

Patterns of third person plural verbal agreement

SILVIA RODRIGUES VIEIRA
ALINE BAZENGA

Abstract

This paper aims to provide a Labovian sociolinguistic description of 3rd person plural patterns of agreement in European (EP), Brazilian (BP) and São Tomé (STP) Portuguese based on very recent samples of speech stratified for age, sex/gender and education. Linguistic and social restrictions for the variation are investigated. Results from statistical analysis indicate that there are two patterns of agreement in Portuguese: a semi-categorical rule, typical of EP, and a variable rule, typical of BP and STP, restricted by specific linguistic and social factors. Additionally, the results indicate that general linguistic constraints – such as the position of the subject, semantic feature of the subject or even discursive parallelism – cannot say anything about historical origin of Portuguese varieties, since they can be concerned with any language. Therefore, besides the quantitative expression of non-agreement, the quality of the occurrences of non-plural marking may support the characterization of each variety.

1. Defining the theme: object and objectives

The presence or absence of morphological marks of agreement has been one of the most important topics to define and characterize varieties of Portuguese. As for Brazilian Portuguese, for example, there has been considerable research and a productive debate concerning at least two distinct conceptions: one assuming that Brazilian rates of absence of agreement marks come from European Portuguese transferred to Brazil, following natural linguistic drift (Naro & Scherre 2007), and another which proposes

that the absence of agreement marks in Brazilian Portuguese (BP), due to the intense linguistic contact, comes from a process of irregular linguistic transmission (Lucchesi, Baxter & Ribeiro, 2009).

Evidence for first conception could be found in data from archaic Portuguese, which would demonstrate the origin of the absence of plural marks, and from contemporaneous European Portuguese (EP), which would also confirm that, regardless of less variation and greater social constraints. Supporters of the second approach propose, however, that (i) the “variation of agreement in PE is, at best, peripheral and residual” (Lucchesi *et al.* 2009: 331), and (ii) in BP, there would be a *continuum* related to the implementation of agreement – from the standard urban educated communities (with high rates of agreement marks) to rural African-Brazilian ones (with lower rates of agreement marks) – which would indicate the effect of linguistic contact.

As we can see, linguistic studies concerning agreement clearly face a problem of empirical validation of the referred positions, since the differences between both approaches are primarily related to the scientific proving of the hypotheses. In fact, whereas a group of researchers takes for granted that data of absence of plural marks in EP are enough to postulate a similar pattern of agreement attested in BP, another group suggests that data do not prove – not only in quantitative terms, but also in qualitative ones – a variable rule of agreement in the case of EP, especially concerning the limited nature of the contexts of the plural marking expression. Moreover, the lack of sociolinguistic controlled sample in EP and the poor knowledge of African varieties and their marks of intense linguistic contact represent serious restrictions on the interpretations of BP origins.

The situation described above shows that contrastive studies of agreement in Portuguese varieties under the same conditions and the same criteria comprise an area in which further investigation is clearly needed. The international Luso-Brazilian Project *Comparative Study of agreement patterns in African, Brazilian and European varieties* aims at fulfilling this gap concerning analyses of the different varieties of Portuguese. As it was mentioned in the introduction of this volume, that project has not only built a database to contrastive analysis, but it has also developed descriptions of this material. As for 3rd person plural agreement, Vieira (2011) provided results which preliminarily attested the existence of variable rules in varieties from Brazil and São Tomé. On the other hand, a very few examples of non-agreement found in Lisbon variety of Portuguese, associated with very specific structural contexts (and sometimes dubious), gave rise to the author’s assumption that the European sample illustrated a semi-categorical rule of agreement. Bazenga (2012), taking into account different criteria for collecting data exclusively produced by speakers in Madeira, presented higher rates of absence of agreement marks than Vieira (2011).

In this paper, the authors, by means of adopting the same criterion of data collection, describe all the occurrences of 3rd person plural in the main sample of the Project in order to achieve the following goals: (i) establish the productivity of 3rd person plural marks in varieties of EP, BP and STP, so that it can be possible to postulate the status of each linguistic rule according to the typology proposed by Labov (2003), (ii) determine the social and structural factors that may constrain the variable rule of agreement if it is the case, and (iii) evaluate the relationship between structural and social factors that constrain the characterization of the agreement's pattern in each variety of Portuguese.

In order to accomplish the goals above, this paper is structured as follows: section 2 offers general information about theoretical and methodological aspects; section 3 brings the results from the analysis based on contemporaneous spoken data of BP, EP and STP, regarding the distribution of occurrences of the 3rd person plural marks and the effects of linguistic and extra-linguistic factors. The final section provides a discussion on the status of each variety in relation to the pattern of agreement, by considering the quantitative and qualitative evidences coming from the present analysis and comparing them with results from other investigations whenever it is possible.

2. Theoretical and methodological aspects

The research was guided by the principles of the Theory of Linguistic Variation (Weinreich, Labov & Herzog, 1968; Labov 1972, 2003), assuming the basic assumption of orderly heterogeneity and the structural and social constraints that act towards the variable phenomena.

As for the typology of linguistic rules, the research adopts the proposal of Labov (2003), under which two types of rules besides the variable one can occur: the categorical and the semi-categorical. In quantitative terms, the author establishes that a rule is categorical when operating with a frequency of 100%; semi-categorical when its violation is so rare that it operates very often (around 95-99%), and variable, when it operates with a frequency between 5 and 95%. Obviously, it is fundamental that the qualitative profile of linguistic contexts, which contain the data concerned with each kind of rule, may be associated to these indexes, so that the extension of the rule types can be evaluated.

As for the phenomenon of verbal agreement, the role of the following extra-linguistic variables has been investigated, according to the stratification of the sample we have studied: age, sex/gender and levels of education. Possible influence of groups of structural factors, especially those which have been tested in other researches, has also been controlled: (i) as to the subject: *position* in relation to the verb, *morph-syntactic configuration*; *distance*

between the SN subject and the verb; parallelism between the marks of the SN subject and the verb (*clausal parallelism*); semantic feature (*animacy*), and (ii) as to verb forms: *verbal phonic saliency* of singular-plural opposition; *transitivity*; parallelism between marks of verbs in series (*discursive parallelism*)¹.

For the data collection, the profile of each occurrence that would be effectively conceived as variable was accurately determined: only the one that contains a subject that semantic and formally refers to the 3rd person plural, expressed by variable forms of noun phrases, overt or null, would be considered. For this reason, any linguistic context which could lead to an ambiguous interpretation (Martins, 2012) as to the plural reference of the subject or the verbal form was disregarded in the analysis of quantitative data so that the comparison of results would be possible.

The criteria for the data collection were based on the assumption that an utterance may refer to different syntactic structures, which implies admitting that the absence of plural marks does not always constitute an effective case of non-agreement. In this sense, the following structures do not comprise the data set: verbal forms of 3rd person plural used as an indetermination feature of the subject (*Quebraram a janela. Não sabemos quem fez isso*. “Broke the window”/ “They’ve broken the window”. “No know who did it”/ “We don’t know who did it”); apparent subjects of 3rd person plural (overt or null) which correspond to topicalized constituents (*Pães de mel, (isso) é meu lanche favorite*. / “Gingerbread, (it) is my snack favorite”/ “Gingerbread, (it) is my favorite snack”); constructions with copulative verbs which can be interpreted as presentational (*os meus filhos é/são o meu problema* / “the my children is/are the my problem” / “my children is/are my problem”); *são/é eles que me incomodam* / “are/is they that me bother/ they are the ones that bother me”); infinitive verb forms, given the frequent ambiguity in its reference, whether personal or impersonal; occurrences of “to have” and “to come” in the present tense (*tem/têm; vem/vêm*), because they are homophones or difficult to identify in most occurrences from the *corpus*; structures in which the determination of the verbal form as singular or plural is dubious for phonetic-phonological reasons (as in the adjacency of the plural ending of verbal form followed by a word with an initial nasal consonant or a vowel, as in *eles cantam na praça* (“they sing in the square”) and *eles cantam uma canção calma* (“they sing a calm song”), respectively); and, finally, several coordination structures, partitive and numerical

¹ Linguistic parallelism is conceived as a principle (Scherre, 1988) according to which there is a tendency for similar forms to occur together, so that marks lead to more marks and zeros lead to more zeros. This principle can be applied to the relation between the subject and the verb (*clausal parallelism*), and between the verb and another verb (*discursive parallelism*).

expressions, titles of works, among others, whose interpretation can be either singular or plural.

Data were collected from all interviews that constitute the basic sample of the *Corpus* (see the introduction of this volume for more details)², which involves 18 participants in each survey point distributed by sex/gender, three age groups (18-35 years, 36-55 years and 56-75 years) and three levels of education (elementary: 5-8 years, intermediate: 9-11 years; high: more than 11 years). For the European varieties, the following regions were considered: Oeiras (OEI), Cacém (CAC), in the Metropolitan Region of Lisbon, and Funchal (FNC), in Madeira Island; for Brazilian varieties, the regions were: Copacabana (COP) and Nova Iguaçu (NIG), in the Metropolitan Region of Rio de Janeiro; and for the variety of Portuguese spoken in Sao Tomé and Príncipe (STP), samples were obtained in São Tomé. In the latter case, other interviews that also fit the criteria adopted in the investigation have additionally been considered, so that it could be possible to investigate the effect of the frequency of using mother tongue over the rates of agreement or non-agreement.

