Null subjects and agreement marks
in European and Brazilian Portuguese

MARIA EUGÊNIA LAMMOGLIA DUARTE
FILOMENA VAREJÃO

Abstract

The purpose of this article is to discuss the relation between the presence of agreement marks and null subjects in contemporary European and Brazilian Portuguese based on very recent samples of speech collected in Lisbon and Rio de Janeiro. In order to bring additional evidence to support the analysis, we will briefly review the issues involving the Null Subject in Romance Languages and some diachronic and synchronic analyses which distinguish Brazilian and European Portuguese as far as null subjects are concerned. We will finally show that recent samples of spoken European and Brazilian Portuguese confirm the stability of EP as a null subject language and the change affecting referential subjects in BP.

0. Introduction

This article provides a discussion related to the presence or absence of person and number morphemes to express the agreement between verb and subject in Brazilian and in European Portuguese. It has been a general belief that standard European Portuguese (EP) exhibits agreement verb marks systematically, whereas in Brazilian Portuguese (BP) it is a variable rule, constrained by structural and social factors, especially as far as third person plural is concerned. Our main focus it to show that null subjects in EP and BP are quantitatively and qualitatively different as the change affecting the representation of referential subjects in BP, in favor of overt pronouns, proceeds.

The paper is organized as follows: section 1 contains a brief review of the issues involving rich agreement and null subjects in the context of Romance languages. This same section brings the main results offered by diachronic and synchronic empirical researches based on Brazilian and European samples coming from different sources. Section 2 introduces the remarkable work developed by Naro and Scherre in the investigation of presence vs absence of agreement marks in BP, with important achievements which have been inspiring a number of researches all over the country. We close this section with Duarte´s empirical studies showing that the presence or absence of agreement marks do not affect the representation of subjects in BP. This leads us to section 3, where we present an analysis based on contemporary spoken European and Brazilian Portuguese in search of empirical evidence for our main claim as regards the role of agreement marks in licensing and identifying null subjects in BP. We conclude the paper with a brief discussion of some of the main differences found in the samples analyzed.

1. Null subjects and "rich" agreement morphology

Discussions related to the Null Subject Parameter (NSP) have been enriched in recent years as more and more studies based on different languages have contributed to the understanding that this is not a Yes/No matter. Holmberg´s (2009) summarizes a series of previous studies which lead to four main types of languages as far as the representation of the subject is concerned: non-null-
-subject languages, radical pro-drop languages, consistent null-subject languages and partial null-subject languages. Under this proposal, EP could be considered a consistent null-subject language and BP, a partial null-
-subject language. Since our purpose is to compare European and Brazilian Portuguese, we will limit our analysis to the inflectional verbal paradigms, which have been a central issue in the discussion related to null subjects in Romance languages.

Officially proposed by Chomsky (1981), when the Generative Grammar turned to the identification of universal Principles in human languages and the limits of variation, or Parameters, responsible for inter-linguistic differences, the NSP has been the one receiving more contributions, discussions and revisions, and, consequently, is the best described among the known/proposed parameters. The original proposal, based on a comparison between English, on the one hand, and Romance null subject languages, particularly Italian and Spanish, on the other, led to the conclusion that the property of "missing" subjects should be related to the "rich" morphological specification of verbal agreement paradigm[1].

---

[1] "The intuitive idea is that where there is overt agreement, the subject can be dropped, since the deletion is recoverable" (Chomsky,1981, 2a. ed.1982:241)

Since Huang (1984), however, the relation between rich agreement and null subjects lost the exclusivity in the licensing and identification of null subjects, since Chinese exhibits no inflections at all[2]. New hypotheses were then raised. Jaeggli & Safir (1989) claimed that what licenses null subjects is not a rich system of agreement but the morphological uniformity of verbal paradigms. A paradigm is considered morphologically uniform if it has only derived forms, which can include person, number, tense endings, for instance, varying across languages, or non-derived forms, consisting only of the root, as is the case of Chinese. The identification in the first case would rely on the agreement marks, and in the second on the presence of a structural or discursive topic[3]. If, however, a paradigm exhibits complex and simple forms, null subjects would not be allowed[4].

Roberts (1993) brings a new contribution to this discussion when he proposes, based on empirical evidence from medieval French, that a "functionally rich" paradigm – with a zero ending and a syncretism – could also license and identify null subjects, just like "formally rich" paradigms.

Even though NSLs are no longer always related only to verb morphology, there is reason to attribute the high rates of null subjects in Italian and European Spanish to a functional principle of grammar whose inflectional paradigms exhibit distinctive endings for each grammatical person, for singular and plural, in most verbal moods and tenses. The option for an overt pronoun would be, therefore, a marked option in terms of frequency, because it is related to contexts of emphasis or contrast, i.e., to avoid ambiguity in the interpretation of a null subject. Such intuition, which associates a functional aspect to a principle of grammar is explicit in Rizzi (1988:15), for whom

> "the use of pronounced material is legitimate only when necessary to convey the intended meaning, within the constraints of UG and of the particular grammar". This implies that, given the existence of a zero pronominal option, in languages like Italian, the overt form will be limited to the cases in which it is necessary, i. e., when the pronominal subject,

---

[2] Cf. Rizzi (1986), for whom the formal licensing and identification of a null subject involves distinct mechanisms.

[3] As a consequence, the status of the null subject would be different according to the mechanisms involved in its identification, a ***pro*** or a ***variable*** (see Figueiredo Silva, 1996), for a proposal and a thorough review of the constraints involving the identification of a null subject in Brazilian Portuguese.

[4] We know that the question is not so simple. In some languages the occurrence of null subjects is restricted to certain verb tenses and grammatical persons. That is the case of Hebrew, for instance (see Borer,1989), which seems to keep a relation between presence of Agr for person and a null subject; we have no information about frequency of null subjects in that language. In this case, the property of exhibiting null subjects would not affect the whole paradigm but specific persons and tenses.

being focal or contrastive must bear stress (evidently, the zero element cannot bear stress)".

Recent empirical researches on spoken European Spanish (Soares da Silva, 2006; 2011) and Italian (Marins, 2009) add the turn-taking among the contexts which favour an overt first person pronoun in NSLs.

