

Revisiting Duarte (1995): for a Gradient Analysis of Indeterminate Subject in Brazilian Portuguese

Juliana Marins^a

Humberto Soares da Silva^b

Maria Eugenia Lammoglia Duarte^c

ABSTRACT

Recent studies on the representation of pronominal subjects in Brazilian Portuguese (BP) show the preference for overt indeterminate referential subjects. A “side effect” of the change in the Null Subject Parameter in BP is the progressive decline in the use of the standard strategies - structures with verb in the third person plural with a null subject and sentences with the pronominal generic clitic *se* - and the appearance of alternative strategies – an increasing frequency of sentences with overt nominative pronouns, especially *você* (you) and *a gente* (the people =we). Our purpose is to refine Duarte’s (1995) data of indeterminate subjects, collected from a sample of the speech of graduates from Rio de Janeiro. Our theoretical framework associates the Theory of Language Variation and Change (WLH, 2006 [1968]) with the Principles and Parameters Theory (CHOMSKY, 1981, 1995), which guides our analysis, from the hypotheses raised to the interpretation of the empirical data (DUARTE, 2016). Our results show that the different strategies are not in direct competition: they can be distributed in three different groups, according to a set of features they share, with respect to arbitrary and generic reference. At one extreme we find [+3rd person/+plural] category, **which excludes the speaker**, represented by the dying arbitrary clitic *se* and the pronoun *eles* (they), preferably overt. At the other extreme, we have a [+3rd person/+singular] category, **which may or may not include the speaker and the addressee**, represented by the generic clitic *se*, the zero strategy (with a 3rd person singular verb form) and *você* (you), which is preferably overt. Finally, we have a [+1st person/+plural] category, **which does include the speaker**, represented by *nós* (we) and *a gente* (the people, the folks=we), with considerable advantage with respect to the former. The variation in each category disposed along our scale is not a stable phenomenon: each point has a strong competitor to represent each degree of indeterminate reference as the change progresses.

KEYWORDS: indeterminate subjects; arbitrary reference; generic reference; Null Subject Parameter

RESUMO

Estudos recentes sobre a representação do sujeito pronominal no português brasileiro (PB) mostram a preferência por sujeitos de referência indeterminada foneticamente representados. Um “efeito colateral”

^a Universidade Federal do Rio de Janeiro – juespmarins@hotmail.com

^b Universidade Federal do Rio de Janeiro – humba6@gmail.com

^c Universidade Federal do Rio de Janeiro – eugenia@brazilmail.com

This research has the support of CNPq, proc. 307158/2014-4.

das mudanças envolvendo a remarcação do Parâmetro do Sujeito Nulo no PB é o declínio progressivo no uso das estratégias tidas como padrão – estruturas com o verbo na terceira pessoa do plural com sujeito nulo e sentenças com o clítico pronominal *se* – e o surgimento de estratégias alternativas, com a frequência crescente do uso do pronome nominativo expresso, especialmente *você* e *a gente*. Nosso objetivo é refinar a análise dos dados de sujeitos indeterminados levantados em Duarte (1995), com base em amostra da fala culta do Rio de Janeiro. Nosso quadro teórico associa a Teoria da Variação e Mudança (WLH, 2006 [1968]) com a Teoria de Princípios e Parâmetros (CHOMSKY, 1981), que guia nossa análise, desde levantamento de hipóteses até a interpretação dos dados empíricos (DUARTE, 2016). Nossos resultados revelam que as diferentes estratégias não estão em competição: elas se distribuem em três grupos, onde ocorre a competição entre as estratégias que os compõem, de acordo com o conjunto de traços que compartilham, o que está relacionado com o caráter arbitrário ou genérico de cada estratégia. Em um extremo dessa escala, encontramos a categoria [+3a pessoa/+plural], **que exclui o falante**, representada pelo agonizante clítico *se_{arb}* e pelo pronome *eles*, preferencialmente pleno. No outro extremo, temos a categoria [+3a pessoa/+singular], representada pelo clítico *se_{gen}*, pela estratégia zero (com o verbo na 3ª pessoa do singular) e *você*, que é preferencialmente pleno, **podendo incluir ou não o falante e o ouvinte**. Finalmente, temos a categoria [+1a pessoa/+plural], **que inclui necessariamente o falante**, representada por *nós* e *a gente*, com considerável vantagem de *a gente* sobre *nós*. A variação em cada categoria disposta ao longo da escala não é um fenômeno estável: em cada ponto há um forte competidor para representar cada grau de referência indeterminada, à medida que a mudança avança. PALAVRAS-CHAVE: sujeitos indeterminados; referência arbitrária; referência genérica; Parâmetro do Sujeito Nulo; variação morfossintática

Introduction

In the past three decades, several papers on the representation of definite reference pronominal subjects in Brazilian Portuguese (BP) – cf. Duarte (1993, 1995, 2012), among others – have shown that BP prefers the structural position of the subject filled by a pronoun, as shown in the examples in (1) for first and second person and in (2) for third person with [+/-human] referents.

- (1) a. Mesmo que **eu** não fizesse o pré-vestibular, **eu** acho
 even that I not take-Past.Subj the preparatory-course I think
 que **eu** passaria por causa da base que **eu** tinha.
 that I would-pass because of-the base that I had
 ‘Even if I had not taken the preparatory course, I think I would have
 passed because of the base I had.’

- b. **você** me disse que **você** está morando em copacabana.
 you me-cl told that you are living in copacabana
 ‘you told me that you are living in copacabana.’
- (2) a. [**Essa minha tia**]_i que mora aqui, **ela**_i é solteirona e eu acho que **ela**_i é
 This my aunt who lives here, she is spinster and I think that she is
 super-feliz, sabe? (...) **Ela**_i é uma pessoa que ajuda os outros pra caramba
 super-happy, know-2SG? She is a person that helps the others a lot.
Ela_i ficou solteira porque **ela**_i quis.
 She stayed single because she wanted
 ‘This aunt of mine who lives here, (she) is a spinster and I think she is
 very happy. She is a person who helps everyone. She remained single
 because she wanted to.’
- b. [**a casa**]_i virou um filme quando **ela**_i teve de ir abaixo.
 the house became a movie when she had of go down
 ‘The house became a movie when it had to be demolished.’

Cases of null referential subjects are restricted to certain structural contexts, such as control structures, with the antecedent as a subject in the previous clause, as in (3), or in an adjacent sentence, as in (4) – cf. Modesto (2004; 2008) –, but they are always in variation with overt subjects, which already outnumber null subjects (cf. DUARTE, 2015):

- (3) a. mas **ele**_i sentiu [que \emptyset _i era o único novo ali, recém-casado]
 but he felt that was the only young there, newly-married
 ‘But he felt he was the only young guy there, newlymarried.’
- b. [**Esse filme**]_i emocionou muita gente [quando \emptyset _i ficou pronto]
 That film touched many people when was ready
 ‘That film touched many people when it was shown.’
- (4) a. **Geisel**_i era realmente um poderoso presidente da república.
 Geisel_i was really a powerful president of-the republic.
 \emptyset _i Demitiu ministros militares...
 fired ministers militaries
 ‘Geisel was in fact a powerful president. He even fired military
 ministers...’

- b. [O Rio de Janeiro]_i é uma beleza! Realmente Ø_i é uma cidade linda
 The Rio de Janeiro is a beauty! Really is a city beautiful
 ‘Rio de Janeiro is a beauty. Really! It’s a beautiful city.’

Such facts have been associated to a change in the setting of the Null Subject Parameter (NSP) in BP, as a consequence of changes in the system of nominative pronouns, which triggered a significant reduction in the verbal inflexional paradigm (see DUARTE, 1993).

Following this path, researches on the representation of indeterminate reference subjects (DUARTE, 1995, CAVALCANTE, 1999, 2007; COUTO 2004, VARGAS 2012), including what we refer today as arbitrary and generic reference (EGERLAND, 2003; HOLMBERG, 2005 among others), propose that, as a “side effect” of the changes in the NSP in BP, the system would have come up with alternative strategies to the ones described by the grammatical tradition, which are restricted to structures with the verb in the third person plural with a null indeterminate subject and sentences with the pronominal “indefinite” clitic *se*, the same strategies characterizing European Portuguese (EP). In fact, these studies have demonstrated a progressive decline in the use of standard strategies and an increasing frequency of sentences with an overt nominative pronoun, especially *você* (you)¹ and *a gente* (the people = we), expanding their use for second and first person plural reference, respectively, to express generic reference, as shown in (5):

- (5) a. Quando **você** é menor, **você** não dá muito valor a essas coisas.
 When you are little you not give much value to these things
 ‘When you are a child you do not value such things.’
- b. mas na época **a gente** não podia acreditar... **a gente** não acreditava
 but at-the time the people not could believe the people not believed
 nisso porque **a gente** era novo.
 in-that because the people was young
 ‘but then we/one could not believe that... we/one did not believe that
 because we were young.’

¹ Regions where *tu* is preferred over *você* to refer to 2nd person also prefer *tu* for generic reference. Even in Rio de Janeiro, where both pronouns are in variation, generic *tu* is usual, being attested in popular varieties. (DUARTE, 2003).

