Universidade de Lisboa

Faculdade de Letras

Vós, *Vocês*, and the Null Subject in European Portuguese

Blake Warren Selph

Dissertação orientada pela Professora Doutora Ana Maria Martins

especialmente elaborada para a obtenção do grau de Mestre em Linguística

2021

Dedicated, *In memoriam,*

to my loving grandparents Wanda and Justus Selph,

who always emphasized and prioritized getting an education.

Rest in peace and live on in my heart.

Thanksgiving

To my dear Grandma Wanda and Grandpa Justus, who always loved their family no matter what and saved money since the day I was born to make sure that I got a good education, I send my heartfelt thanks. I only wish you could be here to see me now.

To all of my family, and especially to my parents Jeanne and Barney, who have always set me up to succeed, and to my brothers James and Steven, and sister Jessica, who are always there for me, I send a big hug and my deepest thanks.

To my partner Camila, whose emotional support made me believe in myself in moments of stress when I would've much rather quit, I send my loving thanks.

To my thesis director Ana Maria Martins, whose knowledge and guidance have been essential to the formation of this work, and whose patience with my sometimes chaotic writing process has been much appreciated, I send my humblest thanks.

To professor João Saramago, who readily made the CLUL's recordings available to me and let me work in his office, I send my thanks.

To my friend Daniel, who has always been a great and understanding friend to me through thick and thin, I send my thanks.

And to you, the reader, I send my thanks for your interest in my work, which is a little part of me.

Resumo

O foco desta dissertação é estudar detalhadamente os pronomes da 2ª pessoa do plural (2PL) *vós*, *vos* e *vosso(s)/vossa(s)*, e os morfemas verbais da 2PL para compará-los com a inovação dominante e crescente que consiste na substituição pela 3ª pessoa do plural (3PL), o que se traduz na expansão de *vocês*, morfemas verbais da 3PL e os pronomes correspondentes *os/as*, *lhes*, *e seu(s)/sua(s)*, no português europeu (PE). Além disso, analisam-se as ocorrências de sujeitos nulos com formas verbais morfologicamente 2PL e 3PL, mas semanticamente 2PL, com o objetivo de comparar a frequência de uso de sujeitos nulos em dialetos que retêm as seis pessoas gramaticais na morfologia verbal e aqueles que retêm apenas cinco, tendo removido a 2PL.

Começando por consultar Cunha & Cintra (1984) como exemplo de gramática tradicional, vê-se que a 2PL é ainda prescrita, porém já nesta gramática se aponta que é pouco usada hoje em dia. Além disso, em Cunha & Cintra (1984) a "mistura de tratamentos", quer dizer, uma combinação de pronomes acusativos/dativos ou possessivos de 2ª pessoa (*vos*, *vosso(s)*) com o pronome nominativo *vocês* não é recomendada. Contudo, os dados disponíveis em Segura (2013), Lara Bermejo (2015) e Aguiar & Paiva (2017), mostram que *vocês* é usado com verbos na 2PL em partes do Norte de Portugal, e, por outro lado, na maioria do Centro-Sul, onde *vocês* substituiu definitivamente *vós*, persistem o clítico *vos* e o possessivo *vosso(s)/vossa(s)*. É com Lara Bermejo (2015) que se apresenta um modelo concreto para explicar estas combinações, mostrando a progressiva inserção da inovação *vocês* e correspondentes flexões verbais e pronomes acusativos, dativos e possessivos no PE. O modelo de progressão da mudança de Lara Bermejo (2015) é: sujeito/reflexivo/verbo > acusativo > dativo > possessivo. Este autor também mostra a presença da inovação *vocês* nos distritos de Aveiro e Viseu, no Norte de Portugal. O presente estudo propõe uma fase a mais do que Lara Bermejo (2015), autonomizando o sujeito e assim ampliando o modelo, que passa a incluir um passo adicional no desenvolvimento faseado da inovação: sujeito > reflexivo/verbo > acusativo > dativo > possessivo. Esta proposta tem originalmente por base alguns dados de Segura (2013) e Aguiar & Paiva (2017) e é depois confirmada por dados do Corpus Dialetal

para o Estudo da Sintaxe/Syntax-oriented Corpus of Portuguese Dialects, que constitui a base empírica central da presente tese.

Uma outra hipótese deste trabalho é que um paradigma verbal com cinco pessoas não resultará em taxas mais baixas de ocorrência de sujeitos nulos, nem afetará a característica do PE de ser uma língua de sujeito nulo (LSN). Enquadra-se esta questão com referência à história do parâmetro do sujeito nulo (PSN), considerando a teoria original de Chomsky (1981) e Rizzi (1982) e os subsequentes refinamentos teóricos propostos por Huang (1984, 1989) e Jaeggli & Saffir (1989). Por fim, este estudo concorda com a ideia de Roberts (1993) de que uma LSN que utiliza a morfologia verbal para identificar o sujeito requer "riqueza funcional" para se manter [+ LSN]. Esta riqueza funcional significa que o paradigma verbal não pode ter mais que um *sincretismo*, um termo usado para descrever a fusão (ingl. *merger*) de duas pessoas gramaticais morfologicamente distintas em uma só. Aplicando este modelo ao PE, este estudo mantém a hipótese de que a perda dos verbos da 2PL não afetará o traço [+ LSN] do PE.

Estudam-se detalhadamente os dados provenientes do Corpus Dialetal para o Estudo da Sintaxe/Syntax-oriented Corpus of Portuguese Dialects (CORDIAL-SIN) com o objetivo de afinar o entendimento da propagação da inovação *vocês*. A análise dos dado permite mostrar que a inovação já se tinha implantado, a Norte, nos distritos de Braga, Porto, Aveiro e Viseu até a década 90, o último período de coleta de dados para o ALEPG (em curso), depois usados para a constituição do CORDIAL-SIN, e que também já estava presente nos dialetos insulares do PE, além naturalmente dos dialetos centro-meridionais. Os distritos do Porto e Aveiro evidenciam estar na segunda fase de expansão da inovação, tendo substituído o pronome sujeito *vós* por *vocês*, e os distritos de Viseu e Braga evidenciam estar na primeira fase, quando o pronome *vós* coocorre com *vocês*. Nos dialetos insulares, os dados são menos conclusivos, mas estando todo o Centro-Sul de Portugal pelo menos na terceira fase, com a excepção da área de Lisboa onde a norma culta travou a inovação após a segunda fase, supõe-se que os dialetos insulares também se encontrem, minimamente, na terceira fase, de acordo com o modelo: sujeito > reflexivo/verbo > acusativo > dativo > possessivo. Após

analisar os dados da localidade Covo, em Aveiro, vê-se que este dialeto já removeu o sujeito *vós*, mas continuam a ocorrer verbos na 2PL, com sujeitos nulos e com o pronome sujeito *vocês*. A localidade Granjal, em Viseu, mostra a coexistência de *vós* e *vocês* como tratamentos informais. Os dados destas duas localidades apoiam o modelo proposto aqui de cinco fases de progressão da mudança, a primeira sendo a perda do sujeito. As localidades do Norte de Portugal também mostram uma forte retenção do imperativo da 2PL, e uma retenção menos forte do indicativo, o que leva à proposta de um modelo conjuntivo/condicional/infinitivo pessoal > indicativo > imperativo para a substituição das formas verbais de 2PL pelas de 3PL. Por fim, fazendo uma comparação quantitativa da frequência de sujeitos nulos, os dados mostram que as taxas de uso de sujeito nulo são semelhantes entre os dialetos com cinco morfemas verbais e os com dialetos com seis, apoiando a hipótese de que se requer uma redução mais significativa no paradigma de flexão verbal para que se perca a riqueza funcional necessária para suportar uma LSN estável.

Este trabalho considera ainda outras fusões e inovações nas pessoas gramaticais do português, incluindo a inovação *a gente* no PE, que não originou a perda da morfologia verbal da 1ª pessoa do plural (1PL). Esta inovação é forte no português brasileiro (PB), como é mostrado por Sória (2013), entre muitos outros, mas não é tão dominante no PE. Tomando em conta os dados de Sória (2013), um estudo paralelo a este que examina *nós*, *a gente*, e o sujeito nulo de 1PL no CORDIAL-SIN, vê-se que a inserção do pronome *a gente* no sistema pronominal do PE não causa uma perda do verbo 1PL, pelo menos com sujeitos nulos. Aliás, *a gente* não funciona como *vocês* porque não aceita um sujeito nulo verdadeiro – o pronome *a gente* apenas pode ser elidido quando tem o antecedente acessível no discurso. Isto significa que o PE, até agora, apenas tem uma fusão no paradigma verbal, entre 2PL e 3PL. O pronome *a gente* só originaria no PE um segundo sincretismo se fizesse perder o uso de 1PL com sujeito nulo, o que não é o caso até agora. Pelo contrário, em alguns dialetos *a gente* pode concordar com o verbo na 1PL.

Comparando o PE com o PB, descrevem-se as fusões morfológicas que ocorreram no PB, e particularmente *tu/você*, o que é um fenómeno do século XX que é muito complexo e bem

documentado. Primeiramente, considera-se a diferença histórica e linguística entre o PB culto e o PB vernáculo, este último tendo fortes raízes no contacto linguístico com aloglotas africanos na colonização do Brasil e apresentando traços advindos de transmissão linguística irregular (Baxter, 1997, 2009; Lucchesi 2001, 2009a, 2009b; Lucchesi, Baxter & Silva, 2009; entre outros). Com estas informações, nota-se que o desenvolvimento do sincretismo *tu/você* não segue percurso idêntico ao do sincretismo *vós/vocês* no PE. No Rio de Janeiro, evidencia-se que o emprego do sujeito *tu* desceu drasticamente no início do século XX juntamente com o possessivo *teu(s)/tua(s)*, mas o acusativo *te* continuava a ser usado (Lopes, Rumeu & Carneiro, 2013; Souza, 2014; Lucena, 2016). Analisa-se depois a perda de *vós* no PB. Esta perda terá começado no início do século XIX, ou mais cedo, mas está pouco documentado e é, portanto, difícil de analisar. Nas cartas estudadas por Lopes & Machado (2005) encontra-se o clítico *vos* em mais de um autor, preservado na despedida epistolar, e o pronome possessivo *vosso* na linguagem do autor mais conservador. Portanto, vê-se que o pronome sujeito e a flexão verbal da 2PL desapareceram primeiro, no PB tal como no PE, mas permanece incerto se o clítico *vos* desapareceu antes ou depois do possessivo *vosso*.

Fazendo uma análise holística do PB no início do século XX, quando se tornou uma LSN parcial, apresentam-se outros fenómenos relevantes além da perda dos verbos da 2SG, em particular a tendência que se manifesta em algumas variedades do PB para a expansão da 3ª pessoa do singular (3SG), originando novas fusões morfológicas. Este fenómeno linguístico parece relacionar-se com fatos de história externa, como o influxo grande de negros libertos para o Rio de Janeiro após a abolição, cujos dialetos apresentariam paradigmas verbais com apenas duas ou talvez três distinções morfológicas de pessoa-número, e também um crescente uso da 3SG com um sentido indeterminado e arbitrário (Nunes, 1990). Relativamente à propriedade de sujeito nulo, estes dialetos corresponderiam a uma LSN parcial ou talvez uma não LSN, o que pode ter impulsionado o PB culto a perder o traço [+ LSN].

Como conclusão geral do estudo, confirma-se a hipótese de que a fusão *vós/vocês* no PE não afeta a estabilidade do PE enquanto LSN, afirma-se que o modelo de implementação desta inovação

é sujeito > reflexivo/verbo > acusativo > dativo > possessivo, e propõe-se que, até aos anos 90, a inovação *vocês* foi implantada em todo o Centro-Sul, nos arquipélagos dos Açores e da Madeira e, a Norte, nos distritos de Braga, Porto, Aveiro e Viseu.

Palavras-chave: *vós*, *vocês*, sujeito nulo, 2PL, concordância verbal, concordância pronominal

Abstract

The focus of this dissertation is a detailed study of the 2nd person plural (2PL) pronouns and verbs, which include *vós*, *vos*, and *vosso(s)/vossa(s)*, and the 2PL verbal morphemes in comparison with the dominant and spreading 3rd person plural (3PL) *vocês* innovation, with the corresponding 3PL verbal morphology and pronouns *os/as*, *lhes*, and *seu(s)/sua(s)*, in European Portuguese. Furthermore, null subject pronouns for both of these semantically 2PL addresses are gathered and analyzed to compare rates of null subjects in dialects which retain six grammatical persons in the verbal morphology and those which retain only five, having removed the grammatical 2PL. The hypothesis is that a five person verbal paradigm will not result in lower rates of null subjects, nor will it affect European Portuguese's trait of being a null subject language (NSL). These are studied primarily referencing the *Corpus Dialectal para o Estudo da Sintaxe*/Syntax-oriented Corpus of Portuguese Dialects (CORDIAL-SIN), and with reference to other existing data and studies of the 2PL, particularly Lara Bermejo (2015).

The present study elaborates on the spread of the *vocês* innovation, showing its presence in the northern Portuguese districts of Braga, Porto, Aveiro, and Viseu by and in the insular Portuguese dialects by the 1990s, the latest period of the data used in the CORDIAL-SIN. Refining the Lara Bermejo (2015) model with a separate step for the subject pronoun, the model subject pronoun > reflexive/verb > accusative > dative > possessive is proposed. The location Covo, in Aveiro, supports this model, showing no occurrences of *vós* in contrast with many occurrences of 2PL verbs with null subjects or the subject *vocês*. The study also finds imperative verbs to be the most resilient, and indicative verbs to be the secondmost, giving a replacement model of subjunctive/conditional/personal infinitive > indicative > imperative. Making a quantitative comparison, the data for null subjects shows that rates of use are similar between dialects with five verbal morphemes and those with six, supporting our hypothesis that the null subject property requires a more significant reduction in the verbal paradigm for a language to shift to a partial or non Null Subject Language (NSL).

Finally, this work considers other mergers and innovations in the grammatical persons of Portuguese, including the *a gente* innovation in European Portuguese, which has not replaced the 1st person plural (1PL) verbal morphology and therefore does not constitute a loss of a grammatical person, and the *tu/você* merger in Brazilian Portuguese, providing a nuanced perspective on Brazilian Portuguese's transition to a partial NSL in light of available evidence. The overall conclusion is that the *vós/vocês* merger innovation is incapable of changing European Portuguese from a NSL to a (BP-like) partial NSL, and that the language retains the functional richness Roberts' (1993) proposes as necessary for the property [+ NSL].

Keywords: *vós*, *vocês*, null subject, 2PL, verbal agreement, pronominal agreement

List of abbreviations

1SG	1st person singular
2SG	2nd person singular
3SG	3rd person singular
1PL	1st person plural
2PL	2nd person plural
3PL	3rd person plural
ALEPG	*Atlas Lingüístico-Etnográfico de Portugal e da Galiza/Linguistic-Ethnografic Atlas of Portugal and Galicia*
BP	Brazilian Portuguese
CORDIAL-SIN	*Corpus Dialectal para o Estudo da Sintaxe*/Syntax-oriented Corpus of Portuguese Dialects
EP	European Portuguese
NSL	Null subject language
NSP	Null subject parameter

List of abbreviations of locations

1. **VPA** Vila Praia de Âncora (Viana do Castelo.)
2. **CTL** Castro Laboreiro (Viana do Castelo)
3. **PFT** Perafita (Vila Real)
4. **AAL** Cast.Vide, Porto da Esp., S. Salv. Aramenha, Sapeira, Alpalhão, Nisa (Portalegre)
5. **PAL** Porches, Alte (Faro)
6. **CLC** Câmara de Lobos, Caniçal (Funchal)
7. **PST** Camacha, Tanque (Funchal)
8. **MST** Monsanto (Castelo Branco)
9. **FLF** Fajãzinha (Horta)
10. **MIG** Ponta Garça (Ponta Delgada)
11. **OUT** Outeiro (Bragança)
12. **CBV** Cabeço de Vide (Portalegre)
13. **MIN** Arcos de Valdevez, Bade, S. Lourenço da Montaria (Viana do Castelo)
14. **FIG** Figueiró da Serra (Guarda)
15. **ALV** Alvor (Faro)
16. **SRP** Serpa (Beja)
17. **LVR** Lavre (Évora)
18. **ALC** Alcochete (Setúbal)
19. **COV** Covo (Aveiro)
20. **PIC** Bandeiras, Cais do Pico (Horta)
21. **PVC** Porto de Vacas (Coimbra)
22. **EXB** Enxara do Bispo (Lisboa)
23. **TRC** Fontinhas (Angra-do-Heroísmo)
24. **MTM** Moita do Martinho (Leiria)
25. **LAR** Larinho (Bragança)
26. **LUZ** Luzianes (Beja)
27. **FIS** Fiscal (Braga)
28. **GIA** Gião (Porto)
29. **STJ** Santa Justa (Santarém)
30. **UNS** Unhais da Serra (Castelo Branco)
31. **VPC** Vila Pouca do Campo (Coimbra)
32. **GRJ** Granjal (Viseu)
33. **CRV** Corvo (Horta)
34. **GRC** Graciosa (Angra do Heroísmo)
35. **MLD** Melides (Setúbal)
36. **STA** Santo André (Vila Real)
37. **MTV** Montalvo (Santarém)
38. **CLH** Calheta (Angra do Heroísmo)
39. **CPT** Carrapatelo (Évora)
40. **AJT** Aljustrel (Beja)
41. **STE** Santo Espírito (Ponta Delgada)
42. **CDR** Cedros (Horta)

List of maps

List of figures

List of tables

Table of contents

1. Introduction

1.1 Presenting the 2PL and the null subject

The continued existence of the 2PL subject *vós*, its conjugations, and its related pronouns is a simultaneously well known and little understood fact. In Brazil, you might be told that we do not use that in Brazilian Portuguese (BP), but they do use it European Portuguese (EP). Then in Lisbon, you will be told that we do not use *vós* here, but it is used in the North. Then in Porto, you could be told it is not that common anymore – they use that in the countryside more. So what is the distribution of the 2PL pronouns and conjugations in Portugal? As these answers show, it is a complicated question. In reality, these pronouns and conjugations are extant in Portugal but are being replaced by the innovative *vocês* form and the 3PL conjugations. In this study, we will elaborate on the places where the innovation exists, the phases that it follows, and why, using previous evidence (Lara Bermejo, 2012, 2015, 2018a, 2018b; Segura 2013; Aguiar & Paiva, 2017) while adding conclusions of our own made from a detailed study of the *Corpus Dialectal para o Estudo da Sintaxe*/Syntax-oriented Corpus of Portuguese Dialects (CORDIAL-SIN).

From this data from the CORDIAL-SIN, we will then take a look at the null subject parameter in Portuguese. EP is firmly a null subject language (NSL) while BP is losing this property (Duarte, 1993, 1995; Kato & Negrão, 2000; Barbosa, Duarte, & Kato, 2001; among others) and is now considered a partial-NSL. The commonly proposed reason for BP's shift to a partial-NSL is the reduction in its verbal paradigm. With too few overtly distinct grammatical persons, the language cannot sustain null subjects as they become too difficult to extract, especially given that spoken BP can present three mergers – *vós/vocês*, *tu/você*, and *nós/a gente*, causing a loss of the 2PL, 2SG, and 1PL verbal morphology. Given that these mergers are the purported reason for the shift, we will examine the *vós/vocês* merger in Portugal to see if it causes a decline in null subjects or if it presents parallels with other mergers or pronoun innovations, which we analyze last. Given that the *a gente* innovation is also present in Portugal (Soria, 2013; Lara Bermejo, 2015, among others), though not

as widespread as in Brazil, we might expect to see some loss in null subject properties in dialects with prominent use of *a gente*, though this is not the case, as Soria (2013) shows.

From this information, we will answer the following questions:

i)　　What is the distribution of the innovative informal *vocês*, corresponding 3PL verbs, and its clitics in Portugal?

ii)　　What are the phases of the *vocês* innovation?

iii)　　Is the *vós/vocês* merger proceding in EP the same way that the *tu/você* merger is spreading in BP? Why or why not?

iv)　　Is the *vocês* innovation in EP proceding the same way as the *a gente* innovation in EP? Why or why not?

v)　　Do dialects of EP with merged 2PL/3PL conjugations present lower rates of null subject usage?

1.2 Layout

This first chapter is a methodological introduction to this study. Here we look at the broader goals behind this thesis, the linguistic subject matter that it tackles, the resources utilized to do so, and the details of how the subject was approached and analyzed in the previous literature. We will explain the CORDIAL-SIN in detail, how it was used for this work, and what its shortcomings were in regards to the interests of this thesis.

The second chapter is an introduction to the principal subject, which is the 2PL in EP. We will discuss the traditional, prescriptivist grammar of Portuguese pronouns and verbs using Cunha & Cintra (1984) as a starting point, and then compare this to the descriptive reality of the language, both EP and BP. BP in particular presents a highly divergent oral reality for subject pronouns, verbal paradigms, clitics, and possessives, as Silva (2013) and Menon (1995) help to show. Then we will take a quick look at the evolution of the innovative address *você(s)* in Portuguese to bring us up to the modern day state of this innovation on the Iberian Peninsula. Looking at Lara Bermejo (2012, 2015, 2018a, 2018b), we will observe the *vosotros/Ustedes* merger in western Andalusia for comparison

and to see its relation to the *vós/vocês* merger in EP. Shifting towards EP, Lara Bermejo (2015) shows the progression of this phenomenon in Portugal. That study uses field work to further flesh out the progression of the innovation in Portugal and incorporates the data from the CORDIAL-SIN as well. We will add to this with a finer detailed analysis of the CORDIAL-SIN data, which is also used in this study.

The third chapter presents the theoretical background of null subject languages, a trait that defines EP, but has changed in spoken BP, which has become a partial null subject language (NSL). The null subject parameter (NSP), originally presented by Chomsky (1981) and Rizzi (1982), is what determines this, and our understanding of it has evolved over time. Elaborating on the evolving analysis of what makes a language [+ NSL], we look at the theoretical analyses of Huang (1984, 1989) and Jaeggli & Safir (1989), and finally Roberts (1993), whose premise of functional richness for NSLs we support, with a slight modification. Then, we will analyze the null subject in EP with the help of Duarte (1995) and Aguiar & Paiva (2017)

The fourth chapter presents all the relevant data that was gathered for this thesis. This will allow us to analyze dialectal EP and observe a variety of phenomena. First, we will look at the distribution of *vós* and 2PL verbs, then *vocês* and 3PL verbs, and then clitics. With this information at our disposal, we will be able to add that the insular Portuguese dialects have taken on the *vocês* innovation just like the central-southern dialects and some northern dialects. We will also see a strong argument for a five stage model of evolution for the *vós/vocês* merger, in which the subject pronoun disappears first. With a quantitative analysis of both the whole North and location 19, Covo, Aveiro, we will see that *vós* tends to dissapear prior to the 2PL conjugation. We also will see that the productivity of null subjects is not affected by the merger. Then, by analyzing the 2PL conjugations, we will see that the 2PL imperative is the form most resistant to change, followed by the indicative. Taking into account our data plus data from Segura (2013), we propose a refined map of the *vocês* innovation in all of the Portuguese territory.

The fifth chapter takes a close look at other innovations in Portuguese subject pronouns. Looking at the work of Soria (2013), whose thesis addresses the same theoretical questions as this study, but with regards to the 1PL and the *a* gente innovation instead of the 2PL and the *vocês* innovation, we will make a detailed comparison of these two phenomena and ultimately conclude that the 2PL merger had no effect on EP and neither does the 1PL *a gente* innovation, because this innovation has yet to constitute a merger. While some locations appear to have removed 1PL subject *nós*, no locations have removed null 1PL verbs, meaning that the 1PL and 3PL verbs remain unmerged in these locations. Furthermore, because the pronoun *a gente* lacks the true null subject properties and is only able to be elided when it is accessible in the discourse, it is unclear as to whether the 1PL verb morpheme will merge at all. Theoretically, however, it should not alter the NSP because neither of these mergers compromise Roberts' (1993) theory of functional richness for the verbal paradigm.

We then direct our attention towards the decline of the null subject in BP. In light of the sharp divide between educated BP and vernacular BP (Lucchesi, 1994, 1998, 2001), we will analyze the myriad differences between BP and EP and show why vernacular BP has been a partial NSL for a long time due to its roots in irregular linguistic transmission, producing a variety of creole-like characteristics and having strong roots in Afro-Brazilian Portuguese (Lucchesi, 2009; Lucchesi & Baxter, 2009; Lucchesi, Baxter, & Silva, 2009; Baxter, 1997, 2009; Parkvall & Álvarez López, 2003, among others). From this perspective, we look at the *tu/você* merger in Brazil and see how it has developed according to a different model than the *vós/vocês* model of EP. Between the 2SG subject pronoun, the imperative form, the clitic pronouns, and the possessive pronouns, we can see myriad differences that do not follow the same pattern as the 2PL merger in EP (Lopes & Machado, 2005; Scherre, 2007; Souza, 2014; Lucena, 2016; Lopes, Rumeu, & Carneiro, 2013, among many others), most notably a retention of the accusative clitic *te* as the most resilient 2SG pronoun. We will also observe the *vós/vocês* merger in BP, but this merger is poorly documented and it is uncertain whether its most resilient form was the accusative, as is the case for the 2SG in BP, or the

possessive, as is true for the 2PL in EP. Ultimately, we will analyze the *tu/você* merger, the indefinite and arbitrary interpretation of Ø3SG, and the influx of Afro-Brazilian verbal paradigms to urban centers as relevant factors in the shift to a partial NSL in BP.