How to situate European Portuguese (EP) and Brazilian Portuguese (BP) in this context? EP exhibits a paradigm with two synchretisms: the first represented by the zero ending for the "indirect"[5] second person singular, *você, o(a) senhor(a)* (you), and the third person singular. It also counts on the exclusive ending <–s> for the "direct" second person singular *tu* (you), used in more familiar relations and with closer friends. The second syncretism refers to second and third person plural <-m>. Such paradigm reflects exactly what Roberts refers to as a "functionally" rich paradigm, as opposed to a formally rich one, and it seems to be responsible for the preference for null subjects in every grammatical person. Duarte (1995; 2000)[6] finds for spoken EP rates of 65%, 76% and 79% of null subjects for 1st, 2nd, and 3rd persons, respectively. Notice that 2nd and 3rd persons do not seem to be affected by the syncretic verb forms, exhibiting similar high rates of null subjects. As for 2nd person singular, null subjects occur regardless of the choice of the address form *você* or *tu*, a clear evidence that the system of oppositions allows the paradigm to function in a balanced way, proving that there is no direct relation between a distinctive inflectional ending and a null subject. In other words, there is no functional compensation at work.

In the case of BP, the question is very different from what is described for Romance NSLs. Our inflectional paradigm has been reduced significantly, either for the occurrence of phonological processes or by changes in the pronominal system, affecting licensing and identification of referential null subjects (cf. Duarte 1993; 1995; 2000). In fact, the use of *você*, either replacing *tu* or in variation with *tu* (generally without the verbal <-s> ending) and the grammaticalization of the former nominal expression *a gente* (the people, we) (see Lopes, 2003; Marcotúlio, Lopes and Vianna in this volume), which also combines with a zero ending verb form, largely preferred over the old variant pronoun *nós* (we), either for definite or indefinite reference, resulted in an extremely reduced paradigm, with three distinctive forms ("if" the speaker uses the canonical endings for second and third person plural). In other words, we lost the functional richness in Roberts's (1993) terms. And even though we still omit subjects, there is clear preference for overt
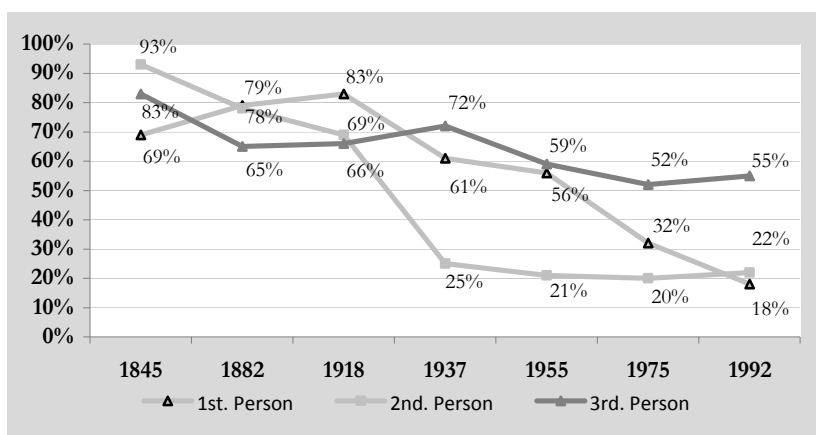
---

[5] Traditional grammars refer to former address forms as "indirect" second person, as was the case of "você".

[6] All the quantitative results presented in this section exclude null subjects in coordinated structures with co-referential subjects, except for the first of a sequence, since a null subject in such structures is not a property of a NSL.

pronouns in the three grammatical persons, and identification of null subjects are no longer related to the inflectional paradigm, as we will see below.

Results coming from theater popular plays, written along the 19th and 20th Centuries (Duarte, 1993) show the reduction in the inflectional paradigms, beginning with the option for *você*[7], in the South-East, from the 1930s on, and later, from the 1970s, with the increasing competition between *nós* and *a gente*. When the pronoun *tu* returns in more recent plays, written in the 1990s, it appears without the inflection <-s>, in variation with *você*. The results obtained by Duarte (1993) show, however, that the implementation of overt subjects affects first and second persons more quickly than third, which shows more resistance to the process of change:

Fig. 1. Null (vs. overt) subjects across seven synchronies (From Duarte 1993:112)

First and second persons start, respectively, with 69% and 93% of null subjects (1845) and reach 18% and 22% in 1992. As for third person, the rates decrease more slowly, from 83% to 45%. Such particular behavior of first person on the one hand and of the third on the other towards an overt category, compared to Cyrinos's (1994; 1997) diachronic results for the change towards null objects, also obtained in the analysis of plays, led Cyrino, Duarte and Kato (2000) to propose the role of a referential hierarchy in processes of change in direction of overt or null representation of pronouns. According to the authors, the hierarchy predicts that more referential items, inherently [+human], such as first and second person pronouns, will be the first to become phonetically expressed, whereas third person expressions, which involve the combination of features [+/- human] and [+/-specific], will show a variable behavior. In the case of subjects, the hierarchy predicts that [+human/+specific] referents will precede other combinations in the process of becoming overt. This explains the more regular behavior of third person in Figure 1: the data collected included all possible combinations of semantic features. A recent analysis by Duarte, Mourão and Mendonça (2012) refine Duarte's (1993) analysis for 3rd. person, on the basis of a larger sample of plays, but keeping the same periods of time established for the previous analysis. The results can be seen in Figure 2:

Fig. 2. Null (*vs.* overt) third person subjects and the combination of semantic features of the referent (Adapted from Duarte, Mourão and Mendonça, 2010: 39)
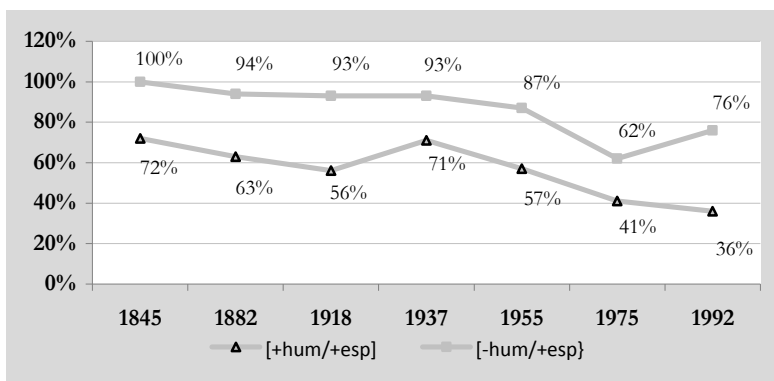
Since the occurrence of [-specific] referents (human or non-humam) is very rare and irregularly distributed in the samples, Figure 2 represents only [+specific] subjects. The line for [+human/+specific] subjects confirms Cyrino, Duarte and Kato's hypothesis. It starts with 72% of null subjects in 1845 and reaches 36% in 1992, clearly leading the change toward overt subjects. Notice, however, that [-human] referents already reveal a change in the last quarter of the 20th Century, which is more expressive in analyses of

spoken language, a behavior which is not attested in prototypical Romance NSLs[8].

