synchronic center of dispersal among them.

³ Nowadays, due to recent migrations, Guaraní languages are currently occupying areas in the South (Rio Grande do Sul, Santa Catarina e Paraná), Southeast (São Paulo e Rio de Janeiro, Espírito Santo), and Midwest (Mato Grosso do Sul) of Brazil.

⁴ Also known as Carijós in some Portuguese sources.

⁵ In the 16th-18th century period, the Jesuit reductions (or missions) used a variety of Guaraní as the *lingua franca* of the south. This variety was considered by non-natives to be the standard variety.

2.2. Founder effects within Tupi

A genetic study conducted by Santos et al. (2015) indicates that among Tupi people, populations located outside MGR present less genetic variation than populations within MGR. Tupian populations were screened for autosomal and uniparental (Y- chromosomes and MtDNA) markers. The following Tupian populations were considered: MGR: *Zoró, Gavião, Suruí, Cinta Larga, Karitiana*; Outside MGR (NonMGR): *Mundurukú, Asuriní do Trocará, Asuriní do Coatinemo, Araweté, Urubú-Ka'apor, Parakanã, Awá-Guajá, Wayampi, Zoé, Aché and Guarani*. Alongside a decay in genetic variability in NonMGR (figure 6), results also suggest recent depopulation in both MGR and NonMGR, which might be a result of the colonization process, but in ancient times MGR experienced a population growth while NonMGR experienced demographic bottlenecks.

Markers	MGR	No-MGR
Classics1	0.193	0.172
STR2	0.683	0.632
Y-STR3	0.253	0.164
Alu insertions4	0.235	0.213
mtDNA sequence5	0,865	0.845

Fig. 6. Variation in genetic markers in Tupi populations (Santos et al. 2015: 4)

Parallel results are presented by Walker et al. (2012), which focused on loss of cultural traits. Based on a linguistic phylogenetic analysis, the authors concluded that NonMGR populations lost some cultural traits present in Tupi, such as uxori-locality (transference of males to live with in-laws after marriage), aggressive warfare, corporate structure (e.g., clans and lineages and moieties), canoe, tattooing, shamanism and lip plugs. This loss of cultural traits was particularly strong in four TG groups, Guajá (northeast of Brazil), Sirinó (Southern - Bolivia), Xetá (Guarani - south of Brazil) and Aché (Guarani - Paraguay).

Rodrigues (2020) investigated linguistic founder effects within Tupi, considering the phonemic inventories of two groups of languages: Southern TG languages (SouthTG) and Tupi within MGR (MGR). Her samples contained the following languages: MGR: *Karitiana (Arikém), Gaivão (Mondé), Suruí Paitér (Mondé), Puruborá (Poruborá), Káro (Ramaráma), Makuráp (Tuparí), Akuntsú (Tuparí), Tuparí (Tuapri), Sakirabiát (Tuparí), Wayoó (Tuparí)*; SouthTG: *Kayowá (Guarani), Nhadeva (Guarani), Xetá (Guarani), Guarani Mbyá (Guarani), Chiriaguano Izoceño (Guarani)*,

Chiriguano Chané (Guarani), Guarayu (Southern), Sirinó (Southern), Paraguayan Guarani (Guarani) and Tapieté (Guarani).

Between-group statistical analyses were conducted, taking into consideration the following factors: number of consonants and number of vowels (oral & nasal) and (b) population size. T-tests and variance tests were not significant for consonants, although a tendency towards a smaller consonant inventory was observed in SouthTG. A significant difference was found for vowel inventory, with MGR languages presenting, among themselves, more variance in vowel inventory size than SouthTG languages (Nasals: $P < 0.00001$, Oral: $P < 0.00001$). No effect on population size was found, which is expected because Tupi languages have a small number of speakers, except for Paraguayan Guarani.

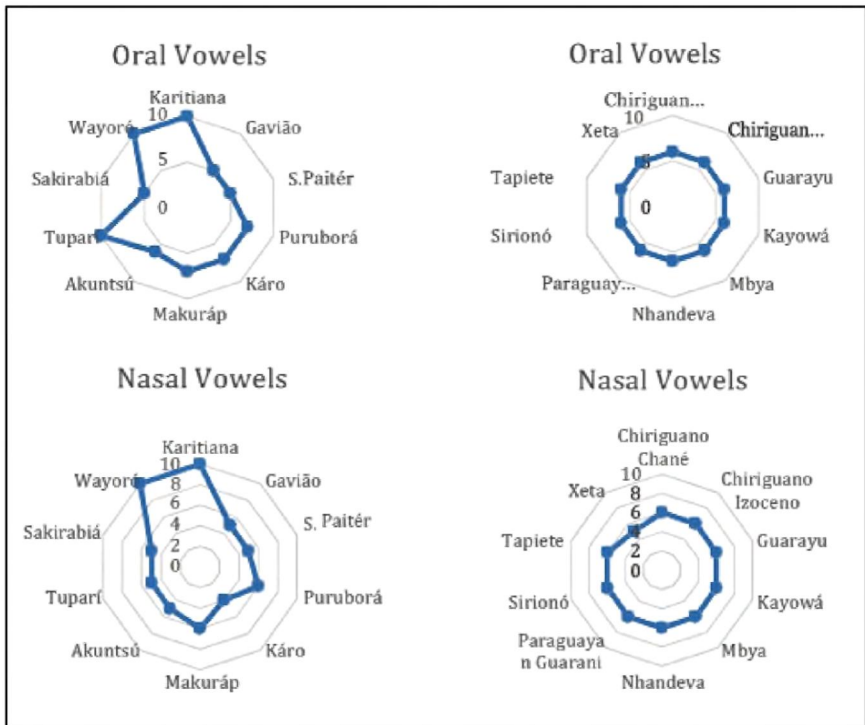


Figure 7. Variation in Vowel inventory size among MGR and SouthTG languages (Rodrigues 2020: 88).

Rodrigues' results indicate no inter-language variation in vocalic phonemes among Guarani languages. All Guarani languages considered in the

study, except for Xetá, present a harmonious 2x6 vowel system (6 oral, 6 nasal). Contrastively, MGR languages exhibit variance among themselves. This therefore suggests a phonemic founder effect in Guarani, with significant reduction in vowel variance.

This conclusion seems to challenge Hay & Bauer's (2007) observation that vowels are more resistant to fluctuations in population size. This contradiction is superficial, however. All reconstruction analyses of Proto-TG suggest that it had the same harmonious 2x6 vowel system observed in Guarani (Lemle 1971, Rodrigues & Dietrich 1997, Schleicher 1998, Meira & Drude 2015). If so, Guarani is highly conservative with respect to vowels, preserving the system inherited from Proto-TG. MGR languages on the other hand, seem to be less conservative. Languages of the Tuparí family for instance, present great variability among themselves (Rodrigues, 2021).

Following Santos et al.'s (2015) conclusion that NonMGR languages suffered a doubled depopulation process before and after Portuguese colonization, in accordance with Hay & Baur, Rodrigues (2020) suggests that the Guarani vowel system was not affected by demographic reduction (but see section 3). In addition, as opposed to Amazonian Tupi, given that most South TG is in constant contact with other languages (e.g., languages of the colonizers (Portuguese and Spanish) and Jê languages), Rodrigues' findings align with Creanza et al.'s (2015) conclusion that languages with too many neighbors present less phonemic variability (we will return to this matter in section 3). It is also compatible with Trudgill (2005) and Hay & Bauer's (2007) observation that small populations have fewer phonemes.

This discussion raises important issues about the role played by vowels and consonants in human language. Why should vowels be more resilient than consonants to external causers of phonemic changes? Investigations on language acquisition point towards a division of labor between vowels and consonants. While consonants are used to build the lexicon, identifying lexical items and highlighting differences among them, vowels signal structural relations and prosodic boundaries. Vowels, not consonants, indicate syllabic structure, boundaries between syntactic constituent and word order, and morphosyntactic process (Nespor et al. 2003, Hochmann et al. 2011, Nespor et al. 2011, among others). This arguably provides a rationale for why languages, with very few exceptions, have more consonants than vowels, the most common being a 4 (20C/5V) ratio (Maddieson & Precoda 2008). In general, vowels are used as structural signalers in Tupi languages. In Guarani Mbyá for example, the

nominalizer morpheme is expressed by -a, as shown by examples in (1), and vowels are used as agreement markers. As shown in fig. 8, contrasts in person agreement are realized by vowels.

(1) Nominalization in Guarani Mbyá (Martins 2003: 123)

ayvu -jopy -a
speech-grab-NMLZer
 'recorder'

Karu-a
 To-eat-NMLZer
 'place to eat'

mba'e-xo-a
 thing-pound-NMLZer
 'plastle'

Person	Active Pronouns	Inactive Pronouns
1PSg	a-	Xe
1PPI (Inclusive)	ja-	nhande
1PPI (exclusive)	oro- (ro)	Ore
2PSg	ere – (re)	nde (ne)
2PPI	pe-	pende (pene)
3PSg/PI	o-	

Fig. 8. Person agreement markers in Guarani Mbyá, adapted from Martins (2003: 36)

The fact that Guarani languages share the same vowel system, thus allow us to form the hypothesis that these languages are grammatically similar, presenting few parametric variations among them. This hypothesis is aligned with Rodrigues' (1985) observation that Guarani languages are dialects of a single language.

3. The effects of language contact on founder languages

An important issue in understanding ancient founder languages is the interaction between ecological factors, population size, language contact and migration. The underlying question is whether human migrations preserve familiar habitats or not. This is an important migratory topic as it converses with questions related to language contact, resulting language changes and linguistic diversity. In this section, we will discuss this issue, focusing on two Guarani languages, Aché and Xetá. In the next section, we will come back to language contact in Paraguayan Guarani.

Based on phylogenetic data from 400 modern Bantu languages, Grollemund et al. (2015) concluded that during their territorial expansion within the African territory (~5.000 B.P.), populations of Bantu speakers explored savannah corridors, avoiding rainforests. Populations that moved into rainforests presented a slow migration rate, delaying the occupation of the new area. As O’Hagan (2014) puts it, “remaining within ecologically similar zones allows for continuity in subsistence practices during migrations”.

Hence, ecological factors might interact with the formation of founder languages in that, migrating groups tend to explore routes that are ecologically similar to their original habitat, and this can shape their languages in different ways. For example, different routes may favor or impede language contact, one of the external driving forces behind language change. It can also result in demographic bottlenecks that might impair language transmission.

The routes taken southwards by pre-Guaranis are rather unclear and it is arguably the case that their territorial expansions occurred in three temporally different migratory waves. First, in Wave 1, Aché went down to Paraguay and Sirinó to Bolivia.⁶ In wave 2, Xetá went to Serra dos Dourados in Paraná/Brazil. Wave 3 consisted of Chiriguano moving to Bolivia, Tapiete and Guarani to Paraguay, and Kaiowa into the region between Argentina, Brazil and Paraguay (Urban 1992, Noelli 2008). Despite this, given the Tupian preference for headwater regions, it is very likely that all pre-Guaranis explored ravines during migration and followed water courses (O’Hagan 2014, O’Hagan et al. 2019).

The location of the center of dispersal of TG languages is also very obscure. Urban (1992, 1996) argues that it is somewhere between the rivers Madeira and Xingu. Recently, using Linguistic Migration Theory (Sapir 1916, Dyen 1956), and based on Michael et al.’s (2015) classification of TG languages, O’Hagan (2014) and O’Hagan et al. (2019) concluded that Proto-TG emerged in the lower part of the Xingu River. First, a branch of Tupi (Proto-Munduruc-Maweti-TG) migrated to northeast Amazon. Then, in the lower Tapajós River, Proto-Maweti-TG split off, migrating latter to lower Xingu, where Proto-TG emerged (Fig. 10).

From lower Xingu, pre-Guaranis might have followed different routes southwards (figure 10) (O’Hagan 2014). They might have followed Tocantins and Araguaia rivers towards Paraná River, which would be the shortest route, but, moving west and then south, they could have explored the watercourses of the Tapajós River or the Madeira River.

⁶ Aché is spoken in Paraguay. It was not included in Rodrigues’ (2021) sample of SouthTG languages.

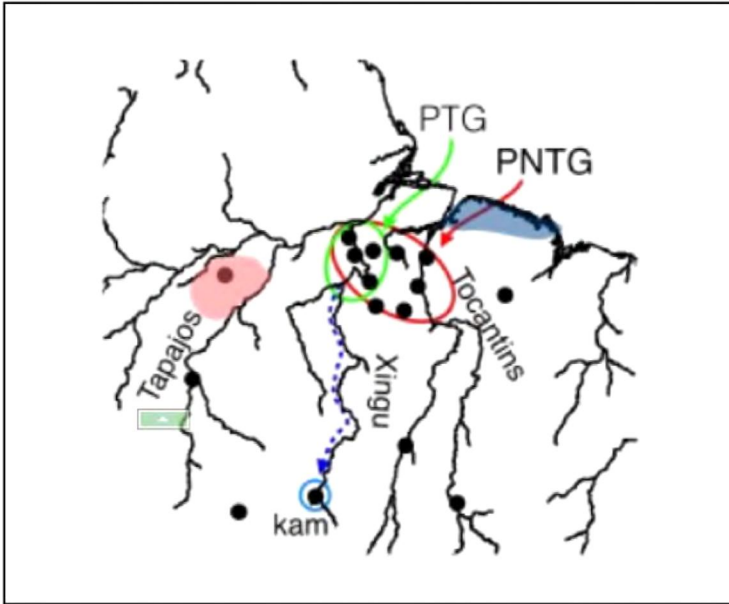


Fig 9. Possible Proto-TG (PTG) birthplace. (O'Hagan et al. 2019: 9)

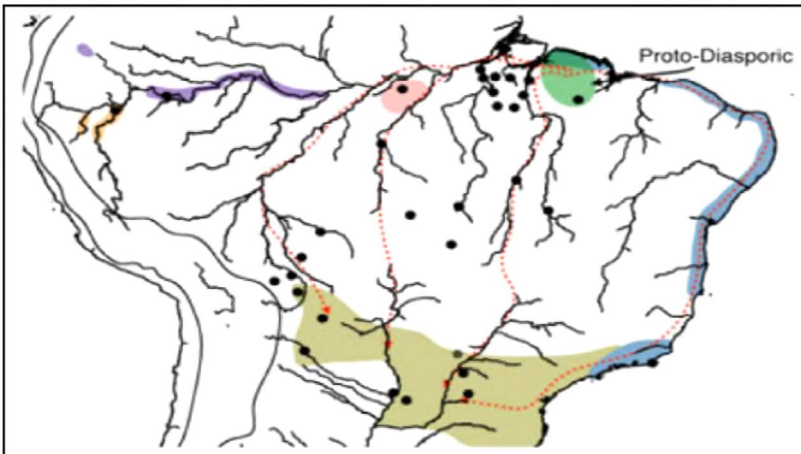


Fig. 10. Possible southward routes taken by pre-Guaranis. (O'Hagan 2014).

Given our current knowledge of this territorial expansion, it is not possible to pinpoint the chosen routes with certainty, although some hypotheses can be raised. One such hypothesis is related to the fact that Proto-Guarani

languages migrated in waves. This is particularly revealing for Xetá and Aché.

Modern Xetá and Aché differ from other Guaraní languages in lexicon, phonology and morphosyntax. According to Rodrigues (1978), taboo nouns and simple nouns denoting mythic factors from Guaraní were substituted in Xetá by nominal phrases. The stress pattern observed in Guaraní, where main stress falls mainly in the last syllable, was altered, being placed in the penultimate syllables. Cabral et al. (2005) observe that this shift in main stress caused weakening and reduction of the last syllables. As a result, modal suffixes disappeared in Xetá, affecting the person agreement system, which displays different forms for indicatives from class II and gerunds in Guaraní, but not in Xetá, where the gerund verbal form is morphologically identical to indicatives from class I. Importantly, Xetá has a reduced vowel inventory. It only has 5 nasal vowels (Rodrigues 2021).

Aché also differs from Guaraní. Röbler (2008) presents a detailed description of Aché, showing that it has around 11 consonants and a 6x2 vowel system, but free variation between [e] and [ɛ] and [o] and [ɔ] (oral and nasal). Aché also permits the use of complex syllabic onsets. In contrast with other Guaraní languages, the main stress has shifted to the penultimate syllable, and long nasal harmony, a feature of Guaraní languages, does not occur. In addition, all prefixes found in Guaraní languages (and in TG in general) disappeared in Aché, including those expressing person and number agreement. Changes are also observed in aspect and temporal markers (Röbler 2008).

Cultural aspects of these groups are also very telling with respect to their migratory history. TG people are known for cultivating their land. In contrast Xetá and Aché seem to have adopted a hunter-gather lifestyle (Hill & Hurtado 1996, Röbler 2008, Rodrigues 1978). Walker et al. (2012) also observed severe loss of other Tupi cultural traits in Xetá and Aché (section 2.2. above). Hence, despite their TG origins, Aché and Xetá exhibit some cultural dissociations from TG.

All in all, our question is: grammatically speaking, why did Xetá and Aché stray away from other Guaraní languages? The facts presented above suggest that the grammatical changes observed are correlated with changes in the vowel systems. This is a possibility that can be formally explored, but there must still be a diachronic event (or a series of events) that caused the observable phonemic changes and consequent grammatical changes.

It has been pointed out that both Aché and Xetá had intense contact with non-Tupi languages spoken in the south, particularly Jê languages, sometimes being mistaken by Jê groups secluded in the forest (Hill & Hurtado 1996, Röbler

2008, Rodrigues 1978). Thus, one hypothesis to be considered (see Röβler 2008), is that language contact is responsible for the changes observed, and it is associated with the fact that Xetá and Aché migrated first (waves I and II).

The migratory routes chosen by Xetá and Aché might have favored language contacts with Jê languages. Also, their route choices might have caused severe demographic bottlenecks forcing them to maintain contact with Jê groups in order to survive. Aché and Xetá are very small languages (see Hill & Hurtado (1996) and Röβler (2008) about Aché, and Rodrigues 2013 for Xetá).

A genetic study conducted on the Aché gene pool indicates descentance from Tupi with considerable transference of genetic material from Jê (Callegari-Jacques et al. 2008).

Although the observations above are mostly based on hypothetical scenarios, they are aligned with what we know about the history of TG groups and their languages.

4. The effects of language attitudes and practices on founder languages

Another important theme on the study of founder languages is related to their social and political use. Take Paraguayan Guarani (PG) as an example. Language contact added to the social and political status of PG set it apart from other Guarani varieties.

In distinct contrast with other Guarani languages, PG has around 5,000,000 speakers. It is the largest native language of South America, with emphasis on the fact that it is widely spoken by non-natives. 46.3% of Paraguayan homes use both Guarani and Spanish and 34%, only Guarani, with bilingualism concentrated in urban centers. Monolinguals (Guarani only) represent most of the rural population (Estigarribia 2020). Hence, PG is the only Guarani language that is not endangered.

However, there is no uniform, single grammatical system that can be called PG. The term refers to a set of linguistic varieties with different degrees of mixture with Spanish. Speakers use one of these varieties depending on their knowledge of, and fluency of Guarani. Jopara is the general term locals use to refer to mixtures between Guarani and Spanish, but according to Estigarribia, locals differentiate these varieties as follows: *Guaraniete*: “true” guarani, *Guarañol*: Jopara mostly based on Guarani and *Castení*: Jopara mostly based on Spanish. As examples of these varieties, consider the data below, extracted from

Estigarribia (2020:17).⁷

- (1) *Guaraniete*
 Chepytyvõmína
 che-pytyvõ-mi-na
1sg.inact-help-plead-req
 ‘Help me, please.’
Guarañol
 Cheayudamína
 che-ayuda-mi-na
1sg.inact-help-plead-req
 ‘Help me, please.’
Castení
 Ayudámena
 ayuda=me-na
help=me-req
 ‘Help me, please.’

The lexicon of these varieties contains words from Spanish and Guaraní, and the amount of Spanish-based lexical influence depends on the speaker’s competence in Guaraní.

With regard to phonemes, PG has the largest inventory among the Guaraní languages. While there is a tendency in Guaraní, as compared to MGR languages, towards a reduced consonant inventory, PG exhibits the opposite pattern. It has the Guaraní vowel system (6x2) and 18 consonants. The mean size of the consonant inventory within Guaraní is 14.11. Hence, the size of the consonant inventory of PG is well above the mean size of Guaraní.

Reconstructions of Proto-TG have posited different numbers of consonants: 12 in Lemle (1971), 13 in Schleicher (1998) and 19 in Rodrigues and Dietrich (1997). Therefore, either PG gained consonants, or maintained the consonants it inherited from TG, whereas the other Guaraní languages reduced theirs. Based on Perreault & Mathew’s (2012) conclusion that phonemes accumulate at a rate between 0.26 and 0.38 per 1,000-year period, and on Creanza et al.’s (2015) observation that languages in contact tend to be more conservative with respect to phonemic inventory size, we may conclude that PG is the most conservative Guaraní language with respect to phonemes, preserving

⁷ See Estigarribia (2015), where it is argued that Jopara is a code-switching system.

the vowel and consonant inventories inherited from TG (see Rodrigues 2020). In accordance with this conclusion, Estigarribia (2020) shows that PG is phonologically faithful to Guaraní, blocking complex onsets within syllables and preferring stress on the final syllable. Also, as to nasal harmony, PG follows the Guaraní pattern (Estigarribia 2017, 2020). As for morphosyntax, although detailed comparative studies on Guaraní languages are to be done, there seems to be no major morphosyntactic differences between PG and other Guaraní Languages, with the exception of Xetá and Aché (section 3).

Given the intense contact between PG and Spanish, comparisons between PG and Aché and Xetá, lead to the conclusion that under language contact, a founder language can either preserve its grammar or have it altered. Thus, the question is why language contact with Spanish did not alter the grammar of PG, which I take to be well represented by the *Guaraniete* variety.

There are two relevant facts here. First, it is probably the case that Xetá and Aché suffered severe depopulations during migration to and settlement in the south, as discussed in section 3. Although the colonization of Paraguay caused an important dissemination of natives, the so-called Paraguayan Guaranis did not experience a depopulation as severe as Aché and Xetá did.⁸ Thus, while intergenerational language transmission might have been impaired in Xetá and Aché, it was not in PG. Second, more recently, positive language policies established by the Paraguayan government and positive national attitudes towards Guaraní, were important factors in preserving the language of the Guaraní founders in Paraguay.

In 1992, PG was declared an official language of Paraguay and education on Guaraní became mandatory. In 2019, it was also listed as an official language of MERCOSUL. This formal status is not observed in other Guaraní varieties, even though many speech communities of Guaraní present a strong ethno-linguistic and cultural awareness. Hence, although PG doesn't have the same prestige that Spanish has in Paraguay, it is taken to be part of the Paraguayan ethno-identity. The existence of a language mixture, Jopará, is an indication of a positive local attitude and practice towards PG.

Bilingualism is very common among speakers of other varieties of Guaraní. For instance, most Brazilian Guaranis are fluent speakers of Brazilian Portuguese as well, but no systematic language mixture with Brazilian Portuguese is

8 According to Röbler (2015), nowadays, Aché has only 250 fluent speakers. Xetá has only 3 speakers (Rodrigues 2013). On the history of the Xetá group of Serra dos Dourados, see the documentary *Xetá*, available at https://www.youtube.com/watch?v=_aUyib-tAlo (last accessed on 02/10/2022).

observed. This is largely due to the fact that the vast majority of non-native Brazilians have no desire to learn a native language.

5. Conclusion

In this paper, we tried to draw together some observations about migrant pluricentric languages, taking Guarani, a group of Tupi-Guarani languages spoken in the lowlands of South America, as an example. These languages exhibit a linguistic founder effect, showing less variation in vowel inventory than languages spoken in MGR, the birthplace of Proto-Tupi. However, given that Proto-Tupi-Guarani had the same vowel system that modern Guarani has, most Guarani languages are rather grammatically conservative. This relates to the role played by vowels in human grammar. Vowels are signalers of grammatical structure and prosodic boundaries. All Guarani languages that preserved the Tupi-Guarani vowel inventory are similar in phonology and morphosyntax. Xetá and Aché differ from Guarani, exhibiting changes in their vowel inventories and phonological and morphosyntactic alterations. We hypothesized that these languages result from language contact (arguably with Jê languages), and severe depopulation events during migration to, and settlement in the south. In contrast to this hypothesis, Paraguayan Guarani did not have its phonemic inventory and grammar altered, despite intense contact with Spanish. This might reflect the following facts: Paraguayan Guaranis did not experience a severe depopulation, positive language policies were established by the Paraguayan government, and among Paraguayans (natives and non-natives), there is a positive national attitude towards Guarani.

To conclude, investigations on migrant languages are an important field of research on language diversity, but in itself, it is an interdisciplinary field, involving knowledge of theoretical linguistics, language as a cognitive system, language typology, language contact, historical linguistics and knowledge about educational and regulatory language policies.

References

- Atkinson, Q. (2011): Phonemic Diversity supports a founder effect model of language expansion from Africa. IN: *Science*, 332. P. 346-349
- Brochado, J. (1984): An ecological model of the spread of pottery and agriculture into eastern South America. Ph.D. dissertation.
- Cabral, A. S. / Rodrigues, A. D. / Vasconcelos, E. A. (2005): Sobre o sistema pessoal da língua Xetá. In: L. Lobato et al. (eds.), *Anais IV Congresso internacional da ABRALIN*. Universidade de Brasília. P. 57-64.

- Callegari-Jacques, S. / Him, K. / Hurtado, A. / Rodrigues, L. (2008): Genetic clues about the origins of the Aché hunters-gatherers of Paraguay. In: *American Journal of Human Biology*, 20/6. P.735-737.
- Campbell, L./ Grondona, V. (eds.) (2012): *The indigenous languages of South America*. Berlin/New York. Mouton/de Gruyter.
- Cavalli-Sforza, L. L. (2001): *Genes, people and language*. University of California Press.
- Champaign, IL & Urbana, IL: University of Illinois.
- Clyne, Micheal (ed.) (1992): *Pluricentric Languages: Differing Norms in Different Nations*. Berlin/New York. Mouton/de Gruyter.
- Creanza, N./ Ruhlen, M. / Pemberton, J. T. / Rosenberg, N. A. / Feldman, M.W. / Ramachandran, S. (2015): A comparison of worldwide phonemic and genetic variation in human populations. *PNAS*, 112. P. 1265-1272.
- Dryer, M. / Haspelmath, M. (eds.) (2005): *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info>, Accessed on 2019-10-26.)
- Dyen, I. (1956): Language distribution and migration theory. In: *Language* 32/4. P. 611-626.
- Estigarribia, B. (2015): Guaraní-Spanish Jopara in a Paraguayan novel. In: *Journal of Language Contact*, 8. P. 183-222.
- Estigarribia, B. (2017): A Grammar sketch of Paraguayan Guarani. In: B. Estigarribia / J. Pinta (eds.), *Guarani Linguistics in the 21st Century..* Leiden: Brill.
- Estigarribia, B. (2020): *A grammar of Paraguayan Guarani*. London: UCL Press.
- Grollemund, R. S. / Brandfor, S. / Bostonen, K. / Meade, A. Venditti, C. / Pagel, M. (2015): Bantu expansion shows that habitat alters the route of human dispersals. In: *Proceedings of the National Academy of Science*, 112. P. 13296-13301.
- Hay, J. / Bauer, L. (2007): Phoneme inventory size and population size. In: *Language*, 83. P. 388- 400.
- Hill, K. / Hurtado, M. (1996): *Aché life history: The ecology and demography of a foraging people*. New York, Routledge.
- Hochmann, J. R. / Benavides-Varela, S. / Nespors, M. / Mehler, J. (2011): Consonants and vowels: different roles in early language acquisition. In: *Developmental Science*, 14. P. 1467-1687.
- Lemle, L. (1971): Internal classification of the Tupi-Guarani linguistic family. In.: Bendor- Samuel D. (ed.) *Tupi Studies*. Norman: Oklahoma: Summer Institute of Linguistics, 107-129.

- Lewis, M. P. (2009): *Ethnologue: Languages of the World*. Sixteenth edition. Dallas, Texas: SIL International. <http://www.ethnologue.com/16>.
- Maddieson, I./ Precoda, K. (1990): *The UCLA Phonological Segment Inventory Database*, 1990. http://web.phonetik.uni-frankfurt.de/upsid_info.html.
- Marques, A. M. / Renesto, E. (2017): Genetic variability and evidence of founder effect in *Hemiodus orthonops* (Characiformes: Hemiodontidae) from the upper Paraná River Basin, Brazil. IN: *Acta Scientiarum, Biological Sciences*, 39. P. 53-58.
- Martins, M. F. (2003): *Descrição e análise de aspectos da gramática do Guarani Mbyá*. PhD. Dissertation, Universidade Estadual de Campinas.
- Mayr, E. (1954): Change of genetic environment and evolution. In.: Huxley J. et al. (eds.), *Evolution as a Process*. London: Allen & Unwin, 157-180.
- Mayr, E. (1963): *Animal species and evolution*. Cambridge, Massachusetts: Harvard University Press.
- Meira, S. / Drude, S. (2015): A summary reconstruction of Proto-mawetí-guarani segmental phonology. In: *Boletim do Museu Paranaense Emílio Goeldi. Ciências Humanas*, 10. P. 275- 69.
- Mello, A. (2000): *Estudo histórico da família linguística tupí-guaraní: Aspectos fonológicos e lexicais*. Ph.D. dissertation, Universidade Federal de Santa Catarina.
- Michael, L., N. / Chousou-Polydouri, B. / Donnelly, E. / Wauters, V. / Meira, S./ O'Hagan, Z. (2015): A Bayesian phylogenetic classification of Tupí-Guaraní. In: *LIAMES*, 15. P. 193-221.
- Muhr, R. (2016): *Pluricentric languages and non-dominant varieties worldwide*. Berlin: Peter Lang. In: Rudolf Muhr et al. (eds.). *Exploring Linguistic Standards in Non-Dominant Varieties of Pluricentric Languages*. Frankfurt am Main, etc. Peter Lang Verlag. 343-363.
- Nespor, M. / Peña. M. / Mehler. J. (2003): On the different roles of vowels and consonants in speech processing and language acquisition. In: *Lingue e Linguaggio*, 2. P. 2003-2229
- Noelli, F. S. (1998): The Tupi: Explaining origin and expansion in terms of archaeology and of historical linguistics. In: *Antiquity*, 72. P. 648-663.
- Noelli, F. S. (2008): The Tupi expansion. In H. Silverman and H. William (eds.) *Handbook of South American archaeology*. New York: Springer. 659-670.
- O'Hagan, Z. (2014): *A Computational-phylogenetic Classification of Tup'í-Guaraní and its Geographical Spread*. Paper presented at Language Variation and Change, University of Chicago. (Slides available at http://linguistics.berkeley.edu/~zohagan/pdflinks/ohagan_lvc_tg_classifica)

- tion_spread_2014.pdf. (Last accessed: 08/02/2022).
- O'Hagan, Z. (2015). A Bayesian phylogenetic classification of Tupí-Guaraní. In: LIAMES, 15. P. 193-221.
- O'Hagan, Z. / Chousou-Polydouri, N. Micheal, L. (2019): Phylogenetic classification a northeastern Amazonian Proto Tupí-Guaraní homeland. In: LIAMES, 19. P. 1-29.
- Perreault, C. / Mathew, S. (2012): Dating the Origin of Language Using Phonemic Diversity. In: PLoS One, 7. e352-89.
- Rößler, E. M. (2008): Aspectos da gramática achê: Descrição e reflexão sobre uma hipótese de contato. MA Thesis, Universidade Estadual de Campinas.
- Rodrigues, A. D. (1964): Classificação do tronco linguístico Tupi. In: Revista de Antropologia, 12: 99-104.
- Rodrigues, A. D. (1978): A língua dos índios Xetá como dialeto Guaraní. In: Cadernos de Estudos Linguísticos, 1. P. 7-11.
- Rodrigues, A. D. (1985): Relações internas na família linguística Tupí-Guaraní. In: Revista de Antropologia, 27. P. 33-53.
- Rodrigues, A. D. (1986): Línguas brasileiras: para o conhecimento das línguas indígenas. São Paulo: Loyola.
- Rodrigues, A. D. (2005): As vogais orais do Proto Tupí. In.: Rodrigues, A. D. / Cabral, A. S. A.C. (eds.) Novos estudos sobre línguas indígenas. Brasília: Editora da UnB, 35-46.
- Rodrigues, A. D. (2007): Tupí languages in Rondônia and in eastern Bolivia. In: Wetzels, Leo (ed.): Language endangerment and endangered languages. Leiden: CNWS Publications, 355-363.
- Rodrigues, A. D. / Cabral, A. (2002): Revendo a classificação interna da família linguística Tupí- Guaraní. In.: A. Cabral / A. D. Rodrigues (eds.), Línguas indígenas brasileiras: Fonologia, gramática e história. Belém: UFPA, 327-337.
- Rodrigues, A. D. / Cabral, A. (2012): Tupian. The indigenous languages of South America: A comprehensive guide. Berlin/New York. Mouton/de Gruyter, 495-574.
- Rodrigues, A. D. / Dietrich, W. (1997): On the linguistic relationship between Mawé and Tupi- Guaraní. In: Diachronica, 14. P. 265-304.
- Rodrigues, C. (2020): Founder effect in Tupian languages. In: Revista Diadorim, 22. P. 65-97.
- Santos, E. / Silva, A. / Ewerton, P. / Takeshita, L. / Maia, M. (2015): Origins and demographic dynamics of Tupí expansion: a genetic tale. IN: Boletim do Museu Paranaense Emílio Goeldi. Ciências Humanas, 10. P. 217-28.

- Sapir, E. (1916): Time perspective in aboriginal American culture, a study in method. Ottawa: Government Printing Office.
- Schleicher, C. (1998): Comparative and internal reconstruction of the Tupi-Guarani language family. PhD Dissertation. University of Wisconsin, Madison.
- Trudgill, P. 2011. Social Structure and Phoneme Inventories. *Linguistic Typology*, v. 15, p. 155-160
- Ramachandran, S. / Deshpande, O. / Roseman, C. / Rosenberg, N. / Feldman, M./ Cavalli-Sforza, L. (2015): Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. In: *Proceedings of the National Academy of Sciences of the United States of America*, 102. P. 15 947.
- Urban, G. (1992): A história da cultura brasileira segundo as línguas nativas. In Cunha, M. C. (ed.) *História dos Índios no Brasil*. São Paulo, FAPESP/SMC/Cia das Letras. 87-102.
- Urban, G. (1996): On the geographical origins and dispersions of Tupian languages. In: *Revista de Antropologia*, 39. P. 61-104.
- Walker, R. S. S. /Wichmann, T./ Mailund, C. J./ Atkisson, Q. (2012): Cultural phylogenetics of the Tupi language family in lowland South America. In: *PLoS ONE* 7/4. e35,025.

Liliana SÁNCHEZ

University of Illinois, Chicago
lesanche@uic.edu

Is Quechua II a *pluricentric* language family?

Abstract

This article explores the extent to which a multinational minoritised family of indigenous languages such as Quechua II can be considered pluricentric. A brief introduction to the complexity of the language family is provided along with an overview of current legislation and policies in the three Andean countries with the largest Quechua-speaking populations. Despite the vitality of some of the language varieties and the use of normalised and standardised forms by the three nation-states, it is difficult to determine whether the standardised varieties or the prestigious ones are the centres of diffusion of the languages, especially in academic settings outside the countries of origins, making it very difficult to determine the extent to which the Quechua II family of languages is pluricentric.

1. Introduction

Pluricentric languages have been defined as languages that comprise two or more “standard” varieties of the language (Ammon 1989) or, at least, two or more centres of diffusion of a prestigious variety of what is conceptualised as the same language (Muhr and Marley 2015).

In this article, I explore the extent to which this concept can be applied to the complex reality of a large family of minoritised languages such as the Quechua languages. Unlike socially and economically dominant languages, minoritised indigenous languages spoken in different countries may indeed have different centers of diffusion. However, given their status as non-dominant languages in those countries and given the fact that they are in most cases threatened by language shift and the loss of intergenerational transmission, it seems appropriate to question the extent to which they can be considered pluricentric.

To determine whether Quechua can be typified as a pluricentric language, it is necessary first to present an overview of the complexity of the Quechua family of languages and the sub-families or varieties of the language

that have been posited since the past century. Parker (1963) and Torero (1964) proposed the currently widely accepted classification of Quechua languages. In this classification Quechua languages are subdivided into two main subfamilies: Quechua I/A and Quechua II/B. All varieties of Quechua I/A are spoken within the boundaries of Peru. A sub-group of varieties of Quechua II/B languages corresponding to the Southern sub-family (QII B-C) (Cerrón Palomino 1988) are spoken across different countries in South America including Argentina, Bolivia, Ecuador, Colombia, and Peru (Cerrón Palomino 1988).¹

This article focuses on Kichwa/Quechua varieties from the Quechua II B sub-family spoken in Ecuador and the Quechua II C varieties spoken in Bolivia and Peru,²³ and the processes of standardisation of the local and regional varieties within each country. Examining these processes is crucial to understand the extent to which the Kichwa/Quechua continuum can be considered a pluricentric family of languages.

Before presenting the current situation regarding standardisation, it is necessary to provide a brief overview of expansion of the Quechua languages. Historically, the Quechua I and II subfamilies experienced complex processes of diffusion in pre-Inka times in coexistence and in competition with other indigenous languages (Sillar 2012). Quechua II varieties had greater levels of geographical expansion during the Inka period as well as in colonial times (Durston 2014). In fact, the Cuzco Quechua variety (Quechua II) was considered one of the main indigenous lingua franca or *Lenguas Generales* in pre-Hispanic and colonial times (Mannheim 1991/2011).

In colonial times, some varieties were privileged as languages of evangelisation (grammars, Dictionaries and Evangelising materials) during the 16th and 17th centuries. In the 17th century the colonial Bourbon reforms that were a reaction to indigenous rebellions resulted in the prohibition of indigenous cultural manifestations and a less tolerant attitude towards indigenous languages (Mannheim 1991/2011).

The 19th century saw the new republics continue the path of general expropriation of indigenous lands in different countries and continuous marginalisation of indigenous peoples, their cultures, and their languages. The

¹ Migration from Bolivia to Chile and Argentina has more recently resulted in communities of Quechua speakers in those countries.

² The name of the varieties spoken in Ecuador is Kichwa, while most varieties in Peru are referred to as Quechua.

³ References to other varieties will be made, but as it will become apparent later in this article, these are the varieties that have had more recent processes of expansion due to international migration.

20th century saw a combination of accelerated language shift towards Spanish (Escobar 2007) and, at the same time, some legislation in favour of indigenous language rights.

As a result of the many layers of diffusion and contact with other languages, different varieties of Kichwa/Quechua languages have received different names. In Bolivia and Peru, most varieties are known as Quechua, but some varieties, such as the Cuzco variety, are also known as Runasimi (Cusihuamán 1976/2001). There are also spelling differences such that some varieties like the ones spoken in the Amazonian region of Peru are spelled Kichwa (Ministerio de Cultura, Peru 2021).

In Ecuador, most varieties are known as Quichua or Kichwa (Limerick 2018), and the term Runashimi is also used to refer to the language (Martínez 2019). These multiple names for the languages reflect the reality of the complex linguistic continuum represented in the Quechua II sub-family.

To better understand the differences found across the three Andean countries with the majority of speakers of Quechua, we need to take into account the percentage of Quechua/Quichua speakers as part of the country's population. The following table from Andrade and Howard (2011) shows the distribution.

Country	Previous Census		Recent Census	
	Total	%	Total	%
Ecuador	499,292	4,10	591,448	4,08
Peru	3'360,331	13,03	3'799,780	13,06
Bolivia	1'542,143	20,81	1'656,978	17,15

Table 1: Population of Quechua/Quichua speakers in Bolivia, Ecuador and Peru (based on Andrade and Howard 2021: 14)

As noted by Andrade and Howard (2021), while the percentages remain apparently stable in Ecuador and Peru, the percentage of Quechua speakers in Bolivia has diminished in the span between the two censuses. The low percentages and the decrease in the percentage of speakers in Bolivia are consistent with the fact that Quechua languages have been historically minoritised in the different countries where they are spoken. This affects to different degrees the levels of social vitality of the languages in the three, making it very challenging to conceptualise them as viable candidates for a pluricentric language group.

Despite the historical marginalisation of the Quechua languages in the three Andean countries, important changes regarding indigenous languages revitalisation and access to intercultural bilingual education have taken place in the legislations of Bolivia, Ecuador, and Peru. These changes have resulted in the continuous growth and evolution of language policies that promote these languages with different levels of success. As we will see in the next section, some of these policies date back to the second half of the 20th century, but they have become more prevalent in the early 21st century.

The organisation of this article is as follows. In section 2, I present current language policies. In section 3, processes of standardisation and issues related to the vitality of Quechua languages in three Andean countries (Bolivia, Ecuador, and Peru) are discussed. In section 4, the expansion of Quechua languages outside South America through immigration and in academic settings is analysed. It also provides a discussion of how the prevalence of some Quechua languages internally to the countries in which they are spoken, and outside them through immigration and diffusion, can illustrate pluricentricism. At the same time, the diffusion of these varieties comes at the cost of the further marginalisation of other varieties within the continuum of Quechua languages.

2. Current language policies, standardisation and vitality of Quechua and Kichwa/Quichua in Bolivia, Ecuador, and Peru

In this section, I present some of the current legislation in Bolivia, Ecuador, and Peru that have significantly changed the legal status of indigenous languages in those countries. As noted by Haboud and Limerick (2017), in the three countries legislation in the area of education has advanced despite difficulties and challenges to the implementation of intercultural bilingual education programs that contribute to language revitalisation.

To understand whether it is possible to think of the Quechua II B and C languages as pluricentric in the sense of having two or more standardised varieties, it is necessary to provide an overview of how legislation on education in indigenous languages has contributed and could contribute to the establishment of “standardised” varieties of the languages in each country.

2.1 Legislation and education in Quechua in Bolivia

Bolivia has a long history of Quechua normalisation. A normalised alphabet was established in 1984 (Albó 1987). This alphabet has gradually

become the norm in intercultural bilingual education programs. One of the most important factors in Quechua normalisation in Bolivia has been the implementation of intercultural bilingual education programs. The relevance of the inclusion of Quechua in the educational system has been the driving force in establishing language rights for Quechua-speaking as well as other indigenous communities. In fact, an important moment in the process of change in the legal and social status of indigenous languages in Bolivia can be traced back to the 1994 Educational reform that resulted in the production of numerous school texts and training for Quechua teachers (Luykx 2011).

Bolivia's 2009 Constitution recognises 36 Indigenous languages, as well as Spanish, as official languages (Art. 5, I). It institutionalises Plurilingual Intercultural Intracultural Education (EIIP, formerly EIB) (Art. 30, II.12) that is guided by the *Suma Qamaña* principle, which in Quechua means "to live correctly and well" (Haboud and Limerick 2017).

The official status of Quechua in Bolivia is also recognised in the 2010 Education Act (Intercultural Education in a plurinational state) and in the 2012 Ley General de Derechos y Políticas Lingüísticas (Law 269) (Haboud and Limerick 2017). Despite this official level of recognition, like in other countries of the region where Quechua is spoken, language shift to Spanish is rapidly extending in urban and even in rural environments.

While there is a normalised alphabet, Bolivia as a country has not established a single unified standardised variety of the language, and dialectal variation is widely acknowledged and recognised. The main Quechua varieties in Bolivia are the northern La Paz variety, and the southern varieties of Chuquisaca, Oruro and Potosí, the latter two in close contact with Aymara (Plaza 2009).

One very salient characteristic of the Bolivian process of recognition of indigenous languages has been its contribution to the promotion of Intercultural Bilingual Education in Indigenous languages in the region. In fact, Bolivia has been at the centre of many efforts to promote Indigenous languages across the Central Andean countries and in Latin America.

The leading institution in the region that has contributed to the education of intercultural bilingual educators and researchers is PROEIB Andes (Programa de Educación Intercultural Bilingüe para los países Andinos). PROEIB Andes has been part of the Faculty of Humanities at the Universidad Nacional de San Simón in Cochabamba, Bolivia, since 1996. It was created in collaboration with the ministries of Culture and several universities from

Bolivia, Argentina, Colombia, Chile, Ecuador, and Peru. From 2000 to 2011, PROEIB Andes was a member of the Intercultural Indigenous University Project, as part of an agreement with the Development Fund for the Indigenous Peoples of Latin America and the Caribbean (PROEIB ANDES 2021).

Among its graduates, PROEIB Andes' MA program boasts an important number of educators and researchers who occupy important positions in government and NGO institutions in Bolivia, Ecuador, and Peru. Graduates from this program have been crucial in promoting Intercultural Bilingual Education in Indigenous languages and Spanish, and their involvement in legislative and educational efforts has resulted in some positive changes, such as the elaboration of educational materials and assessment measures in Quechua in Peru and Bolivia. In that respect, PROEIB Andes has worked as a focal point from which several generations of language planning agents, who are respectful of local and macroregional varieties, have emerged.

2.2 Legislation and education in Quichua/Kichwa in Ecuador

In Ecuador, legislation also recognises Kichwa as an official language. As noted by Haboud and Limerick (2017), Ecuador's 2008 Constitution, in its article 2, recognises Spanish, Kichwa, and Shuar (an Amazonian indigenous language) as official languages of intercultural relations.

While Kichwa is mentioned as a language, there are important differences between Highland and Lowland (Amazonian) varieties of Kichwa (Haboud 2010), and at least eight varieties have been identified: *the Northern, Central and Southern Andean varieties and the Northern and Southern Amazonian varieties* (Aschmann 2007).

As in the case of Bolivia, progress in language rights has taken place mainly in education legislation. The Organic Law of Intercultural Education from 2011, in its Chapter 4, establishes a national level of central administration and three levels of intercultural bilingual education at the regional, district, and circuit or community levels. The latter to be implemented in relevant areas (Organic Law of Intercultural Education, Art. 26-30). As noted by Haboud and Limerick (2017), this establishes a separation between intercultural bilingual education and non-intercultural bilingual education.

Within the intercultural bilingual education system, major efforts have been made to implement the use of a standardised unified variety of Ecuadorian Kichwa, mostly based on Highland varieties (King 2001). These

efforts have encountered different levels of difficulties, ranging from perceptions by members of Kichwa-speaking communities of a lack of cultural appropriateness (Limerick 2020) to the view of standardisation as a threat to indigenous identity by some Kichwa-speaking communities in the Amazonian region of Ecuador (Grzech et al. 2019, Wroblewski 2020).

Even the use of standardised Kichwa greetings by state agents within the intercultural educational system have generated some controversy, as these efforts have been perceived as not culturally appropriate by parents, teachers, and other members of indigenous communities (Limerick 2020). Thus, while among the three countries Ecuador is the only one with a unified standardised variety of Kichwa, its status as a well-accepted variety across the country has not solidified.

2.3 Legislation and education in Quechua in Peru

Peru, like Bolivia, has a long history of officialisation of Quechua. In 1975, a Law Decree establishing Quechua as an official language in the country was approved. In 1979, the officialisation of Indigenous languages in the Political Constitution took place. The Political Constitution stipulated that Quechua and Aymara were official languages, in the forms and in the areas prescribed by the law. The 1993 Constitution established the officialisation of indigenous languages in areas where they predominate. Since 2011, according to the Law of Languages (29735), native languages are official in the districts, provinces or regions where they predominate, as recorded in the National Register of Native Languages.

Article 10 of the Law of Languages of 2011 states that the fact that a native language is official, in a district, province or region, means that the state administration endorses it and progressively implements it in all its spheres of public action, giving it the same legal value and the same prerogatives as Spanish. The official documents issued by the state must be in both Spanish and the official original language when it has rules of writing. It also states that the documents in the official original language have the same legal value as the Spanish one. This particular piece of legislation extends the use of indigenous languages beyond the sphere of intercultural education and is a step towards the actual implementation of a multilingual state.

Like Bolivia and unlike Ecuador, Peru does not have a single unified standardised variety of Quechua, given that, of the three countries, it is the one with more dialectal variation as it has Quechua I and Quechua II varieties. The

Ministry of Culture’s Indigenous data base recognises six Quechua I zones and three macro-variants spoken in Peru that correspond to the Quechua II sub-family. Northern Quechua corresponds to Quechua IIA⁴, Amazonian Quechua corresponds to Quechua IIB⁵, and Southern Quechua corresponds to Quechua IIC⁶.

Educational materials have been created for Southern varieties and to some extent for Northern and Amazonian varieties, as well as for Spanish and other indigenous languages throughout the last decades of the 20th Century and in the 21st Century (Ministerio de Educación 2018). It must be noted that there is one highly prestigious variety that occupies a prominent place in Peru due to its association with the historical Inka period, the Cusco-Collao variety. We will see in more detail in section 3.6 some of the implications of this strong association.

3. The construction of national/regional ‘standardised forms’ vs. the preservation of local varieties

As a result of the new legislation and policies previously discussed, “standardised” forms of local and regional Quechua languages in Bolivia and Peru have become part of state mandated policies, for instance regarding the creation of educational materials in different varieties of Quechua for which some level of normalisation has taken place. These varieties are expected to be used in intercultural bilingual education programs and to a lesser extent in other state services for indigenous populations.

One could think of the normalisation efforts in Bolivia and Peru and the existence of unified Kichwa in Ecuador as evidence in favour of adopting a pluricentric perspective regarding the Quechua languages. While indeed there is an emergence of different “standardised” varieties for educational and other state purposes, as is the case in Peru, the extent to which they are adopted by the indigenous populations seems to be a problematic issue for considering the Quechua II B and C varieties spoken in the three countries as representative of a case of pluricentric languages.

Furthermore, the pluricentric nature of a language or a family of language is based on the social status of their speech communities. The fact

⁴ Northern corresponds in the Indigenous languages database to the Cañaris and Cajamarca varieties.

⁵ Amazonian Quechua in the database includes Pastaza, Napo, Putumayo, Tigre, Alto Napo, Chachapoyas, and San Martín.

⁶ Southern Quechua in the database includes Chanka and Collao, also known as Cusco-Collao.

that indigenous communities are socially marginalised limits the possibility of a strong form of pluricentrism. However, it must be noted that despite linguistic variation internal to each of the countries and differences in legislation, there have been international efforts aimed at promoting Quechua. In 2004, a General Treaty on Integration and Economic Cooperation and Social for the Establishment of a Common Market between the Parties was adopted between Bolivia and Peru. One of its goals was to promote education in rural areas, technical and technological education, and bilingual intercultural education.

3.1 Vitality

Another very important aspect of the discussion of the Quechua II sub-family as a pluricentric one is the issue of language vitality. Despite the fact that the Quechua family of languages is among the most widely spoken indigenous languages of the Americas, not all the varieties have similar levels of vitality.

While most Quechua I languages are endangered or severely endangered or undergo major structural changes when spoken (see Sánchez 2003 for Junín Quechua), the vitality of Quechua II varieties differs within national boundaries (Ministerio de Educación de Perú 2018). For instance, in Peru and Bolivia, the southern varieties such as Cusco-Collao Quechua are considered vital (Eberhard et al. 2019), but even within that context, dialectal variation is great, and some local varieties may be experiencing structural changes and language shift towards Spanish (Kalt and Geary 2021).

In the same vein, some Ecuadorian Quichua varieties are considered vital although language shift is rapidly occurring, despite the symbolic power of these varieties (Haboud 2004). Furthermore, it is very difficult to determine the extent to which the Quechua family of languages enjoys social vitality defined as clear evidence of intergenerational transmission or as the lack of autonomy of the speech communities themselves to determine the future of the language.

3.2 Authenticity, purism, and unification through standardisation

A discussion of pluricentricism regarding Kichwa/Quechua languages cannot avoid issues such as authenticity, purism, and unification. As bilingualism in Kichwa/ Quechua languages and Spanish becomes more extended, the issue of who is an authentic speaker has become more prevalent. For a great number of speakers, an authentic speaker of Kichwa/ Quechua is

one who does not code-switch and generates new Quechua terms. There is a strong association with members of rural communities who view themselves as members of indigenous communities (Povilonis 2021), but there are also urban guardians of language purity (Coronel-Molina 2008).

Authenticity and purism are also linked to the level of prestige that a particular variety holds within a speech community and beyond. In Peru, Cusco Quechua is perceived by many as the language of the glorious Inka past narrative. The defenders of the special status of this variety have been at odds for many years with the state's efforts to adopt a unified alphabet for all Quechua varieties in the country.

This opposition gave rise to the “vowel” debate in the Cusco region in Peru (Hornberger and King 2001). The debate was based on the fact that the standardised alphabet uses the three underlying vowels in all varieties /i/, /u/ and /a/. The existence of allophones in the Cusco-Collao and other Southern varieties is not reflected in the alphabet to allow for intelligibility across Quechua varieties. Thus, while Cusco-Quechua has the allophones /i/ > [e], [ɪ], and /u/ > [o], [ɔ] in the context of a uvular consonant /q/, other varieties, among them some Central varieties, lack them.

Standardisation and unification efforts across all Quechua languages including a common writing system for QI and QII varieties led by linguists and some indigenous educators have been met with strong resistance to the orthographic use of three vowels to symbolise the underlying vowels shared by all Quechua languages. Proponents and users of that system have been labelled ‘trivocalistas’ whereas proponents of the use of five vowels have been labelled ‘pentavocalistas’. Some members of the latter group view themselves as guardians of language authenticity and of its status as a language with the same number of vowels as Spanish (Coronel-Molina 2008).

To summarise the previous two sections, while in each of the three countries most languages of the Quechua II subfamily are considered vital and there are major efforts to normalise and standardise them, the process is an on-going one that has been met with various levels of acceptance by the different speech communities in each country. At the same time, in each country some varieties have greater vitality and are perceived as being more prestigious than others. In the case of Ecuador, unification based on Highland varieties has led to efforts of reclamation among speakers of Amazonian varieties and, in Peru, normalisation of the Quechua alphabet has led to

controversy with speakers of a prestigious variety opposing it on identity and ideological grounds.

Thus, while standardisation efforts have mixed results, the prestige of some varieties over others within national boundaries is a factor to be considered when assessing the Quechua II subfamily potential as a pluricentric language in the future. In the next section, this issue will become apparent when discussing the diffusion of Quechua II beyond South America.

4. The expansion of Quechua languages outside South America

Some of the Quechua varieties spoken in the three countries have expanded outside the boundaries of their original countries of origin, partly due to immigration to other South American countries, North America, and Europe, and to a great extent due to academic interest in them. Academic diffusion of Quechua II languages outside their original countries, especially in North America and Europe, is characterised by the prevalence of the Highland Ecuadorian varieties and the Southern Bolivian and Peruvian varieties. These have become the prestigious varieties taught outside the countries of origin to the exclusion of other varieties with lower numbers of speakers and certainly, for obvious reasons, to the exclusion of endangered varieties such as the Amazonian varieties from Ecuador and Peru.

Kichwa/Quechua academic programs have existed in the United States since the 1960s. A pioneer program was conducted by Prof. Luis Morató, who offered courses in Bolivian Quechua at Cornell University. Elementary Quechua continues to be offered at Cornell since then and can be found at: <https://classes.cornell.edu/browse/roster/FA21/subject/QUECH>.

Another long-standing Quechua program is the one at the University of Illinois of Urbana-Champaign where Prof. Clodoaldo Soto Ruiz taught Ayacucho Quechua, a Southern variety, for more than 25 years.

Currently, Prof. Carlos Molina offers beginner and intermediate Quechua with a wider perspective in terms of dialectal variation although maintaining a focus on Southern Quechua <https://igi.illinois.edu/node/67>.

Southern Quechua courses are also offered at Ohio State University (Bolivian Quechua) by Prof. Elvia Andía-Grágeda (<https://sppo.osu.edu/people/andiagrageda.1>), the University of Michigan (Peruvian Cusco-Collao Quechua) by Prof. Adela Carlos Ríos (<https://ii.umich.edu/lacs/students/language-programs/quechua.html>), and

New York University (Peruvian Cusco-Collao Quechua) by Prof. Odi Gonzáles (<https://as.nyu.edu/clacs/quechua.html>).

At Rowan University in New Jersey, Prof. Marilyn Manley directs the Quechua Linguistics Project, also an expert in Cusco-Collao Quechua. The project's main goal is "To create Quechua language learning resources online to be utilized by Rowan students and others interested in learning Quechua." (<https://chss.rowan.edu/departments/sociology/maru/quechualinguisticsproject.html>)

The University of California Los Angeles offers Highland Ecuadorian Kichwa (<https://www.spanport.ucla.edu/undergraduate/languages-offered/quechua-language-program/>) and the University of Pennsylvania (<https://web.sas.upenn.edu/quechua/>) and the University of Colorado at Boulder (<https://www.colorado.edu/lasc/quechua>) offer Quechua I (Central Quechua) courses.

Quechua teaching materials for US students are beginning to emerge too, such as the Teaching Quechua in the US. Introducing Ayni, a Beginner-level Southern Quechua Textbook project currently under development by DeLoge, Manley, and Molina-Vital (2020). (<https://www.youtube.com/watch?v=XuvZUlwBis>).

As part of a large reclamation effort of the language, Kichwa/Quechua organisations, such as The Quechua Alliance <https://thequechua.org> spearheaded by Prof. Americo Mendoza-Mori, formerly at University of Pennsylvania and now at Harvard University, have started to engage with migrant communities of Kichwa/Quechua speakers in an effort to reclaim and revitalise the language beyond South America.

We can thus speak of diffusion of the Quechua varieties from their original regions to other subcontinents. Furthermore, this diffusion is characterised by a prevalence of some varieties such as Highland Ecuadorian Quechua and Bolivian and Peruvian Southern varieties, and to a lesser extent some Central varieties. Amazonian varieties, on the other hand, do not enjoy vitality in their place of origin and are not part of this process of academic diffusion.

To the extent that this pattern corresponds to that of pluricentric languages one could argue that Quechua-II language family is indeed a pluricentric family. However, as previously mentioned, this characterisation of Quechua II is not a straightforward one as standardisation within each country,

while under way, is still in process, and language shift to Spanish threatens the vitality of these language varieties.

5. Summary

The Quechua family of languages spans across several countries. The demography and vitality of the Quechua languages differ across and within countries. Legislation and policies on indigenous languages vary across countries and new “standardised” varieties of the languages are emerging in each of the countries, especially in those with large numbers of speakers, although sometimes to the detriment of local and regional varieties.

At the same time, the languages of the Quechua II family remain minoritised within the states in which they are spoken as home languages, with low levels of intergenerational transmission for some varieties, making it very difficult for them to become established centres of diffusion. Through academic interest and international migration, some of the most widely spoken varieties (also perceived as more prestigious) have now first and second language speakers in countries such as the United States where Quechua has not been spoken as a home language. However, this academic interest may have no impact on the social vitality of languages of the Quechua II family in their speech communities.

This complex situation forces us to evaluate what the notion of pluricentric languages means for a minoritised family of languages spoken across different countries and in the northern hemisphere diaspora, and how this notion can help or hinder revitalisation and reclamation efforts of the indigenous populations in the countries where the languages are spoken.

References

- Albó, Xavier (1987): Problemática lingüística y metalingüística de un alfabeto quechua: Una reciente experiencia boliviana. In: *Allpanchis*, 19/29-30. P. 431-467.
- Ammon, Ulrich (1989): Towards a descriptive framework for the status/function/social position of a language within a country. In: Ammon, Ulrich. (ed.). *Status and function of languages and language -varieties*. Berlin. Mouton de Gruyter, 21-106.
- Andrade, L. / Howard, R. (2021): Las lenguas quechua en tres países andino-amazónicos: de las cifras a la acción ciudadadana. In: *Káñina*, 45/1. P. 7-38.