The sixth chapter serves to synthesize the results of this thesis, other data, and relevant theories, and then to ask some further questions that this thesis leaves unanswered.

1.3 The CORDIAL-SIN and the ALEPG

The primary source of data for this study is the CORDIAL-SIN – *Corpus Dialectal para o Estudo da Sintaxe*, or Syntax-oriented Corpus of Portuguese Dialects. The CORDIAL-SIN is a corpus of dialectal European Portuguese, a product of 20 years of research coordinated by Ana Maria Martins at the Center of Linguistics of the University of Lisbon (CLUL). It consists of over 600,000 words, and the version used for this study has POS annotations for every word.

The CLUL's collection of recordings for the ALEPG (*Atlas Linguístico-Etnográfico de Portugal e da Galiza*/Linguistic and Ethnographic Atlas of Portugal and Galicia, an atlas still in progress), developed by CLUL's dialectology group, is a much larger body of samples than the CORDIAL-SIN. From the recordings gathered for the ALEPG, available in the CLUL's sound archive, we consulted data for location 8 (in the CORDIAL-SIN, location 87 in the ALEPG), Monsanto, in order to gather more data in this location to verify that it had not begun the *vocês* innovation. In light of some data found at a later stage of completing this project (Segura, 2013; Lara Bermejo, 2015), it would have restrospectively been preferable to examine data for location 28, Gião, Porto, or other locations found to be in the first phase of the *vocês* innovation, but time constraints have prevented this. Still, the data here has led to many interesting conclusions.

1.3.1 Shortcomings

The CORDIAL-SIN has a certain structural shortcoming regarding the research of this thesis; the primary usage of any kind of 2nd person is to refer to the interviewers themselves. In Portuguese, a stranger will normally be treated with a formal address. This is not always the case, as any one

person can have their own idiosyncratic beliefs and preferences surrounding formality in language, but it is the norm for Portugal. This is the case for the speakers interviewed in the corpus; they are older, rural Portuguese, and they use formal addresses when talking to the interviewers. As such, we see this in the data available in the CORDIAL-SIN – the interviewers are often treated with nominal forms, mostly *Os senhores/As senhoras*, sometimes *as meninas*, a polite way for the speakers to refer to women younger than themselves , and in many places there is the form *vossemecês*. The interviewers are also outsiders and strangers, resulting in more formal speech. This data is certainly interesting and informative, and we can say definitively that formal addresses in dialectal Portuguese do not employ 2PL clitics, but it leaves us questioning how these speakers would address someone perceived as an equal, familiar, or subordinate.

Beyond this, because the context of the speech in question is interviews, the interviewees, oftentimes only a singular interviewee, are mostly answering questions for the interviewers. Beyond directly addressing the interviewers, they mostly use a 2nd person address in reported speech when telling stories, or among each other, which is less frequent because it requires a group of at least three speakers. Ultimately we are presented with relatively few instances of informal 2PL addresses because the data is from more formal discursive circumstances instead of a more colloquial conversation.

For the 2PL, there are relatively few examples, which is also because it is the more geographically restricted phenomenon in comparison with 3PL addresses, as the data shows in Chapter 4. Clitic pronouns in general were more infrequent than subject pronouns or conjugated verbs. They simply did not occur often in the given discourse because they were not needed often. Therefore we have few examples of *lhe* or *os/as* and even fewer examples of *vos*.

1.3.2 Our theoretical reorientation with the available data

The original intention of this thesis was to better understand the current state of the traditional 2PL pronouns and conjugations in the Portuguese territory. This grammatical person seemed to be popularly understood to be antiquated and used only in the North, but was the

subject of scant linguistic research and consequently, poorly understood. As such, original aim was to more accurately describe the distribution and syntactic possibilities for 2PL clitics pronouns specifically. Due to a the scarcity of said clitics in the CORDIAL-SIN, this work turned towards the more abundant data yielded for subject pronouns, 2PL verb morphemes, and null subjects with the new goal of studying the null subject parameter (NSP), relating to the reduction of the verbal paradigm from six conjugations to five.

A further development in the focus of this paper was determined by the existing research done by Lara Bermejo (2015). This author carried out a study of the *Ustedes* innovation in western Andalusian Spanish and the *vocês* innovation in EP. For EP, Lara Bermejo (2015) utilized the *Atlas Lingüístico de la Península Ibérica* (Linguistic Atlas of the Iberian Peninsula/ALPI), which had already been analyzed in Lara Bermejo (2012), as well as the CORDIAL-SIN, which he also describes as having scarce data. To supplement the CORDIAL-SIN, then, Lara Bermejo (2015) includes original field work, constituting a geolinguistic corpus of central-southern Portugal, which delves into great detail for 2PL verbs and all 2PL related pronouns – subjects, accusatives, datives, possessives, and obliques. We will discuss this thoroughly in section 2.3. Admittedly, knowledge of that study came at a later phase in this present study, after having extracted 2PL data from the CORDIAL-SIN, but it has not undermined the importance of this work; to the contrary, it provides important and useful knowledge about the *vocês* innovation in EP, to which we have some findings to add, and it allows us to focus more heavily on the null subject.

For an ideal study of the pronouns and verbal morphemes used in Portugal, semantically, as a 2PL, we would need a complete sociolinguistic corpus for the topic of the 2PL that accounts for the variables of age, sex, and education of the speakers and records them in an informal environment across the North, which would complement Lara Bermejo (2015). Even more ideally, such a corpus for all of Portugal would be the ideal work for understanding the current state of the *vocês* innovation. Alas, these projects were not within the scope of this master's dissertation, though I leave the idea open to anyone interested.

1.3.3 Translation

This dissertation being written in English, the data used as examples in the body of this text are translated. Each example is followed by glosses and a free translation, with syntactic annotations on relevant pronouns and verbs in the glosses. *Vocês* is annotated as a 2PL noun, given there is no possible 3PL use for this word, but 3PL verbs and pronouns in agreement with *vocês* are marked as 3PL because these words are originally gramatically 3PL words that have been put in a 2PL semantic context. For the free translation, the common English phrase "you guys", one of many 2PL expressions, is often employed for *vós* or *vocês* to disambiguate plural and singular "you" in English.

1.3.4 Data collection

In collecting data, the focus was collecting all examples of semantically 2PL pronouns and conjugations from the CORDIAL-SIN. We used the POS transcription for this. Our focus was *vós*, *vocês*, and *os senhores/as senhoras*. For the latter, which is a formal, nominal address, there were also cases of *as meninas* as an address. These nominal addresses were ultimately not used in this thesis and are not included in the data. For the subject pronouns *vós* and *vocês*, we collected all occurrences of them, whether functioning as a subject pronoun or in an oblique position. The difference between true subject pronouns with a predicate and subject pronouns in other positions is made clear in the data; the non-subject use of these pronouns is present in the tables that analyze just the frequency of the *vós* and *vocês*, irrespective of verbs.

For verbs, we collected 2PL and 3PL conjugations. For 3PL conjugations, we analyzed all the 3PL conjugations in the corpus with null subjects and carefully determined if they had a 2PL referent or 3PL referent. For 2PL conjugations, there was very little mixing of 2PL verbs with pronouns that were not semantically 2PL, and most had a subject pronoun of *vós*, and a few *vocês*. Some morphologically 2PL verbs were, however, semantically related to another grammatical person, which was not included in our data. This was the case for *vós* as a 2SG address, mostly used in religious context. This was also applicable to the 2SG indicative preterit, which is present in the CORDIAL-SIN as an apparent agreement between 2SG *tu* with the 2PL morpheme *-stes*. This is

semantically 2SG, and furthermore, not a 2PL morpheme at all. This phenomenon, only occurring in the past tense, e.g. *Tu fizestes* instead of *Tu fizeste*, is not an occurrence of the 2PL morpheme. Rather, it is a leveling of the 2SG conjugations. Because all the other 2SG inflected forms have a word final /s/ – *fazes, farás, farias, fazias, fizesses, fizeras, fizeres, fazeres* – there is a leveling pressure to normalize the 2SG *fizeste* by adding /s/, a phenomenon also present in Peninsular Spanish.

1.3.5   Uncertain audio

In the POS annotated version of the CORDIAL-SIN, there are some words that were uncertain to the transcriber. In these cases, the transcription includes the most likely word which is followed by the alternate option annotated as such – <alt> *vos* </alt>. In these cases, we always went with the transcription considered more likely. There was only one such case that would have changed our maps. Nonetheless, the pronoun in question was an example of a dative in the area that allows for dative *vos*, and furthermore would qualify as religious jargon, and as such would not have changed our analysis if it were *vos* instead of *lhe*. The example was:

(1) *Deus   lhe   <alt> vos </alt>   guarde a vocês   de serem como elas.*

God you.DAT.3PL <alt> you.DAT.2PL </alt> protect to you.OBL.2PL from be.INF.3PL like them.

"God save you guys from being like them." (STJ06)

1.3.6   Conjugation anomalies

There were two instances of a 3SG conjugation in place of a 3PL in the corpus. Upon consulting the recordings of these, however, they appear to be 3PL verbs that sound phonetically identical to their 3SG conjugation. In the case of example (2), we are of the opinion that this was a transcription error. For example (3), the speaker does remove the *-em* that should create the word *puderem*, but it appears to be a haplology. From the audio, the speaker produces /pu.der.rẽ.me.djaɾ/ eliding a syllable. As such, these transcriptions are regularized later in our text and counted as 3PL. These occurrences were:

(2) *...e se não **quiser**,   vocês   escolhei...*

And if no want.FUT.SUBJ.3SG, you.NOM.2PL choose.IMP.2PL…

"And if you guys don't want to, choose…" (COV10)

(3) *E   se vocês   **puder**   remediar (...)*

And if you.NOM.2PL can.FUT.SUBJ.3SG remedy…

"And if you guys can fix it…" (COV16)

One curious phenomenon, which in general is included with Ø2PL data except where clearly

mentioned otherwise, was a singular case of a 2SG morpheme being used with the pronoun *vós*. The

example in question is:

(4) *Tendes   de trabalhar para vos   governares.*

Have.PRES.IND.2PL of work   for   you.REFL.2PL govern.INF.2SG.

"You guys have to work to take care of yourselves." (MST, ALEPG)

Prescriptively, this personal infinitive should be conjugated as *governardes*. Because the

example is singular, it could be a dialectal phenomenon of Monsanto or an idiolectal phenomenon of

the speaker. The exclusion of /d/ could be a phonological process. Alternatively, it could result from

morphological interference from the 2SG. Additionally, Monsanto is to the north of the North/South

divide, but is also within a dialectal bubble referred to as a "subdialectal region with highly

differentiated peculiar characteristics" in Map 4, furthering the idea that this may be a phenomenon

related to this isolated center-inland dialectal area.

1.3.7 Abandoned sequences , repetition, and subjects with no verb, an uninflected verb, or several verbs

As is common in impromptu speech, there are instances of abandoned sentences, repetition,

and occasional constructions that use *vocês* without any form of predicate. Abandoned sequences

are marked in the syntactically annotated transcription with (…), and when presented in this text, we

use a simple ellipsis – … –with no parentheses. All of these were counted. Some examples include:

(5) *Ouçam   lá, mas porque é que **vocês**…*

Listen.IMP.3PL there, but why is that you.NOM.2PL.

"Listen here, why is that you guys…" (STJ06)

(6) **Vocês,**   *vocês*   *vejam.*

You.NOM.2PL, you.NOM.2PL see.IMP.3PL.

"You guys, you guys, look." (COV24)

(7) *Ah, não!* **Vocês***!*

Ah, no! You.NOM.2PL!

"Oh no! You guys!" (PIC27)

Subjects sometimes appeared independently with an uninflected infinitive or gerund (inflected gerunds being possible in some dialects). Whenever this happened, the subject was counted but no verb was recorded. For example:

(8) *Era eu a caçar e*   **vocês**   *a amanhar, hem!*

Was I at hunt.INF and you.NOM.2PL at prepare.INF, hey!

"It was me hunting and you guys preparing!" (VPA30)

Furthermore, sometimes a subject would control several verbs. When this was the case, it was counted as an overt subject pronoun for every verb it controlled, but was only counted as one subject pronoun in the count of actual subject pronouns. This is made clear in all tables. Some examples were:

(9) *… e depois* **vocês**   **dai***-o,*   **vendei***-o,*   **dai***-o*   *a quem...*

…and later you.NOM.2PL give.IMP.2PL-it, sell.IMP.2PL-it, give.IMP.2PL-it to whomever…

"… and then give it, sell it, give it to whomever…" (COV02)

(10) *Vocês*   **vieram**   *aqui,* **afincaram***-se*   *a dormir,*

You.NOM.2PL come.PRET.PERF.3PL here, sink.PRET.PERF.3PL-you.REFL.3PL at sleep,

"You guys came here, fell fast asleep," (LUZ37)

1.4 Objectives of this study

The primary goal and hypothesis of this work is to show that the reduction of EP's verbal paradigm from six to five conjugations via the elimination of the historical 2PL does not have a bearing on the NSP. We focus on the 2PL in EP to provide more information on the subject and to then be able to analyze the topic with even greater clarity and perhaps refine our understanding of the characteristics related to a NSL. Because BP is undergoing a shift with respect to the NSP and has changed from a NSL to a partial NSL, another interest of this study is to further shine light on the difference between EP and BP in order to provide a good theoretical explanation for the retention of the characteristic [+ NSL] in EP. We offer a historical analysis of the null subject in BP in Chapter 5, and though we do not undertake new research of BP, we take advantage of the existing ample scholarly data to provide a multi-faceted perspective on BP's shift to a partial-NSL.

In light of the 2PL data for Portugal, this study will show that this verbal merger did not destabilize EP's conjugational paradigm enough to cause a change in the NSP, meaning that there is no significant difference between five and six conjugation paradigms in this regard. There would have to be a much greater reduction in the verbal paradigm to cause this shift, which is in line with existing hypotheses about the NSP, meaning that it was not the catalyst for the shift in BP either. Taking into account the variety of other factors affecting BP, we will make the argument that BP's NSP shift is due to multiple changes in its verbal paradigm, alongside linguistic contact with non-NSL Afro-Brazilian dialects that are the result of irregular linguistic transmission in Brazil.

2 *Vós* and *vocês* in Portuguese: prescriptive grammar, diachrony, and synchrony

2.1 Portuguese grammar

The reality of the 2PL pronouns and conjugations in modern Portuguese is somewhat complex. Its traditional grammatical application is quite consistent, which belies the variable linguistic reality of these pronouns and verbal morphology.

2.1.1 Prescriptive paradigms

If we look at traditional Portuguese grammar manuals, we get a very orderly, clean and complete idea of what subject pronouns in Portuguese are, at least in theory. In almost any scholarly manual we are likely to find a table like the following, in one format or another.

Table 1 Traditional, precriptive subject pronouns in Portuguese

Subject Pronouns	Singular	Plural
1st Person	eu	nós
2nd Person	tu	vós
3rd person	ele, ela	eles, elas

This prescriptivist pronoun table, when taken at face value, is disconnected from the contemporary reality of Portuguese everywhere that it is spoken, even in Portugal. The spoken norm of EP is the central-southern norm based on the speech of Lisbon and Coimbra (Instituto Camões, 2006), which have been centers of higher education historically. In Lisbon, for example, *vós* is absent from the everyday speech of locals, as are the 2PL conjugations (which our data shows in Map 5). Referring to the *Nova Gramática do Português Contemporâneo*, by Cunha and Cintra (1984), which is a comprehensive grammar manual that tends toward a more classical, prescriptivist style, we see the initial presentation of a six conjugation model. Further on, it is clarified that "The pronoun *vós* has practically disappeared from the everyday speech of Brazil and Portugal." (p. 287), excepting ceremonial contexts, and it is used as a 2SG subject in more literary language and in religious jargon.

There is no specific mention of where it may still be used. The manual makes further mention of the forms *você* and *o senhor*, and explains the prominence of *você* in Brazil, where *tu* has largely been replaced, but it does not elaborate on *vocês*. This leaves us to infer that *vocês* is the default informal pronoun, but it is never presented as such. The manual also makes mention of *a gente* as a colloquial 1PL form conjugated in the 3SG, but does not elaborate on its prominence in Brazil nor the non-standard conjugations (i.e. *Nós vai. A gente vamos*) that are often related to the 1PL.

In the case of clitic pronouns, the matter is not quite as complicated as that of subject pronouns. With the exception of accusative *lhe* in BP, this is because there are no innovative clitics replacing the classic clitics in the way that *você*, *vocês*, and *a gente* have come to replace (depending on the dialect) their corresponding 2SG, 2PL, and 1PL pronouns *tu*, *vós*, and *nós*. Once again looking at Cunha & Cintra (1984), we get this basic picture of clitics.

Table 2 Prescriptive Portuguese object pronouns

Object Pronouns	Singular	Plural
1st Person	me	nos
2nd Person	te	vos
3rd person	o, a, lhe	os, as, lhes

The manual's presentation of clitics is not a perfectly descriptive one, however. With regards to *vocês* and the nominal addresses, the book prescribes the 3rd person pronouns that traditionally accompany them and does not recommend using "mixed agreement", which would include using *vos* as the clitic for *vocês*, common in Portugal, or using *te* as the clitic for *você*, common in Brazil, as these phenomena are prescriptively wrong, despite their common use. Furthermore, the manual actively prescribes against the use of what we might call "subject objects", which is a common occurrence in Brazil, i.e. *Eu vi ele* instead of *Eu o vi/Eu vi-o* "I saw him." As such, there is a lack of practical information about how informal address functions in the linguistic reality of Portugal and especially Brazil.

The situation for possessive pronouns is very similar to that of the clitic pronouns. This is to say, the mixing of the 2SG possessive *teu(s)/tua(s)* with the subject *você* is possible in some dialects in Brazil, and mixing the 2PL possessive *vosso(s)/vossa(s)* with the 3PL *vocês* is quite common in Portugal, but these mixtures are not mentioned by Cunha & Cintra (1984).

Table 3 Prescriptive Portuguese possessive pronouns

Possesive Pronouns		Singular Subject	Plural Subject
1st Person	masc.	meu(s)	nosso(s)
	fem.	minha(s)	nossa(s)
2nd Person	masc.	teu(s)	vosso(s)
	fem.	tua(s)	vossa(s)
3rd person	masc.	seu(s)	seu(s)
	fem.	sua(s)	sua(s)

2.1.2   De facto Brazilian paradigms

As mentioned in Cunha & Cintra (1984), the Brazilian pronominal paradigm does not use the 2PL any longer, and many locations do not use the 2SG as well, but this is as far as this pedagogical grammar describes the differences between BP and EP. In the descriptive linguistic reality, BP encompasses up to 4 paradigms for conjugation, which we better see in Silva (2013), which is part of *Gramática do Português*, a highly detailed descriptivist grammar book, more geared towards linguistics than pedagogy. Silva (2013) shows this in Table 4, adapted from page 151 with two slight changes – variable *você/tu* for 4 and 3 form paradigms, because this feature is often regional and we will discuss it thoroughly in section 5.3, and the addition of variable *a gente/nós*, *nós* + 3SG being a distinct possibility as well, seen in Chapter 5 of Sória (2013).

Table 4 Common spoken verbal paradigms in Brazilian Portuguese

	5 forms	**4 forms**	**3 forms**	**2 forms**
1SG	eu falo	eu falo	eu falo	eu falo
2SG	você fala	você/tu fala	você/tu fala	você/tu fala
	tu falas	–	–	–
3SG	ele/ela fala	ele/ela fala	ele/ela fala	ele/ela fala
1PL	a gente fala	a gente fala	a gente/nós fala	a gente/nós fala
	nós falamos	nós falamos	–	–
2PL	vocês falam	vocês falam	vocês falam	vocês fala
3PL	eles/elas falam	eles/elas falam	eles/elas falam	eles/elas fala

For the clitic pronouns, spoken BP is greatly deviant from EP. The dative clitic *lhe(s)* can be used as a semantically 2nd person accusative clitic, and clitics in general, and accusative clitics more so, are losing space to the use of a stressed subject pronoun, as Silva (2013) points out, or alternatively a null object clitic, also common in BP. The 3rd person accusatives *o(s)/a(s)* are rarely used except by highly educated speakers. In Menon (1995), we see confirmation of an accusative *ele*, which the author points to as being widespread among classes. In Ramos (1999, apud Almeida (2009, p. 36)), where we see that any subject pronoun can serve as an object, we can construct this table of object pronoun possibilities for oral BP, ignoring highly educated, prescriptivist phenomena.

Table 5 Common spoken object clitics in Brazilian Portuguese

Subject Pronoun	**Accusative Clitic**	**Stressed Accusative**	**Dative Clitic**	**Stressed Dative**
eu	me	eu	me	para mim/eu
tu	te/lhe	tu	te/lhe	para ti/tu
você	te/lhe	você	te/lhe	para você
ele/ela	–	ele/ela	–	para ele/ela
a gente	–	a gente	–	para a gente
nós	nos	nós	nos	para nós
vocês	lhes	vocês	lhes	para vocês
eles/elas	–	eles/elas	–	para eles/elas

In the case of possessive pronouns in BP, the introduction of both *você* and *vocês* led to greater use of the pronouns *seu(s)/sua(s)* in a 2nd person context. Menon (1995) explains that this pronominal rearrangement consequently led to a disambiguation of these pronouns via the introduction of *dele(s)/dela(s)* for the 3rd person, and for the 2PL, there is such a form as well, *de vocês*. The resultant paradigm for BP is:

Table 6 Common spoken possessive pronouns in Brazilian Portuguese

Subject Pronoun	Possessive Pronoun
eu	meu(s)/minha(s)
tu	teu(s)/tua(s), seu(s)/sua(s)
você	teu(s)/tua(s), seu(s)/sua(s)
ele/ela	dele/dela
a gente	da gente
nós	nosso(s)/nossa(s)
vocês	seu(s)/sua(s), de vocês
eles/elas	deles, delas

These descriptive pronominal paradigms are quite different from the classically prescribed ones seen in 2.1.1. They are the result of a language that has undergone intensive linguistic contact and restructuring, and they offer a great deal of variability, especially regarding the 2SG *tu* and the 3SG *você*. Because the *tu/você* merger in Brazil is a linguistically parallel phenomenon to the *vos/vocês* merger in Portugal, but appears to be having different results, we will discuss these differences in great detail in 5.3. While ideally we would analyze the *vós/vocês* merger in BP in depth and compare it with EP, this merger appears to have been completed in Brazil by the 19th century or perhaps even earlier, and as such we do not have ample linguistic data to make this comparison.

2.2 The 2SG and 3SG in the history of Portuguese

2.2.1 The development of formal addresses

When all the Romance languages were still Latin, there were no formal pronominal addresses, only two second person forms of address – *tu*, the 2SG, and *vos*, the 2PL. With time, an innovation in the 2PL lead to its use to addressing the emperor, seen in Brown and Gilman (1960). They thus divide pronouns into two distinct fields – T-forms (derived from Latin *tu*) for relationships characterized by solidarity, i.e. forms equality, and V-forms (derived from Latin *vos*) for relationships of power. This grammatically 2PL pronoun was semantically used as 2SG formal address and was passed down in the evolution of Romance. The CORDIAL-SIN even shows vestigial use of the 2PL as a singular address in dialectal EP, and it is also used in religious contexts.

Much later, from roughly the 14th century onwards, innovative forms such as *Vossa Mercê*, *Vossa Senhoria*, *Vossa Excelência*, and *Vossa Majestade* were introduced in Portuguese. Faraco (2017) thoroughly explains this era of constant change in forms of address. *Vossa Mercê* appears to be the very first of such in Portuguese, first appearing in writing in 1331 (Luz, 1956). This continual change of forms was a sort of "national preoccupation", which shows in the plays of the time, as Cintra (1972, p. 30) shows. Characters are obsessed with using the correct form of address and thoroughly insulted by being called a "lowly" form like *Vossa Mercê* in the 17th century. As the address evolved, it passed through some intermediate stages like *vossemecê* and *vassuncê* to become *você*, and in Brazil, as Gonçalves (2010) shows, it can be abbreviated all the way to the monosyllabic form *cê*.