Synchronic researches of BP[9] under the parametric perspective start with Duarte (1995). The results for spoken educated speakers, comparable to the sample analyzed for EP, mentioned above (Duarte 2000), reveal for the three persons, respectively, 26%, 10% and 42% of null subjects. An analysis of popular speech, with individuals with 4 to 11 years of school attendance (Duarte 2003), finds an average rate of 20%, regardless of the grammatical person.

As for writing, Barbosa, Duarte and Kato (2005), focusing the expression of third person subjects in a comparative analysis based on interviews transcribed in newspapers and magazines published in Lisbon and Rio de Janeiro, find quantitative and qualitative differences for each variety. The overall results show 78% of null subjects for EP and 44% for BP. The authors refined the structural contexts examined in Duarte (1995), proposing four different structural patterns:

Pattern I: the antecedent of the null subject is the subject of the matrix clause

Pattern II: the antecedent is the subject of the previous adjacent sentence

Pattern III: the antecedent is the subject of a previous, non-adjacent sentence

Pattern IV: the antecedent is in the previous adjacent sentence, but is functionally distinct from the null subject

The results for each pattern can be seen in Table I:

---

[8] It must be said that EP shows overt [-human] pronouns, contrary to what one finds in Italian and European Spanish. However, they are extremely rare.

[9] Even though the results presented here come from the variety spoken in Rio de Janeiro, several other studies on the same phenomenon based on samples from several other regions in the country find similar results, allowing the use of "Brazilian Portuguese".These results such as other variationist researches carried out in different regions of the country bring evidence to Tarallo's (1983:54) claim that dialectal differences in BP are more frequently located at the lexical and the phonetic-phonological levels than at the syntactic level.

Table I – Null (vs overt) subjects in transcribed interviews, according to the structural context (Adapted from Barbosa, Duarte and Kato, 2005: 24)

Pattern	EP	BP
I	39/40 (97%)	18/23 (78%)
II	49/55 (89%)	28/48 (58%)
III	20/28 (71%)	07/28 (25%)
IV	16/24 (67%)	10/23 (43%)

Observe that in EP, null subjects are the preferred realization regardless of the syntactic accessibility of the antecedent: the lowest rate (67%) is obtained when the antecedent bears a different function from the subject (Patterns IV). On the other hand, BP only reveals a clear preference for a null subject in Pattern I: the antecedent bears the same function and is in the previous clause, usually the matrix. Pattern II, which shows the antecedent as the subject of an adjacent clause, exhibits a significant decrease when compared to Pattern I. A distant antecedent or an antecedent with a different function have already surrendered to the change. In other words, a null subject in BP depends on adjacency and of a parallel function of its antecedent. Besides, an overt pronoun in EP is almost categorically [+human] whereas in BP there is no such restriction.

Similar results have been found by Duarte (2007a) for newspaper articles, as shown in Table II:

Table II – Overt third person subjects in newspaper articles, according to structural context – EP and BP – (Adapted from Duarte, 2007a, Table 6)

Pattern	EP	R.W.	BP	R.W.
I	83/87 (95%)	.54	48/66 (73%)	.76
II	100/105 (95%)	.59	54/109 (50%)	.52
III	18/22 (82%)	.20	11/36 (31%)	.29
IV	26/30 (87%)	.34	06/30 (20%)	.16

Once again, null subjects are preferred in EP in every structural context in newspaper writing, which shows, as expected, highest rates than the ones found in transcribed interviews. BP, on the other hand, behaves in a very similar fashion as regards Patterns I and II, but exhibits even lower rates for Patterns III and IV. The importance of such patterns in a process of change is highlighted by the relative weights obtained in the Variable Rule Statistical analysis (Pintzuk, 1988). Since these weights have to be interpreted in relation to the others obtained in the same run, the values for Patterns I and II are higher than the ones obtained for Patterns III and IV in each run. This

reinforces the effect of Patterns I and II in the preference for null subjects as compared to patterns III and IV and explains why the latter are the most favorable contexts to the implementation of the change towards overt subjects.[10] It would be important to emphasize that even though the percentage differences between patterns I and II in EP and patterns III and IV do not seem very significant (which is in accordance with consistent null--subject language) the relative weights difference attest that, if an over subject appears, the preferred patterns will be III and IV.

To sum up, one could conclude that, as the inflectional paradigm becomes unable to identity a null subject in BP, a less referential item and a less accessible antecedent can refrain the process of change.

2. The variable use of agreement marks in Brazilian Portuguese and null subjects

2.1. The studies on variable use of agreement marks

The interest for the lack of agreement nominal and verbal marks in spoken BP has attracted the attention of philologists in the first half of the 20th Century. However, only after Anthony Naro and his students' the first systematic analyses of variable nominal and verbal agreement in BP could we understand the regularity of the phenomenon in the country. Ever since Naro's arrival in Rio de Janeiro in the early 1970s, agreement has been among the main variable processes investigated under his supervision and it is certainly one of the best described variable phenomena in BP and the only morpho-syntactic feature which is subject to strong social stigma. Maybe, such stigma is responsible for a reversal in the direction of the variable process announced by Naro in his 1981 article. According to the author, only negative evaluation and exposure to more prestigious varieties could revert the direction towards the loss of marks.