- Aschmann, Richard P. (2007). Quichua.net. Retrieved from <http://quichua.net/Q/Ec/espanol.html>. December 13 2021.
- Cerrón Palomino, Rodolfo (1988): *Lingüística Quechua*. Cuzco. Centro de Estudios Rurales Andinos Bartolomé de Las Casas.
- Coronel-Molina, Serafín (2008): Language Ideologies of the High Academy of the Quechua Language in Cuzco, Peru. In: *Latin American and Caribbean Ethnic Studies*, 3/3. P. 319-340. DOI: 10.1080/17442220802462477
- Cusihuamán, Antonio (1976/2001): *Gramática Quechua*. Cuzco. Centro de Estudios Rurales Andinos Bartolomé de Las Casas.
- DeLoge, A./ Manley, M./ Molina-Vital, C. (2020): Teaching Quechua in the US. Introducing Ayni, a Beginner-level Southern Quechua Textbook project Conference presentation. Retrieved from <https://www.youtube.com/watch?v=XuvZUlwBis>. December 13, 2021.
- Durston, Alan (2014): Standard Colonial Quechua. In: Salikoko Mufwene. (ed.). *Iberian Imperialism and Language Evolution in Latin America*. Chicago. University of Chicago Press, 225-243.
- Eberhard, D./ Simons, G./ Fennig, C. (2019): *Ethnologue: Languages of the World*, 22nd Edition.
- Escobar, Anna Maria (2007): Lengua y migración en el mundo hispanohablante. In: *Revista Internacional de Lingüística Iberoamericana*, 5. 93-107.
- Grzech, K./ Schwarz, A/ Ennis, G. (2019). Divided we stand, unified we fall? The impact of standardisation on oral language varieties: a case study of Amazonian Kichwa. In: *Revista de Llengua i Dret, Journal of Language and Law*, 71. 123-145. <https://doi.org/10.2436/rld.i71.2019.3253>
- Haboud, Marleen (2004): Quichua Language Vitality: An Ecuadorian Perspective. In: *International Journal of the Sociology of Language*, 167. 69-81, <https://doi.org/10.1515/ijsl.2004.022>.
- Haboud, Marleen. (2010): South America: Andean region. In Christopher C. Moseley (ed.). *Atlas of the World's Languages in Danger* Paris. UNESCO, 95-102.
- Haboud, M./ Limerick, N. (2017): Language Policy and Education in the Andes. In: *Language Policy and Political Issues in Education* Springer International Publishing. 435-447. https://doi.org/10.1007/978-3-319-02344-1_32.
- Hornberger, N./ King, K. (2001): Chapter 7. Reversing Quechua Language Shift in South America. In: Joshua Fishman (ed.). *Can Threatened Languages be Saved?* Bristol, Blue Ridge Summit. Multilingual Matters, 166-194. <https://doi.org/10.21832/9781853597060-009>

- Kalt, S./ Geary, J. (2021): Typological shift in bilinguals' L1: Word order and case marking in two varieties of child Quechua. In: *Languages* 6/42. P. 1-23.
- King, Kendall (2001): *Language Revitalization Processes and Prospects: Quichua in the Ecuadorian Andes*. Cambridge. Multilingual Matters.
- Limerick, Nicholas (2018): Kichwa or Quichua? Competing Alphabets, Political Histories, and Complicated Reading in Indigenous Languages. In: *Comparative Education Review*, 62/1. 103-124. <https://doi.org/10.1086/695487>
- Limerick, Nicholas (2020): Speaking for a State: Standardized Kichwa Greetings and Conundrums of Commensuration in Intercultural Ecuador. In: *Signs and Society*, 8/2. 185-219.
- Luykx, Aurolyn (2011): Paradoxes of Quechua Language Revitalization in Bolivia: Back and Forth along the Success-Failure. In: *Handbook of Language and Ethnic Identity: The Success-Failure Continuum in Language and Ethnic Identity Efforts*. 2. 137-150.
- Mannheim, Bruce (1991/2011). *The language of the Inka since the European invasion*. Austin. University of Texas Press.
- Martínez, Raúl (2019): The verbal art of Kichwa reclamation. In: *Anthropology News*. September 19.
- Ministerio de Cultura de Peru (2021): Retrieved from <https://bdpi.cultura.gob.pe/lenguas/quechua>. December 13 2021
- Ministerio de Educación de Peru (2018): Retrieved from <http://www.minedu.gob.pe/n/noticia.php?id=46036>. December 13 2021.
- Muhr, R./ Marley, D. (eds.). (2015): *Pluricentric Languages: New Perspectives in Theory and Description*. Bern, Switzerland. Peter Lang.
- Parker, Gary (1963): La clasificación genética de los dialectos quechuas. In: *Revista del Museo Nacional*, 32. 241-252.
- Plaza, Pedro (2009): Quechua. In: Crevels, E. / Muysken, P. (eds.). *Lenguas de Bolivia*. Ambito Andino. La Paz, Bolivia. Plural Editores, 217-284.
- PROEIB ANDES (2021) Retrieved from <http://www.proeibandes.org/> December 13 2021.
- Povilonis, Natalie (2021): *Beyond the monolith: Sociolinguistic variation in Chanka Quechua*. Ph.D. Dissertation. New York University.
- Sánchez, Liliana (2003): *Quechua-Spanish Bilingualism. Interference and Convergence in Functional Categories*. Amsterdam. John Benjamins.

- Sillar, Bill (2012): Accounting for the spread of Quechua and Aymara between Cuzco and Lake Titicaca. In: *Proceedings of the British Academy*, 173. 295-319.
- Torero, Alfredo (1964): Los dialectos quechuas. In: *Anales Científicos de la Universidad Agraria*, 2. 446-478.
- Wrobelwski, Michael (2020): Amazonian Kichwa Proper: Ethnolinguistic Domain in Pan-Indian Ecuador. In: *Journal of Linguistic Anthropology*, 22. 64-86.

II. The Pluricentricity of Spanish in the Americas

Sebastian GREUSSLICH

(Bonn University, Germany)
sebastian.greusslich@uni-bonn.de

The Pluricentricity of Spanish in the Americas – Current Perspectives on Variety Dominance, Standardization, and the Media

Abstract

This paper offers a brief outline of several important lines of research that concern the complex mutual interdependence of three key topics related to the current state of the pluricentricity of Spanish in the Americas, as well as its possible future developments. These topics include: the dominance hierarchy among the varieties of Spanish in the Americas; the social conditions as well as the political aims relevant for standardization efforts in different Hispanic Latin American countries; and finally, the contradictory impact of the conventional tertiary mass media on these processes, which is presently overridden in a significant manner by the growing presence of digital media in established social communicative practices. While elucidating their interrelatedness, the systematic tensions that are generated in all areas relevant to linguistic normativity as consequences of globalization dynamics also become more visible in terms of their ubiquity.

1. Introduction – Pluricentricity in 21st Century Latin American Spanish

The pluricentricity of language cultivation in the Hispanic world has developed in a complex as well as dynamic fashion during recent years. At the same time, Spanish-speaking societies are subject to some major overarching transformations related to globalization,¹ even if to different degrees. The aspects that are primarily relevant for the purpose of this paper are: the definitive assertion of large media companies as relevant normative agents in the communicative ecology of Hispanic societies on both sides of the Atlantic (see Greußlich / Lebsanft, 2020, and the contributions therein); the several strategic and pragmatic moves of the Real Academia Española (henceforth RAE)

¹ Cf. for summaries of a wide range of issues concerning globalization, Kühnhardt / Mayer (2017).

and its sister academies in order to account for a changing normative setting (cf. Amorós, 2014; Lauria, 2019); digital media as a recent means of communication that has the potential to transform the very concept of linguistic normativity and correctness (cf. Cabrera / Lloret, 2016).

1.1. Normativity and Space in the Hispanic World

To evaluate and understand the interplay of these socio-communicative dynamics, as well as the motives of their respective agents, it is appropriate to give a brief outline of the normative setting that has guided the reflections and policies on the pluricentricity of Spanish so far (cf. Lara, 2004).² The basic feature of this setting is “hierarchy”, and it unfolds in two domains: space and discourse. This characterization deserves a brief commentary. If we imagine the normative hierarchy of Spanish as a metaphoric triangle, then its apex corresponds to the technical, analytic concept of *norma panhispanica* (cf. Coseriu, 1990; Méndez García de Paredes, 2014). Whereas the empirical content of this norm in terms of linguistic features is difficult to determine, it is, however, defined as the abstract notion of all the common features shared between varieties of Spanish that guarantee its “unity” as a means of communication available for all speakers, regardless of their geographic origin. Inasmuch as this abstract concept of shared normativity is empirically underspecified, it is also an essentially contested concept which invites criticism in terms of potential globalist ideologies (cf. Section 3). Whereas the abstract idealistic nature of the *norma panhispanica* is not called into question, but rather emphasized by theorists who regularly identify it with a rather traditional realm of elaborate cultural products, specifically literature, it is still necessary to account for those norms found on the lower levels of the normative triangle. Indeed, these are located spatially and are characterized by a set of concrete traits which become ever larger as the spatial domain ascribed to the variety at hand narrows. In this sense, the normativity triangle represents a continuum of abstractness / concreteness and the main impact of globalization consists of shifting the preferred *locus* of normativity away from the nation state at its centre in a twofold and divergent manner, towards an

² This historical and systematic account does justice not only to the American stakeholders of the problem at hand, but rather assumes the task to elaborate a historical foundation of contemporary systematic evaluations from a point of view of presumed neutrality. As such, Laras model has gained much credit in the entire Spanish-speaking world (and deservedly so) as the analytic standard in the field. It is interesting to note, however, that we are facing a kind of convergence which is characteristic for the very top of the normative hierarchy, as will be discussed in the following sections. Cf. Zimmermann (2021) for an entirely recent outline.

imagined global Hispanic community that represents an ideal and abstract concept, and towards local communities of practice that are instantiated by social practice and tend to enhance their vindicative stance with regard to abstract norms that are increasingly perceived as hegemonic and imposed by force on a local community (cf. Morgenthaler García, 2008; Amorós / Méndez García, 2018).³ The correlations explained so far are captured in the following figure:

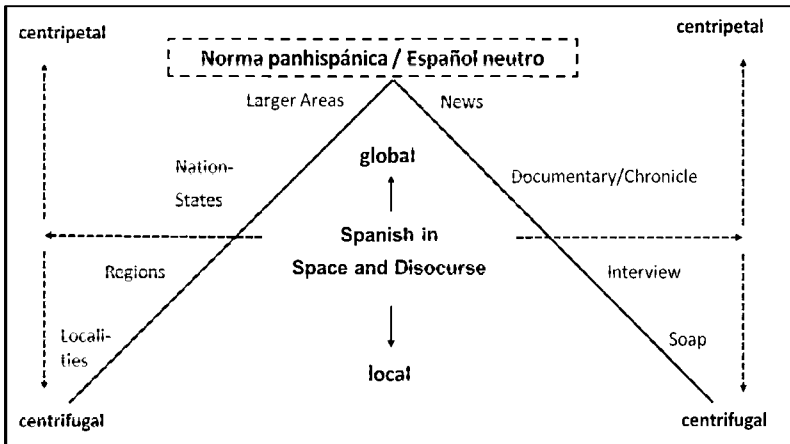


Figure 1

On the left leg of the triangle, we see those spatial categorizations which are most important to every-day perception among speakers as well as to descriptive habits among specialists. As we see on the right leg of the triangle, the relative distance from the above-mentioned *norma panhispánica* also has another dimension, namely degrees of conceptual elaboration, according to different discourse profiles and norms (cf. Garatea Grau, 2008). To the same degree that concrete spatial location of a variety becomes possible, its disposition towards a lower degree of elaboration and an application in less elaborate discourse environments increases.⁴ Here, it becomes clear why it is

³ The references cited above refer to the case of Canarian and Andalusian Spanish respectively. Despite the obvious empirical differences, the structure of the status problem and its social as well as political implications in regard to handling linguistic normativity within national territories are, however, the same. Of course, the evaluation assumed by different specialists can greatly differ, as the references mentioned above show. The recent study Peter (2020) even shows how the perceived status of Andalusian Spanish maintains a mutually dependent relationship with its construction as an abstract entity in metadiscourse.

⁴ This correlation is also expectable as a derivation of the systematic correlations between shared physical space on the one hand and communicative immediacy on the other hand, which has been famously developed in Koch / Oesterreicher (2007).

analytically justified to set a parallel between the *norma panhispánica* as the most abstract category of spatial norm on the one hand and the *español neutro* as the most abstract category of mass media discourse norm on the other hand. In the following section, some remarks on the interrelation between degrees of elaboration in discourse, the communicative setting provided by mass media and space, are in order.

1.2. Normativity and Discourse in the Hispanic world

Interestingly, as far as mass media are concerned, the widely discussed concept of *español neutro*,⁵ which has been largely developed by international media companies interested in expanding their potential customer share and the overall cost revenue ratio (cf. Hesselbach, 2021),⁶ runs systematically parallel to the *norma panhispánica* inasmuch as it occupies the apex of the metaphoric pan-Hispanic norm triangle (see Figure 1 above). Nevertheless, whereas the *norma panhispánica* is ideally space-oriented (thus associated with the left leg of the triangle) in the first place, the *español neutro* is ideally discourse-oriented in the first place (thus associated with the right leg of the triangle). That is also why its empirical content is very different. Far from abstract, it determines the specific variants which, on different levels of language structuring, shall guarantee a wide acceptance of media products destined to a global market (cf. Bravo, 2009; Pöll, 2021). In practice, this almost always implies a decision in favour of traits that are familiar to a quantitative majority of speakers, such as *seseo*, *tratamiento unificado*. Two consequences derive from this observation. Firstly, the quantity of speakers (i.e., potential customers) is the parameter that ultimately decides the acceptance of a trait as

⁵ There have been discussions on the defining differences between three related concepts: *español neutro*, *español internacional* and, more recently and more radically, *español global* as well as *español total* (sic). As a basic result of about nearly twenty years of discussion (of which Gómez Font, 2004 and the so called "Proyecto Rosario" are a relatively early example), it is reasonable to retain the following assumptions: *Español neutro* implies a reference to empirical traits and their application in mass media and the notion of *español internacional*, although its extension is roughly the same, nevertheless emphasizes the presence of a somewhat neutralized variety of Spanish in global institutions, and thus tends to reflect a broader sociological concern (cf. Bravo, 2009). At the same time, the terms "global" and "total" complement each other, as Lebsanft (2020a) explains, approaching the Spanish language in all its potential varieties at once, i.e., whereas "total" highlights the diasystematic level, "global" refers to the social and institutional presence of Spanish in all its varieties. Rather than corresponding to a descriptive reality, these terms imply a heavy load of ideological projection and have consequently been subject to two types of criticism. They have been proven to be empirically void and, even earlier, they have been deconstructed regarding the ideological and practical inconsistencies they convey (Cf. del Valle / Arnoux (2010), Moreno Fernández (2016), Lauria (2019), Arnoux (2020) as well as Lebsanft (2020a) and Pöll (2021) for more.).

⁶ Without being able to cite all the potentially relevant proceedings of the currently well-established *Congresos Internacionales de la Lengua Española* (CILE), it may be mentioned here by way of example Calderón (2010), one main contribution to the panel entitled "Contribución de la lengua a la renta nacional y al empleo" as well as the panel at the Rosario congress (2019) entitled "El rol del lenguaje en la nueva era de la exponencialidad, su impacto en la educación y el emprendedurismo".

part of the *español neutro*. Secondly, as has been pointed out by Amorós (2020), the acceptable traits mostly converge empirically with specific regional or even national standard varieties of Spanish, and the varieties most strictly prone to be reflected in global media productions are just those which pertain spatially to the largest Hispanic Latin American countries, i.e., Mexico, Colombia, and Argentina. One important implication of this finding must be stressed: every norm has its spatial specification, its locus on the map, even if it is posited as abstract and / or compromised between many or all of the speakers potentially involved. Furthermore, if this consequence is applicable, then dominance hierarchies of varieties (cf. Muhr, 2012), even if often severely criticized from a deconstructive viewpoint, effectively tend to be replicated within Hispanic Latin America, though maybe involuntarily so.⁷ Finally, the same kind of space-boundedness shall apply vice-versa to the *norma panhispánica* too. In this case, considering its defining properties, the probable locus of the norm is Castile, which may or may not correspond to the true intentions of the key political agents and institutions involved (cf. Lauria, 2019). As an overarching consequence of the interdependencies just exposed, the role of mass media companies as communicative and social agents becomes evident. It consists, above all, in their functional ability to cover the whole range of discourse profiles from the least up to the most elaborate. Just as public institutions, like schools, whose central role in regard to the implementation of codified norms is undisputed, media companies also reach out even more effectively to virtually every member of society, and, in that sense, their normative impact is comparable or even stronger.⁸

Hence, unlike school teaching and textbooks, media companies and their products indeed are potentially linked to the whole range of varieties that are systematically related to the abstract variational architecture constitutive for a modern standard language. Thus, the selection of a specific variety related to a particular discourse area does not apply so strictly to tertiary media contexts. Rather, tertiary media, prone to spoken language, produce a twofold effect

⁷ Among other intricate aspects, this affirmation also constitutes an interesting solution to the long-time pending question of whether it would be the centripetal or the centrifugal normative dynamics induced by mass media discourse that would ultimately prevail in global re-standardization processes of Spanish (cf. the classic reference Narbona, 2001; recently Greußlich / Lebsanft, 2020).

⁸ As pointed out in Sinner (2017) and taken up in Greußlich / Lebsanft (2020), it is notoriously difficult to measure the impact on individual idiomatic behaviour which is potentially generated by media consumption. However, the significant amount of exposure time, which has been documented for many (Hispanic) Latin American countries, does presumably produce a shift towards oral communication and otherwise distant spoken varieties in speakers' consciousness and away from written communication, which traditionally has been the reference point for codification. This confirmed point is sufficient for our further observations.

regarding re-standardization processes due to the social dynamics of globalization.⁹ Firstly, they enhance the relevance of spoken language in general for the variationist stance and constitution of what is acceptable for speakers as a valid standard variety (cf. Tacke, 2015). Secondly, they practically instantiate the effect of globalization on language cultivation that has been repeatedly diagnosed and described in literature, i.e., a growing visibility of strictly regional varieties, mostly dialects, as well as a growing effective relevance of supra-national media norms for the idiomatic conscience of speakers (though these media norms have national origins and, thus, possibly implicitly replicate national hegemonies). Nevertheless, a third aspect is crucial to add here: national broadcast stations are important public agents even today, and their news communications are driven by national concepts of linguistic normativity (see their central position in the triangle in Figure 1). Their normative impact is only relativized, but not cancelled, by glocalization effects (cf. also Ávila, 2016). It is an essentially political question how speakers and societies are going to deal with this constellation of normative tension (cf. also Garatea, 2020). The answers to this question will most likely be different for different states. In any case, this issue can only be informed by scientific observations, not by scientists.

2. The Language Academies, the Media, and their Normative Stance in Globalization

With fine critical perception, at least three elements characterizing the present ideological stance towards the pluricentricity of Spanish, as far as it is coded in meta discourse, have been identified in recent contributions (cf. Amorós, 2014; Lauria, 2019): the deterritorialization and democratization of Spanish as a shared possession of all its speakers; the dismantlement of the historic memory related to Spanish, most notably with regard to its post-colonial connotations in favour of an idealistic “shared heritage and values” view on the subject; the monetarization of language, stressing its importance as an economic asset, and the opportunities it provides to the idealised Hispanic community. This set of assumptions can be justly interpreted as a commitment to a liberal-democratic political agenda driven by, or at least favourable towards, the constraints provided by globalization. In my view, an observation that is sometimes underestimated should be added here, and that

⁹ Regarding re-standardization as key systematic notion in the given context, cf. Amorós (2014), Lebsanft / Tacke (2020).

is the interrelatedness between the very idea of pluricentric re-standardization, the concept of public space, which is heavily bound to mass media, and democratic citizenship. It is worth considering the possibility that the idea of pluricentricity is not so much a result of scientific progress as it is a side effect of social and, above all, communicative dynamics which form the core tool of globalization. The technical impact of mass media has facilitated the perception of the speech of distant others. Traditional notions of linguistic normativity have started to be called into question as a consequence of this (cf. Tacke, 2015; Eckkrammer, 2021; Greußlich in press). Also, it may be no coincidence that the deployment of a practical, concrete approach to the pluricentricity of Spanish has taken place, precisely after the end of militarized, non-democratic governments not only in Spain but all over Hispanic Latin America. It is clear that science acts as an observer in this setting and, at best, as a counsellor, not as the political promoter of pluricentricity.

At first glance, the RAE, as well as the affiliated American academies and international media companies, shows a largely symbiotic behaviour as far as the ongoing rearrangement of linguistic normativity on the global stage is concerned (cf. Greußlich / Lebsanft, 2020). On the one hand, the RAE has openly embraced the concept of regional standards, which may be understood as a plausible move in order to prevent post-colonial prejudice and preserve its public authority. On the other hand, media companies often publicly consult the academies when it comes to determining the most appropriate solutions in terms of the *buen uso del español*. This is an invited convergence, since it was the RAE itself that founded the department *Español al día* in 1998, getting in touch actively with the mass media sector. Thirdly, it is also true that the Spanish News Agency *Agencia EFE* has found it convenient to promote a growing homogeneity of Spanish in the media worldwide and has managed to establish an independent foundation in 2005, the so called *Fundación del español urgente* (formerly *Departamento del español urgente*), that gives instantaneous advice on linguistic doubts among journalists in the Hispanic world as well as internationally (cf. Lebsanft, 2013). Interestingly, it was in 2020 that the original funding of *Banco Bilbao Vizcaya Argentaria-BBVA*, which had given its original name to the *Fundación del Español Urgente-BBVA*, was replaced by a narrower cooperation between the relevant public institutions: RAE, *Agencia EFE*, and *Instituto Cervantes*. This change is consistent with the overarching strategic plan *España global*, which has been in place since 2018, and is aimed at

the effective development and strengthening of a Hispanic sphere on the global stage, driven by linguistic concerns among other parameters. It is interesting to note the intricate dialectics between a progressive spirit and its regressive effects: undoubtedly, the RAE tradition is grounded in the Enlightenment (cf. Lebsanft, 2013; García de la Concha, 2014; Greusslich, forthcoming) and fosters a sincere inclination to reach out for a harmonious exchange with other parts of the world. However, at the same time, there can be no doubt that present tendencies towards worldwide standardization beyond borders and without predefined limits generate new conflicts, whose outcomes are inevitably hegemonic in nature, even if the actors' intentions are not.¹⁰ *España global* may not be a post-colonial strategy according to intention, but it will end up hegemonic according to fact (cf. Lauria, 2019). This, of course, is not a specific “Spanish” fault, but rather an unavoidable implication of globalization.

The self-alleged credibility of the *Fundación del Español Urgente* among Latin American media professionals does indeed exist, and has recently received a solid institutional backing by establishing two American agencies in the Dominican Republic and Argentina respectively. This means that the necessity of being present physically in America has been recognized, and the consequences of it do not hesitate to become manifest. A revealing example for this is the relief expressed by the *Periódico de México* (03/09/2015), an important Mexican newspaper, in a comment on the re-edition of the *Manual del español urgente* in 2015: “*Vuelve el “Manual de español urgente”, la referencia para escribir en español*”. Here, we can notice how a possible conflict regarding the issue of linguistic norm in public discourse is solved, not by hegemonic force, but by voluntary adherence to the leading role of the most emblematic institution in the field, i.e., the RAE. This, once again, allows for an ambiguous interpretation: as a critical diagnosis, this may indicate a submissive behaviour (cf. Lauria / López García, 2009; Lauria, 2019). Nevertheless, this view is not straightforward. Rather, the historically-gained recognition of the ancient institution of the RAE is also real (cf. Lara, 2004; Garatea, 2020),¹¹ and the fact that this historical coincidence generates hegemonic secondary effects nowadays is contingent in this view. From that perspective, the key point to

¹⁰ Thus, in the end they possibly correspond to radically critical analysis as given in Moreno Cabrera (2015), which are often well-informed and brilliantly executed, but also far from ideologically balanced.

¹¹ Just as Pöhl (2021) duly reminds us, a real and sometimes strong orientation towards Spain also persists in certain areas of Latin America. This view is consistent with the evaluation of Spanish as a significant, yet not fully pluricentric, language (Prieto de los Mozos / / Amorós 2018).

reflect upon would not only be the dominance hierarchy among nations, but also dominance hierarchies within Hispanic Latin American societies.

Finally, the obvious collusion between the state and the media disposes of yet another revealing twist: digitalization (cf. Lauria, 2017a). Not only do important newspapers, among them the Latin American edition of *El País* (cf. Lebsanft 2020a), transform into increasingly digital products whereas traditional paper-bound distribution diminishes just as in other parts of the world,¹² the RAE and ASALE have made constant efforts to enhance their presence in digital media for nearly a decade in order to adapt to the changing character of public space, and to be able to reach out for the non-professional speakers-citizens that are so crucial for the reshaping concept of *buen uso* in globalizing Spanish language culture.¹³ This can be seen, once more, as an appropriate strategic move in order to preserve credibility and relevance as a linguistic authority for the public, as well as to maintain the collusive, symbolic competition with private media companies, which is also a constructive collaboration on another level.

In the next section, we will briefly discuss in which ways the available data impose constraints on the current codification practice. We shall give a brief account of the current state of codification efforts in different areas of the Hispanic world, all of which share their foundation in pluricentricity, but differ in practice due to the scientific record on which they are built in each case, and the ideological stance of the societies they serve.

3. Research Traditions on Variation and Norms and their Impact on Codification

It is fair to say that, since the important speech given by Dámaso Alonso on the Unity of Spanish (1964; cf. also del Valle, 2013) and the set-up of the now famous *Habla culta* project at the IV Congreso ASALE in 1964 (cf. Lope Blanch, 1986), two different research strategies regarding the varieties of Spanish in the Americas have been put in place. A conventional dialectological approach had been the most significant for research on variation of Spanish in the

¹² As a relevant example, *El País* shut down its Latin American print offices in 2019, whereas a daily digital edition persists. This shift can be interpreted twofold: *El País* apparently has not gained sufficient recognition in the American context to achieve economic sustainability (54000 sales compared to 206,000 for *El Clarín* of Buenos Aires and 60,000 for *El Comercio* of Lima); digitalization is obviously a major trend on the newspaper market in general.

¹³ This is true in regard to practically all of the major lexicographic endeavours that are carried by the RAE / ASALE cooperation. Moreover, all the relevant channels of counselling that have been maintained by Fundéu, and, more recently RAE, have been opened to the digital sphere.

Americas during its first stage, which had developed from the foundational article by Henríquez Ureña (1921). On a methodological level, it implies the identification of isoglosses, but notably without offering reliable hypotheses about potential reference points for variation, in order to eventually reconstruct a complex diasystem of varieties in the coserian sense. These methodological considerations, regarding the early and long-lasting “dialectological” approaches to “American Spanish”, are clearly not meant as cheap and unhistorical reproaches, but, rather, are intended as observations of the consequences for research and data available today (cf. Lipski, 2008; Greußlich, 2015). These data are characterized, on trend, by a lack of functional status qualification. They are, to this effect, positivistic, and this is harmful to the key role that variation assumes in recent pluricentric accounts of codification. As the “pan-Hispanic” reference case in point, the *Nueva gramática de la lengua española* shall be taken into consideration in this regard:

Es imposible que en una descripción gramatical se analicen con similar profundidad todos los tipos de variación lingüística [...]. Es plausible, en cambio, resaltar de forma somera los aspectos más notables de cada una [...]. Los especialistas suelen distinguir entre la variación histórica, la geográfica y la social, y acostumbran a dividir esta última en dos tipos: la relativa a los llamados sociolectos [...] y la relativa a los niveles de lengua, es decir, los estilos o registros como los que permiten oponer la lengua coloquial a la empleada en situaciones formales, o diferenciar el habla espontánea de la lengua más cuidada.

La presente gramática atiende en alguna medida a todas estas formas de variación, pero ha tenido que limitarse en muchos casos a consignar su existencia, puesto que analizar en profundidad cada uno de estos factores habría conducido de modo inevitable a un conjunto de estudios monográficos (NGLE, vol. I: XLIII).

[It is impossible to give an equally profound analysis of every type of linguistic variation in a grammatical description. [...]. However, it is plausible to accentuate / to point out the most crucial aspects of each one of them in a sketchy way. [...]. Specialists are accustomed to distinguishing historical variation from geographical and social variation and they commonly classify this last one into two types: the so-called social dialects [...] and the linguistic levels, i.e., styles and registers, as they allow for the establishment of a contrast between colloquial language and language use in formal situations or differentiation between spontaneous speech and more diligent speech.

The present grammar, in some way, reaches out to all these types of variation, but must restrict itself to indicate its existence because analyzing profoundly each one of these factors would have inevitably led to a series of monographic studies.]

The available data often cannot keep up with these sophisticated intentions. Therefore, the descriptive meta discourse, whose prescriptive, codifying effects are mostly limited inasmuch as the selection and mark-up of particular variants is systematically avoided, reflects the evaluative uncertainties that inevitably emerge, not only among the wider public as Moreno Fernández (2012) duly suggests, but even among the specialists who have been in editorial charge (cf. Méndez García de Paredes, 2014; Greußlich, 2015). As has already been diagnosed in the classic study of Berrendonner (1981), a camouflage of these uncertainties occurs regularly in discourse. In the case of the NGLE, the review undertaken in Greußlich (2015) has shown three descriptive strategies to be essential in this regard:

- (1) Zooming in on or recording the act of observation by using verbs which refer either to the observed speech act as such (*usarse*), an involuntary observation (*encontrarse, observarse etc.*), or a voluntary one (*documentarse, registrarse etc.*);
- (2) Positing the data as objects of greater or lesser commonality by using quantitative adjectives: *natural, (in-)frecuente, general, normal, raro, etc.*;
- (3) Linking a specific phenomenon to different levels of variational abstraction (*lengua, habla, variedad, registro*). The status of these qualifications remains systematically fuzzy in all of the aforementioned cases.

In this regard, the new problem that apparently emerges as a consequence of reassessing the interrelation between description and prescription, which is ongoing in order to give more space to the ordinary citizen of democratic societies and to avoid elitism in globalization, is not only technical insofar that it could be mitigated by additional and better suited data sets. Rather, it is systematic and best regarded as a necessary implication of the current ideological stance towards codification that pretends to be neutral in regard to prescription and description, and instead embraces the metaphor of mirroring the speaker's usage (cf. Tacke, 2011). That might be an impossible thing to achieve in practice, even if it is, at the same time, a recommended discursive move in order to regain credibility in societies that have gone through historic periods of dictatorship, as is the case in vast parts of the

Hispanic world. The resulting lack of normative orientation points toward an undermining effect on normativity in general, which is driven by a growing tension between local linguistic habits and global economic incentives that work negatively on the concept of “nation” as a plausible reference point for social identification (see Table 1). In other words, this new approach to regulating the usage of Spanish “by description”, so to speak, theoretically implies a far more detailed acquaintance of the varieties of Spanish, in order to attach credibility to the ideal overarching pan-Hispanic norm.

Before we revise these conditions in some detail (cf. Section 4), a comment on the different types of lexicographic enterprises that are currently developed in the Hispanic world is in order. As the second key text genre in codification, their respective weight and role within the pluricentricity of Spanish ought to be pondered. In order to do so, differential national lexicography, differential Pan-American lexicography, and integral lexicography shall be considered separately and discussed regarding their interrelatedness. Whereas it may be less surprising to recognize the fact that integral lexicography is limited in its applicability to those nations disposing of the required resources, the complementary observation that differential dictionaries of the *diccionario de -ismos* type (cf. Ávila, 2003; Ezcurra, 2020) are available for and very welcome among the broader public in many Hispanic Latin American countries, such as Peru, Colombia, Chile and even Mexico, is all the more revealing. Although a serious debate on the issue of its appropriateness and methodological foundations has been ignited (cf. Greußlich, 2020), the editorial economic success of these enterprises cannot be denied, whereas it is an interesting but complex question where this success and the corresponding public demand might come from, another perspective which shall be highlighted here. In the field of pan-Hispanic lexicography coordinated by the ASALE, we can currently observe a neat systematic tendency towards moving away from the established concept of *americanismo léxico*, replacing it with a new one of a somewhat reversed meaning. It seems that the substantial *Diccionario de americanismos* (DA), published by the ASALE in 2010, is not only a substantial empirical achievement in the field, despite its duly commented technical and conceptual lacunae (cf. Lara, 2012; Lauria, 2017; Ezcurra, 2020; Fajardo, 2021), but at the same time paradoxically restores the gap between an imagined centre (Spain) and a periphery (America) inasmuch as it does not converge with peninsular usage, as the aforementioned authors point out. Why is this the case?

Rather than developing this lexicographic branch further and resolving the diagnosed desiderata, we are presently witnessing a strategic rearrangement which emphasizes the pan-Hispanic perspective as the new reference point for best practices in codification. Concretely, the specifically American content of the DA has gradually been integrated into the general *Diccionario de la lengua española* (DLE), as well as on a digital basis, since 2014. The DLE, on its part, has been intentionally renamed in order to back and reflect the pan-Hispanic ideologeme of Spanish as a commonly shared asset and possession. Certainly, the new label also reflects an undoubted institutional reality, which is manifest in the editorial dynamism of the ASALE, and the collaboration that supports it. However, a crucial quantitative criterion interferes in practice. Being registered in at least 14 countries is necessary and sufficient for an entry from the DA in order to count as part of the general Spanish lexicon. Consequently, the remaining entries of lesser distribution among Latin American countries will figure as *americanismos* in an even stronger sense, therefore being truly stigmatized as the lexicographic leftover with no access to the sphere of general communication. The established text type of *diccionarios de -ismos*, instead of being forced to innovate and develop methodologically, remains in its accustomed position as a mirror of cultural vestiges. This state of affairs is essentially consistent with pan-Hispanic language policy, but it definitely sets the global stage as the genuine locus of codification, whereas the concept of nation appears subordinate. Admittedly, the undermining of the nation as a point of reference is essential to the social dynamics of globalization, so that it is a licit policy in the given context. However, it remains an open question if this “global stage” mode of codification generates sufficient plausibility among the ordinary speakers who are so vital to the liberal worldview underlying this policy.

The essential need for acceptance runs counter to the systematic lack of normative orientation that is part of these recent adaptive moves of codification in globalizing environments, as has been commented above. Thus, it comes as no surprise that a particularly national mode of linguistic self-assertion by independent codification exists, hence the relevant incentives are there. This type of nation-bound codification is best represented by two emblematic endeavours of integral lexicography in recent times, which are the *Diccionario del español de México* (DEM 2010) and the *Diccionario integral del español de la Argentina* (Plager, 2008). These two exemplary dictionaries share several interesting properties (cf. Großlich, 2020, and the references therein): they

have been demanded in their respective societies for a long time; they have been executed by private institutions, independent from the state and government; they have been praised for their innovative methodology by a long row of experts; each of them is embedded in one of the largest speech communities of the Hispanic world. This last circumstance gives credit to the idea that large numbers of speakers and potential clients, concomitant economic incentives, a long historic record of ideological autonomy, and / or relatively stable institutions are criteria that make normative independence possible (cf. Fajardo, 2021). If this statement is correct, the only country lacking in the above list is Colombia, and the reasons for its current absence constitute a relevant topic for discussion, which has to be left for another occasion. Once more, however, the known effect of dialectic indecisiveness emerges. Concretely, it is in both of these cases that the integral lexicographic project co-exists and also competes with the respective exponents of the above-mentioned *Diccionario de -ismos* programme that is supported institutionally by the respective language academy in each of those countries.¹⁴ Consequently, an ironic and somewhat unexpected shift in evaluation and status of those texts can be noticed: although meticulously worked out and methodologically recognized among specialists, they can hardly escape their globally oriented competitors and, on the one hand, have to assume an open ideological confrontation. On the other hand, they instigate by themselves, even if on a minor spatial scale, a type of dominance hierarchy in relation to smaller Hispanic Latin American countries which do not afford or even consider the elaboration of integral dictionaries to be appropriate on their own, rather being geared to their larger neighbouring countries (cf. Quesada Pacheco, 2014; Fajardo 2021). Of course, this effect is neither intended nor explicit, rather, it is due to quantitative prevalence and their consequences in terms of symbolic weight. However, this consequence is real too, not polemically invented, and it shows in which way it is true that the present evasiveness of the notion of nation-state in the context of advanced globalization has not yet been convincingly compensated by another

¹⁴ Cf. Lara (2012) and Company (2012) in order to trace the highly polemic Mexican case; cf. Lauria (2015) for the Argentinian case. Cf. Fajardo (2021) for both.

Part of this complex dialectics is the insight that differential lexicography might still be the methodology of choice, if meticulously applied, as in the dictionaries issued by the famous Augsburg project since 1993. However, it is also true that those dictionaries, although they largely avoid negative connotations which point towards an image as undeveloped and inferior cultures, they have not reached a wider public until today but count as a valuable research tool for specialists.

appropriate concept that would do the same for speech communities seeking appropriate and convincing orientation for usage.

Hence, it is reasonable to suppose that, beyond the lack of data, there are also systematic aspects of linguistic codification in global speech communities, which contribute to transform the very notion of linguistic norm, whose future is unclear at this point in time. The seriousness of this issue involving plausibility and acceptance becomes even more evident if one takes into consideration the most recent large dictionary which has been projected as pan-Hispanic, i.e., the *Diccionario panhispánico del español jurídico* (DPEJ, 2018). This dictionary and its immediate predecessor, the *Diccionario del español jurídico* (DEJ, 2016), arguably form part of a broader editorial strategy, described above as a shift towards pan-Hispanic, as the new normal of codification on the global stage. Moreover, they share the transposition in a digital environment, including the respective user interfaces for direct public access online. As such, they serve important symbolic as well as strategic aims: symbolic aims, inasmuch as the field of law stands emblematically, for the effectivity and transparency of liberal international exchange and commerce, strategic aims inasmuch as it contributes not only to consolidate the pan-Hispanic worldview, but also helps, practically, to foster smooth processes within the Hispanic world as an economic space.¹⁵ Hence, practical and symbolic effects converge, helping to put the *Marca España* strategy into practice and, at the same time, increasingly suggesting the pan-Hispanic level as the new uncontested reference point for social self-awareness, more so than other areas of knowledge could have done. Interestingly, the tendency to give privilege to the universal, global stage, by strengthening the image of Spanish as a tool, and a set of values shared by all its speakers is what public agents, such as the ASALE, and private agents, such as *CNN en español*, share as their ideological basis. Also, their strategic goals are the same inasmuch as they seek a high degree of linguistic coherence and uniformity of their target space (cf. Gómez Font, 2012).¹⁶ However, their social roles and purposes as institutions are different in each case. That is the reason why the empirical, concrete

¹⁵ Fully consistent with this strategy, RAE has launched a special programme called *Lengua Española e Inteligencia Artificial* (Spanish Language and Artificial Intelligence) in order to secure the embeddedness of Spanish in digital media, taking into account the assumption that digital tools are key for global communication, as well as for the future of linguistic normativity. In this same context fits the frequently repeated assertion that Spanish is the second most important language on the web worldwide (cf. Lauria, 2019). Visibly, a practical and a promotional aspect converge here.

¹⁶ This is not contradictory to celebrating diversity at the same time. This is true because the presupposition here consists in rightly assuming the possible degree of linguistic diversity as limited and, thus, convergent with the overarching goal of linguistic unity.

application of their ideological stance towards pan-Hispanic normativity yields systematically diverging results, and these will be discussed in the following section.

4. The Impact of Mass Media on the Pluricentricity of Spanish

In another important research project, undertaken under the patronage of the ASALE, by Mexican specialist R. Ávila (Proyecto 6: *La difusión del español en los medios*), two fundamental findings have been developed since 1998. Firstly, empirically, an ecology of linguistic normativity of Spanish in the media exists which is relatively independent from spoken Spanish outside of the media. The relative independence between these two spheres manifests in a particular way, which can be best reflected by referring, once again, to the triangular figure introduced in Section 1, and will be discussed below. Secondly, it has become evident that an ever-growing convergence with national and regional norms of spoken Spanish (*normas cultas*) does take place going down the elaboration cline, which reflects the relevant media discourse genres, such as evening news (high degree of formality and elaboration), feature (intermediate degree), or soaps (low degree).¹⁷ These relative correlations conversely imply a relative accessibility for dialects in informal genres, like soaps, and accessibility for regional or national standard varieties in elaborated but nation-bound discourse types, like national news broadcasting. A prohibitive environment for the application of concrete, empirical, spatially grounded linguistic norms is to be found solely in globally-oriented, international broadcasting companies, whose likely prototype is *CNN en español*.

To summarise, it is fair to say that, starting from a general level anchored in the analysis of globally oriented media companies and products, the awareness of the growing presence of regional norms in media products, situated on the lower levels of the triangle, has been gradually added over the years (cf. Ávila, 2003 vs. Ávila, 2016). In this vein, starting from the macro-norms Alpha, Beta and Gamma according to Ávila (2003), a subset of contextual variants has been identified by differentiating linguistic media norms according to particular countries. Of course, the observation of these variants implies an empirical shift towards the analysis of relatively more informal discourse genres.

¹⁷ Cf. Greublich / / Lebsanft (2020) for references concerning media genres.

Once more, a dialectic twist becomes manifest at this point. Contrary to the *norma panhispánica*, which is abstract in essence and represents an ideal concept grounded in a consciousness formed by a 400-year history of literary and scientific reference works, the *español neutro* is concrete and applicable in essence. This begs the question: which are, in fact, its concrete properties and from which source do they eventually derive? Contrary to the original idea that *español neutro* should be an artificially manufactured selection of most neutral and thus widely acceptable variants on every level of linguistic structure (cf. Bravo, 2009; Gómez Font, 2012), it has become more evident in recent studies (cf. Amorós, 2020) that those variants effectively selected are typically convergent with the most important national media norms, as described by Ávila, relating to Argentina, Colombia and Mexico. This finding is interesting because it undermines the widespread assumption according to which national categories should be outdated in globalization. Rather, it reinforces the argument advanced in critical analysis of normative meta discourse, which states that national interests might appear disguised as international convergence. Ironically, the recent empirical studies suggest that power relations, as indicated, replicate within Latin America where large countries tend to prevail. Once again, this is possibly not a moral problem which could be solved easily based on sufficient benevolence. Rather, according to the interpretation I tend to subscribe to, these are the new and systematic “challenges for Spanish” in globalization and mediated communication.

Another important outcome of these considerations is a systematic reasoning that makes the plain abstractness plausible and, consequently, the material emptiness of the well-known label of *español global* or *total* (cf. Gómez Font, 2012; Moreno Fernández, 2016; footnote 5). It has been investigated and proven empirically (cf. Lebsanft, 2020a) how in fact, behind the label of globality and / or totality, a spatially bound, concrete variety can be found, as is Castilian in the case of *El País* newspaper, whose protagonists have come to promote it as “the global newspaper” since 2014 (see Section 1.2).

Hence, the sometimes-hidden tensions between empirical norms and abstract norms in the media are one key problem when it comes to reconstructing and evaluating the role of mass media communication for the development of Spanish as a pluricentric language. However, an even more fundamental question arises when it comes to evaluating the impact of digital

media on the notion of linguistic normativity, which is inevitably a key issue for the future of pluricentric Spanish as well.

5. Digital Media, Linguistic Normativity and the Pluricentricity of Spanish – A Brief Outlook on Possible Futures

In 2018, the year when the new *España global* strategy concerning Spanish cultural policy was put into force, an apparently new type of publication was issued by the RAE / ASALE editorial board. Its descriptive title is: *Libro de estilo de la lengua española*. According to the user instructions, its acclaimed purpose is to offer normative orientation for the appropriate usage of Spanish in mediated communicative environments:

“[...] es un manual de corrección y estilo del español donde se muestra la evolución que en los últimos años han experimentado ciertas cuestiones gramaticales, ortográficas y léxicas, con especial atención a la escritura digital [...] Por otra parte, el protagonismo de los medios de comunicación audiovisual en nuestros días hace necesario un manual de pronunciación accesible al gran público” (RAE online).

[...] it is a Spanish style and correction manual the evolution which highlights the evolution that certain grammatical, orthographic and lexical issues have experienced in recent years, paying special attention to digital writing. [...] On the other hand, the leading role of audio-visual media today calls for a pronunciation manual accessible to the general public].

Even if the gap in content between this self-declared manual and the slightly earlier *Libro del buen uso del español* (2013) is actually not too large, the rebranding as an exponent of a specific discourse genre conventionally destined to regulating the language usage inside media companies (cf. Lebsanft, 2020b) is noteworthy. It reflects the growing consciousness of the direct impact individual linguistic behaviour may have on the norm through digital media channels. Equally, it is important to observe how, in this regard, the recent convergence between private global media companies and public institutions once again becomes manifest. This was already the case for a publication guided by similar intentions, which was issued in 2012 by the Fundéu experts as *Escribir en internet. Guía para los nuevos medios y las redes sociales* (cf. also Arnoux, 2020).

However, despite all efforts in terms of potentially normative counselling, it seems impossible to evade the undermining effects that emerge

from this type of individualized mass communication in digital settings (cf. paradigmatically, the apprehensions outlined in Mucci, 2019). As Lebsanft (2020b) duly remarks, the same type of normative underdetermination, which has been diagnosed above and is systematically inherent to pluricentricity, persists:¹⁸

Discussing, for example, the use of the pretérito perfecto simple (canté) vs. the pretérito perfecto compuesto (he cantado), an observation like ‘mientras que en algunas zonas lo normal es decir Ayer comí, en otras se podría decir Ayer he comido’ does not work as a valuable recommendation (Lebsanft, 2020b, 640).

This phenomenon is expected to increase in significance as digital media is arguably becoming the dominant means of mediated communication among younger generations. Thus, it is worth considering the core of it in order to better understand the possible future of linguistic normativity.

Let us have a look at the findings of one recent investigation centred on the so-called *pluscuamperfecto mirativo*, a particular usage of the past perfect tense to encode surprise or unexpected new information.

- (1) Había sido tu novia, entonces.
 [So, she had been your girlfriend. – Implication: I did not know this fact. It is right now that I am realizing it.]

For decades, this usage has been well documented as particularly Bolivian in dialectological literature (cf. Mendoza Quiroga, 2015). Yet, the study of Pfänder, Alcón / Palacios (2020) demonstrates some important additional characteristics of linguistic accommodation and change in computer-mediated communication, based on an innovative tool for digital data gathering.

For example, space as a determinant criterion is still relevant. The phenomenon under scrutiny is of Bolivian origin and is decisively propagated by Bolivian speakers, also in computer-mediated discourse. However, the impact of space is relativized inasmuch as speakers from other parts of the world, and of different origin, participate in discourse and assimilate this usage, which opens access to the dominantly Bolivian discourse community. A second example is that accelerated linguistic innovation and change takes

¹⁸ In this regard, it is appropriate to highlight one particular circumstance: whereas the *Fundéu* manual is exclusively aimed at genres and technical issues related to digital communication, the more recent *RAE / ASALE* publication also includes grammatical problems raised frequently in their digital counselling channels. Whereas this decision may be comprehensible from a practical viewpoint, it undermines the coherence of the entire work inasmuch as the tensions related to pan-Hispanic normativity are brought in.

place. Several new pragmatic functions are added to the semantic scope of the tense form under scrutiny. Above all, the basic value of surprise regarding a newly-appearing bit of information is complemented by a second value, which is in contrast to a prior or common expectation, or even an evaluative implication, which can be positive or negative (as in the following example from Pfänder, Alcón / Palacios, 2020).

- (2) micrero muerto de hambre; cree ser mejor que los dems; oh pero perdón no me había dado cuenta que conducir un micro es lo MÁXIMO.
[Miserable bus driver; he believes himself to be better than others; oh but sorry, I had not realized that driving a bus is THE GREATEST THING TO ACHIEVE]

In summary, we can observe a systematic tension between traditional aspects of normativity and their functional reshaping in new environments. On the one hand, space still counts; the origin of norm-setting speakers and traits is palpable. On the other hand, new virtual boundaries are drawn around particular discourses, not around territories (cf. also Greußlich / Lebsanft 2020). Hence, discourse communities are the new locus of accelerated innovation and possible change. What is more, just as the options for recording spoken discourse and making it durable have multiplied, so has the presence of writing in digital settings, like chatrooms or even email exchange, increased. Even if, in all likelihood, many of these innovations will stay ephemeral (as has been common knowledge in historical linguistics), the multiplication of written discourse in everyday communication generates new opportunities for change from below. The propensity of writing in the style of informal discourse increases over long distances because it is asynchronous. In this regard, it is reasonable to presume that the likelihood of change may increase. Otherwise, less influential speech communities, such as Bolivia, may induce successful innovations among these circumstances, and we can comprehend in what sense global elites see an incentive to take control over these growing dynamics. In essence, a classical scenario of media change, as it has been meticulously described for the invention of the printing press, is currently unfolding. However, whereas the printing press fostered the establishment of implicit norms and the fading out of variants, the opposite is true for linguistic usage and structure in digital media. Here, a de-standardization process is taking place. In this regard, the intervention of the RAE / ASALE can even be evaluated as a democratic intent to avoid de-standardization and the turning

of linguistic normativity into a hermetic elitist project. Once more, we find a dialectic tension in place, which is difficult to evaluate straight-forwardly. This is particularly true because, on the discourse level, a narrowing of potential variation is indeed taking place. As has been pointed out in literature (cf. Lauria, 2021), it is the technical frame which sets the limits of expression in digital communication: the number of different fonts, the number of signs admitted, the space where written text can be placed or inserted etc. Thus, the dialectics of normativity in globalizing language cultivation becomes manifest one last time: rather than defending elitism, as often alleged by critical stances towards the RAE and ASALE codifying works, reinforcing an explicitly linguistic notion of correctness and adequacy in digital environments means defending the legitimacy of linguistic expression against externally imposed limits to form.

In conclusion, the essential present-day issues of normativity and codification of Spanish remain unsolved for the time being, as are many other conflictive social dynamics connected with globalization.

6. Conclusion

Inasmuch as the traditional concepts of linguistic norm do not fully apply to the newly emerging type of unstable short-term normativity, which is characteristic for digital media, the fundamental question consists of a general evaluation of the stance of societies and their members towards globalization and its consequences. To what degree may societies in the Hispanic world shape the social impact of globalization in their realms? Will there be an option for the millions of anonymous, individual users of digital media to intervene in this ongoing process of reshaping the very concept and functioning of linguistic normativity, in a way that could be considered as public decision-making, in accordance with the predictions of liberalism? Or will the new globalized elites of the Digital Era alone, in a dialectic and ironic twist, decide on its outcomes? The possibility that the acceleration of codification processes on digital grounds even contributes to undermining the speaker's orientation should be taken seriously inasmuch as constantly changing versions of metatexts replacing former ones in rapid succession are hardly transparent to the public, as far as their transformations, additions and avoidances are concerned. Even if this may be an acceptable side effect of digitalization, the new instability of digital texts runs counter to standardization as we have known it (cf. Bonnin / Lauria, 2015; Greußlich, 2020).

The cultivation of the Spanish language in a globalized and digitalized setting, whereby both characteristics are intimately connected to each other, can be regarded as a relevant testing ground for the development of possible futures in a globalized world. For that reason, it claims the particular attention of specialized scholars in order to communicate the relevant issues and ongoing reflections to a broader public, eventually relating them to other relevant social and political domains. It is necessary to resolve a series of thorny questions in order to achieve a new consensus on how the cultivation of a standard language should be organized, in order to be significantly useful for society and thus gain the fundamental recognition by its members, which is necessary for any normative order to work effectively.

A close monitoring of how these questions are treated by their stakeholders is mandatory, because the fundamental problem is nothing less than the possible emergence of a new type of digital aristocracy (without necessarily being labeled as such), which would obviously be an ironic and undesired outcome, at least from a democratic standpoint. The apprehension of a continuing post-colonial hegemony, as expressed by critical discourse analysts, might dialectically emerge as part of a new social reality of dominance by data control, even if this is probably not the issue that the critical stance was about initially. A consistent solution of the questions raised and exemplified in regard to Spanish in this paper is, thus, no mere nostalgic matter of taste, but rather is vital for the present and future structure and status of Spanish-speaking societies and possibly beyond.

References

- Alonso, D. (1964): Unidad y defensa del idioma. In: *Boletín de la Real Academia Española*, 44, 387-395.
- Amorós, C. (2014): *Las lenguas en la sociedad*. Madrid. Síntesis.
- Amorós, C. (2020): Los procesos de restandarización lingüística en la hispanofonía: prescripción y norma mediática de la CNN en Español. In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 272-296.
- Amorós, C. / Méndez García de Paredes, E. (2018): The status of Andalusian in the Spanish-speaking world: is it currently possible for Andalusia to have its own linguistic standardization process?. In: *Current Issues in Language Planning*, 20, 1-20.

- Arnoux, E. (2020): De la “unidad en la diversidad” al “español auxiliar internacional” en discursos y dispositivos promocionales panhispánicos. In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 39-60.
- Ávila, R. (2003): La pronunciación del español. Medios de difusión masiva y norma culta. In: *Nueva Revista de Filología Hispánica*, 51 / 1, 57-79.
- Ávila, R. (2016): La pronunciación del español en la radio de veinte países hispánicos. In: *Nueva Revista de Filología Hispánica*, 64 / 2, 495-509.
- Berrendonner, A. (1981): *L'éternel grammairien. Étude du discours normatif*. Bern / Frankfurt a. M. Peter Lang.
- Bonnin, J. / Lauria, D. (2015): Diccionarios online: Hacia una nueva fase del proceso de gramatización de la lengua española. In: *Linguas e Instrumentos Lingüísticos*, 36, 293-322.
- Bravo, E. (2009): Español de América, español internacional. In: María V. Camacho Taboada, José J. Rodríguez Toro, / Juana Santana Marrero (eds.). *Estudios de lengua española: Descripción, variación y uso. Homenaje a Humberto López Morales*. Madrid. Iberoamericana, 77-98.
- Cabrera, M. / Lloret, N. (2016): Español en los nuevos medios. In: Javier Gutiérrez-Rexach (ed.). *Enciclopedia de lingüística hispánica*, vol. 2. London, New York. Routledge, 485-493.
- Calderón, S. M. (2010): El español como lenguaje de la economía global. In: RAE / ASALE (eds.). *V Congreso Internacional de la Lengua Española*. Online: <<https://congresosdelalengua.es/valparaiso/paneles-ponencias/politica-economia-sociedad/calderon-sila.htm>>.
- Company Company, C. (2012): Diccionarios contrastivos e identidad cultural. El *Diccionario de Mexicanismos* de la Academia Mexicana de la Lengua. In: Dolores Corbella et al. (eds.). *Lexicografía hispánica del siglo XXI: nuevos proyectos y perspectivas. Homenaje al profesor Cristóbal Corrales Zumbado*. Madrid. Arco / Libros, 171-190.
- Coseriu, E. (1990): El español de América y la unidad del idioma. In: Facultad de Filología, Universidad de Sevilla (ed.). *I Simposio de Filología Iberoamericana*. Sevilla, 26 al 30 de marzo de 1990. Zaragoza. Libros Pórtico, 43-75.
- DA = Real Academia Española / Asociación de las Academias de la Lengua Española (2010): *Diccionario de americanismos*. Lima. Santillana.
- Del Valle, J. (2013): Linguistic Emancipation and the Academies of the Spanish Language in the Twentieth Century: the 1951 Turning Point. In: José del Valle

- (ed.). *A Political History of Spanish. The Making of a Language*. Cambridge. Cambridge University Press, 229-245.
- Del Valle, J. / Arnoux, E. (2010): Las representaciones ideológicas del lenguaje: discurso glotopolítico y panhispanismo. In: *Spanish in Context*, 7 / 1, 1-24.
- DEJ = Real Academia Española / Consejo General del Poder Judicial (2016): *Diccionario del español jurídico*. Madrid. Espasa.
- DEM = Lara, L. F. (2010): *Diccionario del español de México*. México D.F. El Colegio de México.
- DLE = Real Academia Española (2014): *Diccionario de la lengua española*. Madrid. Espasa.
- DPEJ = Real Academia Española / Consejo General del Poder Judicial (2017): *Diccionario panhispánico del español jurídico*. Madrid. Santillana.
- Eckkrammer, E. (2021): Textos y medios de comunicación. In: Óscar Loureda / Angela Schrott (eds.). *Manual de lingüística del hablar*. Berlin; Boston. De Gruyter, 307-324.
- Ezcurra, Á. (2020): Léxico general y pluricentrismo: aproximación a los americanismos en el diccionario académico». In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 61-80.
- Fajardo, A. (2021): La norma en la lexicografía del español: conflicto, contraste y consenso. In: *Revista internacional de lingüística Iberoamericana*, 19, 17-29.
- Fundación del Español Urgente (2012): *Escribir en Internet: guía para los nuevos medios y las redes sociales*. Barcelona. Galaxia Gutenberg.
- Fundación del Español Urgente (2016): *Manual del español urgente*. Barcelona. Debate.
- Garatea Grau, C. (2008): Unidad y diversidad en el español de América: en torno a variedades y normas lingüísticas. In: Dermeval Da Hora / Rubens Marques (eds.). *Política Lingüística na América Latina*. João Pessoa. Idea. 149-170.
- Garatea Grau, C. (2020): La realidad social del pluricentrismo. En torno a léxico, diccionarios y medios de comunicación. In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 81-94.
- García de la Concha, V. (2014): *La Real Academia Española: Vida e historia*. Barcelona. Espasa.
- Gómez Font, A. (2012): El español global en la prensa del siglo XXI. In: Franz Lebsanft, Wiltrud Mihatsch, / Claudia Polzin-Haumann (eds.). *El español, ¿desde las variedades a la lengua pluricéntrica?.* Madrid; Frankfurt a. M. Iberoamericana / Vervuert, 19-26.

- Greußlich, S. (2015): El pluricentrismo de la cultura lingüística hispánica: política lingüística, los estándares regionales y la cuestión de su codificación. In: *Lexis*, 39 / 1, 57-99.
- Greußlich, S. (2020): Normative Dictionaries (Spanish). In: Franz Lebsanft / Felix Tacke (eds.). *Manual of Standardization in the Romance Languages*. Berlin; Boston. De Gruyter, 605-627.
- Greußlich, S. (forthcoming): Procesos de estandarización y prescriptivismo en la historia del español. In: Steven N. Dworkin, Gloria Clavería Nadal, / Álvaro Octavio de Toledo y Huerta (eds.). *Routledge Handbook of Spanish Historical Linguistics*. London: Routledge.
- Greußlich, S. / Lebsanft, F. (2020): Introducción: Pluricentrismo, estándares regionales, normas implícitas y medios de comunicación masiva. In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 11-36.
- Hesselbach, R. (2021): El español en América ante los procesos de la globalización. In: Eva-Martha Eckkrammer (ed.). *El español en América*. Berlin ; Boston. De Gruyter, 207-224.
- Koch, P. / Oesterreicher, W. (2007): *Lengua hablada en la Romania: español, francés, italiano*. Madrid. Gredos.
- Kühnhardt, L / Mayer, T. (2017): *Bonner Enzyklopädie der Globalität*. 2 Bde. Wiesbaden. Springer.
- Lara, L. F. (2004): *Lengua histórica y normatividad*. México D.F. El Colegio de México.
- Lara, Luis F. (2012): Reseña - ASALE: Diccionario de americanismos (2010). In: *Panace*, XIII / 36, 352-355. Online: <<http://tremedica.org/panacea.html>>
- Lauria, D. (2015): La lengua diccionarizada. Notas sobre la producción lexicográfica del español de la Argentina. In: *El toldo de Astier*, 6, 69-77.
- Lauria, D. (2017a): Avances en el estudio de los instrumentos lingüísticos actuales de la lengua española: los dispositivos normativos híbridos y express. In: *Circula*, 6, 90-113.
- Lauria, D. (2017b): La política lexicográfica actual de las academias de la lengua española: el caso del Diccionario de americanismos (ASALE, 2010). In: *Lexis*, 41 / 2, 263-310.
- Lauria, D. (2019): La institucionalización de la política lingüística panhispánica hoy - Tensiones por la “marca España”. In: *Glottopol*, 32, 209-250.
- Lauria, D. (2021): Discursive Practices Control in Spanish Language. In: *Journal of the Sociology of Language*, 267-268, 143-152.

