



## BRAZILIAN PORTUGUESE: A ‘PARTIAL’ NULL SUBJECT LANGUAGE? <sup>1</sup>

### PORTUGUÊS BRASILEIRO: LÍNGUA DE SUJEITO NULO ‘PARCIAL’?

Maria Eugenia Lammoglia Duarte<sup>2</sup>  
Juliana Esposito Marins<sup>3</sup>

**Abstract:** The aim of this article is twofold. In the first place, we present evidence that the syntactic change towards overt pronominal subjects observed in Brazilian Portuguese is not a stable phenomenon; rather, our empirical results allow to follow the parametric change in course and to identify the progressive loss of crucial properties related to ‘consistent’ null subject languages. The contrastive analysis with European Portuguese shows the stronger and the weaker structural contexts in this continuous battle towards the implementation of overt pronouns. Personal sentences (with definite and ‘indefinite’ – arbitrary and generic – subjects, usually referred as “impersonal”) are analyzed in more detail than those we consider impersonal sentences, which include a variety of structures, with climate, existential and unaccusative verbs, . They are, however, shown to have been deeply affected by the re-setting of the value of the Null Subject Parameter. Then, we will briefly compare Brazilian Portuguese with Finnish null subjects to conclude that Brazilian Portuguese does not seem to fit the group of the so called ‘partial’ null subject languages, which seem to exhibit null subjects in very restricted contexts, have a lexical expletive in apparent variation with null and generic subjects as well as in impersonal sentences, when it seems to be merged to avoid a verb-initial sentence.

**Keywords:** parametric change; ‘consistent’ null subject languages; ‘partial’ null subject languages.

**Resumo:** O artigo se desenvolve em torno de dois pontos principais. O primeiro consiste em apresentar evidências de que a mudança sintática em direção a sujeitos pronominais expressos observada no português brasileiro não é um fenômeno estável; pelo contrário, nossos resultados empíricos permitem acompanhar a mudança paramétrica em progresso e identificar a perda progressiva de propriedades cruciais relacionadas a línguas de sujeito nulo ‘consistente’. A análise contrastiva com o português europeu mostra os contextos estruturais mais fortes e os mais fracos nessa luta em direção a pronomes-sujeito expressos. As sentenças pessoais (com sujeitos definidos e ‘indeterminados’ – arbitrários e genéricos, usualmente referidos na literatura gerativa como “impessoais”) são analisadas em maior detalhe do que as que consideramos impessoais, que incluem uma variedade maior de estruturas, com verbos climáticos, existenciais e inacusativos. Mostraremos, entretanto, que elas são profundamente afetadas pela remarcação do valor do Parâmetro do Sujeito nulo. A seguir, faremos uma breve comparação entre o português brasileiro e o finlandês, concluindo que o português brasileiro não parece se encaixar no grupo das chamadas línguas de sujeito nulo ‘parcial’, que parecem exibir sujeitos nulos em contextos muito restritos, têm um expletivo

---

<sup>1</sup> The section containing the analysis of definite reference subjects has been partially published in DUARTE (2019; 2020). The first DUARTE of this article has the support of CNPq (Proc. No. 311608/2018-3). We would like to thank the reviewers for their careful reading and questions we will try to answer as clearly as we can. Of course, all the shortcomings are our own responsibility.

<sup>2</sup> Full Professor, Faculty of Letters, Federal University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil. [eugenia@letras.ufrj.br](mailto:eugenia@letras.ufrj.br)

Orcid: <https://orcid.org/0000-0001-8329-1226>

<sup>3</sup> Adjunct Professor (Vernacular Letters) at the Federal University of Rio de Janeiro. [juespmarins@letras.ufrj.br](mailto:juespmarins@letras.ufrj.br)

Orcid: <https://orcid.org/0000-0003-1856-4800>

lexical em aparente variação com sujeitos nulos genéricos bem como em sentenças impessoais, que parecer ser inserido para evitar sentenças com um verbo em posição inicial.

**Palavras-chave:** mudança paramétrica; línguas de sujeito nulo ‘consistente’; línguas de sujeito nulo ‘parcial’.

## INTRODUCTION

Even though the relation between visible morphology and syntax has lost the importance it had in previous generative theoretical assumptions, one cannot deny its importance when we are concerned about the group of ‘consistent’ null subject languages (ROBERTS and HOLMBERG, 2010:6), like Italian, Greek and Turkish, which “characteristically show ‘rich’ agreement inflection, i.e. distinct personal endings on the verb” and “all persons and all tenses can feature an unexpressed pronoun”.

Therefore, we will start this article showing some very important differences between European and Brazilian Portuguese inflectional paradigms. Two changes, certainly due to contacts during colonization, illustrated in Chart 1, impacted the inflection verb paradigm in Brazilian Portuguese (BP): the apocope of 2<sup>nd</sup> singular ending <-s> associated with 2PS *tu*, a very frequent phenomenon, not subject to any sort of stigma today; the second is the process of denasalization and reduction of the diphthong [ãw] (orthographic <-am>), with the loss of the posterior glide, which can eliminate the distinctive 2<sup>nd</sup> and 3<sup>rd</sup> person plural inflectional ending, a variable process controlled by structural and social factors and the only variant (pronounced without the ending) subject to stigma in urban areas today.

Chart 1. Personal pronouns and present indicative verbal inflection in Brazilian Portuguese (Verb *trabalhar* ‘to work’)

| PERSON | SINGULAR                                      | PLURAL   |
|--------|---|--|
| 1st    | <i>Eu trabalho</i>                            | <i>Nós trabalhamos</i><br><i>A gente trabalha</i>            |
| 2nd    | <i>Tu trabalha(s)</i><br><i>Você trabalha</i> | <del><i>Vós trabalhais</i></del><br><i>Vocês trabalha(m)</i> |
| 3rd    | <i>Ele/ela trabalha</i>                       | <i>Eles/Elas trabalha(m)</i>                                 |

In addition to these phonological processes, BP set of nominative pronouns has also undergone very important changes: 2<sup>nd</sup> person plural pronoun *vós* and the associated verbal inflection (indicated by the strikethrough in Chart 1) was lost, and, from the 19<sup>th</sup> century on, two new pronouns have been added to the system via grammaticalization: *você* from the address term *Vossa Mercê* ‘Your Grace’ – entering in competition with *tu* (2<sup>nd</sup> person singular) and *a gente* (the people = ‘we’ (similar to French ‘on’), beginning a competition with *nós* (we) (1<sup>st</sup> person plural). Today, both new pronouns outnumber by far the conservative pronouns (*tu* and *nós*), particularly *a gente*, a widespread phenomenon. *Tu* and *você* are neutralized in the regions where both coexist without any distinction concerning dimensions of courtesy<sup>4</sup>. In European Portuguese (EP), on the contrary, they are still in complementary distribution and *você* is rarely pronounced, except in asymmetric relations – from superior to inferior (otherwise it is felt as rude).

<sup>4</sup> This is observed in the suburban area of Rio de Janeiro, where such co-occurrence is attested. It is normal to hear “**Você** me disse que **tu** ia chegar cedo.” (You (você) told me that you (tu) were going to arrive early.)

Both new pronouns occur with 3<sup>rd</sup> person singular verb forms, (*você/a gente trabalha*) which contributed to the reduction or “impoverishment” of the Portuguese verbal paradigm.

On the other hand, pronoun *nós* is extremely strong in EP and still outnumbers *a gente* – which makes some specialists, like Lopes and Brocardo (2016), claim that *você* and *a gente* are not yet fully grammaticalized in EP. *Vós* has also disappeared in EP (although it is still attested in some more isolated areas) and has been replaced by *vocês*, but this pronoun usually occurs in combination with clitic, oblique and possessive pronouns respectively, *vos*, *convosco*, *vosso*, belonging to the obsolescent nominative *vós* paradigm, all lost in BP.<sup>5</sup>

In sum, even though BP pronominal/inflectional paradigm that appears in normative grammars and is attested, to a certain extent, in writing (as a consequence of the school action) seems to be the same as EP’s, this is an illusion. That’s why we still find evidence to claim, following Roberts (1993), that EP still presents a ‘functionally rich’ paradigm (<-o>, <-s>, <-∅>, <-mos>, <-m>) with four distinctive forms for present and past – with two syncretisms for 2<sup>nd</sup> and 3<sup>rd</sup> person plural, and, eventually, two zero endings, which allows licensing and identification of null subjects. BP, on the other hand, has exceeded the number of zero endings, even with paradigms that preserve 2<sup>nd</sup> and 3<sup>rd</sup> person plural agreement inflection (<-o>, <-∅>, <-m>). Thus, we have reasons to assume that this is in the root of the parametric change in course in BP towards a non null-subject language. With such assumption, the aim of this paper is to discuss the inclusion of BP among the ‘partial’ null subject languages (ROBERTS and HOLMBERG, 2010: 6), “languages in which the pronominal subject may remain unexpressed under restricted conditions determined by both the morphological and the syntactic context”.