Table I shows the distribution of 7063 surveys collected in this study by Portuguese varieties:

Table I: Number of tokens by varieties of Portuguese

<u>EP</u>			<u>BP</u>		<u>STP</u>
<u>OEI</u>	<u>CAC</u>	<u>FNC</u>	<u>COP</u>	<u>NIG</u>	<u>ST</u>
1.467	1.185	914	1.395	1.365	737
3.566			2.760		737

3. Third person plural verbal agreement in Portuguese varieties

Data analysis demonstrated that the expression of agreement occurs following different patterns in the varieties of Portuguese. The following table confirms these differences.

² Brazilian students (Juliana Cristina Paula Pires, Vítor Araújo Ferreira, Cristina Marcia Monteiro Lima Correa, Fernanda Villares Vianna Barreto, Karen Cristina da Silva, Rodrigo Cunha da Silva and Daiane Rangel da Silva) collaborated on the task of collecting and processing the data of 3rd person plural marks in Brazilian varieties.

Table II: Distribution of occurrences of third person plural verbal agreement with and without plural marks within urban varieties of Portuguese

Third person plural verbal agreement				
Sample	With plural marks		Without plural marks	
	Tokens	%	Tokens	%
Oeiras	1454/1467	99.1	13/1467	0.9
Cacém	1176/1185	99.2	9/1185	0.8
Funchal	866/914	94.7	48/914	5.3
São Tomé	679/737	92.1	58/737	7.9
Copacabana	1229/1395	88.1	166/1395	11.9
Nova Iguaçu	1067/1365	78.2	298/1365	21.8

The general rates reveal that the EP samples have, in general, a semi-categorical agreement rule for the 3rd person plural, although Funchal demonstrates boundary behaviour between semi-categorical and variable rules in terms of Labov (2003). On the other hand, the rates of plural marks in other urban varieties, both in STP and in BP, show variable rule behaviour, with strong preference for agreement.

Obviously, these rough rates, considering all contexts, do not allow the observation of the data's complexity. For doing so, statistical treatment of the data held by Goldvarb-X (computer software package especially developed for the analysis of variable rules) could be developed, so that one can observe the relevance of the independent variables. In samples from continental EP (Oeiras and Cacém) or insular EP (Funchal), however, it was not possible to establish a list of relevant factor groups from the statistical selection, given the limited expression of non-agreement. Thus, we start from the detailed analysis of the constructions in which the plural marks did not occur, and some percentage rates specifically in Funchal.

3.1. Patterns of agreement in European Portuguese: a variable rule?

In EP, as observed above (Table II), the overall results lead to postulate a semi-categorical verbal agreement rule. Table III below shows all instances of absence of plural marks, in a total of 22 occurrences in the continental varieties of EP.

Table III: Data of non-agreement in Cacém (CAC) and Oeiras (OEI) samples

CACÉM / OEIRAS	Occurrences	FACTORS LANGUAGE					Type of Verb
		I	II	III	IV	V	
		External Sandi	VS	“THAT”	[-anim]		
CAC-A3M	(1) <i>e depois a partir da estação existe os autocarros para os diversos sítios</i>	X	X		X	x inac exist	
CAC-A3M	(2) <i>não existe cinema não existe teatro não existe...exposições esse tipo de não existe</i>	X	X		X	x inac exist	
CAC-C1H	(3) <i>chegou ao fim do mês não pagou aquela/quinhetos euros só pagou trezentos fica os outros dizentos a pagar</i>	X	X		X	x cop	
CAC-C3H	(4) <i>existe vários tipos de população relativamente à classificação que deveremos desenvolver aplicar</i>		X		X	xinac exist	
CAC-A1M	(5) <i>tem umas feições/zinhas que identifica logo que elas são brasileiras</i>			X			
CAC-B1M	(6) <i>vai ver a bola...claro juntamente com outros amigos que se junta a ver a bola</i>	X		X			
CAC-B2M	(7) <i>oitenia mil escudos equivale neste momento a quatrocentos euros</i>	X			X		
CAC-B3M	(8) <i>é assim a de antigamente tinha as suas partes chatas que era aquelas regras completamente absurdas</i>	X	X	X	X	x cop	
CAC-B3M	(9) <i>os meus pais não nos deixavam ir pra rua brincar não nos deixava se calhar receber amigos em casa</i>						
OEI-A3M	(10) <i>saiu algumas gramáticas alteradas mas ninguém está a usar nada</i>	X	X		X	x inac	
OEI-B1M	(11) <i>porque eu já sabia de quem era não sabia de quem era os carros mas sei sei que é dos vizinhos</i>	X	X	X	X	x cop	
OEI-B1M	(12) <i>um bocado a situação que aparece pessoas que a gente pensa às vezes que nem que não precisam</i>		X			x inac	
OEI-C1M	(13) <i>portanto continua os tempos a evoluir cada vez mais não é</i>	X	X		X	x cop	

CACÉM / OEIRAS	Occurrences	FACTORS LANGUAGE					Type of Verb
		I	II	III	IV	V	
		External Sandi	VS	“THAT”	[-anim]		
OEI-B2M	(14) <i>ah ah há coisas que noto que leva muito tempo para se conseguir uma consulta hum</i>	X		X	X		
OEI-C1H	(15) <i>os hábitos que não me agrada muito a é é – neste momento que estamos a atravessar é a falta de respeito</i>	X		X	X		
OEI-C2H	(16) <i>pelo menos por aquilo que me dizem e por aquilo que vejo ah – coisas que – no meu tempo – era muito diferente</i>	X		X	X		x cop
OEI-C3M	(17) <i>e eles são cumpridores (...)quando têm que estudar não – primeiro estudam – o que é bom também (D-isso é ótimo) nunca foge nem pouco mais ou menos não é?</i>	X					x intr
OEI-C3M	(18) <i>os livros dantes eram um bocadinho agora vejo a diferença acho achava que era assim era um bocadinho massudados não é?</i>				X		x cop
OEI-B3M	(19) <i>acho que antigamente as pessoas tinham mais tempo mais tempo porquetava mais em casa</i>	X					x cop
OEI-C1M	(20) <i>para esses há crise e também aquelas pessoas que vieram de fora também chegou aqui também não têm nada não é?</i>	X					x inac
OEI-A3H	(21) <i>estes óculos 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 está em promoção</i>	X			X		x cop
OEI-B1M	(22) <i>as coisas que a gente conhece também não é coisas assim é os carros</i>	X	X		X		x cop
Total		17	12	7	15		16

Observing the rare occurrences with no number marks in data from Oeiras and Cacém, nine and thirteen respectively, we can propose that they constitute specific marked contexts, structures that universally³ favor the absence of plural mark, namely: external Sandhi (column I), when the verb is followed by a [+ vowel] or a [cons, +nasal] at the beginning of the next word, in which there is no anchor of the auto-segment /N/ to the node of rhyme in the syllabic structure (Mota, Miguel & Mendes, 2012: 171), in 17/22 occurrences (77.2%) – a phenomenon which affects mainly verbs of lower phonic saliency; post-verbal subject (column II) in 12/22 cases (54.5%); the presence of the relative “that” – “que”(column III) in 7/22 cases (31.8%); non-animate subject (column IV), in 15/22 cases (68.1%), and type of verb (column V), in particular, intransitive, unaccusative and copula verbs (Cardoso, Carrilho & Pereira, 2011), in 16/22 (72.7%). Besides these linguistic features, we can identify examples with pre-verbal subject, in which the involvement of cognitive factors, typical of oral discourse, are likely to be observed – factors concerning “planning of discourse and quantity of linguistic material between the subject and the verb”⁴ (Mota, Miguel & Mendes, 2012: 169).

The application of verbal agreement rule within varieties of EP leads, however, to a polarity among continental and island varieties, being the continental ones closer to the categorical application of agreement rule and the island variety, in the threshold of application of the variable rule. That polarity inside the European varieties is confirmed by the rates related to the action of linguistic factors, namely those of syntactic-semantic nature, subtler as to the role of the *Semantic feature* of the subject (Table IV) and stronger when the variable *Position of the subject* is concerned (Table V).

Table IV: Third person plural verbal marking in EP and *Semantic feature of the subject*

Variety / Sample		Subject [+animate]	Subject [-animate]
EP	OEI	1249/1253 99.7%	205/214 95.8%
	CAC	970/972 99.8%	206/213 96.7%
	FNC	728/746 97.6%	138/168 82.1%

³ See Brandão & Vieira (2012:1061), for evidence based on Corbett (2000:201), explored at the end of this paper, which supports the universality of some structural variables that act in the use of verbal agreement.

⁴ All the translations presented in this article are of the authors' responsibility.

Table V: Third person plural verbal marking in EP and *Subject position*

Variety / Sample		Pre-verbal subject	Post-verbal subject
EP	OEI	777/782 99.4%	36/40 90.0%
	CAC	731/735 99.5%	40/44 90.9%
	FNC	550/573 96.0%	18/34 47.2%

In this analysis, a particularly relevant trend observed in the data from Funchal emerges, with regard to the absence of the plural marks in the context of the post-verbal subject (47.2%).