- Lauria, D. / López García, M. (2009): Instrumentos lingüísticos académicos y norma estándar del español: la nueva política lingüística panhispánica. In: *Lexis*, 33 / 1, 49-89.
- Lebsanft, F. (2013): Spanische Sprachpflege. In: Sandra Herling / Carolin Patzelt (eds.). *Weltsprache Spanisch. Variation, Soziolinguistik und geographische Verbreitung des Spanischen. Handbuch für das Studium der Hispanistik*. Stuttgart. Ibidem, 57-67.
- Lebsanft, F. (2020a): El español total de las Academias y la lengua de El País: ¿un español global para el “periódico global”? In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V / R unipress, 479-500.
- Lebsanft, F. (2020b): Dictionaries of Language Difficulties. In: Franz Lebsanft / Felix Tacke (eds.). *Manual of Standardization in the Romance Languages*. Berlin, Boston. De Gruyter, 629-650.
- Lebsanft, F. / Tacke, F. (2020): Romance Standardology: Roots and Traditions. In: Franz Lebsanft / Felix Tacke (eds.). *Manual of Standardization in the Romance Languages*. Berlin; Boston. De Gruyter, 3-62.
- LELE = Real Academia Española / Asociación de las Academias de la Lengua Española (2018): *Libro de estilo de la lengua española según la norma panhispánica*. Barcelona. Espasa.
- Lipski, J. (2008): Homeless in Post-Modern Linguistics? (Re / Dis)placing Hispanic Dialectology. In: *Studies in Hispanic and Lusophone Linguistics*, 1 / 1, 211-221.
- Lope Blanch, J. M. (1986): *El estudio del español hablado culto: historia de un proyecto*. México: Univ. Nacional Autónoma de México.
- Méndez García de Paredes, E. (2014): Tensiones entre normatividad lingüística y descripción gramatical en la reciente producción gramatical de la Real Academia Española. In: *Romanistisches Jahrbuch*, 64, 248-285.
- Mendoza Quiroga, J. (2015): El castellano de Bolivia. In: Mily Crevels / Pieter Muysken (eds.). *Lenguas de Bolivia, vol. 4: Temas nacionales*. La Paz. Plural Editores, 21-54.
- Moreno Cabrera, J. C. (2015): *Los dominios del español: guía del imperialismo lingüístico panhispánico*. Madrid. Síntesis.
- Moreno Fernández, F. (1993): *La división dialectal del español de América*. Alcalá. Universidad de Alcalá.
- Moreno Fernández, F. (2012): La dimensión social de la gramática. A propósito de la Nueva gramática básica de la lengua española. In: Tomás Jiménez Juliá et al. (eds.). *Cum corde et in nova grammatica*. Estudios ofrecidos a Guillermo

- Rojo. Santiago de Compostela. Universidade de Santiago de Compostela, 605-615.
- Moreno Fernández, F. (2016): La búsqueda de un español global. In: RAE / ASALE (eds.). VII Congreso Internacional de la Lengua Española. Online: <https://congresosdelalengua.es/puerto-rico/paneles-ponencias/espanol-mundo/moreno-fancisco.htm>
- Morgenthaler García, L. (2008): Identidad y pluricentrismo lingüístico. Hablantes canarios frente a la estandarización. Frankfurt am Main. Vervuert.
- Mucci, C. (2019): El español y la sociedad digital. In: RAE / ASALE (eds.). VIII Congreso Internacional de la Lengua Española. Online: <https://congresosdelalengua.es/cordoba/mesas-redondas/default.htm>
- Muhr, R. (2012): Linguistic Dominance and Non-dominance in Pluricentric Languages. A Typology. In: Rudolf Muhr et al. (eds.). Non-dominant Varieties of Pluricentric Languages. Getting the Picture. In Memory of Professor Michael Clyne. Wien. Peter Lang, 23-48.
- Narbona, A. (2001): Movimientos centrífugos y centrípetos en la(s) norma(s) del español". In: RAE / ASALE (eds.). II Congreso Internacional de la Lengua Española. Online: <https://congresosdelalengua.es/valladolid/paneles-ponencias/unidad-diversidad/narbona-a.htm>.
- NGLE = Real Academia Española / Asociación de las Academias de la Lengua Española (2009-2011): Nueva gramática de la lengua española. 3 vols. Madrid. Espasa Libros.
- Peter, Benjamin (2020): L'andalú – Sprache, Dialekt oder lokale Mundart?. Berlin, Boston. De Gruyter.
- Pfänder, S., Alcón, D., / Palacios, A. (2020): Pluricentrismo digital. Usos y normas emergentes del castellano en la red. In: Sebastian Greußlich / Franz Lebsanft (eds.). El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva. Göttingen. V / R unipress, 407-430.
- Plager, Federico (2008): Diccionario integral del español de la Argentina. Buenos Aires. Voz activa.
- Pöll, B. (2021): Spanish Today. Pluricentricity and Codification. In: Danae Perez et al (eds.). English and Spanish: World Languages in Interaction. Cambridge: Cambridge University Press, 163-183.
- Prieto de los Mozos, E. / Amorós, C. (2018): El grado de pluricentrismo de la lengua española, In: Language Problems and Language Planning, 41 / 3, 254-264.
- Sinner, C. (2017): Language Change through Medial Communication. In: Kristina Bedijs / Christiane Maaß (eds.). Manual of Romance Languages in the Media. Berlin, Boston. De Gruyter, 381-410.

- Tacke, F. (2011): Plurizentrik und normativer Diskurs in der Nueva gramática de la lengua española. In: Romanische Forschungen, 123 / 2, 145-166.
- Tacke, F. (2015): Substituyendo el 'frac' o el 'smoking' por la democrática chaqueta – Zur Kultivierung der Aussprache im Spanischen. In: Michael Bernsen / Elmar Eggert / Angela Schrott (eds.), Historische Sprachwissenschaft als philologische Kulturwissenschaft in der Romanistik. Festschrift für Franz Lebsanft zum 60. Geburtstag. Göttingen. V / R unipress, 241-255.
- Zimmermann, K. (2021): Taxonomía de las variedades hispanoamericanas. In: Eva-Martha Eckkrammer (ed.): El español en América. Berlin, Boston. De Gruyter, 63-84.

Benjamin MEISNITZER & Dennis SCHMECHEL

(Leipzig University, Germany)

benjamin.meisnitzer@uni-leipzig.de / d.schmechel1696@gmail.com

Pluricentricity and the Varieties of Spanish in Central America

Abstract

The aim of this study is to discuss whether or not there is a Central American standard variety of Spanish – or several – based on the analysis of a feature catalogue for Central American phonetics and morphosyntax that was compiled and investigated by Quesada Pacheco (2010 & 2013). The inclusion of the varieties of Central American Spanish is even more important as they have, so far, essentially been represented as a grey area in the discussion on pluricentricity. Perception and linguistic features should be brought into harmony.

1. Introduction – The issue of Central American Spanish

Pluricentricity, according to Clyne (1992: 4), is defined as one historical language with several interacting centres. Studies on Spanish in Latin America usually take a strongly Eurocentric view of national standards (Bierbach 2000) or regional standards (Oesterreicher 2001). The regional standards according to Oesterreicher (2001) include Andean, Mexican, and La Plata Spanish, whilst large areas are left out or form a grey zone. For example, in the regional standards, Central America is not usually considered independently, but instead is included in Mexican Spanish. This unfortunate situation is due to a lack of empirical data. Quesada Pacheco (2010 & 2013) provides more extensive empirical material that is available for the first time, which we will examine before asking about new emerging non-dominant varieties of Spanish. A distinct linguistic variety certainly cannot be defined exclusively by linguistic features but must be able to correlate it with the speaker's attitude and the perception of the speaker's community.

The question of the standard variety or varieties in Central America is anything but trivial, since Spanish has very different imprints due to language contact, and the individual regions were relatively isolated from each other for a long time. Even the status of Spanish is not uniform. For instance, in Belize,

Spanish is generally an L2 at best. There is a particular question about the classification of Panama, because historically, when Central and South America were conquered, Panama was the region where the seafarers went ashore, coming from the Caribbean—Peru, for example, was conquered via Panama. The ships arrived in Panama, were dismantled, then the parts were brought to the Pacific coast by slaves, rebuilt and South America was conquered originally from the Pacific coast. Thus, there is some evidence to suggest that Spanish in Panama tends to form a unit with Caribbean Spanish from a historical point of view and the linguistics of the resulting varieties there (cf. Quesada-Pacheco 2016: 188). Beside Panamanian Spanish, it is also important to distinguish Belizean Spanish, because English is the country's official language and Spanish is mainly an L2. Nevertheless, in the last few decades, a wave of migrants mainly from northern Central America (Guatemala, El Salvador, and Honduras) has reinforced the presence of Spanish, which is nowadays spoken by about 52% of the population (cf. Cardona 2010b). This division into three main varieties is confirmed by Quesada-Pacheco (2016: 187-188). In order to clarify the question of the status of Spanish in Central America, we will first briefly define the concept of standard variety, then present the linguistic material used and followed by the method. In a further section, we will discuss the results of the investigation of linguistic features regarding the question of standard varieties, and finally we will summarize our results.

2. Pluricentricity – Challenges of a theoretical concept

A standard variety must have a recognisable normative centre (usually a city) that serves as a linguistic model; it must serve as a reference in education and the mass media and unite the speakers of the language community in the sense of identity formation, as Bierbach (2000: 144-147) notes. Also, when the standard is fully developed, normative instruments must be available, such as grammars, dictionaries, and pronunciation books, because these serve as a reference and are an important indication to the existence of a collective consciousness of (linguistic) identity and distance from other varieties. Nevertheless, it would be inappropriate to reduce the variety of a historical language to these parameters, because the will of the speakers and the language attitude are also decisive. Moreover, it should not be forgotten that pluricentricity, as Pöll (2012: 40) points out, is a concept with several facets or a multi-layered concept, because in addition to the regional, national, and large-area varieties, there also exists a neutral variety of the mass media in the case of Spanish that

is used throughout Latin America, the so-called *español neutro* (Meisnitzer 2017). In addition, international Spanish could possibly be added, although the status of the last two varieties, which are ultimately artificial to a certain extent, is highly controversial, and they are to some extent parallel to the pluricentricity discussion. Neutral Spanish in particular levels out the variety diversity rather than helping its consolidation, unlike in the case of Brazilian Portuguese, as Arden (2015) states.

The role of speakers and their perception is central (Krefeld/ Pustka 2010), since pluricentricity is based on and determined by mental representations of language variation in the user's minds. However, in order to be able to speak of a (consolidated) standard variety, the existence of a bundle of linguistic features that are to be considered neutral from the perspective of diasystematic marking is necessary, which must be collected descriptively (Oesterreicher 2001: 308).

In the case of Central America, we are dealing with norms that are not strongly codified and that show divergences and convergences with neighbouring varieties. In view of the high variation within individual varieties, some of which are national in character, and some of which are more regional due to historical, societal and socio-economic factors, it is important to note that lectal varieties are prototype categories, "in the sense that some varieties are perceived more prototypical or representative than others and the boundaries between adjacent varieties are not clear-cut" (Silva 2016: 15). A pluricentric language therefore exhibits a network of contiguous lects with several interacting prototypical centres.

In addition, prototype theory can help to understand that all languages are pluricentric to at least some degree, to the extent that they exhibit internal dialectal variation and differing local norms. (Silva 2016: 15)

This assumption developed by Silva (2016) is quite relevant for the discussion about the boundaries between a dialect and a regional standard variety, especially if we are discussing a new emerging non-dominant variety, because according to this definition, the field of tension between dialects as syntopic areal units and national standard varieties or large-areal varieties are conceived as a continuum, so to speak. Their status depends mainly on the mental representation of language variation on the speakers' minds, with a dialect being relational to a standard variety, which serves as reference, and not being used in mass media, nor school, nor formal discourse traditions be-

longing to the field of communicative distance according to Koch/Oesterreicher (1985). For Central America with several Spanish speaking countries, we will begin by attempting to define a large-areal variety or if justified several varieties.

3. Materials

In this paper, we conduct a preliminary analysis of the morphosyntactic and phonetic level of the Central American varieties of Spanish. We try to detect a possible standard variety of Central American Spanish in terms of linguistic features, since its annexation to Mexican Spanish or its omission altogether seems inappropriate in the context of Spanish pluricentricity. The source selected is the extensive investigations of the phonetics and morphosyntax of Central American Spanish that were edited by Quesada Pacheco (2010 and 2013, respectively) as a feature catalogue. Quesada Pacheco himself divides the varieties of Central America into three varieties, based on two linguistic features:

1. No aspiration, use of *vos* – Guatemala and Costa Rica.
2. Aspiration of /s/, use of *vos* – El Salvador, Honduras, Nicaragua
3. Aspiration of /s/, use of *tú* – Panama



Figure 1 – Possible areas of the varieties of Central American Spanish based on two linguistic features (Quesada Pacheco 2016: 194).

We want to expand the number of features taken into consideration, as Quesada Pacheco (2016: 194) himself urges to. The investigations by Quesada

Pacheco proved to be of high value as one general questionnaire was applied to extract a set of morphosyntactic (Quesada Pacheco 2013: 379-382) and phonetic features (ibid. 2010: 211f.) of those countries located in Central America, that is, Belize, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica, and Panama. This allowed for an adequate comparison of the Spanish spoken in Central American countries. However, there are a few limitations with Quesada Pacheco's investigations that ultimately rendered a few aspects too insufficient or unclear for our analysis. Firstly, some authors chose to provide only terminology such as "majority" or "a few informants" instead of precise percentages or numbers of informants. For such instances, we chose reasonably appropriate estimations of the possible percentages of the realizations of a feature. Secondly, although the aim of the investigations was to be as congruent as possible, some authors appeared to have chosen to include or exclude a few aspects of their own accord. Therefore, a few studied features had to be left out in our analysis because there was no overarching theme fit for comparison. Thirdly, for some phonetic aspects, the terminology applied seemed imprecise. In some instances, the fricative [h] was denominated *velar*. Since this phone is in fact glottal, it would be more suitable to either denominate it as such or use the phone [x] instead. For these cases, we had to rely on which realization sounded more logical in the respective context, being either [h] or [x]. Despite these three issues, both works were still fairly fit for a preliminary analysis of Central American Spanish.

While the phonetic investigation was dedicated to the pronunciation of vowels and consonants in certain environments (see Quesada Pacheco 2010: 211f.), the morphosyntactic analysis investigated the gender (e.g., *el sartén* vs. *la sartén*), number (e.g., *pies* vs. *pieses*), forms of address (e.g., whether the husband uses *vos*, *usted* or *tú* to address his wife and vice versa), enclitic pronouns (e.g., *siéntense* vs. *siéntesen* vs. *siéntensen*), derivation (e.g., diminutives *manita* vs. *manito*), verbal variation (e.g., *decimos* vs. *dicimos*), mood and tense (e.g., *fui* vs. *he ido*), and a few phraseological aspects (see Quesada Pacheco 2013: 379ff.).

4. Method

We needed to find a way to determine whether a Central American variety fits into a possible Central American standard or not. We first collected the data of the realizations of all features in each of the seven countries. To illustrate this, the final unstressed /a/ in *masa/casa* may have eight different realizations all across Central America, i.e., *plain*, *short*, *elided*, *polyphonous*, *closed*, *un-*

voiced, opened, and weakened. In Honduras, for instance, 96.4% of all informants realized the /a/ plainly and another 3.6% opted for an opened realization.

After we collected all the data, we then needed a guideline value that would represent the norm and that could be used to identify how closely the realizations of the features in the respective countries related to that norm.

For this we used the arithmetic mean, calculated for each realization, some of which are exemplified in section 5. For the example of the realization of the final unstressed /a/ in *masa/casa*, the mean yielded a value of 87.8% for all plain realizations and a mean of 0.5% for all opened realizations across all seven countries.

In the following step, we calculated the absolute difference (AD) between the percentages of the realizations for each country (CR) and the general mean (GM) $|GM - CR| = AD$. The smaller this value was, the closer the realization of a feature in one country was to the general mean. For Honduras, this means that its deviation from the mean of all plain realizations equals 8.6% and 3.1% for all opened realizations, $87.8\% - 96.4\% = 8.6\%$ and $0.5\% - 3.6\% = 3.1\%$.

Next, the deviation values needed to be evaluated. The number of deviations that each country presents is dependent on where the threshold of the percentages of the absolute difference is set. This means that relatively more deviations from the norm are to be expected the lower this threshold is, since the values of absolute difference ranged from 0% to 79.4% for the phonetic level.

With a threshold set at >5%, Guatemala, for example, showed 106 deviations from the 329 phonetic realizations analyzed (see table 1). The higher this threshold is set, the lower the number of deviations becomes. At >50%, Guatemala only has 12 deviations. Since there was a considerable number of realizations that only reached a mean of lower than 10%, their absolute difference cannot be expected to be any higher than that value. With higher means, the absolute difference can be higher, too. Taking this into account, there is a considerable number of deviations showing at a lower threshold that are most likely of low significance. The higher the threshold is, the more significant the deviations become since their absolute difference can be higher accordingly.

However, it is difficult to determine the threshold from which onwards the deviations become significant. Thus, instead of determining an appropriate threshold, we compared several thresholds, ranging from >5% to >50% in steps of 5%, and observed how rapidly the number of deviations decreased as the threshold increased. This analysis already yielded valuable results, but it ig-

nored the realizations that were specific to only one of the seven countries in Central America. Therefore, we counted the number of realizations for all countries that were specific to only one country. Afterwards, we excluded all country-specific realizations that were used by less than 20% of the informants of the respective country. This would guarantee that random realizations or those made by a relative minority did not influence a possible standard variety negatively.

Going back to the example of the realization of the final unstressed /a/ in *masa/casa*, El Salvador, for instance, is the only country in Central America that realizes the phoneme as either *closed* (3.8%) or *unvoiced* (1.2%), but their percentages are strongly overshadowed by the *plain realization* (93.8%). Hence, the former two are not of any value for El Salvador's uniqueness in terms of the realization of the final unstressed /a/ in *masa/casa*. In summary, two measurements were used to determine whether there might be a possible standard variety of Spanish in Central America.

These measurements are (1) the number of deviations from the Central American mean compared across a range of thresholds of the absolute difference and (2) number of (significant) realizations that are specific to only one country. Both measurements were applied to both the phonetic and the morphosyntactic level of each country in Central America.

5. Results

For both the morphosyntactic and the phonetic level, certain numbers of deviation were to be expected, especially with the lower thresholds, since for instance, deviations of >5% are far more easily reached than deviations of >30%. The peak values, that is, the highest values, of the deviations for each threshold are outlined by dashed lines (see table 1).

It can be observed that Panama peaks in the lower thresholds (>5% to >10%; 126 and 86 deviations, respectively, from the 329 mean values). However, the lower thresholds can be considered as less significant in terms of deviation from a possible standard, as discussed above. In the middle range (>15% to >40%), Honduras clearly peaks in every threshold. These peak values decrease fairly slowly with the increasing threshold compared to the other countries and is only overtaken by Guatemala in the range of >45% to >50%. Additionally, it should be noted that although Panama peaks in the low range thresholds, the values are not higher by much compared to the remaining six countries. On the other hand, Honduras peaks quite considerably in the middle range, closely fol-

lowed by Guatemala. Since the middle range thresholds become increasingly more significant, the deviations that Honduras and Guatemala present are hence to be interpreted as quite significant. The remaining countries show a trend with less deviations from the mean throughout the entire threshold range.

Phonetic Level: Number of Deviations from the Mean							
Threshold of Absolute Difference	Number of Deviations from the Mean per Threshold						
	Belize	Guatemala	El Salvador	Honduras	Nicaragua	Costa Rica	Panama
>5%	113	106	111	115	114	99	126
>10%	82	73	74	81	79	59	86
>15%	54	54	46	65	59	43	55
>20%	43	43	32	55	43	31	39
>25%	30	33	25	48	26	22	25
>30%	19	25	16	40	20	16	16
>35%	12	22	9	34	15	13	8
>40%	7	20	3	25	11	9	5
>45%	4	17	2	11	10	4	0
>50%	3	12	1	9	5	1	0
Total number of mean values (excluding those with insufficient data) = 329							

Table 1: Level of phonetic deviation between the different Spanish Varieties in Central America.

Following the deviations from the mean, the country-specific realizations of the phonetic feature catalogue need to be analyzed. Table 1 demonstrates that Nicaragua and Costa Rica had 14 and 2 realizations, respectively, that were unique. However, neither of these realizations were significant enough to be taken into consideration, i.e., used by more than 20% of the informants. Of the country-specific realizations in Guatemala (9) and El Salvador (24), only one was significant for each country. In Guatemala, 27.2% realized the /s/ before /b/ in *las venas* as the approximant [s]. In El Salvador, 43.8% realized /x/ in *José/abajo* as a lenis [ʰ]. The unique realizations of Belize, Honduras, and Panama proved to be more numerous. Of the unique realizations in Belize (7), three were significant. The prenuclear trill /r/ was realized as the retroflex approximant [ɺ] in intervocalic position, *perro/carreta* (33.3%), and in word-initial position, *rojo* (23%). Although not unique to Belize, this realization is also rather prominent for the postnuclear flap /r/ in contexts such as *viernes* (32.2%), *perla/Carlos* (38.4%), *cárcel/almuerzo* (25.6%), and *enredado* (37.8%). The phoneme /d/ after /r/ as in *muerde(n)* is uniquely realized as the retroflex [d] by 30.7%. Of Honduras' unique realizations (21) four are significant, namely the realization of /x/ as [xʰ] in *cajita* (85.7%), *gente/jeta/cajeta* (~75%), *junio* (38.1%), and *caja/jabón/paja* (50%). Although not unique to Honduras, this realization

can also be found in *José/abajo* (44.1%). Panama (16) has the most significant unique realizations, namely five. The postnuclear /s/ is realized as [h] in *isla* (22.4%), *buenos días* (28.9%), and *las venas* (26.3%). The postnuclear /s/ is also realized as [z] before /t/ as in the United States (51.3%). Lastly, 28.4% realized /b/ as [β] in *jaiba/muy bueno/ceiba*. In analogy to this, two more realizations can be added to that list, though not unique to Panama, the realization of postnuclear /s/ as [h] before /g/ as in *los gatos* (41.1%) and the realization of /b/ as [β] in *curva* (22.7%).

In the morphosyntactic analysis of the deviations from the mean (see table 2), only Panama peaks quite significantly, especially when regarding the higher thresholds of the absolute difference. Again, the peak values for each threshold are outlined by dashed lines. Costa Rica and Nicaragua also peak at >10% and >15%, respectively, although neither peak value is quite as significant when compared to the values that the other countries reached in the respective threshold. Furthermore, the lower thresholds are not to be given that much weight in terms of significance.

Morphosyntactic Level: Number of Deviations from the Mean							
Threshold of Absolute Difference	Number of Deviations from the Mean per Threshold						
	Belize	Guatemala	El Salvador	Honduras	Nicaragua	Costa Rica	Panama
>5%	168	167	163	143	161	174	184
>10%	112	94	108	94	118	119	110
>15%	77	49	75	57	79	77	74
>20%	49	36	53	37	53	56	62
>25%	33	25	36	30	38	36	48
>30%	23	15	29	25	28	25	43
>35%	18	8	19	21	19	20	33
>40%	10	7	12	15	15	16	29
>45%	5	5	8	10	10	7	23
>50%	1	4	6	9	8	5	18
Total number of mean values (excluding those with insufficient data) = 461							

Table 2: Level of morphosyntactic deviation between the different Spanish Varieties in Central America

The number of morphosyntactic realizations specific to only one country is not worth mentioning. Those exceeding 20% are restricted to only one or two per country. Panama is once more an exception as it showed three unique realizations, 44.4% address someone by using *tú* when they are furious, 20.8% vary between *tú* and *usted* in the same situation, and 22.2% use *enante* to express 'hace un momento'. This value does not exceed those of the other countries by much nor do specific realizations appear particularly noteworthy. This may furthermore support the conclusion that the morphosyntax of Panama Spanish is different from the rest of Central America.

6. Discussion

To sum up, Central American Spanish was analyzed for phonetics and morphosyntax. On the phonetic level, approximately three countries stand out in that they are different from what could be considered a Central American norm, namely Guatemala, Honduras, and Panama. The former two show reasonable differences from that norm in terms of their deviations from it, but also in terms of their country-specific realizations. Panama does not deviate significantly from the norm by much. However, the number of its country-specific realizations is quite high. On the morpho-syntactic level, only Panama proves to be different in both its deviations from the norm and its country-specific realizations, although the latter seems to be less significant, given that the other countries are in a lower but similar range.

Based on the data, we claim that two countries can already be excluded from a possible Central American Standard Spanish (see figure 2). First, Panama is different in both phonetics and morphosyntax, following our analysis. Panama Spanish is also often ascribed to Caribbean Spanish, which ties in with our findings. Second, we exclude Belize although it did not peak once and seemed to show a generally close relationship to the norm. Nevertheless, the Belizean political situation needs to be taken into account. English is the official language in Belize and the English rhoticity is a notorious aspect in Belizean Spanish that may render the perception of Belizean Spanish as different. Furthermore, there may be unclear borders regarding Guatemala and Honduras. Both countries peak considerably on the phonetic level, but not on the morphosyntactic level. The extent of their inclusion requires more investigation. This would also be true for Nicaragua since it reaches more or less raised deviation values on both the morphosyntactic and the phonetic level. El Salvador and Costa Rica seem almost ideal, since their deviation values never peak, nor do they have a considerable number of country-specific realizations. Nevertheless, it would be interesting to investigate Costa Rica further because like Belize, it shows fairly high percentages for the realization of the postnuclear /r/ as [ɹ], between 19.4% and 67.3%. This did not particularly influence the statistics since a considerable amount of data for this feature was missing for Nicaragua, which made a reasonable comparison impossible.

Returning to the initial question of our study, we can make the areal subdivision shown in figure (2), in which a large-area non-dominant variety can be identified, although it does not include Panama and Belize (cf. Quesada

Pacheco 2016: 194), both of which are also geographically part of Central America.

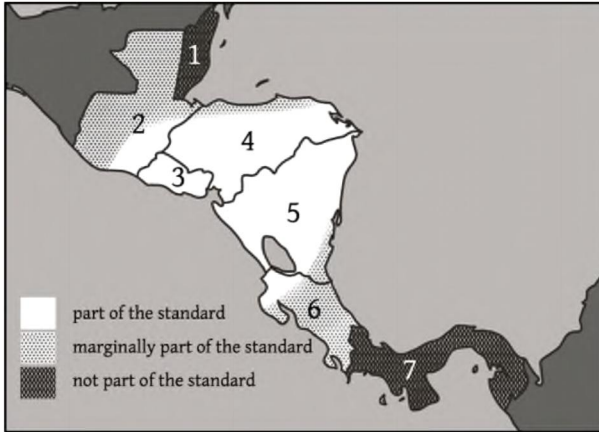


Figure 2 – Central American Emerging Standard Spanish.

Three types of areas are used to outline which countries may be included in a Central American standard variety and which may not (see figure 2). Note that the numbers in figure 2 help identify the respective country. The dark dotted area marks Belize (1), in the north, and Panama (7), in south. These two countries are accordingly excluded from the standard. Guatemala (2) and Costa Rica (6) have some features that may imply that the border regions of these two countries may blend with the core region of the Central American standard. Therefore, they are marked as being both part of the standard and marginally part of the standard, the white area, and the light grey dotted area, respectively. The extent of their membership of the Central American standard is a matter which requires more investigation. The aforementioned core region comprises of El Salvador (3), Honduras (4), and Nicaragua (5). Although these three countries appear clearer in terms of their membership than Guatemala (2) and Costa Rica (6), there is reason to further investigate their features. As a general note, it needs to be made clear that the area of the standard does not necessarily coincide with national borders nor are those areas shown in figure 2 definite – they are merely a first sketch that requires further investigation. One remaining important step would be to make sociolinguistic interviews with speakers from the areas considered in this study in order to figure out if the diatopic variation is relevant to the respective national identity and to

which degree the involved lects are seen as unifiers or dividers by the speakers (Muhr 2012: 29-30).

Pluricentric languages are both unifiers and dividers of people. They unify people through the use of the language and separate them through the development of national norms and indices and linguistic variables with which the speakers identify. They mark group boundaries [...] indicating who belongs and who does not. (Clyne 1992: 1)

Recent studies have shown that “the majority of Central Americans tends to say that they speak Spanish and call their mother tongue “Spanish” or “Castilian [...]” (Quesada Pacheco 2016: 194). This implies that there is a very strong sense of unity in each of the countries. Quesada Pacheco (2016: 195) could also identify a nucleus formed by El Salvador, Honduras, and Nicaragua, where almost all informants agreed on the similarities of their linguistic features, while Costa Rica feels close to Panama. It is also interesting that they apparently feel a particular admiration for their own national variety in Central America and for peninsular Spanish, while no informants expressed admiration for varieties of the neighbouring countries (Quesada Pacheco 2016: 196). The study on language attitudes organized by Chiquito & Quesada Pacheco (2014) confirms our historically and linguistically motivated inclusion of Panamanian Spanish in Caribbean Spanish, since Venezuelans “feel that their variety is close to that of Panamanians” (Quesada Pacheco 2016: 198). It is remarkable that “none of the Central American Spanish features of national varieties are unknown to standard Spanish” (Quesada Pacheco 2016: 200), except *voseo*. However, the core features are often unknown to the Spanish speakers from other varieties as Quesada Pacheco (2016: 200) remarks, which leads him to conclude that “this part of the New World is a kind of invisible section of the Spanish-speaking countries” (Quesada Pacheco 2016: 199).

7. Conclusion

Based on the data we gathered from Quesada Pacheco (2010 & 2013) and our analysis of it, there may very well be a Central American standard variety of Spanish. Two countries are, however, to be excluded from this standard, namely Belize and Panama. With regard to the two countries that we have not included in the Central American standard, it should be noted that Belize can only be attributed to the Hispanophone world to a limited extent, especially since Spanish is generally only L2 for the speakers, and Panama clearly exhibits linguistic features that strongly distinguish it from the other varieties consid-

ered and that place it in the linguistic area of Caribbean Spanish with regard to large-area standards. For the remaining countries it would still be interesting to do a more thorough and precise analysis of their phonetics and morphosyntax to find out whether their national borders do or do not coincide with a possible Central American standard. This would require a more reliable data set that may even provide smaller scale data, i.e., of several regions that were selected beforehand in each of the countries. To make this even more accurate, it may be of interest to also include data of bordering standard varieties. As a final step, a dialectometric analysis similar to the one conducted by Szmrecsanyi (2015) may further clarify whether a Central American Standard would indeed be a reasonable assumption and where its border would be located. In our main division, considering an extended number of variables, the division remains the same in general terms as the one proposed by Quesada Pacheco (2016).

References

- Arden, Mathias (2015): *Inszenierte und elaborierte Mündlichkeit bei TV Globo. Zur soziolinguistischen Modellierung morphosyntaktischer Variablen des brasilianischen Portugiesisch*. Frankfurt a. Main: Lang.
- Arden, Mathias/Meisnitzer, Benjamin (2013): *Plurizentrik und massenmediale Normen: Der Fall des Portugiesischen*. In: Aurelia Merlan & Jürgen Schmidt-Radefeldt (eds.): *Das Portugiesische als Diasystem innerhalb und außerhalb des lusophonen Raums*. Berlin: Lang. p. 19-52.
- Bierbach, Mechthild (2000): *Spanisch – eine plurizentrische Sprache? Zum Problem von norma culta und Varietät in der hispanophonen Welt*. In: *Vox Romanica*, 59. p. 143-170.
- Cardona, Mauricio (2010): *La fonética del español en Belice*. In: Miguel Ángel Quesada Pacheco, (ed.) (2010): *El español hablado en América Central. Nivel Fonético*. Frankfurt: Vervuert. p. 21-48.
- Chiquito, Ana Beatriz/ Quesada Pacheco, Miguel Ángel (eds.) (2014): *Actitudes lingüísticas de los hispanohablantes hacia el idioma español y sus variantes*. *Bergen Language and Linguistic Studies (BeLLS)*, 5. <https://dx.doi.org/10.15845/bells.v5i0.691>.
- Clyne, Michael (1992): *Pluricentric Languages – Introduction*. In: Michael Clyne (ed.): *Pluricentric Languages. Differing Norms in Different Nations*. Berlin/ New York. De Gruyter, p. 1-9.
- Koch, Peter/Oesterreicher, Wulf (1985): *Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte*. In: *Romanistisches Jahrbuch 36 (1985)*. s. 15-43.

- Krefeld, Thomas/ Pustka, Elissa (2010): Für eine perzeptive Varietätenlinguistik. In: Thomas Krefeld & Elissa Pustka (2010): *Perzeptive Varietätenlinguistik*. Frankfurt am Main, etc. Lang (Kommunikative Räume; 8). p. 9-28.
- Meisnitzer, Benjamin (2017): Español neutro im Fremdsprachenunterricht? Potenzial und Grenzen. In: Eva Leitzke-Ungerer & Claudia Polzin-Haumann (eds.): *Varietäten des Spanischen im Fremdsprachenunterricht. Ihre Rolle in Schule, Hochschule, Lehrerbildung und Sprachenzertifikaten*. Stuttgart. Ibidem (Romanische Sprachen und ihre Didaktik; 61). p. 201-221.
- Muhr, Rudolf (2012): Linguistic dominance and non-dominance in pluricentric languages: A typology. In: Rudolf Muhr (ed.): *Non-Dominant Varieties of Pluricentric Languages. Getting the Picture. In Memory of Michael Clyne*. Frankfurt am Main, etc. Lang. p. 23-62.
- Oesterreicher, Wulf (2001): Plurizentrische Sprachkulturen – der Varietätenraum des Spanischen. In: *Romanistisches Jahrbuch*, 51 (2000). p. 287-318.
- Pöll, Bernhard (2012): Situaciones pluricéntricas en comparación: El español frente a otras lenguas pluricéntricas. In: Franz Lebsanft, Wiltrud Mihatsch, Claudia Polzin-Haumann (eds): *El español ¿desde las variedades a la lengua pluricéntrica?* Frankfurt am Main. Vervuert / Iberoamericana, p. 29-45.
- Silva, Augusto Soares da (2016): The cognitive approach to pluricentric languages and the pluricentricity of Portuguese: What's really new? In: Rudolf Muhr, Eugênia Duarte, Amália Mendes, Carla Amorós Negre & Juan A. Thomas (eds.): *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions*. Frankfurt. Lang. p. 13-34.
- Szmrecsanyi, Benedikt (2015): *Grammatical Variation in British English Dialects. A Study in Corpus-Based Dialectometry*. Cambridge. Cambridge UP.
- Quesada Pacheco, Miguel Ángel (ed.) (2010): *El español hablado en América Central. Nivel fonético*. Madrid/Frankfurt am Main. Iberoamericana/Vervuert.
- Quesada Pacheco, Miguel Ángel (ed.) (2013): *El español hablado en América Central. Nivel morfosintáctico*. Madrid/Frankfurt am Main. Iberoamericana/Vervuert.
- Quesada Pacheco, Miguel Ángel (2016): Non dominant-Varieties of Spanish: The Central American Case. In: Rudolf Muhr, Eugênia Duarte, Amália Mendes, Carla Amorós Negre & Juan A. Thomas (eds.) (2016): *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions*. Frankfurt. Lang. p. 187-206.

Juan THOMAS

(Utica University, Utica, NY, USA)

juantomas329@yahoo.com

What taboo Anglicisms say about U.S. Spanish

Abstract

This study describes fourteen Anglicisms that are homonyms with a taboo word (a word which is normally not used in certain social contexts) in Spanish or that give a taboo meaning from English to another word present in U.S. Spanish. The Academia Norteamericana de la Lengua Española (ANLE: North American Academy of the Spanish Language) recommends against the use of those Anglicisms, along with hundreds of others that are characteristic of U.S. Spanish in the three volumes of *Hablando bien se entiende la gente* (HBSELG 1 2010, 2 2014, 3 2021). The ANLE is one of the 24 corresponding language academies of the Spanish-speaking world affiliated with the Real Academia Española (RAE: Royal Spanish Academy), the norm-setting organization for the Spanish language. This study wishes to shed light on how and why these taboo expressions can enter Spanish, if they truly are unique to the Spanish of the U.S., and how they fit into the debate about differential dictionaries of -isms.

1. Introduction

In a recent article about Spanish in Utica, New York (a city of 60000 inhabitants (12.7% Hispanic (U.S. Census 2019)) about 400 kilometers northwest of New York City, the name of a Hispanic restaurant was analyzed in the context of the linguistic landscape of the city (Thomas 2019). Some Utica Hispanics became quite angry about its original name, "El Carajo Restaurant". Even though the city did not force the owner to change the name, he did change it to "El Barajo Restaurant" in order to not offend his customers. This event contrasts with a similar name, but in Italian, or rather, the Southern Italian dialects spoken in Utica for decades. The name of a bar, "The Stroonz", an English adaptation of the Italian *stronzo*, a vulgar word, did not provoke any public outcry similar to "El Carajo Restaurant". Many Uticans did not know what the term meant, and those who did, questioned why someone would name their business with such a word. This difference in attitudes inspired this present study in order to see how English helps to facilitate obscene or vulgar

vocabulary into Spanish and how and why Spanish would or would not tolerate the integration of such loanwords.

Some definitions will help to establish a theoretical framework for this present study. First off, it is necessary to return to Saussure's description of the linguistic sign which consists of two parts: the signifier and the signified. The signifier (an arbitrary sequence of sounds, letters or gestures) is connected to the signified (the concept or mental image). This study deals with loanwords, that is, words which enter into a receptor language after a process of linguistic adaptation. If the etymon comes from English, the loanword is an Anglicism. Bilingual speakers of the source and receptor languages initiate the process of adaptation and integration. It is likely that a second bilingual speaker will understand the new word without difficulty, but it is certain that a monolingual speaker of the receptor language will not. With time, if the word acquires a certain diffusion and frequency of use among speakers, even monolinguals will know and use the word and it will become part of the standard lexicon.

If two different signifiers are associated with the same signified, (they are interchangeable), then they are synonyms (Chamizo-Domínguez 2008 48). If the phonological shape of an Anglicism matches that of another word that is already lexicalized, misunderstandings most likely will result. This is homonymy. Two words are homonyms if they have the same signifier (phonological shape) but two different signifieds (Chamizo-Domínguez 2008 41). The concept of false friends is similar to homonymy. Homonymy exists within one language only. Two words are false friends if they are in different languages and their phonological shapes match, but their signifieds are partially or totally different (Chamizo-Domínguez 2008 28). False friends share signifiers but not signifieds.

Oftentimes, false cognates are associated with false friends. A cognate is a word that shares the same etymological roots with another word (Chamizo-Fernández 2008 3), regardless of the evolution of their signifieds. All false cognates are false friends but not all false friends are false cognates, because matching shapes are not always due to evolution from the same etymon.

Taboo concepts exist in all societies, although taboo situations vary according to the social, cultural and religious norms of each society. A linguistic taboo is the prohibition of vocabulary associated with a taboo situation- it is a word or an expression that must not be said. Oftentimes it is a

profane or vulgar word; other times it is a word that might be awkward in a certain social setting.

The volume *Hablando bien se entiende la gente* (HBSELG1) (Piña-Rosales, Covarrubias, Segura, Fernández 2010) is a guide to help Spanish-speakers clear up spelling and grammar questions which arise because of the influence of English (Piña-Rosales, Covarrubias, Segura, Fernández 2010 2). The volume was published by the Academia Norteamericana de la Lengua Española (ANLE: North American Academy of the Spanish Language) and was intended especially for audiences in the U.S., Canada, and Puerto Rico. The guide is divided into several sections; the first and largest is dedicated to false friends and the following sections deal with expressions, grammatical constructions, verb conjugations, as well as loanwords from languages other than English. Lynch and Potowski (2014) criticized the book's sociolinguistic foundations and proscriptive intentions. Part of their argument rests on examples that they analyzed from the volume that were either not Anglicisms, were part of the Spanish in varieties spoken outside of the U.S., or present in the official dictionary of the Real Academia Española, the norm-setting organization of the Spanish language.

Owing to the popularity of the first volume, a second volume was asked for by the publishing house and *Hablando bien se entiende la gente 2* (HBSELG2) (Piña-Rosales, Covarrubias, Dumitrescu) was published in 2014. The volume is shorter than the first but includes an extensive section on recommendations and then a shorter one on linguistic and grammar points. The third volume, *Hablando bien se entiende la gente 3* (HBSELG3) (Piña-Rosales, Covarrubias, Fernández, Rodríguez, Prunes), came out in 2021 and consists of two sections "No nos fiemos de los falsos amigos [Let's not trust false friends]¹" and "¡Dígalo bien! [Say it well!]"

The Academia Norteamericana de la Lengua Española (ANLE), founded in 1973, is one of the twenty-four corresponding academies of the Real Academia Española (RAE). The panhispanic linguistic policy of the RAE is known as "*velar por que la lengua española, en su continua adaptación a las necesidades de los hablantes, no quiebre su esencial unidad* [to ensure that the Spanish language, in its constant adaptation to the needs of its speakers, does not break its essential unity]" (RAE 2021). The ANLE collaborates with the RAE and with the Asociación de Academias de la Lengua Española (ASALE). Part of the ANLE's

¹ All translations are done by the author.

mission is: "*Cuidar que, en su constante adaptación a las necesidades particulares de los hablantes, el uso de la variante hispanounidense no afecte la unidad y comprensión del idioma en el ámbito hispánico* [To ensure that, in its constant adaptation to the particular needs of the speakers, the use of the Spanish-American variant does not affect the unity and understanding of the language in the Hispanic sphere]" (ANLE 2021). The ANLE publishes bulletins, scientific journals and since 2008, almost forty books. The three volumes of *Hablando bien se entiende la gente* represent the ANLE's efforts in recommending what should and should not be part of Spanish spoken and written in the U.S.

Although the three volumes of *Hablando bien se entiende la gente* recommend against the use many of the expressions that they describe, the consideration of those expressions strongly suggests that they are part of the Spanish spoken in the U.S. and enjoy a certain level of diffusion and frequency of use among U.S. Spanish-speakers. This study focuses on homonymy and the coincidences between taboo words and Anglicisms, and wishes to shed light on how these taboo expressions enter Spanish and if they are unique to the Spanish of the U.S.

Pluricentric theory deals, in part, with a critique of the standards. Various types of lexicographic works for Spanish have caused fierce debate. Differential dictionaries of '-isms' contrast the lexicon of a specific country with what is supposedly common to all (although it is usually only common to Spain) in order to give account of the connotative, denotative and cultural uses in a country. These dictionaries have been debated and criticized (Amorós Negre 2014, Ávila 2003, 2003-4, Ezcurra 2020; Fajardo Aguirre 2011, Lara 1996, Zimmermann 2003, 2013). Americanisms might not be common to the whole territory of Hispanic America. It can also be a theoretic failing to discuss -isms of a specific country without the necessary data to compare the lexicon from individual countries. The methodology of these dictionaries has been questioned (Greußlich 2020). While the three volumes of HBSELG are not dictionaries of *estadounidismos* (words used in the Spanish of the U.S.), we will see how the treatment of Anglicisms by the authors of HBSELG fits in the ideological debate regarding what should constitute the lexicographic norm for Spanish.

2. Methodology

Anglicisms that are homonyms with words already present in Spanish will be studied. The three volumes of *Hablando bien se entiende la gente* (Piña-

Rosales *et al.* 2010, Piña-Rosales *et al.* 2014, Piña-Rosales *et al.* 2021) will be the source of these words. These words will be contrasted with their entries in other reference works of the Real Academia Española: the main dictionary of the Academy (the *Diccionario de la lengua española* (DLE)), the differential dictionary of Americanisms (the *Diccionario de americanismos* (DAA)) produced by the ASALE, and the *Diccionario de anglicismos del español estadounidense* (DAE) produced by the Instituto Cervantes/ Observatorio de la lengua española y las culturas hispánicas en los Estados Unidos.

3. Results

Almost all of the entries in the three volumes of *Hablando bien se entiende la gente* (Piña-Rosales *et al.* 2010, Piña-Rosales *et al.* 2014, Piña-Rosales *et al.* 2021) deal with the influence of English on the Spanish spoken in the United States. The use of some Anglicisms (just by coincidence) expresses sexual situations which are not normally discussed in certain social contexts. Table 1 summarizes the results. *Introducir* is used in United States Spanish like the English 'to introduce', which in most other varieties of Spanish is expressed by *presentar* (Piña Rosales *et al.* 2010 67; Piña Rosales *et al.* 2021 20). *Introducir* in most varieties of Spanish means 'to penetrate or to make enter'. Although the meaning of *introducir* as the Anglicism is not taboo, its use might be confused with the taboo sense of the word in sentences such as "*quiero introducirselo (presentárselo) a mi buen amigo* [I want to make him penetrate (introduce him to) my good friend]". The *Diccionario de americanismos* (DAA) (2010), however, includes the Anglicism *introducir* and gives its use as *presentar* in the U.S., El Salvador, the Dominican Republic and Ecuador. The *Diccionario de anglicismos del español estadounidense* (DAE) (2018: 66) also includes *introducir* with the meaning *presentar*.

Estar excitado (Piña-Rosales *et al.* 2014 38; Piña Rosales *et al.* 2021 16) is usually used in the United States to mean 'to be excited, enthusiastic', as in the English expression 'to be excited'. However, *estar excitado* means 'to be sexually stimulated/to be horny' in varieties of Spanish outside the United States. Just as with *introducir*, the Anglicism *estar excitado* is not taboo, but it could be easily confused with its other entry, the one with the the taboo meaning, already present in Spanish.

Estoy caliente is an Anglicism for '*tengo calor*' [I am warm/hot] (Piña-Rosales *et al.* 2021 18). In most varieties of Spanish *estoy caliente* is charged with sexual connotations and is closer in meaning to the English "I'm horny".

Interestingly, there are contexts even in English where 'hot' also has sexual connotations.

One frequently hears the expression *tener sexo* in the Spanish of the United States to refer 'to having sexual relations' (Piña Rosales *et al.* 2010 96); it is a calque of the English 'to have sex'. For many Spanish speakers, though, *sexo* means 'sexual organs/ the condition of being male or female', and *tener sexo* is to be male or female, (Piña Rosales *et al.* 2010 96). Both the meanings of the Spanish and the Anglicism can be considered to be taboo. The Fundación del Español Urgente (Fundéu 2019), on the other hand, indicates that *tener sexo* is a valid expression for 'to have sexual activity' because the *Diccionario de la lengua española* (DLE) incorporated a fourth meaning for *sexo*, venereal pleasure, in 2001. Yet, Fundéu recognizes that the origin of that meaning could well be from English.

The word *preservativo* (Piña-Rosales *et al.* 2021 24) is an Anglicism for 'preservative'. The word appears in the *Diccionario de anglicismos del español estadounidense* (DAE): "*una sustancia conservante de alimentos* [a substance that preserves foods]" (Moreno Fernández 79), while in standard Spanish it is a condom and is taboo; the Anglicism is not taboo.

'Embarrassed' and *embarazada* are false friends in varieties of Spanish outside the U.S., but U.S. Spanish speakers have incorporated the former into their Spanish. *Embarazada* means 'pregnant' but in the United States, it is also used for 'embarrassed', from the influence of its English etymon (Piña-Rosales *et al.* 2014 73). As with *estar exitado* and *introducir*, the Anglicism *embarazada* is not taboo, but it can easily be confused with the taboo signified.

Although *preñada* also means *embarazada*, normally it refers to a female animal, not a human being. The authors of *Hablando bien se entiende la gente* (2010 80) feel that the English 'pregnant' has helped to promote the diffusion of the expression of *una mujer preñada* (a pregnant woman), instead of *una mujer embarazada*, although they recognize that it is used in that sense in Spanish-speaking countries other than the U.S. Even though *una mujer preñada* is not incorrect, it is a stronger expression (Martin 1979 231) than *una mujer embarazada*.

'To molest' in English means 'to sexually harass'. The Spanish *molestar* is 'to annoy' without reaching the extreme meaning of the English 'molest'. Because of the influence of English, U.S. Spanish favors a semantic extension of *molestar* to coincide with the English sense. This is one of the unique terms in this study because the Anglicism is taboo, while the Spanish usage is not.

However, if a U.S. Spanish-speaker used *molestar* to mean 'to sexually harass', misunderstandings would result with monolingual Spanish-speakers.

Anglicism	English etymon	Meaning of the Anglicism	Spanish etymon / its meaning	Source
Decency: pregnancy				
<i>embarazada</i>	embarrassed	embarrassed	<i>embarazada</i> / pregnant	HBSELG2 73 HBSELG3 16
<i>preñada</i>	pregnant	a pregnant woman	<i>preñada</i> / pregnant animal	HBSELG1 80; DILE
Decency: sexual relations				
<i>tener sexo</i>	to have sex	to have sexual relations	<i>tener sexo</i> / to be male or female	HBSELG1 96; DILE
<i>estar excitado</i>	to be excited	to be excited	<i>estar excitado</i> / to be sexually stimulated	HBSELG2 38 HBSELG3 16
<i>estoy caliente</i>	I am hot	I am hot	<i>estoy caliente</i> /I am horny	HBSELG3 18
<i>molestar</i>	to molest	to sexually harass	<i>molestar</i> /to bother	HBSELG2 76 HBSELG3 23
<i>introducir</i>	to introduce	to introduce	<i>penetrar</i> /to force penetration	HBSELG1 67; DAA; DAE 66; HBSELG3 20
<i>preservativo</i>	preservative	preservative	<i>preservativo</i> / condom	HBSELG3 24; DAE 79
Decency: bodily functions				
<i>lavatorio</i>	lavatory	bathroom	<i>lavatorio</i> /sink	HBSELG3 21
Death and dying				
<i>casquete</i> / <i>casco</i>	casket	casket	<i>casquete</i> /skullcap; <i>casco</i> /helmet	HBSELG3 12
<i>coroner(l)</i>	coroner	coroner	<i>coronel</i> / coronel	HBSELG3 14
<i>hospicio</i>	hospice	hospice	<i>hospicio</i> / orphanage	HBSELG3 18; DILE; DAA
mental health				
<i>insana/o</i>	insane	insane	<i>insana/o</i> / not healthy	HBSELG3 20
general				
<i>obsceno</i>	obscene	obscene/ excessive	<i>obsceno</i>	HBSELG3 23

Table 1: Taboo Anglicisms discussed in this study

Hablando bien se entiende 3 (2021) includes several Anglicisms which refer to the taboo semantic field of death and dying. *Casquete* and *casco* are Anglicisms (Piña-Rosales *et al.* 2021 12) which come from the English *casket*. In standard Spanish, 'casket' is *ataúd*, *casquete* is a skullcap and *casco* is protective headgear or helmet. The Anglicism and its etymon are taboo words from the semantic field of death. Normally, *casquete* is not taboo in Spanish although the expression *echar un casquete* is taboo and refers to coitus.

Also from the semantic field of death and dying is the Anglicism *hospicio* (Piña-Rosales *et al.* 2021 18), from the English *hospice*, which refers to a place where terminally sick patients get care. In most varieties of Spanish, *hospicio* is an orphanage, and as the authors of HBSELG3 state, perhaps the more common word to refer to an old age home is *asilo*, although according to the *Diccionario de americanismos*, *hospicio* can refer to an '*asilo para dementes y ancianos* [an asylum for the insane and elderly]' in the Dominican Republic, Colombia, Ecuador, Bolivia, Chile, Argentina and Uruguay. The DLE also registers that meaning for *hospicio*, although restricts it to Bolivia, Chile and Ecuador.

The authors of HBSELG3 (Piña-Rosales *et al.* 2021 14) cite *coronel/coroner* as an Anglism for *forense*, which means *coroner* in standard Spanish. The homonym is produced because the lateralization of syllable final 'r', a common phonological phenomenon in Spanish (Schwegler/ Kempff 2007 346). The Anglicism and the English etymon are from the taboo semantic field of death and dying. In English, *coronel*, as in Spanish, refers to a military officer.

Insano/a refers to mental health and is an Anglicism from 'insane' which in standard Spanish would be *loco* or *demente* (Piña-Rosales *et al.* 2021 20). In standard Spanish, *insano* means unhealthy. *Lavatorio* in most varieties of Spanish is a sink, and is not taboo. In the United States it comes from *lavatory*, which is a bathroom or restroom (Piña-Rosales *et al.* 2021 21), and could be taboo as it is a place where one performs his or her physiological functions. The Anglicism *obsceno* shares the meaning with its English etymon as 'offensive to morality or decency' (Piña-Rosales *et al.* 2021 23), and in standard varieties of Spanish, *obsceno* only has sexual connotations. The English term 'obscene' has sexual connotations but it is also used to intensify the meaning of a word, such as in 'obscene wealth'.

4. Discussion

The two names, "El Barajo Restaurant" and "The Stroonz" which inspired this study are not homonymous Anglicisms with other words present in U.S.

Spanish or U.S. Italian. *Barajo* is an euphemism (that is, a word or linguistic resource which serves to avoid the use of a taboo word), but it has nothing to do with English, and even though *barajo* is collected in the DLE, its use as a noun is not common. Its similarity to the foul word stands out. However, the anecdote related to the restaurants' names shows that Utica Spanish-speakers will not accept the presence of a taboo word in the linguistic landscape of the city. *Stroonz* could be classified as an Italianism in English, although most, if not all, English-speakers are not familiar with its meaning. Unlike the Utica Spanish-speakers, the Italian-speakers of the city accepted the taboo name of the bar, or at least they did not register any protests. These two examples beg the question about tolerance or intolerance of taboo words in the context of a U.S. city where English is dominant.

In order to understand the homonymy present in this study, it is useful to look at an example. One of the words discussed in a contrastive study of Anglicisms in five Romance languages spoken in the United States (Galician, Italian, Louisiana French and Rumanian) (Thomas 2021) was the adaptation of the English etymon 'grocery' (a food store or food items bought in the store) which gives *grocería* (es) in Spanish. *Grosería* is also present in the *Diccionario de anglicismos del español estadounidense* (DAE) (2018: 63). In both Peninsular and Latin American standard Spanish, *grosería* means 'rudeness'. Anglicisms such as *grosería* turn out to be confusing because they are homonyms with a word already present in Spanish.

Bilingual speakers of Spanish and U.S. English introduce English words into their Spanish. Daily life in the U.S. is the source of these words that are associated with a different culture, as well as with a different dominant language. The words enter U.S. Spanish most likely via code switching and then, over time, are adapted and lexicalized.

If we were dealing with two separate languages, for example, a variety of Spanish spoken outside the U.S. and English, the words discussed in table one would be false friends (two or more words in two or more languages that have the same or almost the same spelling and/or phonological shapes but whose meanings are different). Indeed the authors of *Hablando bien se entiende la gente* consider the Anglicisms in their volumes to be false friends. When the Anglicism coincides with a taboo word, it may not only lead to a misunderstanding but to a breach of social norms, especially when dealing with Spanish-speakers outside of the U.S.

Among the hundreds of entries recorded in *Hablando bien se entiende la gente*, only fourteen taboo words were found. The low quantity is not surprising considering the prohibited nature of taboo concepts as well as the less than likely coincidence that the taboo word would match a word already lexicalized in Spanish. These words are of special interest here because their use in monolingual communities outside of the U.S. is restricted owing to their taboo nature.

Does their integration into U.S. Spanish indicate a degrade and contamination in U.S. Spanish as the authors of *Hablando bien se entiende la gente* suggest? Does their use point to a weakening of Spanish and a strengthening of English, such as the acceptance of the Italianism 'stroonz' in the linguistic landscape of Utica? The study of these words can shed light on the nature of the linguistic contact between Spanish and English in the U.S.

Since the words discussed here are recorded in dictionaries and in *Hablando bien se entiende la gente* they are codified and in a sense, at least part of an informal spoken standard. Even though the intent of the authors of *Hablando bien se entiende la gente* is to call attention against their use, the fact that they are discussed indicate that they are used frequently enough to be noticeable to the authors. Therefore, we are not dealing with false friends, but rather loanwords that are homonyms of a word known to all Spanish-speakers.

Not all of the etyma of the Anglicisms are taboo in English- neither are all of the matching forms in Spanish taboo. The Anglicism, though, when finally integrated into U.S. Spanish, can be taboo or can be confused with a taboo term in Spanish. The words are classified into the following categories of taboo semantic fields: decency (pregnancy): *embarazada* and *preñada*; decency (sexual relations): *tener sexo*, *estar excitado*, *introducir*, *estoy caliente*, *preservativo* and *molestar*; death and dying: *casquete/casco*, *coronel/r*, *hospicio*; bodily functions: *lavatorio*; mental health: *insano/a*; and a general category: *obsceno* (See Table 1).

The English etyma for *embarazada* (embarrassed), *introducir* (to introduce), and *estar excitado* (to be excited) are not taboo; however, the Spanish signified are so. Monolingual Spanish-speakers would interpret the Anglicism as taboo.

The Spanish etymon for *molestar* (to molest) is not taboo; but, upon acquiring the English meaning, it becomes so, although monolingual Spanish-speakers would most likely not interpret the taboo meaning, but rather the meaning 'to irritate'. Neither are *casquete/casco*, *coronel* and *lavatorio* taboo in monolingual varieties of Spanish. 'Pregnant' (which is related to *preñada*) is

taboo in both English and Spanish, although in Spanish *preñada* has a harsh connotation when used for human beings and is considered vulgar (Martín 1979 231).

Some of the entries in *Hablando bien se entiende la gente* are present in Spanish-speaking countries. In addition, the twenty-third edition of the official dictionary (DLE) of the RAE includes some of the words. Such is the case of *preñada*. The DLE does not censure the word and implies humans are included: "*Dicho de una mujer, o de una hembra de cualquier especie* [said of a woman or of a female of any species]".

The authors of *Hablando bien se entiende la gente* also recognize that some Spanish-speaking countries use *preñada* for human beings. *Tener sexo* has a similar story; the meaning of 'venereal pleasure', collected in the DLE, allows the expression to mean 'to have sexual relations'. The *Diccionario de anglicismos del español estadounidense* and the *Diccionario de americanismos* (DAA) include *introducir* (used as *presentar*). The DLE and the DAA include meanings for *hospicio* that are not consistent with the description in HBSELG3, but do indicate semantic overlap with 'hospice'. The DAE also records the use of *preservativo* as a food preservative in the U.S. These five examples indicate some divergence between the ANLE and the RAE/ ASALE, in that the ANLE in *Hablando bien se entiende la gente* speaks more strongly against their use.

The uses documented in *Hablando bien se entiende la gente* for *embarazada*, *estar excitado*, *casquete/casco*, *coronel/r*, *estoy caliente*, *insano/a*, *laboratorio*, *obsceno* and *molestar* do indeed evidence English influence in U.S. Spanish. Their anglicized meanings break the pan-hispanic unity that they share with other varieties of Spanish, resulting in misunderstandings or even offense when used with monolingual speakers outside the U.S. Therefore, as part of its fundamental mission to ensure the unity of the language, the ANLE calls attention to these changes.

The Director of the ANLE and the first author of *Hablando bien se entiende la gente*, Gerardo Piña-Rosales, responded to Lynch and Potowski's article in a letter to the Editor of *Hispania* (Piña-Rosales 2014). He denied that the intention of the volume was to censure the speech of U.S. Hispanics, but rather to recommend alternatives:

"Hablando bien se entiende la gente no tiene otro objetivo que el recomendar equivalencias..."

[Hablando bien se entiende la gente doesn't have any purpose other than to recommend equivalences...]" (Piña-Rosales 2014 255).

Piña-Rosales further states:

Hablando bien se entiende la gente no está orientado a fomentar el estudio del español, sino a mejorar el uso del idioma entre hispanohablantes con un dominio más o menos avanzado del español, a quienes, debido al contacto con el inglés, lengua dominante, se les ha ido anglicando su idioma materno.

[Hablando bien se entiende la gente is not oriented to promote the study of Spanish but to improve the use of the language among Spanish-speakers with a more or less advanced command of Spanish, who, due to the contact with English, the dominant language, have been anglicizing their mother tongue.] (Piña-Rosales 2014 356).

In spite of the sincere intention of the authors of these volumes, there is a purist ideology behind the works. Piña-Rosales' statement about recommending alternatives also highlights the observation that these Anglicisms are indeed being used by speakers of Spanish in the U.S. And, even though the authors claim that their recommendations are alternatives, they call the Anglicisms false friends in the prologue of HBSELG3:

"...ofrecemos nuevos consejos a modo de recomendaciones sencillas y divertidas para evitar la confusión con términos engañosos parecidos del inglés- los "falsos amigos", que confirman el dicho de que las apariencias engañan- y para facilitar la comunicación por medio de un español correcto.

[...we are offering new tips as simple and fun recommendations to avoid confusion with misleading similar terms from English- the false friends which confirm the saying that appearances are deceiving- and to facilitate communication through a correct Spanish] (Piña-Rosales et al. 2021 9).

'False friends' necessarily implies two languages, English and Spanish. In spite of saying that they are offering alternatives, the authors believe that the Anglicisms should not be used in U.S. Spanish.

Although the volumes of HBSELG are not differential dictionaries, the discussion above illustrates how these words are an example of the ideological debate about the dictionaries of -isms in the setting of panhispanic lexicographic norms. We have seen how some words, attributed as Anglicisms in the U.S., were really part of the Spanish of other countries in Hispanic America (Ávila 2003; Ávila 2003-2004; Fajardo Aguirre 2011). This, in part, explains why the *Diccionario de americanismos* includes some of the alternatives against which the editors of HBSELG recommend.

One cannot affirm that a certain word or its meaning are unique to a certain country without having exhaustive studies of all the other Spanishes with which to compare the word, doing so is a failure in methodology (Zimmermann 2003). The editors of HBSELG recommend vocabulary that is closer to the Penninsular Spanish standard than that used in Latin America, which closer to the speech of U.S. Spanish-speakers. This ideology is also similar to that of the authors of differential dictionaries of -isms, because the authors contrast the Anglicism, at least implicitly, with a form that is more common to the lexicon of Spain (Ávila 2003; Ávila 2003-2004; Zimmerman 2003, 2013:106; Adelstein 2016:165).

The volumes of HBSELG differ from dictionaries of -isms in that their intention is not to describe "*lo diferente, lo raro, lo periférico*" [the different, the strange, the peripheral] (Ávila 2003-2004 13), but to recommend against the use of certain forms. Therefore, they are more akin to the *Appendix Probi*, which was a sixth or seventh century list of words and expressions to be avoided when writing Latin along with the recommended forms side-by-side (Penny 1991 4). Even though speakers of vulgar Latin did not heed the recommendations set forth in the *Appendix Probi* to improve their contaminated Latin, the list, just as the three volumes of HBSELG, serve as documentation of linguistic changes in progress.

Languages are constantly changing. These words present innovative uses by Spanish-speakers in the U.S. and could be effective agents for lexical change, thereby shaping the identity of U.S. Spanish.