Our paper is organized as follows. In next three sections will present our methodology and a recent contrastive analysis of European Portuguese (EP) and Brazilian Portuguese (BP), focusing on (a) the expression of referential pronominal subjects; (b) the strategies to represent arbitrary and generic subjects; and (c) the competition between null non-referential subjects, i.e. quasi-argumental and non-argumental expletives, and a number of operations to fill Spec, TP, satisfying the EPP; finally, we will compare our findings for BP, whenever possible, with the information we have about Finnish concerning (a) 1st and 2nd persons, on the one hand, and 3rd person, on the other; (b) the possibility of a null generic subject; and (c) the structure of some impersonal sentences. As we will see, the languages under this label “seem extremely unlikely to constitute a single sub-type” of ‘partial’ null subject languages” (BIBERAUER, 2010: 154).

## **A NEW CONTRASTIVE ANALYSIS OF PB AND BP REFERENTIAL AND INDETERMINATE SUBJECTS METHODOLOGY**

The analysis is based on two samples, recorded in Lisbon and Rio de Janeiro, in 2009-2010, in similar neighborhoods, according to the same social stratification, available at [www.comparaport.lettras.ufrj.br](http://www.comparaport.lettras.ufrj.br). The data include 72 interviews recorded between 2009-2010, 36 in Rio de Janeiro and 36 in Lisbon. Interviewers are stratified according to three age groups, three levels of formal education and gender. They are part

---

<sup>5</sup> See also Bacelar do Nascimento, Mendes and DUARTE (2018) with respect to differences between both paradigms.

of a project about varieties of Portuguese as described in the mentioned link. This allows a quantitative analysis of the phenomena focused in this article.

Sentences expressing contrast or emphasis (which must show an overt pronoun) and the second member of coordinated structures with co-referential subjects (which may have a null subject in non null subject languages) have also been excluded.

The linguistic factors chosen for entry into GOLDVARB-X multivariate analysis (SANKOFF, TAGLIAMONTE AND SMITH, 2005) include person, distinctive inflectional verb ending, tense/mood, the structural patterns, the structure of CP and the cluster of semantic features of 3rd person referents, all tested in previous researches. The social factors consider age, years of school attendance and gender. The definition of the structural factors is not casual. The factor groups have been motivated by Calabrese's (1986) seminal article, showing clearly the behavior of what we refer today as a 'consistent' null subject language, with respect to the patterns that lead to a null subject and those that require an overt subject (for functional reasons, obviously). In his words, 'expected referents' lead necessarily to null subjects. After some empirical analyses, since Duarte (1995), the creation of a factor group that could grasp the possible structural patterns, according to the function of the antecedent and the presence or absence of c-command, improved in Barbosa, Duarte and Kato (2005) and finally arrived to Duarte and Rezende dos Reis (2018) and Duarte (2019; 2020a), when all possible structures, including a distant antecedent could be incorporated. As we will show, an 'expected' referent in a relation of non c-command behaves exactly like the 'expected' referent in c-command relation in a 'consistent' null subject language, such as EP, a property lost in BP E-language. Therefore, the crucial question regarding an 'expected' (or adjacent) referent in a 'consistent' NSL does not apply any longer in BP, except for some residual occurrences (see examples (1) to (5) for the structural patterns in EP and BP in the next section.).

As for the role of null non-animate subjects, the research by Cyrino (1993) and Duarte (1993) made it clear that changes in direction of overt and null pronouns, like overt subjects and null objects in BP are guided by a referential hierarchy (CYRINO, DUARTE AND KATO, 2000): the change towards overt pronouns starts from more referential items, like inherently human referents (1st and 2nd persons – and, as we will show, 'indeterminate' subjects, as well) and, then, affect 3rd person, which interacts with [+/-animate] referents, reaching propositional (neuter) subjects later occurrences (see examples (7) to (9) for the semantic features of 3rd person referents in EP and BP in the next section.). The opposite direction towards null objects starts from less referential items – neuter (propositional) objects and, then, affect [-animate] referents. The extreme of the proposed hierarchy – elements with no reference – is only related to expletive subjects, and, naturally, supposed to be affected in the end of the process of the change towards overt pronouns. As we will show, BP does not follow the hierarchy till its end, i.e. it has not developed, as expected, a lexical expletive, such as medieval French (VANCE, 1998) in the past and the Spanish spoken in part of Dominican Republic (TORIBIO, 1996; MARTÍNEZ-SANZ, 2011) in the present, which developed, respectively, the expletives '*il*' and '*ello*'.<sup>6</sup> However, the expletive subject position in BP tends to be filled with a number of raising operations of constituents and with the insertion of other *quasi*-argumental pronouns.

---

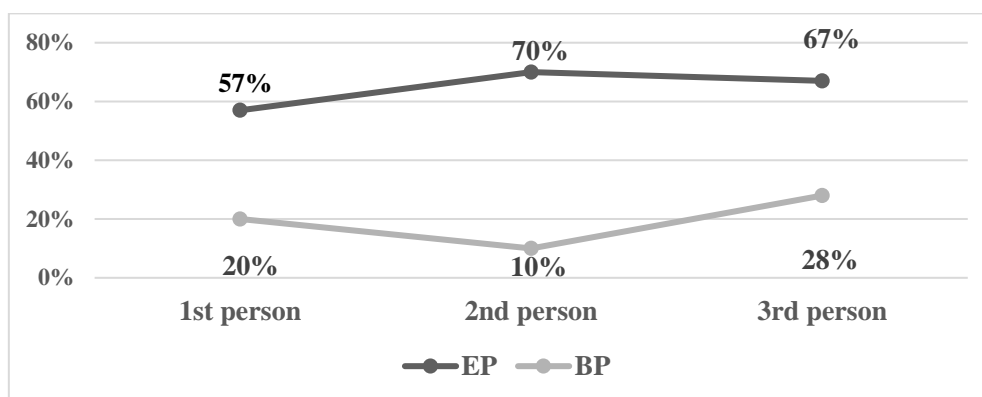
<sup>6</sup> We follow Roberts 1993 and Vance (1996), with respect to the emergence of expletive *il* (attested in Old and Middle French) in the past and Toribio's (1996) and Martínez-Sanz's (2011) about the emergence of expletive "*ello*", in spoken Dominican Spanish in the present. Such emergence followed the change towards overt referential pronouns. It is opportune to remember Labov's (1994) words about the use of the present to understand the past. Likewise, one can say it is useful to use the past to understand the present.

## THE ANALYSIS – REFERENTIAL SUBJECTS IN BP AND EP

In order to demonstrate that BP does not fit in the group of “partial null subject languages” it will be important to compare it with EP so as to show that BP is in continuous process of change, losing some crucial properties that characterize a consistent null subject language, a completely different group from the languages reported as “partial” NSLs.

We collected 2040 definite subjects for EP and 2216 for BP. The first run for each variety confirmed our expectations regarding the status of EP as a ‘consistent’ NSL and the progress of the change in BP. Figure 1 shows the distribution of null subjects for both varieties according to person:

Fig. 1. Null (vs overt) pronominal subjects according to person in EP and BP

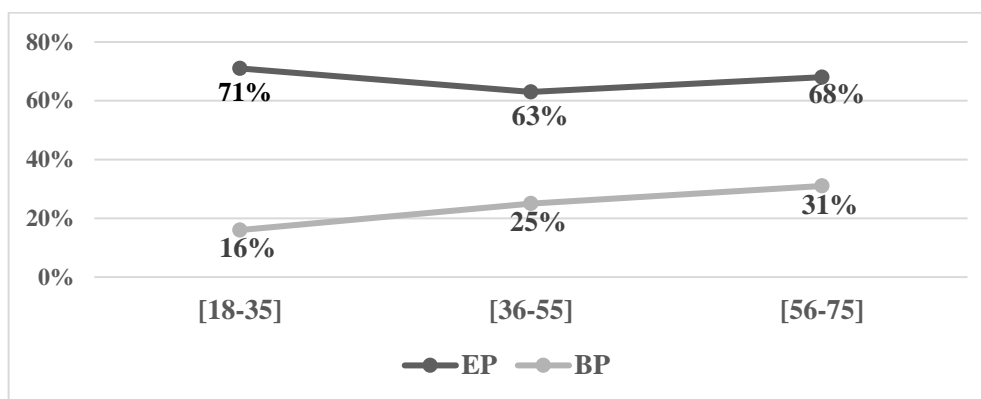


When we compare such results with Duarte’s (1995) findings for Rio de Janeiro speech community, we attest what could be considered a completed change for 2<sup>nd</sup> person, with the same 10% of null subjects, which occur particularly in Yes/No questions, also found in non null subject systems (*Know that man?*). 1<sup>st</sup> and 3<sup>rd</sup> persons show a decrease of null subjects in the lapse of about 16 years separating both samples: an average 25% for 1<sup>st</sup> and 38% for 3<sup>rd</sup> in the 1990s. As for EP, we find a clear preference for null subjects, particularly in 2<sup>nd</sup> and 3<sup>rd</sup> person. The higher rate of 1<sup>st</sup> person singular pronouns confirms Marins’s (2009) results for spoken Italian: 1<sup>st</sup> person subjects tend to be overt when the speaker takes his turn in a conversation.<sup>7</sup>