On the other hand, a new sentence pattern related to indeterminate reference, first noticed by Galves (1987) and later confirmed by other studies (DUARTE, 1995; CAVALCANTE 2007; LUNGUINHO and MEDEIROS JUNIOR, 2009, among others)² would exhibit the use of a null subject with the verb in the third person singular, illustrated in (6):

- (6) a. \emptyset_{gen} não pode entrar de sapato; \emptyset_{gen} fica de meia.
 not can.3sg.pres enter of shoe; stay of sock
 ‘One cannot enter (the museum) with shoes on; one keeps only socks.’
- b. \emptyset_{gen} Não usa mais máquina de escrever.
 not use.3SG.PRES more machine of to-write
 ‘No one uses typewriters anymore.’
- c. \emptyset_{arb} Construiu um Bob’s na Tijuca.
 built.3SG.PAST a Bob’s in-the Tijuca
 ‘They built another Bob’s in Tijuca.’ ‘Another Bob’s has been built in Tijuca.’
- d. \emptyset_{arb} Vai reformar o Centro todinho por causa da
 is-going.3ps.pres to-reform the center whole because of-the
 Olimpíada.
 Olympic-Games
 They are going to rebuild all downtown because of the Olympic Games.’

Lunguinho and Medeiros Jr. (2009), in particular, associate the implementation of such strategy in BP to three facts: (a) changes in the pronominal system, with the consequent weakening of subject/verb overt agreement, as already mentioned; (b) the progressive loss of pronominal third person clitics, especially the reflexives (D’ALBUQUERQUE, 1984)³; and (c) the rigidity of the order SVO. VS order is

² This null generic subject, here referred as “zero strategy”, has been considered a property of partial null subject languages being referred to as “null generic subject” (HOLMBERG, NAYUDU and SHEEHAN, 2009; HOLMBERG (2010). Notice, in addition, that examples (6c) and (6d), pointed out by Lunguinho and Medeiros (2009), in fact, convey arbitrary reference, contrary to the null generic subjects in (6a) and (6b). They resemble a simplified version of a passive structure.

³ Lunguinho and Medeiros Jr (2009) mention the loss of the system of reflexive clitics; however, this is a process of change observed in some regions of the country, which prefer a null reflexive, including

restricted to sentences with unaccusative structures, including unaccusative verbs and lexical passives; in both cases, however, the internal argument bearing nominative case is, in fact, interpreted as a complement, which is evidenced by the lack of verbal agreement.

Taking into account all these changes, the aim of this paper is (a) to show the relation between the changes in the NSP in BP towards an overt subject grammar and the rise of new strategies of subject “indetermination”, considering all the changes related to this new configuration of the system (change in the pronominal system, reduction in the number-person distinctive morphemes in the inflectional paradigm, and the weakening of subject/verb agreement); (b) to evaluate the productivity of the third person singular strategy (which will be referred to as zero strategy as in Duarte (1995) in a system that clearly prefers overt subjects, bearing in mind that its emergence seems to be more recent, and to explore possible motivations for its origin⁴⁵; and (c) to try to dispose the strategies attested along a *continuum*, which has arbitrary reference at one end, and different degrees of generic reference till the other end of the *continuum*. In this way, we will show that they are not in variation as a single set; rather, they have specialized to convey different degrees of indetermination. And, even in each point of this *continuum*, the strategies in competition are morphosyntactic doublets, in the sense of Kroch (1994), following Aronoff (1976), which means that one form is bound to disappear.

Our theoretical framework associates the Theory of Language Variation and Change (WEINREICH; LABOV; HERZOG, 1968) and the Principles and Parameters Theory (CHOMSKY, 1981, 1995) – this necessary grammatical component guides our analysis from the hypothesis raised to the interpretation of the empirical data. This, in turn, will allow us to find answers for the “embedding” problem, crucial to the model of language change used here.

As Duarte’s (1995) empirical research is probably the first to relate overt indeterminate subjects with overt referential subjects (see DUARTE’s chapter 3), our analysis will revisit and review her data in the light of recent discussions concerning the

argumental ones; third person accusative and dative clitics as well as arbitrary/generic clitic *se*, on the other hand, are in process of extinction in spontaneous speech in all regions investigated – see D’Abuquerque (1984) and Duarte and Ramos (2015).

⁴ This strategy has been referred to as a “null generic pronoun” (HOLMBERG, 2005, and subsequent work). As we will show, it can convey generic or arbitrary meaning in BP.

⁵ See Galves’s (1987) hypothesis, according to which it was exactly the impossibility to identify a definite third person null subject that allowed a generic interpretation for the null subject.

NSP. Her analysis was based on a sample comprising 12 graduated speakers, recorded in 1992, stratified for age and gender, as part of the Norma Urbana Culta – Rio de Janeiro (NURC-RJ) *corpus*. The statistical analysis used the Goldvarb X Program (TAGLIAMONTE, 2006). Also, a qualitative analysis was also necessary to identify the discursive contexts, which could be favoring the specialization of the attested indetermination strategies. Therefore, besides extending the data in this “revisit” to Duarte’s sample, discursive factors have also been considered to refine the analysis.

1. The change in the representation of indetermination: relevant facts

1.1. The restructuring of the pronominal system in BP and the changes in the inflectional verbal paradigm

BP experienced a significant restructuring of its pronominal system of nominative forms particularly triggered by grammaticalization processes undergone by the originally nominal address form *Vossa Mercê* – Your Grace – Your Mercy, and the collective DP *a gente* (literally “the people”). Both entered the system as fully grammaticalized pronouns for second (*você* = you) and first person plural reference (*a gente* = we), giving rise to a competition with the former pronouns *tu* (2ps.) and *nós* (1pp.), as shown in Chart 1, adapted from Duarte (1993):

	Pronouns	Paradigm1 19th Century	Paradigm 2 20th Century/1	Paradigm 3 20th Century/2
1PS	eu	estudo	estudo	estudo
1PP	<u>nós</u>	estudamos	estudamos	<u>estudamos</u>
	a gente	-	estuda	estuda
2PS	tu	estudas	estudas	estuda(s)
	você	-	estuda	estuda
2PP	vós	estudais	-	-
	vocês	-	estudam	estuda(m)
3PS	ele, ela	estuda	estuda	estuda
3PP	eles, elas	estudam	estudam	estuda(m)

Table 1. Nominative Pronouns and inflectional paradigms in Brazilian Portuguese (verb “to study” – *estudar*)

It is clear that BP, in earlier stages, exhibited, according to written language (popular plays and letters) six different verbal morphemes. From the 20th century on, with the complete replacement of *vós* (2pp) by *vocês*, the competition between *tu* and *você* (2ps) and between *nós* and *a gente* (1pp), the inflectional paradigm was gradually reduced to four, three and sometimes two oppositions (see DUARTE and SOARES da SILVA, 2016, for a comparison of Romance inflectional paradigms and their role in the expression of definite reference subjects). In the first place, former nominal expressions combine with third person, unmarked verb form; in addition, phonological processes have affected the second person singular morpheme <-s> associated with the verb form combined with *tu* (you) and second and third plural <-m>, the latter being subject to variation according to years of school attendance and monitoring of speech (SCHERRE and DUARTE, 2016).

According to Duarte (1993), until the first quarter of the 20th century, *você* still preserved its original value of an address form and was rarely used in theater plays. Such a situation is confirmed by Souza (2012), who analyzes a sample of letters written in Rio de Janeiro between 1870 and 1970: the pronoun *você* was initially less productive than the pronoun *tu*. Around the 1930s the two pronouns reach similar rates of use, but, more important, they were not in variation, since each form had a correspondent distinct use: *tu* was used in familiar contexts and *você* maintained its original mark of courtesy. From the 1930s on, the use of *você* outnumbers the use of *tu*, which starts to decrease, at least in the samples analyzed⁶⁷. In addition, the complementary distribution is lost: *você* starts to be used in the same contexts as *tu*, an evidence of the fully grammaticalization of *você* as a real second person pronoun.

Considering Chart 1 and Souza's results, we can say that the history of PB exhibits three subsystems for second person reference: one with the pronoun *tu*; another in which *tu* and *você* coexist; and a third one, in which *você* is the predominant form (see LOPES and CAVALCANTE (2011). More recently, Scherre et al. (2015) identify 5 subsystems, refining the type of coexistence of *você* and *tu* in different areas of the country.

With respect to the variation between the innovative pronoun *a gente* and the conservative pronoun *nós* for first person plural reference, we can find a less advanced

⁶ Other address forms have been introduced in the system, such as *o(a) senhor(a)* (the sir/the madam), *Vossa Excelência* (Your excellency), whose use is each time more restricted.