From this evolution, we arrive at an intermediate form for those of equal social status but not intimate enough for *tu*. It serves as a replacement for the singular usage of *vós*, which was used in this capacity concurrently with the more respectful forms, but had become archaic or considered highly provincial by the mid-18th century, according to Cintra (1972). This means that all formal addresses had become grammatically 3rd person. In contemporary EP, Cunha & Cintra (1984) define *você* as a median term of respect, generally used between equals or from superior to inferior, being

less respectful than the 3SG nominal form *o senhor* (literally translated as *the sir* or *the lord*) or any similar nominal addresses (*o doutor* " the doctor", *o professor* "the professor", etc.), but more respectful than the intimate *tu*. In the case of Brazil, however, *você* is the T-form for most of the country, having largely replaced *tu* (more on this in section 5.3.1) and *o senhor* is the V-form.

2.3 The 2PL and 3PL on the Iberian Peninsula

Looking at Spanish and Portuguese from the perspective of European versus American, one of the biggest systematic differences is the 2PL; it is only extant in Europe. In the Americas, the historical Latin paradigm of six distinct conjugations has been reduced to an absolute maximum of five, *vocês* + 3PL in Brazil, and *Ustedes* + 3PL in Hispano-America, these two subject pronouns being cognates derived from the medieval innovation *Vossa Mercê/Vuestra Merced*. In contrast, the paradigm of six conjugations still exists on the Iberian Peninsula.

2.3.1   Spain

There are thoroughly significant differences between the state of the 2PL in Spain and Portugal. In Spain, the use of *vosotros* and the 2PL conjugation is standard. The *Ustedes* innovation only exists in western Andalusia, something historically attested in the ALPI (Lara Bermejo 2012). Contemporaneously, the six-conjugation standard is dominant. Due to the socio-historical dominance of the central-northern dialect based largely in the capital, Madrid, the southern variations of Iberian Spanish are not only considered non-standard, but are even derided. Milla Muñoz (2020) shows in his study that non-Andalusians view Andalusian Spanish negatively, and even Andalusians hold some internalized negative views of their dialect. Anecdotally, I can attest to speakers of central and northern Spanish dialects stating their belief that Andalusian dialects are "bad" or "wrong" Spanish. This is an acute source of sociolinguistic pressure against Andalusian linguistic phenomena.

This 2PL innovation in West Andalusian Spanish shows various stages of evolution. Lara Bermejo (2015) goes on to study this phenomenon in much greater detail, for both Spanish and Portuguese. The hierarchy of progression he proposes for West Andalusia is thus:

(11) Subject pronoun > reflexive > verb > accusative > dative > possessive

The sociolinguistic pressure from the Spanish norm makes for an interesting case when discussing this innovation. While in theory, the West Andalusian innovation should continue to progress according to the model above, the sociolinguistic reality proves otherwise. Because Andalusian dialects are nonstandard, they suffer pressure from normative Castilian Spanish. As such, Lara Bermejo (2015) found that the phenomenon had not expanded in any zones of West Andalusia in the near 100 years between the data from the ALPI and the data from his study.

2.3.2   Portugal

A point of interest highlighted by Lara Bermejo (2012) is that this Ibero-Romance innovation of consolidating the 2PL and the 3PL into the grammatically 3PL *vocês/Ustedes* occurs in a contiguous area, which includes all of southern-central Portugal and western Andalusia, which borders the Algarve, shown in Map 1. This innovation in the 2nd person seems to have arisen in the southwest of the Iberian Peninsula, and Lara Bermejo (2018b) considers the region to be a cross-border *Sprachbund*, sharing lingustic properties such as the *vocês/Ustedes* innovation, which is more clearly visible in Map 2, where we can see that the most advanced stage of the innovation in Portugal is at the border with Huelva, Spain.

Map 1 The *vós/vocês* and *vosotros/Ustedes* merger on the Iberian Peninsula

Lara Bermejo (2012, p. 65)

In Portugal, the linguistic prestige of the 2PL pronouns and conjugations is the opposite of Spain – the innovative *vocês + 3PL* conjugation is considered standard, because central-southern Portuguese is the norm, and the historically predominant *vós* and its 2PL conjugations are located only in dialectal speech in the linguistically conservative North. So, while this innovation may have even originated in Spanish prior to Portuguese, it has stopped spreading in Spain. The *vocês* innovation is in various stages of progression across Portugal and could potentially replace all 2PL grammar entirely, and has already done so in the Southeast, seen in Map 2. Lara Bermejo (2015) proposes the following model from his research of Portuguese.

(12) Subject pronoun/reflexive/verb > accusative > dative > possessive

Map 2 The *vocês* innovation in continental Portugal in Lara Bermejo (2015, p. 443)

Using data from the CORDIAL-SIN and his own field study, Lara Bermejo (2015) furthers the work of Lara Bermejo (2012) and provides the information seen in Map 2. He proposes that the *vocês* pronoun now dominates *vós* up to and including the districts of Aveiro and Viseu, in the sense of using *vocês* as a T-form. Lara Bermejo's (2015) field work for Portuguese constitutes a geolinguistic corpus, comprised of interviews with only elderly speakers, only as far north as Viseu, leaving us with only the CORDIAL-SIN for reference for the rest of the North.

Map 3 Locations of interviews carried out by Lara Bermejo (2015, p. 157)

Segura (2013) uses data from both the CORDIAL-SIN and the ALEPG to show some examples of 2PL verbs and the pronoun *vós* in the North, though she does not offer a detailed regional analysis. Some of these examples help us to further identify 2PL phenomena in the North. Supplementing the CORDIAL-SIN, which is derived from the same data from the same time period, we see concrete examples of *vós* + 2PL in districts where the CORDIAL-SIN was lacking occurrences. Braga, in particular, had no 2nd person data of any kind, but Segura (2013, p. 130) provides some:

(13) ***Ide-vos*** *embora!*

Go.IMP.2PL-you.REFL.2PL out!

"Leave!" (ALEPG, Vila Boa de Bucos, Braga)

(14) ***Vós*** *não **vos** **mexeis**!*

You.NOM.2PL no you.REFL.2PL move.PRES.IND.2PL!

"You guys don't move!" (ALEPG, Gondomar das Taipas, Braga)

Beyond this, there are occurrences of mixed 2PL agreement with the subject, i.e. *vocês* + 2PL. Many of them are in Covo, Aveiro, a location very well represented in the CORDIAL-SIN that we will analyze thoroughly in Chapter 4, and there are also occurrences from the ALEPG showing the phenomenon in Viseu and Braga (Segura, 2013, p. 131).

(15) A   canga   que **vocês**   **tendes**.

The blanket that you.NOM.2PL have.PRES.IND.2PL.

"The blanket that you guys have." (ALEPG, São Romão da Ucha, Braga)

(16) **Vocês**   não **se**   **preocupais**   que eu cá   me   ajeito.

You.NOM.2PL no you.REFL.3PL preoccupy.PRES.IND.2PL that I here myself fix.

"Don't you guys worry, I'll make myself comfortable." (ALEPG, Malhada, Viseu)

Beyond this work, there is further data on the topic from Aguiar & Paiva's (2017) study on Braga speech. Drawn from the *Perfil Sociolinguístico da Fala Bracarense* (Sociolinguistic Profile of Braga Speech, a project carried out from 2012-2014), this work shows that several phenomena are occurring in relation to 2PL verbs and pronouns, and they are tending to disappear. The most obvious phenomenon in the study is the disappearance of the subject pronoun – *vós* is losing usage to the innovative *vocês*, which is predominant among younger generations and those with higher levels of schooling. The mere two examples of *vós* as a subject pronoun were found in the speech of women over 60 with a low level of education. This is not the case for the object pronouns, however, because Aguiar & Paiva (2017) find that *vos* and *convosco* still appear among those with higher levels of education. The use of the null 2PL is still prominent, but can be seen to be decreasing in favor of the null 3PL. While the study finds some 2PL verbs to persist, they are mostly specific verbs with irregular roots like *ter, ir,* and *ser*.

This information is in line with Lara Bermejo (2015) and also Lopes & Machado (2005), showing the 2PL oblique and possessive pronouns in vigor in speakers who no longer use *vós*. It also

leads us to a critique of Map 2 and the model shown in (12) for two reasons. Firstly, this model shows the subject pronoun, reflexive, and verb as the initial transitional phase, and includes the Lisbon area in this phase with the newest transitional zones in the North. Though Lisbon has resisted the next phase – the replacement of the accusative 2PL pronoun – it has already replaced the subject pronoun, reflexive pronoun, and verb morpheme, it should be considered to have at least completed a second phase of transition, ahead of places like Viseu and Braga where the *vós* subject pronoun and 2PL morphology are still present. This indicates a clear phase for the subject pronoun, which is eliminated from a speaker's grammar before the 2PL verbal morpheme, leading us to propose this model:

(17) Subject pronoun > reflexive/verb > accusative > dative > possessive

We will also see evidence for this later in our closer analysis of the data from the CORDIAL-SIN, and particularly Aveiro, where the location Covo offers robust data. This insertion of the subject pronoun phase implies another problem in light of Aguiar & Paiva (2017), which is that the first phase may be spread much farther than Aveiro and Viseu. Taking into account example (15), showing mixed agreement in Braga, we should expect to see the *vocês* innovation in both Porto and Braga. Further study of more detailed sociolinguist data from the North would help solidify this conjecture.

Though we both employ CORDIAL-SIN data, Lara Bermejo (2015) focuses more on T-forms and V-forms and the phases of replacement of the informal 2PL by the informal 3PL. The theoretical explanation for this hierarchy of transition is explained by Wechsler & Zlatic (2000 and 2003, apud Lara Bermejo (2018a)). The agreement of subject and predicate is index agreement, which is contingent upon person, number and gender. Index agreement causes the reflexive clitic and the verb to change soonest in order to agree with the subject in person. The clitics and possessives, however, work via concord agreement, based on case, gender, and number, and as such they are independent of the subject, and furthermore they do not necessarily appear in a sentence with a subject or predicate, explaining why they change last. In general, this present study defers to these

conclusions on the matter, with evidence for a distinct first phase for the subject pronoun in EP being our only point of contention. Because we only have a single example of a mismatched reflexive clitic in example (16), we will not suggest a phase for the reflexive like Lara Bermejo (2015) found for Spanish.

Our proposed progression, which replaces the subject pronoun first, is a model which is inherently favorable to a NSL. Because the subject pronoun can be null, its initial removal does not change the conjugational paradigm, which retains the person in the verbal morphology. Moving into the second phase, however, the morphology begins to change, which brings us to a question: will a reduction affect the null subject parameter of this language? And moreover, why or why not? This question leads us to the main theoretical interest of this thesis – the null subject parameter. Departing from the same data in the CORDIAL-SIN, we will take a closer look at the null subject in EP.

3 Hypothesis: the NSP and its relation to the pronominal paradigms of European and Brazilian Portuguese

3.1 The null subject parameter

The null subject parameter (NSP), or alternatively the pro-drop parameter, is a concept most prominently proposed by Chomsky (1981) and Rizzi (1982), which figures into Chomsky's theory of principles and parameters as a part of universal grammar. Universal grammar is the theoretical basis of language – the innate linguistic faculty of human beings. It manifests itself in principles, which are not language specific, which then lead to parameters, which are defined one way or the other according to the language. He later described this with the metaphor of a switch box .

> We can think of the initial state of the faculty of language as a fixed network connected to a switch box; the network is constituted of the principles of language, while the switches are the options to be determined by experience. When the switches are set one way, we have Swahili; when they are set another way, we have Japanese. (Chomsky, 2000, p. 8)

Our relevant principle is that all sentences must have a subject, which can be explicit or implicit; the parameter, the NSP, is the binary option that is either set to allow null subjects in a language or forbid them. Continuing Chomsky's metaphor, the switch is set to off in French, and it is set to on in Italian, so we can mark French [– NSL] and Italian [+ NSL]. The original explanation for [+ NSL] was that it manifests in languages with a strong morphological paradigm, such as Italian, where the pronoun can be dropped because it is included in the verbal morpheme. French or English, conversely, have an unclear morphological paradigm, and therefore the speaker is required to clarify the subject pronoun, hence [– NSL]. The setting of the parameter was not just related to this one property, however, as Rizzi (1982) proposed a cluster of related properties, which is well described by Roberts & Holmberg (2009), and shown in Sessarego & Gutiérrez Rexach (2017), from whom I have adapted some examples. The original cluster of [+ NSL] properties are the following four.

(18) The possibility of a silent, referential, definite subject of finite clauses

      ∅ Hablo español.   (Spanish, + NSL)

      *Speak Spanish.   (English, – NSL)

      "I speak Spanish."

(19) Free subject inversion

      Está bom assim, acho eu.   (European Portuguese, + NSL)

      *It's fine like this, think I.   (English, – NSL)

      "It's fine like this, I think."

(20) The apparent absence of complementizer-trace effects

      Chi hai detto che há scritto questo libro?   (Italian, + NSL)

      *Who did you say that wrote this book?   (English, – NSL)

      "Who did you say wrote this book?"

(21) Rich agreement inflection on finite verbs

      Yo como /ko.mo/, tú comes /ko.mes/ ,él come /ko.me/, etc. (Spanish, + NSL)

      Je mange /mãʒ/, tu manges /mãʒ/, iI mange /mãʒ/, etc.   (French, – NSL)

      "I eat, you eat, he eats, etc."

The concept of parameters, then, was thought to be binary and to come with a cluster of properties because this facilitates language acquisition for the learner (see Hyams, 1986; Liceras, 1989; among others). Nevertheless, the above properties were shown to be an unfit explanation for Mandarin Chinese, Japanese, and Korean, which use null subjects despite a complete absence of a verbal agreement system, as pointed out in Huang (1984, p. 537) and therefore the property suggested in (21) cannot be said to be part of a mandatory cluster of properties for all [+ NSL] languages because it is not present in these languages. This author argues that [+ NSL] is viable in Italian and Spanish because their rich agreement allows for the pronoun to be identified, and it is valid in Mandarin and Japanese because the complete absence of agreement leads the listener to identify the subject from a superordinate noun phrase. English, however, has verbal agreement, but

it has historically degenerated from a rich paradigm and no longer allows for a successful subject recovery, and is therefore [– NSL]. The availability of pro-drop in languages like Mandarin and Japanese is attributed to a theory of generalized control, and is expanded on in Huang (1989).

As such, we have the proposal that [+ NSL] is possible for languages that are not necessarily morphologically rich but "morphologically uniform", as is supported by Jaeggli & Safir (1989). Their proposal is thus; " An inflectional paradigm P in a language L is morphologically uniform if P has either only underived inflectional forms or only derived inflectional forms." (p. 30). In other words, morphological uniformity means no mixing of stems with morphemes and morphemeless stems. This hypothesis does not hold up in regard to Old French, however, as Roberts (1993) elaborates that it has six verbal morphemes, with a 1SG null morpheme (i.e. an uninflected stem), meaning that for a verb like *chant* (to sing), the infinitive is *chant* and the 1SG is also *chant*. As such, he proposes a different concept, "functional richness". This author supports this with the specific hypothesis:

(22) "A [+ pron] paradigm allows up to one syncretism." Roberts (1993, p. 127).

He uses the term "syncretism" here to describe a convergence of two morphologically distinct grammatical persons into one phonological form. With this in mind, the disappearance of 2PL should have negligible effect on pro-drop in EP because this leaves the verb paradigm with only one syncretism.

This conception of the NSP helps us understand why Romance-based creoles are all [– NSL]. As Lipski (1999) points out, all the Romance creoles have mandatory overt pronouns. Whether the lexifier language is French [– NSL], or Spanish or Portuguese [+ NSL], the resultant creole does not accept null subjects, with the possible exception of a null expletive (i.e. a null subject to convey the phrase "*It is raining*.") Lipski (1999) explains that these languages have an insufficient verbal morphology to license null subjects like their [+ NSL] lexifiers can, and that they lack the discourse oriented properties that license null subjects in Sinitic and other East Asian languages, such as null, nongap, or multiple topics. Without any of these properties, creoles cannot license null subjects, and

the instances where there are null subjects in these languages are not fully independent null subjects, but rather null constants.

3.2 The null subject in European Portuguese

With this information in mind, we can examine EP's status as a NSL. Duarte (1995) examines both EP and BP in her work, and finds that EP remains firmly a NSL.

Table 7 Occurrence of the null subject in European Portuguese

Person	Coordinated phrase Number/Total (%)	Uncoordinated phrase Number/Total (%)
1st	334/561 (60%)	243/459 (53%)
2nd	101/138 (73%)	96/133 (72%)
3rd	303/417 (73%)	194/305 (64%)

(Duarte, 1995, p. 8)[1]

Duarte (1995) does not, unfortunately, differentiate between singular and plural in this data. Nevertheless, this data still shows that EP has a marked preference for the null subject, which is relatively more frequent in the 2nd person than in the 1st or 3rd. She reports that the competing forms of 2nd person address – the grammatical 2nd and 3rd person addresses – do not significantly affect the null subject percentages: the grammatical 2nd person manifests 70% null subjects (49/70), and the grammatically 3rd person but semantically 2nd person address manifests 76.5% null subjects (52/68) (Duarte, 1995, p. 10). Therefore, the *você(s)* forms of address appears to have no effect on the NSP.

Aguiar & Paiva's (2017) data helps elucidate our subject specifically, differentiating between *vós*, *vocês*, and their null subject 2PL and 3PL conjugations. Overall, null subjects dominate the 2nd person in Braga speech, constituting 76.1% of all occurrences, a similar rate to the 72.7% shown by Duarte (1995) for all 2nd person addresses. When we separate the 2nd and 3rd person addresses in Table 8, the rate for *vós* is particularly high, showing 91.3% null subject use, 21/23 instances. In

---

[1]This data comes from a sociolinguistic corpus compiled by the author. The data is drawn from Nascimento & al. (1987), *Português Fundamental, vol. 2 – Métodos e Documentos, tomo 1 –inquerito de frequência*.

contrast, null *vocês* has a rate of only 60.9%. In the context of all the occurrences, this aligns with the model suggested in (17); the subject pronoun is the first to go in EP, and it is disappearing in Braga speech.

Table 8 Semantically 2PL subjects and verbs in Braga speech

Subject	Number/Total	Frequency
Vós	2/46	4.3%
Ø2PL	21/46	45.7%
Vocês	9/46	19.6%
Ø3PL	14/46	30.4%

Adapted from Aguiar & Paiva (2017)

European Portuguese still appears to be firmly [+ NSL]. Both Duarte (1995) and Aguiar & Paiva's (2017) data show EP as having a global percentage of 2nd person null subject in the 70th percentile, specifically 76,1% in Aguiar & Paiva (2017), for 2PL only, and 72,7% (197/271) in Duarte (1995), which does not distinguish between 2PL and 2SG. As such, we move forward with the hypothesis that global percentages of null subjects for 2PL pronominal addresses, these being *vós* or *vocês*, are similar between EP dialects with 2PL verbs and those without them.

Though EP fits neatly within the descriptions of NSLs provided in 3.1, others languages do not, and the NSP in particular and Parametric Theory in general are not without shortcomings. Looking at the property in (18), this can be applied to both EP and BP, but in (19), we see that EP still has free subject inversion, which is not always possible in BP. Furthermore, BP is not a creole language, but has a history of irregular linguistic transmission which appears to have left creole characteristics (see section 5.2; Lucchesi, 1994, 1998, 2001; and Lucchesi & Baxter, 2009), making it not truly [+ NSL], though the influence of creolization on BP is contentious among linguists. Accordingly, our thoughts on parameters have become more refined in light of a larger body of comparative evidence from some languages that are not quite [− NSL] nor [+ NSL], which are now considered to be a partial pro-drop languages.

3.3 Criticism of the NSP and implications from Brazilian Portuguese

The NSP in BP is a fascinating topic because BP appears to be losing its null subject properties and is the subject of much research among linguists. BP contemporaneously is neither strongly [+ NSL] nor [− NSL]. Furthermore, this is also true of Dominican Spanish (Toribio, 2000; Cabrera-Puche, 2008; among others) and Chinchano Spanish (Sessarego & Gutiérrez Rexach, 2017). These being Afro-Hispanic dialects displays a notable similarity with vernacular BP. Many linguists posit that BP, in its educated norm, is in a transition from [+ NSL] to [− NSL] (Duarte, 1993, 1995; Kato & Negrão, 2000; Barbosa, Duarte & Kato, 2001; Camacho 2008, 2013; among others).

In BP specifically, Duarte (1995) shows a sharp contrast in null subject usage compared to her findings for EP. In a diachronic study of null subject usage in popular plays,[2] we can see that the PB had comparable rates of null subject usage up to and including 1918, but between 1918 and 1937, null subject use for semantically 2nd person addresses in the grammatical 3rd person fell drastically from 69% to 25%, coinciding with the loss of the subject *tu*. In the latest period, Duarte (1995) notes that *nós* competes with *a gente*. With this evidence, the author purports that the loss of *tu* was the trigger for BP's shift to a partial NSL, violating Roberts' (1993) model seen in (22). She argues that this is due to the presence of two syncretisms, these being *tu/você* and *vós/vocês* that consolidated into the 3rd person conjugation, and this now violates Roberts (1993) rule of allowing one syncretism. However, two discreet double syncretisms exist in the subjunctive, conditional, and past imperfect of both Latin American Spanish and most Portuguese dialects (1SG/3SG and 2PL/3PL) but they do not necessitate an overt subject, as seen in example (23) from EP. All of the words in bold have a 1SG null subject, which is only evident contextually, despite the words *viesse*, *tinha*, *gostava*, and *evitasse* having an ambiguous morpheme that can be either 1SG or 3SG.

---

[2]The plays used are "O noviço" by Martins Pena (1845), "Como se fazia um deputado" by França Jr. (1882), "O simpático Jeremias" by Gastão Tojeiro (1918),) "O hóspede do quarto nº. 2" by Armando Gonzaga (1937), "Um elefante no caos" by Millôr Fernandes (1955), "A mulher integral" by Carlos E. Novaes (1975), and "No coração do Brasil" by Miguel Falabella (1992). All of these playwrights are natives of Rio de Janeiro.

(23) *Porque me parecia uma coisa detestável, um problema grande de consciência, se amanhã como magistrado ø **viesse.1SG** a verificar que ø **tinha.1SG errado** ou que ø **tinha.1SG julgado** mal em prejuízo de alguém... Hoje ø **gostava.1SG de ser** magistrado porque ø **estou.1SG** convencido que ø **talvez evitasse.1SG** mais injustiças como magistrado do que as que ø **posso.1SG** evitar como advogado.* (Duarte, 1995, p. 11).

Furthermore, *você* is also present in EP's pronominal paradigm, but having competing forms does not affect the rates of null subject use in the language, as mentioned in 3.2. As such, we only consider a convergence of verbal morphology to be a syncretism, and we propose here a rephrasing of Roberts' (1993) idea:

(24) A [+ NSL] verbal paradigm allows any number of double syncretisms, but no triple syncretisms.

This appears to better fits Roberts' (1993) example of Italian, which has a triple syncretism (a phonological convergence of three grammatical persons) in the present subjunctive. A word like *parlare* (to speak) becomes *che (io) parli*, *che tu parli*, *che (lui) parli*, manifesting an overt *tu* for the 2SG[3], which compensates for the irrecoverability of the triple syncretism. So, while the *tu/você* merger would then constitute a triple syncretism of 1SG, 2SG, and 3SG in the subjunctive, conditional, and past imperfect in BP, we consider this unlikely to have triggered the shift to a partial NSL because the method of disambiguating one of the three pronouns, used by Italian, would constitute a much simpler solution. This would be the path of least resistance for the language. Instead, we propose the influence of additional factors in BP's shift, which we will discuss thoroughly Chapter 5 where we compare and contrast other mergers with the *vós/vocês* merger in EP, offering a more nuanced position on BP's shift to a partial NSL in 5.5.

To better understand where these partial NSLs fit within the parameter, Toribio (2000) suggests that the speakers of these dialects are code-switching between competing mental

---

[3] This is generally mentioned as common knowledge without further citation, such as in Roberts (1993, p. 127) and Roberts & Holmberg (2009, p. 6).

grammars as bilinguals do. Sessarego & Gutiérrez Rexach (2017) suggest that these languages

contain competing lexical entries. The status of these languages means that Parametric Theory does

not hold up to scrutiny in its original form. These aforementioned hypotheses suppose that the NSP

is valid but in need of refinement, and take on what Roberts & Holmberg (2009) refer to as

"microparametric" work and associate parameters with lexical entries, which are set when the

entries are learned. Others, like Newmeyer (2004) or Haspelmath (2008) reject parameters entirely.

Parametric Theory aside, linguists have developed another classification to accommodate these

languages, calling them partial pro-drop languages, (Holmberg, 2005, 2009a, 2009b, 2016;

Holmberg, Nayudu, & Sheehan, 2009; among others).