5. Conclusion

The three volumes of *Hablando bien se entiende la gente* (2010, 2014, 2021) recommend against the use of Anglicisms or constructions which the ANLE considers to be unnecessary and/or incorrect in the Spanish of the U.S. Fourteen of those were discussed here and are homonyms with a word already present in Spanish. The ANLE believes that these words are the product of disproportionate influence of English on the Spanish spoken in the U.S., even though some are used in other varieties of Spanish in the Americas.

There are very few of these words, in part, because they are taboo but also because the likelihood of sound coincidences are low. Some of these words illustrate the divergence between the publications of the Real Academia Española/ ASALE and the judgments of the members of the ANLE; this

divergence might be explained by the observation that the authors of these volumes seem to favor the lexical standard of north-central Spain.

These words exemplify the debate which the differential dictionaries of -isms have raised with respect to their suitability and methodological approach. The inclusion of these words in the volumes *Hablando bien se entiende la gente* suggest that they are lexicified at least in an informal, spoken register of U.S. Spanish, in spite of the grounds given against their use.

References

- Amorós Negre, Carla (2014): *Las lenguas en la sociedad*. Madrid, Síntesis.
- Academia Norteamericana de la Lengua Española (ANLE) (2021): *Misión*. Web. <https://www.anle.us/nuestra-academia/mision/>. accessed 16 May 2021.
- Adelstein, Andreína (2016): *Comprehensive Dictionaries and the delimitation of the Argentine variety of Spanish*. In: Muhr, Rudolf, et al. (eds) (2016): *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions*. Wien et al., Peter Lang Verlag. pp. 163-180.
- Asociación de Academias de la Lengua Española (ASALE) (2010): *Diccionario de americanismos*. Madrid, Santillana.
- Ávila, Raúl (2003): *Diccionarios locales, nacionales, internacionales de americanismos*. *Revista Internacional de Lingüística Iberoamericana*. 1,1, pp. 51-66.
- Ávila, Raúl (2003-2004): *¿El fin de los diccionarios diferenciales? ¿El principio de los diccionarios integrales?* *Revista de Lexicografía X*, pp. 7-20.
- Chamizo-Domínguez, Pedro J. (2008): *Semantics and pragmatics of false friends*. New York, Routledge.
- Ezcurra, Álvaro (2020): *Léxico general y pluricentrismo: aproximación a los americanismos en el diccionario académico*. In: Sebastian Greußlich / Franz Lebsanft (eds.). *El español, lengua pluricéntrica. Discurso, gramática, léxico y medios de comunicación masiva*. Göttingen. V & R unipress, pp. 61-80.
- Fajardo Aguirre, Alejandro (2011): *La norma lingüística del español desde una perspectiva lexicográfica: norma nacional versus norma panhispánica*. *Normas*. *Revista de Estudios Lingüísticos Hispánicos*. 1, pp. 53-70. <<https://www.uv.es/normas.>> (17 November 2011).
- Greußlich, Sebastian (2020). *Spanish normative dictionaries*. In: *Manual of standardization in the Romance languages*. Berlin: de gruyter, pp. 605-627.
- Fundéu. (2019). *Fundéu buscador urgente de dudas*. <http://www.fundeu.es/consulta/tener-sexo-5>. accessed 24 July 2020.

- Lara, Luis Fernando (1996): *Teoría del diccionario monolingüe*. México: El Colegio de México.
- Lynch, Andrew/ Potowski, Kim (2014): La valoración del habla bilingüe en los Estados Unidos: Fundamentos sociolingüísticos y pedagógicos en *Hablando bien se entiende la gente*. *Hispania* 97, 1, pp. 32-46.
- Penny, Ralph (1991): *A history of the Spanish language*. Cambridge UK: Cambridge Press.
- Piña-Rosales, Gerardo (2014): Letter to the Editor: En respuesta a un artículo publicado en *Hispania*. *Hispania* 97, 3, pp. 355-356.
- Piña-Rosales, Gerardo/ Covarrubias, Jorge I./ Segura, Joaquín/ Fernández, Daniel (eds.) (2010): *Hablando bien se entiende la gente*. Miami, FL, SantillanaUSA.
- Piña-Rosales, Gerardo/ Covarrubias, Jorge I./ Dumitrescu, Domnita (eds.) (2014): *Hablando bien se entiende la gente 2*. Miami, FL, SantillanaUSA.
- Piña-Rosales, Gerardo/ Covarrubias, Jorge I./ Fernández, Daniel/ Rodríguez, Porfirio/ Prunes, Natalia (eds.) (2021): *Hablando bien se entiende la gente 3*. El Monte, CA, VelázquezPress.
- Martín, Jaime (1979): *Diccionario de expresiones malsonantes del español*. Madrid, Ediciones ISTMO.
- Moreno Fernández, Francisco (2018): *Diccionario de anglicismos del español estadounidense*. Cambridge, MA, Instituto Cervantes at FAS-Harvard University.
- Schwegler, Armin/ Kempff, Juergen (2007): *Fonética y fonología españolas*. Hoboken, NJ, John Wiley and Sons, Inc.
- Thomas, Juan A. (2019): *Imágenes borradas y letras cambiadas: el español en el paisaje lingüístico de Utica, NY*. *Camino real* 11:14, pp. 125-137.
- Thomas, Juan A. (2021) *Anglicismos en las lenguas de migrantes en Estados Unidos: Un estudio contrastivo*. Talk presented at the panel: *Contraste de lenguas y sus ámbitos de aplicación*, XIV Congreso Internacional de Lingüística General. Universidad de Sevilla, Seville, Spain. presented 23 June 2021
- U.S. Census (2019): *American Community Suvey. 2019 Estimates*. <http://data.census.gov>. Accessed 25 August 2021.
- Real Academia Española and Asociación de Academias de la Lengua Española (2014): *Diccionario de la lengua española (DLE)*, 23ra ed. Madrid. accessed 31 August 2021.
- Real Academia Española and Asociación de Academias de la Lengua Española (2021): *La RAE*, Madrid. accessed 4 June 2021

- Zimmermann, Klaus (2003): El fin de los diccionarios de americanismos. La situación de la lexicografía del español de América después de la publicación de los "Diccionarios contrastivos del español de América." *Revista Internacional de Lingüística Iberoamericana* 1,1, pp. 71-83.
- Zimmermann, Klaus (2013): El papel de los diccionarios diferenciales y contrastivos en la estandarización de variedades nacionales en un español pluricéntrico. In: Muhr, Rudolf, et al. (eds) (2013): *Exploring Linguistic Standards in Non-Dominant Varieties of Pluricentric Languages. Explorando estándares lingüísticos en variedades no dominantes de lenguas pluricéntricas*. Wien et. al., Peter Lang Verlag. pp. 99-114.

Soledad CHÁVEZ FAJARDO

(Universidad de Chile, Chile)

schavez@uchile.cl

The (In)Visible (Hispanic) American in Hispanic Lexicography. 19th and 20th Centuries. Some Cases¹

Abstract

In this research, we want to show various ideologising instances present in Spanish-American lexicography (from the mid-19th century to the early 20s of the 20th century), such as some aspects related to the question of Eurocentrism, its problems, some critical perspectives and its reaffirmation in some speech acts. For instance, from the examples it will be possible to appreciate the Spanish lexicographical ignorance of the Latin American reality. Furthermore, the problems that have been generated from the Spanish-American lexicographic discourse are due to the fact that the information present in the Royal Spanish Academy's Dictionary is insufficient or non-existent in relation to the Spanish-American lexicon. At the same time, some reflections and disputes related to how long it takes to "consecrate" a neologism of irregular formation from Latin America within the official lexicographic tradition will be shown.

1. Reading Dictionaries; Dictionaries as Speech; the Ideology Present in a Dictionary.

As they are historical and linguistic objects (Courtine 1981), it is necessary to study dictionaries as discourses and treat these lexicographic discourses as representations that rescue historical, political and social contexts of the community in which the dictionary was written. In other words, the lexicographic discourse refers to an ideological dimension linked to the conditions of production in which it was formulated, as well as the circumstances of its enunciation and, furthermore, its broader socio-historical context¹.

¹ A special acknowledgment to Kilian Kennedy who assisted with translation.

Therefore when reading a dictionary (when critically studying it, dictionaries are read), memories and, above all, ideologies and social imaginaries² are activated. Regarding ideology, understood in the most general way as the "general beliefs (knowledge, opinions, values, truth criteria, etc.), of entire societies or cultures", (Van Dijk 1999: 92), the reflections of Pascual and Olaguíbel (1991) and Camacho (2003-2004 and 2013) are pertinent when it comes to the impossibility of a lexicographer achieving *objectivity* when writing a dictionary², that is, of omitting any ideological reference in the lexicographic exercise. These authors affirm that any effort to erase the traces of ideology in a lexicographical work is practically impossible.³

Although ideological references can be found throughout the dictionary, such as in titles, paratexts, typographical arrangement, markings or in the selection of the *lemario* (index of entries), "the most likely place for the appearance of ideological contents has been – and continues to be – the lexicographical definition" (Azorín 2011: 122), something that I will return to later with the examples selected for this study.

In turn, ideology in lexicographic discourse, can be analysed in various ways; for example, some authors, in order to methodologically organise their study and analysis, have mentioned some outstanding elements, such as Blecua (1990), who recommended looking at personal deictics and their entire sphere, something that was usual in the foundational Spanish-American lexicographic discourse until the dawn of the 20th century. The spatial deictics, such as "Spain and Europe" as references, will be fundamental in the examples selected in this essay; the temporary deictics, as *in the old days*, *in the modern day*, *today*, *now*; the first person plural pronoun – *nosotros* – and its abbreviated form – *nos* – which in the selected examples will mark the difference between Hispanic Americans and Spaniards – or, strictly speaking, with the Spanish entity Royal Spanish Academy (RAE) – the verb forms conjugated in the first person of the plural and the possessive pronouns, where it is also insisted on that unity (be it national, be it continental) of Hispanic America and its reality; adjectives and other evaluative elements, such as affective or evaluative subjective and affective suffixes, which abound and, in the selected examples, will account for linguistic attitudes or critical reflections, such as the question of Eurocentrism

² It's what Pérez (2000:12) calls "To sociologically consider the lexicographic text".

³ Lauria 2011, 2012a and 2012b have worked very well with concerning the foundational Argentine lexicography and the first centenary of the Nation-State and what Casares proposed in his founding study of Spanish lexicography (1950).

or the critical situation of the Spanish-American lexicon within the official lexicography from Spain.

In addition, I would like to use some of the very helpful distinctions made by Camacho (2013: 71-72) who classified the information present in the definitions of the dictionaries of Cubanisms in various ideologies.

First, for this research, I will use for this research the concept of *social ideology*, that is, the statement is presented in relation to individuals in their capacity as members of society, such as can be seen in many of the selected examples in both the Hispanic-American community or a particular American country compared to Spain, the United States, and Europe.

Second, *identity ideology*, where the statement is presented in relation to the features that distinguish or identify the nation, a constant aspect as will be seen in some of the examples, especially in words whose definitions account for and inform about a certain national or continental reality.

Third, *ideology linguistics*, where the statement is presented in relation to the ideal of language and correctness, the founding engine of Spanish-American lexicography, either to validate a use or to censor it.

Finally, I will follow the line of Forgas, who in a foundational study regarding ideology and lexicography, stated that each word, each entry in a dictionary: “is, in reality, an *ideologeme*, since through the lexicographic definition that word is transferred to terms of meaning, or, what is the same, to terms of ideology” (Forgas 1996: 73).

The *ideologeme*, a concept coined by Angenot (1982), is understood as “the representation, in the ideology of a subject, of a practice, an experience, a social feeling. The ideologeme articulates the contents of social consciousness, enabling its circulation, its communication and its discursive manifestation in, for example, literary works”. (Kerbrat-Orecchioni 1980: 35). In fact, in this study some ideologemes will be presented, say, “uncomfortably” within the lexicographic practice that has to do with the Spanish-American lexicography and the characteristic Spanish-American lexicon.

In summary, in this research, we want to show various ideologizing instances present in Spanish-American lexicography (from the mid-19th century to the early 20s of the 20th century), such as some aspects related to the question of Eurocentrism, its problems, some critical perspectives and its reaffirmation in some speech acts.

From the examples it will be possible to appreciate the Spanish lexicographical ignorance of the Latin American reality. Also, the problems

that have been generated from the Spanish-American lexicographic discourse are due to the fact that the information present in the Dictionary of the Royal Spanish Academy is insufficient or non-existent in relation to the Spanish-American lexicon. Some reflections and disputes related to how long it takes to “consecrate” a neologism of irregular formation from Latin America within the official lexicographic tradition will also be discussed.

2. The Question of Eurocentrism: its Criticism and its Acceptance.

The question of Eurocentrism in metalexigraphic studies in the Spanish language has been dealt with, from a critical point of view, quite latterly. It is not until Lara's studies towards the end of the last century when the Eurocentrism that Spanish-American lexicography has had since its dawn, began to be emphasised (cf. Lara 1990) from various angles; for example, constantly subjecting lexicographical planning to academic works, by pointing out which word is or is not in the academic dictionary, or that if such a word is defined in a certain way, its use in another variety is incorrect.

In turn, most of the authors of Spanish-American lexicography until well into the twentieth century handled one monological exemplarity in the Spanish language: the variety regulated by the Royal Spanish Academy (RAE). In this way, the information that exists in these lexicographical articles is usually, directly or indirectly, a dialogue with this entity. There was not, strictly speaking, until well into the twentieth century a knowledge of the concept of linguistic variety and most of the time, this is usually considered erroneous or subordinated to a prestigious variety.

What is presented below are very interesting cases where you can see both sides of the coin. On the one hand, there is a subordinate attitude to an entity such as the RAE, with a constant concern that the said Diccionario entity integrates, incorporates, or modifies definitions of the relevant lexicographical articles. On the other hand, there is a critical attitude towards Eurocentrism, be it in sharp criticism towards the RAE, for example, or towards the ignorance that the European has before the Latin American reality.

3. Some Cases in the Foundational Spanish-American Lexicography Before the Beginning of the 20th Century.

1. “*Este lujo la Real Academia Española solo se ha gastado con la madre patria.*” [This luxury the Royal Spanish Academy has only afforded the mother country.]

There are some cases in which the proposal in some lexicographical articles resides in including and extending the concepts in pursuit of the visibility of an American reality. In other words, the ideologeme in these lexicographical articles has to do with the silencing of the Spanish-American reality. Such is the case of the pair of examples of lexicographical articles of demonyms present in the *Diccionario de Chilenismos y de otras voces y locuciones viciosas* (*Dictionary of Chileanisms and Other Non-Standard Words and Phrases*) of the diocesan priest Manuel Antonio Román (published in 5 volumes, between 1901 and 1918). One is *bonaerense* (“from Buenos Aires”):

Bonaerense, adj. Natural de Buenos Aires. Ú.t.c.s. |Perteneiente o relativo a esta ciudad de América. Adición del último Dicc. ¿Por qué llama este a Buenos Aires ciudad solamente y no también provincia? Porque entendemos que su intención ha sido dar los nombres gentilicios de los Estados de América y de sus respectivas capitales solamente, y no los de provincias, departamentos, ciudades de segundo orden, villas, aldeas, etc. Este lujo solo se ha gastado con la madre patria. (1901-1908)

[Bonaerense, adj. Born in Buenos Aires. [Also use like a noun]. Belonging or relating to this city in America. Addition of the last Dict. Why does he call Buenos Aires only a city and not also a province? Because we understand that his intention has been to give the national names of the States of America and their respective capitals only, and not those of provinces, departments, second-order cities, towns, villages, etc. This luxury has only been afforded to the mother country. (1901-1908)]

Román's claim in relation to the semantic fields incompletely covered in the RAE's dictionary – this case in particular, that of the demonym – goes beyond accounting for an absence. In effect, this absence or insufficiency implies a detriment to the information related to Latin America. The irony also stands out: "This luxury has only been afforded to the mother country." Or in the case of Buenos Aires, what Manuel Antonio Román asks is that he realise the Chilean reality, as in Buenos Aires, because in the Royal Spanish Academy's Dictionary he has not added the reference to the inhabitants of Valparaíso, Chile: "As this is our main port and so beautiful [...] it is right that we call the inhabitant of it porteño and everything that belongs to and is relative to it" (1913-1916). Something that Zerolo did (1895) and that the Royal Spanish Academy's Dictionary began to do in the 1984 edition (for both references, see the NTTLE).

2. “Deje pues a un lado la academia los escrúpulos de monja y el cierto puntillo de amor patrio.” [Put aside, RAE, the scruples of a nun and the certitude of patriotic love.]

The verb *independizar* (*make independent*) brings with it an interesting normative history that caused it to be absent from many dictionaries, something that caused suspicion in some South American intellectuals, precisely because of the historical connection to which the verb refers. The ideogeme, in this case, is a question of the censorship of Spanish-American neologisms, as in the case of this verb.

The morphology of the verb did not correspond to the necessary rules for a verb ending in *-izar*, something that was observed for the first time by Rufino José Cuervo, who vetoed it in all editions of his *Apuntaciones críticas* (*Critical Notes*), always suggesting the use of *emancipar* o *hacer independiente* (*emancipate* or *make independent*) instead of the verb *independizar* (cf. 1867-1872; 1876). In the 1907 edition he affirmed: “a *independizar* no le faltan abogados, pero su formación es a todas luces defectuosa, y solo podría disculparse por una especie de haplogía que hubiera aligerado el teórico inacabable *independentizar*” (1907).

Along these lines, one who did not advocate for the use of the verb is Carlos Gagini: “The Spanish have always said *emanciparse*, *hacerse independiente*, *libertarse*, (*to emancipate themselves*, *to become independent*, *to liberate themselves*), but Americans exclusively use the verb to become independent. As the formation of this neologism does not conform to the derivation principles established by grammar, we advise the use of peninsular terms” (1892). From Spain, Eduardo de Huidobro: “I say the same: neither this verb nor the noun independence are Castilian. You have to say *become* or *be independent*.” (1903).

The RAE began by including the verb in its “manual” dictionary, (published from 1927 to 1989), which was a sort of proving ground for words which might or might not be acceptable in the normative “usual” dictionary. The “usual”, or popular dictionaries could penalise its use: “Useless neologisms *emancipar* or *emanciparse* ” (1927), something that is replicated from Latin America by authors such as the Chilean José Toribio Medina (1928), the Dominican Federico Llaverías (1940) and the Colombian Félix Restrepo (1943) who recognises that the verb is very generalised (cf. RAE file).

In turn, there is another line that advocates its use and, what is more, seeks that the RAE incorporates the term in its dictionary. Along these lines is the Peruvian Ricardo Palma, who also connects the verb with the Hispanic-

American historical process: "Since we became independent from Spain, this insurgent verb has been alive, as well as its reflexive *independizarse*, without any American, learned or uneducated, taking care to find out whether or not it is in the Dictionary" (1896).

Palma will be one of the great defenders of the verb and its officialisation, as well as a great critic because the RAE does not consider it in its index of terms: "The Academy, which has dislike, annoyance and ill will towards such a verb, maintains that it is enough and exceeds with *emancipar*" (1903). Palma's arguments are semantic, not touching on anything related to morphological construction: "we Americans say that the slave is emancipated and the child of the family is emancipated, protected by law and by the civil code; but that the peoples become independent" (1903) and he will not tire of bringing up the historical relevance of the verb in Latin America:

"Whoever emancipates himself does so protected by law, and rare is the case in which emancipation is the result of litigation and sealed paper. On the contrary, independence is almost always obtained by the peoples by dint of fighting and shedding blood [...]. The conquest, which is an imposition of force, does not create rights. The yoke is broken in the same way that it was imposed, by force; and that cannot be called emancipation but liberation, *independizarse*." (1903).

He is followed by the Chilean diocesan priest Manuel Antonio Román, who recognises that it is a verb: "It is widely used, mainly in Latin America" (1913), for whom, following the line of unilateral dialogue with the RAE so that it "accepts it in its dictionary" [...] "and it already seems ripe enough for admission" (1913). Unlike Palma, who advocates for semantic reasons, Román, along with countering Cuervo's position, responds to haplology with euphony:

...pero a nadie puede convencer esta razón, porque al formar los vocablos largos, las lenguas no siguen escrupulosamente las leyes de la fonética, sino las de la eufonía. Por eso a nadie le ocurrió decir *analizar*, de análisis, ni *synthesizar*, de síntesis, ni *irisizar*, de iris, ni *catequesizar*, de catequesis, ni *descatolicizar*, de católico, etc., sino que todos optaron por una forma más breve y eufónica: *analizar*, *synthetizar*, *irizar*, *catequizar*, *descatolicizar*. Y esto sin salir de los verbos en *izar*, para que la réplica sea más concluyente. (1913)

[but nobody can convince this reason, because when forming long words, languages do not scrupulously follow the laws of phonetics, but those of euphony. That is why it did not occur to anyone to say *analyze*, of *analysis*, or *synthesize*, of *synthesis*, or *irize*, of *iris*, or

catechesise, of catechesis, or *de-catholicize*, of Catholic, etc., but everyone opted for a shorter and more euphonious form: analyze, synthesize, irize, catechize, de-catholize. And this without leaving the verbs in “izar”, so that the reply is more conclusive.] (1913)

Then comes one of the characteristic outbursts of the priest who, although he followed the opinions of the RAE, did not shrink at the moment of verifying any inconsistency or neglect of the Spanish entity: “Let the academy put aside the scruples of a nun and the certitude of patriotic love regarding the admission of this verb and do us the gallantry to accept it” (1913).

The verb ended up lemmatised in the RAE Dictionary in its 1956 edition. Regarding how a use censored by the same institution was established, it is worth remembering the words of one of its defenders, Ricardo Palma, who argued:

Si nos echáramos a rebuscar autoridades, nos sería difícil encontrar escritor notable en nuestras repúblicas que no haya usado el verbo *independizar*, contra el cual no hay razón filológica para que no se le estime como de legítima y buena cepa castellana” (1903).

[If we were to search for authorities, it would be difficult for us to find a notable writer in our republics who has not used the verb *independizar* (to become independent), against which there is no philological reason why it should not be considered as a legitimate and good Castilian strain.]

Regarding this, the same RAE *Fichero* (file) brings some emblematic examples of the use of the verb in writing by some important Spanish intellectuals, such as the orientalist and member of the Royal Academy of History Josep Maria Millàs Vallicrosa:

Lo mismo que entre las distintas dinastías en Oriente, de hecho independizadas, en España cada príncipe quería emular el lujo y esplendor de la corte bagdadí” *Poesía hebraio-española*, 1940.

[The same as between the different dynasties in the East, in fact independent states, in Spain each prince wanted to emulate the luxury and splendor of the Baghdad court .]

Also, the writer and academic from the RAE Carmen Conde;

Hace dos años sin justificación alguna determinó independizar su vida de la mía”, *En manos del silencio*, 1945.

[Two years ago without any justification she determined to make her life independent from mine”]

The member of the RAE, writer and academic, Miguel Delibes:

y, al fondo, una escarpada muralla de roca viva que les independizaba del resto del valle”, *El camino*, 1950.

[and, in the background, a steep wall of living rock that made them independent from the rest of the valley.]

The philologist and romanist Martin de Riquer:

con la sola excepción de Guifred, o sea Wilfredo el Velloso, conde de Barcelona, que logró la victoria política de independizarse del rey de Francia”, *Los cantares de gesta franceses*, 1952.

[with the sole exception of Guifred, that is, Wilfredo el Velloso, Count of Barcelona, who achieved the political victory of becoming independent from the King of France.]

3. “Los europeos, que saben tan poco de América que da grima” (“The Europeans, who know so little about America that it is creepy.”

In the case of the *Diccionario de chilenismos y de otras voces y locuciones viciosas* (first volume, 1901-1908), the diocesan priest Manuel Antonio Román directly expresses his annoyance at the lack of knowledge of Hispanic American history: In the case of the *Diccionario de chilenismos y de otras voces y locuciones viciosas* (first volume, 1901-1908), the diocesan priest Manuel Antonio Román directly expresses his annoyance at the lack of knowledge of Hispanic American history:

Bolívar, m. “Moneda de plata de Venezuela, equivalente a una peseta. Es la unidad monetaria”. Admitido en el último Dicc., pero con tres defectos, a nuestro juicio: 1.º no dar la etimología, que, aunque para los americanos sea sabidísima, porque ninguno de ellos ignora quién fue el venezolano Simón Bolívar, no lo es así para los europeos, que saben tan poco de América que da grima; [...]. (1901-1908)

[Bolívar, m. “Silver coin from Venezuela, equivalent to one peseta. It is the monetary unit. Admitted in the last Dict., but with three defects, in our opinion: 1st not giving the etymology, which, although for the Americans it is very well known, because none of them ignores who the Venezuelan Simón Bolívar was, it is not so for Europeans, who know so little about America that it's creepy. [...]. (1901-1908)]

To what extent would there be a lack of knowledge about who Simón Bolívar (1783-1830) was at the dawn of the 20th century? The first references are from encyclopedic lexicography: twenty years after Bolívar's death, the encyclopedic dictionaries of Domínguez and the Gaspar y Roig publishing house (1853) included it in two biographical lexicographical articles.

It is also interesting how Bolívar is characterised: Domínguez qualifies him as a *liberator*, while the Gaspar y Roig *Diccionario* starts by qualifying him as "Founder of the Republic of Bolivia" and then describing him as "dictator and later president of the Republic of Colombia", closing the end of his life with "but accused of aspiring to tyranny, he resigned this position."

Towards the end of the century there is the third encyclopedic reference, that of the encyclopedic dictionary of Zerolo (1895), where the qualifications for Bolívar are substantially modified, since it starts with an "Illustrious general", closing, towards the end of his days with "ingratitude and envy [that] will haunt the last years of his life." In addition, Zerolo includes for the first time the meaning referring to the currency (the bolívar) as the ninth meaning (the only one of language, since the remaining seven have to do with geographical references).

Similar qualifications in another encyclopedic dictionary, already from the 20th century is the one given by Rodríguez-Navas (1918): "Famous American general". On the other hand, Román refers to the term in terms of its monetary meaning and asks the Royal Spanish Academy's Dictionary for an etymological reference for the metonymy in question, which is done, for the first time, in Alemany (1917), followed by the Academy in the 1925 dictionary:

"From the name of Simón Bolívar, who initiated the Independence in America", to modify it in the dictionary of 1992 to: "From S. Bolívar, 1783-1830, Caracas military man who initiated the Independence of America", which remains to this day.

From the name of Simón Bolívar, who initiated the Independence in America", to modify it in the dictionary of 1992 to:

From S. Bolívar, 1783-1830, Caracas military man who initiated the Independence of America", which remains to this day.

This etymological mention has not been exempt from critical reflections, such as that made by the Mexican writer Marcos E. Becerra in his *Rectificaciones y adiciones al Diccionario de la Real Academia Española* (1954):

Con decir ‘de América del Sur’ la R. Academia Española, habría sido más justiciera de lo que quizá intentó serlo en esa anotación. Porque, a cada uno lo suyo. Bolívar, el gran Bolívar, inició la independencia de Sudamérica; la de Norteamérica, Jorge Washington; la de Mesoamérica, Hidalgo” (467-468).

[By saying ‘from South America’ the R. Academia Española would have been more just than perhaps it tried to be in that annotation. Because, to each his own. Bolívar, the great Bolívar, initiated the independence of South America; that of North America, Jorge Washington; that of Mesoamerica, Hidalgo.]

4. Conclusions.

Starting from some examples taken from the foundational Spanish-American lexicography, I wanted to show some aspects of ideology in the Spanish-American lexicographic discourse from the mid-nineteenth century to the early 20s of the twentieth century. I have focused mainly on what is related to the Spanish-American dichotomy, such as some aspects related to the question of Eurocentrism, its problems, some critical views and its reaffirmation in some cases.

Many of the aspects that are usually worked on when ideologically analysing a lexicographic discourse (personal deictics, spatial deictics, first-person singular and plural pronouns, possessive pronouns, evaluative adjectives, among others) are present in these authors. Ideological aspects that have to do with social history, Hispanic American identity and, of course, with linguistic aspects that are detected in a lot of the arguments that are made in many of these lexicographical articles. Precisely, there are many ideologemes that emerge in the Spanish-American lexicographic discourse.

For instance, from the examples it will be possible to appreciate the ignorance that the lexicography from Spain had of the Latin American reality (it can be seen in the case of the lexicographic articles on demonyms). Also, the problems that have been generated, from the Spanish-American lexicographic discourse, due to the fact that the information present in the Royal Spanish Academy’s *Diccionario* is insufficient or non-existent in relation to the Spanish-American lexicon (as in *bonaerense*, *porteño* or *bolívar*). At the same time, some reflections and disputes were shown related to how long it takes to “consecrate” a neologism of irregular formation from Latin America within the official lexicographic tradition (as in *independizar*).

In summary, the majority of the authors of the Spanish-American lexicography until well into the twentieth century handled *one* exemplariness: the variety regulated by the Royal Spanish Academy (RAE). In this way, I emphasise, the information that exists in these lexicographical articles tends to be, above all, a dialogue with this entity. There is not, strictly speaking, until well into the twentieth century a knowledge of the concept of linguistic variety and this, most times, is usually considered erroneous or subordinated to a prestigious variety. However, in these critical reflections one can see the germ of what it is to recognise and identify oneself in the linguistic varieties of the Spanish-American states and work towards their visibility.

References

Primary Sources:

- Cuervo, Rufino José (1907): *Apuntaciones críticas sobre el lenguaje bogotano*. París. A. & R. Roger y F. Chernoviz, editores. 5ª edición.
- Cuervo, Rufino José (1876): *Apuntaciones críticas sobre el lenguaje bogotano*. Bogotá. Imprenta de Echeverría hermanos.
- Cuervo, Rufino José (1867-1872): *Apuntaciones críticas sobre el lenguaje bogotano*. Bogotá. Arnulfo M. Guarín.
- Gagini, Carlos (1892): *Diccionario de barbarismos y provincialismos de Costa Rica*. San José de Costa Rica. Tip. Nacional.
- Eduardo de Huidobro (1903): *¡Pobre lengua! Catálogo en el que se indican más de trescientas voces y locuciones incorrectas hoy comunes en España*. Santander. Imprenta de La Propaganda Católica.
- Palma, Ricardo (1903): *Papeletas lexicográficas*. Lima. Imprenta La industria.
- Palma, Ricardo (1896): *Neologismos y americanismos*. Lima. Imprenta y Librería de Carlos Prince.
- Real Academia Española. *Nuevo Tesoro Lexicográfico de la Lengua Española (NTLLE)*. <http://ntlle.rae.es/ntlle/SrvltGUILoginNTlle>
- Real Academia Española. *Fichero general de la Real Academia Española (Fichero RAE)*. <https://apps2.rae.es/fichero.html>
- Román, Manuel Antonio (1913): *Diccionario de chilenismos y de otras voces y locuciones viciosas*. Tomo III: G, H, I, J, K, L, Ll, M y suplemento a estas ocho letras. Santiago. Imprenta de San José.
- Román, Manuel Antonio (1901-1908): *Diccionario de chilenismos y de otras voces y locuciones viciosas*. Tomo I: A, B, C y suplemento á estas tres letras. Santiago. Imprenta de la Revista católica.

Secondary Sources.

- Angenot, Marc (1982): *La parole pamphlétaire*. París. Payot.
- Azorín, Dolores (2011): *Ideología y diccionario. La mujer en el imaginario social de la época a través del Tesoro de la Lengua Castellana o Española de Covarrubias*. In: *Revista Académica, Boletín de la Real Academia Conquense de Artes y Letras*, 6, 111- 128.
- Blecua, José Manuel (1990): *Diccionario y enunciación*. In: *VVAA: Jornadas de Filología en homenaje al Profesor Marsá*. Barcelona. Edicions de la Universitat de Barcelona, 61-74.
- Camacho, Aurora (2013): *Marcas ideológicas en una muestra del discurso de las definiciones en la lexicografía cubana de los siglos XIX y XX: identificación y funcionalidad*. Tesis en opción al título de Doctor en Lingüística, inédita.
- Camacho, Aurora (2003-2004): *Huellas ideológicas en la lexicografía cubana*. In: *Revista de Lexicografía X*, 21-37.
- Casares, Julio (1950): *Introducción a la lexicografía*. Madrid. Consejo Superior de Investigaciones Científicas.
- Courtine, Jean-Jacques (1981): *Analyse du discours politique (Le discours communiste adresséaux chrétiens)*. In: *Langages*, 62, 9-128.
- Forgas, Esther (1996): *Lengua, sociedad y diccionario: la ideología*. In: Esther Forgas (coord.): *Léxico y diccionario*. Tarragona. Universitat Rovira i Virgili, 71-90.
- Kerbrat-Orecchioni, Catherine (1980): *L'énonciation de la subjectivité dans le langage*. París. Libraire Armand Colin.
- Lara, Luis Fernando (1990): *Dimensiones de la lexicografía. A propósito del Diccionario del Español de México*. México. El Colegio de México.
- Lauria, Daniela (2012 a): *Continuidades y discontinuidades de la producción lexicográfica del español de la Argentina. Un análisis glotopolítico de los diccionarios publicados en el marco del Centenario y en el del Bicentenario de la Revolución de Mayo*. Tesis doctoral. Buenos Aires. Universidad de Buenos Aires.
- Lauria, Daniela (2012 b): *El primer diccionario integral del español de la Argentina: reflexiones acerca del alcance de integral*. In: *Línguas e Instrumentos lingüísticos*, 29, 53-92.
- Lauria, Daniela (2011): *Apuntes para una historia de la producción lexicográfica monolingüe en la Argentina: etapas del proceso de diccionarización y modalidades diccionarísticas entre 1870 y 1910*. In: *Boletín de filología de la Universidad de Chile*, XLVI, 105-151.

- Pascual, José Antonio y Olaguíbel, María del Carmen (1991): Ideología y diccionario. In: Ignacio Ahumada (editor): Diccionarios españoles: contenidos y aplicaciones. Jaén. Universidad de Jaén, 73-89.
- Pérez, Francisco Javier (2000): Diccionarios, discursos etnográficos, universos léxicos. Caracas. Universidad Católica Andrés Bello/Fundación CELARG.
- Van Dijk, Teun A. (1999): Ideología, una aproximación multidisciplinaria. Barcelona. Gedisa.

Meredith CHURCH
(University of Colorado at Boulder, USA)
meredith.church@colorado.edu

Unstressed Vowel Devoicing in Andean Spanish

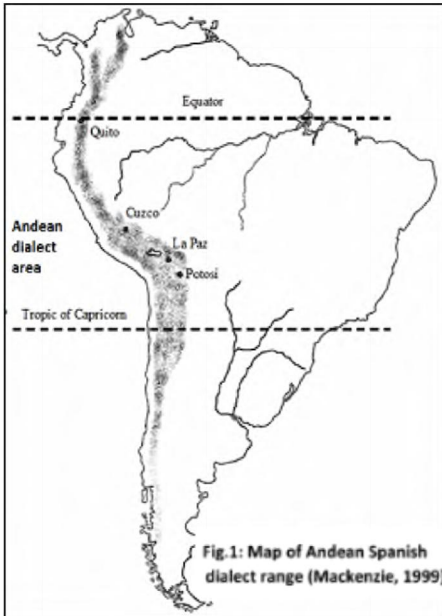
Abstract

There is no general consensus on where a principal linguistic norm-setting center for Spanish as a pluricentric language lies in the Andean region, though Cusco, Peru, is a possible candidate. Andean Spanish is a well-documented variety spoken in this area, and unstressed vowel devoicing is a contact feature in Andean Spanish that can be traced back to Quechua origin. This study analyzes vowel devoicing among sixteen Cusqueñan Spanish speakers, assessing the prevalence and patterning of this phenomenon as a characteristic of a non-dominant variety of Spanish. Findings suggest that the presence of devoicing is best explained by sociolinguistic factors, while the degree of devoicing has more to do with phonetic factors. Overall, Cusqueñan speakers exhibit low rates of devoicing, so this feature may not be representative of an Andean standard of Spanish centered in Cusco.

1. Introduction – Andean Spanish

Andean Spanish is a variety of Spanish spoken throughout the Andes of South America, resulting from centuries of contact with indigenous Andean languages, such as Quechua and Aymara (Cerrón-Palomino, 2003). Figure 1 below shows where Andean Spanish is spoken: the shaded area represents the Andes mountain range, which extends from Chile and Argentina in the south to Colombia in the north, but Andean Spanish is primarily concentrated in Ecuador, Peru, and Bolivia.

Escobar's (1992) comparison of Andean Spanish and Quechua bilingual Spanish has demonstrated that Andean Spanish is a unique dialect in its own right, meaning that monolingual speakers of Andean Spanish will demonstrate linguistic features borrowed from Quechua, but these contact features are not identical to the Quechua influence found in the Spanish of native Quechua-Spanish bilinguals. Escobar's work primarily focuses on the morphosyntactic features of Andean and Quechua bilingual Spanish that can be traced to Quechua: for instance, the double-marking of the genitive in possessive con-



structions (grammatically obligatory in Quechua) is found in both Andean and bilingual Spanish, but the possessor +possessed order of constituents (found in Quechua) in such double-marked constructions is unique to Quechua bilingual Spanish.

Nonetheless, it is important to note that features of Andean Spanish resulting from language contact can be identified at all levels of language. This study addresses a salient phonological feature of Andean Spanish that has been linked to Quechua: unstressed vowel devoicing, whereby vowels in unstressed syllables are devoiced, shortened in duration, and/or fully elided.

For example, *entonces* /en'ton.sɛs/ 'then' may be pronounced [en'ton.sɛs] or [en'ton.s] in Andean Spanish. See Figures 2 and 3 below for examples of devoiced and elided tokens from this study's data set.

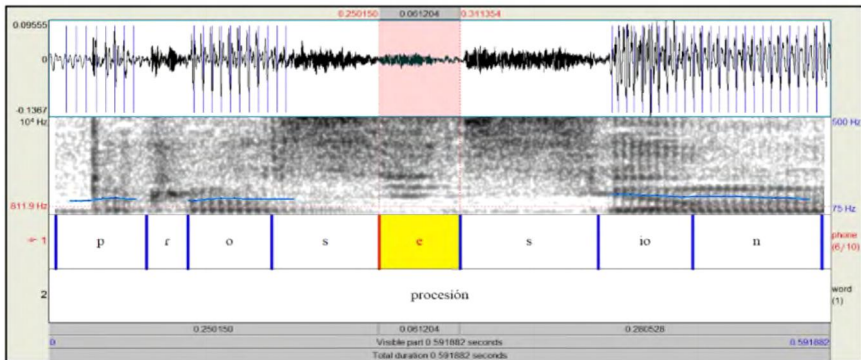


Figure 2: A devoiced but not elided vowel (/e/) between two /s/ segments. Compare its F_0 to the F_0 of the other vowels in this word: the voicing is practically non-existent for the /e/ vowel.

Hardman-de-Bautista (1982) has even noted occasional vowel devoicing (though rarely deletion) after voiced consonants, and several scholars (Hardman-de-Bautista, 1982; Hundley, 1986; Lipski, 1990) have observed that this pattern tends to occur in vowels adjacent to /s/, which is almost always maintained in Andean Spanish.

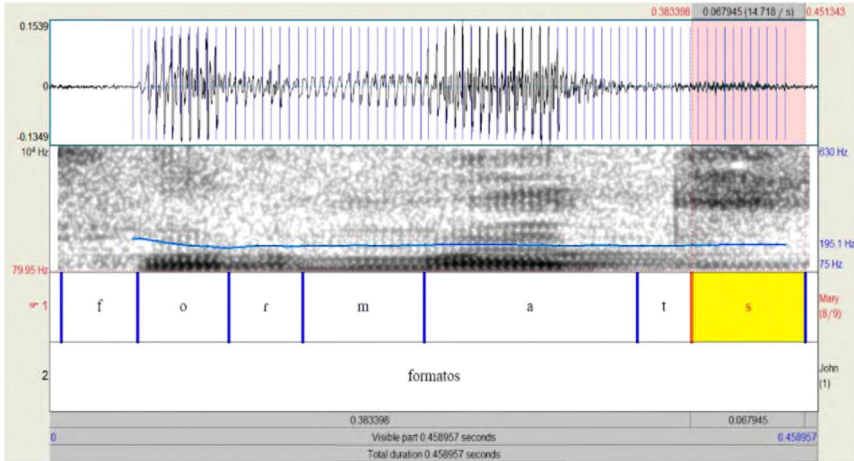


Figure 3: An apparently elided vowel (/o/) between a voiceless dental stop and /s/. Perceptually and acoustically, it is challenging to identify the presence of a vowel or any voicing between these segments.

The combination of these phonological patterns, unstressed vowel devoicing and /s/ maintenance, results in a particularly salient dialectal feature of Andean Spanish, which contrasts with a phonological pattern often found in coastal areas and large cities throughout Latin America, such as Lima, Peru, in which unstressed vowels are maintained while word-final /s/ is aspirated or deleted (Hardman-de-Bautista, 1982; Hundley, 1986). A clear example of this contrast is provided in Hardman-de-Bautista (1982): “A phrase like /juguetes, pues/ (meaning:) ‘well, toys’ would be [hugetsps] in Bolivian, but [hugetepwe] in Cuban (Spanish)” (147).

Hundley (1986) points out that this phenomenon of vowel devoicing and especially deletion adjacent to /s/ in the Andean dialect of Peruvian Spanish actually violates Spanish constraints on syllable structure: namely, the syllable-initial s-clusters that may result from this pattern are prohibited in Spanish. Although Colina (2019) points out that, cross-linguistically, /s/ can appear at syllable boundaries as an exception to the sonority contour, this syllable-initial pho-

notactic restriction is maintained in most varieties of standard, non-Andean Spanish. The restriction can be explained in terms of sonority, as “/s/ does not provide enough of a sonority rise with respect to the following obstruent” (Colina 2019, 137). Furthermore, according to standard Spanish syllable structure, “all syllables have a nucleus (N), which in Spanish must be vocalic;” however, as in the example provided above, the CVC syllable structure of *pues* becomes CC, [ps], in Andean Spanish (Colina 2019, 132). This phonotactic violation calls for a distinct explanation, such as substrata influence. And, in fact, Hundley finds evidence of a similar pattern of vowel devoicing occurring in Peruvian Quechua, which, unlike Spanish, is a stress-timed language. Delforge (2011) confirms this suggested link in a study of monolingual Cusco-Collao Quechua speakers, which demonstrated stronger devoicing of vowels between voiceless consonants. Lastly, this phenomenon has a certain level of social awareness in Spanish-speaking Andean regions: Delforge’s (2012) study of Andean Spanish in Cusco, entitled “Nobody wants to sound like a *provinciano*,” meaning someone from the rural, often poor, and largely Quechua-speaking provinces of Peru, highlights these Spanish-speaking Cusqueños’ awareness of this pattern of vowel devoicing as a salient characteristic of Quechua itself, and as a stigmatized feature of Quechua-accented Spanish. Thus, the link between this pattern in Andean Spanish and Quechua that has been demonstrated in linguistic literature is also widely acknowledged among Spanish speakers in Andean regions, though they associate it with bilingualism.

2. Situating the Andean Variety within Spanish as a Pluricentric Language

Pluricentric languages are understood to have overall standardization and codification, without being monocentric. These languages also have multiple national and regional standards. Thus, Spanish obviously fits this definition of a pluricentric language, though its situation is considerably complex. Oesterreicher (2002) points out that while many other pluricentric languages correspond fairly well to national or regional boundaries, that is not the case for Spanish, which has coexisting, scalar standards: these could be Pan-Hispanic, Pan-American, regional, or local, and these levels are gradient and occasionally overlapping. As an example, Oesterreicher (2002) suggests that what may be a standard feature at the national level for Peru could also correspond to the standard in all Andean countries, or even extend beyond South America.

Thompson (1991) identifies several Spanish-speaking zones of Latin America with their own linguistic norms. Thompson concludes that the Andean zone, including western Colombia, Ecuador, Peru, Bolivia, and northern Chile, has no apparent linguistic norm-setting center. This is due in large part to the contrasting Andean and coastal dialects of Spanish even within the same nation: for instance, the difference between Cusco (Andean) Spanish and Lima (coastal) Spanish within Peru inhibits the national or international codification of a particular variety. Nonetheless, Thompson specifically points out the preservation of post-vocalic /s/ as a salient feature characteristic of the more linguistically cohesive Andean/mountain region and addresses the cultural importance of Cusco (calling it “the old Inca capital and still the capital for many of the Andean Indians”), hinting that Cusco may be a viable norm-setting center for the Andean zone (49).

In trying to place Andean Spanish as a pluricentric variety of Spanish, it best seems to fit Level 5 in Muhr’s (2012) stages of pluricentricity, which encompasses “languages where the status of pluricentricity is acknowledged by the ‘mother’-variety, where the linguistic characteristics are codified including the minor varieties to some degree in dictionaries and reference books,” including some varieties of Spanish. The linguistic features of Andean Spanish are indeed codified to some degree in reference books, such as the 2003 reference grammar of Andean Spanish by Cerrón-Palomino, as well as in numerous linguistic studies. There is widespread acknowledgement of Andean Spanish and its features within and beyond academia, with speakers demonstrating particular awareness of those features that contrast with the coastal dialects and/or have a direct and clear connection to indigenous languages. In some cases, there is stigma attached, as demonstrated in Delforge’s (2012) “Nobody wants to sound like a *provinciano*” study of Cusco Andean Spanish. But, in spite of this acknowledgement and published literature, Andean Spanish has not been officially codified on a national or international level.

The broader research question this work contributes to answering is whether Andean Spanish is simply a dialect, or whether it could be an established non-dominant variety of Spanish as a pluricentric language. Even as a dialect, Andean Spanish may still have a role to play in the development of regional linguistic norms for Spanish as a pluricentric language. Caravedo (2014) discusses the role of perception in normativity, outlining a theoretical shift from language as occupying geographic/geosocial space to mental space. At the heart of this conceptualization of mental linguistic space is the subjective and arbitrary nature of perceptions of linguistic varieties as they correspond to spaces

understood by speakers to be fixed entities. Dialect contact, through migration or other processes, can lead to changes in speakers' perceptions, and, by extension, new and different values for linguistic norms. In this way, continued contact between Andean Spanish and other Latin American varieties of Spanish in Cusco, in Peru, and in other countries, may gradually lead to changes in its status as a non-standard variety of Spanish.

This project is not designed to tackle such a broad question, so the present exploratory study specifically answers the following questions: Using this pattern of devoicing as a representative feature of Andean Spanish, how consistent is it among the sample examined here (is there uniformity in the way the Cusco Spanish speakers in the sample demonstrate this pattern of vowel devoicing that is so characteristic of Andean Spanish), and what does that indicate about the Andes as a pluricentric zone of Spanish with Cusco as a linguistic norm-setting center? Additionally, is the variation in devoicing in the data predictable based on sociolinguistic factors, or is it random?

3. Data

The vowel tokens analyzed in this project were extracted from 16 audio-recorded interviews conducted by the author in Spanish with speakers residing in the Wanchaq district in Cusco, Peru, in May 2019. These formal interviews all followed the same structure, consisting primarily of questions about local rituals, legends, and dreams (adapted from Andrade Ciudad's (2007) master's thesis on Andean Spanish), with some additional tasks used to elicit narration. The audio was recorded on an iPhone 5, usually in the participants' homes or places of work, and the author did not measure or control for the distance between the microphone and the speaker for each recording. While the audio was not normalized for volume or pitch, voicing and silence thresholds were adjusted manually for each speaker. The recorded interviews were later transcribed with the audio/video annotation software ELAN.

The participants in this sample were nine women and seven men, whose ages ranged from 21 to 81 years old, with the average age being 49. All participants were Peruvians currently residing in Cusco, but they were not all Cusco natives. They had diverse occupations and social classes (though this information was not explicitly collected), and most had completed some level of higher education. All of them were native Spanish-speakers, but four of them were also Quechua-Spanish bilinguals, six of them were non-native Quechua-Spanish bilinguals, and the remaining six were Spanish monolinguals. Given Escobar's

(1992) distinction between Andean Spanish and Quechua bilingual Spanish, the initial goal was to limit the sample to Spanish monolinguals (to be able to make claims exclusively about Andean Spanish), but time and resource constraints resulted in the inclusion of Quechua bilingual Spanish speakers in the sample as well.

4. Methods

The recorded and transcribed interviews were first phonetically aligned using Wilbanks' (2021) Spanish forced-aligner, *faseAlign*. Next, a Praat (speech analysis software) script was developed to select all usable tokens from the phonetically aligned audio, and the boundaries of each token were inspected and manually adjusted. Any tokens that contained background noise that acoustically interfered with the identification of the vowel or obscured vowel formants and adjacent segments visually on the spectrogram were rejected. Given the time constraints for this project, it was not feasible to examine all unstressed vowels, so a more limited environment was selected: all vowels, including diphthongs, in unstressed syllables after voiceless consonants and before /s/ were extracted for analysis. This environment was selected because of the general consensus that this pattern occurs between voiceless consonants, and adjacent to /s/ but especially before (Hardman-de-Bautista, 1982; Hundley, 1986; Lipski, 1990), so this environment would be likely to yield devoicing while generating a realistic number of tokens to analyze given the time frame for this project. Diphthongs were included in the sample because previous studies had either explicitly excluded diphthongs (Delforge, 2012) or simply not addressed them in their analyses (Delforge, 2008; Lipski, 1990). Tokens meeting the environmental criteria above across word boundaries (e.g. in the words *las estrellas*, 'the stars', the initial /e/ in *estrellas* is between /s/'s and in an unstressed syllable) were also included, as no previous studies explicitly mentioned analyzing such tokens. These selection criteria yielded 1,611 usable tokens, which were analyzed with a Praat script that generated duration and percent devoicing information for each vowel. The data were compiled and analyzed in the statistical computing software R (R Core Team, 2021), with mixed-effects regression models (using the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages).

The models included the phonetic variables of vowel quality and environment, and the following sociolinguistic explanatory variables: speaker age, gender, Quechua fluency (native bilingual, non-native bilingual, or Spanish monolingual), whether the subject had Quechua-speaking parents, whether the sub-

ject had lived in a rural, Quechua-speaking area outside of Cusco, whether the subject had lived in a coastal, non-Quechua-speaking area, and an interaction term between age and gender, with speaker as a random intercept. Because the data set contained many fully voiced vowels, a logistic regression model including the above factors was first tested on the data set, with full voicing as the response variable.

This model identified which factors explain full voicing vs. any devoicing. This model's fit, as calculated by a theoretical conditional pseudo- r^2 value, was 0.229, indicating that it explained roughly 23% of the variance in devoicing. Subsequently, the fully voiced tokens were removed from the data set, leaving a data set with only vowels that were devoiced to some extent. This data set was not normally distributed and was thus log-transformed. A linear mixed-effects regression model including all of the same explanatory variables was tested on this transformed, only-devoiced data set to identify which factors explain more vs. less devoicing. This model had a conditional pseudo- r^2 value of 0.130, indicating that the model explained about 13% of the variance in devoicing across vowel tokens that show any devoicing. The fact that these fit estimates are lower than those found with the logistic regression model to explain variance in devoicing means that the variables used in both of these explanatory models are better at predicting whether there will be any or no devoicing than at predicting the precise amount of devoicing. Thus, it seems that degrees of devoicing are determined by more complex phenomena which could be identified by better models with additional and different explanatory variables (especially a more robust set of phonetic factors, as discussed below).

5. Results and Discussion

In summary, the primary data analysis was conducted with two models: a logistic regression model that identified the factors associated with full voicing and with any devoicing, and a linear model that identified the factors associated with a higher or a lower degree of devoicing among the vowels that were devoiced to any extent. Turning first to the logistic regression model, the significant effects ($p < 0.05$) with the greatest positive impact on zero devoicing were having a Quechua-speaking parent (estimate = 1.180, SE = 0.221, z-value = 5.327), having lived in a coastal area (estimate = 1.049, SE = 0.257, z-value = 4.080), and the vowel /e/ (estimate = 0.532, SE = 0.158, z-value = 3.360), followed by the significant ($p < 0.01$) negative effect of being a native Quechua-Spanish bilingual

(estimate = -0.894, SE = 0.318, z-value = -2.811). All phonetic environmental factors and all other sociolinguistic factors were insignificant.

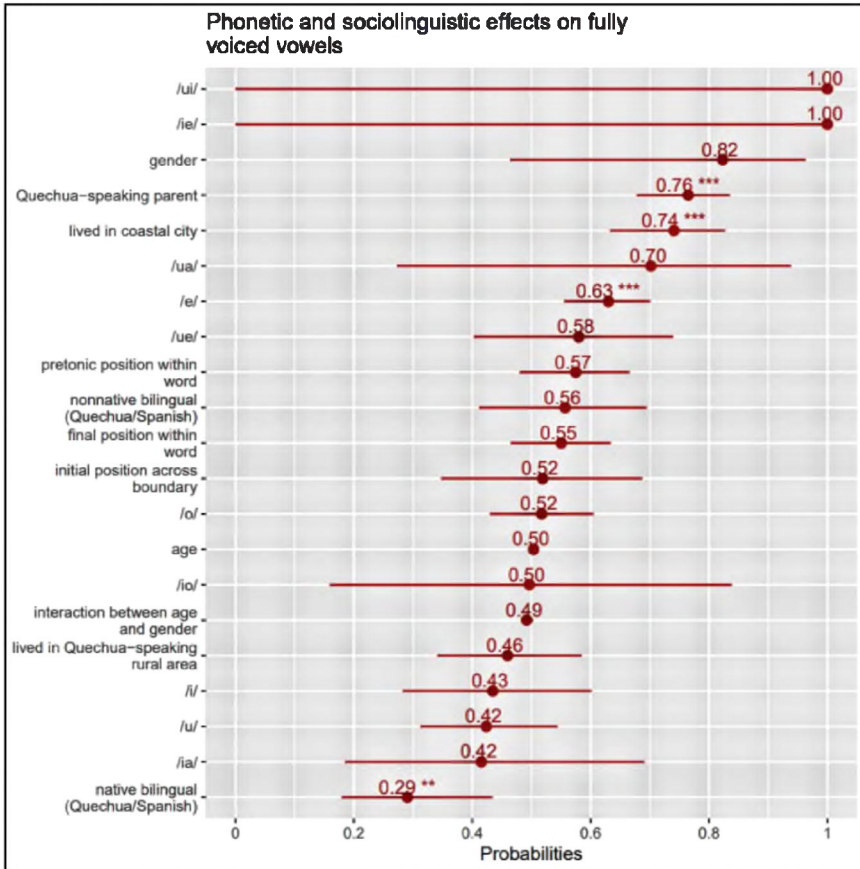


Figure 4: Plot of probabilities and significance of all phonetic and sociolinguistic effects in the logistic regression model on full voicing (a generalized linear mixed model with a logit link function). Horizontal bars represent error. Sample size $n = 1,611$:

Figure 4, created with the *sjPlot* (v2.8.9; Lüdtke, 2021) package, summarizes the probabilities of zero devoicing (i.e. full voicing) and significance of each effect in this logistic regression model on the presence or absence of any devoicing, with horizontal bars representing error.

Beginning with the only significant negative effect, it is quite understandable that being a native Quechua-Spanish bilingual would not be associated with

full voicing, (and, by extension, be associated with devoicing, at least to some degree), as this feature of vowel devoicing has been identified in Quechua and is thought to have become a feature of Andean Spanish due to contact with Quechua (Delforge, 2011; Hundley, 1986). The probability of this effect on zero devoicing is 0.29, which means the likelihood that a native bilingual would fully voice a given vowel in this sample is much less than chance. However, considering Escobar's (1992) claim that Quechua bilingual Spanish is actually a distinct variety from Andean Spanish, this finding suggests that devoicing, to any extent, may be more characteristic of Quechua bilingual Spanish than a standard feature of Andean Spanish.

The only significant phonetic variable was the /e/ vowel, positively correlated with full voicing. This finding is interesting, as other scholars (Lipski, 1990; Delforge, 2008) have claimed that /e/ is the vowel most often devoiced in several positions, and prone to devoicing in general, though they suggest that this finding may be somewhat inflated by the frequent discourse markers *pues* 'well' and *entonces* 'then.' There was not a particularly high rate of occurrence of these discourse markers in the data set analyzed in the present study, as these were interviews conducted by a proficient but non-native Spanish speaker in a formal setting, rather than casual conversational data. The /e/ vowel, the most frequently occurring vowel in Spanish, may be more likely to devoice because it is [+coronal] and thus will have more gestural overlap with adjacent /s/ and /t/, which are the most frequently occurring voiceless consonants in Spanish (Delforge, 2008). Returning to the results of the phonetic variables examined in this study, none of the environment variables was significantly correlated with zero devoicing, which indicates that a given speaker's tendency to devoice vowels at all is better predicted by other factors (largely sociolinguistic; see below) than by the phonetic context of the vowels, and that, in general, phonetic variables are less important factors in explaining whether a vowel is fully voiced or devoiced to any extent.

The significant positive sociolinguistic effects included having a Quechua-speaking parent and having resided in a coastal, non-Quechua-speaking area. Because being a native Quechua-Spanish bilingual is negatively correlated with full voicing, it is an interesting contrast that having a Quechua-speaking parent would positively correlate with full voicing. While one cannot conclusively explain why this correlation appears in the data, Delforge's (2012) article "Nobody wants to sound like a provinciano" discusses the widespread association of this feature of unstressed vowel devoicing with rural, 'backward' Indigenous peoples.

The existence of such stigma may explain the tendency for speakers whose parents are native Quechua speakers to fully voice vowels in this environment conducive to devoicing: aware of the stigma attached to this feature in their parents' speech, speakers may be trying to distance themselves from negative stereotypes by eliminating or reducing the feature from their own speech. This interpretation would certainly argue against devoicing being an emerging standard feature of this variety. Regarding the other significant sociolinguistic effect, the unsurprising finding that having lived in a coastal, non-Quechua-speaking area would positively correlate with full voicing speaks more to the features characteristic of coastal Peruvian Spanish than Andean Spanish. As described previously, while unstressed vowels, especially adjacent to /s/, tend to be devoiced or deleted while the /s/ is maintained in Andean Spanish, a contrasting phonological pattern is often found in coastal areas and large cities, such as Lima, Peru, in which unstressed vowels are maintained while word-final /s/ is aspirated or deleted (Hardman-de-Bautista, 1982; Hundley, 1986).

Thus, with regard to full voicing, it seems that sociolinguistic variables do more of the explanatory work than the phonetic variables. Within the transformed subset of data that contained only devoiced vowels, however, a different pattern emerges. The significant effects ($p < 0.05$) with the greatest positive impact on devoicing were all phonetic: the vowel /i/ (estimate = 8.114e-01, SE = 1.474e-01, t-value = 5.505), the vowel /u/ (estimate = 3.167e-01, SE = 1.024e-01, t-value = 3.092), the word-final position (estimate = 2.999e-01, SE = 8.350e-02, t-value = 3.591), and the word-initial position across word boundaries (estimate = 4.076e-01, SE = 1.734e-01, t-value = 2.351). The only significant ($p < 0.05$) sociolinguistic effect on devoicing was age, which had a small, negative impact (estimate = -2.110e-02, SE = 7.101e-03, t-value = -2.971). Figure 5 summarizes the estimates and significance of each effect in this linear regression model on vowel devoicing.

In this case, it appears that the phonetic variables are more important in explaining the amount of devoicing that occurs in non-fully voiced vowels. The finding that /i/ and /u/ are positively correlated with devoicing confirms findings in previous literature. Delforge's (2011) study of Cusco-Collao Quechua vowel devoicing identifies these same two vowels as the most affected, especially adjacent to uvular consonants, accounting for this devoicing with articulatory predictors. Both Lipski's (1990) study of Ecuadorian Spanish vowel devoicing and Delforge's (2008) study of Cusqueñan Spanish vowel devoicing found that /i/ and /u/, in addition to /e/, were most often devoiced in the word-medial/pretonic

position. Gordon's (1998) typological survey of vowel devoicing across approximately 50 different languages also identifies a cross-linguistic tendency for high vowels, namely /i/ and /u/, to be more affected by devoicing (98).