Lemle and Naro (1977) are responsible for the pioneer investigation, which has been developed by Naro and Scherre, his partner since her first studies in the 1980s. They were able to refine the structural constraints underlying the variation process, such as the phonic salience of the verb plural form (or the nominal forms) and the subject position in relation to the verb, as well as social correlates. Contrary to their expectations, the sequence of analyses carried out in more recent years showed an increase in the use of marks of agreement, either nominal or verbal. Their Panel and Trend Studies (Naro & Scherre, 2003; Scherre & Naro, 2006, among many other analyses)

---

[10] Relative weights should be always interpreted inside the group they have been obtained and always in" relation" to the each weight. For instance: the weights .54 and .59 obtained for Patterns I and II are very close and reveal that, when compared to .20 and .34, they clearly favour a null pronoun.

showed that the increasing use of marks was not only in the performance of speakers who had improved their school attendance in the lapse of time which separated the samples analysed, but also in the speech of those who had not changed their school level. Even though the community study showed flows and counter-flows, the direction towards more marks was evident, and the effect of the structural factors persisted. In the next section we will show that null subjects are no longer related to inflection.

2.2. Null subjects and presence/absence of agreement marks

This section investigates the results obtained for the relation between inflection and null subjects based on two analyses of spoken BP referred in section 1: Duarte (1995), for College educated speakers and Duarte (2003; 2007b), for popular speech[11]. As mentioned in Note 8, coordinated structures with co--referential subjects have not been considered in the analyses related in this section, except for the first of a sequence, since a null subject in such contexts is a more general property among languages rather than a property of NSLs.

Table III shows the results for college graduate speakers:

Table III – Null subjects vs inflection and grammatical person – graduates
(Adapted from Duarte, 1995:56)

Inflec.	<- o> (- i)	<- mos>	<-m>	<-Ø>
Person	N / T (%)	N / T (%)	N / T (%)	N / T (%)
1PS	138/478 (29)			37/153 (24)
2PS.				13/119 (11)
3PS.				165/419 (39)
1PP.		15/53 (28)		3/67 (4)
2PP.			0/ 6 (0)	0/ 2 (0)*
3PP.			42/111 (38)	2/ 16 (13)*

* Verbs which present homophonous forms for third person singular and plural.

The four types of endings which appear in the first line illustrate exclusive first person marks (for singular and plural, respectively) and second and third plural endings. The zero ending can be found in Indicative Imperfect and Subjunctive first person singular; in second and third singular and first person plural, if the speaker uses the pronoun *a gente,* and in third

---

[11] Both samples are available in www.letras.ufrj.br/nurc-rj and www.letras. ufrj.br/peul.
The adjective *popular* is used here as opposed to *graduated*. As we mentioned in section 1, researches related to many syntactic phenomena reveal more similarities than differences when school level is concerned, except as regards the more or less frequent use of explicit agreement marks.

person plural in cases of homophony between singular and plural second and third persons (*ele tem/eles têm* – he has/they have).

The results show that inflection seems to play no significant role in the occurrence of null subjects First person singular shows 29% for distinctive verb inflections and 24% for zero ending forms (as for the imperfect and subjective tenses); first person plural, however, exhibits an interesting correlation between distinctive inflection <–mos>, used with the agonizing pronoun *nós* (we), with 28% of null subjects, and the zero ending, used with *a gente*, only 4%, when in co-reference with the subject of a main clause[12].

In the second person, which seems to be in the root of the pronominal paradigm reduction, as a consequence of the neutralization between *tu* and *você*, null subjects have almost disappeared: 11% for second person singular and no occurrences for second person plural[13].

Third person singular and plural shows the highest rates of null subjects (39% and 38%), which leads us to the conclusion that the presence of agreement marks for third person plural, in the speech of people who usually use them more systematically, especially when the subject precedes the verb and is not distant from it, plays no role in the occurrence of null subjects. Such percentages are reflected in the relative weights obtained by Duarte's analysis[14], showing that the context of resistance of null subjects is correlated to the structural pattern in which they occur and to the semantic features of the referent, as we have seen in section 1.

The results for popular speech do not show different results, as shown in Table IV:

---

[12] The studies referred in Note 9 present similar results.

[13] Even though our sociolinguistic interviews do not favour second person structures, these results are confirmed in contemporary plays, for instance.

[14] In fact, the Variable Rule Program selected the grammatical person as the first group of factors relevant to the representation of subjects, pointing out the significance of third person versus first and second for the occurrence o a null subject. The type of inflectional ending group was not selected.

Table IV – Null subjects vs inflection and grammatical person – non-graduates
(Duarte, 2007b: 270)

Inflec.	<- o> (- i)	<- mos>	<-m>	<-∅>
Person	N / T (%)	N / T (%)	N / T (%)	N / T (%)
1PS.	401/1876 (21)			57/372 (15)
2PS.				35/159 (22)
3PS.				246/1288 (19)
1PP.		29/78 (37)		8/110 (7)
2PP.			1/10 (10)	0/3 (0)
3PP.			59/336 (18)	7/33 (21)

Percentages for null subjects are even lower than those found for graduates, except for instances of first person plural (37%), with the most salient plural ending <-mos>; however, similarly to the previous sample, the preference for *a gente* in place of *nós*, for second person plural reference is a fact in popular varieties as well. Notice that third person plural in popular speech exhibits variation between standard inflection <-m> and a zero ending, but null subject rates are about the same: 18% if there is a distinctive verb mark and 21% with no mark at all. The examples in (1) – (2) illustrate overt subjects with homophonous verb forms for singular and plural (in a) and verbs used with the non-standard zero ending (in b), included in the column labeled "<-∅>" in Table IV above:[15]

(1) a. Vocês <u>têm</u> que saber que cês <u>têm</u> uma filha e um filho...
     you-3PP have-∅ to know that you-2PP have-∅ a daughter and a son
     '*You have to realize that you have a daughter and a son*'

    b. Vocês <u>fica</u> com esses garoto novinho...
     you-3PP stay-∅ with these-PL boy-SG young-SG
     '*You stay with these young boys…*'

(2) a. Nós não temo nada a ver com isso, que eles <u>têm</u> associação.
     we not have nothing to see with that, for they have-∅ association
     '*We have nothing to do with that, because they have their association*'

    b. Bolo, salpicão, eles também <u>adora</u>, né?   Eles <u>come</u> até pedra.
     cake, salad,   they also   love-∅, isn't it? They eat-∅ even stone
     '*Cake, salad, they also love (that), you see? They can eat even stones*'

---

[15] Examples (1)-(7) come from the popular speech sample – PEUL.