⁷ In fact, the pronoun *tu* has not been abandoned in the metropolitan area of Rio de Janeiro. Presently, its use is expanding to more central areas of the city (see PAREDES SILVA, 2003).

process than the one shown for the representation of second person. But the propagation of the change proceeds steadily across age groups, particularly in the function of subject, as pointed out by Lopes (2007, p. 68), “the replacement of *nós* by *a gente* is progressively being carried out in the last 30 years in Rio de Janeiro, not only among graduated (cultured) speakers, but also among non-cultured speakers”⁸. As shown by Callou and Lopes (2004), *a gente* is already implemented in the system since the 1970s⁹, when it outnumbers the overall rates of *nós* in the speech of graduates in Rio de Janeiro, even though older groups still show a competition between the two pronouns. Duarte’s 1992 sample for graduates from Rio de Janeiro reveals the progress of the change: the older group (over 56 years of age) still prefers *nós* (62%), the intermediate group (36-55) reaches 50%, and the younger group (25-35), only 2%. There is no doubt that *a gente* is the new 1st person plural pronoun in spoken BP today – a mapping of this phenomenon in the country can be found in Vianna and Lopes (2015).

1.2. Resetting the Null Subject Parameter

Duarte’s (1993) diachronic research examining theater plays comprising the 19th and the 20th centuries, distributed in seven periods, presents evidence of a parametric change in process. PB is clearly moving from null to overt referential subjects.

Figure 1, adapted from Duarte (1993), shows the progressive decrease of null subjects in the corpus analyzed:

⁸ Our translation from the original: “a substituição de *nós* por *a gente* se está efetivando progressivamente nos últimos 30 anos, seja entre os falantes cultos, seja entre os não-cultos no Rio de Janeiro” (LOPES, 2007, p.68).

⁹ This study was based on a sample of popular speech and college-graduated speakers which belong to PEUL and NURC-RJ projects, respectively; speakers were recorded in two different moments: 1970s and 1990s (NURC), available at: www.lettras.ufrj.br/nurc-rj, and 1980s and 2000s (PEUL), available at: www.lettras.ufrj.br/peul.

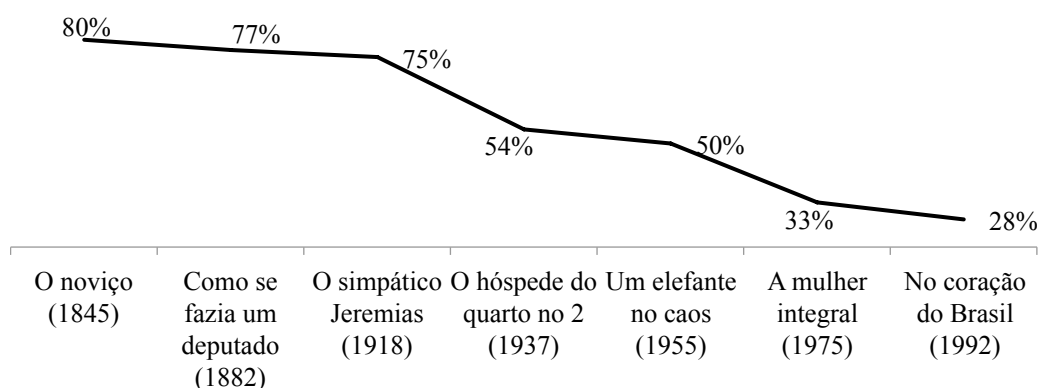


Fig. 1: Null subjects in theater plays along seven synchronies (19th and 20th centuries)¹⁰

Comparing the results in Figure 1 with the three paradigms illustrated in Chart 1, we can visualize three stages in the process of implementation of the change towards overt pronominal subjects: during the first three synchronies, which correspond to Paradigm 1, covering the 19th century and the first quarter of the 20th century – to which the author refers to as a “formally rich” paradigm, using Jaeggli and Safir’s (1989) proposal –, we see a consistent null subject grammar. The next two synchronies coincide with the first significant reduction of the verbal morphology, a consequence of the grammaticalization of *você*, which affects the identification of null subjects. However, null subjects are still identified, in a system with a “functionally rich” paradigm, with a reasonable number of ending distinctions and syncretisms (ROBERTS, 1992). In the last quarter of the century, the insertion of pronoun *a gente*, fully grammaticalized and combining with the unmarked third person verb morphology (zero), would bring about the complete loss of the “Avoid Pronoun” principle, which characterizes Romance Null Subject Languages (CHOMSKY, 1981; RIZZI, 1982).

It is important to add that the change affected first and second person faster than third person; therefore, licensing of null subjects is no longer related to a distinctive ending. Subsequent investigations would show the role played by the semantic feature [+human] in this process (CYRINO, DUARTE and KATO, 2000); it explained the quick implementation of overt first and second person subjects, inherently [+human], and also third person subjects bearing this feature – Duarte, Mourão and Santos (2012), re-examining an enlarged sample of theater plays, confirm that. Subjects bearing [-

¹⁰ Each period is represented by one play (indicated, in the chart, by its title and year of publication or first performance), written in Rio de Janeiro, by very popular playwrights born in the city.

human] feature undergo the change more slowly, but overt non-human subjects already compete with overt ones, a feature that does not fit a null subject Romance language. Besides the semantic feature, a null third person subject is more constrained by co-reference with the subject of the main clause or the presence of an identical antecedent in an adjacent clause bearing the same function – such contexts are illustrated in (3) and (4) in the Introduction to this article. This would be taken as an evidence of topic prominence in allowing null subjects in BP (MODESTO, 2008; HOLMBERG, NAYUDU and SHEEHAN, 2007; HOLMBERG and SHEEHAN, 2010). This is another structural factor to delay the fully implementation of the change in External-Language data. Synchronic samples, however, confirm that overt subjects already outnumber null subjects in every structural context (DUARTE, MOURÃO and SANTOS, 2012). Since indeterminate subjects are inherently [+human], one would expect a side effect of the change in progress in favor of overt indeterminate pronominal subjects. This is what is shown by Duarte's (1995; 2003) synchronic analyses and Vargas' (2012) diachronic study, which motivate the present article, as will be shown in the next sections.

2. Analysis

2.1. The gradation between arbitrary subjects and generic subjects

Up to this point, we have been dealing with subject indetermination strategies without taking into account possible discursive differences, just as it has been made by the researchers who studied subject indetermination using the Generative Theory (DUARTE, 1995; CAVALCANTE, 1999, 2007; COUTO 2004; VARGAS 2012). The referred papers were interested in the preference for overt pronominal subjects to null subjects rather than to establish any kind of distinction among the several possibilities. However, Duarte (1995) already mentions the zero strategy (sentences with the verb in the third person singular and null subject; see note 5) as constrained by the expression of aspect, modality and procedure discourse, which is confirmed by Cavalcante (2007). A refinement of such constraint would be offered by Lunguinho and Medeiros Jr. (2009), who add another possible reading to sentences with the verb in the third person singular and null subject (the zero strategy) to the ones already pointed out. The already mentioned meanings would receive the label “generic” (7a) and “episodic” (7b), which is not an instance of generic subject but could be included in the arbitrary group –

already instantiated in (6c-d) – and would be associated to the tense of the verb¹¹.

- (7) a. Ø Não tá contratando gente ainda não.
 not is hiring people still not
 ‘They are not hiring workers yet.’ ‘Workers are not been hired yet.’
- b. Ø Montou o armário lá de casa semana passada.
 set-up.3SG.PAST the wardrobe there of home week passed
 ‘They set up our wardrobe last week.’ ‘Our wardrobe was set up last week.’

Our concern in this article is not only to observe the zero strategy but the set of indetermination forms so as to find out their discursive motivations and their possible distribution along a gradient between two poles: arbitrary reference and generic reference.¹² By *arbitrary reference*, we mean subjects that cover a finite set of entities or an individual, whose identification cannot be individualized in the discourse, as we can see in (8) below:

- (8) a. Uma vez Ø_{arb} me definiram Austrália como a Inglaterra de
 one time me.CL.1SG defined.3PL Australia as an England of
 bermudas (...) bermudas
 ‘Once they defined Australia to me as England in bermuda shorts.’
- b. Então, (..) Ø_{arb} criou-se em 1920 ou 21, a Universidade do Brasil
 then created.SE.CL in 1920 or 21, the University-of-the Brazil
 ‘Then, in 1920, 1921, the University of Brasil was created.’

¹¹ The authors consider that the distinction is not associated to verb tense only, but with aspect as well, which is not new. In this paper, we will show that the aspect of the verb is really a factor that contributes to characterize the different nuances of subject indetermination.

¹² As far as we can recall, the use of such distinction (which has its origin in several semantic studies in the second half of the 20th Century) to refer to arbitrary and generic indetermination has become common since Holmberg (2005).

- c. Quando você termina o ciclo básico você faz outro vestibular.
 when you finish the cycle basic you tke another examination
 Mas não é um vestibular, né? **Ø_{arb} vai** pelo CR,
 but not is an entrance-examination, right? go.PRES.3PS by-the GPA
Ø_{arb} classifica pelo CR.
 classify.PRES.3PS by-the GPA
 ‘When you finish the basic course you go through another sort of
 entrance examination, see? But it’s not a real examination, they rank the
 student by the GPA.’