In the context of post-verbal subject, data from Cacém (Table III) present three noteworthy occurrences of the unaccusative verbal item *existir* without plural mark (see examples (1), (2) and (4) in table III) in articulation with the subject position (post-verbal, in column II) and its semantic nature ([-animate] in column IV). This phenomenon is widely diffused in insular varieties. Thus, considering only the post-verbal subjects, there is a total of 20 occurrences of *existir* in 34, i.e. 58, 8% (eleven with plural mark and nine without plural mark). Of all eighteen occurrences without plural mark, this item would thus be responsible for 50% of the absence of plural mark in this variety.

The morphophonological conditioning in the context of external Sândhi – reduction of the nasal diphthong of the 3rd person plural morpheme motivated by the phonetic context to the right of the verb –, that operates in most of the data of the two continental varieties (Table III, Column I), also occurs in the insular variety in 68.7% of the cases. As shown in (1), we can observe the lack of plural marking in verbs with grade 1 of *phonic saliency* (concerned only when adding the nasal marks, with or without changes in vowel quality), regardless of other syntactic configurations:

- (1) a. followed by [+vowel] (21 occurrences)
as mercearias na altura fechava às onze (FNC-B1M)
“grocery stores closed at eleven”
- b. followed by [cons, + nasal] (7 occurrences)
houve crianças as que tinha mais (FNC-B1M)
“there were children who had more”
- c. followed by pause (2 occurrences)
os cá de fora que se lixe (FNC-C3M)
“those outside screw them”

Like other continental EP dialectal varieties (Mota, Rodrigues & Soalheiro, 2003; Mota, Miguel & Mendes, 2012; Rodrigues, 2012), non-standard inflectional variants can occur in insular urban varieties. These non-standard variants are of two types:

- i. a variant represented by -EM, corresponding to the nasal diphthong [ẽ̃j] as P6 of verbs with thematic vowel (VT) /e/ and /i/, which occurs also in verbs with thematic vowel /a/, a pattern that reflects the convergence or leveling of variants;
- ii. a variant represented by -U, which may correspond to nasal [ũ] or oral [u] vowels, in which the nasal reduction depends on the phonetic context that follows it (vowel or nasal consonant).

These two patterns of non-standard variants represent about 9% of the data, in other words, 83 out of 866 tokens of verb agreement marks, and it is almost double of occurrences without the standard 3rd person plural mark (5.3%).

The use [ẽ̃j] of the standard form of P6, present in the paradigm of the Simple Present verbs with VT /e/ and /i/, extends to verbs with VT /a/, establishing a convergence in P6 marking. This levelling process in the marking of P6 is also observed in the paradigms of Past Tenses, replacing standard P6 [ẽ̃w̃] in these paradigms (Table VI).

Table VI: Diffusion of non-standard P6 variants in different verbal paradigms (Funchal)

Non-standard P6 variants	Present Ind.	Past continuous			Past perfect			Total
		VT /a/	VT /e/	VT /i/	VT /a/	VT /e/	VT /i/	
-EM variant	19	24	16	9	4	-	3	75 occ.
-U variant	-	5	1	2	-	-	-	8 occ.

The use of the non-standard variant with oral vowel [u] or nasal [ũ] is limited to the *Pretérito Imperfeito*. The paradigms affected correspond mostly to low phonic saliency contexts (concerned only when adding the nasal marks, with or without changes in vowel quality) except for eight occurrences (one of the type – *vaiem-se* by *vãõ* – and seven, corresponding to contrasts of the 3rd person of *Pretérito Perfeito do Indicativo*). In Funchal’s variety, patterns of variants bending P6 are found and that would be, in the case of *Pretérito Imperfeito*, the following: standard [ẽ̃w̃] and non-standard [ẽ̃j] nasal diphthongs; and without nasal diphthong (= standard P3), and with non-standard uses [u] and [ũ] as illustrated by the examples in (2) – (4).

(2) Variant with no mark (variant without diphthong, P3 = P6)

- a. followed by a vowel:
 elas à que chegava à escola (FNC-B1M)
 “they came to school”

- b. followed by nasal consonant:
[problemas] que penaliza mais os seus clientes (FNC-C3H)
“*problems that penalize more and more your customers*”
 - c. followed by a pause:
os cá de fora que se lixe (FNC-C2M)
“*outsiders go to hell/go and get screwed*”
 - d. followed by non-nasal consonant:
mas os dias foi passando (FNC-A1M)
“*but the days were going/passing by*”.
 - e. the verb *existir* (“to exist”) :
mas existe tarifários melhores (FNC-A3M)
“*but there are better fares*”
- (3) –EM “convergent” variant
- a. mas já trabalharem aqui nestas zonas todas (FNC-C1H)
“*but you have already worked in all these areas*”
 - b. pessoas a trabalhar comigo que tinhem vergonha (FNC-C2M)
“*people working with me that feel ashamed*”.
 - c. elas que vaiem-se matando (FNC-A1M)
“*they should go and kill themselves*”
- (4) –U “conservative” variant
- a. eles vinho brincar (FNC-B1M)
“*they come out to play*”
 - b. os meus avós ero padrinhos da minha irmã Gorete (FNC-B1M)
“*my grandparents were the godfathers of my sister Gorete*”

As for the variants without a diphthong, a similar process to that described for a variety of Braga can be at stake (Rodrigues, 2012: 221). In this case, contrary to the standard process, /N/ is associated with the nucleus. [o] and [u] forms result from a process of denasalization of the nasal diphthong [w] followed by the reduction of the diphthong in labial vowels. The form [ũ] would correspond to a process of reduction of the diphthong but without denasalization.

If demonstrated within the insular variety of Funchal, this variant could be considered as “conservative”. It would result from the diffusion of certain aspects observed in some sub-varieties of northern dialects of EP, usually characterized by features that have not undergone some of the changes in the dialect that would constitute the standard EP.

In the case of other variants, without diphthong, on one hand, and with convergence diphthong, in verbs with VT /a/, /e/ and /i/ of the Simple Present Tense (which extends to other verbal paradigms), on the other hand, the association or not of the floating autosegment /N/ – following the proposal made for standard EP by Mateus & Andrade (2000: 133) – may be at stake. According to this proposal, no diphthong variants may result from the

lightness that characterizes final nasal diphthongs without stress, in the 3rd person plural verb forms in standard EP. These diphthongs, but also those that occur in simple nominal forms (*homem* – “man”) or *suffixed words* (*paragem* – “stop”, *lavagem* – “wash”)⁵, are considered post-lexical due to the fact that they occur in words without a thematic constituent (as opposed to lexical nasal diphthongs, generated in the lexicon and that attracts the stress to the final syllable (*irmão* – “brother”) and admits, as the only segment in coda, the fricative /S/ (*irmãos* – “brothers”). In post-lexical diphthongs, the semi-vowel is epenthetic and comes after the process of diphthongization, with both segments occupying a single position in the nucleus, given the lack of a vowel marker of thematic class at the lexical level. The floating autosegment /N/ projects itself over the syllabic nucleus, nasalizing both segments simultaneously. The fact that this nasal autosegment only projects itself in the nucleus prevents any segment in the coda position to be nasalised⁶.

As for extra-linguistic variables, it can be observed certain influence of both *Education* and *Sex/Gender* variables (Table VII and Table VIII, respectively) in the insular variety.

Table VII: Third person plural verbal marking in EP and *Education*

Variety / Sample		Fundamental (5 to 8 years)	Intermediate (9-11 years)	Higher (12 to 15 years)
EP	OEI	411/417 98.6%	386/388 99.5%	657/662 99.2%
	CAC	374/377 99.2%	390/391 99.7%	412/417 98.8%
	FNC	212/236 89.8%	257/268 95.9%	397/410 96.8%

⁵ The reduction of nasal diphthongs of this type also occurs in the insular variety, but with a lower frequency as those observed in the Southern Brazilian varieties (Swindt, Bopp da Silva & Quadros, 2012:353). In these varieties, the lexical class seems to act in the reduction of nasality, where nominal forms are more subject to reduction than P6 verbal forms.

⁶ On the other hand, it may be possible to consider, going beyond the lightness of the diphthongs, another type of instability that can affect oral diphthongs when followed by a fricative in coda, such as [ajf] > [af] (*mais* > *mas*) and [ojf] > [of] (*depois* > *depos*) (Boleo, 1975:338), that also characterizes southern and central and insular dialects of EP (Terceira from the Azores archipelago), and in the south of Brazil (Brescancini, 2009). Also of note is another phonetic phenomenon that consists in the reduction of non tonic final vowel [u] (Lindley Cintra, 1990:181) perceptible as [ɨ] or non expressed, and which affects any words which contains it, independently from its lexical class. Regarding verbs, 1st person singular and plural forms of the present are performed without the vowel [u] (*falo* > *fale*; *tenho* > *tenhe*; *falamos* > *falames*, *temos* > *temes*).

Table VIII: Third person plural verbal marking in EP and *Sex/Gender*

Variety / Sample		Male	Female
EP	OEI	757/760 99.6%	697/707 98.6%
	CAC	523/525 99.6%	653/660 98.9%
	FNC	454/472 96.2%	412/442 93.2%

The statistical effects of extra-linguistic variables⁷ performed in the Funchal sample single out this variety within the EP. A more detailed observation of the data shows that informants with a low education level are responsible for higher rates of the absence of plural marks, and also higher values of non-standard 3rd person plural variants in structural contexts characterized by utterances with pre-verbal subject and non-expressed subjects (Table IX).