This present study does not aim to present strong new ideas either for or against

parameters, though I personally hold doubts about how valid the idea of preinstalled parameters is

considering the great adaptability that we now understand the human brain to have. Nevertheless,

they are useful in the context of languages and their changes, and perhaps if we imagine them as a

lever on a fulcrum, which is more likely to fall to one side or another than remain perfectly balanced

in the middle, instead of a switch, which can only be on or off, they may make more sense. As such,

this study will operate within Parametric Theory for EP, expecting that the small changes in EP's

verbal and pronominal paradigm will not be strong enough to swing the lever towards [− NSL].

4 *Vós*, *vocês*, and null subjects in dialectal European Portuguese

4.1 Geographical distribution

Briefly rehashing some methodological concerns, the CORDIAL-SIN is a corpus focused on

dialectal syntax. It is not specifically geared towards the morpho-syntactic questions pertaining to

pronouns that this thesis is directed towards. As such, the data gathered for this study is relatively

scarce; ideally, we would have at least a double digit number of subject pronouns, conjugated verbs,

accusatives, datives, and possessives in every location, but there are very few clitics and possessives.

We will not analyze possessive pronouns at all, nor will we compare dative and accusative clitics, and

instead defer to the analysis done in Lara Bermejo (2015).

One notable insufficiency is that not all locations have an example of *vocês* (or any cognate

forms), which should in theory be omnipresent, at least as a V-form. A substantial part of the data

gathered was formal nominal addresses, e.g. *os senhores/as senhores/as meninas*. This data, though

very informative and useful in its own context, is not relevant to the analysis of null subject

pronouns because these addresses are not true pronouns. Furthermore, the clitic *vos* is only present

in a minority of locations despite being common in the spoken norm, especially as a dative.

Nonetheless, the data here shows a good panorama for the broader use of *vós* and 2PL

conjugations in Portugal. Though we cannot analyze many specific locales in great detail, the total

number of uses is large enough to be significant for both *vós* and *vocês* and from this, we can see

trends of distribution and usage.

4.1.1 Distribution of *vós* and Ø2PL

Our data reaffirms the isogloss for the North of Portugal, as proposed by Cintra (1971). Even

our locations in Castelo Branco and Coimbra fit into this division. So, as pertaining to the retention of

all six conjugations, the data shows the North to be Viana do Castelo, Braga, Vila Real, Bragança,

Porto, Aveiro, Viseu, Guarda, northern Castelo Branco, and eastern Coimbra, which is shown in Map

5 and coincides with the area of Portuguese septentrional dialects as in Cintra's Map 4.[4]

Map 4 Galician-Portuguese dialect classification, according to Cintra (1971)

---
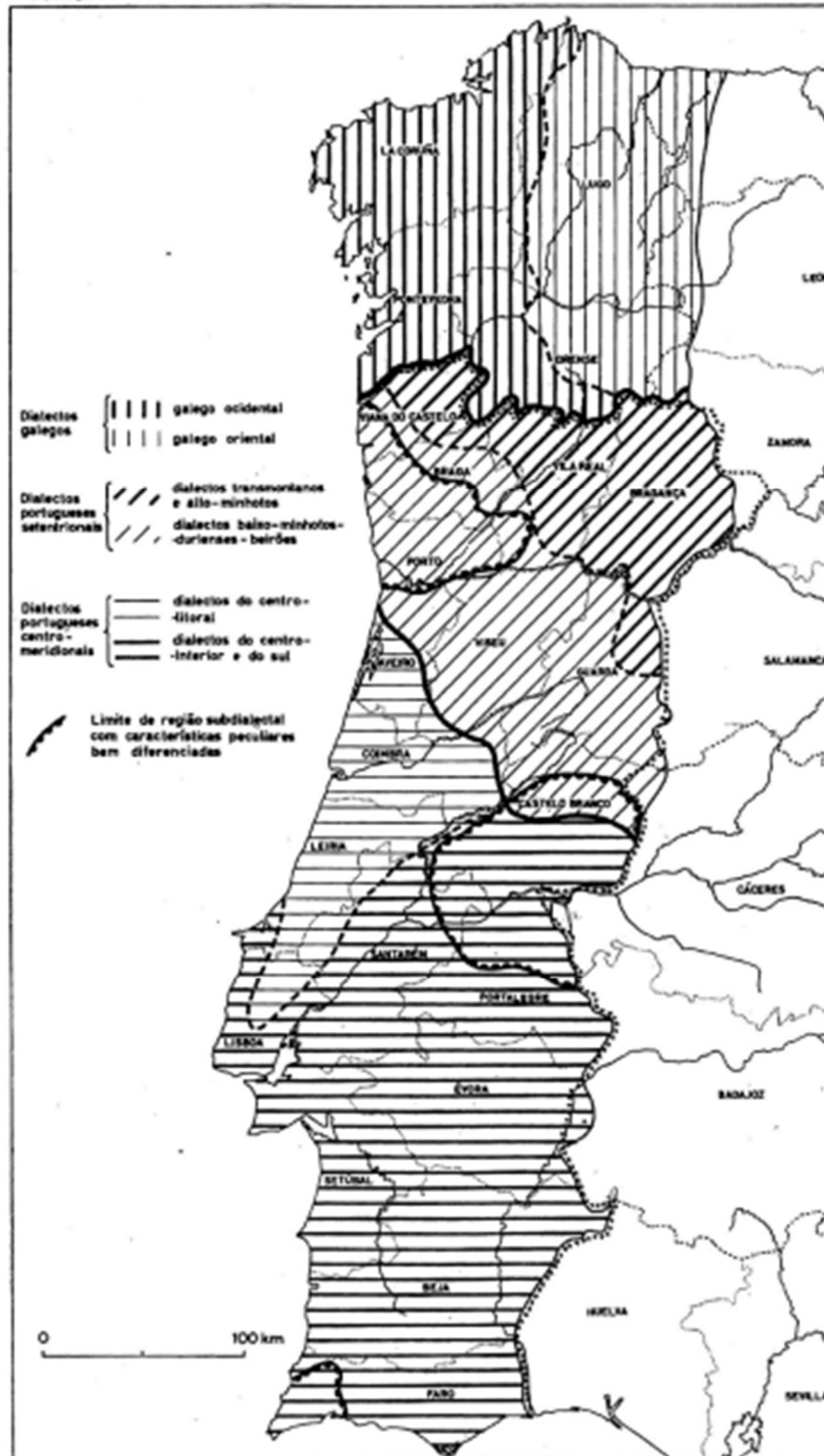
[4] Translated, the solid line between diagonal lines and horizontal lines marks the separation between northern dialects and central-southern dialects.

Map 5 Distribution of *vós* and 2PL conjugations in Portugal
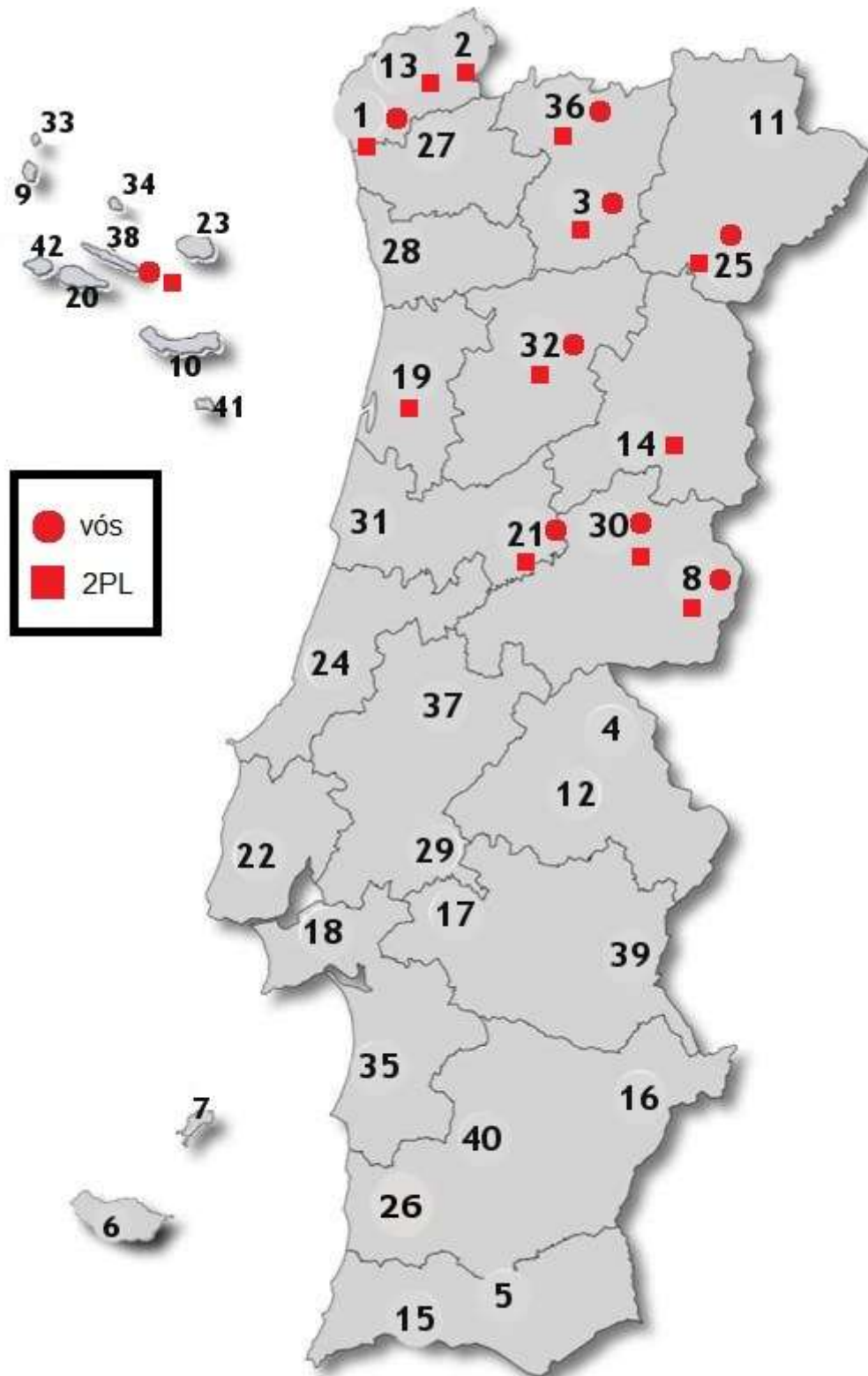
Locations with *Vós* and 2PL

1. VPA Vila Praia de Âncora (Viana do Castelo)
3. PFT Perafita (Vila Real)
8. MST Monsanto (Castelo Branco)
21. PVC Porto de Vacas (Coimbra)
25. LAR Larinho (Bragança)
30. UNS Unhais da Serra (Castelo Branco)
32. GRJ Granjal (Viseu)
36. STA Santo André (Vila Real)
38. CLH Calheta (Angra do Heroísmo)

Locations with 2PL only

2. CTL Castro Laboreiro (Viana do Castelo)
13. MIN Arcos de Valdevez, Bade, S. Lourenço da Montaria (Viana do Castelo)
14. FIG Figueiró da Serra (Guarda)
19. COV Covo (Aveiro)

Map 5 shows us several potential phenomena. The locations all evinced 2PL verbs, though *vós* is not present in every one. They create a well-defined North, and beyond that, we can see a potential divide between a more conservative inland North versus a more innovative coastal North, which is in the first phase of the change, eliminating the subject pronoun *vós*. We will discuss this further in section 4.4.

4.1.2   Distribution of *vocês* and Ø3PL

The distribution of *vocês* (or cognates) covers most of the Portuguese territory. Of all locations in the CORDIAL-SIN, only 27 and 42 provided no data. Only 38 (Calheta), provided 2PL data but no 3PL data (except for one 3PL clitic, clitics being addressed in 4.2). Because of the predominance of *os senhores/as senhoras* in the corpus, several locations did not provide examples of *vocês*, only some Ø3PL examples.

Map 6 Distribution of vocês and 3PL conjugations in Portugal

Locations with *vocês* and 3PL

1. VPA Vila Praia de Âncora (Viana do Castelo)
2. CTL Castro Laboreiro (Viana do Castelo)
3. PFT Perafita (Vila Real)
4. AAL Cast.Vide, Porto da Esp., S. Salv. Aramenha, Sapeira, Alpalhão, Nisa (Portalegre)
5. PAL Porches, Alte (Faro)
7. PST Camacha, Tanque (Funchal)
10. MIG Ponta Garça (Ponta Delgada)
12. CBV Cabeço de Vide (Portalegre)
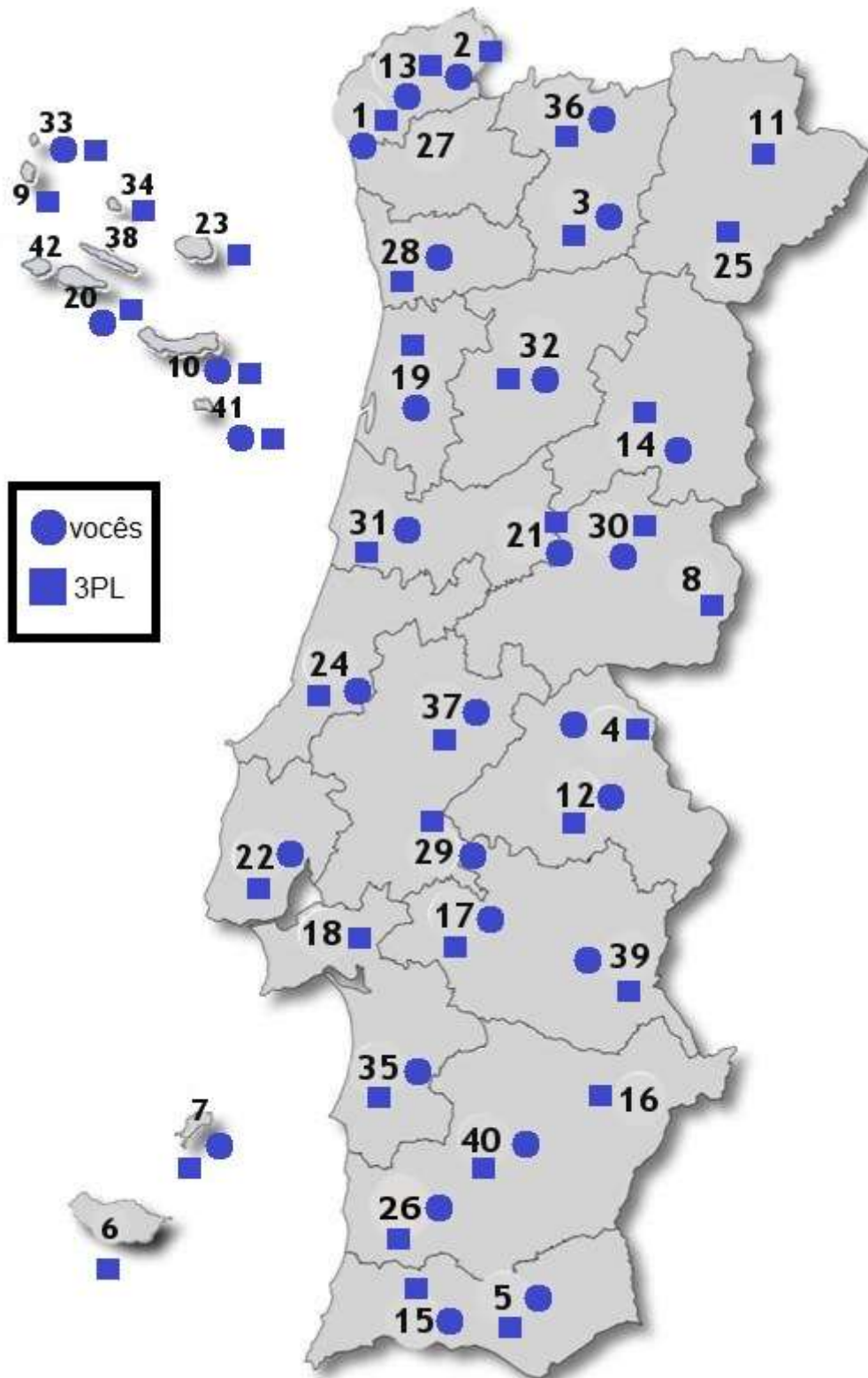13. MIN Arcos de Valdevez, Bade, S. Lourenço da Montaria (Viana do Castelo)
14. FIG Figueiró da Serra (Guarda)
15. ALV Alvor (Faro)
17. LVR Lavre (Évora)
19. COV Covo (Aveiro)
20. PIC Bandeiras, Cais do Pico (Horta)
21. PVC Porto de Vacas (Coimbra)
22. EXB Enxara do Bispo (Lisboa)
24. MTM Moita do Martinho (Leiria)
26. LUZ Luzianes (Beja)
28. GIA Gião (Porto)
29. STJ Santa Justa (Santarém)
30. UNS Unhais da Serra (Castelo Branco)
31. VPC Vila Pouca do Campo (Coimbra)
32. GRJ Granjal (Viseu)
33. CRV Corvo (Horta)
35. MLD Melides (Setúbal)
36. STA Santo André (Vila Real)
37. MTV Montalvo (Santarém)
39. CPT Carrapatelo (Évora)
40. AJT Aljustrel (Beja)
41. STE Santo Espírito (Ponta Delgada)

Locations with 3PL only

6. CLC Câmara de Lobos, Caniçal (Funchal)
8. MST Monsanto (Castelo Branco)
9. FLF Fajãzinha (Horta)
11. OUT Outeiro (Bragança)
16. SRP Serpa (Beja)
18. ALC Alcochete (Setúbal)
23. TRC Fontinhas (Angra-do-Heroísmo)
25. LAR Larinho (Bragança)
34. GRC Graciosa (Angra do Heroísmo)

Locations with no *vocês* do not display much geographic coherence. Their unifying factor is a small data set for the 3PL. Map 6 shows us that 3PL verbs used as a 2PL address are ubiquitous, though the *vocês* pronoun in this data is not present in locations 25 and 11, where the Ø3PL occurs, but there is no occurrence of *vocês*. As, such we presume that *vocês* (or cognate) is indeed present in all of Portugal, at least as a V-form.

4.2 Clitic pronouns

The data for 2PL and 3PL clitics is sporadic. There are many gaps in the map due to lack of data.

Map 7 Distribution of 2ⁿᵈ and 3ʳᵈ person clitics for the semantic 2PL in Portugal
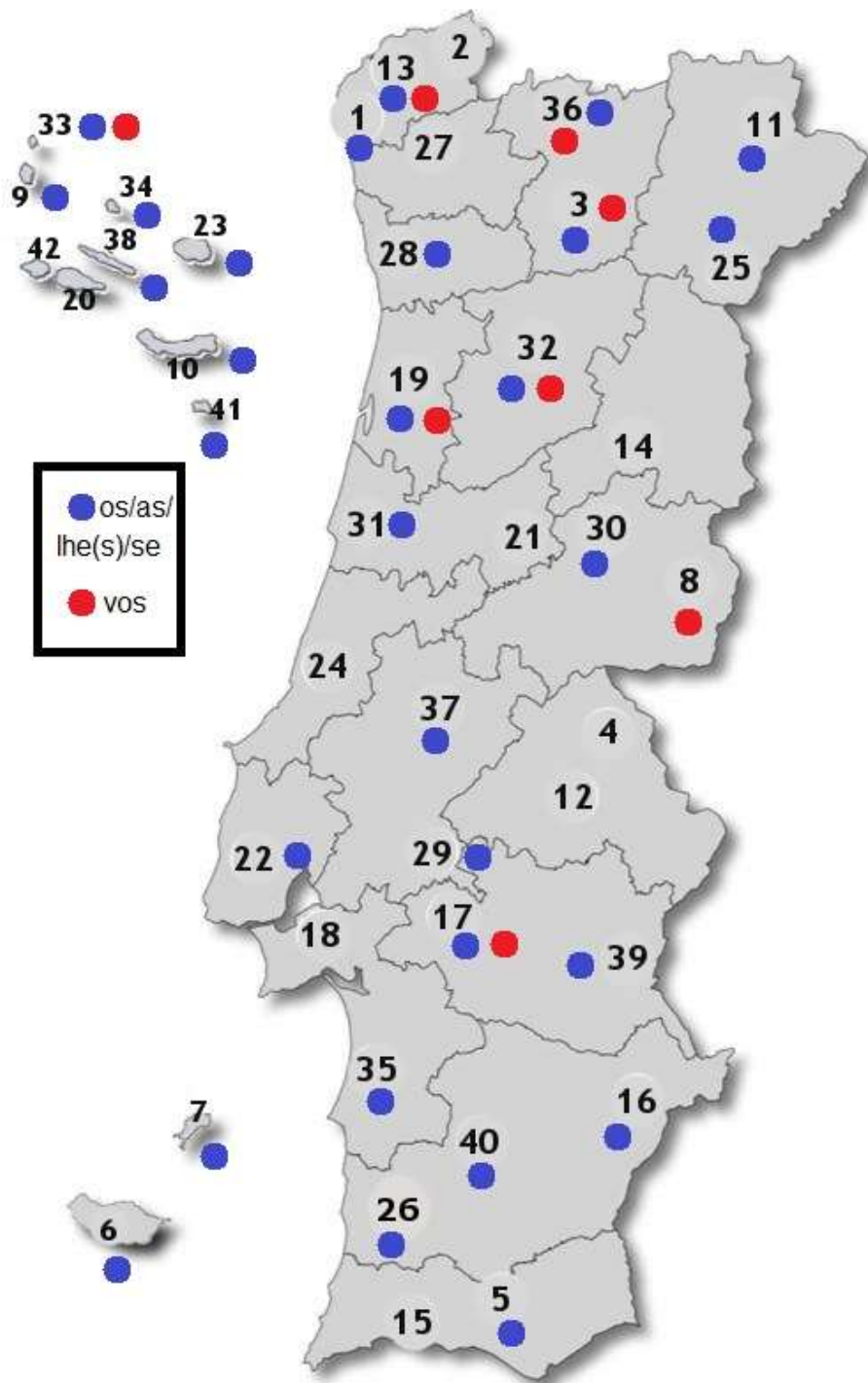
Locations with clitic pronouns

1. VPA Vila Praia de Âncora (Viana do Castelo)
3. PFT Perafita (Vila Real)
5. PAL Porches, Alte (Faro)
6. CLC Câmara de Lobos, Caniçal (Funchal)
7. PST Camacha, Tanque (Funchal)
8. MST Monsanto (Castelo Branco)
9. FLF Fajãzinha (Horta)
10. MIG Ponta Garça (Ponta Delgada)
11. OUT Outeiro (Bragança)
13. MIN Arcos de Valdevez, Bade, S. Lourenço da Montaria (Viana do Castelo)
16. SRP Serpa (Beja)
17. LVR Lavre (Évora)
19. COV Covo (Aveiro)
22. EXB Enxara do Bispo (Lisboa)
23. TRC Fontinhas (Angra-do-Heroísmo)
25. LAR Larinho (Bragança)
26. LUZ Luzianes (Beja)
28. GIA Gião (Porto)
29. STJ Santa Justa (Santarém)
30. UNS Unhais da Serra (Castelo Branco)
31. VPC Vila Pouca do Campo (Coimbra)
32. GRJ Granjal (Viseu)
33. CRV Corvo (Horta)
34. GRC Graciosa (Angra do Heroísmo)
35. MLD Melides (Setúbal)
36. STA Santo André (Vila Real)
37. MTV Montalvo (Santarém)
38. CLH Calheta (Angra do Heroísmo)
39. CPT Carrapatelo (Évora)
40. AJT Aljustrel (Beja)
41. STE Santo Espírito (Ponta Delgada)

Map 7 displays a somewhat incomplete distribution. At first glance, it appears that *vos* is well distributed in the North and near extinct elsewhere, only appearing in locations 17 (Lavre, Évora) and 33 (Corvo, Horta). Looking closer, occurrences of *vos* even in the North present many gaps; only 6 out of 15 potential locations manifested this clitic. Furthermore, locations 17 and 33 are probably not extraordinary holdouts of *vos*. In both locations, the use of *vos* (1 in LVR and 2 in CRV) is dative. As previously mentioned, *vos* is a common clitic in the spoken norm, and Map 2, from the field work conducted by Lara Bermejo (2015) provides a better picture of the clitics. The CORDIAL-SIN data does reaffirm his map, with the only examples of *vos* outside of the North being datives.

From our extended data for Monsanto, we are able to confirm Castelo Branco as being unaffected by the *vocês* innovation so far. In the extra data gathered from ALEPG recordings, *vos* was present in this location as a reflexive, accusative, and dative, which is also what Lara Bermejo (2015) found for two other localities in Castelo Branco, Covilhã and Fundão, in that study's field work.

Despite its gaps, the CORDIAL-SIN data does show a higher productivity of *vos* in the North than in the Center-South. Because most examples are in the North, we can infer that the 2PL clitic has greater productivity there than the Center-South, which uses more 3PL clitics, which corrobates Lara Bermejo (2015).