As shown by the examples in (8), although the subjects highlighted in the sentences cannot be determined by any element in the text/discourse, it is possible to establish a limit for their scope. In (8a), the speaker does not make it clear who defined Australia as “England in bermuda shorts”, but it is possible to infer that there was someone who said that¹³. In (8b), even if they do not tell who created the University of Brazil, there was a person – or an institution – who did it, also limiting the scope of the non-defined subject character. Finally, in (8c), it is interesting to notice the established contrast between the reference of the pronoun *você*, which appears as the subject of the clauses in the previous sentence, and the null subjects of the third person singular verbs in bold that follow. In this case, the speaker refers to the way the institution selects the students who should go from the basic cycle to the professional cycle, without identifying the agent. This makes the reference of the null subject of *ir* (to go) and *classificar* (to classify) more limited, although not precise.

At the opposite pole, we identify the **generic reference** subjects, which correspond to those whose reference cannot be determined by the text/discourse, being unlimited, in the sense that it can refer to an infinite set of entities. The examples in (9) show that:

¹³ It is also possible that this consideration might have been made by more than one person, but even so, there is a finite group of entities.

- (9) a. Mas agora \emptyset_{gen} não se tem mais inverno. Você vê que hoje
 but now not SE.CL has more winter. You see that today
 já estamos no inverno e tá um calor....
 already are.1PL in-the winter and is a heat
 ‘Nowadays one does not have winter any more. You see that we are
 already in the winter and it is so hot.’
- b. E tinha sorveteira que batia em casa. Tinha uma
 and there-was ice-creamer that beat.3SG.PAST at home. Had a
 hélice assim (...). \emptyset_{gen} Fazia sorvete em casa.
 propeller like-this (...) used-to-make.3SG icecream at home
 ‘And there was an appliance to make ice cream at home. There was a
 propeller (...) One used to make ice-cream at home.’
- c. Lá na América também eu cozinhava porque... lá você
 there in-the America also I cooked because...there you
 encontrava tudo pronto ou semi-pronto. Você punha as coisas
 found everything ready or semi-ready. You put.PAST the things
 no forno aquilo ia aparecendo a comida.
 in-the oven that would-go appearing the food
 ‘In the USA I also cooked because... there you already found everything
 ready or semi-ready. You would put the things in the oven and the food
 began do appear.’

In the three cases illustrated in (9), the subjects/agents cannot be identified and show an unlimited reference: in (9a), the winter is no longer noticed by *any carioca* or even by *anyone* who stays sometime or lives in Rio de Janeiro; in (9b), any person could make ice cream at home; and in (9c), the pronoun *você* (you), which does not allude to the interlocutor in this case, refers to anyone who was in America at that time and, therefore, could find ready-made food and put it in the oven.

It is interesting to notice that sentences with the indeterminate pronominal clitic *se* and with verbs in the third person singular and a null subject can convey both arbitrary and generic meaning. What seems to determine one or the other possibility is the interaction of the feature [+3rd person] with the aspect of the verb, as suggested by Lunguinho and Medeiros Jr. (2009). If the verb is perfective, the idea obtained will be

arbitrary, as illustrated by the zero strategy in (8a, b). On the other hand, the interaction with the imperfective aspect, illustrated in (9a, b), provides a generic reading of the subject, similar to the one obtained with *você* in (9c) and in the two first sentences in (8c) above.

In intermediate points of the gradient proposed here, there are subjects whose reference has a certain degree of specificity, which cannot be determined by the text/discourse, although it applies to an infinite set of entities, as in (10):

- (10) a. O que eles vão fazer depois é outra história, né? Brizola tá fazendo coisa pra caramba, a gente tá vendo que ele tá fazendo.
 what they are-going to-do after is another story, right? Brizola is doing thing too much the people (we) are seeing that he is doing
 ‘What they are going to do later is another story, but Brizola is doing a lot of things, we can see it.’
- b. (...) o Rio de Janeiro é uma cidade violenta, todo mundo diz. Nós ficamos nos enganando, dizendo: Ah, Nova York tem também violência
 (...) the Rio de Janeiro is a city violent, every world says. We keep us.CL.REFLEX.1PL deceiving, saying: Well New York has also violence
 ‘... Rio de Janeiro is a violent city, everyone says so. We keep deceiving ourselves, saying: Well, New York also has violence’
- c. (...) a situação mais formal é quando você tem que tratar, com pessoas que hierarquicamente estão, acima, né? Então, se eu vou falar com o Senador, que se eu vou conversar com o Governador, eu tenho que tratar de uma forma, Vossa Excelência e tal (...)
 the situation more formal is when you have to deal with people that hierarchically are above, right? Then if I go to-talk with the senator that if I go to-talk with the Governor I have that to address of a way Your Excellency and so
 ‘The most formal situation is when you have to address people who are hierachically above you.... Then, if I’m talking to the Senator, to the Governor, I have to use a different address form, such as Your Excellency.’

Analyzing the examples in (10), we see that the reference of the subjects highlighted applies to an infinite set of entities, as shown in the examples in (9). The difference resides in the fact that in (10a) and (10b), the pronouns *a gente* (the people) and *nós* (we) include the speaker in the group of possible entities, which becomes the reference more specific. The same can be said about the example (10c), in which the pronoun *eu* (I), which associates the features [+1st person/+singular], used by the speaker to raise a hypothesis about the way any person should address another hierarchically superior using himself as an example, makes the reference even more specific.

It is understood, thereby, that the label “indeterminate subject” comprises a wider range of references, considered here as a scale determined by two opposite poles: the arbitrary reference and the generic reference. Between them, there would be at least two points in which the references that include necessarily the speaker would be located. It is also possible to think that the pronominal strategies to indicate indeterminacy of the subject in BP could be defined by a set of features – combining person and number – producing different forms to cover them. That is the idea defended in this paper. With respect to the interaction of these features and the possible forms of indetermination, we can, therefore, assume that the arbitrary interpretation is conveyed by the presence of the features [+3rd person/+plural] and it seems to be blocked by the features [+1st person/+2nd person]. When the strategy presents the feature [+3rd person], the interaction with the aspect of the verb seems to determine the arbitrary or the generic reference.

2.2. Data analysis related to reference: positioning the strategies on the scale

2.2.1. The standard strategies

Grammatical tradition includes both the third person plural with a null subject and the use of the clitic *se* as the subject indetermination mechanisms (the influence of European Portuguese is obvious). Lunquinho and Medeiros Jr. (2009) suggest, however, that these two strategies would not be interchangeable anymore having become specialized in the system: the third person plural would be associated to the episodic reference and the clitic *se*, to a more generic reference, in the authors' words. We assume that, in some period of BP history, the two strategies could have both types of reference (which is in fact shown in some empirical studies, as in Cavalcante's 1999

- b. Nessa rua que eu moro atualmente, **eles** construíram uma série de in-this street that I live now they built a series of prédios novos, com essas varandinhas que tão na moda. buildings new with these little-balconies that are in-the fashion ‘In the street I live now, they have built a number of new buildings with these small balconies so in fashion.’

A qualitative analysis of the structures with the clitic *se* showed an interesting result, shown in Table 2:

	Arbitrary	Generic	Total
Age group 3	1 occurrence – 4%	24 occurrences – 96%	25 – 100%
Age group 2	5 occurrences – 42%	7 occurrences – 58%	12 – 100%
Age group 1	–	2 occurrences – 100%	2 – 100%

Table 2: Distribution of clitic *se* according to reference in “apparent time”

Table 2 attests the unproductivity of clitic *se* to express arbitrary reference (hereafter referred as *se_{arb}*): there were no occurrences of this type in the speech of the younger group; the six cases attested in the sample occur in the speech of the intermediate age group – 5 occurrences - and only one in the speech of the older group. Such instances of *se_{arb}* can alternate with arbitrary third person plural strategy with the perfective past of the indicative tense (12a) or the perfective past reading of the indicative present form (12b).

- (12) a. Os produtos não vinham por causa da guerra. \emptyset_{arb} **Montou-se** the produces not came because of-the war. Created.3PS-SE.CL às pressas algumas indústrias de alimento, de roupas, de calçados pra quickly some industries of food, of clothes, of shoes to substituir esse produto que a classe brasileira importava. replace this product that the class Brazilian imported ‘The commodities were scarce because of the war. Industries were quickly created to supply the market with food, clothes and shoes imported by Brazil.’
- b. Então é todo um momento histórico muito importante no qual then is all a moment historical very important in-the which

\emptyset_{arb} **se cria** o partido comunista brasileiro (...)

SE.CL create.3SG the party communist Brazilian

‘Then it is a very important historical moment, in which the Brazilian communist party is created.’

On the other hand, the 33 occurrences exhibiting se_{gen} data convey the imperfective aspect, as we can see in (13):

- (13) a. A greve pra mim [...] \emptyset_{gen} só **deve se deflagrar** quando não
the strike for me [...] only should SE.CL deflagrate when not
houver mais recurso nenhum. Mas é um recurso válido que \emptyset_{gen} **se**
there-is more resort none. But is a resort valid that SE.CL
usa.

use.3SG.PRES.