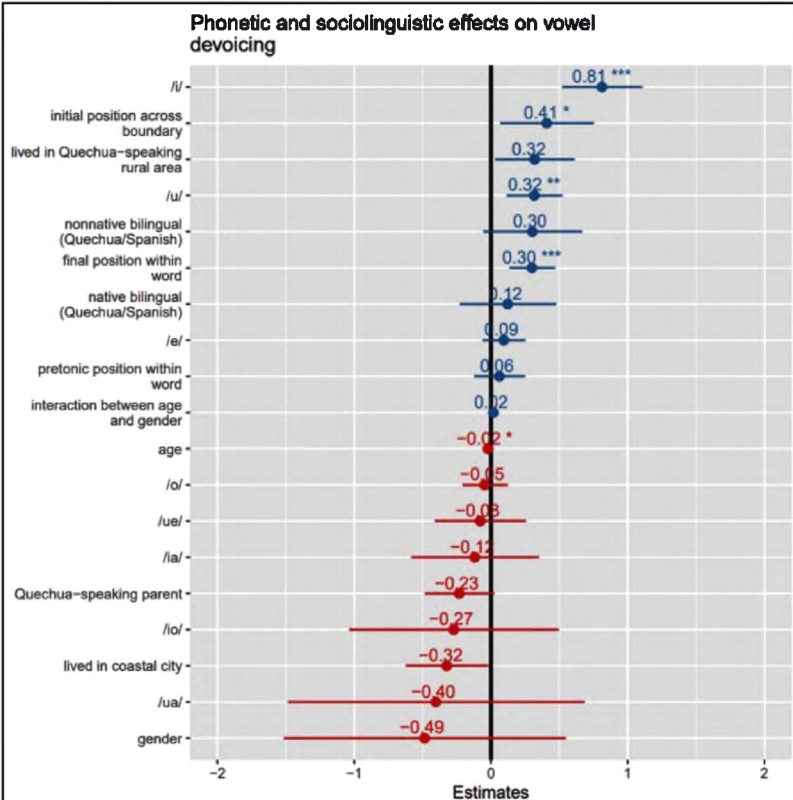


Figure 5: Plot of estimates, significance, and error of all phonetic and sociolinguistic effects in the linear mixed-effects regression model (with speaker as random intercept) on devoicing based on the non-fully voiced data subset ($n = 1,172$).

Delforge (2008) suggests that high vowels' close oral constriction, which raises air pressure in the oral cavity, inhibiting transglottal airflow, may be one reason these vowels are most likely to be affected by devoicing, though Gordon's (1998) theory is more concerned with shortened duration. While of course unstressed vowels are shorter than stressed vowels (pertinent to this phenomenon in Andean Spanish), Gordon also notes that high vowels are shorter than non-high vowels. Thus, these vowels will be more susceptible to gestural overlap—

the increased glottal opening gestures from adjacent voiceless consonants may result in more of the overall duration of these shorter high vowels being devoiced.

The finding in the present study that none of the diphthongs showed any significant correlation with devoicing suggests that previous studies had good reason to disregard them in their analyses of Andean Spanish. With regard to environment, it is unsurprising that the word-final position correlates positively with vowel devoicing, as this environment has been identified as being particularly conducive to devoicing by Hardman-de-Bautista (1982) and Delforge (2008), among others. What is surprising, however, is that the word-initial environment would be such a relatively strong positive effect on devoicing. As mentioned earlier, previous studies did not examine cross-word-boundary environments, so there was no precedent to inform expectations for this variable. Of the sociolinguistic effects on vowel devoicing, age was the only significant effect, with a small, negative impact, indicating that older speakers tend to devoice slightly less of the overall duration of the affected vowel than younger speakers.

This differs from Delforge's (2012) sociolinguistic study of vowel devoicing in Cusqueñan Andean Spanish, which found the lowest rates of vowel devoicing in younger Cusqueñans. These conflicting results may be related to differing criteria for devoicing: whereas this study takes a more gradient approach measuring devoicing as a percentage, in Delforge's (2012) study "all vowels that were identified as affected by the devoicing process, whether partially or completely, were classified as devoiced" (323). Thus, Delforge's (2012) finding regarding age may apply more to our previous question answered with the logistic regression model than this question addressing extent of devoicing. This supports the emerging conclusion of this project that sociolinguistic effects are better suited to explaining whether devoicing will be present at all or entirely absent, while phonetic effects account for more of the explanation behind amount of devoicing present in devoiced vowels. Though age was the only significant sociolinguistic effect identified in this model, the confidence intervals of the rural and coastal residence variables were [0.128, 0.527] and [-0.563, -0.143], respectively.

Thus, it would appear that having lived in a rural, Quechua-speaking area positively correlates with the extent to which a given vowel is devoiced; this confirms previous studies as well as Andean Spanish speakers' stereotypes or at least associations with this feature and rural Quechua-speaking populations. Likewise, if having lived in coastal, non-Quechua-speaking areas negatively correlates with vowel percent devoicing, this would confirm the observations that

coastal dialects in Latin America tend to demonstrate a contrasting phonological pattern in the environments where Andean Spanish vowels are devoiced, as discussed earlier.

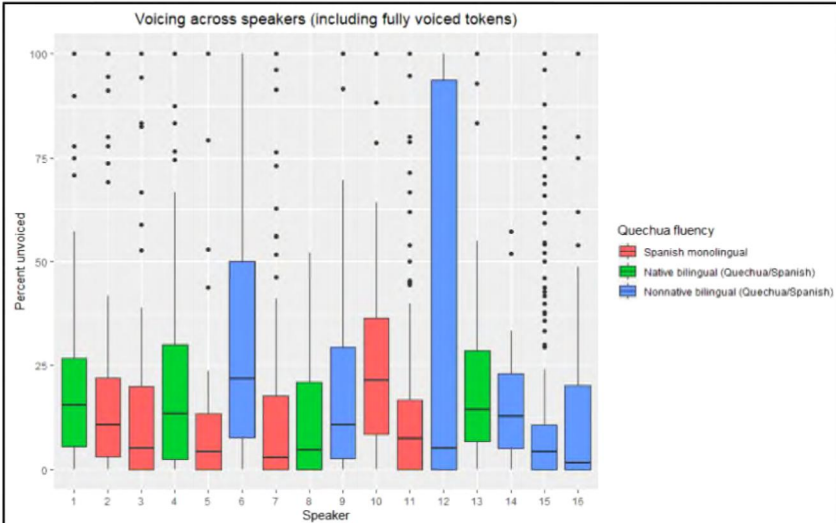


Figure 6: Box plot depicting voicing across speakers based on the full data set. Each bar represents an individual speaker, color-coded by the Quechua fluency variable. Non-native bilingual indicates a native Spanish monolingual who learned Quechua at a later age.

While these variables were not found to be significant at the $p < 0.05$ alpha threshold, these could be found to be significant effects in a larger sample of speakers. Lastly, for a general overview of devoicing in this sample, Figure 6 shows that the median rates of devoicing were fairly consistently low, lacking any major differentiation along the Quechua fluency variable.

6. Conclusion

Drawing together the findings of both models used in the primary analysis of these data, the overall result appears to be that the sociolinguistic factors tend to do a better job of explaining whether vowels will remain fully voiced or be devoiced at all, while the phonetic factors seem to do a better job of explaining to what extent non-fully-voiced vowels will be devoiced. This suggests that the degree of devoicing is not so much sociolinguistically motivated as explained by natural articulatory factors, but that social variables do play a role in whether a given speaker devoices unstressed vowels at all. The findings in this study per-

taining to the coastal residence variable confirm the differentiation between Andean Spanish and the coastal variety. This contrast complicates the possibility of codifying a national variety of Spanish in Andean nations such as Peru, where both Andean Spanish and coastal Spanish are prevalent. The finding that native Quechua-Spanish bilinguals more frequently devoiced tokens in the examined environment may point to this feature being characteristic not necessarily of Andean Spanish, but rather of bilingual Spanish. While Escobar (1992) has documented some morphosyntactic features that differentiate the two varieties, Quechua bilingual Spanish remains an uncoded, less well-documented and established variety than Andean Spanish which likely does not have a norm-setting center. The positive correlation between having a Quechua-speaking parent and fully voicing vowels in these environments that have been shown to be conducive to devoicing suggests the possibility that there could be a social stigma affecting how Spanish speakers in the Andes employ or eliminate this feature from their own speech. Given that speakers appear to be aware of this feature, associating it with rural and specifically Quechua bilingual speech along with certain negative stereotypes, there may be some desire to intentionally reject this feature. This may be a barrier to a variety of Spanish including unstressed vowel devoicing becoming codified in the Andes. Above all, because the sample of Cusqueñan Andean Spanish speakers in this study as a whole did not demonstrate consistently high rates of devoicing, the evidence in this study indicates that most Cusqueñan Spanish speakers associate fully voiced or less-devoiced vowels with the target pronunciation of vowels in Andean Spanish. Thus, if Cusco is a linguistic norm-setting center for Spanish in the Andes, vowel devoicing is not likely a characteristic perceived as standard in this variety.

References

- Andrade Ciudad, L. (2007): Usos de *dice* en castellano andino: Estrategias evidenciales y narrativas en contacto con el quechua. [Master's, Pontificia Universidad Católica del Perú].
- Bates, D. / Maechler, M. / Bolker, B. / Walker, S. (2015): Fitting Linear Mixed-Effects Models Using *lme4*. *Journal of Statistical Software*, 67(1), 1–48.
- Caravedo, R. (2014): Percepción y variación lingüística: Enfoque sociocognitivo. In: *Iberoamericana Vervuert*.
- Cerrón-Palomino, R. (2003): *Castellano andino: Aspectos sociolingüísticos, pedagógicos y gramaticales* (1. ed). Pontificia Universidad Católica del Perú, Fondo Editorial : GTZ, Cooperación Técnica Alemana.
- Colina, S. (2019): Phonotactic constraints on syllable structure. In: Colina, S., & Martínez-Gil, F. (Eds.), *The Routledge Handbook of Spanish Phonology* (1st

- ed.), 131-144. Routledge.
- Delforge, A. M. (2012): 'Nobody wants to sound like a provinciano': The recession of unstressed vowel devoicing in the Spanish of Cusco, Perú. In: *Journal of Sociolinguistics*, 16(3), 311-335. <https://doi.org/10.1111/j.1467-9841.2012.00538.x>
- Delforge, A. M. (2011): Vowel devoicing in Cusco Collao Quechua. In: 17th International Congress of Phonetic Sciences, 556-559.
- Delforge, A. M. (2008): Unstressed Vowel Reduction in Andean Spanish. In: L. Colantoni & J. Steele (Eds.), *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 107-124.
- Escobar, A. M. (1992): El español andino y el español bilingüe: Semejanzas y diferencias en el uso del posesivo. In: *Lexis*, 16(2), 189-222.
- Gordon, M. (1998): The Phonetics and Phonology of Non-modal Vowels: A Cross-linguistic Perspective. In: *Proceedings of the Berkeley Linguistic Society*, 24, 93-105.
- Hardman-de-Bautista, M. J. (1982): The mutual influence of Spanish and the Andean languages. In: *WORD*, 33(1-2), 143-157. <https://doi.org/10.1080/00437956.1982.11435729>
- Hundley, J. E. (1986): The Effect of Two Phonological Processes on Syllable Structure in Peruvian Spanish. In: *Hispania*, 69(3), 665-668. <https://doi.org/10.2307/342779>
- Kuznetsova, A. / Brockhoff, P. B. / Christensen, R. H. B. (2017): lmerTest Package: Tests in Linear Mixed Effects Models. In: *Journal of Statistical Software*, 82(13), 1-26.
- Lipski, J. (1990): Aspects of Ecuadorian vowel reduction. In: *Hispanic Linguistics*, 4(1), 1-19.
- Mackenzie, I. (2020): Andean Spanish. *The Linguistics of Spanish*. <https://www.staff.ncl.ac.uk/i.e.mackenzie/andean.htm>
- Muhr, R. (2012): Linguistic Dominance and Non-dominance in Pluricentric Languages: A Typology. In: R. Muhr (Ed.): *Non-dominant Varieties of Pluricentric Languages. Getting the Picture*. In Memory of Michael Clyne, 23-48.
- Oesterreicher, W. (2002): El español, lengua pluricéntrica: Perspectivas y límites de una autoafirmación lingüística nacional en hispanoamérica. El caso mexicano. In: *Lexis*, 26(2), 275-304. Pontificia Universidad Católica Del Perú. Departamento De Humanidades.
- R Core Team (2021): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Thompson, R. W. (1991): Spanish as a pluricentric language. In: M. Clyne (Ed.), *Pluricentric Languages: Differing Norms in Different Nations*. De Gruyter, Inc. 45-70.
- Wilbanks, E. (2021): faseAlign (Version 1.1.11) [Computer software]. Retrieved from <https://github.com/EricWilbanks/faseAlign>.

III. The Pluricentricity of Portuguese in the Americas and worldwide

Augusto SOARES DA SILVA

(Universidade Católica Portuguesa, Braga, Portugal)

assilva@ucp.pt

Portuguese, pluricentricity and Brazilian Portuguese: A case of a reverted asymmetry?¹

Abstract

This study explores the status of the Brazilian variety of Portuguese within the pluricentricity of the Portuguese language, and more specifically the hypothesis of a possible situation of reverted asymmetry with respect to the European variety. Three issues will be addressed: First of all, the evolutionary relation between all national varieties of Portuguese will be briefly outlined, especially the standardization of the Brazilian variety, the nativization of African varieties, and the increasing pluricentricity of the Portuguese language.

Secondly, the bicentricity of European and Brazilian standards will be highlighted, with particular emphasis on social factors and lectometric indicators that point towards a diachronic divergence between both varieties and a rare state of symmetric pluricentricity.

Finally, we will argue that the great influence of Brazilian culture in Portugal and other Portuguese-speaking countries, resulting from the proliferation of Brazilian audiovisual cultural products, is gradually leading to a situation of reverted asymmetry in the domain of language perception and attitudes and of culture, but not in the domain of language use.

1. Introduction

Portuguese is a pluricentric language with different national standard varieties, namely European Portuguese (EP), Brazilian Portuguese (BP) and African standards in development, especially in Mozambique and Angola. Regarding the two well-established standards, namely EP and BP, Portuguese is one of the few languages that come closest to the rare condition of *symmetric*

¹ This study has been carried out under the research project UIDB/00683/2020 (Center for Philosophical and Humanistic Studies), funded by the Portuguese Foundation for Science and Technology. I would like to thank two anonymous reviewers for their thorough and illuminating comments. Needless to say, the remaining errors are only mine.

pluricentricity. This is mainly due to the balance between the time supremacy of EP and the demographic supremacy of BP, the strong codification of the two national standards, and the increasing awareness of the international importance of the bicentricity of Portuguese in socio-political, economic and cultural terms.

The EP standard is also the reference norm in the Portuguese-speaking African countries (Angola, Cape Verde, Guinea Bissau, Mozambique, São Tomé and Príncipe, and, since 2011, Equatorial Guinea), as well as in East Timor and other Asian territories, especially the Macao Special Administrative Region in China. While Portuguese is receding in Asia, its position in Africa remains strong and is bound to increase in importance in the future. Demographic projection data show a considerable population increase by the end of the 21st century, especially in Angola and Mozambique. This will lead to an even more pluricentric standardization of Portuguese and eventually to a greater multilateral management of national standards.

The BP standard, in turn, is only valid in Brazil. Interestingly though, Brazilian culture, media and language enjoy wide diffusion in Portugal and in the other Portuguese-speaking countries. Moreover, Brazilian media have a great impact on the use of Portuguese in Africa. In contrast, exposure to the EP standard in Brazil is minimal. In fact, few Brazilians have any contact with spoken European Portuguese, meaning that Brazilians in general experience difficulty in understanding the spoken European standard.

In this study, we will address three issues. We will first briefly sketch the evolutionary relation between the different national varieties of Portuguese, with special emphasis on the standardization of BP, the nativization of African varieties (mainly Mozambican and Angolan Portuguese), and the increasing *pluricentricity* of Portuguese.

Secondly, we will focus on the *bicentricity* of Portuguese (that is, the European and Brazilian standards), with particular emphasis on social and linguistic factors that point towards a relatively rare state of *symmetric* pluricentricity, while also tackling some remaining asymmetries. The results of our socio-cognitive and lectometrical research, which pointed to a diachronic divergence between Brazilian and European standards (Soares da Silva 2010, 2014a, b, 2016, 2021a), will be highlighted.

Finally, and as the main point of this study, we will discuss the *reverted asymmetry* hypothesis, which suggests that the Brazilian standard is gradually

becoming the new dominant variety, although the older European standard still benefits from a lot of prestige.

2. From monocentricity to pluricentricity

The standardization of Portuguese started after the establishment of the Kingdom of Portugal in the 12th century, but the first normative codification instruments would only appear four centuries later in the 16th century. From the second half of the 19th century onwards, with the Brazilian Romanticism project of establishing a national literary language after Brazil's independence in 1822 (that is, exactly 200 years ago), a second pole emerged in Brazil, which developed its own relatively independent and divergent standard during the 20th century. As a result, the Portuguese language thus inherited two spelling systems, two sets of grammatical nomenclature, two academies and two standards with important differences in grammar and lexicon, competing with each other on the geopolitical stage. Over the last few years, the Community of the Portuguese-speaking countries (CPLP) and its language agency, namely the International Institute of the Portuguese Language, have tried to transform the Portuguese language policy by switching from a bipolar to a multipolar and multilateral management model (Oliveira 2015).

Oliveira (2016: 39) points out three main stages and periods in the Portuguese standardization process: (i) from a monocentric standardization in Portugal (16th to 19th century); (ii) to a bicentric standardization in Portugal and Brazil (19th to 20th century); (iii) to a pluricentric standardization in the middle of the 21st century in which all members of the CPLP (theoretically) partake.

It is important at this stage to look at some of the sociopolitical and structural characteristics of the standardization process(es) of the Portuguese language. Compared with other Romance languages, such as French and Spanish, the standardization of Portuguese developed at a slower pace and without the support of an actual language academy. Indeed, the Lisbon Academy of Sciences, founded in the 18th century in Lisbon, and the Brazilian Academy of Letters, created at the end of the 19th century in Rio de Janeiro, both played a minor role in the standardization process compared to the Royal Spanish Academy or the French Academy for Spanish and French respectively. The Portuguese and Brazilian academies have a Lexicography Commission and do intervene in orthographic policy, but they do not produce linguistic instruments of normative character. Moreover, the International Institute of the Portuguese

Language (an organ of the CPLP), except for the elaboration of the Common Orthographic Vocabulary (VOC), has not (yet) implemented an active standardization policy.

The standardization process is still fundamentally bicentric, with two major divergent poles in Portugal and Brazil. The two major production centres of linguistic standards are located in the main cities of Portugal – Lisbon and Porto – and in the major capitals of Brazil – Rio de Janeiro and São Paulo. Taking the example of the management of spelling conventions (see Gonçalves 2020), the first simplification of the Portuguese orthographic system took place in Brazil in 1907 and in Portugal, officially, in 1911 and was plagued by disagreement and divergence between two national graphic norms. The “1911 spelling rupture” created, among the Brazilian national elites, an awareness of linguistic independence in relation to Portugal, which stimulated the production of new dictionaries and grammars. In 1943 and 1945, two attempts were made to come to an agreement between both orthographic proposals, but both attempts failed due to legislative problems. The last Orthographic Agreement of 1990 (AO90) also brought about its own share of issues and uncertainties. Only in 2008 was it approved in Portugal and in Brazil, becoming effective in Brazil in January 2009. In both countries, periods of transition were taken into consideration, so that the AO90 became mandatory on May 13, 2015 in Portugal, and January 1, 2016 in Brazil. The AO90 is now official only in four countries of the CPLP (Portugal, Brazil, Cape Verde, and São Tomé and Príncipe), while it is still pending validation by the respective national parliaments in Mozambique and Angola. Crucially, the AO90 is a receptacle of graphic solutions that reflect orthoepic differences between the two well-established national standards, the European and the Brazilian.

The bicentric standardization also manifested itself in large-scale research projects on language variation (e.g. “Educated Urban Standard” begun in 1969, “Grammar of Spoken Brazilian Portuguese” since 1988 in Brazil and “Fundamental Portuguese”, which began in 1970 in Portugal) and in national educational projects (recent schooling projects in Brazil; the National Reading Plan in Portugal, launched in 2007), but also in the increasing number of separate, new, expanded grammars (normative and non-normative) in both Brazil and Portugal over the last fifteen years (see Soares da Silva 2020a); in the divergence between terminologies; in the “language industry” that pits the two national varieties against each other in international markets; in the digital

separation between these two varieties, etc. All these examples are good illustrations of the well-known Lusophone bicentricity (Aguilar e Silva 2007: 20).

Demographic projection data from Portuguese-speaking countries and data on the demolinguistic evolution of Portuguese up to the end of the 21st century show that there will be a major increase in population, hence in the number of Portuguese speakers in the Portuguese-speaking African countries, especially in Angola and Mozambique.

According to the United Nations World Populations Prospects in its most recent version, the 2019 Revision of WPP, the population of the African Portuguese-speaking countries will almost double by the end of the 21st century, increasing from 289,442 million in 2019 to 514,512 million in 2100. There will be a dramatic increase in Angola, from 31,825 million to 188,283 million, and in Mozambique, from 30,366 million to 123,647 million. At the same time, the Brazilian population will decrease from 211,050 million to 180,683 million and Portugal's from 10,226 million to 6,985 million.

Angola and Mozambique together will then represent 60.6% of the population of the Portuguese-speaking countries, while Brazil's percentage will drop to 35.1%. Considering the increase in speakers of Portuguese as L2 or L1 in the African Portuguese-speaking countries over the last three decades, it is expected that the Mozambican and Angolan population will speak Portuguese as L2 or L1 for the next 40 or 50 years.

The number of Portuguese speakers in the world is thus likely to duplicate by 2100, from 260 million currently to 515 million. This, in turn, will lead to a gradual development from a hegemonic bicentric to a truly pluricentric standardization process, with four main national standards – Brazilian Portuguese (BP), Mozambican Portuguese (MP), Angolan Portuguese (AP) and the very small but still prestigious European Portuguese (EP) – the emergence of new norm-setting centres and a truly multilateral management of national standards (Oliveira 2016).

The African varieties of Portuguese, especially MP and AP, emerged in the last two decades of the 20th century and were initially acquired as a second language, in a context marked by language contact with Bantu languages and by inexistent or minimal exposure to the EP standard. Most of the innovations displayed by these non-native varieties regarding EP result from the interference with Bantu languages (Gonçalves 2010, 2013, 2018; Inverno 2011, 2018; Mingas 2000). The current situation of MP and AP can be characterised by a tension between an ongoing process of nativization and EP normative

pressure. On the one hand, the EP standard is a largely unfamiliar variety to the majority of speakers, including teachers (Chimbutane 2018, Inverno 2018). The huge majority of teachers in Mozambique and Angola speak nativized varieties of their respective countries and are locally trained. Moreover, schools in Mozambique, Angola and other Portuguese-speaking African countries function as a place of crystallisation of local characteristics of Portuguese, as part of a process of reinforcing nativization of African varieties. On the other hand, MP and AP, as well as the other African varieties, being subject to a continuing degree of influence from EP, through post-colonial connections and the further development of education, might yet reasonably be expected to bear a stronger stamp of EP influence (Baxter 2018: 308).

3. Standardization in Brazilian Portuguese: between exonormative and endonormative standardization

The process of standardization of BP only emerged in the course of the 20th century and is marked by a tension between exonormative standardization based on the Portuguese literature of the 19th century, and endonormative standardization based on the emergent Brazilian variants. As a result, it is characterised by a considerable distance between the idealised and prescriptive traditional norm and the real norm (or norms) used in big city centres, as well as between written and spoken language, leading to a situation of diglossia. However, a gradual reduction of the marked Brazilian diglossia can nowadays be observed, due to the introduction of patterns of the spoken language into BP written language.

As far as exonormative influence is concerned, a powerful social imagery in Brazilian society still exists that leads journalists, intellectuals and teachers to wage social “wars” in defense of the dogmatic and immutable “standard” calqued on the literary European variety. Three typical social effects are: (i) the conservatism manifested in the style guides used by the main Brazilian newspapers, which merely transcribe the prescriptions laid down in the old grammar books; (ii) the proliferation and success of grammar columns in newspapers, which attempt to root out errors of all kinds; and (iii) the importance given to the European artificial standard in the national high school exams, associated to an elitist linguistic pedagogy that leads teachers to become fixated on “mistakes” and stigmatise all linguistic variants (Faraco 2001, 2008, 2011).

Beneath this linguistic purism lies an ideology of *identity* nationalism, which entails social exclusion, economic discrimination, and white supremacy. This has given rise to linguistic myths and prejudice about the inferiority of the language spoken in Brazil and the respective superiority of the language spoken in Portugal. Some of these myths, constructed in the 19th century and still clearly present today in the Brazilian media and schools, are: “EP is a pure unchanged language, while BP is borrowed and corrupt”; “Brazilians do not know how to speak Portuguese properly”; “Brazilians speak and write wrongly”; “Portuguese is very difficult”; “Brazilian is ungrammatical”; “it is ugly to speak like that because that is how the Indians speak”; “failure to use the ideal standard form of the language will damage one’s job prospects and therefore the opportunity to rise in life” (Bagno 1999, 2000; Scherre 2005). There is also the mythical idea of a single homogeneous language in the immense territory of Brazil and the so-called “veritable Brazilian miracle” – a myth reinforced by the anthropologist Darcy Ribeiro, who argues that the numerous immigrants were irrelevant in establishing the characteristics of Brazilian culture and the amazing “cultural homogeneity” of the Brazilian people (Ribeiro 1997). In short, this is the ideology of “national unity” founded in the romantic vision of “one nation, one culture, one language” (Soares da Silva 2015).

However, a widening gap between written EP and written BP, and a progressive decrease of the distance between oral and written BP, are currently noticeable. Crucially, endonormative Brazilian features are increasingly introduced into the written language. This is the case, for instance, of two syntactic changes, namely the loss of third-person clitic pronouns and the loss of the null pronominal subject. Overt subjects and the replacement of third-person clitics by other strategies are used by Brazilian speakers across social groups, regardless of gender, social class or level of schooling (see Duarte, Gomes & Paiva 2016, 2022).

Other Brazilian variants, though, still predominantly exist in colloquial BP or in the speech of less/uneducated people. This is the case, for instance, of the lack of nominal agreement in number, and of the lack of verbal agreement and the consequent strong simplification of the morphology of the verb. Another set of Brazilian variants only exist in non-urban dialects, chiefly in isolated communities. This is the case of the lack of gender agreement, for instance. The fact that some endonormative features are already incorporated into the written language, whereas others are still restricted to the spoken language, while at the same time exonormative features are still partially used as a reference, means

that current school education has to deal with features of two grammars simultaneously (Mattos e Silva 2004, Faraco 2008, Mendes 2016). It is also expected that an increasing number of endonormative Brazilian features will sooner or later be standardized and incorporated into the written language, including formal register. However, the increase in education level and the attitudes of linguistic purism mentioned above may prevent the standardization of some of these typically Brazilian variants.

4. The symmetric bicentricity of Brazilian and European varieties

Regarding the two established EP and BP standards, Portuguese is one of the few languages that comes closest to the rare condition of *symmetric pluricentricity*. Several social factors have favoured the symmetric bicentricity of BP and EP (see also Joseph 1987: 170 and Clyne 1992: 463). First of all, the historical advantage of Portugal is balanced out by Brazil's much larger population. Secondly, neither Portugal nor Brazil is currently a major political power. Both countries have nevertheless recently gained international prestige for different reasons, Portugal as a member of the European Union, also recently benefitting from an increasing popularity, as is shown by the substantial increase in (cultural) tourism, and Brazil owing to the international popularity of its television soap operas, music and football and as an emerging economic power. Thirdly, within the last 3 decades, dictionaries and grammar books have developed separately in Portugal and Brazil, resulting in an increasingly divergent codification of their respective standard varieties. Fourthly, Portugal and Brazil, as well as other Portuguese-speaking nations, are increasingly aware of the importance of the pluricentricity of Portuguese in socio-political, economic and cultural terms. Finally, the rationalist-utilitarian discourse of "unity in diversity" (or of the "superior unity" of the Portuguese spoken in Portugal and Brazil) has been reinforced by linguists and grammarians in both Portugal and Brazil, whereas the more recent concept of *Lusophony* is now used as a utilitarian political ideology of affirming a supra-national space of economic, cultural and linguistic identity (see Soares da Silva 2015). As the renowned Portuguese linguist Maria Helena Mira Mateus puts it, the Portuguese language is "a valuable investment capital, it is our way of manifesting difference" and "it is through Portuguese that Portuguese and Brazilian people create their own path in the world" (Mateus 2002: 42).

The symmetric bicentricity of BP and EP is not only conspicuous in the culture but also in the language. We have carried out cognitive-sociolinguistic

and lectometric research (Geeraerts et al. 2010) on the EP-BP bicentricity, thereby confirming a *divergent* evolutionary trend between the two national varieties over the last 70 years at the lexical, constructional and attitudinal levels (Soares da Silva 2010, 2014a, b, 2016, 2018, 2021a).

The lexical lectometric analysis (Soares da Silva 2010) was carried out for several thousand observations of the use of alternative terms designating 43 football and clothing nominal concepts, gathered from a large corpus of sport newspapers and fashion magazines from 1950s, 1970s and 2000s (4 million tokens), internet chats related to football (15 million), and labels and price tags displayed in shop windows (1,300 pictures). The database contains 90,202 observations of football terms used in sports newspapers and 143,946 observations of their use in Internet chats; 12,451 observations of clothing terms used in fashion magazines and 2,775 observations of their use on labels and price tags. The lectometric analysis showed that EP and BP diverge over time with respect to clothing terms but converge slightly when it comes to football terms. As clothing terms can be deemed more representative of everyday vocabulary, the results obtained in the domain of clothing are probably closer to the sociolinguistic reality. The slight convergence in the field of football is likely to be an effect of globalization and of the gradual standardization of the football vocabulary worldwide. The lexical lectometric analysis also provided other results: (i) diachronic divergence is observed as much in BP as in EP, which indicates *symmetric* pluricentricity; (ii) BP has undergone more changes than EP since the 1950s, probably as a result of the greater social variation in Brazil and the more recent standardization of BP; (iii) the influence of English and other foreign languages as revealed by the use of loanwords is stronger in BP than in EP (Soares da Silva 2014b); and (iv) the actual distance between the standard and the substandard strata is higher in BP than in EP.

The lexical divergence analysis between EP and BP was replicated at the level of constructional variation. A pilot study focusing on alternations between (i) prepositional constructions such as *falar de/sobre/acerca de/em* ‘speak about’, *interesse em/por* ‘interest in’, (ii) infinitival constructions with causative/perception verbs (*A Maria fez/deixou/viu os miúdos correrem vs correr vs correr os miúdos* ‘Mary made/let/saw the children run-pl/run-sg/run-sg the children’), and (iii) word order in noun-adjective constructions (N+A vs. A+N) show a decrease in uniformity from 1950s to 1970s and from 1970s to 2000s, thus indicating divergence over time (Soares da Silva 2014a).

In a very recent study on the alternation between the presence and the absence of the clitic *se* in the different types of *se* constructions (reflexive/reciprocal, middle, anticausative, passive, and impersonal), the same pattern of constructional divergence between BP and EP was attested (Soares da Silva et al. 2021). By describing the alternative *construals* conveyed by the alternation of overt and null *se* constructions, that is, *energetic* vs. *absolute* *construals* (Langacker 1987, 2008) respectively, through advanced quantitative multifactorial methods, the statistical analysis revealed a lectal difference, with BP characterised by the emergence of a novel construction, the null construction, which is of much more restricted occurrence in EP. It also showed that the focus on the moment of change (i.e., the profiling of the change of state) and language-external factors, such as register, are predictors of the overt/null variation. If the change and, therefore, the energetic elements are profiled, the outcome is usually the overt construction; otherwise, the event is conceptualised as a self-contained whole and the null construction is preferred. In a study in progress on the variation in prepositional relative constructions (i.e., pied-piping, chopping and resumptive relatives), the statistical analysis also points to BP/EP constructional divergence: the greater standardization of the chopping construction (i.e., without the preposition) in BP entails a restructuring of the conceptual and discursive space of relatives, particularly the cognitive accessibility, discursive relevance and conceptual prominence of the relativized head noun.

Attitudinal variables have also been considered. A survey was conducted concerning the use of alternative clothing terms and knowledge about their typical usage and origin. For both Brazilian and Portuguese respondents, the proportion of the total number of terms considered endogenous and binational is higher than the proportion of terms characterised as exogenous and binational, thus revealing a divergent attitude (Soares da Silva 2014a). A divergent attitude is manifested by a preference for endogenous terms (i.e. BP terms in the intentions of Brazilian respondents and EP terms in the intentions of Portuguese respondents) and binational terms. Another attitudinal study concerns *romantic* vs. *rationalist* cultural cognitive models towards the unity/diversity of EP/BP, which are valid for both Brazil and Portugal. The attitudes underlying discourse and the instruments of Portuguese standardization reveal the *romantic separating* ideology of claiming the Brazilian language as distinct from Portuguese and the *rationalist separating* attitude that has been present in the process of divergent standardization of European and

Brazilian varieties over the last century, in contrast with the *romantic unifying* attitude of reclaiming the purity of the (European) Portuguese language and the *rationalist unifying* ideology of promoting the superior unity of Portuguese in the current transcontinental global context and its economic and political benefits (Soares da Silva 2015, 2020a).

Moving from *formal* onomasiological variation (the choice of different expressions for the same conceptual category) to *conceptual* onomasiological variation (the choice of different conceptual categories for a referent), we have conducted research on the conceptual variation of emotions in BP and EP. The multifactorial profile-based analysis revealed that the differences in conceptualising anger and pride in BP and EP are intrinsically related to cultural collectivism vs. individualism differences, especially between the more individualistic, indulgent, and emotionally expressive culture of Brazil and the more collectivistic and restrained culture of Portugal (Soares da Silva 2020b, 2021b, 2022). Hofstede's (2001) cross-cultural comparison model shows cultural differences between Portugal and Brazil. Moreover, research carried out by Brazilian sociologists, anthropologists, and historians highlight emotionality as a defining feature of Brazilian identity and culture, i.e., Brazilian culture is characterised by a greater social acceptance of the display of emotions compared to other countries (Rezende & Coelho 2010).

5. Brazilian-European Portuguese asymmetries

Yet the symmetric bicentricity of BP and EP is not perfect. There are asymmetries in both the culture and the language, especially regarding language intelligibility.

First and foremost, the Brazilian variety benefits from high degrees of visibility on a popular level in Portugal and in the other Portuguese-speaking countries, due to the popularity of Brazilian television soap operas, music and football, as well as to the sharp increase in the number Brazilian emigrants since the beginning of this century, especially in the past four years. In contrast, the exposure of Brazilian speakers to the European variety is minimal. Brazilians in general have difficulty in understanding the spoken European standard, while the same cannot be said of the Portuguese in relation to the spoken Brazilian standard. Besides comprehension problems related to EP lexical, syntactic and pragmatic features, the greatest difficulty for Brazilians is to understand EP phonetics, especially the EP unstressed vowels, which underwent a strong rise, reduction and even disappearance in some cases. For example, Brazilians

perceive the EP pronunciation of the verb *telefonar* ‘to phone’ as only one syllable, although it contains four of them. Some concrete examples of non-comprehension of the EP standard from the Brazilian media include the subtitling of interviews with Portuguese youngsters on the MTV television channel, and the subtitling and dubbing into BP of Portuguese movies such as *Capitães de Abril*, based on the Revolution of 25 April 1974, and the television series *Equador*.

A second asymmetry, although to some extent embedded in the first one, concerns the fact that there is a very unequal relationship in the circulation of cultural goods (Lagares 2021: 73). During the 19th century and the beginning of the 20th century, cultural circulation took place from Portugal to Brazil. From the middle of the 20th century onwards, this direction was reversed with the proliferation of Brazilian cultural products conquering the global market and entering Portugal in large numbers. In comparison, there are nowadays very few Portuguese cultural goods that enter the Brazilian market. The asymmetry is also evident with respect to the press and the circulation of information: the Portuguese know a lot about Brazil, whereas Brazilians know very little about Portugal. Underlying this lack of knowledge is a negative view of Portugal deeply rooted in Brazilian society and still conspicuous in the press, in textbooks and even in cultural productions, such as films and soap operas. Brazil seems to be ashamed of its Portuguese heritage (see Fino’s 2021 communication science and historiographic study, which also points out a sense of mutual strangeness or discomfort that permeates the relationship between Portugal and Brazil). It is not by chance that Brazil does not have a public holiday to mark the arrival of the Portuguese on its territory.

A third asymmetry concerns the weight that the Portuguese tradition still has on the Brazilian space. A good illustration of this is the strong linguistic purism in Brazil mentioned above, which imposes a standard calqued on literary EP, condemns the use of foreign words and typically Brazilian constructions, spreads linguistic prejudice about the inferiority of the language spoken in Brazil, and uses the EP standard as an instrument of social exclusion.

A fourth asymmetry concerns the fact that Portugal founded the Camões Institute in 1992, whose mission is to promote the Portuguese language and culture worldwide, but there is still no equivalent institution in Brazil.

Finally, the language stratification disparity in Brazil also counts as an asymmetry in the BP-EP bicentricity. As demonstrated in lectometric terms in the previous section, the BP standard remains much more distant from the

reality and diversity of its spoken language than the EP standard (Soares da Silva 2010, 2014a).

6. The pluricentricity of Portuguese: A Brazilian reverted asymmetry?

The answer to the question about whether a process of reverted asymmetry between both national varieties can currently be observed, with the Brazilian standard becoming the dominant variety is: Yes, in language perception and attitudes as well as culturally, and No, with regard to language use. The Yes answer, however, can only be asserted with some reservations, since Portugal too still enjoys some prestige and cultural influence.

Four indicators can be pointed out to account for the Brazilian reverted asymmetry in culture and in language perception and attitudes.

First of all, the Brazilian cultural impact on Portugal and other Portuguese-speaking countries since the middle of the 20th century is evident in the proliferation of Brazilian audiovisual cultural products, especially music and television series and soap operas, as well as in the success of Brazilian football, food, and print media entering the Portuguese market massively. It is in fact an effect of the internationalisation of Brazilian cultural products, which is stronger in Portugal. The great knowledge and consumption by the Portuguese of Brazilian cultural goods is intrinsically related to the Portuguese familiarity with the Brazilian cultural market and to Brazil's international cultural diffusion through the massive export of cultural products, especially those related to mass entertainment. Brazilian popular culture is therefore well-known and widely consumed by the Portuguese, while the reverse cultural influence is currently minimal. More recently, watching Brazilian videos on YouTube has become a growing trend among Portuguese children due to the lack of a similar offering in Portugal, and has started to become prominent among teenagers as well ("Brazil is invading the vocabulary of young people" was the alarmist title of an article published in the Portuguese newspaper *Notícias Magazine* in September 2021, thereby expressing the feeling of fear prevailing in Portuguese society). There are, however, a few counterforces to the Brazilian cultural influence in Portugal. Portuguese television started producing successful Portuguese soap operas and series in the early 2000s, and the Brazilian soap operas, which have been extremely popular and influential since their entering Portugal in 1977, from now on must coexist and even compete with Portuguese soap operas. The production of Portuguese soap operas more than doubled from the 1900s to the 2000s, while the presence of Brazilian ones in Portugal in the same period of

time fell by half. However, the decrease in the broadcasting of Brazilian soap operas was not very significant, diminishing from 62% to 48% (see Torres & Burnay 2014). Another aspect that could decrease the Brazilian cultural influence in Portugal is the fact that the new Portuguese generation is less influenced by Brazilian music and soap operas than the previous generations. Conversely, it should be noted that Portuguese football coaches have been increasingly and successfully working with Brazilian teams, although this is but a specific and punctual phenomenon.

Secondly, the number of Brazilian immigrants in Portugal increased dramatically in the last four years, now estimated at around 200.000 – an increase of 43% from 2018 to 2019 and of 22% from 2019 to 2020. The entrance of more Brazilians into Portugal started at the beginning of the new century, in part due to the greater difficulty in accessing the USA after 2001. Brazilians are currently the largest foreign community in Portugal, representing 34% of the total number of immigrants. This number is likely to be even greater, because Brazilian immigrants with Portuguese citizenship (or citizenship from another European country) and Brazilians in an irregular situation are not included in these data. The profile of these Brazilian emigrants is heterogeneous, ranging from workers with few academic qualifications to graduates such as lawyers, engineers, doctors, psychologists, advertisers and even university students. The linguistic and cultural affinity, the quality of life (health, education and mainly safety), the ease of obtaining a visa, and the relatively low living costs compared to other European countries are the main factors that attract Brazilians to Portugal. And the number of Brazilian immigrants continues to rise despite the Covid-19 pandemic. Importantly, there is no evidence that Brazilian immigrants try to adapt and conform to the standard EP norm to be better integrated, even considering the negative attitude found against the Brazilian immigrants in several situations.

Thirdly, Brazilian language forms are attitudinally perceived in Portugal as “invaders” and even “killers” of the language. A collective book entitled “Murdering the Portuguese language!” was published in Portugal in 1983, in which the “language killer” is the Brazilian soap opera (Moura 1983). More recently, aversion towards the Brazilian language variety has re-emerged in the context of the recent Orthographic Agreement, which entailed more changes in EP than BP. Many Portuguese regard the Agreement as representing the unacceptable submission of Portugal to Brazil, a “national disaster” (Moura 1998), an operation of “Brazilianization” of the language or even the “sale of the

language” to Brazil, or as an act of cultural imperialism undertaken by the former colony. The strong increase of Brazilian immigrants mentioned above has furthermore strengthened the negative attitude of many Portuguese schools towards the Brazilian language variety, considering it as a hindrance or misleading factor in the acquisition of speech by Portuguese children. Moreover, the negative attitude displayed by the Portuguese towards Brazilian people, especially the existence of discrimination against Brazilian immigrants seeking work in Portugal, and the concern that Portuguese children would start speaking Brazilian Portuguese, does not reduce, let alone diminish, the insecurity felt by the Portuguese towards the perceived threat of Brazilian influence, but rather reinforces it.

Finally, the Afro-Brazilian linguistic continuum (Petter 2009, Álvarez López et al. 2018), manifested in a set of similar linguistic features diverging from the EP standard in both Brazilian and African varieties of Portuguese, which have emerged from a shared Bantu substratum and contact-induced changes, could be another important factor accounting for the influence of the Brazilian variety on the African varieties of Portuguese and, consequently, for the reverted asymmetry.

Yet, although the Brazilian cultural influence in Portugal is considerable, the linguistic influence of the Brazilian variety on European Portuguese language use is minimal. For instance, the linguistic impact of the very popular Brazilian soap operas, which was stronger in the 80s and 90s, is negligible: very few Brazilian expressions entered colloquial EP in the course of the last four decades, and the few that did started fading away or even disappearing from the 2000s onwards.

A similar scenario is likely to happen with the so-called ‘invasion’ of the Brazilian variety through the influence of Brazilian YouTubers, who are especially popular among Portuguese children and teenagers these days. Crucially, our lectometric research on the BP-EP bicentricity, referred to in section 4, showed that the anticipated growing influence of BP on EP is not confirmed. Interestingly, only 2% of Brazilian terms are commonly used in the European variety regarding the vocabulary of football, which was considered as one of the lexical fields most likely to be influenced by BP (Soares da Silva 2010). Examples of Brazilian terms introduced in EP are *bandeirinha* ‘assistant referee’, *atacante* ‘forward’, *falta* ‘foul’, and *plantel* ‘team’. There are many Brazilian terms, however, for which no occurrence was found at all in EP: for instance, *goleiro*

'goalkeeper', *arquivo* 'goalkeeper', *avante* 'forward', *escanteio* 'corner', *impedimento* 'offside', *gol* 'goal', and *zagueiro* 'back'.

7. Conclusion

A few conclusions can be drawn from this discussion about the (a)symmetry of Portuguese pluricentricity, based on lectometric data of language use, perceptual and attitudinal language data and cultural data. To begin with, Portuguese is still fundamentally a *bicentric* language, owing to the linguistic, cultural and political primacy of Portugal and Brazil, which still are the two major production centres of language standards and linguistic norms. The two countries currently remain the only actors in the *formal* or deliberate, top-down process of standardization in all dimensions of language use, not only the traditional triad consisting of orthography, lexicon and grammar, but also orthoepy, the treatment of linguistic doubts and difficulties, as well as formal and informal registers.

Secondly, Portuguese is a bicentric language presenting a divergent evolutionary trend. Moreover, the divergent bicentricity of the European and Brazilian varieties is fundamentally *symmetric*, with some minor asymmetries. The symmetric bicentricity of EP and BP was lectometrically and sociocognitively confirmed, both objectively through corpus-based lexical and constructional variables, and subjectively through attitudinal indicators obtained by means of a survey submitted to Portuguese and Brazilian students (see Soares da Silva 2010, 2014a, b, 2016, 2021a).

Thirdly, Portuguese is slowly moving from bicentricity to pluricentricity, with the newly emergent African varieties, especially Mozambican and Angolan Portuguese, reinforced by the increasing demographic, economic and political importance of Portuguese-speaking African countries. The great challenge is now to reach a truly *pluricentric* standardization, as well as a multilateral, rather than bilateral, management of national standards to accurately reflect the pluricentric situation of Portuguese, with two dominant varieties, the European and the Brazilian standards, two main non-dominant varieties, the Mozambican and the Angolan standards, and other African and Asian non-dominant varieties. Rather than a pan-Lusophone supranational model to which some romantic and rationalist language policies seem to aspire, there should be more comparative studies on the established and emerging standards, concentrating efforts on undertaking a real pluricentric codification of Portuguese.

Fourthly, we saw that the great influence of Brazilian popular culture in Portugal and other Portuguese-speaking countries is gradually leading to a potential situation of *reverted asymmetry*, with regard to culture and language perception and attitudes. The main manifestations of this reverted asymmetry are

- (i) the proliferation of Brazilian audiovisual cultural products, especially music, television series and soap operas, and, more recently, YouTube videos aimed at children, entering Portugal and making a huge impact (although Brazilian soap operas have seen their influence gradually diminishing in the last 20 years due to the growing success of Portuguese soap operas);
- (ii) the dramatic increase in the number of Brazilian immigrants in Portugal in the past four years;
- (iii) the perception by Portuguese people of a *Brazilianization* of the European standard and the attitudinal aversion towards the Brazilian language forms viewed as invaders and even killers of the language; and
- (iv) the shared sociohistorical characteristics and a set of similar linguistic features of the Brazilian and African varieties of Portuguese, due to their common Bantu basis. There are, however, counterforces and forms of resistance to these Brazilian influences, so the situation of reverted asymmetry may still be in its early stages. Importantly, we do not (or not yet) observe a reverted asymmetry with respect to language use: the influence of the Brazilian variety on the European one is minimal.

Finally, the complex and dynamic nature of the pluricentricity of Portuguese, especially the dynamicity of Brazilian Portuguese, may shed light on the gradient concept of pluricentricity and on the complicated process of pluricentric standardization, as well as on the blurred distinction between dominant and non-dominant varieties of pluricentric languages.

References

- Aguiar e Silva, Vítor (2007): *Ilusões e desilusões sobre a política da língua portuguesa*. In: Manuel Gama (ed.). *A Política da Língua Portuguesa*. Braga. Universidade do Minho, 13-26.
- Álvarez López, Laura / Gonçalves, Perpétua / Avelar, Juanito (eds.) (2018): *The Portuguese Language Continuum in Africa and Brazil*. Amsterdam/Philadelphia. John Benjamins.
- Bagno, Marcos (1999): *Preconceito Lingüístico: O que é, como se faz*. São Paulo. Edições Loyola.

- Bagno, Marcos (2000): *Dramática da Língua Portuguesa: Tradição gramatical, mídia e exclusão social*. São Paulo. Edições Loyola.
- Baxter, Alan N. (2018): Epilogue. In: Laura Álvarez López, Perpétua Gonçalves & Juanito Avelar (eds.). *The Portuguese Language Continuum in Africa and Brazil*. Amsterdam/Philadelphia. John Benjamins, 293-313.
- Chimbutane, Feliciano (2018): Portuguese and African languages in Mozambique: A sociolinguistic approach. In: Laura Álvarez López, Perpétua Gonçalves & Juanito Avelar (eds.). *The Portuguese Language Continuum in Africa and Brazil*. Amsterdam/Philadelphia. John Benjamins, 89-110.
- Clyne, Michael (ed.) (1992): *Pluricentric Languages. Different Norms in Different Countries*. Berlin/New York. Mouton/de Gruyter.
- Duarte, Maria Eugênia L. / Gomes, Christina Abreu / Paiva, Maria da Conceição (2016): Codification and standardization in Brazilian Portuguese. In: Rudolf Muhr (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions*. Frankfurt am Main, etc. Peter Lang Verlag, 51-65.
- Duarte, Maria Eugênia L. / Gomes, Christina Abreu / Paiva, Maria da Conceição (2022): Beyond the dichotomy dominant and non-dominant varieties of pluricentric languages: The case of Brazilian Portuguese. In: Rudolf Muhr, Eugenia Duarte, Cilene Rodrigues & Juan Thomas (eds.). *Pluricentric languages in the Americas*. Graz/Berlin. PCL-Press.
- Faraco, Carlos Alberto (2008): *Norma Culta Brasileira - Desatando alguns nós*. São Paulo. Parábola Editorial.
- Faraco, Carlos Alberto (2011): O Brasil entre a norma culta e a norma curta. In: Xoán Lagares & Marcos Bagno (eds.). *Políticas da Norma e Conflitos Linguísticos*, São Paulo. Parábola Editorial, 259-275.
- Faraco, Carlos Alberto (ed.) (2001): *Estrangeirismos. Guerras em torno da língua*. São Paulo. Parábola Editorial.
- Fino, Carlos (2021): *Portugal-Brasil: Raízes do Estranhamento*. Lisboa / São Paulo. Lisbon Press.
- Geeraerts, Dirk / Kristiansen, Gitte / Peirsman, Yves (eds.) (2010): *Advances in Cognitive Sociolinguistics*. Berlin/New York. Mouton/de Gruyter.
- Gonçalves, Filomena (2020): Portuguese: Orthography and Orthoepy. In: Franz Lebsanft & Felix Tracke (eds.). *Manual of Standardization in the Romance Languages*. Berlin/Boston. Mouton/de Gruyter, 651-677.
- Gonçalves, Perpétua (2010): *A gênese do português de Moçambique*. Lisboa. Imprensa Nacional/Casa da Moeda.

- Gonçalves, Perpétua (2013): O Português em África. In: Eduardo Paiva Raposo, Maria Fernanda Bacelar do Nascimento, Maria Antónia Coelho da Mota, Luísa Segura & Amália Mendes (eds.). Gramática do Português. Vol. I. Lisboa. Fundação Calouste Gulbenkian, 157-178.
- Gonçalves, Perpétua (2018): Research on L2 varieties of European languages. From descriptive to formal grammars. In: Laura Álvarez López, Perpétua Gonçalves & Juanito Avelar (eds.). The Portuguese Language Continuum in Africa and Brazil. Amsterdam/Philadelphia. John Benjamins, 43-63.
- Hofstede, Geert (2001): Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations. Thousand Oaks. CA. Sage.
- Inverno, Liliana (2011): The restructuring of Portuguese morphosyntax in interior Angola – Evidence from Dundo (Lunda Norte). (Unpublished doctoral dissertation). Coimbra. Universidade de Coimbra.
- Inverno, Liliana (2018): Angolan Portuguese: Its historical development and current sociolinguistic setting. In: Laura Álvarez López, Perpétua Gonçalves & Juanito Avelar (eds.). The Portuguese Language Continuum in Africa and Brazil. Amsterdam/Philadelphia. John Benjamins, 111-133.
- Joseph, John E. (1987): Eloquence and Power. The Rise of Language Standards and Standard Language. New York. Basil Blackwell.
- Lagares, Xoán Carlos (2021): Which pluricentrism? Tensions and conflicts on the construction of the Spanish and Portuguese international linguistic space. In: WORD, 67/1. P. 60-81.
- Langacker, Ronald W. (1987): Foundations of Cognitive Grammar. Vol. 1: Theoretical prerequisites. Stanford. Stanford University Press.
- Langacker, Ronald W. (2008): Cognitive Grammar. A basic introduction. Oxford. Oxford University Press.
- Mateus, Maria Helena Mira (2002): A Face Exposta da Língua Portuguesa. Lisboa. Imprensa Nacional-Casa da Moeda.
- Mattos e Silva, Rosa Virgínia (2004): “O português são dois ...”. Novas fronteiras, velhos problemas. São Paulo. Parábola Editorial.
- Mendes, Edleise (2016): The Portuguese language and its non-dominant varieties: how to teach them? In: Rudolf Muhr (ed.). Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions. Frankfurt am Main, etc. Peter Lang Verlag, 85-97.
- Mingas, Amélia (2000): Interferência do kimbundu no português falado em Lwanda. Luanda. Edições Chá de Caxinde.

- Moura, Vasco Graça (ed.) (1983): *Estão a assassinar o português!* Lisboa. Imprensa Nacional-Casa da Moeda.
- Moura, Vasco Graça (1998): *Acordo Ortográfico: A perspetiva do desastre*. Lisboa. Aletheia Editores.
- Oliveira, Gilvan Müller de (2015): O Instituto Internacional da Língua Portuguesa e a gestão multilateral da Língua Portuguesa no âmbito da CPLP. In: *Revista Internacional de Lingüística Iberoamericana*, 13/25. P. 19-34.
- Oliveira, Gilvan Müller de (2016): The system of national standards and the demolinguistic evolution of Portuguese. In: Rudolf Muhr (ed.). *Pluricentric Languages and Non-Dominant Varieties Worldwide. Part II: The Pluricentricity of Portuguese and Spanish. New Concepts and Descriptions*. Frankfurt am Main, etc. Peter Lang Verlag, 35-48.
- Petter, Margarida M. (2009): O continuum afro-brasileiro do português. In: Charlotte Galves, H. Garmes & F. Ribeiro (eds.). *África-Brasil. Caminhos da Língua Portuguesa*. Campinas. Editora da Unicamp, 158-173.
- Rezende, Claudia Barcellos / Coelho, Maria Cláudia (2010): *Antropologia das emoções*. Rio de Janeiro. Editora Fundação Getúlio Vargas.
- Ribeiro, Darcy (1997): *O Povo Brasileiro: A formação e o sentido do Brasil*. São Paulo. Editora Schwarcz.
- Scherre, Maria Marta Pereira (2005): *Doa-se lindos filhotes de poodle: Variação lingüística, mídia e preconceito*. São Paulo. Parábola Editorial.
- Soares da Silva, Augusto (2010): Measuring and parameterizing lexical convergence and divergence between European and Brazilian Portuguese. In: Dirk Geeraerts, Gitte Kristiansen & Yves Peirsman (eds.). *Advances in Cognitive Sociolinguistics*. Berlin/New York. Mouton/de Gruyter, 41-83.
- Soares da Silva, Augusto (2014a): The pluricentricity of Portuguese: A sociolectometrical approach to divergence between European and Brazilian Portuguese. In: Augusto Soares da Silva (ed.). *Pluricentricity: Language Variation and Sociocognitive Dimensions*. Berlin/New York. Mouton/de Gruyter, 143-188.
- Soares da Silva, Augusto (2014b): Measuring and comparing the use and success of loanwords in Portugal and Brazil: A corpus-based and concept-based sociolectometrical approach. In: Eline Zenner & Gitte Kristiansen (eds.). *New Perspectives on Lexical Borrowing: Onomasiological, methodological and phraseological innovations*. Berlin/New York. Mouton/de Gruyter, 101-141.
- Soares da Silva, Augusto (ed.) (2014): *Pluricentricity: Language Variation and Sociocognitive Dimensions*. Berlin/New York. Mouton/de Gruyter.

- Soares da Silva, Augusto (2015): Cultural cognitive models of language variation. Romanticism and rationalism in language policy debates about the unity/diversity of European and Brazilian Portuguese. In: Jocelyne Daems, Eline Zenner, Kris Heylen, Dirk Speelman & Hubert Cuyckens (eds.). *Change of Paradigms – New Paradoxes. Recontextualizing Language and Linguistics*. Berlin/New York. Mouton/de Gruyter, 253-274.
- Soares da Silva, Augusto (2016): The cognitive approach to pluricentric languages and the pluricentricity of Portuguese: What's really new? In: Rudolf Muhr (ed.). *Pluricentric languages and non-dominant varieties worldwide. Part II: The pluricentricity of Portuguese and Spanish. New concepts and descriptions*. Frankfurt am Main, etc. Peter Lang Verlag, 13-34.
- Soares da Silva, Augusto (2018): O português no mundo e a sua standardização: entre a realidade de uma língua pluricêntrica e o desejo de uma língua internacional. In: Henrique Barroso (ed.). *O Português na Casa do Mundo, Hoje*. Braga. Centro de Estudos Humanísticos da Universidade do Minho / Húmus, 111-132.
- Soares da Silva, Augusto (2020a): Portuguese: Normative Grammars. In: Franz Lebsanft & Felix Trache (eds.). *Manual of Standardization in the Romance Languages*. Berlin/Boston. Mouton/de Gruyter, 679-700.
- Soares da Silva, Augusto (2020b): Exploring the cultural conceptualization of emotions across national language varieties: A multifactorial profile-based account of *pride* in European and Brazilian Portuguese. In: *Review of Cognitive Linguistics*, 18/1. P. 42-74.
- Soares da Silva, Augusto (2021a): Revisiting the Cognitive Sociolinguistic approach to pluricentricity. In: Gitte Kristiansen, Karlien Franco, Stefano De Pascale, Laura Rosseel & Weiwei Zhang (eds.). *Cognitive Sociolinguistics Revisited*. Berlin/Boston. Mouton/de Gruyter, 477-489.
- Soares da Silva, Augusto (2021b): Measuring the impact of (non)figurativity in the cultural conceptualization of emotions in the two main national varieties of Portuguese. In: Augusto Soares da Silva (ed.). *Figurative Language: Intersubjectivity and Usage*. Amsterdam/Philadelphia. John Benjamins, 387-438.
- Soares da Silva, Augusto (2022): Metaphor, emotion, and intralinguistic cultural variation: metaphors of ANGER in European and Brazilian Portuguese. In: Ulrike Schröder, Milene Mendes de Oliveira & Adriana Maria Tenuta (eds.). *Metaphorical Conceptualizations. (Inter)Cultural Perspectives*. Berlin/Boston. Mouton/De Gruyter, 189-222.

- Soares da Silva, Augusto / Afonso, Susana / Palú, Dafne / Franco, Karlien (2021): Null *se* constructions in Brazilian and European Portuguese: Morphosyntactic deletion or emergence of new constructions? In: *Cognitive Linguistics*, 32/1. P. 159-193.
- Torres, Eduardo Cintra / Burnay, Catarina Duff (2014): A Telenovela em Portugal: estreias, importação e exportação (1993-2012). In: Moisés de Lemos Martins, Rosa Cabecinhas, Lurdes Macedo & Isabel Macedo (eds.). *Interfaces da Lusofonia*. Braga. Centro de Estudos de Comunicação e Sociedade, Universidade do Minho, 126-138.

Maria Eugênia L. DUARTE
Christina A. GOMES
Maria da Conceição PAIVA

(Federal University of Rio de Janeiro, Brazil)
eugenia@brazilmail.com
christina-gomes@uol.com.br
paivaa@club-internet.fr

Beyond the dichotomy Dominant and Non-Dominant varieties of Pluricentric Languages: the case of Brazilian Portuguese

Abstract

This paper compares four different varieties of Pluricentric Portuguese that have been characterised as having a symmetric or an asymmetric relationship with European Portuguese (EP) – considered the standard model by the first Brazilian grammarians in late 19th Century. We show that Brazilian Portuguese (BP), Mozambique Portuguese (MP) and São Tomé Portuguese (STP), in relation to EP, have developed convergent linguistic patterns regarding the expression of 3rd person clitics, preferring null objects, and therefore, diverging from EP; but they differ from BP in relation to the overall rates of overt referential pronominal subjects, since BP is set apart from the other two varieties, although MP and BP converge, to a certain extent, depending on the discourse person of the subject. These results suggest a very complex picture, since the development of their norms follows opposite trends for the two syntactic features approached, which renders it difficult to characterise them as varieties that adopted or rejected the dominant national norm.

1. Introduction

The varieties of a Pluricentric language have been approached in terms of their mutual relationship, considering the dominance of one of them regarding to the others. Then, they are treated as belonging to either a dominant (DV) or a non-dominant (NDV) variety group. This distinction between DV and NDV is based on several features (Clyne, 1992; Muhr, 2012; 2013) in order to provide a way to capture the specificities of their linguistic

development (how dependent/independent), and the types of their processes of standardisation and codification. In this paper, we claim that the dichotomic distinction between DV and NDV languages (Clyne, 1992; Muhr, 2012) can be argued, if we take into account that the criteria that subsume this distinction are external and based on a static rather than a dynamic concept of linguistic norm.

Several years of sociolinguistic studies in Brazil have already challenged this conception of linguistic norm with strong empirical evidence. Hence, in order to address our main issue, we take into consideration the fact that linguistic norms are dynamic not static. Additionally, we argue that the dichotomic distinction between NDV and DV is political/ideological rather than linguistic or structural, since the adoption of a general norm does not hold for all different varieties of the same language through the rejection of local norms and the adoption of exogeneous ones. This attitude disregarded not only the fact that, in the middle of the 18th Century, Brazilian Portuguese had already acquired a local identity and was beginning to exhibit changes that would deviate from the mother tongue, but also had developed a model of writing that distinguished it from EP. The model sought by Brazilian grammarians was the language of the great 19th century writers of Portuguese. The consequence of such a “decision” would create a big difference between the local and the prescribed standard descriptions in grammars and would increase the huge gap between real speech and a presumed standard writing in Brazil (Duarte, Gomes and Paiva, 2016:53). Furthermore, evidence from different studies has shown that not only different norms or standards may evolve in the course of language development but that such norms are inherently variable as well (Meyerhoff; Niedzielski, 2003; Kerswill, 2006), which means that the degree that varieties differ cannot be measured in terms of only one dimension or a specific linguistic structure.