Table IX: Non-standard and absence of third person plural verbal marking in pre-verbal subject contexts and null subject and *Education* – Funchal (EP)

<i>Education</i>	With non-standard plural marks		With zero plural marks	
	Tokens	%	Tokens	%
Fundamental (5 to 8 years)	51/82	62.2	20/30	66.6
Intermediate (9-11 years)	31/82	37.8	5/30	16.7
Higher (12 to 15 years)	–		5/30	16.7

When considering the total values of non-standard inflectional variants of the 3rd person plural and homophones variants of the 3rd person singular (resulting from the absence of verbal agreement mark), the effect of the *Sex/Gender* variable is quite evident. Women are responsible for strategies that do not conform to the standard EP (Table X).

⁷ The 5.3% rate of non-mark occurrences in the Funchal sample allowed a variationist analysis which demonstrates the statically relevant performance of these two extra-linguistic variables. Goldvarb analysis ordered all variables by statistical relevance as follows: *position of the subject, education, animacity, gender and discursive parallelism*.

Table X: Non-standard and absence of third person plural verbal marking and *Sex/Gender* – Funchal (EP)

<i>Sex/Gender</i>	Non-standard plural marks (–EM ; –U)		Zero marks	
	Tokens	%	Tokens	%
Women	65/82	79.3	30/48	62.5
Men	17/82	20.7	18/48	37.5

These values allow us to consider, within the insular urban variety, the existence of “a **socio-linguistically polarized** situation” (emphasis added), reinforced by the fact that “the level of education (...) is associated with other variables such as socio-economic status, greater or lesser access to cultural goods, more or less contact with individuals from different social strata.” (Brandão & Vieira, 2012a: 1054). It seems clear that within the Madeiran society there is a boundary that separates speakers with little schooling, stigmatized by their vernacular, and the elites that use consolidated standard EP. The city of Funchal has over time been source of internal migration, where rural populations converge. These populations, concentrated in the periphery, contribute to economic activities linked to trade and services (Census 2001 and 2011).

In contrast with the widespread proposal by Labov (1990: 205), where women in the same social group tend to be more conservative with regards to their preference for standard variants in relation to stigmatized ones (non-standard), the Funchal sample (Table XI) demonstrates precisely the opposite: in the community of speakers with low education, women demonstrate marginal linguistic behaviours, leading the use of non-standard variants (vowel -U and nasal diphthong –EM) with a percentage of 69.2% and, with lower projection, the variant without plurality mark (56%).

Table XI: Non-standard and absence of third person plural verbal marking and *Sex/Gender* in *Fundamental level Education* – Funchal (EP)

Fundamental Education (5-8 years)	With Non-standard marks (–EM; –U)		Zero marks (no diphthong variant, P3 = P6)	
	Tokens	%	Tokens	%
Women	36/52	69	14/25	56
Men	16/52	30.8	11/25	44

This empirical data, as advocated by Lucchesi (2012) when confronted with the role of *Gender* in urban varieties of BP, should be understood in the light of the social role assigned to the genera in the specific socio-historical settings of each community. In the case of Brazil, Lucchesi (2012: 799)

further states that there has been an ongoing linguistic leveling, due to the influence developed by large urban centres over other regions, and in this context:

The processes of marked social variation (such as the nominal and verbal agreement) point to processes of “top down” and “out to in” of these rural peripheral communities of major Brazilian cities. Thus, men, who on average are better placed in the labour market and move further out of the community, tend to better assimilate the forms of urban variants, quickly abandoning the use of deviant rural and popular language forms.

The weak implementation of the verbal agreement variable in informants with low education levels has not been statistically attested in continental EP varieties, but it is widely studied in varieties of BP, and parallels another study by Vianna (2011) on the implementation of *a gente* (form that corresponds to “we”) in varieties of Portuguese. In Vianna’s (2011) study, the *locality* factor is statistically relevant, clearly distinguishing the insular variety from the two continental varieties of EP (Cacém and Oeiras). The *education level* is a determinant factor in the sample of Funchal, which has a 52% occurrence of *a gente* in informants with a low level of education compared with the 32% and 13% observed in Cacém and Oeiras, respectively. The *gender* factor, as observed in the implementation of the third person verbal agreement, also conditions the choice of variant *a gente*, contrasting the behaviour of informants from Funchal with other urban varieties of EP. The contrast in gender manifests itself in the female preference for the *a gente* (51%) (Cacém, 30%, Oeiras 8%), while males prefer the prestigious “nós” (we) form (91%), values that are similar to those obtained in other varieties (Cacém, 86%; Oeiras, 91%).

Andrade (1990), in his analysis of the non-standard palatalization phenomenon of /l/, identified as one of the characteristic features of the island variety, whose optionality is conditioned by the presence of preceding oral and nasal [i] or [j], whether belonging or not to the word (*quilo* – “kilo” > [kiʎu]; *vi logo* – “I saw immediately” [viʎɔgu]). This author also demonstrated the influence of extra-linguistic factors in this process. In this case, and taking only gender into account, it seems that “overall, women palatize more than men in all styles considered” (Andrade, 1990: 28).

It is important to note that in the context of this study, the existence of P6 inflectional variant patterns may result from the influence of extra-linguistic social variables (*Education* and *Sex/Gender*), whether or not related to historical aspects of dialect construction, which begins in the fifteenth century with the settlements in Madeira Island, in the “cycle of language expansion” (Castro, 2006), and in discontinuity with the processes of formation of standard and dialectal varieties in continental Portugal. In parallel to the development of the standard variety of EP, based on distancing from “not only the Castilians but also from the northern dialects spoken by

the ancient Portuguese nobility” (Cardeira, 2009: 111) and with the contribution of Portuguese expansion (through linguistic contact in non-European areas but also within the Portuguese territory with the presence of African slaves, and whose “presence in society would, inevitably, have some impact in fostering linguistic simplification and smoothing, strengthening the drift of the meridional Portuguese.” (Cardeira 2009: 112)), the Portuguese spoken by the first settlers, which was succeeded by other waves of immigrants from various parts of Portugal, was in contact⁸, since the beginning of its implementation in the insular space, with many people of great ethnic diversity (European settlers, attracted by the possibilities of commercial development, Genovese Florentines, French, Spanish, English, but also, Guanches, Moors and African slaves) (Vieira, 1997, 2001).

To sum up this section, it is important to say that the attention given to data analysis of verbal agreement variants in the insular EP variety in this paper is an attempt to increase the knowledge of non-standard dialects from EP, a task deemed as necessary by Galves (2012). Finally, regarding the extra-linguistic constraints, gender and education variables seem to be differentiating markers between insular and continental EP varieties explained by specific historical-political and social conditions during the insular community construction. Looking forward to meeting Galves (2012: 134) request, we sought to deepen in this section the analysis of phonetic factors which supposedly affect verbal agreement in EP. A detailed examination of the three samples of the EP confirms that “the phonic saliency itself does not seem to play no role in dialectal Portuguese”. Indeed, increasing instances of the absence of plural marks, for example, could not be observed (as if it would avoid examples like *eles vai* / “they goes”; *eles cantou* / “they singed”) – a result that seems to contribute to the interpretation of the agreement rule in varieties of Portuguese.

3.2. Patterns of agreement in BP and STP: restrictions for variation

Just as occurred for the nominal agreement (see Brandão, this volume), a variable 3rd person plural agreement rule is observed in the communities outside of the current territorial domains of Portugal.

In the case of BP, the rates of agreement marks (78.2% in NIG; 88.1% in COP) are consistent with trends verified in other studies on verbal agreement

⁸ About the settlement of Madeira Island and the relative characterization data of the population in the XV to XVIII centuries, see Pinto & Rodrigues (1993) and Ferraz (1990). In both studies, percentages of slaves living in Madeira following this period are presented. This linguistic contact also occurred in Continental Portugal specially in the city of Lisbon, but it is probable that the impact of this contact was larger in the smallest population of Funchal and due to its distinct social structure (Lower number of Noblemen, also called “homens bons” – good men) (Gonçalves, 1958).

in urban communities' speech. This speech usually shows a quite distinct linguistic behaviour from those attested in rural areas. The high rates of plural marking in Brazilian samples are characteristic of urban varieties and are primarily practised by educated individuals. These rates sharply contrast with those obtained in the speech of other communities from rural or peripheral areas (see Bortoni-Ricardo, 2005).

Brazilian social stratification shows striking differences concerning the degree of education and social classes, on one hand, and the opposition rural vs. urban, on the other. These differences reveal themselves linguistically in a *continuum* of greater or lesser use of prestigious variants, especially regarding agreement marks. In communities of speakers with little education level and / or more rural characteristics, the rates of plural marks are often quite different from those of highly educated speakers who come from urban (or "rurban") areas.

In STP sample, it seems that high rates of plural marks (92.1%) result from the strong influence of the European standard agreement norm, which serves as a model of learning the Portuguese language as L2.

Despite of socio-historical and linguistic differences between São Tomé and Brazilian varieties, it is important to point out that the variable agreement rule seems to be constrained by very similar structural and social restrictions. The selection of relevant variables according to the statistical analysis, as can be seen in the table below, demonstrates that similarity.