Looking at the data in Table 9, we see that no location reaches a double digit number of clitics; the most productive location is 17, with 8 clitics. Thusly, we do not undertake any numerical analysis of clitics for any specific location because the data is too scant. Analyzing the total numbers, the global percentage of 3PL clitics is 81.7% (85/104) versus 18.3% (19/104) for the 2PL clitic. This number shows the 2PL clitics as pronouns with diminishing use in the Portuguese territory, at the very least in rural dialects. The 3PL clitic used as a V-form, accompanying the formal nominal addresses directed towards the interviewers, may skew this data somewhat.

Table 9 Semantically 2PL 2nd and 3rd person clitics in Portugal

Locale	os/as/ lhe(s)/se	vos	Location	os/as/ lhe(s)/se	vos
1	5	0	26	7	0
3	4	2	28	1	0
5	1	0	29	5	0
6	2	0	30	1	0
7	1	0	31	1	0
8	0	4	32	2	3
9	2	0	33	2	2
10	1	0	34	1	0
11	6	0	35	3	0
13	3	1	36	7	1
16	1	0	37	5	0
17	7	1	38	1	0
19	2	5	39	4	0
22	1	0	40	5	0
23	1	0	41	2	0
25	1	0	Total	85	19

4.3 Analysis of the insular dialects

For most locations, this present study does not closely analyze the use of *vocês* as a V-form versus a T-form, considering that this is done with great attention in Lara Bermejo (2015). However, that study does not include the insular dialects of the Azores and Madeira archipelagos, only continental dialects. As such, we will take a closer look at the *vocês* innovation in these locations.

There is one location outside of the North which retains the 2PL, 38 (Calheta, Angra do Heroísmo). There are two possible explanations for this. One is that the examples in question are reported speech from an old story, not a part of the speaker's grammar. All 3 examples are reported speech from a speaker telling a story, and only represent one sentence, which was repeated with a

small modification as the speaker recalls the story. The second possibility is that the linguistic

isolation of the dialect on the island of São Jorge has conserved the use of 2PL. This is possible, and

the lexical analysis of Brissos, Gillier, & Saramago (2017) shows that the dialects of the Azores are

first and foremost conditioned by the island they are located on, the natural geographical isolation

of an island leading to more unpredictable variation even between islands that are in close

proximity. We consider the retention of *vós* somewhat less likely, however, because of the context

of the occurrences of *vós* and the presence of the *vocês* innovation on other islands.

Analyzing the subjects, there are occurrences of *vocês* as a T-form 10, 20, 33, and 41,

representing the islands of São Miguel, Pico, Corvo, and Santa Maria. Corvo and Santa Maria in

particular are the respective northwest and southeast extremes of the Azores archipelago,

suggesting that the innovation is widespread. The island of Corvo (33, CRV) provides us with an

example of mixed agreement with a 2PL dative.

(25) *Não, a gente não* **vos** *dá que a gente levamos muita pressa*

No, the people no you.DAT.2PL give that the people carry much hurry

e **vocês** *não* **têm** *pratos para* **vos** *deitar a comida.*

and you.NOM.2PL no have.PRES.IND.3PL plates for you.DAT.2PL lay the food.

"No, we won't give you any because we're in a hurry and you don't have plates to put

the food on." (CRV30)

Since the *vocês* innovation is present here, the smallest and most remote island, it is likely

present throughout the entire archipelago. With this in mind, we consider the 2PL subject and verb

in CLH to be an archaicism, and it is not paid much attention in this study. The data is included in the

tables and totals of this study, but any calculation of data which is for the North excludes it.

For the Madeira archipelago, we unfortunately have only one occurrence of *vocês* at our

disposal in location 7, PST, on the island of Porto Santo, which is much smaller than the main island,

Madeira, and provides this example:

(26) *Olha,* **vocês** *que* **levem** *tudo!*

Look.IMP.2SG, you.NOM.2PL that carry.PRES.SUBJ.3PL all!

"Look, may you guys take it all!" (PST24)

The subject *vocês* in this context appears to refer to strangers, but the context of the story suggests that it is a T-form. The informant is talking about an old windmill that he used to run. When it stopped functioning, people broke in to take things, after which a man came with an offer to purchase the millstone, and received the response above, which we are interpreting as a non-respectful *vocês*. This minimal data and the absence of *vós* and the 2PL leads us to believe that the *vocês* innovation has spread to the Madeira archipelago as well.

Though the clitic data from these two regions is scarce, making it hard to propose which phase of the innovation these locations are, let us take note of Lara Bermejo's (2015) observations for the Lisbon area. It is the only area which has resisted changes in the clitic pronouns due to its status as an educational, economic, and political center, subjecting it to greater normative pressure. It has completed the replacement of subjects, verbs, and reflexive and not progressed to the next stage. Since these factors do not apply to either of the insular regions in the Atlantic, we propose that they are in the third of our five proposed stages. We have no data for accusatives in the Azores, and of all the clitics shown in Map 7, only two locations display clitics used as a T-form: two occurrences of the 2PL *vos* dative in location 33, and one occurrence of the 3PL *se* reflexive in location 41. As such, the *vocês* innovation has replaced the 2PL subject, reflexive, and verb, which is clear from our data, and we additionally propose that accusative *vos* is undergoing replacement, but this is less clear. As such, our proposal is that the Azores are minimally in the third phase of the *vocês* innovation and could be in a more advanced stage. A more detailed study of all nine islands is necessary to confirm this and to verify the phase of the innovation on each island.

## 4.4 The state of 2PL pronouns and conjugations in the North in the CORDIAL-SIN

Not all the locations in the North showed evidence of the 2PL, for distinct reasons. Location 27 (Fiscal, Braga) provided no data whatsoever. Location 28 (Gião, Porto) only displays 3PL forms, which could be due to insufficient data (only 6 conjugated verbs and 3 subject pronouns), but could

also be the legitimate extinction of the 2PL forms in the area due to normative pressure. The use of *vocês* in this locale is as a V-form, directed towards the interviewers, and this suggests that the corresponding T-form could be *vós*. However, Porto is Portugal's second largest metropolitan area and the largest center of commerce in the North, and therefore experiences greater contact with the spoken norm, which would increase the normative pressure of the innovative form. The *vocês* innovation was probably present at the time of these recording roughly 30 years ago because Segura (2013) shows examples of *vocês* + 2PL in the district of Braga with occurrences from the ALEPG. Aguiar & Paiva (2017) show it in Braga as well, and seeing as the phenomenon spreads in waves, it would have to first pass through the district of Porto[5]. In the case of 11 (Outeiro, Bragança), there is insufficient data. There are 6 Ø3PL conjugations, no pronouns, and many occurrences of the nominal address in this locale. Due to its geographical isolation in far northeastern Portugal, the local dialect should be conservative and retain the 2PL. Additional research is necessary to confirm this.

Location 19, however, clearly retains the 2PL conjugation but not the subject pronoun. In light of all our other evidence, this furthers the idea that the many coastal dialects of the North have begun the first phase, corresponding to the loss of *vós*, while the varieties further inland have not. This area's denser population, larger cities, and greater amount of commerce cause more linguistic contact than the inland North.

---

[5] Lara Bermejo (2012, 2015, 2018a) demonstrates that the spread of the *vós/vocês* and *vosotros/Ustedes* innovation follows an epidemiological pattern of waves, spreading outward from the origin of the innovation to the geographically adjacent areas. This is why the *vocês* innovation is most advanced in Beja, its origin, and is incipient in the North of Portugal, where it has arrived more recently after having gradually spread northward over time.

Map 8 Occurrences of *vós* in Portugal (excluding the Azores)
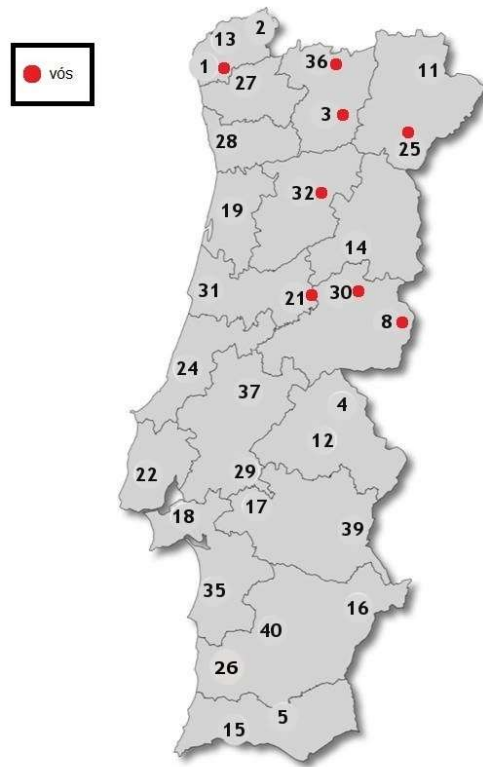
Map 9 Hypothetical *vós* distribution in the districts of Portugal
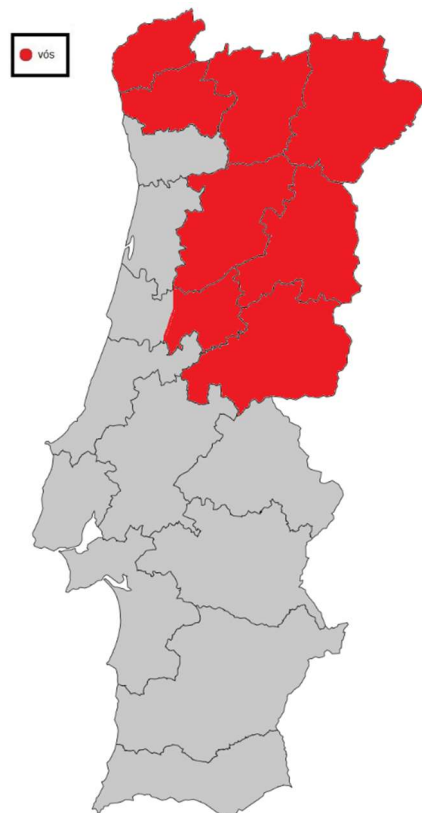
Taking into consideration Segura (2013), we see occurrences of the subject pronoun *vós* in the district of Braga, and it is included in Map 9. The district of Guarda is included in this map because it is surrounded by districts with occurrences of *vós*, deductively, it ought to retain the pronoun. The district of Porto did not present any evidence for *vós*, but it is necessary to consult all the locations in the ALEPG to confirm this.

Remembering that Aguiar & Paiva (2017) showed only 4.7% use of the *vós* subject pronoun in Braga, let us consider this idea as we look at a location that has removed *vós*.

4.4.1   *Vocês* + 2PL

Location 19 (Covo, Aveiro) displays a robust set of data, seen in Table 10. It has no *vós* and displays 33 instances of the pronoun *vocês*, 31 with conjugated verbs and 2 without. There are 34 conjugated verbs with these *vocês*, which is because an overt pronoun occasionally controls multiple verbs; see example (29) below. Of these verbs, 29.4% are 2PL, and 70.6% are 3PL. By comparing conjugations alone, Covo presents us with 40.4% for 2PL versus 58.4% for 3PL. If we assume that *vós* is extinct and all Ø2PL therefore have an implicit *vocês*, then we can interpret Covo to have successfully completed the first phase of the model (17) that was presented at the end of Chapter 2, eliminating the subject pronoun.

Table 10 Semantically 2PL verbs and subjects in Covo, Aveiro

Location	*Vocês* + conjugated verb		Null	
	2PL	3PL	2PL	3PL
19	18.2% 10/55	43.6% (24/55)	22.2% (12/55)	14.8% (9/55)

Covo stands out as the only locale in the CORDIAL-SIN with the mixed *vocês* + 2PL phenomenon; nonetheless, its high number of occurrences gives us 10/155 (6.5%) of non-null *vocês* conjugations for the whole Portuguese territory and a rather significant 10/65 (15.4%) for the North, where 2PL conjugations occur. These 10 examples make for 14.5% of all 2PL verbs (including our 3 examples from CLH, São Jorge, Azores).

Additionally, of the 12 Ø2PL verbs, 8 are grammatically null but discursively mixed due to an explicit subject pronoun elsewhere in a nearby clause. Including these 8, then, 26.1% of 2PL verbs are discursively mixed with *vocês*. We will not go beyond this and assume the other 4 Ø2PL verbs are discursively connected to *vocês* because this would become too speculative, especially when applied to other locations with unclear data. The below sentences are examples of discursively mixed agreement. Non-null *vocês* + 2PL are in bold, and *vocês* that are discursively connected to a Ø2PL are highlighted in black.

(27) *Olhai   lá,   que é que* **vocês**   *andam   a fazer?!*

Look.IMP.2PL there, what is that you.NOM.2PL go.PRES.IND.3PL at do?!

"Look here, what is that you guys are going around doing?" (COV01)

(28) **Vocês   andais   *a gastar dinheiro...***

You.NOM.2PL go.PRES.IND.2PL at spend money…

"You guys go around spending money…" (COV01)

(29) *E então assim,* **deixai**-*me   morrer e* **deixai**   *morrer a velha e depois* **vocês**

And then so, let.IMP.2PL-me die and let.IMP.2PL die the old and after you.NOM.2PL

***dai*-o,   **vendei**-*o,* **dai**-*o   a quem vocês*

give.IMP.2PL-it, sell.IMP.2PL-it, give.IMP.2PL-it to whom you.NOM.2PL

quiserem   porque nada   disso me incomoda.

want.FUT.SUBJ.3PL because nothing of this me bothers.

"And like that, let me die and let the old lady die and after give it, sell it, give it to whomever you guys want because none of that bothers me." (COV02)

(30) *Ó meus amigos, agora* **escolhei**   *o   que* **vocês**   *quiserem.*

Hey my friends, now choose.IMP.2PL that which you.NOM.2PL want.FUT.SUBJ.3PL.

"Look my friends, choose what you want now." (COV10)

(31) *Se* **quereis**   *à cega   a tirar uns bilhetes assim, e se não quiserem,*

If want.PRES.IND.2PL at the blind at pull some tickets so,   and if no want.FUT.SUBJ.3PL,

vocês **escolhei,** **'lei'** *as folhas como vocês*

you.NOM.2PL choose.IMP.2PL, read.IMP.2PL the sheets like   you.NOM.2PL

quiserem, *diante do advogado, e* **vocês** **escolhei,** *e eu*

want.FUT.SUBJ.3PL, in front of the lawyer, and you.NOM.2PL choose.IMP.2PL, and I

fico com aquilo que vocês *não quiserem".*

stay with that that you.NOM.2PL no   want.FUT.SUBJ.3PL.

"If you guys want, blindly drawing lots like that, and if you don't want that, choose, read

the papers as you like, in front of the lawyer, and choose, and I will keep what you don't

want." (COV10)

(32) *E se* **vocês** *não quiserem,* **ide** *falar com o padre.*

And if you.NOM.2PL no want.FUT.SUBJ.3PL, go.IMP.2PL speak with the father.

"And if you guys don't want to, go speak with the priest." (COV13)

(33) *E se* **vocês** *puderem* *remediar (...) sem trabalhar, sem agricultura,*

And if you.NOM.2PL can.FUT.SUBJ.3PL remedy… without work, without agriculture,

fazeis *bem.*

do.PRES.IND.2PL well.

"And if you guys can fix it without working, without agriculture, you're doing well"

(COV16)

(34) *Mas se* **vocês** *logo se virem naufragados,*

But if you.NOM.2PL then you.REFL.3PL see.FUT.SUBJ.3PL shipwrecked,

agarrai-*vos* *à terra."*

grab.IMP.2PL-you.REFL.2PL to the land.

"But if you see yourselves shipwrecked, cling to the land." (COV16)

(35) **Vocês,** **lembrai**-*vos.*

You.NOM.2PL, rembember.IMP.2PL-you.REFL.2PL

"You guys, remember." (COV19)

(36) **_Vocês_**   **_'sendes'_**   _uns   burros!_

You.NOM.2PL be.PRES.IND.2PL some donkeys!

"You guys are dumbasses!" (COV29)

In these examples we see numerous occurrences of _vocês_ directly controlling a 2PL verb. In (29), (31), and (35) there is even an explicit _vocês_ with an imperative 2PL verb, despite an imperative conjugation never requiring a non-null subject pronoun. This clear mixed agreement and complete absence of _vós_ strongly indicates the elimination of the 2PL subject pronoun, despite the prominent 40.4% use of 2PL conjugations. This supports our refined version of Lara Bermejo's (2015) model, which was proposed in 2.3.2 – the subject pronoun disappeared first in Covo.

While Granjal, Viseu (location 32) did not provide examples of mixed agreement, it was found in Segura (2013), shown in (16) in section 2.3.2. Furthermore, Granjal provided this juxtaposition of occurrences from one speaker in a piece of dialog:

(37) _Eu às vezes digo para os meus:_ "**_Vocês_**   **_haviam_**   _de ser criados..."_

I at the times say to the mine: "You.NOM.2PL have.PRET.IMP.IND.3PL of be raised…"

"I sometimes say to mine: 'You kids had to have been raised…'" (GRJ68)

(38) "_Ai, meus filhos, meus filhos! Se_ **_vós_**   **_passassens_**   _o que_

Ah,   my sons,   my sons! If you.NOM.2PL pass.PRET.IMP.SUBJ.2PL what

nós passámos, **_comiens_**   _até pedras!_"

we passed, eat.PRET.IMP.IND.2PL even rocks.

"Oh, my children, my children! If you'd gone through what we went through, you'd eat

rocks!" (GRJ68)

Here we can see the _vocês_ innovation in Viseu, though not in an explicitly mixed agreement. This speaker begins to address her grandchildren, addressees who would receive the most informal treatment, as _vocês_. She then alternates to the classic _vós_ at the end of the anecdote. With the examples from Segura (2013), we can see that mixed agreement is a possibility in Viseu as well, as well as in Braga, though the CORDIAL-SIN provides no data for the latter district.

To further analyze subject pronouns, we shift our view to global percentages.

Table 11 Occurrence of *vocês* and *vós* (subject pronoun, oblique position, and otherwise)

Location	Vocês	Vós	Location	Vocês	Vós
1	87.5% (7/8)	12.5% (1/8)	24	100% (1/1)	
2	100% (1/1)		25		100% (2/2)
3	88.9% (8/9)	11.1% (1/9)	26	100% (4/4)	
4	100% (1/1)		28	100% (3/3)	
5	100% (4/4)		29	100% (45/45)	
7	100% (1/1)		30	66.7% (2/3)	33.3% (1/3)
8		100% (2/2)	31	100% (2/2)	
10	100% (2/2)		32	44.4% (4/9)	55.6% (5/9)
12	100% (10/10)		33	100% (1/1)	
13	100% (3/3)		35	100% (6/6)	
14	100% (3/3)		36	88.9% (8/9)	11.1% (1/9)
15	100% (1/1)		37	100% (13/13)	
17	100% (1/1)		38		100% (3/3)
19	100% (33/33)		39	100% (3/3)	
20	100% (4/4)		40	100% (6/6)	
21	75% (3/4)	25% (1/4)	41	100% (2/2)	
22	100% (3/3)		Total	91.6% (185/202)	8.4% (17/202)

In Table 11, most locations show a majority for *vocês*. For four locations, *vós* is the dominant pronoun, but in three of these, it is the only pronoun present, and in small numbers, and in one location, 32, it represents 5/8 pronouns. In total, *vocês* is vastly preferred in Portugal, representing 91.5% (184/201) of subject pronouns compared to *vós*'s 8.5% (17/201). Reducing our scope to the North, it is 18.9% (14/74). This percentage shows a resilient *vós* which has yet to disappear but is losing space to *vocês*.

This information pertains to the *vós* pronoun alone, so to see further evidence of its decline,

we must compare the null and non-null subject conjugations.

Table 12 Conjugated verbs grouped by controlling subject pronoun

Location	Vós	Vocês	Ø2PL	Ø3PL
1	12.5% (1/8)	50% (4/8)	0	37.5% (3/8)
2	0	7.7% (1/13)	23.1% (3/13)	69.2% (9/13)
3	5.9% (1/17)	47.1% (8/17)	11.8% (2/17)	35.3% (6/17)
4	0	0	0	100.0% (1/1)
5	0	27.3% (3/11)	0	72.7% (8/11)
6	0	0	0	100.0% (1/1)
7	0	25.0% (1/4)	0	75.0% (3/4)
8	16.7% (2/12)	0	50.0% (6/12)	33.3% (4/12)
9	0	0	0	100.0% (2/2)
10	0	50.0% (2/4)	0	50.0% (2/4)
11	0	0	0	100.0% (6/6)
12	0	53.3% (8/15)	0	46.7% (7/15)
13	0	33.3% (1/3)	66.7% (2/3)	0
14	0	20.0% (2/10)	10.0% (1/10)	70.0% (7/10)
15	0	16.7% (1/6)	0	83.3% (5/6)
16	0	0	0	100.0% (1/1)
17	0	10.0% (1/10)	0	90.0% (9/10)
18	0	0	0	100.0% (1/1)
19	0	63.0% (34/55) (10 occurrences of 2PL verbs)	22.2% (12/55)	14.8% (9/55)
20	0	40.0% (2/5)	0	60.0% (3/5)
21	25.0% (1/4)	50.0% (2/4)	0	25.0% (1/4)
22	0	75.0% (3/4)	0	25.0% (1/4)
23	0	0	0	100.0% (5/5)
24	0	25.0% (1/4)	0	75.0% (3/4)
25	80.0% (2/5)	0	40.0% (2/5)	20.0% (1/5)

26	0	83.3% (5/6)	0	16.7% (1/6)
28	0	50.0% (3/6)	0	50.0% (3/6)
29	0	50.0% (36/72)	0	50.0% (36/72)
30	11.1% (1/9)	0	33.3% (3/9)	55.6% (5/9)
31	0	25.0% (2/8)	0	75.0% (6/8)
32	12% (3/25)	16% (4/25)	48% (12/25)	24% (6/25)
33	0	20.0% (1/5)	0	80.0% (4/5)
34	0	0	0	100.0% (1/1)
35	0	27.8% (5/18)	0	72.2% (13/18)
36	4.5% (1/22)	31.8% (7/22)	4.5% (1/22)	59.1% (13/22)
37	0	24.3% (9/37)	0	75.7% (28/37)
38	100.0% (3/3)	0	0	0
39	0	11.8% (2/17)	0	88.2% (15/17)
40	0	31.6% (6/19)	0	68.4% (13/19)
41	0	28.6% (2/7)	0	71.4% (5/7)
Total	3.2% (15/462)	33.8% (156/462)	9.5% (44/462)	53.5% (247/462)

In Table 12, it is revealed that *vós* + 2PL is the most uncommon address possibility of the four, representing only 3.3% of verbs and only 15 occurrences. Ø2PL also has a minor percentage – 9.5%. This is expected – *vós* is only used in parts of Portugal while *vocês* is ubiquitous. Looking at the North alone, the panorama changes.

Table 13 Conjugated verbs grouped by controlling subject for just the North

Vós	*Vocês* + **3PL**	*Vocês* + **2PL**	**Ø2PL**	**Ø3PL**
6.2% (12/195)	28.7% (56/195)	5.1% (10/195)	22.6% (44/195)	37.4% (73/195)

We can see that the null subject is the predominant subject for either pronoun, but the relative percentage of *vós* is significantly smaller, only constituting 18.2% of all 2PL subject positions, while overt *vocês* occupies 43.4% of 3PL subject positions. Because *vocês* replaces *vós* as the subject for 10 2PL verbs in location 19, Covo, *vocês* as the subject pronoun represents 33.8% (66/195) of all subject positions in the North. With only 12 examples, non-null *vós* is barely more frequent than

vocês + 2PL, with 10 examples; this is to say that the *vocês* innovation is roughly equally distributed with the historical *vós* for non-null 2PL verbal subjects. So, while the 2PL verbal morphemes constitute 33.8% of conjugations, *vós* only constitutes 15.4% of non-null subjects.

Since Granjal (location 32) also shows itself to be a location beginning to shift, let us look at its data. Granjal shows a stronger resistance to *vocês* than Covo – 60% 2PL verbs versus 40% 3PL verbs. Use of null subjects is much higher here, but in this case "much higher" is 72%, which is within the range of other studies, and Covo presented a significant number of overt subjects with an imperative, a potentially dialectal feature that led to fewer null subjects. Granjal appears to have commenced phase 1 of the *vocês* innovation (whereas Covo has concluded it) and it has not affected the use of null subjects in this location.

In light of this data, there is a strong argument for an ongoing elimination of *vós*, which corroborates the findings of Aguiar & Paiva (2017) for Braga, in which *vós* constitutes only 4.3% of subjects, despite Ø2PL subjects having an abundant 45.7%. What's more, the two occurrences of *vós* as a subject pronoun correspond to women in their 60s with a low level of education, the latter two factors lending to linguistic conservatism. The CORDIAL-SIN does not provide data for young speakers, and the informants for the corpus are older, rural speakers with low levels of education. This means that our number 6.2% may skew towards an over-representation of overt *vós*. Overall, this data strengthens the argument that *vós* is being eliminated in northern Portuguese dialects. Taking into account our data from the CORDIAL-SIN and the relevant examples (see (13)-(16) in section 2.3.2) from the ALEPG seen in Segura (2013), we propose a new map, Map 10, that includes the insular dialects and shows 5 phases of transition.