Based on data from oral modality, we will compare three different varieties of Portuguese that have been characterised as having a symmetric or an asymmetric relationship with European Portuguese (EP) – the standard model. We will show that, even under different historical and linguistic contexts, during and after the colonial period, Brazilian Portuguese (BP), Mozambique Portuguese (MP) and São Tomé Portuguese (STP), in relation to EP:

- have developed convergent linguistic patterns regarding the expression of 3rd person clitics, diverging from EP;

- differ in relation to the overall rates of overt referential pronominal subject, since BP is set apart from the other two varieties, although both MP and BP converge, to a certain extent, depending on the discourse person.

Considering the relationship among these different varieties of Pluricentric Portuguese, BP has been characterised as a DV due to its symmetric relationship with EP¹ (Kretzenbacher et al., 2013:261; Soares da Silva, 2013:80; Duarte, Gomes, Paiva, 2016:61), whilst MP and STP have been classified as NDV, whose relationship with EP is asymmetric (Kretzenbacher et al., 2013:261; Mendes, 2016:86).

Taken together, the compared results discussed in this paper will show that, despite their different status in the DV/NDV approach, African varieties have developed similar linguistic patterns to EP, concerning null subjects, and to BP, regarding null objects. Therefore, their distinction in the NDV set of characteristics is mainly related to political and ideological aspects rather than to grammatical ones, since NDV varieties do not necessarily keep DV patterns”, in the present case, EP and BP patterns. A close examination of the figures will make this clear.

We argue that the comparison among BP, MP and STP regarding syntactic features provides some evidence that challenges the idea that ND nations tend to succumb to norms from the D nations(s), the notion of “linguistic cringe” according to Clyne (1992).

2. Methodology

The analysed data were collected from samples organised in a way to obtain spontaneous speech of speakers that differs in relation to social characteristics such as age, level of education and gender.

BP, EP and MP data come from samples recorded by COMPARAPORT Project2, with interviews including 36 speakers born in Rio de Janeiro, 36 in Lisbon, 18 in Maputo, stratified according to age (18-35 / 36-55 / 56-75), school level (Fundamental, High School and College Graduation) and gender. It is worth to mention that one third of the MP sample is composed by speakers

¹ The fact that BP is considered symmetric to EP in spite of all the differences pointed out in a number of comparative analyses confirms our claim in this paper.

² Projeto COMPARAPORT : Estudo comparativo de variedades africanas, brasileiras e europeias do Português), available at www.corporaort.letras.ufrj.br.

that acquired Portuguese as L1 and the individuals from Brazil and Lisbon are native speakers of Portuguese.

The 17 interviews for São Tomé and Príncipe belong to the sample VARPOR (African Varieties of Portuguese/Varieties Africanas do Português) from the Center of Linguistics at the University of Lisbon. The social stratification is the same as the COMPARAPORT Project. A missing interview is due to the fact that a woman with a College Graduation belonging to the older group was not located. In São Tomé and Príncipe, STP is the first language for the majority of the population that also speaks Forro, a creole variety.

The four varieties of Pluricentric Portuguese are compared in relation to two syntactic variables: the expression of 3rd person anaphoric accusative complements and the null versus overt pronominal subjects.

3. Results

3.1 The anaphoric accusative complements

This section starts with the presentation of the results for data concerning the variation in the expression of 3rd person anaphoric accusative complements. The envelop of this variation consists of three variants: the accusative clitic, the standard variant, the null object and the nominative 3rd person pronoun in accusative function, an extremely rare variant in EP.

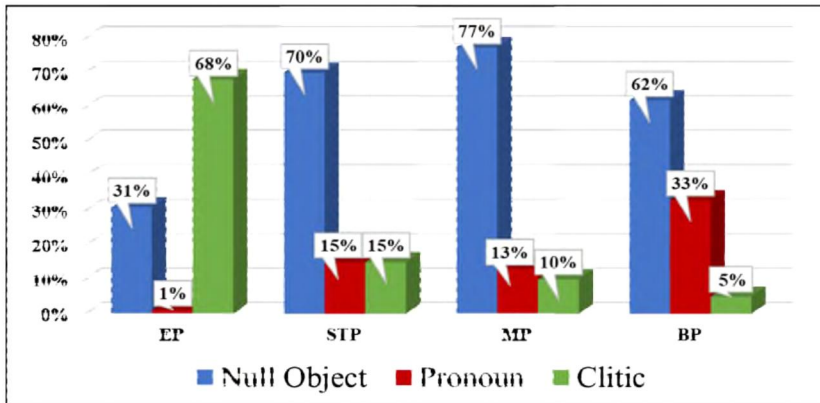


Fig. 1: Anaphoric accusative complements in varieties of Portuguese (Adapted from Table 11, by Marques de Sousa, 2021: p. 157)³

³ Abbreviations: EP=European Portuguese; STP= São Tomé and Príncipe Portuguese ; MP=Mozambique Portuguese; BP= Brazilian Portuguese.

The overall rates⁴ for the anaphoric accusative complements in the four varieties of Portuguese are presented in Figure 1, followed by examples in (1), (2) and (3)⁵ and comments about the three variants attested in each variety:
 Figure 1- Anaphoric Accusative complements in varieties of Portuguese
 (Adapted from Table 11, by Marques de Sousa, 2021: p. 157)

(1) Accusative clitic (the preferred strategy in EP)

- a. [Ele]_i é pequenino pra estar sozinho em casa. Então eu decidi matriculará-[lo]_i aqui. (EP)
 he is young to be alone at home then I decided to-enroll -CL here
"He is too young to stay home by himself. Só I decided to enroll him here."
- b. Eu falo [a língua portuguesa]_i. Tenho-[a]_i como a língua materna (STP)
 I speak [the language portuguese] Have-CL as the tongue mother
"I speak the Portuguese language. I have it as my mother tongue"
- c. Cuidemos bem [dos nossos mais velhos]_i, não podemos apartá-[los]_i (PM)
 let's-take-care [of-the our more elder]_i, not can keep-CL-away
"Let's take care of our elders, we can't stay far from them."
- d. Eu acho que [nossa Constituição]_i está acabada, não temos mais como utilizá-[la]_i. (BP)
 I think that [our constitution]_i is over, not have more how to-use-CL
"I think our Constitution is useless, we cannot use it anymore."

As shown in (1) above, EP presents 68.3% of clitics, whereas this variant is the least used in the analysed samples of STP, MP and BP, showing, respectively, 15%, 10%, and 5%, a result that contrasts with those obtained for the null object.

In fact, EP does exhibit the null object, but with a lower rate (31%) and it is more constrained than the other varieties, particularly by the semantic feature of the referent ([–animate]), normally mentioned outside the limits of the sentence where the null object is attested. The three other varieties, on the other hand, are characterised by the prevalence of the null object, 70%, 77% and 62%, respectively, as one can see in (2):

⁴ The results presented here are shown in percentages, but the logistic regression obtained by the Goldvarb Program (Sankoff, Tagliamonte, Smith, 2005) for Relative Weights, guarantee the significance of the rates shown in the Figures.

⁵ All the examples for the anaphoric accusative come from Marques de Sousa (2021)

(2) Null accusative objects in EP, STP, MO and BP (examples from Marques de Sousa, 2021)

- a. Gostei muito dos cânticos, comprei [Cds]_i, levei [Ø]_i para casa. (EP)
 enjoyed_{-1ps} much the chants bought [Cds]_i, took_{1ps} [Ø]_i to home
I enjoyed the chants, bought CDs and took them home.
- b. Eu colhia [as informações]_v, escrevia [Ø]_i e passava [Ø]_i às meninas (STP)
 I collected [the Information] wrote [Ø]_i and passed [Ø]_i to-the girls.
I collected the Information, wrote them and passed them to the girls.
- c. A família que fecha-se no mundo, cria [seus filhos]_i e impede [Ø]_i de sair. (MP)
 the family that closes-itself in-the world, raises [their children]_i and forbids [Ø]_i to leave
A family that closes itself in the world, raises their children and forbids them to leave.
- d. [Ela]_i tem o mundo nas mãos. Mas é preciso conduzir [Ø]_i. (BP)
 [she]_i has the world in-the hands. But is necessary to guide [Ø]_i
She has the world in her hands, but it is necessary to guide her.

In relation to the third strategy, the use of the nominative pronoun *ele/ela /eles/ elas* (he/she/they) in accusative function, we can point out that (a) in STP, MP and BP its frequency is much lower than the null object, whilst EP presents the lowest overall rate of this variant, 1%, corresponding to three occurrences. Examples for the varieties can be seen in (3):

(3) Nominative pronoun in accusative function

- a. Mas é assim: ajudei [ele]_i a reeleger, o primeiro ministro. (EP)
 but is so: helped_{-1ps} he to re-elect the prime minister
"But, that is it: I helped him to re-elect the Prime Minister."
- b. Quando [meu filho Álvaro]_i vai embora, vou lá na roça ver [ele]_i. (STP)
 when [my son Álvaro]_i leaves, go there in-the field to-see [he]_i
"When my son Álvaro leaves, I go to the field to see him,"
- c. [As meninas]_i quiseram ficar comigo e fui educando [elas]_i. (MP)
 [the girls]_i wanted to-stay with-me and (I) was educating they;
"They girls wanted to stay with and I kept raising them."
- d. [Ele]_i mora sozinho, minha mãe morreu, não consigo abandonar [ele]_i. (BP)
 [he]_i lives alone, my mother died not can abandon [he]_i
"He leaves alone, my mother died, I can't abandon him."

The nominative pronoun is more frequent in BP than in the other two African varieties. All the examples show nominative pronouns used as direct objects, with verbs that select one internal argument (as in (3b, c,d) or two internal arguments, such as *help* (in (3a)). Verbs like *help*, *oblige*, *persuade*, *convince*, etc, which normally select a [+human] direct object, favours the use of the nominative pronoun as an accusative complement. Two of the three occurrences in EP fall into this syntactic pattern (S V OD OBL). The third occurrence, also rare in EP, but more common in the other varieties, although clearly preferred in BP, illustrate cases of Exceptional Case Marking, structures, where the subject of the embedded clause, an infinitival complement of a causative, a sensitive and a verb of permission, receives accusative Case from the verb of the main clause. This is what is shown in (4a) for EP, a single occurrence. The same holds for the subject of small clauses (4b, c), also Case-marked by the verb:

- (4) a.tinha [**uns clientes**]_i, e eu fui ver [**eles**]_i, a actuar. (EP)
 (I) had [some clients] and I went to-see they to perform
I had some clients and I went to see them act.
- b. (Eu) não gosto [**das novelas portuguesas**]_i. Eu acho [[**elas**]_i, muito muito quietas. (MP)
 I not like of-the soap-operas portuguese. I find [they] much, much quiet
I don't like the Portuguese soap-operas very much. I find them very, very quiet.
- c. Você tem que casar pra você fazer [**ele**]_i, feliz.
 you have to marry for you to-make [he] happy
You must get married to make him happy.

As shown, the three varieties differ deeply from EP according to the use of the accusative clitic, which means that they are characterised by the use of strategies that are rare in EP, particularly the nominative pronoun in accusative function and, more frequently, the null object.

It is clear that the acquisition of third person accusative clitics in varieties of Portuguese fails, regardless of the acquisition as L1 or as L2, contrary to what happens with 1st and 2nd person clitics, the deictic ones, which are normally acquired (see Duarte et al. 2021 about the loss of 3rd person accusative and dative clitics and the indefinite clitic in the history of BP and their strength in EP).

3.2 Null and overt pronominal subjects

Let's now turn our attention to the alternation between null and overt pronominal subjects in the varieties under analysis. The data comes from the same samples mentioned in note 1 (See Methodology). EP has been considered a "consistent" null subject language, a system that licenses null subjects because of its "rich" verbal morphology, like Greek, for instance (Roberts and Holmberg, 2010). BP, on the other hand, is included by the authors among "partial" null subject languages, like Finnish, for instance, a label that does not seem to fit BP, which, contrary to what has been described for Finnish, is losing null subjects in every syntactic environment, at least as far as referential pronominal subjects are concerned⁶. The variationist analysis used the structural and social factor groups refined in Duarte and Rezende dos Reis (2018) and Duarte (2019).

Figure 2 shows the overall rates for null pronominal subjects (vs overt pronouns) in the four varieties.

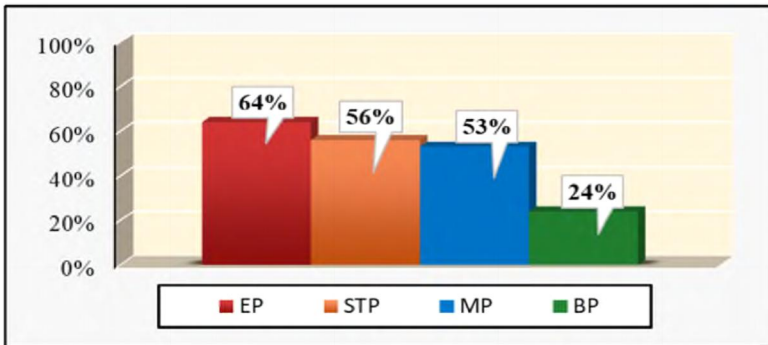


Fig. 2: Null and overt pronominal subjects in four varieties of Portuguese

According to Figure 2, the African varieties are much closer to EP, whereas BP is clearly apart from them. When we consider the distribution of null subjects according to the discourse person (1st sg/pl, 2nd sg/pl, 3rd sg/pl), the overall rates, as shown in Figure 3, allow a deeper comprehension of the data behaviour in each variety.

⁶ Duarte and Marins (2021) show that Finnish allows null subjects in very restricted contexts, whilst BP is in a process of losing them in every context. The completion of the change affects less syntactically accessible and [+human] antecedents first.

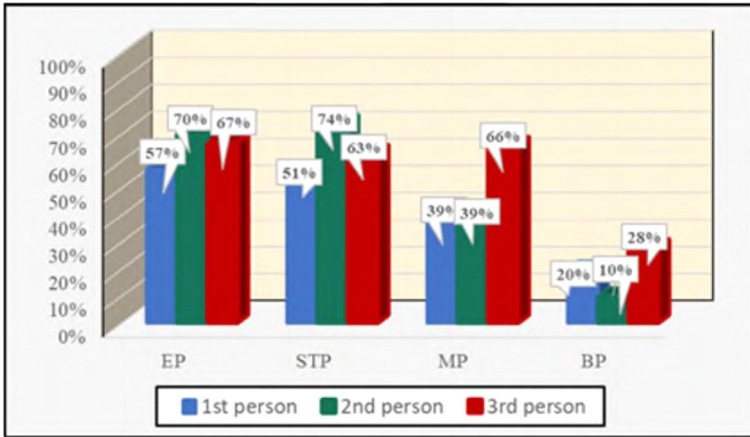


Figure 3- Null (vs overt) pronominal subjects according to person
(from Duarte and Rezende dos Reis, 2021)

We can observe that EP shows a more regular distribution in the preference for null subjects in the three persons, which is expected from a consistent null subject language, as we will show below. Among the African varieties, STP is closer to EP than MP. However, they behave similarly with respect to the 3rd person. The pattern observed for BP is consistent with the conclusion of an advanced process of change, with the lowest rates of null subjects in all persons.

PERSON	SINGULAR	PLURAL
1st	Eu trabalho	Nós trabalhamos A gente trabalha
2nd	Tu trabalha(s)~ Você trabalha	Vós trabalhais Vocês trabalha(m)
3rd	Ele/Ela trabalha	Eles/Elas trabalha(m)

Chart 1. Personal pronouns and indicative present verbal inflection in Brazilian Portuguese (verb to work = trabalhar)

It is obvious that the inflectional verb paradigm can help explain the differences in the rates in Figure 3. Two important changes, shown in Chart 1, impacted the pronominal system in BP: the first is usually related to contacts during colonisation – a) the apocope of 2nd singular verbal morpheme <-s> in the varieties that use pronoun *tu*, a very frequent phenomenon, and is not subject to stigma; the other is the process of denasalisation and reduction of

the diphthong [ãw] (orthographic <-am>) with the loss of the posterior glide, which can eliminate the distinctive 3rd person plural morpheme, a variable process controlled by structural and social factors and the only variant (with the reduction) subject to stigma today.

As for the set of personal pronouns, two new pronouns entered the paradigm from the 19th century on, via grammaticalisation: *você/vocês* ‘you, you all’, and *a gente* ‘we’. Both *você* and *a gente* come from DPs - *Você* is the full grammaticalisation of the address form *Vossa Mercê* ‘Your Grace’ and entered in competition with *tu*; *a gente* comes from a DP meaning ‘the people’ (like French *on*) and is also fully grammaticalised in contemporary BP as a 1st person plural pronoun and outnumbers by far the conservative pronoun *nós*. Since both new pronouns occur with 3rd person singular verb form, this accelerated the reduction of the inflectional paradigm. Today we have the following situation: *Tu* and *você* are in variation in the regions that keep both pronouns (Scherre; Yacovenço; Paiva, 2019), *nós* is close to obsolescence, particularly among younger generations (Omena 1996; Lopes 2002; 2003).

As for EP, there have been no cases of apocope, *Tu* and *Você* are in complementary distribution – *tu* for family and close friends, *você* usually unpronounced except when a superior addresses an inferior. Otherwise, it is considered rude.⁷ *Nós* is much more frequent in EP when compared to *a gente*.⁸

In sum, we can say that the EP paradigm keeps its “functional” richness (<-o>, <-s>, <-Ø>, <-mos>, <-m>) with two syncretisms and eventually two zero morphemes (See Roberts 1993). As far as we know *vós*, although with low frequency, it is still attested in some more conservative areas of Portugal (Paiva; Aguiar, 2017). Moreover, the clitic, oblique and possessive pronouns belonging to *vós* paradigm (*vos*, *convosco*, *vosso*) are frequently used in association with nominative *vocês* (Paiva; Aguiar, 2017).

BP, on the other hand, has exceeded the number of zero morphemes, even with paradigms that preserve 2nd and 3rd person plural inflection (<-o>, <-Ø>, <-m>).

As for the two African varieties presented here, they show a different picture. Mozambique keeps pronoun *tu* with the canonical inflection but the use of *você* outnumbers *tu* – in any case both pronouns show the same rates of null subjects. They prefer *nós* to *a gente*, which favours the null subject because of the distinct inflection.

⁷ Curiously, *vocês* (for 2nd person plural, replacing obsolescent pronoun *vós* is not felt as rude.

⁸ According to Lopes and Brocardo (2016) *você* and *a gente* are not fully grammaticalised in EP.

São Tomé and Príncipe shows an expressive preference for *ocê*, but behaves like EP, keeping the pronoun preferably null; the use of *nós* also outnumbers *a gente* and null 1st person plural subjects are predominant; finally, they exhibit a frequent occurrence of denasalisation of 3rd person <ão> but this does not seem to prevent the preference for null 3rd person subjects, certainly identified by an antecedent

Therefore, we could say that Mozambique and São Tomé are closer to EP than to BP, because *tu* is preserved to a certain extent and *nós* is largely preferred to *a gente*.

We will restrict our analysis to the 3rd person, a very important context, because, contrary to 1st and 2nd, besides the interaction of [-/+animate referents], it can exhibit the same structural patterns found for 1st and 2nd persons, and it seems to be a more revealing context in Figure 3. In a “consistent” null subject language belonging to the Romance group, a null subject in an embedded clause or in a main clause occurs regardless of c-command. Examples in (5) illustrate a 3rd person subject in the main clause c-commanding the subject of the embedded clause in each variety. EP shows 94% of null subjects in this context, MP, 86%, STP, 62%, whilst BP shows 41%. This is the most resistant context in BP for the ongoing change detected by Duarte (1994) and confirmed by Duarte (2019), based on the sample described in the present paper.⁹

(5) The null subject antecedent is in the main clause; it c-commands the null subject in the embedded clause

- a. [O pobre]_i; continua com essa mentalidade porque \emptyset_1 só pensa nele. (EP)
 the poor guy keeps with this mentality because only thinks in-himself
The poor guy keeps this mentality because he is only concerned about himself.
- b. Como eu disse, [católico]_i; está a ver que \emptyset_1 está a perder muitos cristão. (STP)
 as I said, [catholic] is_{2ps} to see that \emptyset_1 is to lose many christian
As I said, catholic, you can see that they are losing many Christians.
- c. [Os professores]_i; também pensam que \emptyset_1 são mal pagos. (MP)
 the teachers also think that are badly paid
Teachers also think that they are underpaid.
- d. [Meu marido]_i; foi quase preso porque \emptyset_1 foi mergulhar (BP)

⁹ The other resistant context refers to [-animate] subjects, with an average 50% of null subjects. In any case, the very emergence of a paradigm of [-animate] personal pronoun in BP means another loss of a crucial property of “consistent” null subject languages, which have such personal pronouns.

my husband was almost arrested because went to-dive
My husband was almost arrested because he went to dive.

The same preference for a null subject is also attested when there is no c-command, i.e., the subject of the preceding adjunct clause **is not in a c-command relation** with the subject of the postposed main clause. The examples in (6), illustrate the pattern:

(5) **The null subject antecedent doesn't c-command it**

- a. Quando **ele**_i mudar, \emptyset _i ainda vai estar mais um ano nesta escola (EP)
 when he moves still is-going to-be more one year in-this school
When he moves, he is still going to stay one more year in this school.
- b. À medida que [**a pessoa**]_i cresce \emptyset _i também toma os rumos delas. (MP)
 in-the measure that the person grows-up \emptyset _i also takes the paths of-theirs
As the person grows up he/she takes his/her own paths.
- c. Como [**o meu filho**]_i vive lá com avó paterno, \emptyset _i passa mais tempo com ela. (STP)
 as the my son lives there with grandma paternal \emptyset _i passes more time with her.
As my son lives with his paternal grandmother, he spends more time with her.
- d. Se [**o aluno**]_i tem problema, \emptyset _i vem pra gente conversar. (BP)
 if the student has problem \emptyset _i comes for we to-talk
if the student has a problem he comes to talk to us. (Examples from Duarte, 2019)

The results for this pattern confirm what was expected for EP, with 93% of null subjects, the same result with c-command; it is followed by MP and STP, respectively with 87% and 67%. Thus, the condition of c-command for the antecedent of the null subject is not a constraint for its occurrence in these three varieties of Portuguese. On the other hand, for BP, only 11% of null subject in this context were attested, which means that a crucial property of “consistent” null subject languages is almost lost in BP. Such results seem to be related to the inflectional verb paradigm. The African varieties seem more successful in the acquisition of the European pronominal system and inflectional paradigm, even though important differences have been pointed out. However, one could say that they are closer to EP, but the fact is that there seems to be a continuum, which still places STP followed by MP closer to EP.

4. Conclusion

Our results show dynamic not static systems, for the three varieties that evolved from EP, since they reveal: a) the same trend of language change, regarding the expression of 3rd person anaphoric direct object, since the three languages have developed a similar linguistic pattern that differs from EP: the loss of the accusative clitic and its replacement by a null object, a change that is below consciousness; since the null object is not subject to negative social evaluation, as it is also observed among graduate speakers; b) a similar behaviour to EP, for MP and STP, with respect to the Null Subject Parameter - the African varieties are closer to EP, certainly because of their pronominal system. BP is no longer a [+NS] language.

The comparison outlined in the previous section showed that the norms that evolved in each of the new varieties, BP, emerging from the transplanted EP to Brazil, and MP and STP, emerging from the adoption of EP as an official language and taught as a second language in school and now beginning to develop as a first language, followed opposite trends, when considering the two syntactic features approached in this paper. This suggests a very complex picture for different varieties of a Pluricentric Language, since there is not a clear effect of the codification policy in order to assure the maintenance of the main linguistic patterns of the “original” language in the new varieties. MP and STP are similar to EP with respect to the realisation of pronominal subjects, but they are distant from EP in relation to the loss of the accusative clitic system for anaphoric reference, which renders it difficult to characterise MP and STP as varieties that have adopted or rejected the dominant nation norm. If specific linguistic features are not enough to establish the degree of submission of the new varieties to the older or original linguistic norms, what criteria would support the classification of MP and STP as NDVs in relation to EP? Are they non-dominant varieties, even though they share features with EP and BP, which is considered symmetric to EP in spite of their deep differences? It seems that the criteria that subsumes the distinction among Pluricentric Languages as Dominant and Non-Dominant Varieties are external, more related to the historical relationship among the nations, and based on a static, rather than a dynamic, concept of linguistic norm.

References

- Aguiar, Joana / Paiva, Maria da Conceição (2017): "Vocês tenham cuidado, sois educadas para isso": second person pronouns in Braga speech. In: Pilar Barbosa / Maria da Conceição de Paiva / Celeste Rodrigues. *Studies on variation in Portuguese*. Amsterdam/Philadelphia. John Benjamins, 135-152.
- Clyne, Michael (ed.) (1992): *Pluricentric Languages. Different Norms in Different Countries*. Berlin/New York. Mouton/de Gruyter.
- Duarte, Maria Eugenia L. (1994): *A perda do Princípio "Evite Pronome" no português brasileiro*. PhD Dissertation, Federal University of Rio de Janeiro.
- Duarte, Maria Eugenia L. (2019): *O sujeito nulo referencial no português brasileiro e no português europeu*. In: Charlotte Galves et al. (eds.) *Português Brasileiro: Uma segunda viagem diacrônica*. Campinas. Ed. da Unicamp, 93-126.
- Duarte, Maria Eugenia L. / Gomes, Christina A. / Paiva, Maria da Conceição de (2016): *Codification and standardisation in Brazilian Portuguese*. In: Rudolf Muhr et al. (eds.). *Pluricentric Languages and Non-Dominant Varieties Worldwide: The pluricentricity of Portuguese and Spanish: New concepts and description*. Wien/Frankfurt am Mein. Peter Lang Verlag, v. 2, 45-60.
- Duarte, Maria Eugenia L./ Rezende dos Reis, Eduardo P. (2018): *Revisitando o sujeito pronominal vinte anos depois*. *REVEL*, v. 16, n. 30, 173-197.
- Duarte, Maria Eugenia L. / Rezende dos Reis, Eduardo P. (2021): *The structural position of the subject in varieties of Portuguese*. Paper presented at ALFAL conference, 2021.
- Duarte, Maria Eugenia L./Cardoso, Marianna C./Fernandes, Ulli B./Marques de Sousa, Antonio A. (2021): *A redução no quadro de clíticos de 3a. Pessoa no Português brasileiro: um estudo diacrônico*. *Revista Laborhistórico*, v. 4, especial issue, 154-187.
- Duarte, Maria Eugenia L./Marins, Juliana E. (2021): *Brazilian Portuguese: a 'partial' null subject language?* *Cadernos de Estudos Linguísticos*, v. 63, 1-21.
- Kerswill, Paul (2006): *Migration and language*. In: Klaus Mattheier / Ulrich Ammon / Peter Trudgill (eds.). *Sociolinguistics/Soziolinguistik. An international handbook of the science of language and society*, 2nd ed., vol 3. Berlin. De Gruyter, 2271-2285.
- Lopes, Célia. (2002): *Vossa Mercê > você e Vuestra Merced>usted: o percurso evolutivo ibérico*. *Linguística*, v. 14, 173-190.
- Lopes, Célia. (2003). *A inserção de 'a gente' no quadro pronominal do Português. Madrid. Iberoamericana*.

- Lopes, Célia / Brocardo, M. Teresa. (2016): Main morphosyntactic changes and grammaticalization processes. In: Leo Wetzels et al. (eds.). *The handbook of Portuguese linguistics*. Maldon. John Wiley & Sons, 471-486.
- Kretzenbacher, Heinz L./Hajek, John./Norrby, Catrin. (2013): Address and introductions across two pluricentric languages in intercultural communication. In: Rudolph Muhr et al. (eds.). *Exploring Linguistic standards in non-dominant varieties of pluricentric languages*. Frankfurt am Main. Peter Lang Verlag, 259-274.
- Marques de Souza, Antonio A. *Representação do Acusativo Anafórico em Variedades do Português*. PhD Dissertation, UFRJ, 2021.
- Meyerhoff, M. / Niedzielski, N. (2003): The globalization of vernacular variation. In: *Journal of Sociolinguistics*, no. 7(4), 534-555.
- Muhr, Rudolph (2012): Linguistic dominance and non-dominance in pluricentric languages: A typology. In: Rudolph Muhr (ed). *Non-dominant varieties of Pluricentric Languages. Getting the picture*. Frankfurt am Main, Peter Lang Verlag, 23-49.
- Muhr, Rudolph (2013): Codifying linguistic standards in non-dominant varieties of pluricentric languages – adopting dominant or native norms? In: Rudolph Muhr et al. (eds.). *Exploring Linguistic standards in non-dominant varieties of pluricentric languages*. Frankfurt am Main. Peter Lang, 11-44.
- Mendes, Edleise (2016): The Portuguese language and its non-dominant varieties: how to teach them? In: Rudolf Muhr et al. (eds.). *Pluricentric Languages and Non-Dominant Varieties Worldwide: The pluricentricity of Portuguese and Spanish: New concepts and description*. Wien/Frankfurt. Peter Lang Verlag, v. 2, 85-97.
- Omena, Nelize P. de. (1996): A referência à primeira pessoa do discurso no plural. In: Giselle M. de Oliveira e Silva / M. Marta Scherre (eds.) *Padrões Sociolinguísticos*. Rio de Janeiro. Tempo Brasileiro, 183-215.
- Roberts, Ian (1993): *Verbs and diachronic Syntax*. Dordrecht. Kluwer.
- Roberts, Ian / Holmberg, Anders (2010): Introduction: parameters in Minimalist Theory. In: Teresa Biberauer et al. (eds). *Parametric Variation: null subjects in Minimalist theory*. Cambridge. Cambridge University Press, 1-57.
- Sankoff, David / Tagliamonte, Sali / Smith, Eric (2005): *Goldvarb Programs*, University of Canada, Ottawa.
- Scherre, Maria Marta P. / Yacovenco, Lilian / Paiva, Maria da Conceição (2019): An overview of Luso-Brazilian sociolinguistics: second person pronouns. In: Gabriel Rei-Doval / Ferando Tejedo-Herrero (eds.). *Lusophone, Galician and*

Hispanic linguistics: Bridging frames and traditions. London/NewYork. Routledge, 246-262.

Soares da Silva, Augusto (2013): El pluricentrismo del portugués: aspectos generales y elementos del enfoque sociolectométrico. In: Rudolph Muhr et al. (eds.). Exploring Linguistic standards in non-dominant varieties of pluricentric languages. Frankfurt am Main. Peter Lang, 79-90.

Dinah CALLOU

(Rio de Janeiro Federal University, Brazil;
National Scientific and Technological Development Council)
dinah@letras.ufrj.br

Carolina SERRA

(Rio de Janeiro Federal University, Brazil)
carolinaserra@letras.ufrj.br

Aline FARIAS

(Rio de Janeiro Federal University, Brazil)
alinefarias@ufrj.br

On R-deletion in final coda position: regional diversity in Brazilian Portuguese and syllable phonology¹

Abstract

We focus on the process of R-deletion in final coda position, regional diversity, and phonological syllable configuration, preceded by an overview on the origins of Brazilian Portuguese. The theoretical framework rests on the principles of language change and phonological syllable theories. Also, the Laboratory Phonology apparatus is used; a sample of 288 SVO sentences containing words ending in ‘ar’ (*celular* ‘mobile phone’) and ‘or’ (*professor* ‘teacher’), recorded in reading tasks performed by eight 20 to 40 year old female speakers with higher education from Rio de Janeiro. The analysis revealed that 1) R-deletion in final coda is one example of ongoing changes that present regional diversity and 2) traces of the loss of the segment are detectable in vowel duration, the time unit being re-associated with the syllable nucleus.

1. Introduction – On the history of Brazilian Portuguese

Brazilian Portuguese is spoken today by a population estimated of 213,317,639 speakers² distributed over a territory of eight and a half million square kilometres – as a result, it cannot be linguistically homogeneous. While the heterogeneity of Brazilian Portuguese is no longer a subject for recurrent discussion, the language’s origins do continue to rouse heated debate.

¹ The study was supported by grants from the Brazilian research foundations CNPq (process No.310528/2020-8) and CAPES (CAPES-PRINT Programme – process No.88887.508095/2020-00).

² Reference data: 1 November 2021. Source: Brazilian Official Statistics Bureau (Instituto Brasileiro de Geografia e Estatística, <https://www.ibge.gov.br/>).

Portuguese is the official and most spoken language³ in Brazil, coexisting with a set of other languages, which have very different histories of contact from that of the language first transplanted from the Iberian Peninsula in Brazil in 1500 by the first Portuguese colonisers. At present, that involves some two hundred indigenous languages, heritage languages (primarily European and Asian), and, in former *quilombo* areas, remnant speech from African languages, which has survived the language extermination sponsored by Portuguese-speakers.

For the other countries where Portuguese is also the official language (Angola, Mozambique, Cabo Verde, Guinea-Bissau, São Tomé and Príncipe, and, more recently, East Timor), quite recent studies have endeavoured to unveil the directions that Portuguese is taking. Unlike the Brazilian case, however, in other former colonies, the process of political independence occurred only a short time ago and European Portuguese continues to be the prestige language. In Angola and Mozambique (which became independent in 1975), the language situation is complex and Portuguese is expanding primarily through schooling. In these countries then, the extinguishing of the local Bantu languages and the nativisation of Portuguese (which is not the mother tongue of most of the population) are more recent phenomena.

Differently from what has happened in Africa, economic interests played a fundamental role in the spread of Portuguese in Brazil (Faraco, 2018) and, during the 18th century, the metropolis set about applying language policies to introduce and spread Portuguese, which, until then, had not been the most spoken language in Brazil. This was done by way of royal decrees and orders between 1722 and 1759; the most emblematic one is the *Directorio dos Índios* (Indian Directorate), published in 1757 (Mariani, 2001, 2004). From then on, meaning from the late 18th century and most importantly over the course of the 19th century, an authentic grammar of Brazilian Portuguese took shape (Tarallo, 1996; Pagotto, 1998). Since the 19th century, Brazil's national identity has been fixed linguistically on Portugal's Standardisation, sometimes at odds with the local language system, and a paradoxical relationship has grown up: the more the dominant elites agreed on the rules of the disciplinary language Portuguese, the more Brazilian it became.

The fact is that the changes that have taken place over the last the centuries, both in Brazil and in Portugal, have distanced these two varieties of

³ In 2002, Law No. 10.436/2002 recognised Brazilian Sign Language as a legal means of communication and expression by the hearing impaired.

Portuguese from each other, to the point where their grammars display significantly different morpho-syntactic characteristics (Galves, 2001; Duarte, 2007, 2011; Duarte & Serra, 2015; Mendes, Duarte, Nascimento, Pereira & Estrela, 2015; Duarte, Gomes & Paiva, 2018). In a way, Brazilian Portuguese (BP) has also become a variety of reference, even for the European Portuguese. The extent of that influence, however, has yet to be investigated.

While the morpho-syntax of BP offers abundant material for description, this is even more true of its phonetics and phonology, the field addressed by this study. The aim is to establish a correlation between the phenomenon of variable R-deletion in final coda position (*celular* ~ *celula*∅ ‘mobile phone’; *cantor* ~ *canto*∅ ‘singer’) and regional diversity. It also endeavours to ascertain, from a phonological standpoint, how the grammar is reorganised as an effect of ongoing linguistic change. More precisely, it asks what vestiges the R-deletion phenomenon may be leaving in its wake, while the change is not yet entirely complete. The study hypothesis is that the syllable’s duration is reconfigured when the segment is deleted: *di.ta.doR* → *di.ta.do*∅ → *di.ta.do:* (‘dictator’).

2. State of the art

A number of studies have focussed on variability in the production of coda rhotics in BP (Callou, 1987; Callou, Leite & Moraes, 1996; Brandão, Mota & Cunha, 2003; Hora & Monaretto, 2003; Monaretto, 2010; Serra & Callou, 2013; Xavier, 2020; Kailer & Almeida, 2020; Serra, Callou, Korol & Martins, 2021, and many others). These studies show that implementation of R in final coda position ranges from an alveolar trill to a voiceless laryngeal fricative (aspiration), leading to variable loss of the segment in the final stage of a weakening process and to syllable simplification (CVC → CV).

In fact, R-deletion is a very old phenomenon in Portuguese and was initially a marker of social stratification, detected in non-dominant varieties. During the 16th century, as can be seen in Gil Vicente’s plays, it was used to signal black slaves’ speech. From that century on, the phenomenon has gradually expanded to all social classes and levels of education, and (i) it is no longer stigmatized and (ii), according to Labov (1994), it is a linguistic change below the level of the speaker’s conscious awareness.

The authors’ analyses of recent speech samples in the Language Atlas of Brazil project (*Projeto Atlas Linguístico do Brasil*, ALiB), drawn from towns in the interior of the northeastern, southeastern, and southern regions of Brazil

(Oliveira, Santana & Serra, 2014; Serra & Callou 2013, 2015; Callou, Serra & Cunha, 2015; Oliveira, Santana, Xavier & Serra, 2018; Farias, Caldas & Serra, 2018; Serra, Callou, Korol & Martins, 2021), based on the theory of language change (Weinreich et al., 1968), and making use of multivariate analysis software (Goldvarb X, Sankoff et al., 2005) and sociolinguistic analysis (Labov, 1994), reveal that R-deletion is basically affected by three factors:

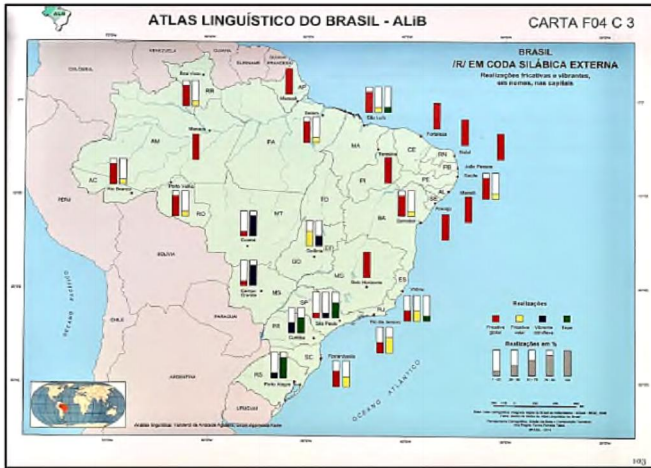
- (i) morphological word class (R-deletion is always more frequent in verbs), i.e., this is a linguistic factor that drives change⁴;
- (ii) word length measured by the number of syllables (related to the phonic salience principle), since the process is less frequent in monosyllabic words (mar ‘sea’; flor ‘flower’);
- (iii) the speaker’s region of origin, since frequency of R-deletion in coda is related to the R-variant produced in each region of Brazil, when the segment is not deleted.

The geolinguistic investigation of the ALiB Project⁵ gave rise to the mappings of final coda R shown below, which contemplate 25 Brazilian state capitals. Note that, in the North, Northeast, and part of the Southeast and Mid-West, the more innovative fricative productions ([+back]) predominate, while in the South, the more conservative vibrant productions ([-back]) are more frequent (Figures 1 and 2). As regards weakening of R in final coda position, the regions where the [+back] variants predominate are also where final R-deletion is more frequent (Figures 3 and 4)⁶.

⁴ The deletion process operates differently by morphological category: in verbs (in the infinitive *querer* ‘to want’ and subjunctive *quiser* ‘want’), the syllable coda is a redundant morphological marker co-occurring with the word accent, which does not occur in words of other morphological classes. When two markers exist to indicate the verb morphology, one of them – the R – is deleted.

⁵ The *Atlas Linguístico do Brasil* (Cardoso et al., 2014) presents current trends in Brazilian speech from all parts of the country, taking into account information on the diastatic component, secondary and higher education; the generational component, contemplating age groups I (18 to 30 years) and II (50 to 65 years); and the sex component, using both male and female informants. The survey of 1,100 informants of the ALiB Project in the five regions of Brazil (25 state capitals + 225 sites in the interior) gave rise to phonetic, semantic-lexical, and morpho-syntactic maps that form part of the publication and reveal a rich scenario of language diversity.

⁶ In the figures 1 and 2, the colours represent the production of R as follows: Red– glottal fricative, Yellow–velar fricative, Dark blue– retroflex flap, and Dark green– tap. In the figures 3 and 4, the Yellow column represents R-deletion and the Red column, implementation of the R segment.

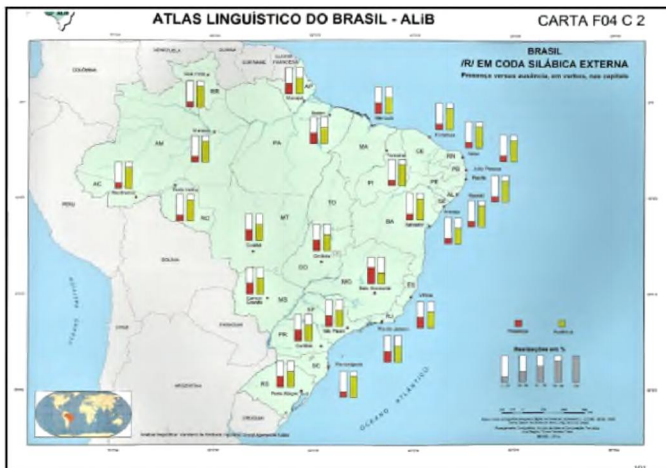


Figures 1 and 2 – Distribution of production of R in final coda position in nouns (above) and in verbs (below) – Letters F04 C3 and C4 (Cardoso et al., 2014).





Figures 3 and 4 – Distribution of R-deletion in final coda position in nouns (above) and in verbs (below) – Letters F04 C1 and C2 (Cardoso et al., 2014)



The phenomenon occurs at different stages in different communities in Brazil and the assumption is that the process is quite advanced in the northeastern region as compared to the southern region. The authors were able to confirm that, like all kinds of linguistic change, the overall process of R-

loss involves both social and linguistic stimuli, and constraints. More recently, the hypothesis has been investigated that when the coda segment is dropped, mora is maintained, possibly by lengthening the preceding vowel (Farias, 2018), in line with the conclusions of acoustic studies on acquisition of the coda constituent (Mezzomo, 2003). This study has tested this hypothesis in a corpus of controlled speech, considering certain variables.

3. Theoretical framework, corpus, and methodology

In order to investigate the relation between variable R-deletion in final coda and the possibility that the time unit of the deleted coda is re-associated with the syllable nucleus, this study applied the theoretical and methodological toolset of Variationist Sociolinguistics (Labov, 1994), using computer software for multivariate analysis (Goldvarb X – Sankoff et al., 2005), the syllable structure proposed by Hyman (1985: 87), and the principles of Laboratory Phonology, in order to constitute a corpus and to perform acoustic analysis of syllable time units (moras) (Praat software – Boersma & Weenink, 2012-2017).

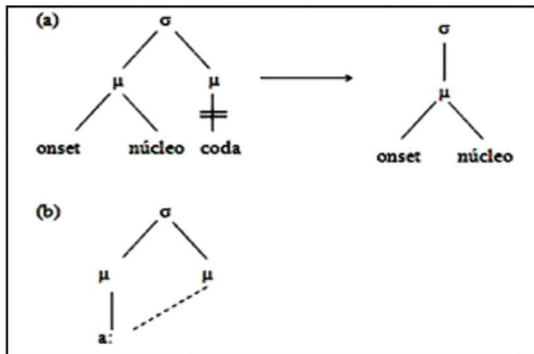


Figure 5: Mora deletion (a) and mora re-association (b)

The dataset analysed was obtained from a reading task conducted using a PowerPoint presentation with a set of phrases to be read one by one, which was recorded in August 2019 in a soundproofed room at the Acoustic Phonetics Laboratory of the Faculdade de Letras, Universidade Federal do Rio de Janeiro (UFRJ). The 36 target sentences were randomised with another 72 distraction sentences. The participants were eight female speakers, born in Rio de Janeiro, with higher education and 20 to 40 years old at the time the recording was made. The option for beginning the analysis with female speakers was due to

practical reasons, but similar analysis is expected with men (the dataset with male speakers is not yet complete), although, till now, there is no sociolinguistic hypothesis to justify the procedure, besides the fact that it would be necessary to adjust the acoustic values, if we consider both male and female individuals.

In building the experimental sentences, the following conditions were met: SVO (Subject-Verb-Object) sentence structure pattern, the least artificial semantic content possible, and the next word not starting with a vowel, in order to avoid re-syllabification. In the variationist analysis, the following variables were tested statistically: 1) words with stressed final syllables and nuclear vowel [a] or [o] (in Portuguese, the most productive words with R in coda contain these syllable nuclei); 2) words with two or three syllables, all nouns; 3) the context following the R, whether consonant or pause; 4) the quality of the following consonant; 5) the lexical item containing the R in coda; 6) words in three different positions in the prosodic hierarchy, prosodic word (Pw), phonological phrase (PhP), and intonational phrase (IP). The 36 sentences read by the 8 speakers are shown in Table 1.

Prosodic Phonology Theory (Nespor & Vogel, 1986, 2007) posits that a speech stream can be divided into hierarchically organised constituents, viz., syllable (σ) < foot (Σ) < prosodic word (Pw) < phonological phrase (PhP) < intonational phrase (IP) < phonological utterance (U). Each prosodic constituent works as a domain of (non-)application of segmental processes and occurrence of diverse melodic markings (Nespor & Vogel, 1986, 2007; Frota, 2000; Tenani, 2002; Serra & Callou, 2015). Accordingly, as proposed by Serra & Callou (2013, 2015), the hypothesis investigated was that deletion is applied differentially by the prosodic boundaries of Pw, PhP, and IP⁷, and that the vowels tested behave differently with regard to their duration. As sentence end (IP final) is an important context for determining melodic patterns in Portuguese, this context may possibly preserve more the segment and, when the R is deleted, syllable lengthening may be greater at this prosodic boundary.

Data set: 36 sentences (x 8 speakers)

⁷ In Portuguese, a Pw contains a single primary stress (Vigário, 2003); the PhP is formed by a lexical head and the functional elements on its non-recursive side (the left side, in Portuguese) within the same maximal syntactic projection and, optionally, the non-branching complement phrase of this lexical head, if within the same maximal projection (Frota, 2000; Tenani, 2002); the IP is formed by every sequence of adjacent phonological phrases of a root sentence or by every sequence of adjacent phonological phrases that is not incorporated structurally into the syntactic tree of the sentence (Nespor & Vogel, 1986, 2007; Frota, 2000).

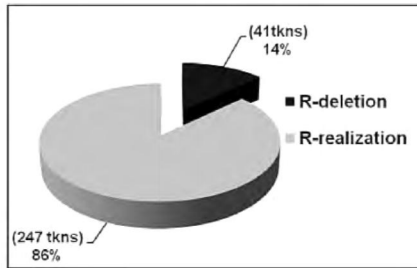
Vowel	Number of syllables	Prosodic boundary
/a/	2 syllables <i>ba.zar</i> 'sale' <i>co.lar</i> 'necklace' <i>lug.ar</i> 'place'	As freiras trabalham no bazar]Pw semanal. Minha irmã comprou um colar]Pw preto. O presidente inaugurou um lugar]Pw perfeito.
		As freiras trabalham no bazar]PhP todo dia. Minha irmã comprou um colar]PhP para mim. O presidente inaugurou um lugar]PhP semana passada.
		As freiras trabalham no bazar]IP Minha irmã comprou um colar]IP O presidente inaugurou um lugar]IP
	3 syllables <i>cel.u.lar</i> 'mobile' <i>mi.li.tar</i> 'soldier' <i>ti.tu.lar</i> 'regular'	Meu pai comprou um celular]Pw preto. O exército convoca o militar]Pw competente. O técnico escolheu o titular]Pw permanente.
		Meu pai comprou um celular]PhP para mim. O exército convoca o militar]PhP todo ano. O técnico escolheu o titular]PhP semana passada.
		Meu pai comprou um celular]IP O exército convoca o militar]IP O técnico escolheu o titular]IP
/o/	2 syllables <i>a.tor</i> 'actor' <i>tu.mor</i> 'tumour' <i>can.tor</i> 'singer'	Meu pai encontrou o ator]Pw famoso. O médico descobriu um tumor]Pw perigoso. O programa premiou o cantor]Pw famoso.
		Meu pai encontrou o ator]PhP semana passada. O médico descobre um tumor]PhP todo ano. O programa premiou o cantor]PhP semana passada.
		Meu pai encontrou o ator]IP O médico descobriu um tumor]IP O programa premiou o cantor]IP
	3 syllables <i>pro.fe.ssor</i> 'teacher' <i>di.ta.dor</i> 'dictator' <i>tra.i.dor</i> 'traitor'	A escola contratou um professor]Pw fantástico. A polícia prendeu o ditador]Pw cruel. O povo encontrou o traidor]Pw fugitivo.
		A escola contrata um professor]PhP todo ano. A polícia prendeu o ditador]PhP semana passada. O povo encontrou o traidor]PhP semana passada.
		A escola contratou um professor]IP A polícia prendeu o ditador]IP O povo encontrou o traidor]IP

Table 1: Data set – variables tested

The durations of all 288 vowels produced in the target words were measured in seconds, using the Praat software. The vowel duration measurements were recorded in Excel.

4. Results and discussion

The overall results pointed to an R-deletion rate of 14% (41/288), as shown in Graph 1. This result is not surprising, given that it is a corpus obtained from reading and the target words are nouns, which tend to retain the segment more. The statistical software Goldvarb X (Sankoff; Tagliamonte; Smith, 2005) indicated no considerable differences by vowel type and, accordingly, deletion rates were similar for /a/ and /o/.



Graph 1: R-deletion with final -ar and -or

The only variable found statistically to be active in R-deletion was the number of syllables in the word. Once again, it was found that in longer words, there was a greater tendency for the segment to be suppressed in final syllable coda, as shown in Table 2.

Number of syllables	Tkns/Total	%	PR
2	9/144	6%	0.16
3	32/144	22%	0.83

Table 2: R-deletion by number of syllables (Input: 0.10)

The results of the acoustic analysis showed that, in all contexts, vowel duration was greater following R-deletion, as can be seen in Tables 3 and 4, below.

Prosodic Boundary	Number of syllables	Target word	Mean vowel duration	% Lengthening
Prosodic word	2	/a/ with R (<i>co.laR</i>)	0.096s	
		/a/ without R (<i>co.la∅</i>)	0.124s	29%
Prosodic word	3	/a/ with R (<i>ce.lu.laR</i>)	0.110s	

		/a/ without R (<i>ce.lu.la∅</i>)	0.135s	23%
Phonological phrase	2	/a/ with R (<i>co.laR</i>)	0.108s	
		/a/ without R (<i>co.la∅</i>)	0.142s	31%
Phonological phrase	3	/a/ with R (<i>ce.lu.laR</i>)	0.113s	
		/a/ without R (<i>ce.lu.la∅</i>)	0.137s	21%
Intonational phrase	2	/a/ with R (<i>co.laR</i>)	0.127s	
		/a/ without R (<i>co.la∅</i>)	0.178s	40%
Intonational phrase	3	/a/ with R (<i>ce.lu.laR</i>)	0.131s	
		/a/ without R (<i>ce.lu.la∅</i>)	0.168s	28%

Table 3: Mean duration of /a/ in words with 2 and 3 syllables, by prosodic boundary

The results pointed to at least 21% lengthening of the vowel /a/ when the speaker did not pronounce the R. At the phonological phrase (PhP) boundary, the duration of the vowel /a/ in two-syllable words –in *cola∅* as compared to *colaR*, for example– was found to be 40% greater. The duration of the vowel /o/ was found to be at least 17% greater when the R is deleted.

Prosodic Boundary	Number of syllables	Target word	Mean vowel duration	% Lengthening
Prosodic word	3	/o/ with R (<i>can.toR</i>)	0.089s	
		/o/ without R (<i>can.to∅</i>)	0.109s	22%
Prosodic word	3	/o/ with R (<i>di.ta.doR</i>)	0.096s	
		/o/ without R (<i>di.ta.do∅</i>)	0.122s	27%
Phonological phrase ⁸	2	/o/ with R (<i>can.toR</i>)	0.099s	
Phonological phrase	3	/o/ with R (<i>di.ta.doR</i>)	0.094s	
		/o/ without R (<i>di.ta.do∅</i>)	0.111s	18%
Intonational phrase	2	/o/ with R (<i>can.toR</i>)	0.106s	
		/o/ without R (<i>can.to∅</i>)	0.136s	28%
Intonational phrase	2	/o/ with R (<i>di.ta.doR</i>)	0.121s	
		/o/ without R (<i>di.ta.do∅</i>)	0.142s	17%

Table 4: Mean duration of /o/ in words with 2 and 3 syllables, by prosodic boundary

⁸ There were no cases of R-deletion in this prosodic context.

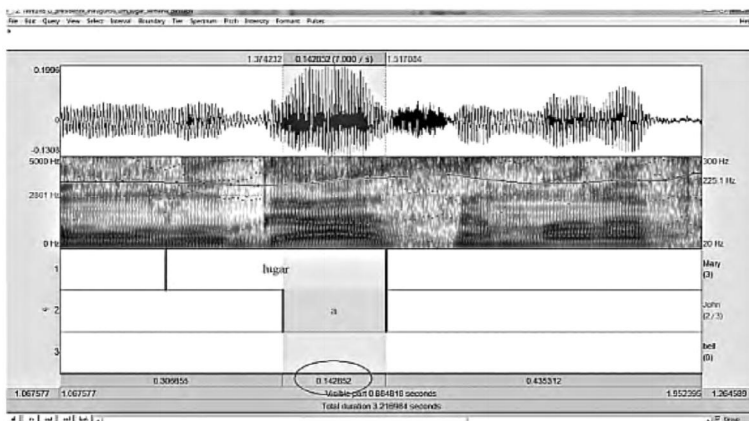


Figure 6: Duration of /a/ in the word *lugaØ* where the R is deleted (0.142s)
 O presidente inaugurou um *lugaØ* semana passada. (Sp 1, PhP).

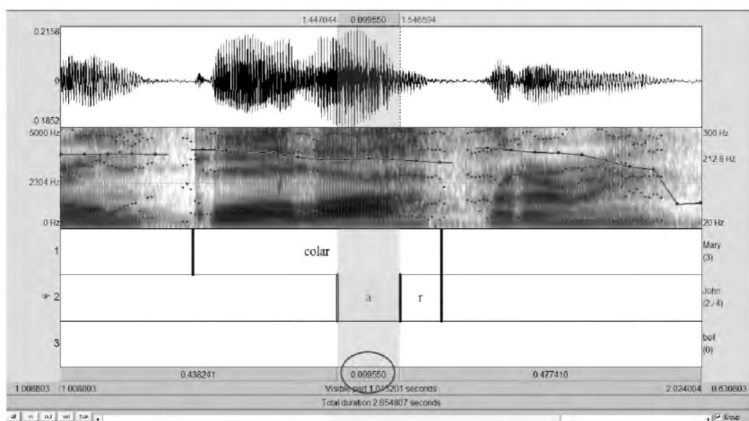


Figure 7: Duration of /a/ in the word *colaR* where the coda is produced (0.099s).
 Minha irmã comprou um *colaR* para mim. (Sp 1, PhP).

Figures 6 and 7 taken from Praat illustrate the differences in duration between /a/ in the word *lugaØ*, where the R is deleted, and in the word *colaR*, where the coda is produced (0.142s and 0.099s respectively). Previous research (Farias, 2018) using a small sample of spontaneous speech (ALiB) has pointed to compensatory lengthening when R is deleted. The acoustic analysis found average compensatory lengthening of close to 75%. Those results and those of this study offer input to support the hypothesis that the time unit (mora)

associated with the coda is re-associated with the nucleus when the segment is deleted.

5. Conclusion

To summarise, the R-deletion rate in reading and in nouns is still low, but it could be found in the speech of individuals born in Rio de Janeiro city. Also, word length once again proved relevant for analysing the R-deletion process, corroborating prior findings regarding this phenomenon. Given both the vitality of the phenomenon, which is extremely diverse in different dialectal areas, and the multiple paths of innovation found in Brazilian speech, a great deal has been written on consonantal processes and still more remains to be discovered through new analytical approaches, both in reading and in spontaneous speech. Acoustic analysis revealed a tendency to compensatory vowel lengthening when R is not produced. More in-depth studies on the subject are required but the findings of this study agree with those of Farias (2018) of the spontaneous speech of Teresina and Correntes, in the northeastern Brazilian state of Piauí: vowel lengthening is present as a strategy and seems to be a vestige of the ongoing change that is leading to R-deletion in final coda.

References

- Boersma, P. / Weenink D. (2012-2017): Praat: doing phonetics by computer [computer software]. Phonetic Sciences, University of Amsterdam. Available at <http://www.praat.org>.
- Brandão, S. F. / Mota, M. A. / Cunha, C. S. (2003): Um estudo contrastivo entre o português europeu e o português do Brasil: O -R final de vocábulo. In: Brandão S. F. & Mota, M. A. (org.). *Análise Contrastiva de Variedades do Português*. Rio de Janeiro: In- Fólio, pp. 163-180.
- Callou, Dinah Maria Isensee (1987): *Variação e Distribuição da Vibrante na Fala Culta do Rio de Janeiro*. Rio de Janeiro: UFRJ/PROED.
- Callou, D. / Leite, Y. / Moraes, J. (1996): *Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil*. In: *Gramática do Português Falado*, v. VI, I. Koch, (ed.), pp. 465-493. Campinas: UNICAMP.
- Callou, D. M. I / Serra, C. R. / Cunha, C. (2015): *Mudança em curso no Português Brasileiro: o apagamento do R no dialeto nordestino*. *Revista da ABRALIN*, v. 14, pp. 195-219.
- Cardoso, S. et al. (2014): *Atlas Linguístico do Brasil*. *Cartas linguísticas*, v. 2. Londrina: EDUEL.