Table XII: Operating variables for implementing third person plural verbal marking by sample/variety

Copacabana	Nova Iguaçu	São Tomé
<i>Education</i>	<i>(Discursive and clausal) Parallelism</i>	<i>(Discursive and clausal) Parallelism</i>
<i>(Discursive and clausal) Parallelism</i>	<i>Configuration of the subject</i>	<i>Subject position</i>
<i>Verbal Phonic Saliency</i>	<i>Verbal Phonic Saliency</i>	<i>Education</i>
<i>Age</i>	<i>Education</i>	<i>Sex/Gender</i>
<i>Sex/Gender</i>	<i>Subject position</i>	<i>Semantic feature of the subject</i>
<i>Subject position</i>	<i>Semantic feature of the subject</i>	
Significance: .00 Input: .97	Significance: .02 Input: .90 – .88	Significance: .01 Input: .97

Although in different order, several variables – *Discursive and Clausal parallelism*, *Subject position*, and *Education level* – operate both in BP and STP. Furthermore, *Semantic features* of the subject and *Sex/Gender* were

highly significant in two out of three samples. Finally, Brazilian varieties particularize themselves on the diversity of contexts in which the singular-plural alternation occurs, so that specific structural variables, such as *Phonic salience* (COP and NIG), operate. In the following sections, we do not include results concerning variables which operated only in one sample, such as *Age* (COP) and *Configuration of the subject* (NIG).

3.2.1. Social Variables

a) Education

The effect of education level in African and Brazilian varieties is quite enlightening, as it points out social conditions. It is quite different from what we can observe in continental EP, where being educated or not does not change the semi-categorical rule of agreement.

Table XIII: Third person plural verbal marking in EP and *Education* by variety of Portuguese

Variety / Sample		Fundamental (5 to 8 years)	Intermediate (9-11 years)	Higher (12 to 15)
BP	COP	299/410 72.9% .17	340/382 89.0% .41	590/603 97.8% .79
	NIG	333/457 72.9% .32	243/362 67.1% .35	491/546 89.9% .68
STP	ST	179/215 83.3% .14	220/236 93.2% .49	280/286 97.9% .81

In general, one can say that the effect of *Education level* surely favors agreement in the three samples, though in a slightly different way in NIG. While the increasing effect of education could be clearly observed in increasing rates of plural marks in COP, a region of high social-economic prestige, in NIG, a region of low socio-economic prestige, the factor that effectively demarks the preference for agreement is the access to Graduation level. The influence of education has been strongly demonstrated in several investigations, either to sustain the Brazilian sociolinguistic polarization (Lucchesi *et al.*, 2009), or to investigate the stability and / or change of agreement patterns (Naro & Scherre, 2003). In this regard, authors have shown that urbanization and access to school lead to increase the use of agreement marks by all social classes.

In the case of STP, where the variable *Education level* acts almost identically to Copacabana, it should also be noted that access to education implies, in reality, a greater use of the Portuguese language itself, instead of the mother tongue acquired by the informants. For this reason, the more or less contact with the Portuguese language was controlled in our analysis

through the declarations produced by the informants themselves. Results below (Table XIV), which refers to the total of interviews available in *Corpus* VARPOR (STP), show that the frequency of Creole use is closely correlated to the use of agreement marks: the higher this frequency, the lower the use of plural marks.

For clarification, it is important to say that all individuals from ST sample declared themselves speakers of Portuguese as L1. However, regarding the linguistic plurality in São Tomé society, where three creoles and other minority languages coexist (Hagemeier, 2009) and where the use of the Creole language is considered non-prestigious, the present study controlled the possible influence of contact with Portuguese following the proposal by Brandão (in this volume). According to it, there are individuals who claim that they (a) only use Portuguese and use one or no Creole language (zero or low frequency), (b) use Portuguese and one or more Creole with lower frequency, or (c) use Creole more often – in the market, with friends (high frequency).

Table XIV: Third person plural marking and *Frequency of Creole use* in PST (all informants)

Frequency of Creole use	Tokens	Perc.	RW
Zero / Low	465/503	92.4%	.57
Intermediate	572/634	90.2%	.48
High	71/89	79.8%	.35

Considering also the intimate relationship between access to education and greater use of the Portuguese language in São Tomé society, a cross tabulation of data concerning *Frequency of Creole use* and *Education level* variables (Table XV) was developed.

Table XV: Third person plural marking in PST (all informants): *Frequency of Creole use* versus *Education*

<i>Frequency of Creole use</i>	<i>Education</i>					
	Fundamental (5 to 8 years)		Intermediate (9-11 years)		Higher (12-15 years)	
Zero / Low	122/139	88%	740/222	92%	138/142	97%
Intermediate	181/217	83%	237/260	91%	154/157	98%
Higher	71/89	80%	----	----	----	----
Total	374/445 84%		442/482 – 92.7%		292/299 – 97.7%	

The data show that none of the informants of the intermediate and higher levels of education reported using a high frequency of Creole, which confirms the utterance that education is a sort of passport to mastering the Portuguese language in São Tomé. Low educated informants' speech confirm

that greater frequency of use of Creole leads to lower rates of verbal 3rd person plural marks.

It must be emphasised here that the correlation between these variables constitutes, in a certain extent, an empirical evidence for the hypothesis that the greater the linguistic contact, the lower the use of agreement marks. This hypothesis has been employed in several studies which claim that the non-agreement in BP arises from the strong linguistic contact amongst African, Indian and European populations.

b) Sex/Gender

As stated earlier, there is apparently no difference between the agreement rates reported for men and women in the NIG samples, where this difference is inferior to 3%. In STP and COP, as observed in the study of nominal agreement, a relevant statistical difference between men and women can be verified (Table XVI).

Table XVI: Third person plural verbal marking and *Sex/Gender* by variety of Portuguese

Variety / Sample		Male	Female
BP	COP	761/882 86.3% .43	468/513 91.2% .63
	NIG	483/607 79.6%	584/758 77.0%
STP	ST	380/424 89.6% .37	299/313 95.5% .68

In the cases of COP and ST, women (.63, .68 respectively) tend to produce greater agreement marks than men (.43, .37). This result confirms the Labovian hypothesis that, although women are often the most innovative in processes of linguistic change, they are very conservative when dealing with a variant of low prestige.

The fact that this trend was not observed neither in Nova Iguaçu, in which men and women behave approximately in the same way, nor in Funchal, where an opposite trend was found, may indicate different gender patterns inside social groups. Apparently, although the hypothesis should be validated with specific sociological research, the higher the inclusion of women in social, economic, professional activities (as seems to be the case of Copacabana and São Tomé), the greater the tendency to effectively use plural marks, which is a strong indicator of high cultural and education levels.

3.2.2. Linguistic variables

a) *Discursive Parallelism (from verb to verb)*

The effect of the parallelism principle within the verbal agreement – from both subject to verb (*Clausal Parallelism*), and verb to verb (*Discursive Parallelism*) – has been widely attested in samples of BP. The results shown in the following table not only confirm such an effect in Brazilian data, but also attest it in the STP sample.

Table XVII: Third person plural verbal marking and *Discursive parallelism* by variety of Portuguese

Variety Sample	/	Plural marks in the preceding verb	No plural marks in the preceding verb	Isolated verbs or first verb of a series
BP	COP	359/359 100%	6/74 8.1%	864/962 89.8%
	NIG	269/277 97.1% .92	4/121 3.3% .01	794/967 82.1% .52
STP	ST	171/171 100%	15/15 0%	508/551 92.2%

We can observe the clear effect of presence / absence of plural marks in verbs in series in the BP and STP samples. This effect is evidenced by the categorical effect, in COP and ST, or by contrasting relative weights, in NIG (.92 vs. .01).

Apparently, in any sample in which an alternation between singular and plural forms occurs, one can observe the effect of parallelism, because this group of factors hypothetically is of a more cognitive-procedural nature than a strictly linguistic one. In this sense, although it was not possible to observe its operation in continental EP, due to the lack of effective variable data, Funchal has shown this effect in a small group of occurrences.

b) *Clausal parallelism (from subject to verb)*

Unlike European samples – in which the effect of subject marks on verb marks could not also be observed, due to the categorical nominal agreement (see Brandão, this volume) –, the *Clausal parallelism* is also highly active in BP and STP samples.

Table XVIII: Third person plural verbal marking and *Clausal parallelism* by variety of Portuguese

Variety / Sample		Plural marks in the subject	No plural marks in the subject	Numeral in the subject
BP	COP	918/991 92.6% .54	13/33 39.4% .14	5/7 71.4% ---
	NIG	807/954 84.6% .54	29/65 44.6% .14	7/11 63.6% ---
STP	ST	462/490 94.3% .52	8/15 53.3% .17	1/1 100% ----

As indicated by the relative weights, the parallelism effect of the presence or absence of subject plural marks is evident both in BP and STP samples. According to it, non-marked subjects (.14, .17) favor non-plural marking of the verb, when compared with marked subjects (.54, .52).

c) *Subject position*

As Table XIX shows, while post-verbal subject disfavors the use of plural marks (.16 in COP; .17 in NIG, and .08 for ST), pre-verbal subject, with values above the neutral point of .50, favors it.

Table XIX: Third person plural verbal marking and *Subject position* by variety of Portuguese

Variety / Sample		Pre-verbal subject	Post-verbal subject
BP	COP	916/1007 91.0% .52	37/44 84.1% .16
	NIG	830/1003 82.8% .52	18/39 46.2% .17
STP	ST	462/492 93.9% .66	29/39 74.4% .08

The relationship between this variable and the *Transitivity* of verbs is always conceived as a hypothesis: it seems that there is a logical correlation between post-verbal subjects and unaccusativity and/or intransitivity. Statistically, however, not only in this study but also in other researches (Scherre & Naro, 2007), it has been proven that, in Brazilian samples, the effect of subject position is usually higher than that of transitivity: “(...) independently from the type of verb, any argument or NP on the right of the

verb tends, relatively, to reduce the marks of overt agreement” (Scherre & Naro, 2007: 312).