The use of non-standard verb forms obeys the constraints pointed out by Naro and Scherre in their several analyses: less salient verb endings are more easily lost (*fica ~ ficam* – stay); with those forms showing more phonetic salience, such as irregular verbs, like *ser* and *estar* (to be), which are inflected as *é* and *está* (is) for second and third person singular and *são* and *estão* (are) for second and third plural, and other forms, such as *fez*--3PS/*fizeram*-3PP (did), the use of plural marks is more frequent. See examples (3)-(4):

(3) a. Vocês_i <u>são</u> do Brasil? [∅]_i <u>São</u> brasileiros? Vocês <u>falam</u> muito bonito.
You-2PP are-2PP from Brazil? Are-2PP Brazilians. You-2PP speak--2PP very beautiful
'*You are from Brazil? You are Brazilians? You speak so beautifully!*'

b. Vocês <u>estão</u> muito bem. Vocês <u>vão</u> ser então feliz.
You-2PP are-2PP very well. You-2PP are-2PP going to-be happy
'*You are very well. You are going to be happy.*'

(4) a. Eu não sei o que eles <u>vão</u> mandar, né?
I not know what they-3PP are-3PP going to send, isn't it?
'*I don't know waht they are going to send, you see?*'

b. Mas eles <u>fizeram</u> uma costeira que quase encostou nos fundos das casa dos particular.
But they-3PP made-3PP a wall which almost reached the backyard of the-PL resident-SG
'*But they built a wall which almost reached the backyard of the residents.*'

If the rates of null subjects were not so low, one could think of a sort functional compensation. But they are low for each person, even in cases where a null subject would not lead to ambiguity. See the second sentence in example (5):

(5) Eles <u>trata</u> meu filho bem; <u>eles</u> <u>trata</u> ele bem.
They-3PP treat-3PS my son well; they-3PP treat-3PS he well.
'*They treat my son well; they treat him well*'

In the same way, it is not easy to relate presence or absence of marks with overt or null subject. In the two sequences below we see opposite tendencies:

(6) a. <u>Eles</u>_i que <u>manda</u>;[∅]_i não <u>mandaram</u> ainda.
They-3PP that send-3PS; not sent-3PP yet
'*It's them who send (it); They haven't sent (it) yet.*'

b. <u>Eles</u>_i nem se <u>lembram</u> do que [∅]_i <u>prometeu</u>.
They-3PP not even remember-3PP what promised-3PS
'*They don't even remember what they have promised.*'

In coordinated structures, this lack of correlation is even more clear:

(7) a. <u>Eles</u>ᵢ <u>vai</u>   ver e [∅]ᵢ <u>vai</u>   gostar.
     They is-3PS going to-see and is-3PS going to-like
       '*They are going to see (it) and are going to like (it).*'

  b. Aí <u>eles</u>ᵢ <u>vai</u>   ligar   e [∅]ᵢ <u>vão</u>   corrigir.
     Then they is-3PS going to-call and   is-3PS going to-fix
     '*Then they are going to call and are going to fix (it).*'

  c. Aí, [∅] <u>jogam</u>   pedra, [∅] <u>quebra</u>   telha...
     Then   throw-3PP stones,   break-3PS tiles....
     '*Then, they throw stones, break tiles…*'

It is no surprise that the Variable Rule Statistic Program has selected only a social factor in the analysis for this sample: the number of years of school attendance. Among the structural factors, the syntactic type of clause, the structural pattern, the semantic feature of the referent have been selected as relevant for the expression of the subject. The grammatical person and the presence/absence of distinctive ending have been ignored. In the next section, we will examine two recent samples of speech, focusing our analysis on third person plural sentences only.

3. A new comparative analysis: null subjects and agreement marks in EP and BP

The analysis presented in this section will be restricted to one sample collected in Oeiras, Lisbon, and another collected in Copacabana, Rio de Janeiro for the Project[16]. Considering the results obtained for different samples, described in the previous sections, and the results obtained by Vieira and Bazenga for verbal agreement (this volume), we believe that we will reach reliable results for the relation between null subjects and use of agreement marks.

Our focus is null third person plural subjects in finite sentences. We disregarded cases of assertive answers, as in (8a,b) and cases of repetitions caused by hesitation, counted as a single occurrence, as in (9a,b), and clear cases of arbitrary reference, as in (10):

(8) a. Entrevistador: e os professores eram bons?
     Interviewer: and the teachers were-3PP good?
               '*And were the teachers good?*'

---

[16] See the Editorial statement for the description of the *corpora.*

Entrevistado: **era** (COP-B-1-M)
Interviewee: was-3PS
                              'yes, they were'

b. Entrevistador: e os avós sempre mimam, né?
   Interviewer: and the grandparents always spoil-3PP, don't they?
               'and grandparents always spoil, don't they?'

Entrevistador: **Ø$_i$ sempre mimam-3PP** (COP-A-2-M)
Interviewer: **Ø$_i$** always spoil-3PP
               'yes, they do'

(9) a. eles **sofrem** muito… **sofrem**… arquitetura é dificil (COP-A-2-M)
    they suffer-3PP a lot …suffer-3PP architecture is hard
    '*They suffer… suffer… architecture is hard.*'

b. muitas pessoas elas... estudavam, **tinha**....**tinham** essa visão assim...
   (COP-A-2-M)
   many persons they ...studied-3PP, had-3PS…had-3PP this vision
   like-that.
   '*Many people suited, they had… had… this kind of perspective.*'

(10) tem essas coisas que a gente sabe que aqui **Ø$_{arb}$ não falam** assim
   (COP-A-3-H)
   has these things that the people knows that here don't speak-3PP
   this-way
   '*There are those things which one knows they they don't speak like
   that here*'

Contrary to the previous analyses reviewed in sections1 and 2, however, we decided to include coordinated structures because we believe that this will be the only structural context which will present similar results for each variety. We consider coordinated sentences those which occur from the second clause on in a sequence with a co-referential subject, as illustrated in (11):

(11) a. **eles$_i$** se conheceram naquela praça. **Ø$_i$** Namoraram, **Ø$_i$ casaram e
   Ø$_i$ foram** morar na Tijuca, na Adolfo Mota (COP-C-3-M)
     they SE-CL met-3PP in that square. Dated-3PP, got married-3PP
   and went-3PP to live in Tijuca, on the Adolfo Mota
   '*They met in that square. They dated, got married, and moved to
   Tijuca, at Adolfo Mota street.*'

b. há outra coisa que também não funciona – a lei d[os médicos]$_i$: Ø$_i$
   Trabalham no privado e **Ø$_i$ trabalham** no público (OEI- A-3-H)
   there is another thing that does-not work – the law of the doctors:
   work-3PP in the private and **Ø$_i$ work-3PP** in the public (service)
   '*There is another thing that doesn't work – the doctor's law: they can
   work in private and public service.*'