Map 10 The *vocês* innovation in all of Portugal according to the CORDIAL-SIN and ALEPG
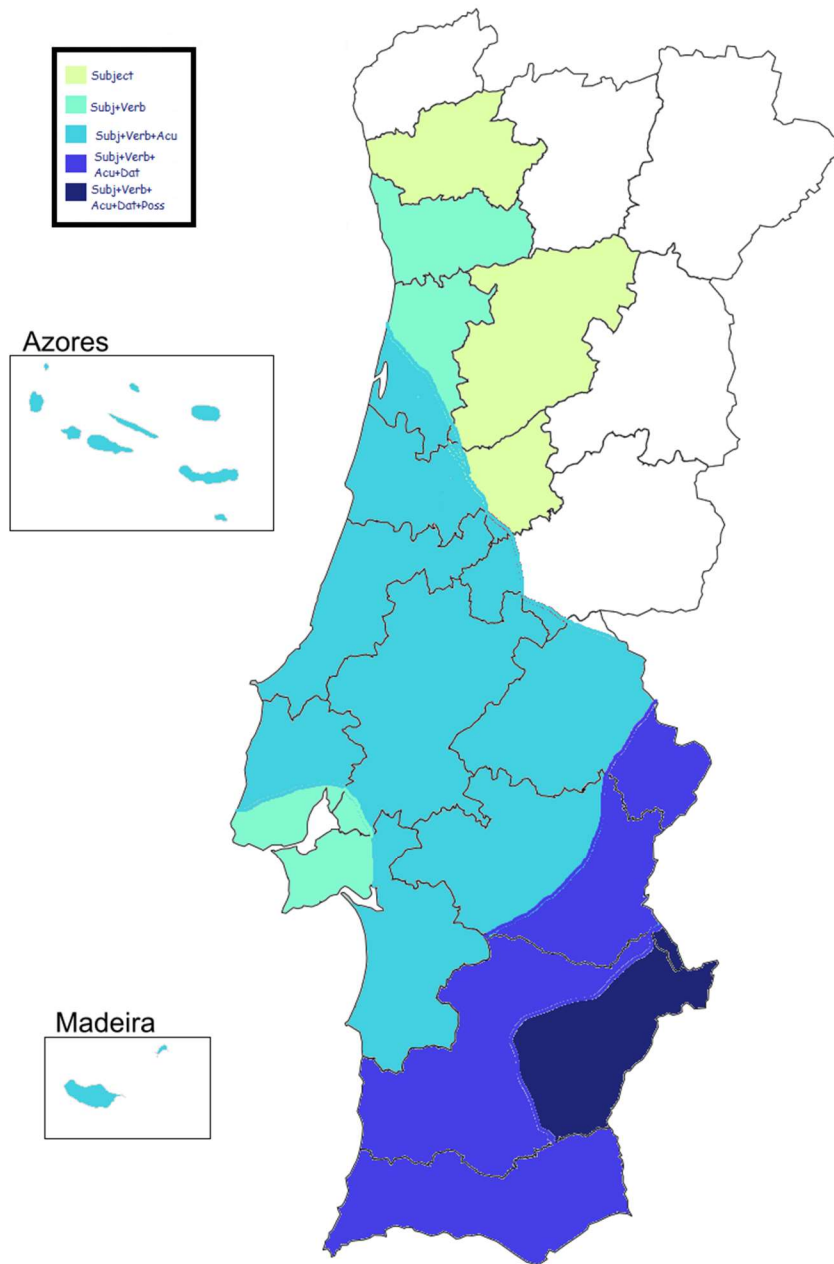
4.4.3 Implications for the NSP

As discussed previously, EP is considered to be firmly [+ NSL], as we saw in Duarte (1995),

shown in Table 7, in section 3.2. Furthermore, this was the finding in Sória's (2013) study of the

CORDIAL-SIN for the 1PL. Her study has an abundance of data in comparison with this study, due to

the discursive context of an interview leading to the informants speaking in the 1PL with high

frequency, and it showed that preference for *nós* or *a gente* in any given locale did not make a significant difference in use of Ø1PL. As such, EP remained notedly [+ NSL].

Looking back at Table 12 in the previous section, the global preference for the 2PL is for a null subject. When we divide between the North and the rest of Portugal, the North displays 60% (117/195) null subjects versus 65.2% (174/276) for all other locales. This difference of 5.2% is relatively small and both numbers fall within the range of null subjects we saw in Duarte (1995) for EP. This confirms our hypothesis that the reduction of from a six person verbal paradigm to a five person verbal paradigm will not alter the NSP in EP (nor any other Romance language), with dialects retaining the 2PL conjugations showing similar global percentages of null subjects to those that have eliminated it. If anything, this data shows a potentially higher use of the null subject pronoun in areas with a five conjugation paradigm, though this is probably due to variation in the samples.

Interestingly, Covo, which is analyzed thoroughly here, shows only 38.2% (21/55) null subject usage, much lower than the average. The occurrences from Covo have many overt subjects with commands, which is uncommon in normative Portuguese, and could be a peculiarity of the location. Even if Covo does indeed have high amounts of overt subjects, this does not mean that it has become or is becoming [– NSL]; a shift to [– NSL] requires a change in grammatical properties (as were mentioned in Chapter 3), not simply fewer occurrences of null subjects. Furthermore, Lobo (2016) points out that overt subjects can be employed for such strategies as emphasis, focus, or disambiguation and since locations like Covo do maintain null subjects in positions that are not coreferential to previous overt subjects, more overt subjects do not necessitate a change in the null subject property. Overall, the North presents itself as being clearly [+ NSL].

4.5 The case for crystallized 2PL forms

Another facet of Aguiar & Paiva's (2017) study was the prominence of irregular verbs in 2PL conjugations. 84.2% of irregular verbs were Ø2PL. While it is not entirely clear what the criteria were for an irregular verb, our criterion was an irregular stem in the conjugation. For example, *querer* would be a regular verb in the present indicative – *quero, queres, quer* – despite the irregular 3SG

quer because there are no changes to the stem, but would be irregular in the future subjunctive –

quiser, quiseres, quiser – due to the stem change. This study did not find such high indices of 2PL for

irregular verbs in the North – only 22.6% (19/84) of irregular verbs were 2PL conjugations. This does

not counter Aguiar & Paiva's (2017) observation that verbs with irregular stems are the most

common 2PL conjugations in Braga, but it does put it into a broader context. This may be the case

for other locations in the North that we do not have sufficient data to analyze. Alternatively, it could

be a feature of Braga speech specifically – northern Portuguese dialects are not a monolith and

Braga may preserve these certain 2PL verbs due to linguistic variation.

What our study reveals is that 2PL conjugations appear to favor certain verb paradigms.

Most of all, the 2PL imperative is used. This is the only case in which the 2PL conjugations comprise

the majority, constituting 72.3% of commands. The interjective nature of a command perhaps

affords it some resistance to change. Commands, unlike most verb tenses, never require an overt

subject in any context (though they can take one), and they are relatively common in speech. Below

in Figure 1, we see that the imperative is the second most common in total, behind the more

common present indicative. The present indicative is the second most common 2PL conjugation, but

its relative percentage is much lower than that of the 3PL, only 24.1%. For the imperfect preterit,

2PL verbs constitutes 50% of occurrences, while in the perfect preterit, it only constitutes 14.3%. For

the future indicative and the present and preterit subjunctive there is very little data.

Figure 1 Semantically 2PL verb conjugations in the North
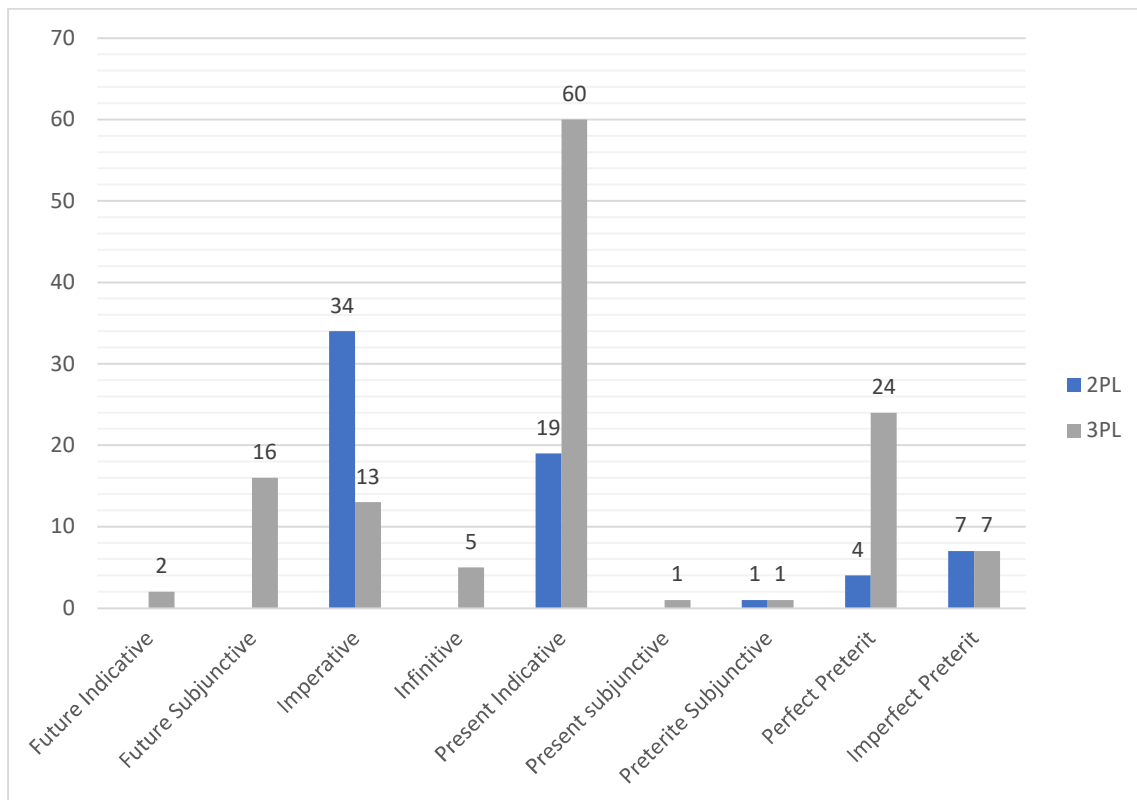
The future subjunctive presents 16 3PL examples and 0 2PL examples in Figure 1. This conjugation appears to have consolidated entirely into the 3PL. The clearest implications from the data in Figure 1 is that the 2PL imperative is the most resilient 2PL conjugation. All past and present indicative conjugations also manifest 2PL forms, but appear to be losing space to 3PL forms. The subjunctive and infinitive moods show 24 examples, but only present one 2PL conjugation in the preterit subjunctive. We did not find evidence for crystallized irregular forms, but the relative crystallization of the 2PL imperative was notable.

Since other locales do not present robust enough data for a more meticulous analysis, we will only examine our most robust locale, Covo, where we can look at this trend fairly clearly, as well as the imperative and indicative conjugations, all found in Figure 2.

Looking back at examples (27)-(36) again in section 4.4.1, all the future subjunctive verbs are conjugated in the 3PL, often in contrast with other imperative and present indicative 2PL verbs in

the same sentence, such as in (31), repeated below. In this sentence, the commands are 2PL, the

present tense *quereis* is 2PL, and the future subjunctive *quiserem* is 3PL all three times that it occurs.

(31) "*Se **quereis** à cega a tirar uns bilhetes assim, e se não **quiserem**,*

If want.PRES.IND.2PL at the blind at pull some tickets so, and if no want.FUT.SUBJ.3PL,

*vocês **escolhei**, 'lei' as folhas como vocês*

you.NOM.2PL choose.IMP.2PL, read.IMP.2PL the sheets like you.NOM.2PL

***quiserem**, diante do advogado, e vocês **escolhei**, e eu*

want.FUT.SUBJ.3PL, in front of the lawyer, and you.NOM.2PL choose.IMP.2PL, and I

*fico com aquilo que vocês não **quiserem**".*

stay with that that you.NOM.2PL no want.FUT.SUBJ.3PL.

"If you guys want, blindly drawing lots like that, and if you don't want that, choose, read

the papers as you like, in front of the lawyer, and choose, and I will keep what you don't

want." (COV10)

The evidence shows a marked preference for 3PL in the future subjunctive in Covo, with a

total of 10 examples, 0 2PL examples, and a clear preference to use this conjugation even in

sentences which use 2PL conjugations elsewhere. The 2PL future subjunctive is likely extinct in this

location.

For the 3PL perfect preterit, it is not as clear if it has completely overtaken the 2PL preterit in

Covo. All 7/7 of this conjugation are 3PL, but there are no such juxtapositions of 2PL and 3PL to

provide further insight, like with the future subjunctive. Given that there are 7 3PL examples in Covo,

and 7 2PL examples in the rest of the North, the two forms are probably in competition in this

location and in the North at large, much like the more abundant present indicative, with 26.7% use

in Covo.

Figure 2 Semantically 2PL verb conjugations in Covo, Aveiro
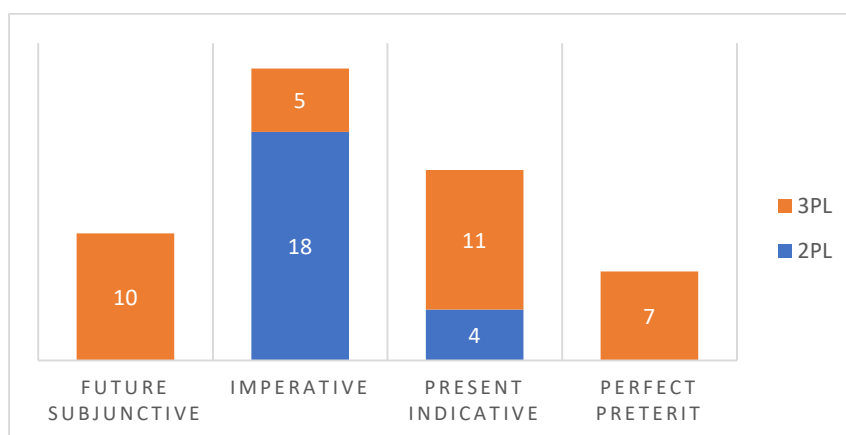
This information pertaining to the 2PL imperative creates an interesting parallel with the 2SG imperative in BP. As to be detailed in 5.3.2, the 2SG command has persisted in BP, outlasting all other 2SG verbs, even in dialects with no other 2SG morphology. In our data, there was no evidence for 2PL commands lingering in areas that had otherwise completed phase 2, but the similarity supports the idea that the affirmative latinate 2nd person commands are the most resilient verbal form. In light of our data, we can propose a verbal replacement model based on mood.

(39) Subjunctive/Conditional/Personal Infinitive > Indicative > Imperative

This model, curiously, stands in contrast to the model for West Andalusian Spanish, in which Lara Bermejo (2015) shows this dialect to adopt the 3PL imperative, specifically the affirmative imperative, first. There is a potential explanation for this difference. The form of imperative adopted in the Andalusian innovation is identical with the infinitive verb form in Spanish, a form used as an imperative far outside of western Andalusia. Lara Bermejo (2015) considers it to be a 3PL form because the default interpretation of infinitives in Spanish are in the 3rd person, which is to say, an infinitive like *sentar* "to seat" would receive the 3rd person *se* reflexive clitic for the reflexive infinitive form *sentarse* "to seat oneself". However, the clitic *se* can also be interpreted as an impersonal, corresponding to "oneself" in English, and as such, this form's lack of a true grammatical person is probably why it was the first verb to change in the dialect. An infinitive command, while possible, is significantly less common in EP, and as a result, we are faced with the subjunctive 3PL

command versus the classic 2PL command, with no such option for an imperative that could be interpreted as either 2PL or 3PL. Beyond this difference, our model shows the subjunctive to disappear prior to the indicative as well.

4.6 Summary

From our study of 2PL pronouns and verbs in the CORDIAL-SIN, we have been able to make several conclusions. The data from Covo, Aveiro, supports our model with five phases of transition – subject pronoun > reflexive/verb > accusative > dative > possessive – with phase 1 of the *vocês* innovation being the replacement of the subject pronoun *vós*, which is completed prior to the phase 2, the replacement of the 2PL reflexive clitic and verbal morphology. The data from Granjal, Viseu, shows a location in the first phase of replacement, in which *vós* and *vocês* coexist, while Covo shows the second phase, in which *vós* has been eliminated and 2PL and 3PL verbs and reflexives coexist. The data from the Azores and Madeira archipelagos has led us to propose that the insular dialects are at least in the third phase of replacement, in which the 2PL and 3PL accusative forms coexist, though there could be potential dialectal variation within these regions.

By analyzing verbal morphology, we found a hierarchy of replacement for verbs based on mood – subjunctive/conditional/personal infinitive > indicative > imperative, with the imperative being the most resilient 2PL mood.

Based on the data from our study, Lara Bermejo (2015), and supplemental data from Segura (2013), we have proposed the distribution of the *vocês* innovation seen in Map 10, showing five phases. For the North, we see that the innovation is taking hold in the coastal regions, while the more inland, northeastern regions remain unaffected by the innovation, conserving the traditional 2PL in its entirety.

Table 14 All 2nd person subjects, conjugations, and clitics

Location	Conjugated verbs			Pronouns				
	Vocês	Vós	Ø2PL	Ø3PL	Vocês	Vós	Vos	Lhe/se / os/as
1	4	1	0	3	7	1	0	5
2	1	0	3	9	1	0	0	1
3	8	1	2	6	8	1	2	4
4	0	0	0	1	1	0	0	0
5	3	0	0	8	4	0	0	1
6	0	0	0	1	0	0	0	2
7	1	0	0	3	1	0	0	1
8	0	2	6 (1 in 2SG)	4	0	2	4	0
9	0	0	0	2	0	0	0	2
10	2	0	0	2	2	0	0	1
11	0	0	0	6	0	0	0	6
12	8	0	0	7	10	0	0	0
13	1	0	2	0	3	0	1	3
14	2	0	1	7	3	0	0	0
15	1	0	0	5	1	0	0	0
16	0	0	0	1	0	0	0	1
17	1	0	0	9	1	0	1	7
18	0	0	0	1	0	0	0	0
19	34 (10 in 2PL)(only 31 *vocês* – some control multiple verbs	0	12	9	33	0	5	2
20	2	0	0	3	4	0	0	0
21	2	1	0	1	3	1	0	0
22	3	0	0	1	3	0	0	1
23	0	0	0	5	0	0	0	1

24	1	0	0	3	1	0	0	0
25	0	2	2	1	0	2	0	1
26	5 (only 4 *vocês* – 1 controls 2 verbs)	0	0	1	4	0	0	7
28	3	0	0	3	3	0	0	1
29	36	0	0	36	45	0	0	5
30	0	1	3	5	2	0	0	1
31	2	0	0	6	2	0	0	1
32	4	3	12	6	3	5	3	2
33	1	0	0	4	1	0	2	2
34	0	0	0	1	0	0	0	1
35	5	0	0	13	6	0	0	3
36	7	1	1	13	8	1	1	7
37	9	0	0	28	13	0	0	5
38	0	3	0	0	0	3	0	1
39	2	0	0	15	3	0	0	4
40	6	0	0	13	6	0	0	5
41	2	1	0	5	2	0	0	2

5 Comparisons with other innovations and mergers in the pronominal paradigm of Portuguese

In EP, another significant innovation in the pronominal paradigm is the semantically 1PL form *a gente*. Soria's (2013) study of *nós, a gente* and the 1PL null subject is an important companion and predecessor to this study, as it investigates the effects of the insertion of *a gente* in EP on the NSP and uses data from the CORDIAL-SIN in order to make conclusions about the insertion of *a gente* in BP. This thesis has taken the same question – that of pronoun insertion – and applied it to the older *vocês* innovation. A comparison of these two studies, their findings, and their implications is apt because *a gente* represents a potential reduction in the pronominal paradigm, and so does *vocês*. Both are innovative pronouns that often have a "mismatched" agreement due to a semantic person that is different from the grammatical person. *A gente* is a 1PL pronoun that should remove 1PL desinences in favor of the 3SG, and *vocês* removes the 2PL pronoun *vós* and its desinences in favor of the 3PL.

In the case of BP, there are two widespread mergers, which are *tu/você* and *vós/vocês*. Because our study has shown that the *vós/vocês* merger has not affected the null subject properties of EP, we expect that this merger (the first merger in BP) will also have not affected the null subject properties of BP. Based on our interpretation of Roberts' (1993) concept of functional richness, seen in model (24), repeated below, the 2SG merger of *tu/você* should not compromise the viability of null subjects in BP.

(24) A [+ NSL] verbal paradigm allows any number of double syncretisms, but no triple syncretisms.

As such, we closely examine this merger in this chapter to examine how this merger may have contributed to the loss of [+ NSL] in BP alongside other relevant factors that do not apply to EP.

5.1 Agreement with subject pronouns *nós*, *a gente*, *vós*, and *vocês* in European Portuguese

For the 1st person plural, *a gente* has the most agreement possibilities. Sória (2013) shows *a gente* to most commonly coincide with its grammatical person, the 3SG morpheme. *A gente* + 1PL, which is a semantic agreement, is less frequent but notable, and finally, *a gente* + 3PL is a rare agreement possibility. For the case of *nós*, it is mostly paired with its grammatical 1PL morpheme and only rarely is coupled with the 3SG morpheme. This same phenomenon holds for the 2PL; *vocês* + 3PL is the most common, and occurrence with the 2PL morpheme is less common. For *vós*, this study found no examples of *vós* + 3PL, e.g. **Vós estão errados!* In both studies, the results show that the introduction of an innovative pronoun does not alter the characteristic [+ NSL] for EP. For *a gente*, there was no relation found between high productivity of this subject pronoun and the Ø1PL. Similarly, the elimination of *vós* and 2PL verbs does not lead to a reduction in null subjects, but distinctly, the *vocês* innovation replaces all 2PL verbs in many dialects, while the *a gente* innovation does not.

A conclusion of Soria (2013) is that the replacement of a subject pronoun does not necessarily lead to the replacement of its corresponding conjugation. Even in locations with no attested *nós*, the Ø1PL is not eliminated. We have seen a similar result in Covo in particular, where the elimination of *vós* appears complete, but there is still considerable use of 2PL conjugations, both with *vocês* and with null subjects. The distinction it that in the case of *nós/a gente*, it is uncertain if the Ø1PL will eventually be replaced, as many dialects maintain Ø1PL verbs and 1PL conjugations matched with *a gente*. This means that it does not create any syncretism because it leaves five conjugations in place even when it replaces *nós*. For it to become a syncretism, it would have to replace Ø1PL conjugations as well. As such, the model (17) may not apply to *nós/a gente*. If *a gente* replaced the Ø1PL, it would constitute a triple syncretism in the subjunctive, conditional, and preterit imperfect conjugations with the 1SG and 3SG conjugations, but seeing that *a gente* does not accept a null subject unless the subject is accessible, it would disambiguate itself even if it fully

replaced 1PL verbs, allowing null subjects for the other subject pronouns. This would be similar to how the mandatory overt expresson of the 2SG *tu* in the Italian subjunctive allows 1SG *io* and 3SG *lui/lei* to continue be expressed as null subjects. This means that, according to our modification of Robert's (1993) hypothesis seen in (24), even a complete replacement of *nós* by *a gente* should not provoke a shift in the null subject properties of EP.

From a historical context, *vocês* is a much older innovation that has significantly replaced the traditional 2PL in most of Portugal. *A gente* is a newer form in Portugal, and it is uncertain the extent to which it will replace the 1PL. On the contrary, with *vós/vocês*, the replacement of the 2PL verbs appears to only be a matter of time and is following a clear model, repeated here.

(17) Subject pronoun > reflexive/verb > accusative > dative > possessive

A further difference between these two innovations is apparent in Lara Bermejo (2015). While the *vocês* innovation is found to follow a progressive, hierarchical model that expands in waves, *a gente* does not have clear phases. The region of the Algarve shows predominant use of the *a gente* subject, while neighboring Beja, the region in the final phase of the *vocês* innovation, shows conservative retention of *nós*. This is then inverted for dative *a gente* – the Algarve retains *nós* and Beja prefers *a gente*. In light of this, the *a gente* innovation does not appear to follow the model of the *vocês* innovation, and we can conclude that not all innovations in the pronominal paradigm follow the same rules – not even in the same language and dialect.