Since 1<sup>st</sup> person null subjects seem to obey all the other constraints observed for 3<sup>rd</sup> person subjects, we will restrict our analysis to 3<sup>rd</sup> person. The factor groups selected by Goldvarb-X for the realization of null versus overt subjects in EP and BP – **in the same order of relevance** – were (a) the structural patterns (related to the antecedent position/function), (b) the cluster of semantic features of the referent and (c) the structure of CP (a more refined group than the syntactic function of the clause). The age group has not been selected, but the distribution of 3<sup>rd</sup> person subjects across the three age groups in Figure 2 reinforces the stability of EP and confirms the change in progress in BP, attested in DUARTE (1995).

<sup>7</sup> This could suggest that such overt pronouns in EP and in Italian could be considered strong pronouns, such as French “Moi” and English “Me” and the subject in Spec TP would be null. In fact, in BP the paradigms of weak and strong pronouns are homophonous (Kato 1999) – “EU, eu não faço isso (Me, I don’t do it). We thank one of the reviewers for reminding us that.

Fig. 2 Null (vs overt) 3<sup>rd</sup> person pronominal subjects according to age groups in EP and BP.



Notice that the behavior of EP younger group reveals a slight preference for null subjects, while BP continues to suggest what we call a ‘change in apparent time’ in progress (LABOV, 1994), the same descending curve for null subjects among the younger group, confirming the results in Duarte (1995), with a significant decrease (the sample of Rio de Janeiro speech community recorded in 1992 showed, 19%, 20,6% and 34% of null subjects, respectively, for each age group).

Let’s now turn our attention to the first group selected, which refers to the structural patterns, illustrated in the following examples for both EP and BP:

**Pattern A** – Antecedent in preceding main clause with same function c-commands the subject in the following subordinate clause

- (1) a. [O pobre]<sub>i</sub> continua com essa mentalidade porque Ø<sub>i</sub> só pensa nele. (EP)  
the poor guy keeps with this mentality because only thinks in-himself  
‘The poor guy keeps this mentality because he is only concerned about himself.’  
b. [Os pais]<sub>i</sub> passam aos filhos o que eles<sub>i</sub> têm, né? (BP)  
the parents transmit to-the children what they have, right?  
‘The parents transmit to their children what they know, right?’

**Pattern B** – Antecedent in preceding adjunct clause with same function does not c-command the subject in the following main clause

- (2) a. Quando ele<sub>i</sub> mudar, Ø<sub>i</sub> ainda vai estar mais um ano nesta escola (EP)  
when he moves still is-going to-be more one year in-this school  
‘When he moves, he is still going to stay one more year in this school.’  
b. Se ele<sub>i</sub> tem medo, alguma coisa ele<sub>i</sub> fez. (BP)  
if he has fear something he did  
‘If he is afraid, something wrong he did.’

**Pattern C** – Antecedent is in the adjacent clause with same function

- (3) a. [O dono]<sub>i</sub> é um amigo meu. Ø<sub>i</sub> Tem tido problemas. (EP)  
the owner is a friend my has.3ps had problems  
‘The owner is a friend of mine. He has had problems’  
b. [César Maia]<sub>i</sub> era mais tranquilo. Ele<sub>i</sub> investia mais. (BP)  
‘César Maia was more relaxed. He invested more.’

**Pattern D** – Antecedent is in a different function

- (4) a. Tinham medo d[o pai]<sub>i</sub> porque Ø<sub>i</sub> batia-lhes.(EP)  
 had.3pp fear of-the father because beat.3ps-them  
 ‘They were afraid of their father because he beat them.’
- b. Eu não posso ter sentado do lado de [um cara bonitinho]<sub>i</sub>  
 I not can have-been sat of-the side of a guy handsome  
 sem saber que ele<sub>i</sub> era superperigoso. (BP)  
 without to-know that he was super-dangerous  
 ‘I cannot have been sat beside a handsome guy without noticing he was so dangerous’

**Pattern E** – Antecedent is distant – intervening sentences with different subjects

- (5) a. Havia [amigos meus]<sub>i</sub>, até colegas de escola, que jogavam à bola na rua.  
 there-were friends my, even friends of school, who played ball in-the street.  
 Eu não podia porque o meu pai não deixava. Sempre estava à janela,  
 I not could because the my father not would-let. usually was.1ps at-the window  
 Ø<sub>i</sub> chamavam, mas eu não podia. (EP)  
 called.3pp but I not could  
 ‘There were friends of mine, even school friends, who used to play ball in the street. I couldn’t, because my father would not let me go. Many times, I was at the window, they called me, but I was not allowed to go.’
- b. [O meu filho]<sub>i</sub> tava chegando em casa – que nós trabalhamos com festa, como eu  
 The my son was arriving in home – ‘cause we work with party, as I  
 te falei, né – e ele<sub>i</sub> tinha ido comprar bolas. Aí, não tinha as bolas  
 you.CL told, see – and he had gone to buy balloons. So, not had the balloons  
 que nós queríamos. **Ele<sub>i</sub>** trouxe o dinheiro de volta. (BP)  
 that we wanted. He brought the money of back  
 ‘My son was arriving home – cause we work with parties, as I told you – and he had gone to buy balloons. But there were not the ones we wanted. He brought the money back’

**Table 1. Null (vs overt) 3<sup>rd</sup> person subjects according to the sentential pattern**

| Antecedent               | EP Input: 0.756                                  |     |              | BP Input: 0.248                                   |       |              |
|--------------------------|--|-----|--------------|---|-------|--------------|
|                          | N / T  | %   | R. W.        | N / T   | %     | R. W.        |
| A (c-command)            | 78/83  | 94% | <b>0.930</b> | 19/46   | 41%   | <b>0.765</b> |
| B (no c-command)         | 13/14  | 93% | <b>0.854</b> | 15/116  | 11.5% | 0.197        |
| C (in adjacent sentence) | 402/515  | 78% | <b>0.588</b> | 225/586   | 38%   | <b>0.646</b> |
| D (in another function)  | 77/153   | 50% | 0.274        | 37/175  | 21%   | 0.428        |
| E (distant)              | 66/183   | 36% | 0.183        | 35/241  | 14.5% | 0.330        |
|                          | <b>range</b>                                     |     | <b>0.747</b> | <b>range</b>                                      |       | <b>0.435</b> |
|                          | Log likelihood = -44.136<br>Significance = 0.000 |     |              | Log likelihood = -624.928<br>Significance = 0.000 |       |              |

The percentages and the relative weights<sup>8</sup> obtained for EP reflect the exact behavior of a “consistent” Romance NSL: a null subject is the unmarked choice regardless of c-command between a subject and its antecedent, reaching 94% and 93% with and without c-command. Same function is, therefore, crucial, which is also confirmed when the antecedent is in an adjacent sentence (78%). Notice, however, that even in less favorable structural conditions, for functional reasons – antecedent in another function or distant – EP still reaches 50% and 36% of null subjects. Relative weights give us a picture of the strength of patterns A and B in relation to patterns D and E.

When we examine the results for BP, the scenario is completely different: we see a system in change and identify that the first significant loss is the possibility of identification of a null subject in a non-c-command relation with its antecedent (pattern B). This is a crucial loss in the process of change. Percentage as well as relative weight, 11.5% and 0.197 are the lowest, followed by patterns D and E (0.428 and 0.330, respectively). As expected, patterns A and C still show some resistance, but already announce the preference for overt subjects (41% and 38% of null subjects). Adjacency is still important to license a null subject in BP – as shown by Saab (2016) –, but if and only if the antecedent c-commands the subject. The ranges for both varieties reinforce the progress of the change in BP. As we can see, the distribution is not different – BP is losing null subjects in every structural pattern, but there is not a single context where an overt subject is categorical. However, it is true that “control” by an antecedent in the main clause or in an adjacent sentence will be the last contexts to be beaten.