- Duarte, Maria Eugênia (2007): Sujeitos de referência definida e arbitrária: aspectos conservadores e inovadores na escrita padrão. In: *Revista Lingüística – Revista do Programa de Pós-graduação em Lingüística* – v. 3, n. 1, UFRJ, pp. 89-115.
- Duarte, Maria Eugênia Lammoglia (2011): A língua portuguesa carioca na virada do século XIX; reflexões sobre a constituição na norma culta no Brasil. *Revista de Ciências da Vida*, v. 33, pp. 79-90.
- Duarte, M. E. / Serra, C. R. (2015): Gramática(s), ensino do português e 'adequação linguística'. *Matraga* (Rio de Janeiro), v. 22, pp. 31-55.
- Duarte, M. E. L. / Gomes, C. A. / Paiva, M. C. A. (2018): The implementation of endogenous syntactic features in Brazilian standard writing. In: Murh, Rudolf; Meisnitzer, Benjamin. (Org.). *Pluricentric Languages and Non-Dominant Varieties Worldwide: New Pluricentric Languages - Old Problems*. 1ed. Berlin: Peter Lang, pp. 429-442.
- Faraco, Carlos Alberto (2018): Aspectos da história socioeconômica e linguística do Brasil. *Diadorim*, Rio de Janeiro, v. 20 – Especial, pp. 23-52.
- Farias, Aline (2018): O apagamento do rótico na (re)organização silábica. *Dissertation (Masters in Portuguese Language) Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro*.
- Farias, A. / Caldas, V. / Serra, C. (2018): Sobre o processo de apagamento do rótico em coda silábica: diversidade regional. *Diadorim*, Rio de Janeiro, v. 20 – Especial, pp. 365-389.
- Frota, Sonia (2000): *Prosody and focus in European Portuguese. Phonological phrasing and intonation*. New York: Garland Publishing.
- Galves, Charlotte (2001): *Ensaio sobre as gramáticas do português*. Campinas, Editora da Unicamp.
- Hora, D. / Monaretto, V. (2003): Enfraquecimento e apagamento dos róticos. In: D. Hora & G. Collischonn (orgs.). *Teoria Linguística: fonologia e outros temas*. João Pessoa: Editora Universitária, pp. 114-143.
- Hyman, Larry (1985): *A Theory of Phonological Weight*. Dordrecht: Foris.
- Kailer, D. A. / Almeida, É. F. (2020): As variantes róticas em coda silábica no interior de Santa Catarina. In: Cláudia Regina Brescancini & Valéria Neto de Oliveira Monaretto. (Org.). *Sociolinguística no Brasil: textos selecionados*. 1ed. Porto Alegre: BC-PUCRS, v. 1, pp. 1-339.
- Labov, William (1994): *Principles of linguistic change. Internal factors*. Cambridge, Blackwell.
- Mariani, Bethania (2001): "A Institucionalização da língua, história e cidadania no Brasil do século XVIII: o papel das academias literárias e da política do

- Marquês de Pombal”, in Orlandi, E.P. (org.) História das idéias lingüísticas no Brasil. Campinas, Unemat Editora/Pontes, pp. 36-58.
- Mariani, Bethania (2004): Colonização linguística. Campinas, Pontes.
- Mendes, A. / Duarte, M. E. L. / Nascimento, M. F. B / Pereira, L. / Estrela, A. (2015): Pronominal Constructions and Subject Indetermination in Varieties of Portuguese: a global view on norms. In: Muhr, R.; Marley, D. (org.). Pluricentric Languages: New Perspectives in Theory and Description. 1ed.Frankfurt am Main: Peter Lang, v. 17, pp. 109-125.
- Mezzomo, Carolina (2003): A análise acústica como subsídio para a descrição da aquisição do constituinte coda. In: Letras de Hoje. Porto Alegre, v. 38, n. 2, pp. 75-82.
- Monaretto, Valéria (2010): Descrição da vibrante no português do sul do Brasil. In: Bisol, L & Collischonn, G. (Orgs). Português do Sul do Brasil. Porto Alegre: EDIPUCRS, pp.119-127.
- Nespor, M. / Vogel, I. (1986): Prosodic Phonology. Dordrecht: Foris.
- Nespor, M. / Vogel, I. (2007): Prosodic Phonology: with a new foreword. Walter de Gruyter GmbH & Co., Berlin.
- Oliveira, I. C. / Santana, M. / Serra, C. R. (2014): Apagamento do rótico em coda silábica interna e externa: a região serrana do Rio de Janeiro. Anais da 25ª Jornada Nacional do GELNE, pp. 1-11.
- Oliveira, I. C. / Santana, M. / Xavier, K. / Serra, C. R. (2018): O rótico em coda silábica final na região Sul do Brasil: variação e mudança no Corpus do ALiB. Diadorim, Rio de Janeiro, v. 20 – Especial, pp. 334-364.
- Pagotto, Emilio Gozze (1998): Norma e condescendência: ciência e pureza. Línguas e Instrumentos linguísticos 2, pp. 49-68.
- Sankoff, D./ Tagliamonte, S. / Smith, E. (2005): Goldvarb X: a variable rule application for Macintosh and Windows. University of Toronto, Department of Linguistics.
- Serra, C. / Callou, D. (2013): A interrelação de fenômenos segmentais e prosódicos: confrontando três comunidades. XXVIII Encontro Nacional da Associação Portuguesa de Linguística. 1ed. Coimbra: APL, v. 1, pp. 585-594.
- Serra, C. / Callou, D. (2015): Prosodic structure, prominence and /r/-deletion in final coda position: Brazilian Portuguese and European Portuguese contrasted. In: De Dominicis, A. (org.). pS-prominences: Prominences in Linguistics International Conference. Viterbo: DISUCOM PRESS, pp. 96-113.
- Serra, C. R. /Callou, D. M. I. /Korol, C. /Martins, L. B. (2021): Variação e mudança do rótico em coda final: a região Sul resiste (como pode?). In: Marins, J. E.; Orsini, M. T.; Cavalcante, S. R. O. (org.). Contribuições à descrição e ao

- ensino do Português Brasileiro: da fonética ao discurso, com parada obrigatória na sintaxe: uma homenagem a Maria Eugênia Lammoglia Duarte. 1ed. São Paulo: Pimenta Cultural, v. 1, pp. 20-55.
- Tarallo, Fernando (1996): "Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX". *Língua e cidadania*. Campinas, Pontes.
- Tenani, Luciani Ester (2002): *Domínios prosódicos do português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos*. Campinas: UNICAMP, 2002. Tese (PhD in Linguistics), Programa de Pós-Graduação em Linguística, Instituto de Letras, UNICAMP, Campinas.
- Vigário, Marina (2003): *The prosodic word in European Portuguese*. Mouton de Gruyter.
- Weinreich, L. / Herzog, M. (1968): "Empirical Foundations for Theory of Language Change". In: Lehmann, Paul; Malkiel, Yakov. (eds.) *Directions for Historical Linguistics*. Austin: University of Texas Press: 95-188. [Fundamentos empíricos para uma teoria da mudança linguística. Trans.: Marcos Bagno; Technical review: Carlos Alberto Faraco. São Paulo: Parábola, 2006.]
- Xavier, Karilene da Silva (2020): *As múltiplas pronúncias do rótico na música popular brasileira do século XX: da vibrante à fricativa e ao apagamento*. 249 f. Thesis (PhD in Portuguese Language) Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Christina GOMES

(Federal University of Rio de Janeiro, Brazil)
christinagomes@letras.ufrj.br

Marcelo L. MELO

(Federal University of Rio de Janeiro, Brazil)
malmelo.lopes@letras.ufrj.br

Linguistic and Social Patterns of Sound Variation in two Contemporary regional varieties of European and Brazilian Portuguese¹

Abstract

In this paper, we compare the results for linguistic and social patterns of the same sound change, with data collected from two samples of contemporary regional varieties of European and Brazilian Portuguese, from Lisbon and Rio de Janeiro respectively. We argue that a full comprehension of the varieties of a pluricentric language demands a dynamic view of language, according to which each variety may develop its own patterns of linguistic and social variation since, depending on their sociohistorical relationship, they may or may not influence each other (Labov, 2007). By evaluating the similarities and dissimilarities of the observed patterns of the competing variants of coda (s), such as in *antes* (before), that have been developed in each variety, we aim to address whether the outcomes observed in varieties of pluricentric languages are totally independent or still linked by features connected to the variety from which they have originated.

1. Introduction

According to Muhr (2018: 48), Non-Dominant (ND) varieties of pluricentric languages must be described on the basis of their own standards, instead of taking the exogenous standards of the Dominant Variety (DV). In this paper, we argue that a full comprehension of the varieties of a pluricentric language demands a dynamic view of language, according to which each variety may develop its own patterns of linguistic and social variation since, depending on their sociohistorical relationship, they may or may not influence each other.

¹ Christina A. Gomes is also supported by the Brazilian Funding Agency Conselho Nacional de Pesquisa/CNPq.

The issue of how the varieties of a pluricentric language may influence each other, and the way they develop their linguistic patterns regarding their relationship with the DV, has been addressed by many authors in different studies about different groups of pluricentric languages. For instance, according to Meyerhoff and Niedzielski (2003), several studies about New Zealand English (NZE) identified a shift from old, more British-like norms, to newer, more American-like ones, implemented in the last half century, even in the absence of face-to-face transmission channels, due probably to globalization, or, more specifically, to Americanization. These studies showed the increasing use by younger New Zealanders of words from U. S. English, such as *napkin*, *can*, *movies*, *pants*, and *truck*, that compete with words inherited from the transplanted British English, such as *serviette*, *tin*, *pictures*, *trousers*, and *lorry*. The studies also documented two sound changes in progress related to the dental alveolar /t/. The change towards a flapped intervocalic variant *r*, in words such as *butter*, is attributed to the influence of U. S. English, whereas the change towards a word final glottal *ʔ*, as in *get lost*, is attributed to British English influence. The use of *be like* as a quotative verb (*I was like 'oh my God this isn't happening'*) is also attributed to the influence of U. S. English in New Zealand English. In relation to the lexical variation, the authors conducted a pilot study to assess the perception of some of the words that show variability between the ones inherited from British English and the new U. S. English forms. They observed a consistent agreement among participants on the evaluation of *lift*, *jersey* and *biscuit* as absolutely and very New Zealand on one hand, and of *cookie* as not at all New Zealand, while others showed the same pattern but not consistently among all the responses. Then, it was possible to detect that some words are spreading and are perceived as exotic, while others are being nativized or perceived as local.

Meyerhoff and Niedzielski (2003: 545-546) argued that the possibility of an independent, parallel development cannot be ruled out for all the detected changes. Still, according to the authors, even if the hypothesis of the globalization of U. S. English is an issue, it doesn't mean that these forms are aligned to a self-conscious non-Zealand identity. On the contrary, they are part of the vernacular in Labov's terms, defined as the variety used in the speaker's most unself-conscious states (Labov, 1972). However, the authors don't rule out the possibility of an influence of the American and the British varieties in Contemporary NZE, due to the contact through different types of media (television, internet etc). They argued that these changes may be completely

independent, since parallel development in the varieties of pluricentric languages cannot be ruled out for all the detected changes. On the other hand, in relation to the sound changes, they can be due to the adoption of an exogenous pattern, and, in this case, they state that patterns adopted from another variety or language tend to keep the same constraints observed for the original variety or language, which would consist of evidence of the relationship between the change in NZE and in American and British English.

However, according to Labov (2007), changes resulting from the adoption of patterns from another variety not spoken locally, defined as diffusion across communities, implies in the weakening of the original pattern and the loss of structural features in opposition to changes that are transmitted within the speech community, which faithfully reproduce the patterns transmitted through generations, a pattern characteristic of the family tree model. And then, for Labov, since adults are the ones that tend to adopt exogenous patterns, the outcome of a change characterised as diffusion across communities is attributed to different learning abilities of children and adults. Adults do not learn and reproduce linguistic patterns with the same accuracy and speed displayed by children, since a decline in learning abilities from the ages of 9 to 17 years old has been observed, especially after the critical period (Johnson; Newport, 1989; Newport, 2000; Scovel, 2000).

Considering this theoretical background, in this paper we compare the results for linguistic and social patterns of the same sound change, with data collected from samples of two contemporary regional varieties of European (EP) and Brazilian Portuguese (BP), respectively Lisbon and Rio de Janeiro. By discussing the similarities and dissimilarities of how the patterns of the competing variants of coda (s), such as in *antes* (before), developed in each variety, we aim to address the following main research question: in what manner are the outcomes observed in varieties of pluricentric languages totally independent or still linked by features connected to the variety from which they have originated? This is an important issue since, due to advances in technology of mass media communication, it is possible that, in the globalized world, different varieties of a pluricentric language may influence each other even in the absence of face-to-face communication, as it is considered by Meyerhoff and Niedzielski (2003) in relation to the influence of British and American English on contemporary NZE.

Although it is outside of the purpose of this study, we point to another question that we considered important to address in the studies of pluricentric

languages: how is the influence of the DV on NDV, if it exists, perceived by local speakers of the NDV variety? More recently, the influence of Brazilian YouTubers among Portuguese children has been reported in relation to the use of some Brazilian words instead of the correspondent European Portuguese word. According to the newspaper *Diário de Notícias*, Portuguese children, influenced by Brazilian YouTubers, have been replacing some EP words by their corresponding BP word, for instance, such as *grama* instead of *relva* (grass), *ônibus* instead of *autocarro* (bus), and *listras* instead of *riscas* (stripes). The extent of this kind of linguistic loan and the way it is perceived by different EP speakers is an interesting line of investigation.

Regarding the status of Contemporary EP and BP as varieties of pluricentric Portuguese, several authors have pointed out that there is a symmetrical relationship between EP and BP, and that both languages constitute Dominant Varieties of pluricentric Portuguese (Silva, 2013:88; Kretzenbacher/Hajek/Norrby, 2013: 261; Duarte/Gomes/Paiva, 2016: 61). It has been noted that, regarding the phonological organisation of both languages, the developmental path of different aspects of phonology has led to outstanding differences between both languages (Massini-Cagliari/Clagliari/Redenbarger, 2016). In relation to the coda (s), as in *nós* (we) and *cesta* (basquet), Rodrigues and Hora (2016) presented an overview of the observed differences between EP and BP. Therefore, by comparing more deeply varieties of each language that share some similarities but differ in relation to others, this paper seeks to deepen the understanding of how the sound change observed for coda (s) had developed in each variety and to what extent they are linked due the previous historical relationship between the two languages or are independent processes.

2. Comparing linguistic and social patterns of the coda (s) in EP and BP

In this section, we develop a comparison between what was observed in relation to coda (s) in Lisbon and Rio de Janeiro, based mainly on the results obtained in Rodrigues (2005) and Melo (2012, 2017), both studies with data collected from samples of spontaneous speech.

Rodrigues (2005) conducted a study about the varieties spoken in Lisbon and Braga based on data of spontaneous speech. She collected samples formed by local residents of each city whose speech was obtained using the sociolinguistic interview. She interviewed 78 individuals from Lisbon, 49 females and 29 males, distributed according to levels of age (13-19; 20-25; 26-39; 40-55, 56-up) and schooling (illiterates, elementary school, high school, university). She

collected 2,628 tokens of word final coda (s). Table 1 presents the distribution of the variants of coda fricative according to the following phonetic context.

Variants	[ʃ]		[ʒ]		[z]		∅		Total
	N	%	N	%	N	%	N	%	
Unvoiced Consonant/Pause	1146	100	0	0	0	0	-	-	1146
Voiced Consonant	89	13.7	558	86.2	0	0	-	-	647
Vowel, Glide	144	21.2	24	3.5	510	75.2	-	-	678
/S/ or other sibilant	75	85.2	13	14.7	0	0	-	-	88
Other Contexts	-	-	-	-	-	-	69	100	69
Total	1454	55.3	598	22.6	510	19.4	69	2.6	2628

Table 1. Distribution of the variants of final (s) in Lisbon according to following context
Source: Rodrigues (2005: 234)

It is observed that final coda (s) is mainly produced as a post-alveolar fricative, whose voicing tends to match the voicing of the following context, 100% of unvoiced postalveolar before unvoiced consonant, and 86.2% of voiced postalveolar when the following context is a voiced consonant. However, a low percentage of unvoiced codas was observed when followed by a voiced consonant (13.7), a vowel or a glide (21.2). When the following context is a vowel or glide, the coda is mainly realised as the voiced fricative [z], although the occurrence of voiced and unvoiced fricatives is possible in this context. These distributions indicate that there is not a categorical assimilation of voicing in the contexts where the coda is followed by a voiced consonant or a vowel or a semivowel. The place of articulation is variable in these contexts as well. Finally, absence of coda was observed only in cases of total assimilation of contiguous similar segments, such as another fricative as in *caos social* (social chaos) (Rodrigues, 2005: 233-240).

According to Rodrigues (2005: 244), any kind of relationship was not observed between the different variants and speakers' social profiles, such as level of education and age, which means that, in the speech community of Lisbon, the variants of coda (s) are not related to any social attribute of the speaker. In other words, the variants of coda (s) are not used in social indexicality.

Regarding Brazilian Portuguese, Melo (2012) conducted a study about the variety spoken in the city of Rio de Janeiro based on data collected from a subset of speakers from two samples of spontaneous speech: a) 8 individuals, with a

middle and low-middle class profile, from Censo 2000 Sample, with 31 speakers distributed according levels of age (15-25; 26-49; 50-up), schooling (first half of elementary school, second half of elementary school, high school) and sex; and b) 8 individuals from EJLA 2008, with 16 socially-excluded slum-living adolescents with no access to formal education, who were in a reformatory when they were interviewed. These adolescents have very fragile ties with their families, no access to the formal labour market, and, thus, they are usually associated with illegal activities. They also have no access to formal education in a regular basis, and most of them had never finished the first years of schooling. Both samples were obtained using the sociolinguistic interview method.

Melo (2012) analysed all occurrences of syllables ending on (s), whether in the middle or at the end of the word. Previous studies about the BP variety spoken in Rio de Janeiro have shown that the n of coda (s) in BP include alveolar [s/z], post-alveolar [ʃ/z], back fricatives (velar and glottal) [x/χ, h/h] and the absence of coda, as in me[z]mo, me[ʒ]mo, me[h]mo, me∅mo, and that the post-alveolar variants are the most frequent phonetic occurrence of the coda (Scherre/ Macedo, 2000; Callou/Brandão, 2009). The absence of coda is possible when the following context is different from a pause, except when followed by a vowel. Another important characteristic of coda fricatives in BP is that, differently from EP spoken in Lisbon, the voicing of the following consonant determined the voicing of the coda (s), in a way that there is a categorical matching of voicing of the coda and the following consonant (Câmara Jr., 1977: 80; Cristóforo-Silva, 1999: 144). Table 2 shows the distribution of the tokens of each variant per each sample.

Variants	[j]/ [ʒ]		[s]/ [z]		[x,h]/[χ,h]		∅		Total
	N	%	N	%	N	%	N	%	
Sample									
Censo 2000	3949	74	1016	19	243	5	148	3	5356
EJLA 2008	1517	53	349	12	850	30	125	4	2841

Table 2. Distribution of the variants of coda (s) per sample in Rio de Janeiro
Source: Melo (2012: 68)

As can be observed in more recent speech samples of Contemporary BP, the post-alveolar variants are the most frequent phonetic realisation of the coda, even when the samples differ deeply in terms of the social-economic features of the speakers.

According to Callou and Brandão (2009), its prevalence in the speech community is the result of the implementation of a change (from alveolar to

post-alveolar fricative). The data obtained in Melo (2012) also showed that, when the syllable is in word final position followed by a vowel, only alveolar and back fricatives are possible. In this context, the alveolar is the most frequent variant (99% of the occurrence in Censo 2000 data, and 98,5%, in EJLA 2008 data). In relation to the back fricative variant, when socially-excluded lower-class speakers with irregular schooling are considered, the difference of overall rates in both samples stands out: 30%, for EJLA Sample, and 5% for Censo 2000 Sample speakers. It is around 4% for university speakers (Callou; Brandão, 2009). Taken together, the overall frequencies of back fricatives in each level of schooling, considering Melo (2012), and Callou and Brandão (2009) research, show that this variant is correlated to level of education. Melo (2017) conducted an experiment to observe the perception of the post-alveolar and the back fricatives by participants with different levels of education. The results obtained showed that the back fricatives are stigmatized by university, and high-school participants.

From what can be extracted from Rodrigues (2005) and Melo (2012), we can observe similarities and differences regarding the expression of coda (s) in EP and BP. These characteristics are organised in table 3.

	Comparison	Lisbon	Rio de Janeiro
Variants	partially similar	[ʃ], [ʒ], [z], ∅	+ [s], [x, γ, h, ð]
Overall rate of post-alveolar	Similar	prevalence of post-alveolar	
Coda Voicing assimilation	Different	variable in voiced contexts	Categorical
Following context – vowel	partially similar	prevalence of [z], and occurrence of both post-alveolars	Prevalence of [z], occurrence of back fricatives
Social stratification	Different	Absence of social constraints	Back fricatives more frequent among social-excluded low class speakers

Table 3: Similarities and differences between Lisbon and Rio regarding coda (s)

By the comparison between the results obtained in Rodrigues (2005) and Melo (2012), it is clear that the variety of European Portuguese spoken in Lisbon and the variety of Brazilian Portuguese spoken in Rio de Janeiro (also referred to as carioca) have evolved in the same direction toward the predominance of the post-alveolar fricative for the expression of coda (s).

However, the studies allow us to note that both varieties share partially the same variants, since, in the carioca variety, the differences between both

varieties are more than only different inventories of phonetic possibilities of the coda. The effect of the following context and the categorical assimilation of the voicing of the following context deepen the differences between both varieties. In EP, final word codas followed by a vowel can be produced, although less frequently, as an unvoiced or voiced post-alveolar fricative as well. This linguistic pattern was not observed in BP, although variation is also observed in this phonetic context, since Melo (2012) registered back fricatives in this environment. While in BP, the proportion of back fricatives is higher among low-class socially-excluded speakers and receives negative social evaluation among high school and university speakers, as shown by Melo (2017), in EP, none of the phonetic possibilities of coda (s) receive any kind of social evaluation.

3. Conclusion

The expression of coda (s) as a post-alveolar fricative is an innovation in the variety of Portuguese transplanted in the beginning of the 16th century. It is a strong feature of the variety spoken in the city of Rio de Janeiro, and it is the most frequent variant in other dialects from the North, Northern and South regions of Brazil (Marroquim, 1934; Silva Neto, 1956; Jesus/Mota, 2007), if compared to the ones with higher percentages or even categorical realisation of the alveolar fricatives. In relation to EP, Rodrigues (2005), examining data from Braga, provides further evidence of the non-obligatory assimilation of voicing at least when the coda is expressed by the unvoiced fricative and for the occurrence of both post-alveolar followed by a vowel or a glide.

According to Silva Neto (1956), it is unclear whether the presence of the post-alveolar in the city of Rio de Janeiro is an independent process or the consequence of the influence of the Portuguese spoken by the Portuguese nobles and the Royal Portuguese family that came to Brazil in 1808.

Despite the lack of historical evidence, if this pattern of prestige inherent to the variety spoken by the Royal Family served as a model for the variety spoken in Rio de Janeiro at the beginning of the 19th century, the development of the change that is observed in Rio is consistent with the type of change outlined by Labov (2007), diffusion across communities.

According to Labov, this type of change implies the adoption of an exogenous pattern by adults, whose learning abilities differ from that of children, leading to a different sociolinguistic pattern, since adults do not learn and reproduce linguistic patterns with the same accuracy and speed displayed by children. The sociolinguistic pattern developed in the carioca variety involves

the incorporation of a different variant (back fricative), a different effect of the following phonetic context in the distribution of the variants, including the obligatory/non-obligatory effect of voicing of the following segment. Then, the comparison between linguistic constraints of the variants in both cities showed that not only different patterns of conditioning, related to the following phonetic context, have developed, but also a new variant, the velar/glottal fricative, emerged in the variety spoken in Rio de Janeiro.

This situation is comparable to the ones outlined in Labov (2007) for the spread of the *short-a* system from New York to New Jersey, Albany, Cincinnati and New Orleans. He provided evidence that the complex conditioning of the tense and lax variants of *short-a*, that includes phonetic, morphological and lexical constraints, and that was developed in the variety spoken in New York in the eighteenth century, is partially observed in the contemporary English spoken in the mentioned cities.

According to Labov, the patterns observed are characterised by the loss of some structural details in the spread of the NYC *short-a* system to the dialects, due to the contact between them and/or the historical process of colonisation, since the adoption of a new variety in this specific context occurs through the contact among adults who change their own speech in a sporadic and inaccurate manner.

So, the differences observed between Rio de Janeiro and Lisbon can be related to the adoption of an exogenous pattern: this specific feature (postalveolar fricatives in coda) of the European Portuguese brought by the Royal family, at the beginning of the nineteenth century, was adopted by adults who changed their own speech in a sporadic and inaccurate manner, yielding a different pattern of phonetic constraints.

In addition, the specificities of the Brazilian society also added the social value component for the back fricative variants of coda (s) in the social ecology. It is worth mentioning that the research question presented in the Introduction “how is the influence of the DV onto NDV, if it exists, perceived by local speakers of the NDV variety?”, although very important in the study of pluricentric languages, is outside of the purpose of this study because there is no means to assess the social values of the variants spoken in Rio when the change towards the postalveolar began.

On the other hand, if the patterns observed in both varieties are parallel and independent processes, they provide interesting evidence about issues

related to the effect of the following context in triggering change processes that lead to different phonetic outcomes.

The results presented in this paper brought evidence of the dynamic character of pluricentric Portuguese varieties, although it was not possible to bring consistent evidence to decide whether or not the current characteristics of the Portuguese spoken in Lisbon and in Rio de Janeiro are linked by features present in EP at the beginning of the 19th century.

References

- Câmara Jr. Joaquim M. (1977): *Para o Estudo da Fonêmica Portuguesa*. Rio de Janeiro. Padrão.
- Cristófaros Silva, Thaís (1999): *Fonética e Fonologia do Português: roteiro de estudos e guia de exercícios*. São Paulo. Contexto.
- Callou, Dinah / Brandão, Silvia (2009): *Sobre o /S/ em coda silábica no Rio de Janeiro: falas culta e popular*. In: Ana Claudia Petters / Mônica M. G. SAVEDRA BARRETTO (eds.) *Sociolinguística no Brasil: uma contribuição dos estudos sobre línguas em/de contato: homenagem ao Prof. Jürgen Heye*. Rio de Janeiro. 7 Letras, 27-34.
- Diário de Notícias. „Há crianças portuguesas que só falam brasileiro“, downloaded from: dn.pt/sociedade/ha-criancas-portuguesas-que-so-falam-brasileiro-14292845.html, access: December 10th 2022.
- Duarte, Maria Eugénia L. / Gomes, Christina A. / Paiva, Maria da Conceição de (2016): *Codification and standardisation in Brazilian Portuguese*. In: Rudolf Muhr et al. (eds.). *Pluricentric Languages and Non-Dominant Varieties Worldwide: The pluricentricity of Portuguese and Spanish: New concepts and description*. Wien/Frankfurt am Main. Peter Lang Verlag, v. 2, 45-60.
- Kretzenbacher, H. / Hajek, J. / Norrby, C. (2013): *Address and introductions across two Pluricentric Languages in intercultural communication*. In: Rudolf Muhr et al. (eds.). *Exploring Linguistic standards in non-dominant varieties of pluricentric languages*. Frankfurt am Main. Peter Lang Verlag, 259-274.
- Johnson, J. / Newport, E. (1989): *Critical period effects in second-language learning: The influence of maturational state on the acquisition of English as a second language*. In: *Cognitive Psychology*, no. 21. 60-99.
- Jesus, C. dos S. de / Mota, J. A. (2006): *A variação fonética no português do Brasil: diferenças diatópicas na realização do /S/ em coda silábica, a partir dos dados do Atlas Linguístico do Brasil (AliB)*. Apresentado no Seminário Estudantil de Pesquisa do Instituto de Letras da Universidade Federal da Bahia, Salvador.

- Labov, William (1972): *Sociolinguistic Pattern*. Philadelphia, Pennsylvania. University of Pennsylvania Press.
- Labov, William (2007): Transmission and diffusion. In: *Language*, no. 83. 344-387.
- Marroquim, Mario (1934): *A língua do Nordeste (Alagoas e Pernambuco)*. São Paulo. Companhia Editora Nacional.
- Massini-Cagliari, G. / Cagliari, L. C. / Redenbarger, W. J. (2016): A Comparative Study of the Sounds of European and Brazilian Portuguese. In: Leo Wetzels / Sergio Menuzzi / João Costa (eds). *The Handbook of Portuguese Linguistics*. Oxford. Willey-Blackwell, 56-68.
- Melo, Marcelo A. S. L. (2012): *Desenvolvendo novos padrões na comunidade de fala: um estudo sobre a fricativa em coda na comunidade de fala do Rio de Janeiro*. Dissertação (Mestrado em Linguística), Universidade Federal do Rio de Janeiro.
- Melo, Marcelo A. S. L. (2017): *Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social*. Tese (Doutorado em Linguística), Universidade Federal do Rio de Janeiro.
- Meyerhoff, M. / Niedzielski, N. (2003): The globalization of vernacular variation. In: *Journal of Sociolinguistics*, no. 7(4). 534-555.
- Muhr, Rudolf (2018): Misconception about pluricentric languages and pluricentric theory – an overview of 40 years. In: Rudolph Muhr / Benjamin Meisnitzer (eds.). *Pluricentric Languages and Non-Dominant Varieties Worldwide: New Pluricentric Languages: Old Problems*. Frankfurt am Main. Peter Lang, 17-56.
- Newport, Elissa L. (2002). Critical periods in language development. In: L. Nadel (ed.) *Encyclopedia of Cognitive Science*. London. Macmillan. 737-740.
- Rodrigues, Maria Celeste (2005) *Lisboa e Braga: Fonologia e Variação*. Lisboa: Fundação Calouste Gulbenkian.
- Rodrigues, M. C. / Hora, D. da (2016): Main Current Processes of Phonological Variation. In: Leo Wetzels / Sergio Menuzzi / João Costa (eds). *The Handbook of Portuguese Linguistics*. Oxford. Willey-Blackwell, 504-525.
- Scovel, Thomas. (2000). A critical review of the critical period research. In: *Annual Review of Applied Linguistics*, no. 20. 213-223.
- Scherre, M. M. P. / Macedo, A. (2000): Restrições fonético-fonológicas e lexicais: o -S pós-vocálico no Rio de Janeiro. In: Maria Cecília Mollica / Mário Eduardo Martelotta (eds.). *Análises linguísticas: a contribuição de Alzira Macedo*. Rio de Janeiro. Serviço de Publicações – Faculdade de Letras – Universidade Federal do Rio de Janeiro, 52-64.

- Silva, Augusto S. da (2013): El pluricentrismo del portugués: aspectos generales y elementos del enfoque sociolectométrico. In: Rudolph Muhr et al. (eds.). Exploring Linguistic standards in non-dominant varieties of pluricentric languages. Frankfurt am Main. Peter Lang, 79-90.
- Silva Neto, Serafim. (1956) Introdução ao estudo da língua portuguesa no Brasil. Rio de Janeiro. Acadêmica.

Raquel Meister Ko. FREITAG

(Federal University of Sergipe, Brazil)

rkofreitag@academico.ufs.br

Mobility and higher education in grammatical patterns of Brazilian Portuguese

Abstract

Seeking evidence of the pluricentric nature of Brazilian Portuguese, this paper explores the hypothesis that a trigger for this process is the effect of the changes in higher education on linguistic patterns. Due to public policies for inclusion, the Brazilian educational system has expanded in the past decade and has changed the profile of students. The migration of students to a new community expanded their contact with different varieties of Portuguese. A study in a sample comprised of sociolinguistic interviews with undergraduates from the Federal University of Sergipe, Brazil, controlled three morphosyntactic variable features. Results suggest the role of the university as an inclusive and integrative field for linguistic variation in Brazilian Portuguese, with effects on the normative orientation of grammatical patterns. The observation of a set of grammatical shows different grammatical patterns, and this can support a pluricentric hypothesis.

1. Introduction

The recognition of Portuguese as a pluricentric language is not something new (Clyne, 1992, Baxter, 1992), with a quite peculiar situation in which a non-dominant variety, Brazilian Portuguese, has been occupying the place of the dominant language (Silva, 2010): if Brazilian Portuguese were considered as an independent language, by its number of speakers, it would be ranked #10 in the Etnologue, while the Portuguese spoken in the rest of the world would be ranked #38, together with Filipino (cf. Freitag, *to appear*).

However, Brazil is a country of continental dimensions, where only an apparent monolingualism is prevalent: there is different evidence of a process that raises Brazilian Portuguese to a pluricentric language, without a single dominant variety. This evidence can be seen in the results of descriptive linguistic documentation projects, such as the Norma Urbana Culta (NURC) project, which has provided the database for contemporary grammars of Brazilian Portuguese, and in sociolinguistic and dialectological studies. Even

stronger evidence of the pluricentric character is found in official documents for education, such as the *Base Nacional Curricular Comum* ('Common National Curricular Base') – BNCC, which recognises diversity without a polarizing axis of linguistic norm. The BNCC, by recognising the diversified part of the curriculum, gives legitimacy to regional varieties to establish themselves in a non-dominant relationship. The pluricentric notion of Brazilian Portuguese is also present in the *Programa Nacional do Livro Didático* ('National Textbook Program'), which considers the “prestigious urban norms” of Brazilian Portuguese.

Batoreo (2014) notes that Portuguese grammars (produced in Brazil and Portugal) do not point to a pluricentricity in the direction of considering the two national varieties, without attributing a dominant status to one of them. However, more and more Brazilian Portuguese grammars are taking into account regional diversity without a dominant normative axis. As Castilho (*to appear*) says: “With the change of the capital to Brasília and the development of other regions, a situation of cultural polycentrism, clearly foreseen by Rossi (1968/1969), began to occur in Brazil. Today, it is a useless task to seek in the speech of Rio, São Paulo or any other region a standard valid for the whole country. Throughout this grammar, I have repeated that we have several linguistic patterns whose prestige is valid for the regions in which they are practiced.”

Thus, today, Meisnitzer's (2020) view that “the normative center of the variety of the communicative distance is located in the axis [...] that goes from the urban centers of Rio de Janeiro and São Paulo” is not sustained. What Silva (2010) considers a problem of diglossia due to the great regional and social variation that is characteristic of Brazilian reality can be interpreted as evidence of a pluricentricity without a dominant variety.

Despite the empirical evidence and the planning in educational documents, the pluricentric reality of Brazil is still an obstacle to the teaching of a first language. Even though it is diverse, normative standards stemming from a Lusitanian grammatical tradition still prevail (Faraco, 2008): As Batoreo (2014) points out, there are still few linguistic instruments that fully adhere to this conception of Brazilian Portuguese as a pluricentric language, without a dominant variety.

Arguments for the standardisation of Brazilian Portuguese have been provided by sociolinguistic studies, especially those that compare variable behavior patterns according to the speakers' educational level; this is shown by the results of Duarte, Gomes and Paiva (2013), who observed the exposure of

social evaluation of linguistic features according to education in Rio de Janeiro. Higher levels of education in the same speech community reduced the frequency of features considered stigmatized, such as rhoticism. In some cases, usage patterns and prescriptive effects clash, as in the case of clitic collocation (Martins, Meisnitzer, 2016).

Education has, therefore, a crucial role in the linguistic configuration of Brazilian Portuguese; while linguistic diversity has been recognised as a learning right on a geographic dimension, speakers attribute social value to linguistic features due to their education. Therefore, it is natural to expect that in higher education institutions where speakers from different Brazilian sociolinguistic realities are gathered together, linguistic standards have been set. The recent changes in higher education have led to the observation of the coexistence of linguistic patterns of different communities in the same space, which has generated effects on teaching. The expansion of higher education in Brazil over the last few decades has resulted not only in a significant increase in registrations, but also in a diversification of regions and social origins of university students.

There is a new profile of undergraduates who have attended higher education: young people from lower classes who studied in public schools, but only after the creation of new educational policies did they reach public universities. These changes lead to contact between different linguistic norms and can trigger conflicts in higher education. The Faculty said new undergraduates are deficient in reading and writing skills, and they recognise that their difficulties affect their performance in classes and assessments.

This paper explores the situation of contact of different linguistic norms in higher education. To do so, a new approach to analyse variation is proposed, considering the covariation among different features.

2. Mobility and higher education in Brazil

For several decades, higher education in Brazil was a privilege. Although education is not equivalent to professional success, the probability of better-remunerated jobs is associated with higher education (Nonato, Pereira, Nascimento 2012): Before decades of dictatorial and neoliberal governments, President Lula's government took over the role of the State as an inductor agent of public policies at the educational level (Pereira, Silva, 2010, Carvalho, 2014).

Public policies, such as the *Expansão com Interiorização de Universidade Pública* ("Expansion with Interiorization of Public Universities") – EXPANDIR

program, created in 2005, and the support program for *Reestruturação e Expansão das Universidades Federais* plan ('Restructuring and Expansion of Federal Universities') - REUNI, were implemented with the objective of reducing asymmetries, especially regional and social inequalities in the offer of public higher education in Brazil. In addition to the creation of 18 new federal universities, six of which were in the Northeastern region of Brazil, a region characterised by social inequality in contrast with Southeast and South regions, the EXPANDIR and REUNI programs led to the creation of 182 new university campuses (Andifes, 2015): The landscape of public higher education in the Northeastern region of Brazil changed significantly between 2003 and 2010, with public policies for the expansion and countrywide extension of public universities, leading to an increase in the number of admissions in higher education, and contributing to the reduction of regional and social inequalities (Fig. 1).

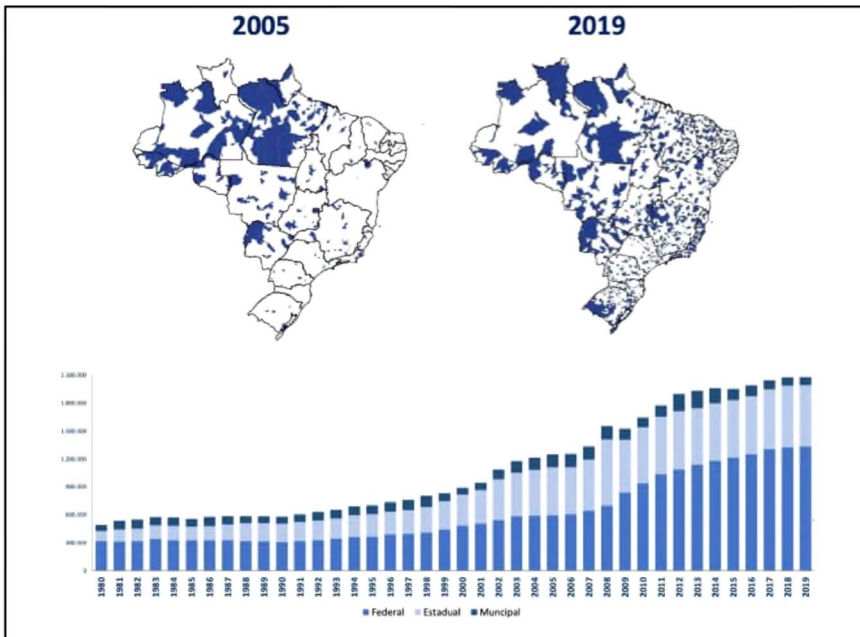


Fig. 1: Expansion of Brazil's public higher education. Data provided by Censo da Educação Superior ('Higher education census') 2019 (MEC/INEP):

Sistema de Seleção Unificada ('Unified Selection System) – SiSU and racial quotas laws are other policies that have contributed the most to student mobility and have made the expansion of regional, cultural, and ethnic diversity in Brazil's higher education institutions possible.

Considering student mobility, the new regional configuration of education in Brazil has affected the migration flows of undergraduate and graduate students in the country. This migratory flow leads to linguistic contacts, which modify relations with communities, social movements, and local and regional culture (Sankoff, 2001, Akar, 2010, Matias, Pinto, 2020).

The effects of student mobility are a multiplicity of social profiles inside universities, students from rural and urban areas from several states of Brazil, and students from different socioeconomic classes in the same living space. Students from different regions, in contact with each other, adapt their way of speaking according to the needs that this new role demands; and the demands of language adjustments are related not only to the new region, but also to the place where the contact happens, the university campus, which has social expectations regarding their new role of being undergraduates (Britto, Silva, Castilho, Abreu, 2008, Marques, Cepêda, 2012, Ristoff, 2014, Tomás, Silveira, 2021).

For instance, in this process, the Federal University of Sergipe expanded from 2 to 6 campuses, all new in the countryside. Plus, the adoption of SiSU in 2010 led to the enlargement of internal positions, with expansion moving to the countryside, and the promotion of policies/measures to attract students from other Brazilian states (expansion by mobility): *A língua do universitário: fala, leitura e escrita para o letramento acadêmico*¹ ('The University Student's Language: Speaking, Reading, and Writing for Academic Literacy') is an integrated project in progress at the Federal University of Sergipe, Brazil, which aims to describe these new linguistic patterns that are being configured at the university as a result of the linguistic contact arising from the expansion policies of Brazilian higher education.

¹ This is a research and teaching project that involves undergraduate (pedagogical support and scientific initiation) and graduate studies, with masters and doctoral research that provides a descriptive panorama of the linguistic reality of the institution's students, contributing to public policies and affirmative actions aimed at academic success and permanence, as well as the description of Brazilian Portuguese. This project was supported by *Programa de Apoio Pedagógico* ('Pedagogical Support Program'), (05/2020/UFS/PROGRAD/PROEST), which aims to grant scholarships to students who have committed themselves to taking part in complementary academic activities, providing support in subjects or issues related to their degree area.

The origin of this project stems from the evidence that the linguistic impacts of the expansion interfere with undergraduates' organisation and performance; difficulties with reading and writing are recurrent problems pointed out by Brazilian undergraduates in general. This difficulty becomes more pronounced with the recent democratisation of access to higher education and the entrance of a new profile of undergraduate, formed by those who, until then, were denied access to universities, such as young people from state-run schools and low-income households, and who, only after the creation of the most recent policies, have had the possibility to be admitted, creating a "new undergraduate" profile (Britto *et al.*, 2008):

These undergraduates have to fit their course together with a job, working during the day and studying at night. Getting a higher education degree allows people to advance to middle-level job positions. Meanwhile, the "classic undergraduate", who has historically enrolled in courses at the university, has family support to enroll in the more prestigious and daytime courses. Higher education offers them a way to maintain their social status.

These changes have not only interfered with the social structure of the academic communities and the neighbourhoods where they are located; they have also interfered with the language. Due to the access of socially disadvantaged classes and the academic mobility of students from different regions, each student carries with them values, cultures, norms, and linguistic features typical of their place of origin and social space. In a social space like the university, these exchanges lead to a scenario of linguistic conflict, which demands student retention policies that can reduce inequalities.

This is not an exclusive Brazilian Portuguese reality: Matias and Pinto (2020) have developed the *Trovoada de Ideias* ('Brainstorming') project, which aims to provide linguistic and social inclusion of students from Portuguese-speaking African countries in Portuguese higher education. This action recognises that there is a deeper understanding of the factors of (in)comprehension between students fluent in different norms of Portuguese and the university host community in Portugal.

3. Data collection in the context of mobility

To measure the effect of mobility on and in language in higher education, a set of specific methodological procedures for data collection were developed in the scope of the *A língua do universitário* ('The language of undergraduate') project, an action to reduce linguistic asymmetry in higher education. For this

goal, it was necessary to conduct linguistic documentation of students at the Federal University of Sergipe, constituting a database to support linguistic description, which is the basis for the preparation of grammars, dictionaries, and language teaching programs. The linguistic documentation procedures for this project present some differences from the classic sociolinguistic data collection procedure.

Instead of the concept of speech community, this data collection assumes a broader concept of “community of practice” (a group of people engaged in sharing social practices, who do not necessarily share linguistic behavior): In practical terms, in this data collection, the territorial/geographic dimension is an important feature in the recruitment and selection of participants for the sample.

Where and with whom the student lives and how they move up to campus are strong predictors of their social profile, and can improve information about their social background.

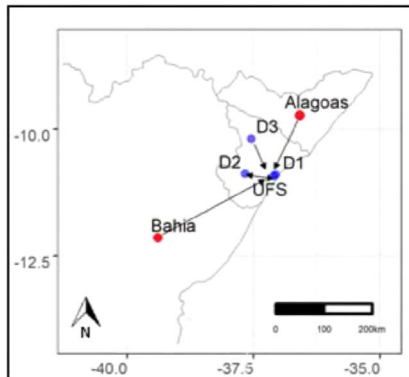


Fig.2: Geographical distribution of displacements in the sample.
Source: Siqueira and Freitag (2022):

Called *Deslocamentos* ('Displacements') sample,² this sociolinguistic documentation approach assumes an extended concept for a community of practices, coordinating stable and fluid sociodemographic stratification such as sex, age, education, origin, and vulnerability (in the same direction that is adopted by the university to offer financial assistance to students without

² The design, collection, and audio transcription of sociolinguistic interviews from the *Deslocamentos* sample complied with ethical requirements for research involving humans and were funded by Grants 02/2015 SENACON/MJ and CAPES/FAPITEC/SE 10/2016 PROMOB.

resources to support themselves during their undergraduate course) in a sample of sociolinguistic interviews in which participants are recruited as volunteers for convenience samples until cell saturation. It means that the field researchers go to a specific place where the speakers are in social practice, like a university. The result was a semi-controlled sample of speech, associated with social practice, with agency and stance.

Displacement sample was composed of the 60 sociolinguistic interviews balanced by gender, and number of years in the course, and also controls the students' degree of mobility; displacements D1, D2, D3, Bahia and Alagoas (fig. 2): D1-3 concerning students born and raised in the same state as the Federal University of Sergipe; Alagoas and Bahia concerning students from other dialectal areas. D1 students are from the capital, live with their parents or relatives, have their lunch on campus or at home, and do not participate in academic projects, like assistant research or assistant teacher, because they have work. D2 students are from the countryside of the State and move from home to campus every day. They also live with their parents or relatives and are engaged in academic projects. D3 students are from the state's countryside as well, but they live in the capital with roommates. They are also engaged in academic projects. Students from Alagoas and Bahia show similar behaviour to D3 students.

It is also important to discuss methodological issues in how to provide empirical evidence of the pluricentrism of certain languages. How does one measure the proximity or the distance between two or more varieties of the same language? Based on previous studies on Portuguese, different techniques have been used to show the pluricentric character. Silva (2010) proposed measuring the convergence/divergence of European and Brazilian Portuguese through the observation of lexical patterns, which he defines as the choices between denotational synonyms. Castro, Souza and Oliveira (2016) combined different techniques of automatic discrimination of varieties on a sample of Twitter.

However, especially in Brazil, sociolinguistic studies have contributed the most to provide evidence for pluricentrist hypotheses. For example, Duarte, Gomes and Paiva (2013) look into three variable phenomena in two samples characterised by the speakers' level of education. Each of the phenomena is analysed independently. This type of sociolinguistic analysis provides patterns of use of the phenomenon, not the community. Patterns of use of the phenomenon are important for deducing rules; the comparison of

convergence/divergence of varieties requires the identification of the patterns of use of each speaker in the community.

The type of approach that describes joint patterns of more than one variable is called covariation, and different strategies have been used, such as building a model from single studies (Guy, Oushiro, 2015), and correlation matrix (Beaman, 2021): In this paper, another strategy is tested: the clustering of patterns.

For this purpose, three grammatical processes described in the *Deslocamentos* sample are considered together at the same time: the variation in verbal agreement (Novais, 2021), in locative preposition for movement phrases (Rodrigues, 2021) and in determiner before possessive pronouns (Siqueira, 2020): These grammatical phenomena were described by observational studies for each grammatical feature with data from the *Deslocamentos* sample (n = 60), and the datasets of each one feature were joined for this analysis. From a sociolinguistic point of view, verbal agreement is certainly one of the most studied variables in Brazilian Portuguese (Mendes, Oushiro, 2015): The findings systematically point to internal constraints (phonic salience and verb position) and external constraints, such as the speakers' formal education. Non-agreement realisation is highly stigmatized (Scherre, 2005):

Prepositions governing locative complements of movement verbs in Brazilian Portuguese are in variation [a ~ para ~ em]: *Vou na casa da minha mãe* ('I'm going to my mother's house') or *Vou para a casa da minha mãe* or *Vou à casa da minha mãe* are attested realisations. As in agreement, there are prescriptivist effects: *a* and *para* are considered socially well-evaluated forms, while *em* is considered an incorrect form by grammarians. The same direction is shown by sociolinguistic studies (Vieira, 2009, Wiedemer, 2010):

The presence or absence of an article in determiner position (D) before possessives in NP as in *Eu vi o seu irmão* ('I saw the your brother') and *Eu vi ø seu irmão* ('I saw ø your brother') behaves as a dialectal marker. Although there is no prescriptive effect in grammar, a dialectal effect is observed; varieties from the South and Southeast of Brazil tend to fill in the determiner position, while varieties from the Northeast tend to omit it. In *Deslocamentos* sample, the absence of articles in D-position is the pattern; in other regions, such as Alagoas, D-presence is the pattern (Siqueira, Freitag, 2022): Combining the use of the variables can help us to understand dialectal patterns and differentiate them from social patterns. Considering the effects of mobility, an observational multivariate study to identifying the clustering patterns by speakers in

Deslocamentos sample can provide evidence for understanding and recognition of language varieties and dialects of Brazilian Portuguese, and contribute with evidence to the pluricentric hypothesis.

4. Data analysis

A new analysis was carried out with the dataset consisting of the combination of the datasets of each individual study (Siqueira, 2020, Novais, 2021, Rodrigues, 2021), shared following the Open Science principles for reproducible science.³

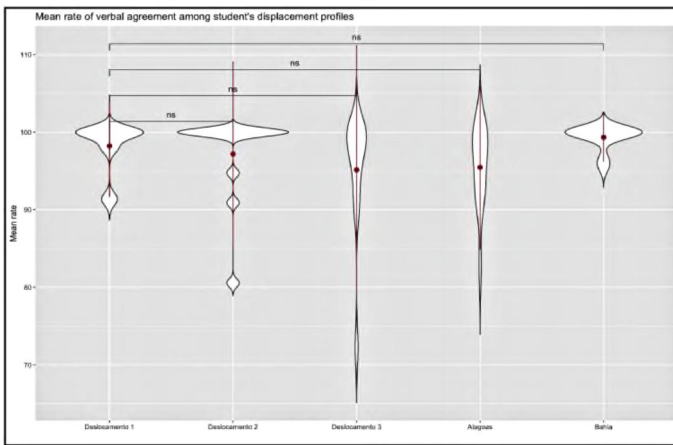


Fig. 3: Mean rates of verbal agreement in 3pp.

Each variable was observed in a separate model and the mean rates of absence of article before possessive pronouns (D-absence), verbal agreement (Vagr) and standard preposition in locative (PrepP) by each of 60 students were computed and compared with the effect of students' gender, time in the course, and displacement profile.⁴ Differences between the global mean rate for verbal agreement ($M = 97.07$, $SD = 5.39$) and the mean rate by displacement profile are not significant (fig. 3): The global mean rate for D-absence ($M = 57.60$, $SD = 16.17$) is affected by displacement profile group (fig. 4): there are no significant

³ The dataset, and the full reproducible code were shared in Open Science Framework.

Rodrigues (2021): <https://osf.io/d4qp7/>

Siqueira (2020): <https://osf.io/p6zdt/>

Novais (2021): <https://osf.io/c9btf/>

⁴ Statistics were done using R 4.0.3 (R Core Team, 2021), the rstatix v 0.6.0 (Kassambara, 2020) and the factorextra v 1.0.7 (Kassambara, Mundt, 2020) packages. Data visualization were done using ggplot2 (Wickham, 2016): The full reproducible code is available in <https://osf.io/9vqtq/>.

differences among the mean rates for Sergipe displacement profiles, that are significantly greater than the mean rates for Alagoas ($M = 55.68$, $SD = 8.46$), and Bahia ($M = 44.36$, $SD = 17.06$).

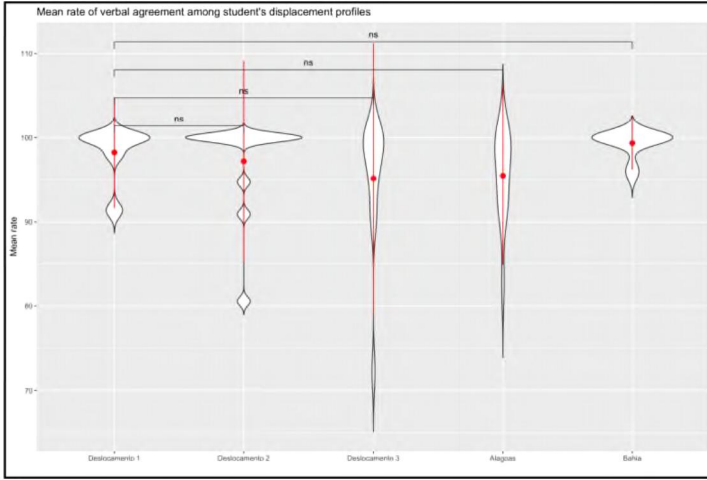


Fig. 4: Mean rates of D-absence

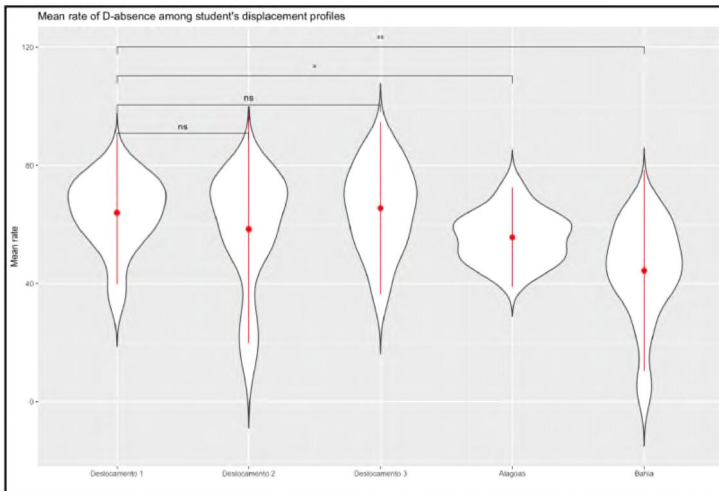


Fig 5: Mean rates of standard preposition in PrepP.

The results rate for standard preposition in PrepP (Fig. 5) follow the same direction: there is no significant difference between the global mean rate ($M = 85.63$, $SD = 17.74$) and the Sergipe and Alagoas displacement profiles, only Bahia ($M = 74.01$, $SD = 20.53$).

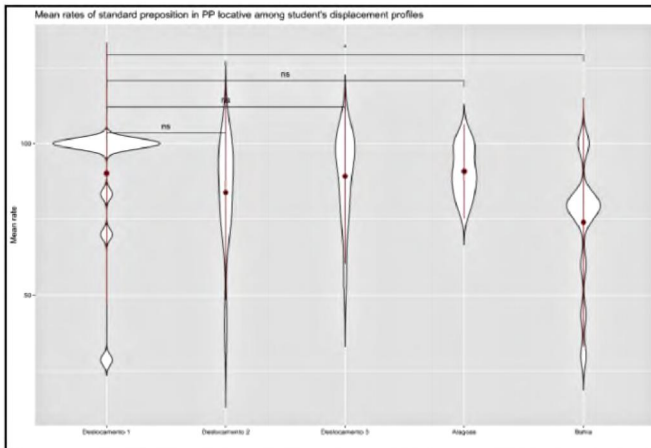


Fig. 6: Mean rates of standard preposition in PrepP.

A linear model for each one of the variables (Fig. 6) shows that time in the course is not significant. For D-absence and standard preposition in PrepP the effect of the Bahia displacement profile is statistically significant and negative.

Seeking patterns of behaviour of individuals concerning the three variables, a cluster analysis using K-medoids technique to cluster the data set of mean rates objects into four clusters determined by silhouette tool (Fig. 7): This process is more robust than k-means, considering the range of variance among mean rates.

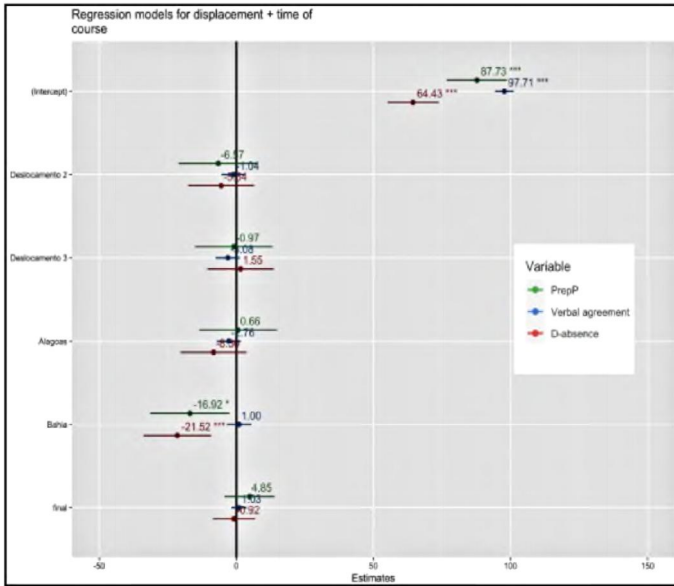


Fig. 7: Linear model for mean rate ~ displacement + time.

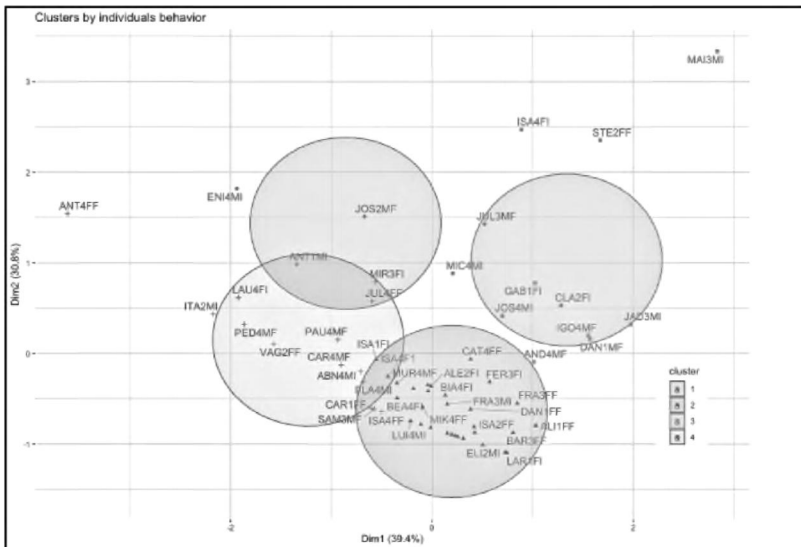


Fig. 8: Clustering analysis.

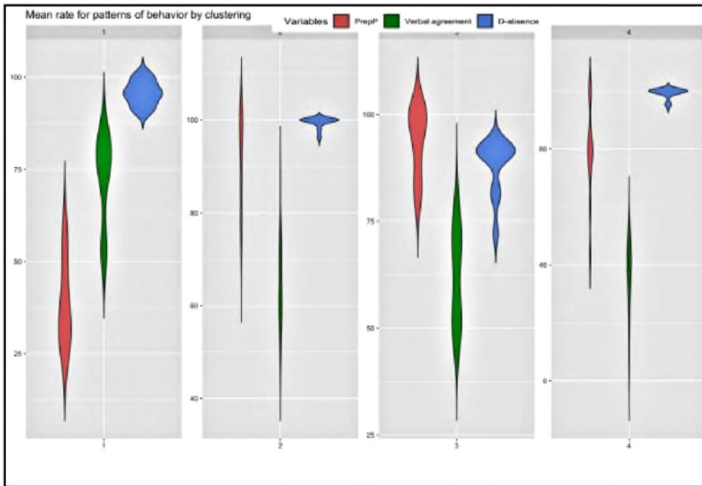


Fig.9: Group patterns by clustering.

The clustering analysis (Fig. 8) identified four groups that follow different grammatical patterns.

Group 1 is the smallest; four students comprise this group, joined by the smallest mean rate of standard preposition in PrepP ($M = 38.53$, $SD = 12.44$), D-absence ($M = 72.39$, $SD = 13.39$) and verbal agreement ($M = 95.65$, $SD = 3.63$) above global rates. The single common feature for these students is that they are beginners and none of them are from Alagoas.

Group 2 is the biggest, with thirty students joined by a high mean rate of standard preposition in PrepP ($M = 91.73$, $SD = 9.88$), D-absence ($M = 64.16$, $SD = 8.79$) above the global rate, and a verbal agreement mean rate nearly categorical ($M = 99.45$, $SD = 1.19$): This group is balanced in gender and time in the course. Most students are from Sergipe, rather than Alagoas and Bahia.

Group 3 computes eleven students joined by the highest mean rate of standard preposition in PrepP ($M = 93.79$, $SD = 8.03$), D-absence ($M = 62.07$, $SD = 11.26$) above the global rate and the smallest mean rate for verbal agreement ($M = 88.07$, $SD = 6.72$): In this group, male beginners from Alagoas are predominant.

Group 4 includes fourteen students who have the lowest mean rate of D-absence ($M = 35.69$, $SD = 13.21$), the lowest mean rate of standard preposition in PrepP ($M = 79.58$, $SD = 15.37$), which is lower than the global rate, and a nearly categorical mean rate of verbal agreement ($M = 99.23$, $SD = 1.70$): Most of the students in this group are from Bahia and are finalists.

5. Conclusion

By examining the integration index associated with each of the groups, certain profiles emerge, and it suggests that social identity and social interactions in constitution affect grammatical patterns in this sample. These results highlight the role of the university as an inclusive and integrative field for the linguistic variation in Brazilian Portuguese, with effects on the normative orientation of grammatical patterns.

The splitting of three variable processes into three grammatical patterns, combining dialectal patterns and social patterns, suggests that the analysis and recognition of language variants and dialects of Brazilian Portuguese is a more complex task than supposed by the sociolinguistic approach and requires large datasets and different statistical approaches.

However, the findings also highlight the importance of grammatical features in different varieties of Brazilian Portuguese, lending credence to the evidence that it is a pluricentric language. Although preliminary, these results reinforce that the differences among varieties of Brazilian Portuguese go beyond phonology: grammatical features that, combined, can lead to different grammars with problems of comprehension, such as those emerging in Brazilian higher education after expansion.