As shown by the data of this study, this effect can be observed even in continental EP (Cardoso, Carrilho & Pereira, 2011) and especially in Funchal sample. In fact, as shown by Corbett (2000), the position of the subject (besides its semantic *animacy* feature) changes the agreement patterns in languages from different genealogies, so that it seems that this influence does not mean an effective variability but a grammatically distinct structural behaviour.

For this reason, this research chose to group separately the data by the presence or absence of post-verbal subjects. It should be noticed that all the other relevant variables act in the presence or absence of such structures. However, the exclusion of post-verbal subject was fundamental to attest the particularity of European samples, in the sense that the apparent variation in Funchal, for example, results largely from the presence of post-verbal subjects, as already demonstrated in the previous section.

d) Semantic feature of the subject

The results below (Table XX) show the influence of semantic feature of the subject regarding *animacy* of the referent in NIG and ST samples.

Table XX: Third person plural verbal marking
and *Semantic feature of the subject* by variety of Portuguese

Variety / Sample		Subject [+animate]	Subject [-animate]
BP	COP	1035/1175 88.1%	194/220 88.2%
	NIG	948/1187 79.9% .53	119/178 66.9% .36
STP	ST	575/612 94.0% .56	104/125 83.2% .26

Relative weights, which are similar, confirm the influence of the semantic feature of the subject: non-animate subjects disfavor verbal plural marks (.36 and .26, in NIG and in ST, respectively), while animate subjects encouraged these marks (.53 in NIG and in ST .56, respectively).

The semantic-discursive nature of this variable conveys, we suppose, a sort of saliency in information processing, according to which it would be highlighted, by agreement marks, the dynamic feature (of control) of verbal predicators. Not coincidentally, this effect can also be minimally observed in

European samples and was attested in different languages, as evidenced by Corbett (2000).

e) Verbal phonic saliency

The principle of phonic saliency (Lemle & Naro, 1977) – whereby the greater the phonic differentiation between singular and plural forms, the higher the tendency for agreement – does not reach the apparent “universality” which is observed in the groups of linguistic factors previous analysed. The results of this investigation show that the effect of phonic salience occurred exclusively in Brazilian varieties.

Table XXI: Third person plural verbal marking and *Verbal phonic saliency* by variety of Portuguese

Variety / Sample		Grade 1	Grade 2	Grade 3	Grade 4
BP	COP	669/795 84.2% .35	160/173 92.5% .51	180/192 93.8% .79	220/235 93.6% .70
	NIG	592/812 72.9% .29	193/245 78.8% .55	144/156 92.3% .78	138/152 90.8% .74

Except for grades 3 and 4, in which there is no increase in the relative weights of plural markings, the general principle was demonstrated: increasing phonic material differentiation between singular and plural forms of verbs increases probability of overt plural marker (in COP: .35> .51> .79/.70; in NIG: .29 > .55 > .78/.74).

Several studies have presented the relevance of this variable in Brazilian samples. It is curious to note, however, as demonstrated in the study by Scherre & Naro (2006), that this group of factors is not usually pointed as statistically significant in samples with educated informants’ speech. Apparently, as occurs in the European and São Tomé varieties, the massive preference for agreement makes the cases of non-plural marking mostly concentrated in the initial levels of phonic differentiation, so that increasing degrees of salience are unobservable.

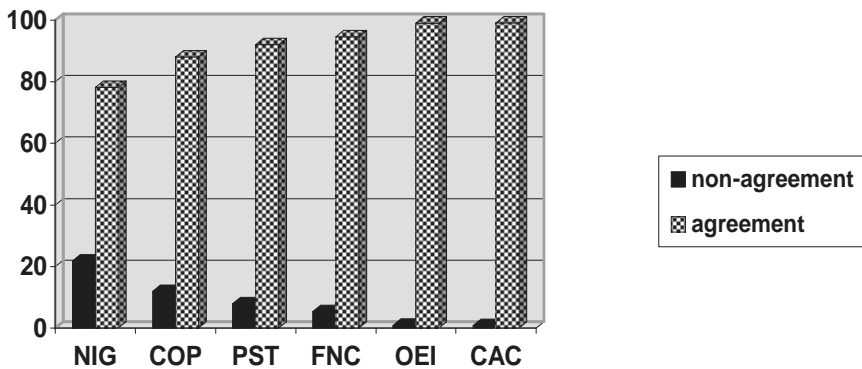
4. Systematization of the results and contrastive approach of third person plural agreement in the varieties of Portuguese

Based on the results of this paper, associated, wherever possible, with conclusions from other studies on the subject, it is proposed in this section a contrastive approach between the 3rd person plural agreement patterns observed in Portuguese varieties.

4.1. *The use of plural verbal marks in varieties of Portuguese: quantitative aspects*

Frequencies of 3rd person plural marks in the samples of this study show that, of all literate and educated urban varieties, the European ones differ from many others. This is easily observed in the following figure.

Figure 1: Non-agreement and agreement in third person plural verbs by variety of Portuguese



In this graph, a *continuum* which goes from the end of plural marking variation, represented by Brazilian varieties, to the end of a clear semi-categorical rule, represented by varieties of continental EP, can be stated. The varieties of São Tomé and Funchal are in intermediate position, the former alternating forms of expression in a variety of contexts far more extensive than the latter.

Evaluating if these results are compatible with other scientific descriptions is an important task to avoid incipient conclusions. For this reason, it is important to compare, although briefly, the general rates obtained here with those of other investigations which describe different samples.

In the case of STP, results have shown that agreement patterns are similar to the ones verified in BP. This similarity, not only in verbal agreement but also in the noun phrase, is very relevant, especially to evidence that the loss of inflectional marks is typical of situations of intense linguistic contact. In this section, however, the little knowledge on the expression of 3rd person in African varieties of Portuguese as a whole does not allow comparative approach nor further discussion.

Otherwise, the variable rule of agreement has been widely described in representative samples of different regional and social Brazilian speech communities. Although detailed presentations of the studies are not possible, due to the limits of the present article, we can summarize (see Table XXII)

the general rates obtained in studies that demonstrate the broad spectrum of variation from more urban to more rural populations in Rio de Janeiro, which also correlates to the education level of the informants.

Table XXII: Frequency of plural marks in different studies on third plural verbal agreement in BP

Variety		Sample/Study	Education level	Plural marking
BRAZIL	URBAN	Sample Mobral – 1970 (Naro, 1981)	All data Illiterate	<u>48%</u>
		Peul – 1980 (Scherre & Naro, 2006)	All data	<u>73%</u>
			1-4 years	63%
			5-8 years	78%
		9-11 years	81%	
		NURC / RJ – 1970/80 (Graciosa, 1991)	Higher	<u>94%</u>
	Peul – 2000 (Naro & Scherre, 2005)	All data	<u>84%</u>	
1-4 years		75%		
5-8 years		85%		
9-11 years	93%			
Copacabana – 2010 (Vieira & Bazenga, this volume)	All data	<u>88.1%</u>		
	5-8 years	72.9%		
	9-11 years	89.0%		
12-15 years	97.8%			
Nova Iguaçu – 2010 (Vieira & Bazenga, this volume)	All data	<u>78.2%</u>		
	5-8 years	72.9%		
	9-11 years	67.1%		
12-15 years	89.9%			
NO URBAN	Corpus APERJ – 1980 (Vieira, 1995)	All data (0-4 years)	<u>38%</u>	

Naro (1981) and Vieira (1995) investigated the illiterate speech; while the first study focuses on speech in the metropolitan area, with the MOBREAL

sample, in the 1970s, the latter investigates the speech of fishermen in the Northern Region of the state of Rio de Janeiro, with the *Corpus* APERJ, of the 1980s. As for the speech of individuals who accessed formal education, Scherre & Naro (2006), besides the present research, investigated the speech of individuals with elementary and intermediate level of education in Rio de Janeiro, from approximately 1980 and also from approximately the end of the twentieth century, and Graciosa (1991) deals with the speech of individuals with higher education, based on *Corpus* NURC-RJ.

Third person plural occurrences from the speech of Rio de Janeiro, in the second half of the twentieth century or earlier, reveal increasing rates of agreement in a *continuum* that goes from rural areas to more urban ones and from the least educated speech to the most educated one. This *continuum* is proposed by Lucchesi, Baxter & Silva (2009:348), who, using results from other Brazilian studies and including those related to the speech of the isolated Afro-Brazilian community of Helvecia/Bahia, interpret the distribution of data as revealing a Brazilian sociolinguistic polarity. The following figure, which includes general agreement rates shown in Table XXII and further results from Helvecia, also demonstrates this *continuum*.