Let’s now see some examples that illustrate, the second factor group selected as significant: the cluster of semantic features:

#### **[-animate / – specific]**

- (6) a. Nós temos que pensar que [a escola]<sub>i</sub> não é uma ilha. Ø<sub>i</sub> Está dentro  
 we have to think that the school not is an island. Is.3ps inside  
 de um contexto, dentro de um determinado tipo de actuação. (EP)  
 of a context, inside of a certain type of actuation  
 ‘We have to realize that the school is not na island. It is in a context, in a certain kind of action.’
- b. [Escola pública]<sub>i</sub> nunca é boa opção porque elas<sub>i</sub> são ruins. (BP)  
 school public never is good option because they are bad  
 ‘Public schools are never a good option because they are bad’

#### **[-animate / + specific]**

- (7) a. [A globalização em si]<sub>i</sub> não é boa nem é má, ou seja, como tudo, Ø<sub>i</sub> tem coisas  
 the globalization in itself not is good nor bad, that is, like everything, has things  
 boas e tem coisas más, não é? (EP)  
 ‘The globalization is not good or bad, just like everything. It has good and bad things, see?’

---

<sup>8</sup> ‘Relative weights’ do not have analytical significance isolated; what matters is their ordination and the difference between the higher and the lower. This is why the logistic program came to replace the probabilistic statistical program in order to follow language change.



- b. [**O sistema público**]<sub>i</sub> é totalmente diferente de empresas privadas.  
 the system public is totally different of companies private  
 Ø<sub>i</sub> não funciona da mesma maneira. (BP)  
 not works of-the same way  
 ‘The public system is totally different from private companies. It does not work in the same way.’

[+animate / – specific]

- (8) a. Eu costumo dizer: quando há [**uma criança**]<sub>i</sub>, alguém tem de ser o pai.  
 I use to-say: when there-is a child, somebody has to be the father  
 Ela<sub>i</sub> não saiu por obra e graça do espírito santo. (EP)  
 she.FEM not come by work and grace of-the spirit saint  
 ‘I used to say: when there is a child, somebody has to be the father. It was not born by work and bless of the holy spirit.’
- b. Acredito que [**a pessoa**]<sub>i</sub> tendo um talento, a partir do momento que **ela**<sub>i</sub>  
 believe.1ps that uma person having a talento to start of-the moment that she  
 desenvolve aquilo, **ela**<sub>i</sub> vai ser sensacional.(BP)  
 develops that, she.FEM will be sensational  
 ‘I believe that a person, having a talent, from the moment she develops that, he/she will become sensational’

[+animate / + specific]

- (9) a. Quando [**as crianças**]<sub>i</sub>; muitas vezes são deixadas sem orientação, Ø<sub>i</sub> acabam  
 when the children many times are left without guidance end-up  
 por escolher caminhos que não são os mais correctos. (EP)  
 by to-choose ways that not are the most correct  
 ‘When children are left without guidance they end up choosing the wrong ways.’
- b. [**A criança**]<sub>i</sub>; acaba abandonando as escolas. Ø<sub>i</sub> Perde o interesse. (BP)  
 the child ends-up abandoning the schools, loses the interest  
 ‘The child ends up abandoning school. He/she loses interest.’

When we look at the role of the cluster of semantic features in Table 2, again we see EP as a well behaved ‘consistent’ null subject language.

**Table 2. Null (vs overt) 3<sup>rd</sup> person subjects according to the semantic features**

|            | EP Input: 0.756                                   |       |              | BP Input: 0.248                                  |       |              |
|------------|---|-------|--------------|--|-------|--------------|
| FEATURES   | N / T   | %     | R.W.         | N / T  | %     | R. W.        |
| -ani/-spec | 12/12   | 100%  | ---          | 7/12   | 58%   | <b>0.863</b> |
| -ani/+spec | 137/142   | 96.5% | <b>0.942</b> | 73/173   | 42%   | <b>0.692</b> |
| +ani/-spec | 191/246   | 78%   | 0.562        | 62/191   | 32.5% | 0.555        |
| +ani/+spec | 308/559   | 55%   | 0.307        | 189/803  | 23.5% | 0,437        |
|            | <i>range</i>                                      |       | <b>0.635</b> | <i>range</i>                                     |       | <b>0.426</b> |
|            | Log likelihood = -440.725<br>Significance = 0.000 |       |              | Loglikelihood = -624.928<br>Significance = 0.000 |       |              |

Non-animate referents associated with the feature [-specific] are categorically null and those associated with the feature [+specific] reach 96.5% of overt pronouns (five

occurrences). In fact, EP differs from Italian and Peninsular, South American and part of the Caribbean Spanish in this respect (see Marins, 2009 and Duarte and Soares da Silva, 2016). These varieties do not have personal pronouns for [-animate] referents. However, if EP does allow such pronouns, they are extremely rare. Our five occurrences in the sample refer to “the globalization”, “the Brazilian soap operas”, “the colonial war”, “liberty and democracy” and “these courses”. Referents bearing a [+animate] feature can exhibit overt pronouns but, as shown in Table 2, the interaction with specificity plays an important role. The lowest rate of null subjects, as expected, is attested with [+animate/+specific] referents (always considering the interaction with the structural pattern).

As for BP, the development of [-animate] personal pronouns is another crucial property to drive BP away from consistent NSLs. As expected, null subjects with [-animate] referents show more resistance, with 58% and 42% of occurrence, but it is clear that the system is halfway to overt [-animate] pronouns. As for [+animate subjects], they already reach 32.5% and 23.5% if associated with [-specific] and [+specific] referents. Again, the observation of the relative weights is extremely important to have a picture of the strength of animacy and specificity in the process. In fact, we have a gradience when we examine the weights. And, once again, we can see that the ranges revealed for both varieties attest the weakening of animacy in BP, just like we observed for the antecedent function in Table 2.

The third factor group selected refers to the structure of CP. More than distinguishing the syntactic function of the clause, we can group them according to the presence of a pronoun (relative, interrogative) in Spec, CP, a complementizer (in the head of CP) or no element in CP. Table 3 shows the relevance of this factor group.

**Table 3. Null (vs overt) 3<sup>rd</sup> person subjects according to the structure of CP**

| CP             | EP Input: 0.757                                   |     |              | BP Input: 0.248                                  |     |              |
|----------------|---|-----|--------------|--|-----|--------------|
|                | N / T   | %   | R.W.         | N / T  | %   | R. W.        |
| Empty CP       | 438/632   | 69% | <b>0.578</b> | 252/901  | 28% | <b>0.517</b> |
| Element in C'  | 161/238   | 68% | <b>0.422</b> | 74/217   | 34% | <b>0.561</b> |
| El. in Spec,CP | 40/81   | 49% | 0.179        | 05/61  | 08% | 0.130        |
|                | <i>range</i>                                      |     | <b>0.399</b> | <i>range</i>                                     |     | <b>0.426</b> |
|                | Log likelihood = -440.725<br>Significance = 0.000 |     |              | Loglikelihood = -624.928<br>Significance = 0.000 |     |              |

The importance of this group is that it reveals that even in a consistent NSL, the presence of a relative pronoun (more than an interrogative pronoun) disfavors null subjects when compared to sentences headed by a complementizer or with an empty CP (cf. Marins, 2009 for Italian and Duarte and Soares da Silva, 2016 for several Romance languages). Percentages and relative weights reinforce the two extremes, and they show why the change with elements in Spec, TP is close to completion in BP.<sup>9</sup>

## THE ANALYSIS – INDETERMINATE/INDEFINITE SUBJECTS IN EP AND BP

The previous section has shown that important properties of consistent NSLs, as far as definite referential subjects are concerned, have been lost or are in the course of the

<sup>9</sup> Interviews do not favor wh-questions in the speech of interviewers, but other analyses (Duarte, 1995 and Marins 2009) show that clauses headed by relative pronouns do favor an overt subject in EP and Italian.

change towards completion. An expected consequence of the change reaches the indeterminate/indefinite subjects. We owe to Egerland (2003:76) the formal distinction between arbitrary and generic subjects, considered by the authors as impersonal subjects<sup>10</sup>: the first label refers “to a non-specific group of individuals, close to ‘some people’, unspecified ‘they’, or ‘someone’; the other denotes a quasi-universal set of individuals, and is roughly equivalent to ‘people’ (in some cases ‘everyone’ or ‘anyone’)”. In ‘consistent’ null subject languages, arbitrary subjects can be represented by a null subject with the verb in the 3<sup>rd</sup> person plural form or associated to the clitic ‘*se*’, excluding the speaker. Generic subjects are represented by the clitic ‘*se*’, and may exclude or include the speaker. The distinction mentioned above has been improved in recent years including degrees of indeterminacy. In addition, Holmberg (2005; 2010) pointed out the possibility of a null arbitrary and a null generic subject in Finnish with no morphological mark, with the verb in the unmarked 3<sup>rd</sup> person singular verb form. Inspired in Egerland (2003) and Holmberg (2005, 2010, among others), Marins, Soares da Silva and Duarte (2017), revisited the data comprising ‘indeterminate/indefinite’ subjects collected by Duarte (1995), whose sole purpose was to check the tendency to use overt nominative pronouns to express them. In their review, they proposed a gradient analysis of such subjects using phi-features, such as [+3<sup>rd</sup> p / + pl] (for ‘arbitrary’, in Egerland’s terms) and [+3<sup>rd</sup> p / + g] (for ‘generic’, in Egerland’s terms), in such a way that the group of entities belongs to the *quasi*-universal set of individuals, and is roughly equivalent to ‘people’ (in some cases ‘everyone’ or ‘anyone’). They also included a middle point in the scale between these extremes: [+1<sup>st</sup> p / + pl], which includes the speaker and not necessarily the hearer. This gradient proposal is published in Marins, Soares da Silva and Duarte (2017). In the same issue, Holmberg and Phimsawat (2017) present a tripartite proposal for the classification of indeterminate/indefinite subjects, using a semantic criterion: *exclusive*, *semi-inclusive* and *inclusive* reference indeterminate subjects. Both have been proved extremely useful to re-group indeterminate/indefinite subjects as we will show here using the same samples analysed for definite reference subjects in the previous section.