‘In my opinion, the strike should be used as a last resort. One may only spark a strike when there is no appeal anymore. But it’s a valid resort one can use.’

- b. Quando eu era criança, \emptyset_{gen} **punha-se** a mesa pra tomar lanche.
When I was child set.3SG.PAST-SE.CL the table to have snack
‘When I was a child one would set the table to have an afternoon snack.’

As we have shown, the data does confirm the specialization of the two subject indetermination strategies regarded as standard: the strategy that bears the features [+3rd person/+plural] has specialized to represent arbitrary reference whereas the clitic *se*, bearing the features [+3rd person/+singular], clearly specializes for the generic reference; the rare uses of se_{arb} confirm that they are residual. This indicates that both strategies are no longer in variation, possibly constituting a case of complementary distribution in BP. However, we must not forget that even se_{gen} is disappearing in the speech of younger generations, as shown in Figure 2 and in Table 1.

2.2.2. The non-standard strategies

Considering that third person plural represents arbitrary reference and that clitic se_{gen} decreases significantly, we should expect the rise of an alternative form to represent this

type of reference. And also considering that third person clitic paradigm in BP (including accusative, dative and “indefinite” clitics) is clearly disappearing, one would expect overt nominative pronouns to represent indeterminate subjects. However, taking into account our approach, the new strategies would fit in different categories located in one point of our scale: (a) [+3rd person/+plural] category represented by *pro* + 3rd person plural verb or overt pronoun *eles* (they); (b) [+3rd person/+singular] category, represented by *se*, *pro* + 3rd person singular verb and *você* (you); and finally (c) [+1st person] and [+singular] or [+plural] category, represented by *eu* (I), *nós* (we) and *a gente* (the people = we). This makes it clear that we do not consider the different strategies as a single set. The idea is that indetermination, generally speaking, can be represented by the three categories characterized by a combination of person and number features mentioned in (a)-(c) and that the different strategies belonging to each category are in fact in variation. Considering that each group encodes different information from the ones that do not exhibit the same set of features, we start analyzing sentences with *pro* and verb in the third person singular – hereby referred to as *zero* strategy – and the pronoun *você*, in the face of the use of the conservative clitic *se_{gen}*.

- **The [+3rd person/+singular] group**

Similarly to what happens to the clitic *se* (as shown in Table 1), the zero strategy can express both arbitrary and generic reference. From the 40 occurrences of zero, 32 conveyed generic reference and only 8, arbitrary reference (see (8c) above). On the contrary, the use of *você* – in 387 occurrences – is always associated with generic reference, which may be explained by the presence of the [+2nd person] feature present in the original address form that underwent grammaticalization as mentioned in 1.1. Figure 3 presents the distribution of the three strategies with generic reference – *se*, zero and *você* –, through the three age groups.

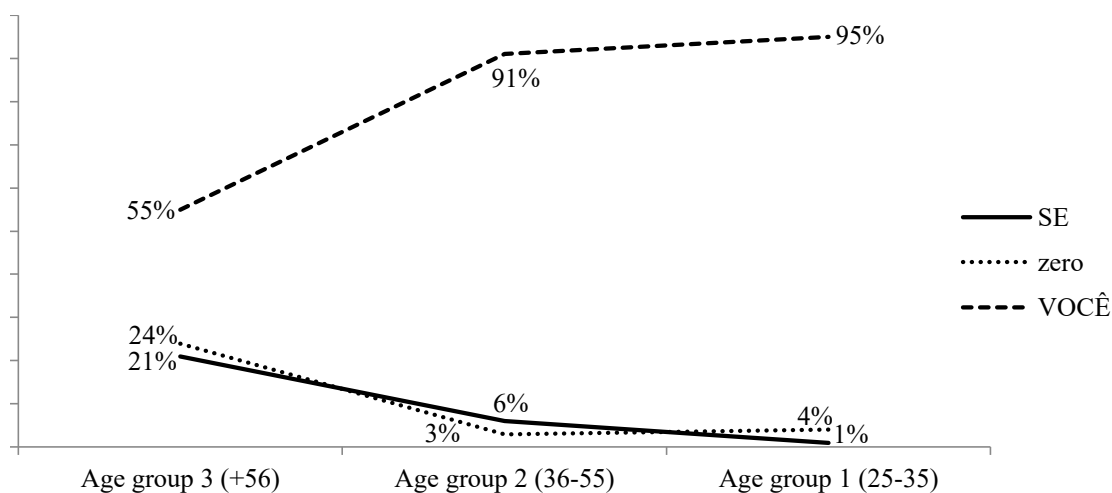


Fig. 3: Distribution of generic *se*, zero and *você* in “apparent time”

The distribution clearly reveals that the preferred strategy to replace *se_{gen}* is *você* and not the zero strategy, as suggested by recent analyses. Both *se* and zero show a significant decline, ranging from 21% and 24%, respectively, in the speech of the older group to an average of 4,5% and 2,5% among the speakers in the two younger groups. *Você*, on the other hand, with the highest rates among all speakers in the sample, shows an expressive increase across the three generations, reaching 95% in the speech of age group 1.

To explain such scenario, we should turn our attention to the possible origins of the zero strategy for subject indetermination. It is possible that, at some point in the evolution of the ongoing change, the identity between “synthetic” passive voice sentences and active voice sentences, both with the clitic *se*, associated with the reduction of the clitic paradigm, have favored the emergence of the zero strategy¹⁴. On the other hand, this situation would not fit a system that progressively presents clear preference for overt subjects. Therefore, it is possible that the expansion of the use of *você* is the solution the system has found to assure the generic interpretation and to avoid the empty subject position. Considering that *você* presents the second person feature, the arbitrary interpretation would be blocked, being restricted to the zero strategy, taking into account the interaction with the verbal aspect.

With respect to the context in which the form *se* still resists in the speech of older

¹⁴ We take Cavalcante’s (2007: 72) Table 3, which opposes only the zero and the *se* strategies (disregarding *você*), as an empirical evidence for this. Such table shows that the zero strategy rates increase in apparent time, replacing *se*, which decreases from older to younger speakers, an expected result. This does not mean that it outnumbers generic *você*

generations, the transitivity of the verb seems relevant. All occurrences of *se* (39 tokens) – either generic or arbitrary – exhibit a transitive verb, selecting an accusative or an oblique complement, as we can see in (14):

- (14) a. \emptyset_{gen} **Jogava-se** futebol na Rua Visconde Silva
 used-to-play.3SG-SE.CL soccer in-the Street Visconde Silva
 ‘One used to play soccer on Visconde Silva St.’
- b. Você colocava as suas jóias pra ir trabalhar, pra ir passear, pra ir
 you put-on the your jewelry to go to-work, to go for-a-walk, to go
 à rua, normalmente, sem problema nenhum. Nem \emptyset_{gen} **se**
 to-the street normally, without problem any. Not SE.CL
pensava em roubo
 used-to-think in theft
 ‘You put on your jewelry to go to work, to take a walk, to go out, without
 any kind of fear. One would not worry about theft.’

Taking into account the perspective we are adopting, based on Lunguinho and Medeiros Jr. (2009), the expectation would be that the tokens with the zero form would also occur with “direct” transitive verbs, as described by the authors. Therefore, if the zero strategy occurs only with “direct” transitive verbs, *você* would cover the other kinds of verbs. And it is what we find: considering the 40 tokens of the zero strategy with both types of reference, 36 appear with transitive verbs, most of them “direct” transitive, but there were still some occurrences of the zero form with transitive verbs which select an oblique internal argument, as in (15):

- (15) a. Tinha orquestra, \emptyset_{gen} não **precisava** daquele pra tocar.
 had.3SG orchestra, not needed.3SG of-that to play
 ‘There was an orchestra, one did not need that to play.’
- b. E as casas eram muito generosas de espaço, dava pra você
 and the houses used-to-be very generous of space, gave for you
 fazer festas, quando terminava tudo, \emptyset_{gen} **chegava** na cozinha,
 to-make parties when finished all, arrived.3SG in-the kitchen,
 ainda tinha feito uma areazinha pra lavar roupa
 still had sort-of a little-area to wash clothes

‘And the houses used to be more spacious, you could make parties (...); after the kitchen, there was still a little space serving as a laundry’

Contrary to the expectations, however, we found four instances with generic zero and non-transitive verbs, as exemplified with the copula in (16).

- (16) Depende, entendeu? Porque eu tô namorando há quatro anos e meio.
 depends, see? Because I am dating for four years and a-half
 E aí \emptyset_{gen} fica meio dependente de namorado.
 and so gets.3SG sort-of dependent of boyfriend
 ‘It depends. Because I’ve been dating for four years and a half. Therefore, you get sort of dependent on your boyfriend.’

The situation described seems to reinforce our claim: the strategy with *se* is progressively replaced by zero and, especially, by *você*; the few occurrences of zero show a transitive verb. On the other hand, *você* spreads in the system, appearing with all types of verbs.