Based on previous evidence, we will take into account two factors: the cluster of semantic properties of the referent ([+/-human] and [+/-specific]) and the structural patterns discussed above. Before we present the results obtained, let's illustrate the four patterns proposed by Duarte, Barbosa, and Kato (2005), as mentioned in section 1:

Pattern I – the antecedent of the null subject is the subject of the matrix clause

(12) eles $_i$ aproveitaram   **o que Ø$_i$ podiam** daqui**...** (COP-C-2-M)
      [they] $_i$ enjoyed-3PP what **Ø$_i$ could-3PP** here…
      '*They enjoyed as much as they could here*'

Pattern II – the antecedent is the subject of the previous adjacent sentence

(13) O problema é haver [cada vez mais políticos corruptos]$_i$ que não fazem nada pelo país. **Ø$_i$ Querem** só meter dinheiro ao bolso (OEI- A-1-H)
      the problem is that there are [more and more corrupt politicians]$_i$ who do nothing for the country. **Ø$_i$ Want-3PP** only to-put money in the pocket
      '*The problem is that there are more and more corrupt politicians, who do nothing for the country. They only want to put the money in their pockets.*'

Pattern III – the antecedent is the subject of a previous, non-adjacent sentence

(14) por exemplo se eu fosse sair à noite os meus pais mandavam estar por exemplo à meia noite em casa e eu às vezes não conseguia estar a essa hora e não tinha meios de comunicar. **Ø$_i$ Ficavam** preocupados (OEI- A-1-H)
      for instance, if I went out at night [my parents]$_i$ ordered to be for instance at midnight back home and I sometimes could not be at the time and had no means to communicate – **Ø$_i$ Got-3PP** concerned
      '*If I went out at night, my parents said when I should be back, but sometimes I couldn't arrive at the time agreed and had no means to communicate with them.*'

Pattern IV – the antecedent is in the previous adjacent sentence, but is functionally distinct from the null subject

(15) a. vida está fácil... há máquinas para fazer tudo também, para os adultos também, está fácil para os miúdos também: já ao computador **Ø$_i$ estão** sentados (OEI- A-3-H)
      life is easy...there are machines to do everything...., for the adults also is easy, is easy [for the children]$_i$ also: already **Ø$_i$ are-3PP** sitting at the computer
      '*Life is easier...there are machines to do everything, even for adults; it's easy for the children too. They are already sat at the computer.*'

b. eu já tive oportunidade de vê-[los]ᵢ de perto e tudo... E **Øᵢ são** pessoas que falam bem (COP-A-3-H)
 I had the opportunity to see [them]ᵢ personally and all… And **Øᵢ are- -3PP** persons who speak well
'*I had the opportunity to see them closely. They speak well.*'

Table V presents the total of occurrences, which will be the focus of our analysis:

Table V – Third person null subjects vs inflection in EP and BP

Inflection	<-m>	<-Ø>
Origin	Occ. (%)	Occ. (%)
EP (Oeiras)	507 (99)	5 (1)
BP (Copacabana)	195 (79)	53 (21)

The first point to mention is that the total of null subject sentences in EP is almost three times the total attested in BP, which is no surprise, since BP prefers overt subjects. In fact, Vieira and Basenga (this volume) found 174 occurrences of overt personal pronouns in the Oeiras sample (EP) and 410 in the Copacabana sample (BP). In EP, null subjects are almost categorical in the presence of inflection: only 1% of the data, corresponding to five occurrences exhibit null subjects. BP, on the other hand, presents 21% of null subjects with absence of agreement marks. The examination of the two structural factors will help us reach some generalizations about the two varieties. Let's start with the effect of the semantic features in EP (Table VI) and BP (Table VII):

Table VI – Third person null subjects and the semantic features – EP – Oeiras

Inflection	<-m>	<-Ø>
Features	Occ. (%)	Occ. (%)
[+hum/+esp]	172 (99)	1 (20)
[+hum/-esp]	307 (99)	1 (20)
[-hum/+esp]	11 (74)	3 (60)
[-hum/-esp]	17 (100)	0

It is important to notice that the significant number of occurrences of [+hum/-spec] referents has a lot to do with the format of the interviews, whose main purpose was to obtain third person plural subjects; therefore, questions about raising children and adolescents today and people's behavior in the present as compared to the past, for instance, were necessary. That explains why, in both samples, [+hum/-spec] referents are more frequent,

followed by [+hum/+spec] referents. The occurrence of [-hum] subjects is rare. The absence of marks in BP is concentrated in [+hum] subjects.

Table VII – Third person null subjects and the semantic features – BP – Copacabana

Inflection	<-m>	<-Ø>
Features	Occ. (%)	Occ. (%)
[+hum/+esp]	88 (74)	31 (26)
[+hum/-esp]	91 (81)	21 (19)
[-hum/+esp]	9 (90)	1 (10)
[-hum/-esp]	7 (100)	0

Let's examine the effect of the structural patterns. Since we included coordinated sentences, they will appear as pattern V:

Table VIII – Third person null subjects and the structural patterns – EP – Oeiras

Inflection	<-m>	<-Ø>	**Total**
Patterns	Occ. (%)	Occ. (%)	Occ. (%)
I	95 (98)	2 (2)	97 (18.8)
II	203 (100)	0 (0)	203 (39,5)
III	51 (95)	3 (7)	54 (10,4)
IV	11 (100)	0	11 (2,2)
V	147 (99)	(0)	147 (29)

The total distribution of null subjects in the last column shows that all patterns present null subjects. Knowing that anaphoric subjects in patterns I and II are typically null in consistent null-subject languages, the total rate of 58,4 for these patterns this is an expected result. Notice that patterns III and IV, are related to distance or a functionally different antecedent, reach only 13% of null subjects, a very distant result from the rates for writing, which can be explained by the role played by the modality (speech vs writing) in recovering less accessible antecedents. Coordinated sentences respond for 29% of null subjects.