Figure 2: *Continuum* of third person plural verbal marks according to rural / urban features and educational levels in Brazilian varieties based on Lucchesi, Baxter & Silva (2009:348)

[+marks]				▶	[- marks]
RJ	RJ	RJ	RJ		Helvecia/Bahia
Urban	Urban	Urban	Rural		Rural
Highly educated	Moderately educated	Illiterate	Illiterate		Afro-Brazilian communities
94 / 97.8%	73 / 89%	48%	38%		16%

With regard to the few descriptions of EP samples, we can observe the following works⁹: Varejão (2006) and Gandra (2009)¹⁰, both based on *Corpus* CORDIAL-SIN, which includes dialectal (popular) varieties of spoken Portuguese in the entire length of Portugal and islands; Monguilhott (2009)¹¹, that considers speakers with elementary or intermediate school

⁹ Only the studies of other researchers have been selected, given that the previous investigation developed by the authors of this article relied also on part of the basic sample investigated here, which was reconfigured in this article.

¹⁰ The results have been categorized in the “rural” axis because all or most of the points considered by Gandra and by Varejão are really situated in rural area.

¹¹ The author conducts interviews with informants from Cascais, Sintra, Belém (Lisbon neighborhood) and central Lisbon. While considering the first two points as rural, we chose to represent the data set in Table XXI in urban axis, since mostly represented speech can be considered at least “rurban” in the sense employed in Bortoni-Ricardo (2005).

education levels and born in the Metropolitan Region of Lisbon; Rubio (2012), who describes the speech of informants from several regions of Portugal, from the Reference Corpus of Contemporaneous Portuguese (CRPC), distributed according to four grades of education level; and Monte (2012), which considers illiterate informants, with secondary education (Young adults education – EJA) and higher education levels, in the region of Évora. The table below summarizes the percentage of 3rd person plural marks for each study.

Table XXIII: Frequency of plural marks in different studies on third person plural verbal agreement in EP

Variety	Sample/Study	Education level	Plural marking	
PORTUGAL	URBAN	All data	<u>91%</u>	
		Metropolitan Region of Lisbon (Monguilhott, 2010)	Fundamental Higher	89-90% 92-94%
		Évora (Monte, 2012)	All data Illiterate Education (EJA) Secondary education	<u>93.1%</u> 91.2% 93.5% 94.4%
		Oeiras / Lisbon, Cacém (Vieira & Bazenga, this volume)	All data Fundamental Intermediate Higher	<u>99.2%</u> 98.9% 99.6% 99.1%
	Funchal (Vieira & Bazenga, this volume)	All data Fundamental Intermediate Higher	<u>94.7%</u> 89.8% 95.9% 96.8%	
	NO URBAN	Several regions – Cordial-SIN (Varejão, 2006; Gandra, 2009)	All data (0-4 years)	<u>91%;</u> <u>96.5%</u>
		Several regions – CRPC (Rubio, 2012)	All data 1-4 years 5-8 years 9-11 years 12 or more	<u>93.9%</u> 91.9% 93.9% 93.7% 96.2%

As can be seen in the table above, there is a high achievement score of the plural marks in all studies. Independently of the region and its profile (more urban or more rural communities), all samples present rates of marks over 90%. In view of the thresholds used in this study for differentiating a variable rule (up to approximately 94%) and a semi-categorical one (from 95% on), based on Labov (2003), it is necessary to discuss the results ranging around 91-94%, that would slightly differ the results of this article from those mentioned works.

Concerning the results related to education levels, the significant proximity between them prevents a safe proposal of this variable's effect. Except for the results from Funchal, whose particular characteristics have been previously discussed, the results are very similar (with variation below 5%). Considering all studies, the preference for agreement can be confirmed regardless of the speaker's education level.

Observing specifically the methodological guidelines of each investigation, it is clear that quantitative differences result from procedures adopted regarding data collection. Confirming this statement is the fact that studies which use the same *corpus* (though not always with equivalent locations), as is the case of Gandra (2009) and Varejão (2006), can obtain different results: 91 and 96.5%, respectively¹². Undoubtedly, the inclusion and exclusion of such data shows different interpretation of structures involving the 3rd person plural. For this reason, in the next section, a qualitative approach of contexts that have been verified in studies of each variety is presented.

4.2. The relationship between structural and social variables and the definition of social varieties: qualitative aspects

In general, while in Brazilian samples the achievement of presence or absence of plural marks occurs in diverse structural contexts, and in similar amounts, due primarily to education level and to rural vs. urban opposition, as already demonstrated, it is necessary to evaluate, in detail, the results for European Portuguese, which have demonstrated to diverge in certain aspects. Based on this assessment, it is possible to debate the proposal, conveyed in studies such as Naro & Scherre (2007), that the differences between Brazilian and European varieties of Portuguese are more quantitative than qualitative.

Therefore, in order to define the qualitative extension of the variable contexts in European samples, three procedures were adopted: (i) a survey of structural types included in data collection, (ii) the observation of the

¹² Bazenga (2012), who also drew upon data from the *Corpus Concordância*, had previously registered more than 10% of non-agreement in the Madeiran variety, rates that relates primarily to the fact that data collection included *has/have* verb and constructions which was preceded by topicalized Noun Phrase.

occurrences of unmarked verbs, and (iii) to check the profile of variables that were statistically significant in works with Labovian orientation.

4.2.1. Data collection for the study of the third person plural agreement

In the table below, it can be observed the contexts that – according to the authors – were excluded from the data collection of works with European samples.

Table XXIV: Criteria for excluding data in studies on third person plural verbal agreement in EP

Structural context excluded from data collection	Sociolinguistic studies
Third person plural verbs indicating arbitrary reference of the subject	Monguilhott (2010) Monte (2012) Rubio (2012)
Passive constructions with <i>se</i>	Monguilhott (2010) Varejão (2006)
Semantic agreement with singular name (<i>o pessoal</i>)	Monguilhott (2010) Monte (2012)
Homophone verbal forms (like <i>tem / têm</i>)	Monte (2012) Varejão (2006) Rubio (2012)
Sentences produced by the informant which repeats the verbal form of question asked by the interviewer	Monte (2012) Varejão (2006)
Sentences with 'to have' conveying existential value	Monte (2012)
Subject represented by the indefinite pronoun 'tudo' referring to a third person plural noun phrase (for ex. <i>os filhos e seu marido são tudo Benfica</i>)	Monte (2012)
Partitive expressions in the subject	Monte (2012)
Personal infinitive verbs	Monte (2012)
Coordinated and post-posed subject with the first constituent of the noun phrase in the singular (<i>morreu a senhora e o patrão</i>)	Monte (2012) Varejão (2006)
Interrogative pronouns in the subject	Monte (2012)
Sentences in which there is a possible agreement with a expletive pronoun (<i>ele falta-me umas peças</i>)	Varejão (2006)

Firstly, the differentiated degree of detail regarding the data exclusion criteria employed in each investigation draws the attention in these results. It is undoubtedly the work of Monte that best explains such criteria. Strictly speaking, if it is taken for granted that the contexts not mentioned by the authors were considered in the data collection, the global rates could not be safely compared. It happens that, by observing the examples given in these studies, it can be assumed that, although some exclusion contexts have not

been mentioned, other works most probably have not considered them too. Meanwhile, there are in fact structures interpreted as variables in one study, and non-variables in others.

Rubio (2012), for example, unlike Monte (2012), usually considers occurrences such as the following, where the verb also relates to the generic reference *tudo* (“everything”) in the context after the verb:

- (5) *é. e depois há, há isto que, que parece-me, parece-me que tem importância, é que, dantes a farmácia era a farmácia oficina, on[de], onde ha(...), havia... como sabem... agora os medicamentos é quase tudo especializado [CRPC-1082-20]*
“And then there is, as it seems to me, it seems to me what matters is that the chemist was once a workshop before, where there was, as you know... now medication is specialized in almost anything”. [CRPC-1082-20]

In general, the basis for the exclusion of data is the impossibility of determining the reference of the subject as effectively plural, which ultimately generates an ambiguous construction by its very nature, where the referent can be singular or plural (a kind of semantic agreement), or unspecified/generic (indeterminate 3rd person plural; infinitive cases)¹³.

Comparing these works with the present article, this investigation – although it initially contemplated synthetic passive – is particularized to extend the data exclusion criterion also to the following cases: apparent 3rd person plural subjects (overt or null) which refer to topicalized constituents; copula constructions, which can be interpreted as presentational (*os caras é isso aí...* – “the guys is that's it ...”); constructions whose determination of the verb as singular or plural is dubious for phonetic-phonological reasons. Apparently, the justification for lower rates of agreement in other studies is the interpretation of such ambiguous contexts, especially with regard to the first two structural types mentioned here. If all mentioned works eliminate or treat in a particular way, *a priori*, the contexts considered in this paper as ambiguous or as difficult to determine the status of the singular or plural verb, certainly all researches would register higher rates of plural marking, which would confirm the status of the semi-categorical agreement rule in EP.

As observed in the data collection criteria, the absolute amounts presented in the studies already represent *a priori* the qualitative delimitation of the agreement phenomenon, so that the elements of the quantitative-qualitative pair are interdependent. Accordingly, the supposedly variable status of this

¹³ The various contexts of apparent agreement variation are treated in detail in Martins (2012), which validates the exclusion of virtually all cases of Table XXIV, except for the infinitive data (whether personal or impersonal), which are not mentioned by that author, and the interrogative pronouns, which the author himself (Monte, 2012) admits that were primarily excluded because they were very few in number.

phenomenon in EP is defined by the structural interpretation of the occurrences in question.

In order to better observe the quality of contexts found in each study, one must examine the occurrences of non-agreement usually found in European samples.