- **[+1st person] [+/-plural]**

As we said above, the feature [+1st person] seems to block the arbitrary reading. All the 134 tokens with indeterminate subjects bearing this feature, receive generic interpretation, as we can see in (17):

- (17) a. Quando **a gente** levanta a coisa é que **a gente** vê tudo que
 when the people investigate the thing is that the people see all that
 aconteceu
 happened
 ‘When investigate it all, we understand all that really happened,’
 b. **Nós** ficamos nos enganando, dizendo: Ah, Nova York
 we keep us.CL.REFLEX.1PL deceiving saying: Well, New York
 tem também violência.
 has also violence

- ‘We keep deceiving ourselves, saying: Well, New York also has violence.’
- c. É muito caro vir para o Rio de Janeiro. (...) Agora o que is very expensive to-come to the Rio de Janeiro Now what vamos fazer?
go.1PL to-do?
Ø_{gen} **Vou** ver a Amazônia, tem que pegar um avião,
go.1SG to-see the Amazon, have.3SG that to-take a plane, quarto horas e tal
four hours and so
‘It’s very expensive to visit Rio de Janeiro. (...) What are we going to do then? I want to go to the Amazon, you must take a plane. It’s about a four-hour flight away.’

We can notice that *a gente* and *nós* share the same set of features ([+1st person], [+plural]), belonging to the same group or category. As expected, this competition, showed in Figure 4, is bound to disappear, with the victory of the innovative strategy, following exactly the same tendency observed for definite 1st person plural (LOPES, 1993, 2003, among others).

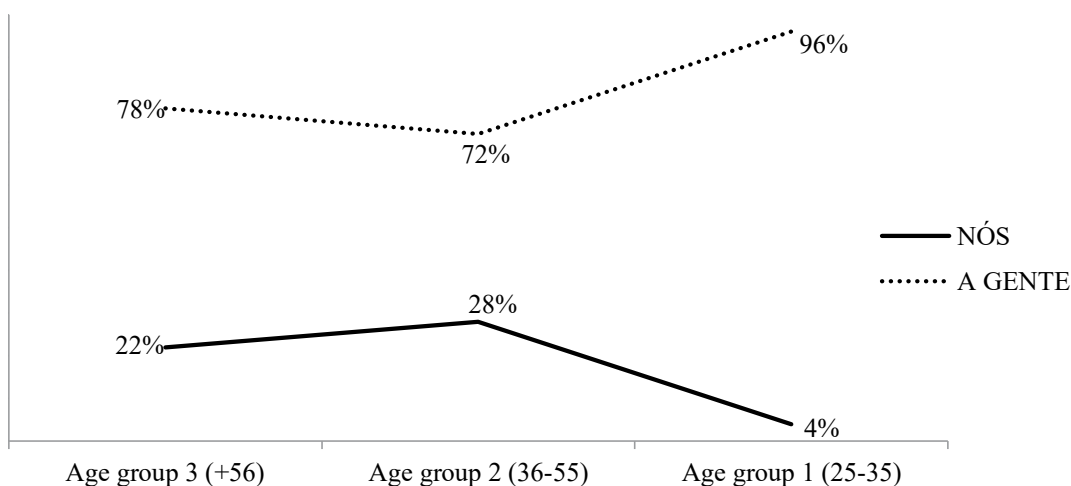


Fig. 4. Evolution of first person plural strategies in “apparent time”

As we can observe, the use of *a gente*, although completely implemented in the speech of all three age groups, already reaches 96% in the speech of the younger group.

This result confirms what was already attested by Duarte (1995), both for definite and indeterminate reference subjects, and by Vargas (2012), for indeterminate subjects along the 19th and the 20th centuries. This suggests that pronoun *nós* is in process of extinction in spontaneous speech.¹⁵

As for the use of [+1st person], [+singular], materialized by the pronoun *eu*, it cannot be explored here because we found only 4 tokens, all of them in the older age group, shown in (10c) and (17c). Therefore, it is not possible to make any sort of generalization about the use of this strategy, although the data show its generic character, with a higher degree of specificity than the first person strategies that combine with [+plural] feature.

Even though based on four tokens, 1st person singular *eu* seems to be used in situations in which the speaker offers an example to the proposed situation in the discourse, making the experience even more real, showing it could be lived by any person, including himself/herself, therefore limiting the generic reference.

Summarizing what has been said up to this point, we have proposed that BP has moved from a stage in which there were four possibilities to represent indeterminate pronominal subjects to a stage with seven possibilities. The strategies usually presented to express indeterminacy do not constitute a group nor can they be explained by the distinction between arbitrary *versus* generic reference only. We have a first stage with different categories disposed along a *continuum* as shown in Diagram 1, when BP exhibited four strategies (phase 1). As innovative strategies entered the system – *a gente*, *você* and zero –, what we find is variation in some points of the scale (phase 2). Our data suggest that in each point one of the forms will replace its competitors.

¹⁵ Writing recovers the clitic *se* and the 1st person plural *nós* (usually null) for indetermination (COUTO, 1994; DUARTE, 2007).

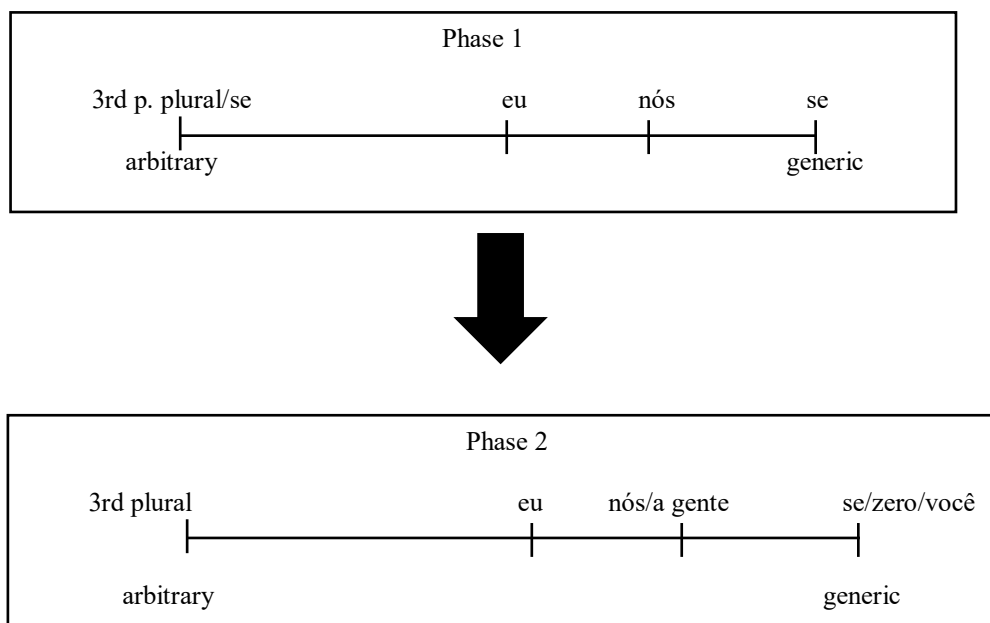


Diagram 1. Change in the representation of subject indetermination in two stages

Therefore, if our predictions are correct, it is possible that, in the future, BP will return to a reduced set of strategies for indetermination (phase 3), as illustrated in Diagram 2:

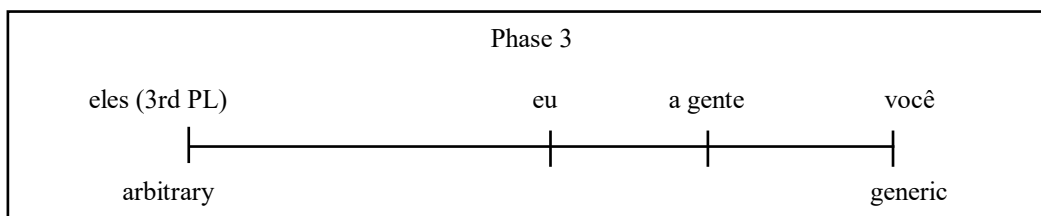


Diagram 2. Possible set of forms for subject indetermination in the future

2.3. Overt non-defined reference subjects: effects of the changes in the NSP

Another relevant factor concerning the observation of the data is the representation of the non-defined reference subjects. Duarte (1995) had already pointed out that, similarly to what happens to definite reference subjects, the “indeterminates” also presented higher rates of overt subjects (65% against 35% of null pronouns), especially among speakers in the younger group. In our revision of Duarte’s sample for the present analysis, we attested 84% of overt pronouns, an even more expressive percentage than that shown by Duarte

(1995).¹⁶ Table 2 exhibits the distribution of occurrences according to the strategy, including 3rd person plural (they):

Age Groups	G3 (+56)	G2 (36 -55)	G1 (25-35)
<i>Eles</i> (they)	13/58(22%)	19/26 (73%)	23/37 (62%)
<i>Você</i> (you)	59/65 (91%)	73/81 (90%)	223/241 (93%)
<i>Nós</i> (we)	9/12 (75%)	3/5 (60%)	2/2 (100%)
<i>A gente</i> (the people=we)	33/44 (75%)	17/17 (100%)	47/50 (94%)