The examples of the strategies in each point of the gradient are illustrated as follows (examples come from BP not only for limitation of space but because EP has no null arbitrary nor generic subjects with the unmarked 3<sup>rd</sup> person verb form). We start with arbitrary reference subjects.

**[+3<sup>rd</sup> p / + pl] – [exclusive reference] – (*eles* ‘they’) – null or overt; clitic *se*; Ø (3<sup>rd</sup> ps verb)**

(10) a. Agora que **eles** ‘**tão arrumando** tudo... É muito difícil  
 now that they are repairing all.... Is very difficult  
 vaga na escola.  
 vacancy in-the school

‘Now that they are repairing everything. A vacancy in the school is very difficult.’

---

<sup>10</sup> We prefer to consider such subjects as indeterminate or indefinite because the sentence predicate selects an argument that is “suspended”. The tripartite classification naming them arbitrary as exclusive (excluding the speaker – an argument), and the generic as semi-inclusive or inclusive, including the speaker or optionally including him/her (i.e. the external argument) is proposed by Holmberg and Phimsawat (2017) in the same journal Marins, Soares da Silva and Duarte (2017) propose the features [+3p/p], [1p/p] and [3p/s] to express the same refinement between arbitrary, on the one hand and generic, on the other. A sentence like “this chair sits comfortably” a recurrent example offered by Holmberg is not an impersonal sentence in our view. The subject of “sit” is suspended” – someone can sit comfortably in this chair)

b. Hoje tudo que se pede é um segundo grau, né?  
 today all that CL ask.3SG is a high school, see?  
 ‘Today all they ask is a high school level, see?’

c. Olha... na televisão  $\emptyset_{arb}$  fala muito isso, né?  
 look... on-the television speak-3SG a-lot this see?  
 ‘Listen, they often talk about that on television, see?’

The distribution of the 290 e and 412 data for EP and BP, respectively, can be seen in Table 4.

**Table 4. [+3<sup>rd</sup> p / + pl] – [exclusive reference] according to age groups**

|                    | EP                         |                           | BP                         |                          |                          |
|--------------------|----------------------------|---------------------------|----------------------------|--------------------------|--------------------------|
|                    | (ELES)                     | SE                        | (ELES)                     | SE                       | Null arb                 |
| [18 a 35 years]    | 46 (23%)                   | 32 (35%)                  | 93 (25%)                   | 3 (18%)                  | 9 (47%)                  |
| [36 a 55 years]    | 79 (40%)                   | 24 (26%)                  | 124 (33%)                  | 8 (47%)                  | 4 (21%)                  |
| [56 a 75 years]    | 74 (37%)                   | 35 (39%)                  | 159 (42%)                  | 6 (35%)                  | 6 (32%)                  |
| <b>TOTAL</b>       | <b>199 (100%)</b>          | <b>91 (100%)</b>          | <b>376 (100%)</b>          | <b>17 (100%)</b>         | <b>19 (100%)</b>         |
| <b>% EP and BP</b> | <b>199/290<br/>(68.6%)</b> | <b>91/290<br/>(31.4%)</b> | <b>376/412<br/>(91.3%)</b> | <b>17/412<br/>(4.1%)</b> | <b>19/412<br/>(4.6%)</b> |

Arbitrary subjects in EP show exactly the same strategies characterizing ‘consistent’ null subject languages. In addition, their distribution for the age groups is regular. Considering the total of data, 3<sup>rd</sup> person plural (very rare with an overt pronoun) outnumbers the clitic *se*. BP, on the other hand, shows a clear preference for 3<sup>rd</sup> person plural (with a significant use of null subjects in variation with overt nominative *they* – 42%) followed by the clitic *se* and the null arbitrary, both with regular distribution, except for a small rise among younger speakers. However, when we examine the final line of the table, we see that 91.3% correspond to 3<sup>rd</sup> person plural strategy, with 4.1% and 4.6% to clitic *se* and the null arbitrary strategies. In any case, the fact that BP does license a ‘null arb’ and its absence in EP cannot be disregarded. And the obsolescence of the arbitrary ‘clitic’ is also a fact.

Generic subjects in the next (middle) point of the proposed gradient are illustrated as follows:

[+1<sup>st</sup> p / + pl] – [semi-inclusive reference] – (*nós* ‘we’) – null or overt; (*a gente* ‘the people’) = *we*, French ‘on’) – null or overt<sup>11</sup>

(11) a. Eu não sei quantos mil habitantes **nós** temos.  
 I not know how-many thousand inhabitants we have  
 ‘I don’t know the population we have /what is our population.’

<sup>11</sup> As for *nós* and *a gente*, it must be said that they can be perfectly replaced by inclusive strategies in BP and are in fact mentioned by Holmberg (2010) (‘*A gente não pode/nós não podemos ser felizes neste país*’ = **we** cannot be happy in this country; *Você não pode ser feliz neste país* = **you** cannot be happy in this country)

- b. Aqui **a gente** trabalha com diversas áreas culturais.  
 Here the people work with several areas cultural  
 ‘We work with several cultural areas here.’

**Table 5. [+1<sup>st</sup> p / + pl] – [semi-inclusive reference] according to age groups**

|                    | EP                   |                     | BP                     |                        |
|--------------------|----------------------|---------------------|------------------------|------------------------|
|                    | (NÓS)                | (A GENTE)           | (NÓS)                  | (A GENTE)              |
| [18 a 35 years]    | 141 (42%)            | 17 (19%)            | 37 (30%)               | 219 (43%)              |
| [36 a 55 years]    | 113 (33%)            | 52 (58%)            | 43 (35%)               | 149 (30%)              |
| [56 a 75 years]    | 86 (25%)             | 21 (23%)            | 43 (35%)               | 136 (27%)              |
| <b>TOTAL</b>       | <b>340 (100%)</b>    | <b>90 (100%)</b>    | <b>123 (100%)</b>      | <b>504 (100%)</b>      |
| <b>% EP and BP</b> | <b>340/430 (79%)</b> | <b>90/430 (21%)</b> | <b>123/627 (17,6%)</b> | <b>504/627 (80,4%)</b> |

Even though the distribution is irregular across the age groups for EP with respect to the use of each pronoun, the total distribution confirms the preference for pronoun *nós*, particularly null, reinforcing the strength of the “functionally” rich paradigm (Roberts, 1993). BP, on the contrary shows a regular distribution of *nós* (preferably overt) in the three age groups but, with respect to *a gente*, one can observe an increasing preference for it in the speech of the younger age groups. The general distribution in the final line exhibits an opposite situation when we compare EP and BP. Remember that *a gente* is associated to a 3<sup>rd</sup> person singular verb form, unmarked for inflection.

Generic subjects in the opposite extreme of our gradient, are illustrated below:

[+3<sup>rd</sup> p / + sg] – [inclusive reference] – (você/tu ‘you’) – null or overt; clitic ‘se’; Ø (3<sup>rd</sup> ps verb)

- (12) a. Se **você** não tiver sonho, **tu** não é nada.

If you not have dream, you not is nothing  
 ‘If you don’t have a dream, you are nothing.’

- b. É o que mais **se** vê aí fora... violência, né?

Is what more CL see-3SG outside violence, right?  
 ‘That’s what one sees/you see all around..violence, right?’

- c. Ø<sub>gen</sub> tá precisando de políticos mais sérios.

is needing of politicians more serious  
 ‘One needs/you need more serious politicians.’

- d. Pra beber Ø<sub>gen</sub> **tem** que **ter** noção.

to drink must-3SG to-have notion  
 ‘To drink one has / you have to be wise.’