A close examination of the occurrences of non-agreement, only five cases, show 2 null subjects in pattern I, one with a [+hum/-spec] referent and the other with a [-hum/+spec] referent, illustrated in (16), both with the verb to be (*estar* and *ser*):

(16) a. as pessoas tinham mais tempo porque **Ø_i tava** mais em casa (OEI-
-B-3-M)
[the persons]_i had more time [because **Ø_i was-3PS** more at home]
'*People had more time because stayed longer at home*'

b. os livros dantes eram um bocado... agora agora vejo a diferença.
Acho, achava [que Ø_i era era um bocado massudos] não é? (OEI-
-C-3-M)
[the books]_i in-old-times were-3PP a little... now, now (I) see the
difference. (I) believe, believed that Ø_i was-3PS very boring, isn't it?
'*In old times, books were… now I see a difference. I think they
were too boring.*'

and 2 occurrences in Pattern III (one with a [+hum/+spec] and 2 with [-
-hum/+spec] referents, in (17): three instances are found in Pattern III, with a
distant antecedent, one with [+hum/+spec] and two with [-hum/+spec]
referents, as in (17):

(17) a. Entrevistado: e eles são responsáveis, sobretudo quando Ø_i têm que
estudar primeiro estudam, o que é bom também...
 Interviewee: and [they]_i are responsible, especially when **Ø_i** have
to study. **First Ø_i study-3PP**, which is good too…
'*And they are responsible, specially with their studies.*'

Entrevistador: isso é óptimo
Interviewer: That is very-good
           '*That's good.*'

Entrevistado: **Ø_i nunca foge** .... (OEI-C-3-M)
Interviewee: **Ø_i never try** to escape**,**
           '*They never try to escape.*'

b. os interesses são diferentes, em termos das coisas que eu fazia e
faço agora. Ø_i É completamente diferente né? (OEI-A-2-H)
[the interests]_i are different, in terms of the things I used-to do and
(I) do now. **Ø_i is-3PS** completely different
'*Interests have changed, in terms of what I used to do and do now.
They are completely different.*'

c. [estes óculos]_i estão a estão a cinquenta euros mas querem
aumentá-los para sessenta e cinco; então metem o preço em cima
sessenta e cinco euros e depois metem riscado os os cinquenta
euros tipo a dizer que Ø_i está em promoção (OEI-A-3-M)
these glasses are costing fifty euros but (they) want to raise them
to sixty-five; then, (they) major the price to sixty-five and then
(they) stripe the price so as to mean that **Ø_i is-3PS** on sale
'*These glasses cost 50 euros but they want to raise the price to 65;
so, they write 65 and then they stripe the price meaning they are
on sale.*'

The comparison of the figures in Table VIII with the ones in Table IX show a different situation:

Table IX – Third person null subjects and the structural patterns – BP – Copacabana

Inflection	<-m>	<-Ø>	Total
Patterns	Occ. (%)	Occ. (%)	Occ. (%)
I	19 (95)	1 (5)	20 (8)
II	63 (86)	10 (14)	73 (29)
III	1 (50)	1 (50)	2 (0,8)
IV	2 (50)	2 (50)	4 (1,6)
V	110 (74)	39 (26)	149 (60)

Coordination is the most favorable pattern for null subjects, when one considers the total of occurrences (60%); among the cases of presence and of absence of marks they also lead the ranking. See the example in (18):

(18) ...minha filha entra entra chorando mas [elas]$_i$ brinca, **Ø$_i$ bate palma**, **Ø$_i$ canta** musiquinha, **Ø$_i$ bota** música, **Ø$_i$ dança**... (COP-A-1-M)
... my daughter comes in crying, but [they]$_i$ play-3PS, **Ø$_i$** clap-3PS hands, **Ø$_i$** sing-3PS songs, **Ø$_i$** play-3Ps music, **Ø$_i$** dance-3PS
'*My daughter comes crying, but they play, clap their hands, sing songs, play music, dance.*'

We should also mention 14% of absence of marks in sentences found for Pattern II, with all combinations of semantic features, as shown in (19):

(19) a. [eles]$_i$ aproveitaram o que podiam daqui... e iam [funcionários]$_i$ daqui pra São Paulo. Ø$_i$ Ficava de segunda a sexta e voltava... **Ø$_i$ Ficava** sábado e domingo com a família...(COP-C-2-M)
they enjoyed what they could here and used-to go-3PP employees to São Paulo; **Ø$_i$ stayed-3PS** from Monday to Friday and Ø$_i$ returned-3PS. **Ø$_i$ Stayed-3PS** Saturday and Sunday with the family…

b. os hospitais precisa de uma grande reformulação; **Ø$_i$** precisa de pessoas profissionais (COP-B-2-H)
the hospitals need-3PS a big reformulation; **Ø$_i$ needs**-3PS persons professional
'*The hospitals need a big reform: they need professional personal*'

The unexpressive rates in Patterns III (0,8%) and IV (1,6%) and even in Pattern I (8%) confirm that the most resistant context for null subjects in speech is pattern II, which exhibits a null subject identified by a clear antecedent in the adjacent context, typical of partial null-subject languages (Holmberg, 2009).

3. Some final remarks

Our brief analysis allows the following considerations: EP and BP do not exhibit the same behavior as regards the occurrences of null subjects. It is true that we cannot consider EP as a radical pro-drop system; rather, one could say it behaves like a consistent null-subject language, and agreement plays an important role in the preference for null subjects. BP, on the other hand, is in process of change and its partial pro-drop nature is confirmed by the importance of an external element (a topic in the previous, adjacent context) in the identification of a null subject (see Modesto, 2008). But, as shown by previous analyses, this seems to be a stage in the process of change, since researches which compare overt and null subjects show that the typical contexts for null referential subjects (definite or arbitrary) already exhibit preferably an overt subject, regardless of the presence of an accessible subject.[17]

As for the occurrence of non-agreement third person plural verb marks, our results show that they are very rare in SV sentences in EP and tend to decrease in BP as well. Future work should deepen the investigation on the importance of relational verbs in EP. A review of Varejão's (2006) analysis of dialectal EP (Varejão and Duarte, in preparation) will show the main patterns of non-agreement, which seem to be in the origin of instances of EP absence of agreement marks shown in this article and in Vieira and Bazenga (this volume), most of which involving relational verbs in SV patterns.

References

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

Borer, H. (1989) Anaphoric AGR. In *The Null Subject Parameter* (O. Jaeggli & K. J. Safir, editors), pp. 69-110. Dordrecht: Kluwer.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris.