Table 2. Overt pronouns (vs. total of occurrences) according to strategy and age group

Figure 5 shows the distribution of overt versus null pronominal strategies (*você*, *nós* and *a gente*), across the three generations:

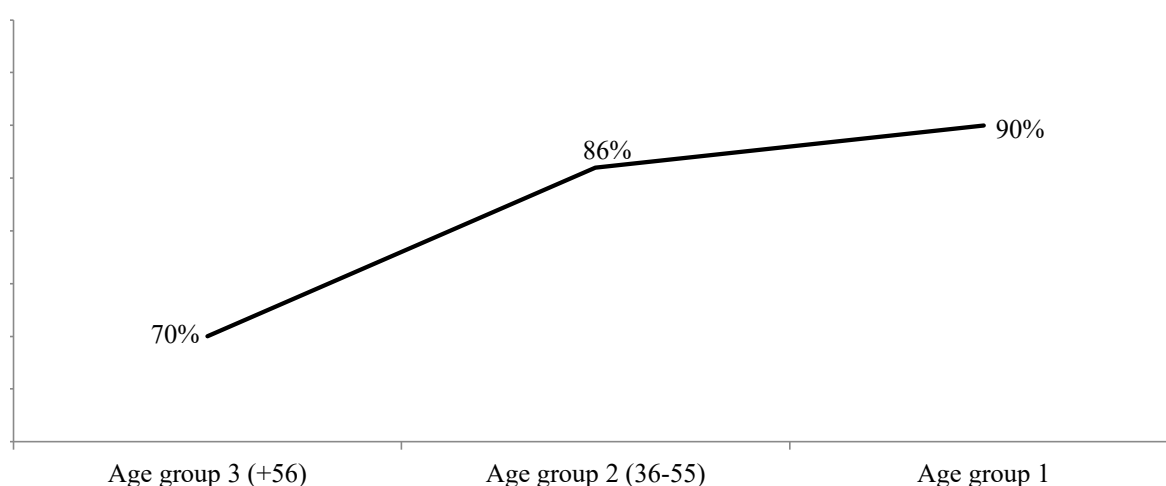


Fig. 5. Overt pronominal strategies (vs null) in apparent time

The percentages show that, in the early 1990s, not only overt subjects were already preferred to null pronouns, but also that the process of change in the speech of the younger group was close to completion. While for group 3 the mean percentage of overt subjects is 70% – 117 occurrences in a total of 167 –, the younger group reaches 90% – with 295 occurrences out of 330.

Figure 6 represents the results in Table 2 for overt pronouns. With respect to the category combining the features [+3rd person], [+plural], we can see the most significant change: overt 3rd person subjects, with 22% in the speech of speakers in the older group, rises to 73% and 62% among the speakers of the intermediate and younger groups.

¹⁶ Duarte (1995) did not include the clitic *se* as an overt strategy as we do in this analysis.

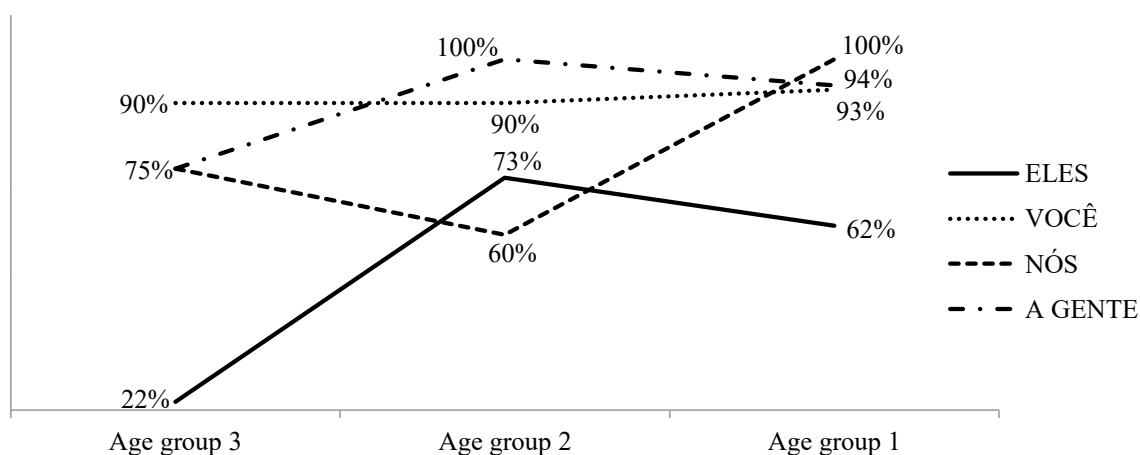


Fig. 6. Distribution of overt subjects in “apparent time”

As for the combination [+3rd person/+singular], we can see that overt *você* is already predominant in all age groups, with a percentage around 90%. Not only does this confirm Duarte’s (1995) results, but it also constitutes an important argument in favor of the victory of *você* as the strategy for [+3rd person/+singular] group over *se* and zero strategies, as previously mentioned. Therefore, it seems natural that *você* stands as the principal strategy in the extreme of the proposed gradient.

Finally, the combination of features [+1st person/+plural], with the competition between *nós* and *a gente*, presents an interesting picture. The pronoun *a gente*, which, as seen before, is replacing *nós*, for definite and indefinite reference, presents an increase in the rates of overt subjects – from 75% among the older group to 100% and 94% in the speech of intermediate and younger groups, respectively. Considering the use of *nós*, one can say that, in spite of its overt expression being preferred, reaching 100% in Group 1, our sample attested only 19 occurrences of such strategy (shown in Table 2) and the younger group shows only two instances of this pronoun, attesting its almost complete extinction in spontaneous speech.

3. Final considerations

This article aimed to contribute to studies concerning the representation of indeterminate subjects in BP, trying to relate the change between the reset of the NSP in this variety of Portuguese to the appearance of new strategies – preferably overt – as alternatives to the conservative forms described in traditional grammars and consistent with European Portuguese, which has been taken as a model to still current descriptions (see CARDOSO, 2017, for the diachrony of indetermination strategies in European Portuguese). Our

analysis made it possible to show that the several strategies under the label “subject indetermination” reveal different types of reference, which led us to dispose them along a scale or a *continuum*: in one pole, we find arbitrary reference, which corresponds to an indeterminate subject in the space of the discourse, but, in a way, limited to a set of entities, and, in the opposite pole, we have the generic reference, which also corresponds to an indeterminate subject in the space of the discourse, but whose scope is unlimited. In intermediate points in this scale, we find indeterminate subjects, including necessarily the speaker, whose scope can be wider or narrower, being closer to one pole or the other.

Through the revision of the corpus previously analyzed by Duarte (1995), consisting of the speech of graduates born in Rio de Janeiro, distributed in three age groups, in a pioneer empirical research bringing evidence of the resetting of the NSP in BP, we have seen that the different strategies attested are not in direct competition: it was possible to propose three different groups, with three strategies would be in competition in each one, defined by a set of features. One of them, characterized by [+3rd person/+plural], is represented by the pronoun *eles* (they), null or overt, and the clitic “se”. The [+3rd person/+singular] cluster of features contains the strategies with the clitic *se*, the zero strategy (3rd person singular) and *você* (you), which, in the analyzed sample, is almost categorically overt. Still concerning this group of features, it was interesting to notice that the obsolescence of the clitic *se* seems to have opened the way to the forms zero and *você*, but, differently from recent researches suggest, the zero strategy neither appears as an emerging resource, nor represents a strong candidate to replace *você*. The reason for that lies exactly the fact that the zero strategy decreases among the younger groups and this fact is consistent with a system moving towards overt pronominal subjects. It would be surprising to find it moving to an opposite direction. This zero strategy shows a more restricted use, being constrained with respect to aspect, modality and transitive verbs, which can explain its marginality in the system. Finally, the [+1st person/+plural] group can be represented by the strategies *nós* and *a gente*, with considerable advantage for the latter, which presents rates over 70% in the three age groups. The pronoun *nós* rarely appears, but when it does, it is predominantly overt.

In sum, the variation existing in each group or category disposed along our scale is not a stable phenomenon. In fact, each point has a strong competitor to represent each degree of indeterminate reference as the change in progress moves towards completion.