**Table 6. [+3<sup>rd</sup> p / + sg] – [inclusive reference] across three age groups**

|                    | EP                              |                                  | BP                                 |                                 |                                 |
|--------------------|---------------------------------|----------------------------------|------------------------------------|---------------------------------|---------------------------------|
|                    | VOCÊ/TU                         | SE                               | VOCÊ/TU                            | SE                              | Null <sub>gen</sub>             |
| [18 a 35 years]    | 16 (44%)                        | 73 (43%)                         | 439 (30%)                          | 13 (26%)                        | 24 (42%)                        |
| [36 a 55 years]    | 7 (20%)                         | 48 (28,5%)                       | 577 (40%)                          | 21 (42%)                        | 10 (18%)                        |
| [56 a 75 years]    | 13 (36%)                        | 48 (28,5%)                       | 439 (30%)                          | 16 (32%)                        | 23 (40%)                        |
| <b>TOTAL</b>       | 36 (100%)                       | <b>169 (100%)</b>                | <b>1455 (100%)</b>                 | 50 (100%)                       | 57 (100%)                       |
| <b>% EP and BP</b> | <b>36/205</b><br><b>(17.6%)</b> | <b>169/205</b><br><b>(82.4%)</b> | <b>1455/1562</b><br><b>(93.1%)</b> | <b>50/1562</b><br><b>(3.2%)</b> | <b>57/1562</b><br><b>(3.7%)</b> |

Generic subjects denoting “a quasi-universal set of individuals, roughly equivalent to ‘people’ (in some cases ‘everyone’ or ‘anyone’), repeating Egerland’s words, also exhibit a completely different behavior in EP and BP. Nominative *você* or *tu* as generic subjects, usually null in EP, correspond to 36 occurrences or 17.6% in confront with the clitic *se*, largely preferred by each group, reaching an overall of 82.4%. Interestingly, the youngest group produced 15% more clitics than the two older groups, a suggestion that the clitic is, in fact, robust in EP and the preferred strategy to represent generic subjects. In BP, nominative *você* and *tu* (the first much more frequent than the second in Rio de Janeiro, although they can be attested in variation without the canonical verbal inflection associated with pronoun *tu*, as shown in example (12a)), outnumber by far the other two strategies. The generic clitic *se*, just like the arbitrary one, reaches 3.2% and the null generic, 3.7%. If one can confirm the obsolescence of the clitic, we cannot make predictions about the future of the null generic (or arbitrary), usually mentioned as an argument to include BP among partial null subject languages. Its distribution in this more recent sample is the same for older and younger speakers. We will return to this strategy when we compare Finnish and BP.

## NON-REFERENTIAL SUBJECTS

As for non-referential subjects, we will start this section with a very important change in BP with respect to the replacement of *haver* (there to be) by *ter* (to have), the possessive verb, which keeps its original meaning but has become the preferred existential verb in spoken BP. On the emergence of existential *ter* in BP and its change into a more grammatical verb, see, among many others, Avelar (2006), Callou and Avelar (2000), Avelar and Callou (2007) Marins (2013). Based on popular theatre plays written along the 19th and the 20th centuries in Rio de Janeiro and Lisbon, Marins (2013) finds high rates of *haver*, ranging between 88% and 98% and low rates of *existir* (to exist) and *ocorrer* (occur), never higher than 12%. As for BP, *haver* is the existential verb throughout the 19th century and the three first decades of the 20th century. From the 1930s on, existential *ter* enters the system, reaching 48% in the plays written between 1953-1967, and, in the end of the century, *ter*, corresponds to 81%, *haver* to 13% and *existir* to 6%. More recently, Marins (2019) confirmed the full implementation of *ter* in the recente sample for spoken BP analyzed in the previous sections. Based on the sample

of Rio de Janeiro (2009-2010), the results reveal that the rate of *ter* is around 90% in existential sentences in all three age groups considered in the sample.

The rise of existential *ter* follows the implementation of overt pronominal subjects and we have reasons to claim that this is not a coincidence; as a first step, we have the re-setting of the NSP and the consequent changes in the agreement syntax, allowing the replacement of existential *haver* by *ter*; as a second step, existential sentences with *ter* allow personal sentences and accept the raising of non-arguments to the structural position of the subject without a preposition (see AVELAR, 2009 and MARINS, 2013):

- (13) a. [ $\emptyset_{\text{expl}}$  Tem leite [na porta da geladeira]  
has milk [in-the door of-the fridge]  
b. [Na porta da geladeira]<sub>i</sub> tem leite [t]<sub>i</sub> ]  
[in-the door of-the fridge] has milk  
c. [A porta da geladeira] tem leite [t]<sub>i</sub> ]  
[the door of-the fridge] has milk.  
d. [A geladeira]<sub>i</sub> tem leite [na porta t]<sub>i</sub> ]  
[the fridge has milk [in-the door]  
*'There is milk in the door of the fridge.'*

Other impersonal structures, such as the sentences with unaccusative verbs still allow a null expletive, but there has been an increasing tendency to avoid null expletives either by raising the DP or part of it, when it is modified by a genitive. Duarte and Fernandes (2016) collected data from internet “complaint sites”. The patterns they found are illustrated in (14) and (15):

- (14) a. [TP  $\emptyset_{\text{exp}}$  entrou [DP uma enfermeira]]  
entered a nurse  
b. [TP [uma enfermeira]<sub>i</sub> entrou [ t ]<sub>i</sub> ]  
a nurse entered  
*'A nurse came in.'*
- (15) a. [TP [ $\emptyset_{\text{exp}}$  rachou [a tela do meu celular]]  
cracked the screen of-the my cell phone  
b. [TP a tela do meu celular]<sub>i</sub> rachou [ t ]<sub>i</sub> ]  
the screen of-the my cell phone cracked  
c. [TP [O meu celular rachou [a tela [ t ]<sub>i</sub> ]  
the my cell phone cracked [DP the screen [ t ]<sub>i</sub> ]  
*'The screen of my cell phone cracked.'*

The 115 data with a simple DP (14a,b) showed 28% of V DP order and 72% of DP V. The 128 structures containing a DP genitive modifier (15a,b,c) showed the following distribution: 4% of V DP, 38% of DP V and 58% of Genitive V DP, as shown in (15a,b,c), respectively.

A final illustration of the competition between null expletives and impersonal sentences can be seen in locative raising with weather verbs:

- (16) a. [TP [ $\emptyset_{\text{exp}}$  Venta muito por essas janelas]  
winds a lot through these windows  
b. [TP [Essas janelas]<sub>i</sub> ventam muito [ t ]<sub>i</sub> ] (apud Pontes, 1987)  
These windows wind.3PP.a lot  
*'A lot of wind comes through these windows'*

- (17) a. [TP [ $\emptyset_{\text{exp}}$  Chove sempre nas florestas tropicais]  
           rains often in-the forests tropical  
       b. [TP [As florestas tropicais chovem sempre [ *t* ]<sub>i</sub> ]  
           The forests tropical rain.3PP often  
           *'It often rains in tropical forests.'*

Many other examples, including the restructuring of unaccusative verbs as transitives can be seen in Duarte (2004; 2017); Kato e Duarte (2017; 2018, among others). The absence of a lexical expletive in BP and the raising of prominent elements to Spec, TP in examples (13) to (17) suggest that the change we are witnessing has been triggered not only by the reduction of the inflectional paradigm but by the topic prominence attested in Brazilian Portuguese grammar (PONTES, 1987; KATO and Duarte, 2008; MODESTO, 2008, among many others). Since topic prominent languages do not have lexical items without semantic content (LI and THOMPSON 1976), null expletives tend to be replaced by referential items. BP has developed a paradigm of weak pronouns (Kato, 1999), projecting TP and optionally SubjP (Kíss, 2002; Cardinaletti, 2004), and not only thematic constituents, but non-thematica as well, can occupy Spec, TP, which confirms Avelar and Galves's (2011) hypothesis according to which, in PB, Spec, TP is phi-independent. Therefore, it is not surprising to find left dislocated referential (definite or indeterminate/indefinite) subjects, which correspond to about 10% of overt subjects in the sample analysed, illustrated in (18), but also non-thematic subjects with a dislocated element (19):

- (18) a. [SubjP Um robô]<sub>i</sub> [TP ele<sub>i</sub> vai tirar as medidas.  
           a robot<sub>i</sub> he<sub>i</sub> (it) is-going to-take the measurements  
           *'A robot will take the measurements'*  
       b. [SubjP VOCÊ] [CP quando você viaja [TP você passa a ser turista...  
           you when you travel you come to be tourist  
           *'When you travel you become a tourist.'*
- (19) a. [SubjP [o meu celular]<sub>i</sub>] [TP [a tela]<sub>k</sub> rachou [[ *t* ]<sub>k</sub> [ *t* ]<sub>i</sub> ]  
           the my cell phone the screen cracked  
           *'The screen of my cell phone cracked.'*  
       b. [SubjP As florestas tropicais] [TP elas chovem sempre [ *t* ]<sub>i</sub> ]  
           The forests tropical they rain.3PP often  
           *'It often rains in tropical forests.'*