Cyrino, S. (1994) *O objeto nulo no português do Brasil – um estudo sintático-diacrônico*. PhD Dissertation, UNICAMP, São Paulo, Brazil.

Cyrino, S. (1997) *O objeto nulo no português do Brasil – um estudo sintático-diacrônico*. Londrina: Editora da UEL.

Cyrino, S.; Duarte, M. E.; Kato, M. A. (2000) Visible subjects and invisible clitics in Brazilian Portuguese. In *Brazilian Portuguese and the Null Subject Parameter (*M.A. Kato & E.V. Negrão, editors), pp. 55-104. Frankfurt-Madrid: Vervuert-Iberoamericana.

---

[17] BP may preserve some specialized null subject contexts, such as arbitrary subjects with third person singular verb forms, in sentences expressing modality, but nominative arbitrary pronouns are, by far, the preferred strategies.

Duarte, M. E. L. (1993). Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In *Português brasileiro: uma viagem diacrônica (*I. Roberts I & M. A. Kato, editors), pp. 107-128. Campinas: Ed. da Unicamp.

Duarte, M. E. L. (1995). *A perda do princípio "Evite Pronome" no português brasileiro*. PhD Dissertation, UNICAMP, São Paulo, Brazil.

Duarte, M. E. L. (2000). The loss of the Avoid Pronoun principle in Brazilian Portuguese. In *Brazilian Portuguese and the Null Subject Parameter* (M.A. Kato & E.V. Negrão, editors), pp. 17-36. Frankfurt-Madrid: Vervuert-Iberoamericana.

Duarte, M. E. L. (2003). A evolução na representação do sujeito pronominal em dois tempos. In *Mudança linguística em tempo real* (In M. C. Paiva & M. E. L. Duarte, editors), pp. 115-128. Rio de Janeiro: Contra Capa/Faperj.

Duarte, M. E. L. (2007a) Sujeitos de referência definida e arbitrária na escrita padrão: aspectos conservadores e inovadoras. *Revista Linguística*, 3 (1), 89-115.

Duarte, M. E. L. (2007b) Sujeito nulo/pleno e marcas de concordância. In *Anthony Julius Naro e a linguística no Brasil* (S. Votre & C. Roncarati, editors), pp. 265--277. Rio de Janeiro: 7Letras. 265-277.

Duarte, M, Mourão, G. & Mendonça, H. (2012) Os sujeitos de terceira pessoa: revisitando Duarte (1993). In *O sujeito em peças de teatro (1833-1992): estudos diacrônicos* (M. E. L. Duarte, editor), pp. 21-44. São Paulo: Parábola Ed.

Figueiredo Silva, M. C. (1994). *La position sujet en Portugais Brésilien (dans les phrases finies et infinitives)*. PhD. Dissertation, Université de Genève.

Holmberg, A. (2009) Null subject parameters. In *Parametric Variation: Null Subjects in Minimalist Theory* (T. Biberauer *et al*., editors), pp. 88-124. Cambridge: Cambridge University Press.

Huang, C.T-J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry,* 15, 531-574.

Jaeggli, Osvaldo & Safir, Kenneth, J. 1989. The Null Subject Parameter and Parametric Theory. In *The Null Subject Parameter.* (O. Jaeggli & K. J. Safir, editors), pp. 1-44. Dordrecht: Kluwer.

Lemle, M. & Naro, A. (1977) *Competências básicas do português*. Rio de Janeiro: Mobral/MEC.

Lopes, C. (2003) *A inserção de 'a gente' no quadro pronominal do português*. Frankfurt/Madrid. Vervuert/Iberoamericana, 18.

Marins, J. (2009) O Parâmetro do Sujeito Nulo: uma análise contrastiva entre o português e o italiano. MS. Thesis, Universidade Federal do Rio de Janeiro, UFRJ. www.letras.ufrj.br/posvernaculas

Modesto, M. (2008) Topic prominence and null subjects. In *The limits of syntactic variation.* (T. Biberauer, editor), pp. 375-409. Amsterdam/Philadelphia: John Benjamins.

Naro, A. (1981) The social and structural dimensions of a syntactic change. *Language*, 57, 63-98.

Naro, A. & Scherre, M. (2003) Estabilidade e mudança linguística em tempo real: a concordância de número. In *Mudança Linguística em Tempo Real.* (M. C. Paiva & M. E. L. Duarte, editors), pp. 47-62. Rio de Janeiro: Contra Capa/Faperj.

Pintzuk, S. (1988) *Varbrul Programs*. Unpublished.

Scherre, M. & Naro, A. (2006) Mudança sem mudança: a concordância de número no português brasileiro, *SCRIPTA*, 9, 18, 109-131.

Soares da Silva, H. (2006) *O parâmetro do sujeito nulo: confronto entre o português e o espanhol.* MS. Thesis. Universidade Federal do Rio de Janeiro, Brazil. www.letras.ufrj.br/posvernaculas

Soares da Silva, H. (2011) *Evidências da mudança paramétrica em dados da Língua--E: o sujeito pronominal no português e no espanhol.* PhD. Dissertation, Universidade Federal do Rio de Janeiro, Brazil. www.letras.ufrj.br/posvernaculas

Rizzi, L. (1986) Null objects in Italian and the theory of pro. *Linguistic Inquiry*, 17, 501-557.

Rizzi, L. (1988) The new comparative syntax: principles and parameters of universal grammar. ms.

Roberts, I. (1993) *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.

Tarallo, F. (1983) *Relativization Strategies in Brazilian Portuguese*. PhD. Dissertation, University of Pennsylvania, USA.

Varejão, F. (2006) *Variação em estruturas de concordância verbal e em estratégias de relativização e no português europeu popular.* PhD. Dissertation, Universidade Federal do Rio de Janeiro, Brazil.

Varejão, F. & Duarte, M. E. L. (In preparation) Patterns of non-agreement marks in Dialectal European Portuguese.

Maria Eugênia Lammoglia Duarte
(Departamento/Programa de Pós-graduação
em Letras Vernáculas) da Universidade
Federal do Rio de Janeiro
eugenia@brazilmail.com

Filomena Varejão
(Departamento/Programa de Pós-graduação
em Letras Vernáculas) da Universidade
Federal do Rio de Janeiro
varejaofilomena@gmail.com