References

- ARONOFF, Mark. *Word formation in generative grammar*. MIT Press, 1976.
- CALLOU, Dinah; LOPES, Célia. “Contribuições da sociolinguística para o ensino e a pesquisa: a questão da variação e mudança linguística.” *Revista do GELNE*, 5 (1-2), p.63-74, 2004.
- CARDOSO, Marianna. *Estratégias de indeterminação em peças brasileiras e portuguesas: uma análise contrastiva*. Master Thesis – FL, UFRJ, Rio de Janeiro, 2017.
- CAVALCANTE, Silvia. *A indeterminação do sujeito na escrita padrão: a imprensa carioca dos séculos XIX e XX*. 1999. Master Thesis – FL, UFRJ, Rio de Janeiro, 1999.
- CAVALCANTE, Silvia. “O sujeito nulo de referência indeterminada na fala culta carioca”. *Diadorim: Revista de Estudos Linguísticos e Literários*. Rio de Janeiro, v.2, p. 63-82, 2007.
- CHOMSKY, Noam. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.
- CHOMSKY, Noam. *The Minimalist Program*. Cambridge, Mass: The MIT Press, 1995.
- COUTO, Marco Aurélio de Souza. *Estratégias pronominais de indeterminação do agente*. Master Thesis – FL, UFRJ, Rio de Janeiro, 2004.
- CYRINO, Sônia; DUARTE, Maria Eugênia L.; KATO, Mary. A. “Visible subjects and invisible clitics in Brazilian Portuguese”. In: KATO, M. A.; NEGRÃO, E. (eds.). *Brazilian Portuguese and the Null Subject Parameter*. Madri: Iberoamericana/Frankfurt: Vervuert, 2000, p. 55-73.
- D’ALBUQUERQUE, Alair da Cruz Reis Cavalcanti. “A perda dos clíticos num dialeto mineiro.” *Sociolinguística e ensino do vernáculo*, 78/79, Rio de Janeiro, Tempo Brasileiro, p. 97-120, 1984.
- DUARTE, Maria Eugênia L. “Do pronome nulo ao pronome pleno: a trajetória do sujeito nulo no português do Brasil”. In: ROBERTS, I.; KATO, M. A. (eds.). *Português brasileiro: uma viagem diacrônica*. Campinas: UNICAMP, p. 107-128, 1993.
- DUARTE, Maria Eugênia L. *A perda do princípio Evite Pronome no português brasileiro*. 1995. PhD Dissertation – IEL, UNICAMP, Campinas, 1995.
- DUARTE, Maria Eugênia L. “A evolução na representação do sujeito pronominal em dois tempos”. In: PAIVA, M. C.; DUARTE, M. E. L. (eds.). *Mudança linguística em tempo real*. Rio de Janeiro: ContraCapa/FAPERJ, p. 115-128, 2003.
- DUARTE, Maria Eugênia L. “Sujeitos de referência definida e abstrata: aspectos *Diadorim, Rio de Janeiro, Revista 19 — Volume Especial 2017*.”

- conservadores e inovadores na escrita padrão.” *Revista Linguística*, UFRJ, v. 3, n. 1, 88-115, 2007.
- DUARTE, Maria Eugênia L. (ed.). *O sujeito em peças de teatro (1833-1992): estudos diacrônicos*. São Paulo: Parábola, 2012.
- DUARTE, Maria Eugênia L. “Avanços no estudo da mudança sintática associando a Teoria da Variação e Mudança e a Teoria de Princípios e Parâmetros”. *Cadernos de Estudos Linguísticos*, Unicamp, v. 57, n. 1, 85-111, 2015.
- DUARTE, Maria Eugênia L. “Sociolinguística ‘Paramétrica’”. In MOLLICA, M. C.; FERRAREZI JR. C. (eds.) *Sociolinguística, sociolinguísticas – uma introdução*. São Paulo: Contexto, 2016, p. 33-44.
- DUARTE, M. Eugênia L.; MOURÃO, Gabriela; SANTOS, Heitor. “Os sujeitos de terceira pessoa: revisitando Duarte 1993”. In: DUARTE, M. E. L. (ed.) *O sujeito nulo em peças de teatro (1843-1992): estudos diacrônicos*. São Paulo: Parábola Editorial. 2012, pp. 21-44.
- DUARTE, M. Eugênia L.; SOARES da SILVA, Humberto. “Microparametric variation in Spanish and Portuguese: The Null Subject Parameter and the role of the verb inflectional paradigm.” In KATO, M. A.; ORDOÑEZ, F. (eds.) *The Morphosyntax of Portuguese and Spanish in Latin America*, p. 1-26, 2016.
- EGERLAND, Verner. “Impersonal pronouns in Scandinavian and Romance.” *Working Papers in Scandinavian Syntax*, 71: 75-102, 2003.
- GALVES, C. “A sintaxe do português brasileiro”. *Ensaio de Linguística*, 13, 31-49, 1987.
- HOLMBERG, Anders. “Is there a little *pro*? Evidence from Finnish.” *Linguistic Inquiry*, 36, 533-564, 2005.
- HOLMBERG, Anders. “The null generic pronoun in Finnish: a case of incorporation.” In: BIBERAUER, T. et al. (eds.) *Parametric Variation: Null Subjects in Minimalist Theory*. Cambridge: CUP. p: 200-230, 2010.
- HOLMBERG, Anders; NAYUDU, Aarti; SHEEHAN, Michelle. “Three partial null subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi.” *Studia Linguistica*, 63, 59-97, 2007.
- HOLMBERG, Anders; SCHEEHAN, M. Control into finite clauses in partial null subject languages. In T. BIBERAUER et al. (eds) *Parametric Variation: null subjects in Minimalist theory*. Cambridge: Cambridge University Press. 125-152, 2010.

- JAEGGLI, Osvaldo; SAFIR, K. “The Null Subject Parameter and Parametric Theory”. In: JAEGGLI; SAFIR (eds.). *The Null Subject Parameter*. Dordrecht: Kluwer, 1989, p. 1-44.
- KROCH, Anthony. “Morphosyntactic Variation”. In BEALS, K. et al. (eds.) *Papers from the 30th Regional Meeting of the Chicago Linguistic Society v. 2: The Parasession on Variation in Linguistic Theory*. 1994, 1-23.
- LABOV, William. *Principles of Linguistic Change: Internal Factors*. Cambridge MA: Blackwell, 1994.
- LOPES, Célia. “Nós por a gente no português falado culto.” In: *Anais do II Congresso da ASSEL/Rio*. Rio de Janeiro: Faculdade de Letras, 1993, p. 75-81.
- LOPES, Célia. *A inserção de a gente no quadro pronominal do português: percurso histórico*. 1999. PhD Dissertation, FL, UFRJ, Rio de Janeiro, 1999.
- LOPES, Célia. *A inserção de 'a gente' no quadro pronominal do português*. Frankfurt/Madri. Vervuert/Iberoamericana, vol. 18, 2003.
- LOPES, Célia. “A gramaticalização de 'a gente' em português em tempo real de longa e de curta duração: retenção e mudança na especificação dos traços intrínsecos”. *Fórum Lingüístico UFSC*, v. 4, p. 47-80, 2007.
- LOPES, Célia; CAVALCANTE, Sílvia. “A cronologia do voçamento no português brasileiro: expansão de você-sujeito e clítico-te”. *Linguística*, Madrid, v. 25, p.30-65, 2011.
- LUNGUINHO, M. Vinicius; MEDEIROS JUNIOR, Paulo. “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, 147-191, 2009.
- MODESTO, Marcello. “Sujeitos nulos em línguas de tópico proeminente”. *Revista da ABRALIN*, vol. 3, 121-148, 2004.
- MODESTO, Marcello. “Topic Prominence and Null Subjects”. In: BIBERAUER, T. (Ed.) *The limits of syntactic variation*. Amsterdam/Philadelphia: John Benjamin Publishers, 2008, p. 375-410.
- PAREDES SILVA, Vera L. “O retorno do pronome *tu* à fala carioca.” In RONCARATI, C.; ABRAÇADO, J. *Português Brasileiro – contato linguístico, heterogeneidade e história*, p.160-169, 2003.
- RIZZI, Luigi. *Issues in Italian syntax*. Dordrecht: Foris, 1982.
- ROBERTS, Ian. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer, 1992.

- SCHERRE, Marta et al. “Variação dos pronomes *tu* e *você*.” In: MARTINS, M. A.; ABRAÇADO, J. (eds.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, p. 133-173, 2015.
- SCHERRE, Marta; DUARTE, Maria Eugênia L. “Main Current Processes of Morphosyntactic Variation.” In WETZELS, L; MENUZZI, S; COSTA, J. (eds) *The Handbook of Portuguese*. Malden, MA: John Wiley & Sons Inc. p. 526-544, 2016.
- SOUZA, Janaína P. F. de. *Mapeando a entrada do você no quadro pronominal: análise de cartas familiares dos séculos XIX-XX*. Master Thesis. Rio de Janeiro: UFRJ, 2012.
- TAGLIAMONTE, Sali. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press, 2006.
- VARGAS, Amanda de S. C. “A evolução na representação das estratégias pronominais de indeterminação”. In: DUARTE, M. E. L. (ed.). *O sujeito nulo em peças de teatro (1843-1992): estudos diacrônicos*. São Paulo: Parábola Editorial. p. 45-68, 2012.
- VIANNA, Juliana Barbosa S.; LOPES, Célia. “Variação dos pronomes *nós* e *a gente*.” In: MARTINS, M. A.; ABRAÇADO, J. (eds.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, p. 109-131, 2015.
- WEINREICH, Uriel; LABOV, William.; HERZOG, Marvin. “Empirical foundations for a theory of language change”. In: LEHMAN, W.; MALKIEL, Y. (eds.). *Directions for Historical Linguistics*. Austin: University of Texas, p. 97-195, 1968.