References

- Akar, Hanife. (2010): Challenges for schools in communities with internal migration flows: evidence from Turkey. In: *International Journal of Educational Development* 30/3, P.263-276.
- Associação Nacional dos Dirigentes das Instituições Federais de Ensino Superior. Programa de expansão, excelência e internacionalização das universidades federais. Brasília, 2012. http://www.andifes.org.br/wp-content/files_flutter/1360930928PEEXIU.pdf.
- Batoréo, Hanna. (2014): Que gramática (s) temos para estudar o Português língua pluricêntrica? In: *Revista diadorim* 16, P. 01-15.
- Baxter, Alan Norman. (1992): Portuguese as a pluricentric language. In: *Pluricentric languages: differing norms in different nations*, Mouton de Gruyter, P. 11-23.
- Beaman, Karen V. (2021): Exploring an approach for modelling lectal coherence. In: *Language Variation – European Perspectives VIII*, P.135-160.
- Britto, Luiz Percival Leme, Silva, Edineuza Oliveira Silva, Castilho, Katlin Cristina de Castilho, Abreu, Tatiane Maria Abreu (2008): *Conhecimento e formação nas*

- IES periféricas perfil do aluno “novo” da educação superior. In: Avaliação: Revista da Avaliação da Educação Superior, 13/1, P. 777-791.
- Carvalho, Cristina Helena Almeida de (2014): Política para a educação superior no governo Lula: expansão e financiamento. In: Revista do Instituto de Estudos Brasileiros, P. 209-244.
- Castilho, Ataliba Teixeira de (2010): Nova gramática do português brasileiro. São Paulo, Contexto.
- Castilho, Ataliba Teixeira de (to appear): Nova gramática do português brasileiro, second edition.
- Castro, Dayvid, Ellen Souza, and Adriano de Oliveira. (2016): Discriminating between Brazilian and European Portuguese National Varieties on Twitter Texts. In: 5th Brazilian Conference on Intelligent Systems (BRACIS), P. 265-270.
- Clyne, Michael. (1992): Pluricentric languages - introduction. In: Pluricentric Languages: Different Norms in Different Nations, Mouton de Gruyter, P. 1-9.
- Faraco, Carlos Alberto. (2008): Norma culta brasileira: desatando alguns nós. São Paulo, Parábola Editorial.
- Freitag, Raquel Meister Ko. (to appear): Linguistic repositories as asset: Challenge for sociolinguistic approach in Brazil. In: Digital Archives.
- Guy, Gregory, and Oushiro, Livia (2015): The effect of salience on covariation in Brazilian Portuguese. In: University of Pennsylvania Working Papers in Linguistics 21/2, P.1-18.
- Kassambara, Alboukadel (2020): rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.6.0. <https://CRAN.R-project.org/package=rstatix>
- Kassambara, Alboukadel / Mundt, Fabian (2020): factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>
- Maria Eugênia Lammoglia Duarte, Christina Abreu Gomes, and Maria da Conceição Paiva (2013): Patterns of variation in non-dominant varieties of Brazilian Portuguese. In: Exploring linguistic standards in non-dominant varieties of pluricentric languages, Peter Lang, P. 331-342.
- Marques, Antonio Carlos Henriques, and Vera Alves Cepêda (2012): Um perfil sobre a expansão do ensino superior recente no Brasil: aspectos democráticos e inclusivos. In: Perspectivas: Revista de Ciências Sociais 42, P.161-192
- Martins, Marco Antonio, and Meisnitzer, Benjamin (2016): The use of clitics in Brazilian Portuguese: the development of an endogenous standard variety. In: Muhr, Rudolf et.al. (2016): Pluricentric Languages and Non-Dominant

- varieties Worldwide. Volume 2: The pluricentricity of Portuguese and Spanish: New concepts and descriptions. Peter Lang, P. 67-84.
- Matias, Ana Raquel, and Paulo Feytor Pinto (2020): Overcoming linguistic barriers in Portuguese higher education: The case of international African students. In: Portuguese Journal of Social Science 19.2-3), P. 189-214.
- Meisnitzer, Benjamin. (2020): Interfaces entre didáctica y lingüística: las variedades europea y brasileña del portugués en la didáctica del portugués como lengua extranjera y la competencia variacional receptiva. In: Quaderns de Filologia: Estudis Lingüístics XXV, P. 185-201.
- Mendes, Ronald Beline, and Oushiro, Livia (2015): Variable number agreement in Brazilian Portuguese: an overview. In: Language and Linguistics Compass 9/9, P. 358-368.
- Nonato, Fernanda, Pereira, Rafael Moraes, Nascimento, Paulo Meyer., and Araújo, Thiago Costa (2012): O perfil da força de trabalho brasileira: trajetórias e perspectivas. Boletim de mercado de trabalho-conjuntura e análise. Report. Instituto de Pesquisa Econômica Aplicada (IPEA): <http://repositorio.ipea.gov.br/handle/11058/3884>
- Novais, Viviane Silva de (2021): Variação na concordância verbal de terceira pessoa do plural na fala de universitários sergipanos. 2021. Dissertação (Mestrado em Letras), Universidade Federal de Sergipe.
- Paiva, Maria da Conceição de, and Scherre, Maria Marta Pereira. (1999): Retrospectiva sociolingüística: contribuições do PEUL. In: DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada, 15/spe, P. 201-232.
- Pereira, Thiago Ingrassia, and Silva, Luís Fernando Santos Correa da (2010): As políticas públicas do ensino superior no governo Lula: expansão ou democratização? In: Revista Debates 4/2. P. 10-31.
- Ristoff, Dilvo (2014): O novo perfil do campus brasileiro: uma análise do perfil socioeconômico do estudante de graduação. In: Avaliação: Revista da Avaliação da Educação Superior 19/3, P. 723-747.
- Rodrigues, Fernanda Gabrielle Costa (2021): Variação na regência de complementos locativos de verbos de movimento na fala de universitários da UFS. Dissertação (Mestrado em Letras), Universidade Federal de Sergipe.
- Sankoff, Gillian. (2001): Linguistic Outcomes of Language Contact. In: Handbook of Sociolinguistics. Basil Blackwell, P. 638-668.
- Scherre, Maria Marta Pereira. (2005): Doa-se lindos filhotes de poodle: variação lingüística, mídia e preconceito. São Paulo, Parábola Editorial.

- Silva, Augusto Soares da (2010): Measuring and parameterizing lexical convergence and divergence between European and Brazilian Portuguese. In: *Advances in cognitive sociolinguistics* 45, P. 41-84.
- Siqueira, Manoel (2020): *Variação no preenchimento da posição determinante antes de possessivos pré-nominais: padrões dialetais e contatos*. Dissertação (Mestrado em Letras), Universidade Federal de Sergipe.
- Siqueira, Manoel, and Freitag, Raquel (2022): Can mobility affect grammar at the syntax level? A case study in Brazilian Portuguese. In: *Organon*, 37/72, P.1-20.
- Tomás, Maria Carolina, and Leonardo Souza Silveira. (2021): Expansão do ensino superior no Brasil: diversificação institucional e do corpo discente. In: *Revista Brasileira de Sociologia* 9/23 P.149-177.
- Vieira, Maria José Blaskovski. (2009): *Variação das preposições em verbos de movimento*. In: *Signum: Estudos da Linguagem* 12/1, P.423-445.
- Wickham, Hadley (2016): *ggplot2: Elegant Graphics for Data Analysis*. New York, Springer-Verlag.
- Wiedemer, Marcos Luiz. (2010): A atuação dos fatores sociais na seleção das preposições de regência do verbo ir (movimento) na fala de Santa Catarina. In: *Revista Estudos Linguísticos* 39/2, P. 640-655.dsdd

IV. The Pluricentricity of French and Dutch in the Americas

Karine GAUVIN

(Université de Moncton, New Brunswick, Canada)
karine.gauvin@umoncton.ca

On the matter of an endogenous norm in Acadia (New Brunswick, Canada)

Abstract

This contribution focuses on the reasons why an endogenous norm has not yet emerged in French Acadia. Under review are two dominant vernaculars, firstly a “mixed” variety called Chiac, and secondly, Acadian French. Typical of urban areas, Chiac incorporates lexical items borrowed from English into the local French matrix (mostly nouns, adjectives, and verb radicals, as well as some adverbs), while Acadian French is characterized by archaisms that date back to the first settlements of the 17th century. Unlike Chiac, Acadian French has more positive connotations and remains in use by the elderly and in rural areas, but both varieties are perceived to be antithetical to Standard French practices.

1. Introduction

For a long time, standard language ideology dominated the Francophone community, in which speakers of French would strive to speak a homogeneous language, regardless of their geographic origins or social conditions. However, now we are witnessing a gradual erosion of this hegemony, especially regarding terminology standardization, dictionary production and feminization practices. The larger centres such as Quebec, Belgium and Switzerland are experiencing a growing legitimization of prestige varieties in their respective territories, but what about smaller, “fringe”, communities? For instance, can French in Acadia lend itself to some form of standardization? This question stems from a broader ambition to describe what I have called “Acadian Standard French” (*le français standard acadien*), a field of study that remains, to this day, largely unexplored (Gauvin, 2014). Belgian linguist Michel Francard, who is well acquainted with the Acadian sociolinguistic situation noted, “There is an urgent need to carry out a double effort in Acadia: to describe the Acadian variety and its norms following the example of what has

been undertaken for Quebec French, and to obtain a social consensus to recognize and disseminate this variety, notably on the part of teachers” (Francard, 2005:382)¹. This call to action has, to this day, remained unanswered.

2. Historico-political context of Acadia

[t]hese events helped to shape the political situation that led to the imposition of the famous oath of allegiance to the British crown, which the Acadians refused to take, and which served as a pretext for the Deportation, a tragic event that left an indelible mark on the social, political, economic, and linguistic evolution of Maritime Acadia. As a result of these deportations, the Acadian community went from being a community that felt it could influence its own destiny to a collection of groups that were scattered throughout the Maritime provinces and were excluded for nearly a century from any participation in social and political life. (2005: 82)

Collège Saint-Joseph and *Le Moniteur Acadien* both played a leading role in the genesis, development, and social organization of Acadian nationalism. Around their core, what traditional historiography would call “the leaders of the Acadian nation” would be grafted. We know that several organizers and speakers of the first great National Conventions, to mention only Pascal Poirier, Pierre Landry, Father Camille Lefèbvre and Father Philius Bourgeois, were directly linked to Collège Saint-Joseph. As for *Le Moniteur Acadien*, the national elite of the time, classically trained and eager to discuss and display their thoughts in a French – and Catholic – inspired newspaper, will find it a godsend. More so since the newspaper could bring together and rally to a common cause people who were not from the Collège Saint-Joseph, such as Father Marcel-François Richard (1986: 75). This period of reconstruction of Acadia was marked by the emergence of an ideological discourse that revolved around a sense of national pride:

Journalists and orators sought to convince their compatriots that their age-old isolation was over and that their present state of inferiority was due, after all, only to the Deportation. Acadia had had a glorious past before the “Grand Dérangement”, they said, and the deportees were heroic; today’s Acadians are no less valiant than their fathers.

¹ All citations in this paper have been translated by the paper’s author.

Therefore, if they remain faithful to the traditions, the language, and the faith of their ancestors – with the help of Providence – the national survival is assured. Acadia has been and remains promised to a bright future. (Boudreau and Maillet, 1993: 716)

As for standard French practices, they have been increasingly present in the media (including national radio and television, some private radio stations, and the print media) and, to a lesser extent, in public signage. At present, Acadians rely on numerous reference works produced in Paris (such as the *Petit Robert* and the *Petit Larousse illustré*) and increasingly, in Montreal (for example, the online dictionary *Usito*) for their didactic needs. However, in professional settings, a correction software called *Antidote* is commonly used. This Montreal-based tool from *Druide informatique* includes a general dictionary of approximately 127,000 words (www.druide.com/fr). Finally, recommendations on terminology by the Office Québécois de la langue française are typically adopted in Acadia.

3. Chiac

In the southeastern part of New Brunswick, there exists a variety of French called Chiac that incorporates – and even transforms – English words, sounds and grammar into a French matrix itself characterised by the presence of archaic features. Chiac is usually associated with the city of Moncton, a more urban area where the Acadian community is well represented and where many of its important institutions are situated. According to Boudreau (2011), Chiac draws on the cruder elements of urbanity, such as disorder, mixture, and impurity, to emphasize its role as an element of counterculture associated with assimilation and even contamination with English.

Once heavily stigmatised and reserved for private use, Chiac has now infiltrated various public spheres and is more widely accepted than it was 50 years ago: it is showcased by artists internationally, whether they are filmmakers, poets, or novelists; most notably singers have exported their brand internationally and done so with ease. This is the case of Lisa LeBlanc, a well-known Acadian singer who calls her music “trash folk” and uses Chiac unapologetically. An example is LeBlanc’s stance on the set of the show *On n’est pas couchés*, known for both its familiar tone and the irreverence of its host, Laurent Ruquier. The program is watched by many in France and elsewhere, and the interaction between guest and host was viewed over 84.000 times on social media. Boudreau describes it here:

In October of the same year [2013], Lisa LeBlanc, an Acadian singer, made a name for herself in France and won the *France Inter/Télérama* award for best first francophone album. She was the guest of Laurent Ruquier in December. When asked where she started playing music, she answers “in a bar” (*bar* being pronounced in English). Someone says: *un bar?* pronounced in French, but instead of nodding, she stops talking, gives him a piercing look and resumes her own pronunciation – she repeats *bar* in English several times – and the people on the set start pronouncing *bar* like Lisa LeBlanc (Boudreau, 2019: 120).

In the larger public sphere, *Chiac* is also used more freely, as is the case of a local insurance company who uses it to advertise its services (“Worry pas”), or the coffee shop who offers a “trilingual” advertisement on their counter explaining payment options. In that sense, “French speakers have now reappropriated their vernacular in order to give it meaning (in terms of identity) and to reduce its stigmatization” (Boudreau and Perrot, 2010: 53). However, *Chiac* is still perceived as an inappropriate vernacular for formal contexts, an attitude that is reinforced both at home and in the national media, especially the Québec media. Arrighi and Urbain (2017) explain that the anglicisation of practices are of particular concern nationally:

Wake up Quebec" (Haché, 2014) is the leitmotif of the columnists. From one text and author to another, the argument is similar, as is the rhetorical construction of the comments, or the lexical fields used. Vocabularies borrowed from risk, threat and disease are used to remind us that anglicization is spreading in Quebec, just as it has already overwhelmed other French-speaking communities on the continent. It is then the mixed linguistic practices of the speakers of these communities that are mainly mentioned to support the point. (Arrighi and Urbain, 2017: 106)

Hence, Standard language ideology (Lodge, 1997) still dominates the linguistic landscape today, and *Chiac*, for its liberal use of English and bold attitude towards French, is a poor choice for an endogenous norm in Acadia. Too many still consider it a sign of the degeneration of the French language.

4. Acadian French

4.1 Acadian French glossaries

The language spoken today by those [the Acadians] of Nova Scotia and Cape Breton, of New Brunswick and the Magdalen Islands, is the same language

spoken by their elders with Razilly and D'Aulnay de Charnisay, when the country was founded, in the first half of the 17th century [...] It has been preserved intact and perpetuated without noticeable alterations, except in recent times because of the introduction of Anglicisms, which are, unfortunately, too numerous. The language of Acadia is, therefore, the language of the centre of France, at the beginning of the reign of Louis XIV (Poirier, 1928: 83). Poirier further states what Acadian French is not a half-language related to dialects and argues that it is even less so a *patois*. In the title of the book, reference is made to a 'language' (*Le parler franco-acadien et ses origines*), and this oral dimension of the variety under study is generally the only one considered:

This is an attempt to rehabilitate the Franco-Acadian language. [/]

Or if they speak French, their French, it will be with embarrassment, almost with blushing. [/] The idea of rehabilitating, within my means, the Acadian language, with the words of which it is composed, came to me at the Collège Saint-Joseph (New Brunswick), while I was taking classes there. (Poirier, 1928: 7)

The strategy adopted by Pascal Poirier in his attempts to legitimize Acadian French was also known and used in the rest of French Canada at that time: Scholars in Quebec, including Dunn (1880), Clapin (1894), Dionne (1909), and the Société du parler français au Canada (1930), would adopt the same approach, that is, to legitimize Canadian usages "because they originated in the French of the seventeenth and eighteenth centuries, or, as a new argument, in the languages of the provinces of France" (Poirier and Saint-Yves, 2002: 67). As in Quebec, "archaisms are particularly valued at the end of the century since they attest to the link to France, while English loanwords are seen as the manifestation of a certain assimilation to the Anglophone culture of the North American continent" (Boudreau and Urbain, 2013: 38). For Boudreau and Urbain, "the discourse thus creates tension between a desire for legitimacy, for recognition in the Francophonie and for national distinction. Language crystallizes these tensions" (2013: 40), which explains and justifies the obsession in finding words used originally, the quest for originality and the suppression of English loanwords.

In a few years, the Acadian language of Rivière-Bourgeois, which has been spoken for more than two hundred years, will have disappeared forever. The reason is that this parish is irrevocably anglicized. Few

people fifty years old and younger still speak the old French, the French of the first [settlement]. (1998: 21)

For the teaching of French, teachers cannot count on any lexicographical tool that would allow them to consider the current local variety in the learning of the language. It is in response to this need that we have prepared this dictionary which aims not only to describe the words of Acadian French, but also to relate them as much as possible to those used in France and Quebec. [/] The important thing for me [in the creation of this dictionary] has been to give a contemporary picture of the French lexicon of Acadia, one that is based on recent surveys, while insisting on the portion that characterizes it. (1999: 18)

Surprisingly, and despite this, the dictionary's nomenclature still relates to the various manifestations of rural life, whether they are terms relating to seasons, North American fauna, and flora, as well as fishing and agriculture, etc., and excludes words that reflect an industrialized society. Hence, demonstrating the juxtaposition of *Acadian French* as announced in the title, and the content, firmly rooted in traditional practices, leads us to believe that it is virtually impossible to shed the perception of Acadian French as anything other than a variety that is fixed in time, that has not evolved since the first settlements of the 17th century. To this day, there is still no book, glossary or dictionary that describes the current lexical practices of the Acadian community.

4.2 Scientific papers

Linguists are not exempt from reproducing commonplace linguistic ideologies, as we shall illustrate. Papers of two well-known Acadian French linguists, Karin Flikeid and Louise Péronnet, were examined, to establish which specific Acadian variety of French was studied, and which speaker was surveyed. These papers were selected because they have been instrumental within the academic institution and as such, are still cited today. Linguists not only describe traits of the languages they study, but they also put forth their perceptions of these languages and hence, “construct” them as well. The resulting picture highlights the fact that a “variety of language” is always dependent on the works that describe it and thereby, create it.

The names used by linguists to refer to the languages studied helps us understand their perspective. The following is a non-exhaustive list of the terms used in the papers: *Franco-Acadian*, *Acadian*, *Acadian French*, *Acadian*

speech(s), variety(ies) of Acadian French, N.B. Acadian speech, N.S. Acadian speech, etc. What do these terms refer to?

In the case of Louise Péronnet, she is essentially considering “traditional” Acadian French behind, which she places the practice “of the older generation which, because it has been little influenced by the English language or standard French, remains the faithful representative of traditional speech” (1989: 10).

Both researchers define Acadian French in genealogical terms. For instance, Péronnet begins one of her articles by saying that “Acadian French is the French spoken by the Acadians, i.e., the descendants of the French colonists who founded Acadia in 1604” (1995: 399). Flikeid defines Acadian French along similar lines: for her, Acadian French refers to historical Acadia, the Deportation, and the Acadian diaspora. The linguistic material that these two linguists study is therefore drawn from the pool of elderly, rural and poorly educated speakers.

In the illustration part of her article, Péronnet uses three documents from the 17th, 18th, and 19th centuries presenting “Acadian French through the centuries” to illustrate contemporary uses in their “regional diversity” (1995: 423). Flikeid, for her part, chose to “focus on the characteristics that still persist today” (1994: 322). It should be noted that the author’s objective is to put forward salient and ancient features and to measure the rate of conservation or abandonment according to the communities, based on “contemporary databases and studies” (1997: 256). These studies focus on vernacular languages. It is also understandable that Flikeid, seeking for her article the presence or absence of “traditional Acadian features” in the languages of different areas, has concentrated on works that prioritize rural and already ancient ways of speaking.

Considering these observations, we can now return to the initial question: how can the concept of “Acadian French” be used to describe the current practices of the Acadian community? How can we reconcile this perception of Acadian French within a contemporary view of the concept?

For the authors studied, this notion does not refer to a geographical conception of the term, as one might think. It is generally accepted, according to the naturalist approach, that the inhabitants of a territory are named according to that territory and that they speak the language named according to its inhabitants. In our corpus, the word *Acadian* does not seem to reflect the name of the territory, *Acadia*, but rather a variety of the older language,

traceable to its origins and passed down genealogically to the present.² The glossary specialists and the linguists featured here do not include in the object studied current urban or youth features, or even English words. As was first purported by Pascal Poirier, the stance adopted is the same: the object is posited as unchanged, as if frozen in time, taking us back to the first settlements. But do Acadians today not speak Acadian French?

The archaisms put forth by the nationalist movement of the 19th century was made central to the variety, as it showcased its ‘uniqueness’ and differentiated it from other French varieties; it also allowed for the revival of moribund words and the recovery of a ‘purer’ state of the language (Boudreau, 2011: 78-79). This ideology still permeates discourse today, putting forth the ‘originality’ of Acadian French – but what if, by doing so, Acadians feel their variety has no linguistic legitimacy relative to both other French varieties and Standard French? By emphasizing the marginality of French in Acadia, glossarians and linguists are placing it squarely outside the realm of any form of legitimate use, especially seeing as Standard language ideology has long convinced the Acadian community that its language is deficient. Similarly to Chiac, but for different reasons, Acadian French is not a suitable representative of Acadian practices and cannot be mobilised as an endogenous norm.

5. Conclusion

Though some archaisms are still in usage today, mostly by the elderly or in rural areas, they are vastly overrepresented in Acadian lexicographical works, and still regularly discussed in articles produced by experts on Acadian French. On the other hand, English loanwords, the defining characteristic of Chiac, are typically ignored in descriptive works but are put forward elsewhere as examples of the mediocrity of French in Acadia. Neither vernacular is apt to be used as a viable endogenous norm for the community.

In terms of language planning, the Acadian community of New Brunswick has exclusively focused on obtaining linguistic equality with the anglophone community, i.e., the ability to use French in most areas of public life. The Acadian community has worked diligently to create autonomous and

² For my part, I will keep using the term *Acadian French* to designate the variety of French spoken today in Acadia, in an attempt to restore the original meaning of the word *Acadian*, by both situating it in its larger geographical context and by contrasting it with other national French languages, including those of France, Belgium, Switzerland and Quebec. The term *standard Acadian French* thus refers to the more sustained linguistic practices of the Acadian community, the one that is currently situated in the Maritime Provinces of Canada, and more specifically, in the Province of New Brunswick.

homogeneous institutions, meaning unilingual French spaces within the predominantly English-speaking society that are open to English speakers but that function in French. Gains on this front have been made, but Acadia lags behind other French communities like Québec, for instance, where the regular production of descriptive works has been ongoing since the 1990s.

In the absence of a workable endogenous norm, describing the standard lexical practices of the Acadian community should be the first step as it would establish a benchmark against which other registers could then be measured. It would also be a positive step for Acadians' own perception of their linguistic practices, as it would at once affirm their ability to speak French, and have their specific traits recognized within this register. It would also highlight commonalities within the community, instead of emphasizing its differences. Finally, it would serve as a reminder that Acadian traits can be found in all types of discourse, including prestigious varieties.

References

Corpus

- Boudreau, Éphrem. (1988): *Glossaire du vieux parler acadien. Mots et expressions recueillis à Rivière-Bourgeois (Cap-Breton)*. Montréal. Éditions du Fleuve.
- Brasseur, Patrice and Jean-Paul Chauveau (1990): *Dictionnaire des régionalismes de Saint-Pierre et Miquelon*. (Canadiana Romanica, vol. 5). Tübingen. Max Niemeyer Verlag.
- Cormier, Yves (1999): *Dictionnaire du français acadien*. [Saint-Laurent]. Fides.
- Cyr, Sébastien (1996): *Le sel des mots: Glossaire madelinot*. Fatima. Le Lyseron.
- Flikeid, Karin (1994): *Origines et évolution du français acadien à la lumière de la diversité contemporaine*. In: Raymond Mougeon and Édouard Beniak (eds.), *Les origines du français québécois*. Québec. Presses de l'Université Laval. 275-326.
- Flikeid, Karin (1997): *Structural aspects and current sociolinguistic situation of Acadian French*. In: Albert Valdman (ed.), *French and Creole in Louisiana*. New York. Plenum Press. 255-286.
- Héon, U. (1943): *Quelques mots et expressions en usage aux Îles de la Madeleine*. Université de Moncton. Centre d'études acadiennes.
- Massignon, Geneviève [1962]: *Les parlers français d'Acadie. Enquête linguistique*. Paris. Librairie C. Klincksieck.
- Naud, Chantal (1999): *Dictionnaire des régionalismes du français parlé des îles de*

- la Madeleine. L'Étang-du-Nord. Les Éditions Vignaud.
- Péronnet, Louise (1989): Le parler acadien du sud-est du Nouveau-Brunswick: éléments grammaticaux et lexicaux. New York. Peter Lan.
- Péronnet, Louise (1995): Le français acadien. In: Pierre Gauthier and Thomas Lavoie (eds.). Français de France et français du Canada: les parlers de l'Ouest de la France, du Québec et de l'Acadie. Lyon. Centre d'études linguistiques Jacques Goudet. 399-439.
- Poirier, Pascal (1928): Le parler franco-acadien et ses origines. Québec. Imprimerie franciscaine missionnaire.
- Poirier, Pascal (1993): Le glossaire acadien. Éd. critique établie par P.M. Gérin. Moncton. Éditions d'Acadie et Centre d'études acadiennes.
- Thibodeau, Félix E. (1988): Le parler de la Baie Sainte-Marie (Nouvelle-Écosse). Le vocabulaire de Marc et Philippe. Yarmouth (N.-É.). Les Éditions Lescarbot.

Cited works

- Blanchard, Étienne (1914): Dictionnaire de bon langage. Paris. Librairie Vic et Amat.
- Boudreau, Annette (2011): La nomination du français en Acadie; parcours et enjeux. In: J. Morency, J. de Finney and H. Destrempe (eds.). L'Acadie des origines: mythes et figurations d'un parcours littéraire et historique. Sudbury. Prise de parole. 71-94.
- Boudreau, Annette (2016): À l'ombre de la langue légitime: l'Acadie dans la francophonie. Paris. Classique Garnier.
- Boudreau (2019): Cachez cet accent que je ne saurais entendre! In: Ed. Szlezák and Kl. S. Szlezák (eds.). Sprach-und Kulturphänomene in der Romania. Festschrift für Ingrid Neumann-Holzschuh zum 65. Geburtstag. Berlin. E. Schmidt Verlag. 111-125.
- Boudreau, Annette, and Marie-Ève Perrot (2010): Le chiac, c'est du français: représentations du mélange français/anglais en contexte inégalitaire. In: H. Boyer (ed.). Parloires hybrides. Paris. L'Harmattan, 51-82.
- Boudreau, Annette, and Émilie Urbain (2013): La presse comme tribune d'un discours d'autorité sur la langue: représentations et idéologies linguistiques dans la presse acadienne, de la fondation du Moniteur acadien aux Conventions nationales. In: Francophonies d'Amérique, 35. P. 23-46.
- Boudreau, Raoul, and Marguerite Maillet (1993): Littérature acadienne. In: J. Daigle (ed.). L'Acadie des Maritimes: études thématiques des débuts à nos jours. Moncton. Chaire d'études acadiennes. Université de Moncton. 707-748.

- Canac-Marquis, Steve, and Claude Poirier (2005): Origine commune des français d'Amérique du Nord: le témoignage du lexique. In: A. Valdman, J. Auger and D. Piston-Hatlen (eds.), *Le français en Amérique du Nord: état présent*. Sainte-Foy. Les Presses de l'Université Laval. 517-538.
- Caron, Napoléon (1880): Petit vocabulaire à l'usage des Canadiens-français, Trois-Rivières, *Journal des Trois-Rivières*.
- Chaudenson, Robert (1998): Variation, koïnésation, créolisation: français d'Amérique et créoles. In: P. Brasseur (ed.), *Français d'Amérique: variation, créolisation, normalisation, actes du colloque Les français d'Amérique du Nord en situation minoritaire (Université d'Avignon, 8-11 octobre 1996)*. Université d'Avignon. Centre d'études canadiennes. 163-180.
- Clapin, Sylva (1894): *Dictionnaire canadien-français ou Lexique-glossaire des mots, expressions et locutions ne se trouvant pas dans les dictionnaires courants et dont l'usage appartient surtout aux Canadiens-Français*. Montréal. C.O. Beauchemin, et Boston, chez l'auteure; réimpr. en 1974 aux Presses de l'Université Laval de Sainte-Foy.
- Dionne, Narcisse-Eutrope (1909): *Le parler populaire des Canadiens français ou Lexique des canadianismes, acadianismes, anglicismes, américanismes, mots anglais les plus en usage au sein des familles canadiennes et acadiennes françaises, comprenant environ 15,000 mots et expressions, avec de nombreux exemples pour mieux faire comprendre la portée de chaque mot ou expression*. Québec. Laflamme & Proulx; réimpr. en 1974 aux Presses de l'Université Laval de Sainte-Foy.
- Dubois, Lise (2005): *Le français en Acadie des Maritimes*. In: A. Valdman, J. Auger, D. Piston-Hatlen (dir.), *Le français en Amérique du Nord: état présent*. Sainte-Foy. Presses de l'Université Laval. 81-98.
- Francard, Michel (2005): Attitudes et représentations linguistiques en contexte minoritaire: le Québec et l'Acadie. In: A. Valdman, J. Auger et D. Piston-Hatlen (eds.), *Le français en Amérique du Nord. État présent*. Québec. Presses de l'Université Laval. 371-388.
- Gauvin, Karine (2021): *Le français en Acadie du Nouveau-Brunswick: un standard à définir*. In: *la Revue de l'Université de Moncton*. Vol. 50, n^{os} 1-2, p. 355-382.
- Gauvin, Karine (2014): *L'activité lexicographique en Acadie des Maritimes: bilan et perspectives*. In: A. Boudreau (ed.), *Minorités linguistiques et société/Linguistic Minorities and Society*. n^o 4, p. 42-81.
- Hambye, Philippe et Michel Francard (2008): Normes endogènes et processus identitaires. Le cas de la Wallonie romane. In: Cl. Bavoux, L.-F. Prudent and

- S. Wharton (eds.) Normes endogènes et plurilinguisme. Aires francophones, aires créoles. 61-84.
- Lodge, R. Anthony (1997): *Le français. Histoire d'un dialecte devenu langue*, trad. de l'anglais par Cyril Veken, Paris, Fayard.
- [Maguire, Thomas] (1841): *Manuel des difficultés les plus communes de la langue française, adapté au jeune âge, et suivi d'un Recueil de locutions vicieuses*, Québec, Fréchette.
- Manseau, J.-A. (1881): *Dictionnaire des locutions vicieuses du Canada avec leur correction suivi d'un dictionnaire canadien*, Québec, J.A. Langlais.
- Milroy, James, et Lesley Milroy (1985): *Authority in Language: Investigating Language Prescription and Standardization*, Londres, Routledge.
- Pépin-Filion, Dominique (avec la coll. de Josée Guignard Noël) (2018): *La situation linguistique au Nouveau-Brunswick: des tendances préoccupantes et quelques signes encourageants. Rapport présenté au commissaire aux langues officielles*.
- Poirier, Claude, et Gabrielle Saint-Yves (2002): *La lexicographie du français canadien de 1860 à 1930: les conséquences d'un mythe. Cahiers de lexicologie. Tome 80, p. 55-76*.
- Pöll, Bernhard (2005): *Le français, langue pluricentrique? Étude sur la variation diatopique d'une langue standard*. Peter Lang.
- Rinfret, Raoul (1896): *Dictionnaire de nos fautes contre la langue française*. Montréal. Librairie Beauchemin.
- Société du parler français au Canada (1930): *Glossaire du parler français au Canada*. Québec. L'Action sociale; réimpr. en 1968 par les Presses de l'Université Laval de Sainte-Foy.
- Statistique Canada (2016): *Statistiques sur les langues officielles au Canada, recensement de 2016*.
- Valdman, Albert (1979): *Créolisation, français populaire et le parler des isolats francophones d'Amérique du Nord*. In: Al. Valdman (ed.), *Le français hors de France*. Paris. Champion. 181-197.

Eric MIJTS

(University of Aruba, Aruba)
eric.mijts@ua.aw

Pluricentric languages in the Americas: the case of Dutch in the Dutch Caribbean

Abstract

This paper focuses on the characteristics of Dutch in the Caribbean. It introduces the concept that a non-dominant Caribbean variety of Dutch has developed with clear and distinct characteristics, but these characteristics are described mostly as deviations from the standard rather than regional characteristics of a rather homogenic variety of Dutch. Speakers of Dutch in the Caribbean generally accept Netherlandic Dutch as their norm. We hypothesize that that homogeneity originates at the convergence of the Caribbean diaspora in several urban centres in the Netherlands. Extensive research into the process of circular migration and language development is necessary to develop a full comprehension of the origin of the characteristics of Caribbean Dutch. Such research may lead to a more generic understanding of the development of postcolonial development of language varieties of the former colonizer's language.

This paper introduces the Caribbean part of the Kingdom of the Netherlands as a linguistic conundrum and continues, in the second section of the text, with the geographic, political, and linguistic context of the islands. A theoretical discussion on dominance and multilingualism in pluricentric language areas follows next, in the third section, and the fourth section is a reflection on current insights on the Caribbean Dutch variety in the Dutch Caribbean, to finally conclude that a non-dominant Caribbean variety of Dutch has developed, which is strongly influenced by the iterative migration of large groups of Dutch Caribbean citizens to the European Netherlands.

1. Introduction – The Caribbean part of the Kingdom of the Netherlands as a linguistic conundrum

De Ridder (2020) describes the gradual recognition of national varieties of Dutch by the Dutch language planning body, the *Taalunie*, formerly the *Nederlandse Taalunie* (*The Dutch Language Union*): Up until the 1990s and even at the

beginning of the 21st Century, little attention was paid to variation in the Dutch language and the norm for speakers of Dutch was the European Netherlandic Dutch norm (van der Sijs 2017): Gradually, recognition of the pluricentricity of the Dutch language started to grow when two national varieties of the Dutch language were identified, Netherlandic and Belgian Dutch, and Subsequently, Surinam and Caribbean Dutch (De Ridder 2020 p. 65-66): However, Caribbean Dutch is a complex concept due to the geographic nature of the region, with the Caribbean Dutch-speaking population spread out over a total of 6 islands. Each of these islands has a different linguistic and demographic composition, and they are spread out over the north and south of the Caribbean, up to almost 1000 kilometres apart. Furthermore, to gain a better understanding of the characteristics of Caribbean Dutch, the circular migratory patterns need to be understood. These patterns were a result of the relations within the Kingdom of the Netherlands, causing many of the inhabitants of the Dutch Caribbean to migrate to the Netherlands –mainly for study purposes– and to come back at a certain point in time. It is not only in the contexts of the Caribbean islands that we can find the origins of the Caribbean Dutch variety but also, if not mainly, in the circular migration movements of the diaspora, that cause the speakers of the Caribbean variety of Dutch to converge when they come into contact with Dutch society –but also with the Surinamese diaspora– and, consequently, Caribbean Dutch develops through such twofold language contact.

The existence of Caribbean Dutch, or Antillean Dutch as it used to be called, is exemplified not only by the abovementioned reference to the existence of a Caribbean Dutch variety by the Taalunie, but also by the fact that several researchers have made an inventory of Caribbean Dutch lexical items (Joubert 2005): Likewise, the etymology of this Caribbean Dutch lexis has been studied (Joubert & van der Sijs 2020) as well as specific characteristics of language use by Dutch Caribbean citizens (Vervoorn 1976, Kester & Fun 2012, Fun 2014, Kester, Buijinck, & Hortencia 2017): Nonetheless, the existence of a separate Caribbean variety has thus far not sufficiently been studied and it remains under discussion whether there actually is a Caribbean Dutch regional variety. One of the reasons for this, is that Dutch is not a native but a foreign language for the majority of the populations of the island territories (see table 1) (Mijts & van Oostendorp 2017, Mijts 2017, Mijts 2008).

The adoption of about 500 Caribbean Dutch words in the official word list of the Dutch language (Van Dale/Taalunie 2015) in 2015 constituted the recog-

nitition of the existence of a specifically Caribbean Dutch vocabulary. In a first glossary Sydney Joubert and Nicoline van der Sijs (2020) provide the origins of 447 Antillean Dutch words, many of which have been included in the official word list of the Dutch language.

2. Context

The former Dutch Caribbean consist of the six islands that used to constitute the Dutch colony Curaçao and Dependencies (Curaçao en onderhorigheden), which became the country of the Dutch Antilles (Nederlandse Antillen) in 1954. The Netherlands Antilles disintegrated in two phases: in 1986, Aruba segregated from the Netherlands Antilles and in 2010, the Netherlands Antilles dissolved altogether. Currently, the six islands of the Dutch Caribbean consist of three countries, Aruba, Curaçao, and Sint Maarten, that, together with the Netherlands, constitute the Kingdom of the Netherlands, and the three remaining islands have become the so-called ‘special municipalities’ of the Netherlands: Bonaire, Saba, and Sint Eustatius. All six countries are multilingual. On the three islands in the North (Saba, Sint Eustatius, and Sint Maarten), Caribbean English is the main language of communication, while on the islands in the South, Papiamentu is the main language of communication. On all six islands, Dutch has official status, however, due to more recent immigration waves, Spanish has also started to play a substantial role as a means of communication.

As the different island territories fall under different jurisdictions, language policies, education policies, and governmental policies vary making a uniform description of the six islands very difficult. However, due to the fact that the Charter for the Kingdom of the Netherlands prescribes concordance and because of the shared constitutional past, there is legal and governmental cohesion among the territories. As stated before, this cohesion, however, is not reflected in the linguistic composition of the societies, that differs significantly due to their different characteristics, including geographic location and size, as well as the differences in economic development and the ensuing demographic development. Due to the multilingual nature of these island communities, not only a variety of Dutch is present in them, but also varieties of Papiamentu and Papiamentu, Spanish, and English. On the one hand, the 1986 secession from the Netherlands Antilles led to the adoption of an alternative orthography, leading to the codification of varieties of the language with the name Papiamentu in Aruba and Papiamentu in Curaçao. On the other hand, when it comes

to English, the Caribbean varieties of English that are spoken in Aruba (Richardson & Devonish 2016), Saba (Myrick 2014), Sint Eustatius (Aceto 2006), and Sint Maarten (idem) are also recognized as creoles in their own right.



Figure 1: the position of the ABC-islands in the South of the Caribbean and the SSS-islands in the North of the Caribbean (dialectloket.be).

The creoles are under continuous pressure from the European English norm that is introduced in schools –where Dutch foreign language speakers of English often teach British English to English-based Creole speaking children-- and from the influence of the use of American English, because of the dominant American tourism industry and media. In none of these societies does Dutch play a significant role in people’s daily or cultural life. An explanation of the sociological, educational, and cultural implications of the complexity of these multilingual constructs would lead too far for the purpose of this paper but has been extensively described by Farclas, Kester, and Mijts (2019), Pereira (2018), and Mijts (2021):

When it comes to the use of Dutch, put simply, speaking Dutch with Caribbean characteristics is frowned upon in many settings and stigmatises speakers as uneducated.

	Aruba	Bonaire	Curaçao	Saba	St Eustatius	St Maarten
Number of inhabitants	111.083	21.745	153.671	1.918	3.142	37.224
Most spoken	Papiamentu 72%	Papiamentu 63.8%	Papiamentu 80%	English 93%	English 85%	English 67.5%
2nd language	Spanish 14%	Dutch 15.4%	Dutch 11%	Dutch 4%	Spanish 7%	Spanish 12/9%
3rd language	English 8%	Spanish 15.2%	Spanish 6%	Spanish 3%	Dutch 6%	Creole 8.2%
4th language	Dutch 6%	English 4.5%	English 3%	Not reported	Other 2%	Dutch 4.2%
Percentage multilingual citizens	>90% (CBS Aruba, 2000)	90	Estimated at >90%	57	68	Unknown
Official languages	Dutch Papiamentu	Dutch Papiamentu English	Dutch Papiamentu	Dutch Papiamentu English	Dutch Papiamentu English	Dutch English

Table 1: The reported first home languages in the Dutch Caribbean according to territory, according to ranking in combination with demographical data and reported official languages (CBS Aruba 2000, 2020, CBS Curaçao 2022, CBS Nederland 2014, 2021, World Factbook 2022)

As the numbers presented in table 1 reveal, Dutch is not frequently mentioned by the respondents as the most spoken home language. Moreover, despite the fact that the majorities of the populations report being multilingual, Dutch does not play a role in their private lives. Dutch is the language learned by most in the schools, in which the Netherlandic Dutch norm for pronunciation and grammar is imposed on children. For many children, the first structural encounters with the Dutch language are in the classroom. A very high percentage of the populations of these territories are 'quite fluent' multilinguals. Note that, despite the fact that multiple languages are listed as official languages for every territory, the status of these languages is not equal, as Dutch is structurally prescribed as the sole language to be used in specific domains such as education, administration, and legislation.

3. Theoretical discussion on dominance and multilinguality in pluricentric language areas

The presence of Dutch as a de facto minority language, albeit with official status, and the language contact situation of Dutch with several other languages in the island territories of the former Dutch Caribbean have led to the

development of specific characteristics that are typically recognized as being Caribbean Dutch characteristics. The dominance of the Dutch language in certain domains in these territories should not lead to the conclusion that Dutch is the dominant language in these societies. On the basis of Clyne (1992), Muhr & Delcourt (2001), and Muhr (2003, 2005), Muhr has developed a comprehensive overview of characteristics of dominance of pluricentric languages (Muhr 2012): These characteristics will be applied to the Caribbean variety of Dutch in the next section. However, first, the concept of dominance needs to be considered within the framework of multilingual decolonialised societies –and the consequences for access to education, governance, or the judiciary– like those of the former Caribbean islands within a broader categorization of multilingual states (Mijts 2021).

Aruba's history of colonization and economic development has resulted in a rich and heterogeneous ethnic and linguistic melting pot, in which four languages play a dominant role in different domains of language use. A majority of the Aruban born population speak Papiamentu, Dutch, English, and Spanish. All four languages can be considered to be dominant languages. First, Papiamentu is an official language and the home language of 68 percent of the population. Second, Dutch –home language of six percent of the population– is the other official language and the dominant language of the Kingdom of the Netherlands, to which Aruba belongs. Third, English –spoken as a home language by eight percent– is the dominant language of tourism, the predominant industry in Aruba. Fourth, Spanish –the reported home language of 14 percent of the population– is also the language of nearby Venezuela and Colombia, and as a result, it is the regionally dominant language (Mijts and Waterman 2016).

It is sufficiently clear that the Dutch language and the assumed Caribbean variety of Dutch operate in a complex relationship with the other languages in the multilingual realities of the former Dutch Caribbean. In order to understand the specific multilingualism in decolonialised societies, we need to contrast it with other, more traditionally recognized and studied forms of multilingualism. For that purpose, we may distinguish four major categorizations that reflect different contextual variables determining the relations between the different languages spoken in these territories as well as the access to education in the home language of people.

First of all, in multilingual countries with several official languages, like Canada, Switzerland, Belgium, and South Africa, language provisions have not always been made to facilitate education, and access to government and the

judiciary in all languages of that nation, in all regions. This leaves population groups either the option to 'abandon' their home language or to move to regions where their language is a language of instruction for their children. Second, migration, voluntary or forced, is one of the main reasons why children end up in situations where they go through an education system, far from home, in which the language of instruction is not their home language. However, in many of these cases, the language of education will be the dominant language in that society and also the majority language as well as the language of governance and the judiciary. This is a second group that does not only include labour migrants, but also political and economic refugees, often with a complex and temporary status. A third group would be the speakers of so-called minority or regional languages, as described in the European Charter for Regional or Minority Languages (ECRML): The ECRML describes these languages as "languages that are: i. traditionally used within a given territory of a State by nationals of that State who form a group numerically smaller than the rest of the State's population; and ii. different from the official language(s) of that State; it does not include either dialects of the official language(s) of the State or the languages of migrants" (ECRML, 1992): Facilities for speakers of these languages are provided, however, these provisions are constantly under dispute. Finally, the last and often forgotten group are the speakers of majority languages in countries and states in which the home language of the majority of the population is not the language of instruction in schools, the judiciary, or governance, despite the fact that the languages that are used for these official functions are foreign languages to the vast majority of the people of that region.

In the global North, challenges that arise in multilingual countries or from migration and the existence of minority and regional languages are well known and well-studied. Both the first, second, and third group in the categorization above have been the centre of attention of a multitude of academic and policy studies, both in Europe and in the US. Time and again, attempts have been made to capture the complexities of the position and challenges of the fourth group under the umbrella of minority languages and regional languages. It is within the framework of the fourth category of multilingual settings that we need to understand Dutch in the Dutch Caribbean, as the foreign language that is used as language of instruction, governance, and the judiciary, while it is a foreign language for the vast majority of the populations of the Dutch Caribbean. Despite the fact that the majority of the families in the Dutch

Caribbean are multilingual, Dutch is usually not used at home, nor in daily life, but only in formal settings. It is in the diaspora of Dutch Caribbean citizens in the Netherlands –where these migrants belong to the second group in this categorization– that Dutch becomes a language that is present in daily life and much more frequently used by these Dutch Caribbean citizens, not only in interaction with Dutch people, but also in interaction with speakers of Surinamese Dutch and often amongst each other as well.

4. Current insights on the Caribbean Dutch variety in the Dutch Caribbean

The previous paragraph contained an overview of the multilingual development of the former Dutch Caribbean. Thus far, many, including myself, have argued that there is no such phenomenon as a Caribbean variety of Dutch in the Caribbean, as Dutch is clearly a foreign language for the vast majority of the populations of these islands. As contact with the Dutch language in the Dutch Caribbean is limited to a specific and limited time and space in people's lives: the classroom and, if unfortunate enough in life, in the courtroom. In that view, the typically Caribbean characteristics of Dutch would be constructed as consequences of the learning and use of Dutch as a foreign language –influenced by the home language of those learners of Dutch– and, as such, as "errors" in language learning that ought to be "remedied". This approach, however, actually ignores a couple of distinct characteristics that may be more indicative of the existence of a Caribbean variety of Dutch, including the existence of vocabulary that has been identified as Caribbean and has been adopted by the *Woordenlijst der Nederlandse Taal*, and an understanding of a most uncommon and insufficiently studied process of development of a language variety through the postcolonial diaspora. That analysis may also be a starting point for the study of a very interesting form of genesis of pluricentric languages in which the formerly colonized come together in the former colonizers' country and a new variety takes shape in that meeting of the diaspora.

A little more background should be provided on this phenomenon. At critical moments in Caribbean –including Surinamese– development, considerable migration has taken place from the formerly colonized territories to the Netherlands. Examples of these migration waves are the Surinamese migration of 1975, the Aruban migration of 1984-1986, and the Curaçao migration of 1968. These migration waves led to the development of a considerable migrant community in cities like Dordrecht, Amsterdam, The Hague, and Rotterdam,

and also to the recognition of Papiamentu as a minority language under the ECRML in 2021 (De Groot 2019, Van der Velden 2021): In these communities, a specific variety of Dutch has developed, and is also adopted in the Dutch language and projected in popular culture. Apart from these migration waves, another process takes place, in which Dutch Caribbean youngsters migrate to the Netherlands for higher education. Every year hundreds of young people from the Caribbean leave the islands and start their higher education in the Netherlands. This is the brain drain from the Caribbean islands as the most talented ones will be scouted and hired in Europe and many will only come back after a prolonged stay in Europe, where they have had their first jobs and have paid off (part of) their study debts. This process also bereaves the islands of a young, creative, and activist generation that could help shape politics, innovation, and the creative industries in the islands. An interesting aspect of this temporary migration is that in their transition from the Caribbean to the Netherlands, the status and use of their language also changes: while Dutch used to be a foreign language only used in a very limited number of domains for them in the Caribbean, it becomes a second language, widely present and used in the society in which they will spend a substantial formative time of their lives. These Caribbean youngsters go through a multi-year immersion in Dutch language and culture, in continuous contact with their peers –often speaking Papiamentu amongst each other, yet often finding themselves in a position where they have to speak Dutch to accommodate others.

When we reflect on the dominance of the Caribbean variety of Dutch according to the parameters as summarized in Muhr (2012), we have to conclude that, within the Kingdom of the Netherlands, the Caribbean variety of Dutch is a non-dominant variety. Caribbean Dutch bears all the general features of a non-dominant variety, including its number of speakers, the fact that the Dutch variety is the norm in the educational settings in the former Dutch Caribbean (hence also the persisting presence of Dutch teachers), the colonial nature of the origins of Dutch in the Caribbean, and having low status. When it comes to attitudes and beliefs, uncertainty about the proper standard norm, exclusion of the variety in schools, and self-devaluation of non-dominant forms are clearly present in the discourse of the Dutch Caribbean societies as well as in the Netherlands. Also, when it comes to language loyalty –and identification by the elites with the European Netherlands Dutch norm– and the absence of language and status planning, it is clear that the Caribbean Dutch variety is non-dominant. In short, in spite of the growing awareness of the potential ex-

istence of a Caribbean variety of Dutch, currently the status of the Caribbean Dutch variety –and even the existence of a generally identifiable Caribbean Dutch variety– is questionable.

As stated earlier, the genesis of a Caribbean variety of Dutch finds its roots partly in the Caribbean and in the Netherlands. The context-related vocabulary in Caribbean Dutch finds its roots in the need of terminology that describes specific landscapes and architectural, administrative, culinary, and biological features. Examples of these words are *knoek*, for the description of the barren fields of the southern Caribbean, that derived from *knoekenhuis*, the typical pastoral dwellings that were built in these territories, and *cessantia*, a typical example of administrative terminology for a contract-termination construct that does not exist in the Netherlands or Belgium. In the culinary domain, examples like *pastechi*, a typical Southern Caribbean pie, *snacktruck*, for the roadside food trucks, and *stobá*, for stews of goat meat or other ingredients, demonstrate that the culinary culture is the origin of linguistic variation in the Dutch Caribbean. Finally, an abundance of bird and plant species have found their way into the vocabulary, such as *suikerdieffe* (bananaquit), *lionfish* –and of course the *lionfishinvasie* (Lionfish invasion), *kenepa* –also spelled as *knippa* (a typical regional fruit), and *zuurzak* (soursop):

As can be observed from the overview of the Caribbean Dutch words above, some of these words stem from Papiamentu, only some from English, and none from English creoles as spoken throughout the Dutch Caribbean territories. One would expect words like *journey cake* or *Johnny cake*, and many others that are derived from English, also to be recognized as part of the characteristics of Caribbean Dutch. This may have to do with the status of English creoles that are given low appreciation by the populations that speak the language themselves, even referring to it as *bad English* or *bush English*. There are also distinct phonological markers such as the overexaggerated rendering of /s/-sounds and diphthongs, pronunciation of the /v/ as /w/, and a whole series of morphemic and syntactic characteristics that are usually considered deviations from Standard Dutch rather than characteristics of a Caribbean variety of Dutch.

5. Conclusion: commonalities and differences

Despite indications of recognition of a Caribbean variety of Dutch by the *Taalunie*, according to their own recent publication on language variation in Dutch, there are "three, partially different norms for standard Dutch for the

Netherlands, Flanders, and Surinam" (Taalunie 2019, p.3): As such, the characteristics of the Antillean varieties of Dutch appear not to be recognized –yet– as norms, however, there is clear evidence of uncodified normativity. Insufficient research has been done to provide grounds for a full overview of characteristics, however, a number of characteristics can be provided as a starting point for the claim of the existence of a Caribbean Dutch variety (Mijts 2008, Mijts & Waterman 2015) that, on the one hand, originates in the Caribbean itself and, on the other hand, further develops in the continuous cycle of migration in both directions across the Atlantic, where communities from the different islands come together in student communities and in specific neighbourhoods in the Netherlands, especially the Hague and Rotterdam, where the Caribbean varieties of Dutch come together and actually become more homogenic.

We can conclude that a non-dominant Caribbean variety of Dutch has developed with clear and distinct characteristics, but that these characteristics are mostly described as deviations from the standard rather than as regional characteristics of a rather homogenic variety of Dutch. Speakers of Dutch in the Caribbean generally accept Netherlandic Dutch as their norm. We can hypothesize that that homogeneity originates in the convergence of the Caribbean diaspora in several urban centres in the Netherlands. This conclusion underscores the need for extensive research into the process of circular migration and language development. This can lead to a more generic understanding of the postcolonial development of language varieties of the former colonizer's language.

References

- Aceto, M. (2006): Stavian Creole English: An English-derived language emerges in the Dutch Antilles. *World Englishes*(25 (3-4)), 411-435.
- CBS Aruba. (2000): Census 2000. CBS Aruba.
- CBS Aruba. (2020): Quarterly Demographic Bulletin 2020 - 4th quarter. CBS Aruba.
- CBS Curaçao. (13. 01 2022): CBS Curaçao. Downloaded from <https://www.cbs.cw/population>
- CBS Curaçao. (13. 01 2022): Censo 2011. Downloaded from CBS Curaçao: <https://www.cbs.cw/demographics-characteristics>
- CBS Nederland. (2014): Papiaments de meest gesproken taal op Bonaire en Engels op Saba en Sint Eustatius. Von CBS Nederland: <https://www.cbs.nl/nl-nl/nieuws/2014/51/papiaments-de-meest-gesproken-taal-op-bonaire-en-engels-op-saba-en-sint-eustatius-abgerufen>

- CBS Nederland. (2021): Caribisch Nederland; bevolking, geslacht, leeftijd en geboorteland. CBS Nederland.
- Clyne, M. (1992): *Pluricentric languages, Differing norms in different countries*. Berlin: Mouton-de Gruyter.
- De Groot, G. R. (2019): European Charter for Regional and Minority Languages. In T. Schoonheim, & J. Van Hoorde, *Language variation. A factor of increasing complexity and a challenge for language policy within Europe (Contributions to the EFNIL Conference 2018 in Amsterdam)* (S. 115-128): European Federation of National Institutions for Language.
- De Ridder, R. (2020): Dutch national varieties in contact and in conflict. In R. Muhr, *European Pluricentric Languages in Contact and Conflict* (p. 65-80): Wien: Peter Lang Verlag.
- Dialectloket.be (n.d.): Nederlands in het Caribisch gebied. Ghent University. <https://www.dialectloket.be/tekst/nederlands-in-de-wereld/nederlands-als-officiële-taal/nederlands-in-het-caribisch-gebied/>
- ECRML. (1992): *European Charter for Regional or Minority Languages*. EU.
- Faraclas, N., Kester, E.-P., & Mijts, E. (2019): *Community based research in language policy and planning*. Cham: Springer.
- Fun, J. (2014): *Es style di alles mag: Een onderzoek naar het taalgebruik van Arubanen* (diss.): Utrecht: Universiteit Utrecht.
- Joubert, S. (2005): Curaçaos-Nederlands. In N. v. Sijs, *Wereldnederlands: oude en jonge variëteiten van het Nederlands* (S. 31-57).
- Joubert, S., & van der Sijs, N. (November 2020): Antilliaans-Nederlandse woorden en hun herkomst. Trefwoord.
- Kester, E.-P., & Fun, J. (2012): Language use, language attitudes and identity among Aruban students in the Netherlands. In N. Faraclas, R. Severing, C. Weijer, & E. Echteld, *Multiplex cultures and citizenships: multiple perspectives on language, literature, education, and society in the ABC-islands and beyond*. (S. 231-248): Curaçao: Fundashon pa plnifikashon di idioma.
- Kester, E.-P., Buijink, J., & Hortencia, T. (2017): Language use, language attitudes and identity among Aruban and Curacaoan students in the Netherlands. In N. Faraclas, R. Severing, C. Weijer, E. Echteld, W. Rutgers, & R. Dupey, *Archaeologies of Erasures and Silences - Recovering othered languages, literatures and cultures in the Dutch Caribbean and beyond* (S. 13-19): Curaçao/Puerto Rico: University of Curacao & University of Puerto Rico.
- Mijts, E. (2008): Nederlands is een Vreemde Taal . *Ons Erfdeel*, S. 94-101.
- Mijts, E. (2008): Nederlands is een vreemde taal. *Ons Erfdeel*(1), 94-101.

- Mijts, E. (2017): Waar de uitzondering niet de regel zou moeten zijn: Nederlands in het Caribisch deel van het Koninkrijk der Nederlanden. *Vaktaal*, S. 6-7.
- Mijts, E. (2021): *The situated construction of language ideologies in Aruba: a study among participants in the language planning and policy process (diss.)*: Antwerp & Ghent: University of Antwerp and Ghent University.
- Mijts, E., & van Oostendorp, M. (13. 01 2022): *Nederlands op Aruba*. Downloaded from [neerlandistiek.nl: https://www.youtube.com/watch?v=iI9OLicfYjs](https://www.youtube.com/watch?v=iI9OLicfYjs)
- Mijts, E., & Waterman, K. (2015): Wat is de positie van het Nederlands in de multilinguale samenleving. In L. Echteld, E. Mijts, W. Rutgers, & R. Severing, *Taalbeleid in het aribisch gebied: heden en in de nabije toekomst*. (S. 125-131): Curaçao: University of Curaçao.
- Muhr, R. (2003): Die plurizentrischen Sprachen Europas - ein Überblick. In E. Gugenberger, & M. Blumberg, *Vielsprachiges Europa. Zur Situation der regionalen Sprachen von der iberischen Halbinsel bis zum Kaukasus*. (S. 191-231): Frankfurt & Wien: Peter Lang Verlag.
- Muhr, R. (2005): Language Attitudes and language conceptions non-dominating varieties of pluricentric languages. In R. Muhr, *Standardvariationen und Sprachideologien in verschiedenen Sprachkulturen der Welt. / Standard Variations and Language Ideologies in different Language Cultures around the World*. (S. 11-20): Wien: Peter Lang Verlag.
- Muhr, R. (2012): *Non-dominant varieties of pluricentric languages, getting the picture*. Frankfurt am Main: Peter Lang.
- Muhr, R., & Delcourt, C. (2001): *Les Langues Pluricentriques. Variétés nationales des langues européennes à l'intérieur et à l'extérieur de l'espace européen*. (Bd. Numero thematique 79/2001 de Revue Belge de Philologie et Histoire:Langues et Litteratures Modernes).
- Myrick, C. (2014): Putting Saban English on the map: A descriptive analysis of English language variation on Saba. *English world-Wide*, 161-192.
- Pereira, J. (2018): *Valorization of Papiamentu in Aruban society and education, in historical, contemporary and future perspectives (diss.)*: Willemstad - Curaçao: University of Curaçao.
- Richardson, G., & Devonish, H. (2016): *The English creole of Aruba: a community-based description of the San Nicolas variety*. San Nicolas - Aruba: Instituto Pedagogico Arubano. Von Instituto Pedagogico Arubano. abgerufen
- Taalunie. (2019): *Visie op taalvariatie en taalvariatiebeleid*. Taalunie.
- Van Dale/Taalunie. (2015): *Het Groene Boekje: Woordenlijst der Nederlandse Taal*. Van Dale.

- van der Sijs, N. (13. 05 2017): Op naar een Surinaamse en Antilliaanse editie van de Atlas van de Nederlandse taal! Downloaded from Neerlandistiek.nl: <https://neerlandistiek.nl/2017/05/op-naar-een-surinaamse-en-antilliaanse-editie-van-de-atlas-van-de-nederlandse-taal/> abgerufen
- Van der Velden, B. (2021): Nos ta papia Papiamentu: Over de erkenning van het Papiaments in Nederland. Nederlands Juristenblad, 1736-1740.
- Vervoorn, A. (1976): Antilliaans Nederlands. Den Haag: Kabinet voor Nederlands-Antilliaanse zaken.
- World Factbook. (13. 01 2022): Sint Maarten. Downloaded from CIA World Factbook: <https://www.cia.gov/the-world-factbook/countries/sint-maarten/#people-and-society>

This is the first of two volumes that present part of the outcome 9th World Conference on Pluricentric Languages that took place from August 26-28 2021. The volume contains fourteen contributions. The conference was organised by the “Working Group on Non-Dominant Varieties of Pluricentric Languages” (WGNDV), this time hosted by the Austrian German Research Centre in Graz. The conference had three main topics: (1) Pluricentric languages in the Americas; (2) The localisation of global audiovisual and print media in pluricentric language areas. (The papers of this section will be published in a separate volume); (3) Pluricentric languages worldwide. The papers of this volume refer to topic (1) while the papers of theme (3) are published in volume (2).

All articles of this volume deal with pluricentric languages that exist in the Americas. It is the first volume of this kind that has been ever published on this topic. Section (1) contains two papers that concern the indigenous language Quechua and the Guarani-Tupi language family. Section (2) presents five papers about the pluricentricity of Spanish in the Americas. There are five papers in section (3) that deal with Portuguese in the Americas and worldwide. And finally in the last section there is a paper about French in Acadia (Canada) and another one about Dutch in Aruba which is part of the Dutch Caribbean territories.

The editors:

Rudolf Muhr is Professor em. of Graz University and head of the Austrian German Research Centre in Graz. He is also Coordinator of the “International Working Group on Non-dominant Varieties of Pluricentric Languages”. His research activities are focused on researching pluricentric languages and the description of Austrian German.

Maria Eugênia Lammoglia Duarte is full professor at the Federal University of Rio de Janeiro (UFRJ, Brazil) and holds a research grant from the National Council of Research (CNPq). She studies syntactic change in varieties of Portuguese and other Romance Languages.

Cilene Rodrigues is an Associate Professor at Pontifical Catholic University of Rio de Janeiro. Her research centers around comparative syntax, particularly on languages of South America. In addition, it contemplates grammar & neuro diversity, with focus on grammar in psychotic disorders.

Juan Thomas is Associate Professor of Spanish at Utica College in Utica New York, US. The focus of his research is on Spanish in the US, Languages in contact; Attitudes in Second Language Acquisition; phonology; philology (Hispanic, Italian, Galician).