5.2 Portuguese in contact with Africa and America

In order to examine the syncretisms present in BP, namely the *tu/você* and *vós/vocês* mergers, let us step back to look at the broader context of the Portuguese language in Brazil. BP is highly distinct from EP because Portuguese is a global language, spread by empire, which has been and remains in contact with a variety of languages, most significantly from Africa and South America. Contemporaneously, Portuguese is pluricentric, meaning that it has different codified norms in different places, most notably in Portugal and Brazil. The concept of a norm is very significant to Portuguese, especially BP, because the norm is often in stark contrast with the reality of orality.

Despite its pluricentricity, which might lead to believe that Portuguese would have many acceptable varieties, the Brazilian and Portuguese norms are the two dominant norms.

The Brazilian norm is often times highly divergent from the daily reality of spoken, vernacular BP. Vernacular BP first arose in the circumstances of colonization, and is likely the result of what Lucchesi & Baxter (2009) refer to as irregular linguistic transmission. In the initial phases of colonization, there was extensive contact with Native Amerindian languages, as Silva (2013) shows. However, the population of Natives rapidly decreased from both intentional genocide or epidemics of European diseases. Inversely, the African impact on BP only rose with the arrival of more and more enslaved Africans well into the 19th century. After the official abolition of the transatlantic slave trade in 1830, enslaved Africans were internally moved from the Northeast to the Southeast, part of an economic shift in the country. In the time period of 1801-1850 (Mussa, 1991, p. 63), Europeans and Euro-Brazilians constituted 31% of the population, Africans, Blacks, and miscegenated Blacks constituted 65%, and integrated Natives a mere 4%. Black voices were then the main vehicle of the new Portuguese vernacular in Brazil.

So then, when we speak of prescribed, normative, or educated BP, we are looking at a Luso-descendent variety, most prominently spoken and written by the country's European rulers. The prescribed verbal model in pedagogical grammar books maintains all six persons to this day, although de facto regional norms largely maintain a four person model, having removed the 2nd person pronouns *tu* and *vós*. For vernacular BP however, today we see a three conjugation paradigm retaining 1SG for *eu*, 3SG for *tu/você, ele/ela*, and *a gente/nós*, and 3PL for *vocês* and *eles/elas*. An even more reduced two conjugation paradigm retains only the 1SG morphology with *eu*, 3SG morphology for all other persons, in addition to nominal syntagma that only mark the plural in the first word, which can be seen in example (40)-B. These are often strongly Afro-Brazilian descended dialects and sociolects of those with the lower socioeconomic status and education, these two being related; the continued lack of economic, social and therefore linguistic prestige afforded to Afro-

Brazilians is a part of slavery's racist legacy in Brazil, creating the divide for these two main varieties of BP (see Lucchesi, 1994, 1998, 2001, among others).

(40) A. *Estes meninos vão.* (Educated BP)

This-PL boy-PL go.PRES.IND.3PL

"These boys go."

B. *Estes menino vai.* (Vernacular BP)

This-PL boy-SG go.PRES.IND.3SG

"*These boy goes"

In regular circumstances, a language is passed generationally to children by their parents and community who acquired the language in the same fashion. In the irregular circumstances of Africans in contact with the Portuguese, however, Portuguese was first acquired as a second language due to necessity, often by a community of speakers of various unintelligible languages. The subsequent generation was then transmitted a language by a community of L2 speakers who had an incomplete knowledge of the language compared to a native speaker. This is why Lucchesi (1994, 1998, 2001) conceives of two major strains of BP – the educated, more Luso-descended norm of the urban elite and the popular, vernacular form born of irregular linguistic transmission in rural Afro-Brazilian communities.

When irregular transmission occurs in a context with a much larger number of alloglots than native speakers, it can birth a creole language with a restructured grammar that is morphologically simplified and grammatically distinct from Portuguese, the lexifier, which was the case on the islands of Cabo Verbe and São Tomé, for example. These circumstances exist on a continuum, however, and in a less extreme context when there is a greater input from the lexifier language, the result is not a creole, but the dialect will still receive creolizing characteristics, as Lucchesi (2001) puts. These are distinct linguistic features from the grammatical restructuring during contact, some transferred from

the substrate languages of the original alloglots.[6] Baxter (1997) argues this as well, noting that Afro-Brazilian verbal paradigms have creole-like characteristics. Taking a particular look at the well-presereved speech of Helvécia, Bahia, speakers here could produce (though not always) a completely leveled verbal paradigm, extending to even the 1SG, something uncommon in any BP dialect, with such utterances as:

(41) *Eu   fala.*

    I.NOM.1SG speak.PRES.IND.3SG

    "*I speaks." (Baxter, 1997, p. 268)

BP is not a creole,[7] but some highly notable features like the aforementioned reduced verbal paradigms, described by Lucchesi, Baxter, & Silva (2009), and a simplification of nominal syntagma to only require one plural marker, shown by Baxter (2009), result in sentences like example (40) and are the result of irregular linguistic transmission.[8] Both the two and three person verbal paradigms are so reduced as to no longer be able to maintain the null subject property, and this study proposes that this feature in popular PB is a contributor to this shift to a partial NSL. Because these extreme circumstances have not existed in EP, contact between NSLs and partial NSLs would not have existed to cause a parameter shift.

5.3 The 2SG and 3SG in Brazil

The *tu/você* merger in Brazil is a largely 20th century phenomenon that has had vast grammatical repercussions for BP. Because this merger is the 2SG counterpart to the 2PL merger of plurals *vós/vocês*, we will analyze it in depth in this section to see how it is similar and different to

---

[6] The topic of creoles is a vast field of study that we will not delve into in great detail, though it is certainly relevant; for the broader topic of pidgins and creoles, see Bickerton (1973), Mühlhäusler (1986), and Siegel (2008); for Portuguese-based African creoles, see Couto (1993), Kihm (1994), Pratas (2002), Hagemeijer (2009), and Cardoso, Hagemeijer, & Alexandre (2015), among many others.

[7] Some argue that BP has origins as a semi-creole, see Parkvall & Álvarez López (2003) for a balanced insight into this hypothesis.

[8] A further example which is not relevant to this topic of verbal agreement, is generalized proclisis of clitic pronouns in spoken BP. Generalized proclisis is mostly likely a Bantu influence, as shown by Santos (2018, p. 44).

the *vós/vocês* merger in EP, in regard to both the phases of the innovation and its implications for the NSP.

5.3.1   Subject pronouns

Although the treatment *você* is naturally related to the plural treatment *vocês*, the pronoun we have examined in this study, they are distinct in their application. In spoken Portuguese norms worldwide, the plural address *vocês* has become a standard informal address. The place of the singular *você* is more variable – in Portugal, it is generally a median form of respect, but in Brazil it is a standard informal address that has largely replaced *tu*, at least in theory. This 2SG merger presents some parallels to the replacement of *vós* by *vocês* that we are examining in Portugal; an innovative address, originally a V-form, is taking the place of the traditional T-form. Therefore, knowledge of the Brazilian 2SG may further our understanding of pronoun replacement for the 2PL in EP and beyond.

Lopes & Machado (2005), in a study of 19th century family letters, illustrate the relatively advanced level of replacement of *tu* by *você* as an intimate address. At this point in time, *você* was becoming grammaticalized in BP as a pronoun instead of a nominal phrase. One remaining nominal property was that it was explicitly expressed 83% of the time during an era in which BP was still a null subject language (Lopes & Machado, 2005, p. 57), while *tu* is a null subject 98% of the time. The lesser educated grandmother's letters show a preference for *você* over *tu*, the former constituting 57% of 2SG addresses, the latter, 43%. The more linguistically conservative and more highly educated grandfather prefers *tu* systematically, at a rate of 96%. In the grandmother's writing, there is mixed agreement when using *você*, mixing the grammatical 2SG pronouns *teu(s)/tua(s)* and *te* with *você*. For the possessive pronoun, there are 14 uses of the 2SG against 2 uses of the 3SG, and for the clitic pronoun, 100% of 9 instances are *te*. As such, we can see a tendency for the elimination of the subject pronoun and its verbal morphology, which have index agreement, prior to the elimination of clitic and possessive forms, which have concord agreement.

When we look at the subject pronoun *tu* contemporaneously, however, we see that it has not disappeared. Scherre (2007, pp. 201-204) shows that there are actually five subsystems of 2SG pronoun and verb agreement.

(42) Exclusive use of *você*

Seen in Minas Gerais, all the Central-West except Brasília, almost all of Paraná, and in Salvador, Bahia

(43) Predominant use of *tu* with low rates of 2SG verbal agreement

Seen in the South

(44) Predominant use of *tu* with greater rates of 2SG verbal agreement

Seen in the South, North, and Northeast

(45) Mixed use of *tu* and *você* with no 2SG verbal agreement

Found in the Southeast, in the Cities of Rio de Janeiro and Santos, São Paulo, and in the bilingual regions of Paraná and Rio Grande do Sul, in rural Bahia, and in Brasília

(46) Mixed use of *tu* and *você* with mixed 2SG verbal agreement

Seen in the South and in the Northeast

The persistence of *tu* in the fourth subsystem, (45), is significant in that the 2SG verbal morphology has been eliminated but the subject pronoun remains. It exists in Rio de Janeiro, one of the country's cultural and economic centers, and is in becoming a regular part of the speech in Brasília, the political capital of the country. Its dialect is still in formation due to the city being a planned city populated by migrants from other regions starting in 1960, and appears to be adopting *tu* quite aggressively as an informal address. Lucca (2005, p. 83, apud Scherre (2007)) shows young male speakers to produce *tu* at a rate of 78% when they did not know they were being recorded, and Dias (2007, p. 76, apud Scherre (2007)) found a rate of 45% for when they were partially aware of being recorded. As such, we can see *tu* in vigor in parts of the country that have completely abandoned the 2SG verbs, with one exception – the 2SG imperative.

5.3.2 Verbal morphology

Many BP dialects still employ the 2SG imperative. The historically correct imperative for the pronoun *tu* is synchronically identical to the 3SG present indicative, and for the pronoun *você*, it is identical to the 3SG present subjunctive. The 2SG command is historically derived from the Latin infinitive verb with the infinitive morpheme removed, not the 3SG present indicative (Scherre, 2007, p. 197), but because the 2SG command and 3SG present indicative are synchronically identical (excepting the verb *ser*) and this is how speakers understand the imperative, we will refer to it as the indicative command, and to the *você* imperative form as the subjunctive command.

Despite subsystems (43) (44) (46) being the only ones that allow for 2SG verbal morphology, all five of these systems can employ the indicative command, despite most of them having low to nonexistent 2SG agreement for the present indicative. A dialect such as that of Rio de Janeiro is the third model, using *tu* as the most intimate of three form but with no 2SG agreement except for the imperative, and has low use of the possessive *teu*. What is more, this appears to be an expanding phenomenon; Scherre (2007, 2004) analyzes the commands used in *Turma da Mônica*, a popular comic strip in Brazil, and finds rising use of the indicative command, going from 7% in 1970-1971, to 72% in 2001-2005 (Scherre, 2007, p. 211).

5.3.3 Possessive and object pronouns

The order of replacement for 2SG possessives and clitics for BP, is not entirely clear. Returning to Lopes & Machado (2005), we can see this combination:

(47) *Com muitas saudades **te** abraça **sua** dindinha do coração.*

With many longings you.ACC.2SG hugs you.POSS.3SG dindinha from the heart.

"Missing you lots, your Dindinha sends a heartfelt hug."

(48) *Lembrança a os meninos e**te** abraça Com muitas saudades **Sua** Mai*

Reminder to the boys   and you.ACC.2SG hugs with many longings   you.POSS.3SG mom

e Amiga Barbara.

and friend Barbara.

"Remind the kids, sending you a hug and missing you lots, your mother and friend

Barbara."

Such an order of pronoun replacement is different from the models (11) and (12) that Lara

Bermejo (2015) observed for Peninsular Spanish and EP, discussed in 2.3. So then, though there

appears to be a clear paradigm for the progressive elimination of pronouns in the 2PL/3PL merger on

the Iberian Peninsula, it is not so clear from the perspective of BP. While 2SG clitics and possessives

are theoretically becoming obsolete in the modern BP norm, the is quite complicated and varies by

region and in time. Lucena (2016) studies a diachronic sample of letters written in Rio de Janeiro,

which are classified into three phases, 1870-1899, 1900-1939, and 1940-1979. Only in the first phase

is *seu* predominantly a V-form; in the next two, it is amply used as a T-form. This was not the case for

the clitics related to *você*, however, as seen in Souza (2014), who analyzes letters from Rio for a

similar time period – 1880-1980. This author compared *te* with the other clitic options *lhe*, *o/a*, and

clitic *você*, and showed that *te* was the predominant form in every time period, in every part of a

letter, and for the variables of sex and verbal morphology. The only circumstance in which it was not

the preferred clitic was in personal correspondences, as compared to family and romantic

correspondences, showing that *lhe* had retained a function as a V-form while *te* had persisted as a T-

form. Most pertinently to our interest in the order of pronoun substitution in a merger, *te* was the

dominant clitic in letters with exclusive use of *você*.

Table 15 Clitic use in Rio de Janeiro letters using only the subject pronoun *você* (1880-1980)

| Te	Você	O/A	Lhe	Null
50.8%	16.1%	22%	8.5%	2.5%
(60/118)	(19/118)	(26/118)	(10/118)	(3/118)

Adapted from Souza (2014, p. 96)

Additionally, this data is for all time periods studied. For the period 1956-1980, *te*

constitutes 56% of clitics, and at this same time there is no longer exclusive use of the subject

pronoun *tu*, a 95.5% use of *você* as the exclusive subject, and only 4.3% mixed *tu/você* use [*sic*, the

missing .2% is probably a rounding anomaly.] (Souza, 2014, pp. 110-112). More supporting evidence

comes from Lopes, Rumeu, & Carneiro (2013, p. 204), whose study of letters shows 75% (3/4) use of

te with exclusive use of the *você* subject. This sits in stark contrast to the findings of Lucena (2016, p.

163) for *teu* and *seu*, showing 4.5% (5/112) occurrence of *teu* and 95.5% (107/112) occurrence of *seu*

for the period of 1940-1979. Summarily, *você-te-seu* was the dominant paradigm in Rio de Janeiro

for the period of 1940-1980, distinguishing itself from our model (17) and Lara Bermejo (2015) that

show the possessive as the final pronominal holdout of the 2PL in EP.

Beyond this historical analysis, there is evidence for rising 2SG forms in Brazil today. Almeida

(2009) examines the *lhe* clitic, used as both an accusative in the city of Salvador, Bahia, and finds

that its use is being substituted by the *te* clitic. This innovation is being led by women and younger

generations, who employ *te* regularly. The factor of higher or lower levels of education had relatively

little impact on the choice of clitic.

Figure 3 Relative weight of *lhe* use in Salvador
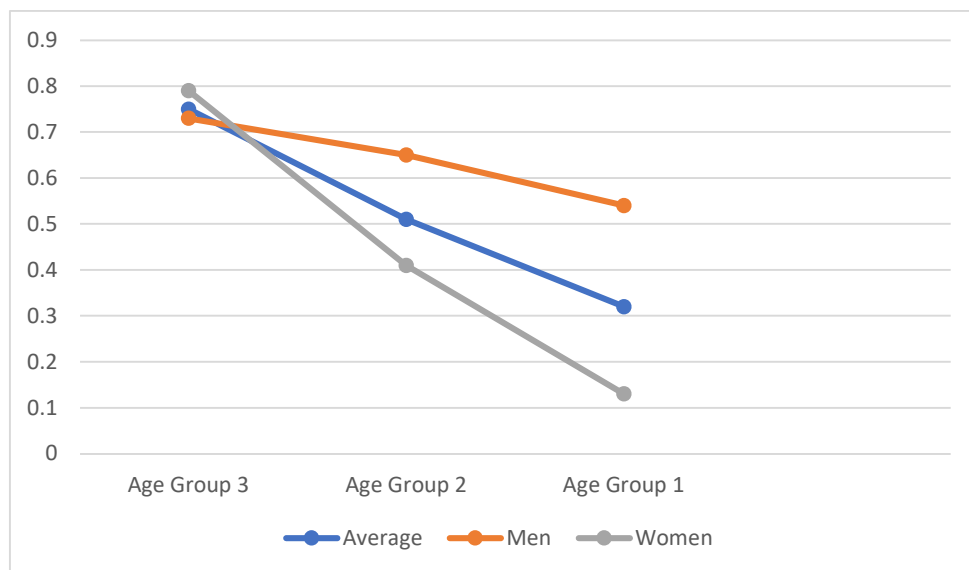
Adapted from Almeida (2009, pp. 144-147)

The implication of this is that in the history of Brazil, and certainly Salvador in particular, the

clitic *te* fell into decline with the introduction of *você* and its grammatically corresponding clitic *lhe*,

but then in more recent history the process reversed and the clitic *te* began to rise in use, replacing

lhe. This is evident upon comparing Figure 3, showing the contemporary rise in *te* usage, with Lopes,

Rumeu, & Carneiro (2013) study of letters from Bahia from 1880-1950, seen in Table 16. Considering

that *te* is both an accusative and a dative, upon combining accusative *o/a* and dative (and sometimes accusative) *lhe*, we see the 3SG clitic pronouns occurring at a rate of 85.2% versus *te* at only 3.7%.

Table 16 Clitic use in letters from Bahia using only the subject pronoun você (1880-1950)

Te	Você	O/A	Lhe	Null
3.7%	3.7%	29.6%	55.6%	7.4%
(1/27)	(1/27)	(8/27)	(15/27)	(2/27)

Adapted from Lopes, Rumeu, and Carneiro (2013, p. 204)

5.3.4   Summary

While this section has detailed some of the distribution of 2SG pronouns and verbs in BP, it has by no means been exhaustive. The topic of the *tu/você* merger is extensive and the focus of much scholarship and analysis. Beyond those cited already, further work on subject pronouns include Menon (2000), Modesto (2006), Mota (2008), Santana & Teixeira (2009), Franceschini (2015), and Babilônia & Martins (2015); more studies of imperatives include Scherre (2012), Lamberti & Schwenter (2018), Rumeu (2019), and Carvalho (2021); for work on 2nd person clitics, Araújo & Carvalho (2015), Oliveira (2015), and Schwenter, Hoff, Dickinson, Bland, & Lamberti (2018); and on possessive pronouns, Arduin (2005) and Assis & Marcotúlio (2015), among many others.

Because a complete review of 2SG pronouns and verbs throughout Brazil would require a book or even multi-volume project, our brief analysis of Brasília, Rio de Janeiro, and Salvador has served an illustrative purpose, showing key differences between BP's *tu/você* merger and EP's *vós/vocês* merger. These dialects show us that the restructuring of the pronominal paradigms in BP is following a different course than EP, and the 2SG/3SG merger is not following the same model as EP's 2PL/3PL merger.

In summary, the hierarchy of subject pronoun > reflexive/verb > accusative > dative > possessive cannot be said to apply to BP. The *te* clitic has proved more resilient than the *teu* possessive in Rio de Janeiro, as a clear example, and is even experiencing a resurgence in Salvador. The classic 2SG imperative is also experiencing a resurgence, and as such it appears this merger may

never be "complete" in the way that the *vós/vocês* merger is complete in BP and appears to be headed towards completion in EP.

5.4 The 2PL and 3PL in Brazil

Contemporaneously, the use of the grammatical 2PL is extinct in Brazil.[9] The history of 2PL pronouns in Brazil would make an interesting comparison with the 2PL in EP, as it was replaced progressively by the 3PL, but this merger is poorly documented, appearing to have begun in the early 19th century or even earlier. There is proof of this in the aforementioned 19th century letters, which show a complete categorical replacement of *vós* by *vocês*. Lopes & Machado (2005) assert that it was considered archaic by this time, attributing this to Faraco (2017, p. 123), however this is an incorrect interpretation; Faraco was citing Cintra (1972, pp. 29-30) who states that *vós* had become archaic as a singular form of address, not in its plural usage. As such we can see that the subject pronoun *vós* had disappeared from BP by the time of the late 19th century, but the timeline of its disappearance is uncertain. Although the 2PL subject and verbs are absent, the corresponding 2PL *vosso* and *vos* are still present as pronouns that semantically match the grammatically 2PL *vocês*. Both of these pronouns form part of the style of letter writing at the time, as they appear mostly in the closing line of the letters they are present in.

In the case of *vos*, it only appears in closing lines, and the grandmother (the more innovative speaker) presents mixed agreement. Notably, she utilizes the *vocês* possessive form *sua*, but in one instance uses the accusative *os,* and in the next uses the accusative *vos*.

(49) *...que de   ca   de longe **sua**   avó   **os**   abraça e abençoa*

...that from here from far you.POSS.3SG grandmother you.ACC.3PL hug and bless

*com muitas saudades e que **vos**   quer bem.*

with many longings and that you.ACC.2PL loves well.

---

[9] With the exception of religious jargon.

"…that from far away, your grandmother sends you hugs and blessings, missing you guys

and loving you lots." Lopes & Machado (2005)

In the case of *vosso*, it appears in the closing lines of letters, and also in other parts of the

letters of the most conservative speaker, Christiano. His letters present variation between the *seu*

and *vosso* possessives in the body of the text, while only using the clitic *vos* in closing lines,

suggesting that his retention of the 2PL possessive is a variable part of his idiolectal grammar, but his

use of the 2PL clitic is strictly related to a stylistic norm of the time used in closing letters.

A fascinating point in Lopes & Machado (2005) is that while *você* has the nominal

characteristic of overt use at an 83% rate, *vocês* shows a much lower rate, only 43% overt expression

to 57% null. This suggests that the insertion of *vocês* into the BP pronominal paradigm occurred

before the insertion of *você* because it lost the nominal characteristic of overt expression first, and it

confirms this study's view that the merger of *vós* and *vocês* did not cause a shift in null subject usage

in Brazil.

These examples show two things. One is the possibility of a piece of grammar's

crystallization in sayings or in certain contexts, such as the closing line of a letter. Secondly, they

establish an incomplete model for the *vós/vocês* merger in BP. The subject pronoun and 2PL verbs

disappeared first, and the clitic and possessive pronouns remained in use longer, semantically

matched with the grammatically 3PL *vocês* and conjugations, which is in line with the first two

phases of the replacement model for EP, shown in (17) in section 2.3.2. With regard to the

accusative, dative, and possessive, however, the model is unclear. For the grandmother, the 2PL

clitic was present as a resistant, crystallized form as a part of a closing line in a letter, outlasting the

2PL possessive pronoun, but for the grandfather, the 2PL possessive coexisted more freely with the

3PL possessive throughout his writing while the *vos* clitic was relegated to the closing line. This

leaves the final phases of the *vós/vocês* merger unclear.

The *tu/você* merger in Brazil does not follow the same model as the *vós/vocês* merger in

Portugal, with the retention of *tu* in some places, the retention of the 2SG imperative, and locations

that retain the *te* dative clitic more strongly than the *teu(s)/tua(s)* possessive pronoun. As such, we cannot definitively say how the *vós/vocês* merger occurred in BP; it may have followed the same course as EP is following, or it may have presented a different order, particularly regarding the accusative, dative, and possessive, like the *tu/você* merger.

5.5 Why Brazilian Portuguese is no longer [+ NSL]

The grand question that BP's mergers put forth is understanding why the language has shifted away from [+ NSL]. The linguistic context of BP is quite different from EP, and as such, we should reflect on the circumstances which facilitated this change. Duarte (1995) makes a strong case for the erosion of the verbal paradigm proceeding from the loss of the 2SG verbal morpheme, combining *tu* and *você* under the 3SG. This precipitated a drastic drop in 2SG from 69% null subject usage in 1918 to 25% in 1937 (p. 20).

When we consider that Duarte's (1995) timeframe for this shift in educated urban BP is in early 20th century, and that Brazil abolished slavery in 1888, we see that vernacular dialects born of irregular transmission necessarily predate this shift. While this elimination of the 2SG is relevant and impactful, it occurred in the early 20th century concurrently with other relevant phenomena.

One factor complicating the understanding of 3SG verbs at this time was the indefinite and arbitrary interpretation given to the Ø3SG. Nunes (1990, p. 111) shows that the passive *se* clitic had become mostly used as an indeterminate by the 19th century, and in the 20th century, the suppression of this clitic, creating a Ø3SG with an indeterminate value, surpassed 50%. This feature contrasts EP, which prefers to use the *se* clitic to construct sentences with an indeterminate subject. As such, when the *tu/você* merger occurred, the Ø3SG as an indeterminate was a concurrent phenomenon. While arbitrary pronominals (an English example being "They stole my bike." in which "they" has an arbitrary interpretion, the same as "someone") are not a distinct grammatical person and do not affect the NSP (Jaeggli, 1986), the same may not be true of an indeterminate null subject, and the interpretation of the Ø3SG in BP is both arbitrary and indeterminate. Furthermore, Holmberg (2009b) finds this null generic pronoun (preferring the word generic to indeterminate) to

be a common feature among other partial NSLs, including Finnish, Marathi, Hebrew, and Icelandic, and this author points out that none of these languages allow 3rd person determinate null subjects without an accessible antecedent. As such, we propose that the indeterminate Ø3SG in BP could not coexist in a verbal paradigm with determinate Ø2SG and Ø3SG, and this was a factor in the drastic drop in Ø2SG subjects in the early 20th century.