## BRAZILIAN PORTUGUESE: A “PARTIAL” NULL SUBJECT LANGUAGE?

The inclusion of BP among ‘partial’ null subject languages (Holmberg, Nayudu and Sheehan, 2007) takes into consideration (a) the occurrence of null 3rd person subjects in embedded clauses if ‘controlled’ by the antecedent in the higher clause<sup>12</sup> and (b) the possibility of a null generic subject with the verb in the 3rd person singular, contexts that

<sup>12</sup> We are not taking into consideration the possibility of null 1st and 2nd person subjects in Finnish in root sentences (Holmberg and Nikanne 2003) because, as the authors say, that they are possible **especially in writing**. Since writing is very conservative, in part because of school action, we prefer to work with speech when we are concerned with L1 grammar.



also allow null subjects in Finnish (Holmberg and Sheehan, 2010). However, our data show that BP is losing null subjects in root and in embedded clauses, but still exhibits null 1st, 2nd and 3rd person subjects in every syntactic structural pattern<sup>13</sup>, and, if it is true that null generic (and also arbitrary) subjects are allowed, as we showed in this article, they are extremely rare when compared to other strategies with overt nominative pronouns. As for Finnish, (a) we have no idea how frequent null 3rd person embedded subjects are in this grammar; (b) contrary to BP, arbitrary subjects in Finnish are overt (21) and “generic pronouns can, **and must, be null**” (HOLMBERG, 2005:540, except when the verb appears in first position and an overt expletive must be inserted (HOLMBERG, 2010), as in (22a):

(21) Ne puhuu Itävallassa saksaa.  
 they speak Austria-INE German  
 ‘*They speak German in Austria.*’

(22) a. \*Istuu mukavasti tässä.  
       sits comfortably here  
 b. Tässä istuu mukavasti.  
       here sits comfortably  
 c. Sitä istuu mukavasti tässä.  
       EXP sits comfortably here  
       ‘*One sits comfortably here.*’

and (c) the same seems to be true for non-thematic predicates and impersonal sentences (HOLMBERG and NIKANNE 2002):

(23) Ulkona sataa.  
 outside rains  
 ‘*It’s raining outside*’

(24) a. \*leikkii lapsia kadulla.  
       play children in-street  
 b. Sitä leikkii lapsia kadulla.  
       EXP play children in-street  
 c. Kadulla leikkii lapsia.  
       in-street play children  
 d. Lapsia leikkii kadulla.  
       children play in-street  
 ‘*Children are playing in the street./ There are children playing in the street.*’

The use of the expletive in Finnish does not seem ‘optional, as emphasized by the authors, because in the absence of a DP or an adverbial constituent in first position it becomes obligatory, which suggests a sort of V2 grammar. BP has no lexical expletive, and, although, as we have shown, it tends to avoid a V-initial sentence, we still attest a

---

<sup>13</sup> If we were considering Brazilian standard writing, we would, of course, have another picture. Writing does recover, partially, obsolescent forms, mixing EP grammar, BP grammar and exhibiting forms that belong to neither grammar (DUARTE, 2020b).

very strong competition between null expletives in V1 sentences and raising of referential constituents to first position. Examples (14a), (15a), (16a), (17a) show that there is no restriction to V initial sentences in BP. Besides, the apparent free order with intransitive verbs shown for Finnish in (24) is not possible in BP, a system that still allows VS with unaccusatives in declaratives and Wh-interrogatives if the internal argument is a lexical DP (DUARTE, 1992, among others), but not with inergatives or transitives.

## TO CONCLUDE

To sum up, we must remember Biberauer's (2010: 154) claim about 'partial' null subject languages: "recent research has shown that it is extremely unlikely that languages under this label constitute a single (sub)type of languages. They are extremely unequal, allowing null referential subjects under very restricted circumstances". And we would add that the methodology used for the different analyses is completely different. While data from BP come from empirical research and show the propagation of the change, allowing one to observe "reflexes of grammars in patterns of language change", using Kroch's (1989) words, descriptions for other languages seem to come from intuition, which is good but not sufficient to offer a clearer and revealing comparison. The impression one has is that they are stable grammars, which is not true for BP. The same seems to be true for languages included in the group of 'consistent' null subject languages. Roberts and Holmberg (2010: 12) recognize that the differences presented "to establish that partial null-subject languages have a range of properties distinguishing them from consistent null-subject languages", but they also remember that "it is very likely that many languages that have been seen as 'consistent' null-subject languages are in fact partial null-subject languages" (see DUARTE and SOARES DA SILVA, 2016 for a gradient distribution of some 'consistent' null subject languages). The emergence of overt non-animate/non-specific personal pronouns in European Portuguese, even though extremely rare, seems to give support to Roberts and Holmberg's hypothesis. More research based on speech should be crucial to offer more realistic descriptions. As for writing, it is well known that it is not a very reliable source for the information we seek because of its conservative character and resistance to implement innovative features, even though contemporary Brazilian Portuguese writing shows a mixture of both.

---

## REFERENCES

- AVELAR, J. O. Gramática, competição e padrões de variação: casos com *ter/haver* e *de/em* no português brasileiro. *Revista de Estudos da Linguagem*, 4. 2006. p.99-144
- AVELAR, J. De verbo funcional a verbo substantivo: uma hipótese para a supressão de *HAVER* no português brasileiro. *Letras de Hoje*, Porto Alegre: PUC-RS, v. 143, 2006. p. 49-74.
- AVELAR, J.; CALLOU, D. Sobre a emergência do verbo possessivo em contextos existenciais no português brasileiro. In: CASTILHO, A. et al. (Orgs.) *Descrição, História e aquisição do português brasileiro*. Campinas: Pontes, 2007. p. 375-402.
- AVELAR, J.; GALVES, C. Tópico e concordância em português brasileiro e português europeu. In: COSTA, Armanda; FALÉ, Isabel; BARBOSA, Pilar (Orgs.) *Textos Seleccionados: Actas do XXVI Encontro Nacional da Associação Portuguesa de Linguística*. Porto: APL, 2011. p.49-65.
- BACELAR DO NASCIMENTO, M. F; MENDES, A.; DUARTE, M. E. L. Sobre formas de tratamento do português europeu e brasileiro. *Diadorim: Revista de Estudos Linguísticos e Literários*, UFRJ, v 20, n. Especial, 2018. p. 245-262.
- BARBOSA, P.; DUARTE, M. E. L.; KATO, M. A. Null Subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, 2005. p. 11-52

- BIBERAUER, T. Semi pro-drop languages, expletives and expletive pro reconsidered. In: BIBERAUER, T. et al. (eds.) *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press. 2010. p. 153-199.
- CALABRESE, A. Pronomina: some properties of the Italian pronominal system. In: FUKUI, N.; RAPAPORT, T.; SAGEY, E. (eds.) *MIT Working Papers in Linguistics*, 8, 1-46.1986.
- CALLOU, D.; AVELAR, J. Sobre *ter* e *haver* em construções existenciais: variação e mudança no português do Brasil. *Revista Gragoatá*, 9, 2000.
- CARDINALETTI, A. Towards a cartography of subject positions. In: L. Rizzi (ed.) *The Structure of CP and IP: the cartography of syntactic structures, vol 2.*, New York: Oxford University Press, 2004. p. 115-165.
- CYRINO, S. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: Roberts, I.; KATO, M.A. (Orgs.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Pontes, 1993. p. 163-184.
- CYRINO, S.; DUARTE, M. E. L.; KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (orgs.) *Brazilian Portuguese and the Null Subject*. Frankfurt am Main: Vervuert Verlag, 2000. p. 55-73.
- DUARTE, M. E. L. A perda da ordem V(erbo) S(ujeito) em interrogativas-q no português do Brasil. *D.E.L.T.A.*, v. 8, n. Especial, 1992, 37-52.
- DUARTE, M. E. L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: *Português Brasileiro: uma viagem diacrônica*. Campinas: Pontes, 1993. 107-128.
- DUARTE, M. E. L. *A perda do princípio "Evite Pronome" no português brasileiro*. Tese (Doutorado em Linguística), Instituto de Estudos Linguísticos, UNICAMP, Campinas, 1995.
- DUARTE, M. E. L. On the 'embedding' of a syntactic change. In: GUNNARSSON, B-L. (ed.) *Language Variation in Europe: Papers from the Second International Conference on Language Variation in Europe, ICLaVE 2*, Uppsala University, Sweden, June, 2003. Uppsala: Universitetstrifkeriet, 2004. p.145-155.
- DUARTE, M. E. L. Analyzing a parametric change in Brazilian Portuguese: a sociolinguistic investigation. In: BARBOSA, P.; PAIVA, M. C.; RODRIGUES, C. (eds.). *Studies on variation in Portuguese*, v. 14. John Benjamins, 2017. p. 234-253
- DUARTE, M. E. L. O sujeito nulo referencial no português brasileiro e no português europeu. In: GALVES, C.; KATO, M. A.; ROBERTS, I. (orgs.) *Português Brasileiro: uma segunda viagem diacrônica*. Campinas, Ed. da Unicamp, 2019. p.93-126.
- DUARTE, M. E. L. Sociolinguistics as a powerful tool to follow the course of a parametric change. *University of Pennsylvania Working Papers in Linguistics*, 25 (2): 2020a, p. 9-19.
- DUARTE, M. E. L. A sintaxe do português brasileiro: entre a fala espontânea e a escrita padrão. In: SALOMÃO, S. (ed.) *Temas da Língua Portuguesa: do pluricentrismo à didática*. Roma: Edizione Nuova Cultura. 2020b, p. 131-152,
- DUARTE, M. E. L.; SOARES DA SILVA, H. Microparametric variation in Spanish and Portuguese: the null subject parameter and the role of the inflectional paradigm. In: KATO, M. A; ORDOÑEZ, F. (eds.) *The morphosyntax of Portuguese and Spanish in Latin America – Oxford studies in comparative syntax*. Oxford: Oxford University Press, 2016. p.1-26.
- DUARTE, M. E. L.; FERNANDES, U. S. B. Construções de tópico-sujeito em contexto de variação e mudança. *Web-Revista Sociodialeto*, v. 6, 2016. p. 347-371
- DUARTE, M. E. L.; REZENDE DOS REIS, E, P. Revisitando o sujeito nulo vinte anos depois. *ReVEL – Revista Virtual de Estudos da Linguagem*, vol. 16, n. 30, 2018. p. 173-197.
- EGERLAND, V. Impersonal pronouns in Scandinavian and Romance. *Working Papers in Scandinavian Syntax* 71, 2003. p. 75-102.
- HOLMBERG, A. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry*, 36, 2005. p. 533-564.
- HOLMBERG, A. The null generic pronoun in Finish: a case of incorporation. In: T. BIBERAUER, et al. (eds.) In: *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: CUP, 2010. p. 200-230.
- HOLMBERG, A.; NIKANNE, U. Expletive, Subjects and Topics in Finish. In: SVENONIOUS, P. (ed.) *Subjects, expletives and the EPP*. Oxford: OUP, 2002. p. 1-45.
- HOLMBERG, A.; NAYUDU, A; SHEEHAN, M. Three partial null subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63, 2007. p. 59-97
- HOLMBERG, A.; SHEEHAN, M. Control into finite cluses in partial null subject languages. In: T. BIBERAUER et al. (eds). *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press, 2010. p. 125-152.
- HOLMBERG, A.; PHIMSAWAT, On-Usa. Truly minimal pronouns. *Diadorim: Revista de Estudos Linguísticos e Literários*, UFRJ, vol. 19, n. Especial, 11-36, 2017.

- KATO, M. A. Strong and weak pronouns in the null subject parameter. *PROBUS*, v. 11(1), 1999. p. 11-38.
- KATO, M. A.; DUARTE, M. E. L. *Semantic and phonological constraints on the distribution of null subjects in BP*. Paper presented at NAWVE32, University of Pennsylvania, October, 2003.
- KATO, M. A.; DUARTE, M. E. L. *Mudança paramétrica e orientação para o discurso*. Paper presented at the 24th meeting of the Portuguese Association of Linguistics, University of Minho, Braga, November, 2008.
- KATO, M. A.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Revista Veredas*, 18 (1), 2014. p. 1-22.
- KATO, M. A.; DUARTE, M. E. L. O sujeito no português brasileiro e sua tipologia. In: E. Pilati; H. Salles; R. Naves (Orgs.) *Novos Olhares para a gramática do português brasileiro*. Campinas: Pontes, 2017. p. 13-42.
- KATO, M. A.; DUARTE, M. E. L. Pre-verbal position in BP: a reinterpretation of the ‘avoid pronoun’ principle. *Diadorim: Revista de Estudos Linguísticos e Literários*, UFRJ, vol. 20, n. Especial, 2018. p. 610-626.
- KÍSS, K. E. The EPP in a Topic prominent language. In: SVENONIUS, P. (org.) *Subjects, Expletives and the EPP*. New York: Oxford University Press, 2002. p. 107-124
- KROCH, A. Reflexes of grammar in patterns of language change. In: *Language Variation and Change*, 1, 1989. p. 199-244.
- LABOV, W. *Principles of Linguistic Change*. Oxford, UK/Cambridge, MA.: Blackwell. 1994.
- LI, C. N.; THOMPSON, S. Subject and topic: a new typology of language. In: LI, C. N. (org.). *Subject and Topic*. New York: Academic Press Inc, 1976. p. 457-489.
- LOPES, C. R. S.; BROCARD, M. T. Main Morphosyntactic Changes and Grammaticalization Processes. In: WETZELS, L.; COSTA, J.; MENUZZI, S. (Eds.). *The Handbook of Portuguese Linguistics*. Maldon: John Wiley; Sons, Inc, 2016. p. 471-486.
- LUNGUINHO, M. V.; MEDEIROS JR. P. Inventou um novo tipo de sujeito: características sintáticas e semânticas de uma estratégia de indeterminação do sujeito no português brasileiro. *Revista de Estudos em Língua e Literatura*, v. 23, 2009. p. 147-191.
- MARINS, J. *O Parâmetro do Sujeito Nulo: uma análise contrastiva entre o português e o italiano*. Dissertação (Mestrado em Letras Vernáculas), UFRJ, Rio de Janeiro, 2009.
- MARINS, J. *As repercussões da remarcação dos parâmetros do Sujeito Nulo: um estudo diacrônico das sentenças existenciais com ter e haver no PB e no PE*. Tese (Doutorado em Letras Vernáculas), UFRJ, Rio de Janeiro, 2013.
- MARINS, J. *Contextos de resistência de haver em sentenças existenciais no PB: o tempo verbal e o traço semântico do argumento interno*. Paper presented in the I Forum Internacional de Sociolinguística, Rio de Janeiro, UFRJ, november, 2019.
- MARINS, J.; SOARES DA SILVA, H.; DUARTE, M. E. L. *Revisiting DUARTE (1995): for a gradient analysis of indeterminate subjects in Brazilian Portuguese*. Paper presented in the conference of the working group of Theory of Grammar, Belo Horizonte, UFMG, July, 2015.
- MARINS, J.; SOARES DA SILVA, H.; DUARTE, M. E. L. Revisiting DUARTE (1995): for a gradient analysis of indeterminate subjects in Brazilian Portuguese. In: *Diadorim: Revista de Estudos Linguísticos e Literários*, UFRJ, v. 19, n. Especial, 2017. p. 140-172.
- MARTÍNEZ-SANZ, C. *Null and overt subjects in a variable system: the case of Dominican Spanish*. Tese (Doutorado). Universidade de Ottawa, Canadá, 2011.
- MODESTO, Marcelo. Topic prominence and null subjects. In: T. Biberauer (ed.) *The limits of syntactic variation*. Amsterdam: John Benjamins. 2008. p.375-406.
- PONTES, E. *O tópico no Português do Brasil*. Campinas: Pontes. 1987.
- ROBERTS, I. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer. 1993.
- ROBERTS, I; HOLMBERG, A. Introduction: parameters in Minimalist theory. In: T. BIBERAUER et al. (eds) *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press. 2010, 1-57.
- SAAB, A. On the notion of partial (non) *pro*-drop in Romance. In: KATO, M. A.; ORDOÑES, F. (Eds.) *The morphosyntax of Portuguese and Spanish in Latin America – Oxford studies in comparative syntax*. Oxford: Oxford University Press. 2016. p.49-77.
- SANKOFF, D.; TAGLIAMONTE, S; SMITH. E. Goldvarb X: *A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto. 2005. Available for download from: <http://individual.utoronto.ca/tagliamonte/goldvarb.html>.
- TORIBIO, J. Dialectal Variation in the licensing of null referential expletive subjects. In: PARODI, C.; QUICOLI, C. A.; SALTARELLI, M.; ZUBIZARRETA, M. L. (Eds.) *Aspects of Romance Linguistics*. Washington, DC: Georgetown University Press, 1996. p. 409-432.

VANCE, B. *Null subjects and syntactic change in medieval French*. Tese (Doutorado), Cornell University, Nova York, 1989.

Recebido: 10/11/2020

Aceito: 1/6/2021

Publicado: 19/8/2021