Returning to the topic of creoles and Afro-Brazilian Portuguese, there is a presumptive distinction with the educated norm of BP. As discussed in 5.2, Afro-Brazilian dialects are the result of irregular linguistic transmission, and while we cannot know the exact grammar of the original L2 speakers, Lucchesi and Baxter (2009) note that a reduction in inflectional morphology is inherent in any process of irregular linguistic transmission. This means that these varieties, especially those with more radical restructuring, would have morphologically defective paradigms that would require overt pronouns because they no longer fit the model in (24), meaning that vernacular BP should show low use of null subjects.

As seen in Sessarego & Gutiérrez Rexach (2017), many Afro-Hispanic dialects are also partial pro-drop languages. This also applies to BP's vernacular dialects, which are more likely to reduce plural morphemes in nominal syntagma, noting the plural in only the first word of the syntagma and replacing verbal morphology with the 3SG (Baxter, 2009; Lucchesi, Baxter, & Silva, 2009). This paradigm reduces verbal morphology to only the 1SG for self-reference and the 3SG for all other subjects, as seen in example (40) in section 5.2. Furthermore, in one of the most radical and well-preserved Afro-Brazilian dialects in Helvécia, Bahia, Baxter (1997) shows that even the 1SG conjugation can be leveled by the 3SG, seen in (41) in section 5.2.

Consequently, Lucchesi (2009) examines the NSP in Afro-Brazilian dialects, expecting lower use of null subjects, but the result is similar rates of null subjects in both the educated urban speech studied by Duarte (1995), at 29%, and the rural Afro-Brazilian dialects, at 27%. Lucchesi (2009) proposes that this is most likely a convergent phenomenon between the two varieties, with educated BP reducing null subjects and vernacular BP augmenting null subject use.

Taking this into account, the early 20th century rise in overt subjects that observed in the letters and plays of educated, literate Brazilians of high socioeconomic standing did not exist in a vacuum; it was counterposed to the reduced verbal paradigm of poorer and Blacker speakers, whose reduced verbal paradigm required overt subjects. Despite the conscious stigmatization of this "incorrect" verbal paradigm by speakers of educated BP, these sociolects coexisted in the same space and necessarily had some form of contact with another. Even more significantly, the abolition of slavery led to a large migration of free Blacks to urban centers, particularly Rio de Janeiro, whose population increased rapidly from 1890 to 1920, and continued to receive significant numbers of freed Blacks throughout the 20s, peaking in the 1930s (Costa, 2015). It is between 1918 and 1937 that Duarte (1995) notes a precipitous drop in null 2SG in Rio de Janeiro. The arrival of hundreds of thousands of Afro-Brazilians, with a vernacular BP paradigm, was the sociopolitical backdrop of this time.

Although the reduced verbal morphology of only two persons was and still is commonly considered incorrect by normative standards, using an overt pronoun instead of a null pronoun is not stigmatized, something that Lucchesi (2009) highlights. This is a subtler grammatical feature that is neither correct nor incorrect in the eyes of prescriptivist grammar, potentially allowing its influence to spread more easily. In other words, the difference between *"Nós vai embora"* and *"Nós vamos embora"* is a clearly maintained sociolectal boundary, and a speaker of educated BP would consciously recognize the simplified agreement of the former as incorrect grammar, refuse to use it, and perhaps even correct the vernacular BP speaker who did use it. The difference between "*Nós vamos embora*" and "*Ø vamos embora*" is not stigmatized, however, and said educated speaker would probably not consciously recognize this variation at all, considering it a perfectly grammatical option, and would therefore be susceptible to adopting the phenomenon subconsciously.

In conclusion, the [+ NSL] trait of the Brazilian educated norm existed in a linguistic space filled with other phenomena related to the null subject, while dialectal EP does not. The rise of non-null subjects in BP coincides with the *tu*/*você* merger, which should not have led to a shift alone, and

also coincides with the rise of an indeterminate Ø3SG and the social context of mass Afro-Brazilian migration to urban centers, particularly Rio de Janeiro, bringing partial null subject paradigms that merged five of six conjugations into the 3SG. Looking at BP in this light, we see that the drastic decline in 2SG null subjects occurred in the context of the *tu/você* merger, three possible interpretations for the Ø3SG, and contact between NSL dialects and partial-NSL dialects, factors which do not apply to EP.

6 Conclusions

6.1 The elimination of the 2PL and why European Portuguese is [+ NSL]

From the information gathered here, the *vós/vocês* merger is the first complete verbal syncretism in EP. The *vocês* innovation originated in the district of Beja, in southeastern Portugal, where it forms part of the cross-border *Sprachbund* with western Andalusian Spanish's *vosotros/Ustedes* innovation (Lara Bermejo, 2018b). The innovation is well-established in the Center-South of Portugal, the entirety of this region having at least eliminated the subject *vós* and 2PL verbs. In the North of Portugal, the *vocês* innovation is in its incipient phases, and it is spreading progressively from region to region. By the 90s, the latest date for the data used in this study, the innovation had spread into the dialectally northern areas of eastern Coimbra, northern Aveiro, Viseu, Porto, and Braga. Under our model of five phases, Porto and Aveiro have commenced the second phase, and Braga, Viseu, and eastern Coimbra are in the first phase of the innovation. From our global percentages, we have confirmed that this innovation has not affected EP's status as a NSL, as both dialects with five and six conjugations maintain high rates of null subject use.

The data from this study sheds light on a variety of phenomena and potential phenomena. In short, these are the major conclusions of this study:

i)   The model subject pronoun > reflexive/verb > accusative > dative > possessive is correct for Portugal. The 2PL subject pronoun *vós* is being eliminated more rapidly than 2PL conjugations in the North, where it constitutes 18.9% (14/74) of all instances of *vós* and *vocês* and only 6.2% (12/195) of all verbal subjects in the North.

ii)   V*ocês* + 2PL is a viable agreement in the North, specifically the districts of Viseu, Aveiro, Braga, and potentially Porto. *Vós* + 2Pl competes with *vocês* + 2PL until *vós* is removed.

iii)   The coastal North presents more locations with no *vós* subject pronoun, suggesting

that it is has begun the 2PL innovation while much of the inland North conserves the

2PL and has yet to adopt the innovation.

iv)   Overt *vós* is greatly inferior to Ø2PL use, but the global percentage of null subjects in

the North is 60% (117/195). The global percentage for all other Portuguese

territories is 65.2% (174/276). This corroborates our hypothesis that the 2PL/3PL

syncretism in Portuguese does not alter its [+ NSL] trait.

v)   2PL verbs are most commonly an affirmative imperative conjugation and are more

frequent than 3PL imperatives. 2PL commands appear to be more resilient than

other conjugations, and the indicative is the next most resilient, creating this

replacement model: subjunctive/conditional/personal infinitive > indicative >

imperative.

vi)   The *vocês* innovation is present in the insular Portuguese dialects.

Further implications from our data include the following, but they are less certain due to

relatively scant data.

vii)   The *vocês* innovation in the Azores is likely in the third stage of innovation. Further

study is necessary to confirm this and to verify the specific stage of innovation for

each island.

viii)   The Madeira archipelago appears to have adopted the *vocês* innovation, at the very

least having completed the second phase of our model.

Finally, the ultimate frustration and joy of any research project is that it leaves us with more

questions than we ever started with. Some questions that this thesis has left unanswered are as

follows.

What is the more exact extent of the *vós/vocês* merger in Portugal? Furthermore, what is

the state of the *vós/vocês* merger in the insular Portuguese dialects? For the insular dialects, we

have made an educated guess that they are in the third stage, but this is uncertain. These are

questions that can be answered with the ALEPG, but this atlas is not yet published, and the hundreds of hours of audio recordings from dozens of locations require further study for more detailed conclusions.

Why and how is Brazil's *tu/você* merger progressing? We have provided some limited insight into this topic, showing that the clitic *te* is the most resilient form, but we do not have a theoretical explanation for this, and the *tu/você* merger is a gargantuan question, and involves every type of linguistic variable imaginable. Diachronic, diastratic, diageneric, and diatopic information is necessary to truly have a grasp on this continuing phenomenon.

What we can conclude with certainty from this study is that the *vós/vocês* innovation is in progress in Portugal and as it progresses, it is not posed to make in change in Portuguese's characteristic as an NSL.

References

Aguiar, J., & Paiva, M. C. (2017). Vocês tenham cuidado, sois educadas para isso: second person pronouns in Braga speech. *Studies on Variation in Portuguese*, 136-152.

Almeida, G. S. (2009). *Quem te viu quem lhe vê: A expressão do objeto acusativo de referência à segunda pessoa na fala de Salvador*. Master's dissertation, Universidade Federal da Bahia.

Araújo, F. J. M., & Carvalho, H. M. (2015). TE e LHE como clíticos acusativos de 2ª pessoa em cartas pessoais cearenses. *LaborHistórico, 1* (1), 62-80.

Arduin, J. (2005). *A variação dos pronomes possessivos de segunda pessoa do singular teu/seu na região sul do Brasil*. Master's dissertation, Universidade Federal de Santa Catarina.

Assis, D. M. S., & Marcotulio, L. L. (2015). Língua e Ensino: os possessivos da segunda pessoa do plural no português brasileiro. *Línguas & Ensino, 1* (1), 167-185.

Babilônia, L., & Martins, S. A. (2015). A influência dos fatores sociais na alternância dos pronomes tu/você na fala manauara. *Guavira Letras, 1* (13), 49-60.

Barbosa, B., Duarte, M. E. L., & Kato, M. A. (2005). Null subjects in european and Brazilian Portuguese. *Journal of Portuguese Linguistics, 4* (2), 11-52.

Baxter, A. (1997). Creole-like features in the verb system of an Afro-Brazilian variety of Portuguese. *Creole Language Library, 19*, 265-288.

Baxter, A. (2009). A Concordância de Número. In D. Lucchesi, A. Baxter, & I. Ribeiro (Orgs.), *O Português Afro-Brasileiro* (pp. 269-294). EDUFBA.

Bickerton, D. (1973). The nature of a creole continuum. *Language*, 640-669.

Brissos, F., Gillier, R., & Saramago, J. (2017). Variação lexical açoriana: estudo dialetométrico do Atlas Linguístico-Etnográfico dos Açores. *Aproximacións á variación lexical no dominio galego-português, 11*, 11-28.

Brown, R., & Gilman, A. (1960). The Pronouns of Power and Solidarity. In *Style in Language* (pp. 253-76). MIT Press.

Cabrera Puche, M. J. (2008). *Null Subject Patterns in Language Contact: The Case of Dominican Spanish.* PhD dissertation, Rutgers University.

Camacho, J. (2008). Syntactic Variation: The Case of Spanish and Portuguese Subjects. *Studies in Hispanic and Lusophone Linguistics, 1* (2), 415-433.

Camacho, J. (2013). *Null Subjects.* Cambridge University Press.

Cardoso, H. C., Hagemeijer, T., & Alexandre, N. (2015). Crioulos de base lexical portuguesa. In M. Iliescu, & E. Roegiest (Orgs.), *Manuel des anthologies, corpus et textes romans* (pp. 670-692). De Gruyter.

Carvalho, L. F. (2021). A expressão variável do imperativo no português brasileiro. *Domínios de Lingu@gem*, 1-38.

Chomsky, N. (1981). *Lectures on Government and Binding* (5th ed.). Foris Publications.

Chomsky, N. (2000). *New Horizons in the Study of Language and the Mind.* Cambridge University Press.

Cintra, L. (1971). Nova Proposta de Classificação dos Dialectos Galego-Portugueses. *Boletim de Filologia, XXII*, 81-116.

Cintra, L. F. L. (1972). *Sobre Formas de Tratamento na Língua Portuguesa.* Livros Horizonte.

Costa, C. E. C. (2015). Migrações negras no pós-abolição do sudeste cafeeiro (1888-1940). *Topoi, 16* (30), 101-126.

Couto, H. H. (1993). The genesis of Portuguese creole in Africa. *Atlantic Meets Pacific: A global view of pidginization and creolization*, 381-389.

Cunha, C., & Cintra, L. F. L. (1984). *Nova Gramática do Português Contemporâneo.* Edições João Sá da Costa.

Duarte, M. E. L. (1993). Do Pronome Nulo ao Pronome Pleno: A Trajetória do Sujeito no Português do Brasil. In I. Roberts, & M. A. Kato (Orgs.), *Português Brasileiro: Uma Viagem Diacrônica* (pp. 107-128). UNICAMP.

Duarte, M. E. L. (1995). *A Perda do Princípio 'Evite Pronome' no Português Brasileiro.* PhD dissertation, UNICAMP.

Faraco, C. (2017). O tratamento você em português: uma abordagem histórica. *LaborHistorico, 3* (2), 114-132.

Franceschini, L. T. (2015). Variação pronominal tu/você em Concórdia/SC: o papel dos fatores sociais. *Signótica, 27* (2), 265-286.

Gonçalves, C. R. (2010). De Vossa Mercê a Cê. *Caderno do CNPF, 14* (4, tome 3), 2535-2550.

Hagemeijer, T. (2009). As línguas de S. Tomé e Príncipe. *Revista de crioulos de base lexical portuguesa e espanhola, 1*, 1-27.

Haspelmath, M. (2008). Parametric Versus Functional Explanations of Syntactic Universals. In *The Limits of Syntactic Variation* (pp. 75-107). John Benjamins.

Holmberg, A. (2005). Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry, 36*, 533-64.

Holmberg, A. (2009a). Null Subject Parameters. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Orgs.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 88-124). Cambridge University Press.

Holmberg, A. (2009b). The Null Generic Subject Pronoun in Finnish. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Orgs.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 200-230). Cambridge University Press.

Holmberg, A. (2016). Null Subjects in Finnish and the Typology of Pro-drop. In A. Tamm, & A. Vainikka, *Uralic Syntax.* Retrieved from https://ling.auf.net/lingbuzz/003138

Holmberg, A., Nayudu, A., & Sheehan, M. (2009). Three Partial Null-Subject Languages: A Comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica, 63*, 59-97.

Huang, C. T. J. (1984). On the Distribution and Reference of Empty Pronouns. *Linguistic Inquiry*, 531-574.

Huang, C. T. J. (1989). Pro-Drop in Chinese: A Generalized Control Theory. In *The Null Subject Parameter* (Vol. 15, pp. 185-214). Kluwer Academic Publishers.

Hyams, N. (1986). *Language Acquisition and the Theory of Parameters.* D. Reidel Publishing Company.

Instituto Camões. (2006). *A Pronúncia do Português Europeu*. Retrieved December 2020, from http://cvc.instituto-camoes.pt/cpp/acessibilidade/capitulo1_2.html

Jaeggli, O. A. (1986). Arbitrary plural pronominals. *Nat Lang Linguist Theory, 4*, 43-76.

Jaeggli, O. A., & Safir, K. J. (1989). The Null Subject Parameter and Parametric Theory. In *The Null Subject Parameter* (Vol. 15, pp. 1-44). Kluwer Academic Publishers.

Kato, M. A., & Negrão, E. V. (2000). *Brazilian Portuguese and the Null Subject Parameter.* Vervuert/Iberoamericana.

Kihm, A. (1994). *Kriyol syntax: the Portuguese-based creole language of Guinea-Bissau* (Vol. 14). John Benjamins.

Lamberti, L., & Schwenter, S. A. (2018). Testando o papel da referência temporal na forma do imperativo em português brasileiro. *Revista Linguística, 14* (2), 231-258.

Lara Bermejo, V. (2012). Ustedes Instead of Vosotros and Vocês Instead of Vós: an Analysis Through the Linguistic Atlas of the Iberian Peninsula (ALPI). *Dialectologia: Revista Electrònica, III*, 57-93.

Lara Bermejo, V. (2015). *Los tratamientos de 2PL en Andalucía occidental y Portugal: Estudio geo- y socio-lingüístico de un proceso de gramaticalización*. PhD dissertation, Universidad Autónoma de Madrid.

Lara Bermejo, V. (2018a). Relaciones de caso y extensión de la concordancia en el español peninsular y el portugués europeo. *Revista de Filología Española, 98* (1), 85-110.

Lara Bermejo, V. (2018b). *La Cortesía en la Península Ibérica: Dialectología del Sprachbund Suroccidental.* Peter Lang International Academic Publishers.

Liceras, J. (1989). On some properties of the pro-drop parameter: looking for missing subjects. In S. Gass, & J. Schacter (Orgs.), *Linguistic Perspectives on Second Language Acquisition* (pp. 109-133). Foris Publications.

Lipski, J. (1999). Null subjects in (Romance-derived) creoles: routes of evolution. *Annual Meeting Proceedings of the Society for Pidgin and Creole Linguistics, Los Angeles*.

Lobo, M. (2016). Sujeitos Nulos: Gramática do Adulto, Aquisição de L1 e Variação Dialetal. In A. M. Martins, & E. Carrilho (Orgs.), *Manual de Linguística Portuguesa* (pp. 558-580). De Gruyter.

Lopes, C. R. S., & Machado, A. C. M. (2005). Tradição e inovação: indícios do sincretismo entre a segunda e a terceira pessoas nas cartas dos avós. In *A norma brasileira em construção: fatos lingüísticos em cartas pessoais do século XIX* (pp. 45-66). Universidade Federal do Rio de Janeiro.

Lopes, C. R. S., Rumeu, M. C. B., & Carneiro, Z. (2013). A configuração diatópico-diacrônica do sistema de tratamento do português. *Revista do GELNE, 15*, 191-216.

Lucchesi, D. (1994). Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa, 12*, 17-28.

Lucchesi, D. (1998). A constituição histórica do português brasileiro como um processo bipolarizado: tendências atuais de mudança nas normas culta e popular. In S. Grosse, & K. Zimmermann (Orgs.), *"Substandard" e mudança no português do Brasil* (pp. 73-100). TFM.

Lucchesi, D. (2001). As duas grandes vertentes da história sociolinguística do Brasil. *Delta, 17* (1), 97-130.

Lucchesi, D. (2009). A realização do sujeito pronominal. In D. Lucchesi, A. Baxter, & I. Ribeiro (Orgs.), *O Português Afro-Brasileiro* (pp. 167-183). EDUFBA.

Lucchesi, D., & Baxter, A. (2009). A transmissão linguística irregular. In D. Lucchesi, A. Baxter, & I. Ribeiro (Orgs.), *O Português Afro-Brasileiro* (pp. 101-124). EDUFBA.

Lucchesi, D., Baxter, A., & Silva, J. A. A. (2009). A Concordância Verbal. In D. Lucchesi, A. Baxter, & I. Ribeiro (Orgs.), *O Português Afro-Brasileiro* (pp. 331-372). EDUFBA.

Lucena, R. P. O. (2016). *Pronomes possessivos de segunda pessoa: a variação teu/seu em uma perspectiva histórica*. PhD dissertation, Universidade Federal do Rio de Janeiro.

Luz, M. S. (1956). Fórmulas de tratamento no português arcaico. *Revista Portuguesa de Filologia, 7*, 251-363.

Martins, A. M. (Org.) (2000-). CORDIAL-SIN: Corpus Dialectal para o Estudo da Sintaxe / Syntax-oriented Corpus of Portuguese Dialects. CLUL. Retrieved from
http://www.clul.ulisboa.pt/en/10-research/314-cordial-sin-corpus

Menon, O. P. S. (1995). O sistema pronominal do português do Brasil. *Letras*, 91-106.

Menon, O. P. S. (2000). Pronome da Segunda pessoa no Sul do Brasil: tu/você/o senhor em Vinhas da Ira. *Letras de Hoje, 35* (1), 121-164.

Milla Muñoz, Á. (2020). *Beyond Stereotypes and Visual Cues: Assessing Language Attitudes toward Andalusian Spanish.* PhD dissertation, Indiana University.

Modesto, A. T. T. (2006). *Formas de tratamento no português brasileiro: a alternância tu/você na cidade de Santos-SP*. Master's dissertation, Universidade de São Paulo.

Mota, M. A. (2008). *A variação dos pronomes" tu" e" você" no português oral de São João da Ponte (MG)*. Master's dissertation, Universidade Federal de Minas Gerais.

Mühlhäusler, P. (1986). *Pidgin and Creole Linguistics.* Blackwell.

Mussa, A. B. N. (1991). *O papel das línguas africanas na história do português do Brasil*. Master's dissertation, Universidade Federal do Rio de Janeiro.

Newmeyer, F. J. (2004). Against a Parameter-Setting Approach to Language Variation. *Linguistic Variation Yearbook*, 181-234.

Nunes, J. M. (1990). *O famigerado se: uma análise sincrônica e diacrônica das construções com se apassivador e indeterminador.* Master's dissertation, UNICAMP.

Oliveira, T. L. (2015). Os pronomes dativos de 2ª pessoa na escrita epistolar carioca. *LaborHistórico, 1* (1), 81-98.

Parkvall, M., & Álvarez López, L. (2003). Português Vernáculo Brasileiro e a Hipótese da Semi-crioulização. *Abralin, 2* (1), 111-152.

Pratas, F. (2002). *O Sistema Pronominal do Caboverdiano*. Master's dissertation, Universidade Nova de Lisboa.

Rizzi, L. (1982). *Issues in Italian Syntax.* Foris Publications.

Roberts, I. (1993). Verbs and Diachronic Syntax: A Comparative History of English and French. *Studies in Natural Language and Linguistic Theory, 28*.

Roberts, I., & Holmberg, A. (2009). Introduction: Parameters in Minamalist Theory. In T. Biberauer, A. Holmberg, I. Roberts, & M. Sheehan (Orgs.), *Parametric Variation: Null Subjects in Minimalist Theory* (pp. 1-57). Cambridge University Press.

Rumeu, M. C. B. (2019). A inserção do você no português brasileiro escrito dos séculos XIX e XX: reflexos nas construções imperativas de 2SG. *LaborHistórico, 5* (Special), 15-38.

Santana, J. C. D., & Teixeira, E. S. P. (2009). Uma abordagem variacionista do uso do TU por falantes cultos de Feira de Santana. *SEMOC-Semana de Mobilização Científica-Segurança: A paz é o fruto da justiça*.

Santos, L. M. S. (2018). *A influência africana no português vernáculo brasileiro: reflexões sobre clíticos de objecto.* Master's dissertation, Universidade de Lisboa.

Saramago, J. (Ongoing). *Atlas Linguístico-Etnográfico de Portugal e da Galiza*. CLUL. Retrieved from http://www.clul.ulisboa.pt/projeto/alepg-atlas-linguistico-etnografico-de-portugal-e-da-galiza

Scherre, M. M. P. (2007). Aspectos sincrônicos e diacrônicos do imperativo gramatical no português brasileiro. *Alfa: Revista de Linguística*, 189-222.

Scherre, M. M. P. (2004). O imperativo gramatical no português brasileiro: reflexo de mudança lingüística na escrita de revistas em quadrinhos. *Actas do XX Encontro Nacional da Associação Portuguesa de Linguística* (pp. 817-829). Lisbon: APL.

Scherre, M. M. P. (2012). Padrões sociolinguísticos do português brasileiro: a importância da pesquisa variacionista. *Tabuleiro de Letras, 4*.

Schwenter, S. A., Hoff, M., Dickinson, K. V., Bland, J., & Lamberti, L. (2018). Experimental evidence for 2SG direct object pronoun preferences in Brazilian Portuguese. *Revista Linguística, 14* (2), 259-290.

Segura, L. (2013). Variedades dialetais do português europeu. In E. B. P. Raposo, M. F. B. Nascimento, M. A. C. Mota, L. Segura, & A. Mendes, *Gramática do Português* (pp. 85-143). Fundação Calouste Gulbenkian.

Sessarego, S., & Gutiérrez Rexach, J. (2017). Revisiting the Null Subject Parameter: New Insights from Afro-Peruvian Spanish. *Isogloss*, 43-68.

Siegel, J. (2008). *The Emergence of Pidgin and Creole Languages.* Oxford University Press.

Silva, R. V. M. (2013). O português do Brasil. In E. B. P. Raposo, M. F. B. Nascimento, M. A. C. Mota, L. Segura, & A. Mendes, *Gramática do Português* (pp. 145-154). Fundação Calouste Gulbenkian.

Sória, M. V. P. (2013). *Nós, a Gente e o Sujeito Nulo de Primeira Pessoa do Plural*. Master's dissertation, Universidade de Lisboa.

Souza, C. D. (2014). *Eu te amo, Eu lhe adoro, Eu quero você: a variação das formas acusativas de 2ª pessoa em cartas pessoais (1880 – 1980)*. Master's dissertation, Universidade Federal do Rio de Janeiro.

Toribio, A. J. (2000). Setting Parametric Limits on Dialectal Variation in Spanish. *Lingua: International Review of General Linguistics, 110* (5), 315-341.