4.2.2. Patterns of unmarked verbal contexts in EP

In the present study, it was possible – by listing all occurrences without the plural mark in Lisbon and Cacém, and by categorizing data without explicit ending verbal form or without the standard ending verbal form in Funchal – to control the quality of the absence of plural marking occurrences, similarly to the work by Vieira (2011). From what has been detailed in the description of the results, one can list the most common constructions: (i) VS structures composed mostly of unaccusative verbs; (ii) presence of subjects with semantic feature [- animate]; (iii) the presence of the verb followed by a phoneme (nasal consonant, vowel) that can change the quality of the final vowel of the verb and not setting a verbal form identical to the singular.

Besides these contexts, the examples found in this or other studies report the presence of the relative pronoun *que* (“that”) – that favor the singular form, by inaugurating a relationship of the verb with its referent, mediated by a non-inflected number form – or the presence of complex noun phrases – influence that was safely detected in other investigations into the phenomenon.

Moreover, comparing the results of this study with those of others, it appears that the mark cancellation rates correspond to constructions with the verb *ser* (“to be”) or with other verbs preceded by noun phrase which can be interpreted as topic, if not eliminated by the data collection criterion.

It is important to note that, even when collecting occurrences of this nature, scholars admit that their rates may be closely related to these specific structural types. Rubio (2012: 372) states that “for the third person plural verbal agreement, the variation is semi-categorical due to being restricted in only a few contexts”. Gandra (2009: 153) warns that, in occurrences with the verb *ser* (“to be”), “an uncertainty between singular and plural sense of the element to which it is referring” constantly occurs. Monguilhot & Coelho (2012: 313) admit that, in data analysis of Lisbon samples, the syntactic context with copula verbs presents the agreement with the predicative, and analysed the hypothesis that there would be in the position of the subject a neutral null pronoun (like *isso* – “that”). Monte (2012), recognizing the specificity of the phenomenon in constructions with the verb *ser* (“to be”), isolates these constructions from other verbal occurrences.

Since these are the types of structures that would indicate number mark cancellation in EP, it is essential to interpret the variable or invariable character of such structures. The present paper reaffirms the position taken by Vieira (2011) and echoes the interpretations that these utterances do not

signal a lack of agreement, at least in the same form of non-marked constructions such as those occurring in Brazilian (and African) sentences like *os menino comeu maçãs, mas nós não come essa fruta* (“the boy ate apples, but we do not eat this fruit”). The particularities of the contexts found in the European samples are located in three levels: syntactic, in the case of the so-called post-verbal subject; semantic, in cases involving, to some extent, ambiguity in singular or plural references; and phonetic-phonological, utterances in which the achievement of the number-personal endings questions the equivalence or difference between the singular and plural forms.

For cases of supposedly post-verbal subjects, there is a vast literature that questions the status of noun phrase postponed to the verb and argues that this syntactic constituent has features more comparable to a direct object, than actually to a subject. Monte (2009: 76-77) demonstrated that many scholars defended this thesis. Generative studies have shown that this construction results from the association of an unaccusative verb with an internal argument, and not really a subject, which in this case would be a null expletive. In these theoretical field or not, it is essential to take into account the differentiated status of these structures, which may occur in any variety of Portuguese and even in other languages.

According to Galves (2012), in an article that proposes a thorough reflection on the empirical facts and theoretical notions involved in the debate about the origins of BP, if it is the result of a “continuing process of EP (drift) or even if it is a process of rupture (contact)” (p. 123), the analysis of unaccusative constructions also extends to those with the verb *ser* (“to be”):

The analysis of accusative verbs extends easily to the verb *to be*, with the difference that this verb also takes clauses as complements. All cases in which the verb is singular can *be* analysed as constructions in which its subject is a null expletive, and the agreement is made with this expletive. (op.cit: 131).

Besides the differentiated syntactic status of the structures commented, we must resume the cases of ambiguity in the interpretation of the subject as singular or plural not only with *ser* (“to be”) but also with other verbs. Martins (2012: 97), in an article that details many cases where the same utterance may materialize different syntactic structures, setting cases of structural ambiguity, argues that:

They were ambiguous as to the distinction singular/plural pronoun subjects who return a topicalized constituent (as being distinctive when it comes to the topicalized number form as a noun phrase, on the one hand, and the subject pronoun, on the other). It is the pronoun that takes back the topic, with the function of the subject, laying the subject-verb agreement, whether it is phonologically realized, or a null pronoun.

In this sense, the author also argues – in conformity with what Mota & Vieira (2008) suggest and Mota (2012) and Martins (2012) proposes – that structures like *Peras em vinho é a minha sobremesa favorita* (“Pears in wine is my favourite dessert”) may receive the interpretation equivalent to “pears in wine, [null pronoun] is my favourite dessert.” Galves (2012) also deals with such construction – as in the sentences *os nossos agasalhos [exp] é estes; duas canas [exp] dá para pescar* (“our coats [exp] is these; two reeds [exp] gives to fish”) – comparing it with phrases from French. Doing so, she admits the analysis that the Noun Phrase before the verb would be the topic of the sentence, not the subject.

Although the alternative forms (singular or plural) may be apparent, it would not be considered a variation in the subject-verb agreement in this and other cases. Rather, it would be considered the postulation of the Luso-Brazilian grammatical tradition, in which the subject-verb agreement can be made sometimes with the idea and the meaning as in the so-called “ideological agreement”.

Finally, as to the utterances in which the occurrences of number verbal endings, in terms of phonetic-phonological, question the equivalence or difference between the singular and plural forms, the recognition of the influences that the following context imposes to the verb (especially consonants and nasal vowels), illustrated in this study, as well as a detailed description of the pronunciations of the EP verbal number endings (see Mota & Vieira, 2008; Mota, this volume) make clear the need to describe more accurately the increasing degree of phonic differentiation between singular and plural verb forms in varieties of Portuguese.

4.2.3. Analyzing the statistically significant variables’ profile

The verification of the variables’ profile indicated as statistically significant in Brazilian and European studies can justify the recognition of variable contexts of agreement. Considering three of the studies based on Brazilian samples constituted in the second half of the twentieth century, we can observe relevant linguistic variables at each level of education.

The Table XXV shows the variables relevant in all Brazilian social varieties: order of the subject in relation to the verb (which, in some studies, includes the presence of the relative *que* (“that”)), distance between the subject and the verb (or correlated variables, such as the presence of pause), discursive parallelism, semantic feature of the subject. Except for highly educated speech, in the Brazilian samples the phonic saliency and clausal parallelism factors also act. In the present study, safeguarding the characteristics of the *corpus* – covering elementary, intermediate and high education levels together –, the effect of these variables is confirmed.

Table XXV: Relevant variables in studies of third person plural verbal agreement:
Brazilian samples – different levels of education

Higher education (Graciosa, 1991)	Intermediate education (Scherre & Naro 2006)	Non-educated speech (Vieira, 1995)
<i>Subject position</i>	<i>Verbal phonic saliency (except for 9-11 years)</i>	<i>Verbal phonic saliency</i>
<i>Distance: subject-verb</i>	<i>Subject position</i>	<i>Clausal and discursive parallelism</i>
<i>Discursive parallelism</i>	<i>Clausal and discursive parallelism</i>	<i>Subject position</i>
<i>[Animacity – or not controlled in the data]</i>	<i>Animacity of subject referent</i>	<i>Animacity of subject referent</i>
	<i>[Presence of relative que – “that”]</i>	<i>Distance: subject-verb</i>
	<i>Presence of pause: subject and verb</i>	

It is relevant to realize that the influences of social order – especially education level – accelerate or retract the implementation of the linguistic restrictions. Apparently, the difference in behaviour observed in samples of individuals with high education relates to the assessment of agreement in each speech community. Since it is considered a stereotype (Labov, 1972), the absence of plural marks occurs in the Brazilian educated speech only in constructions that can be considered non-marked or slightly prominent, those that rise to the level of consciousness only in the most monitored situations: (i) post-verbal “subject” (*chegou os livros*; “the books arrived”); (ii) pre-verbal subject when distant from the verb, represented by the relative *que* (“that”) with discrete phonic differentiation between singular and plural forms (*o rapaz e a moça que sempre de manhã pode ver o sol a nascer* / “the boy and girl who always can see the sunrise in the morning”). Thus, in the BP, what constitutes the indexical feature of illiteracy / uneducated is the absence of plural marks in salient contexts: pre-verbal subject (especially when not marked), next to the verb with a high degree of phonic saliency (i.e. *os menino saiu* / “these boy went out”).

The patterns observed in data produced by Brazilian graduates are similar in quantity and quality to data observed in European samples. The table below (XXVI), which presents only the results of studies that took advantage of the statistical selection of groups of factors possibly acting in EP, confirms this similarity.

Table XXVI: Relevant variables in studies of third person plural verbal agreement: European samples

Monguilhott (2009)	Rubio (2012)	Monte (2012)
<i>Semantic feature of the subject</i>	<i>Subject position</i>	<i>Subject position</i>
<i>Subject position</i>	<i>Semantic feature of the subject</i>	<i>Semantic feature of the subject</i>
<i>Type of verb</i>	<i>Structural type of noun phrase subject</i>	<i>Structural type of noun phrase subject</i>
		<i>Type of verb ('to be' vs. other verbs)</i>
		<i>Verbal phonic saliency</i>

The statistical treatment in works that have adopted the variationist method shows that, in EP, the groups of factors that act systematically are the ones related to the position and the semantic feature of the subject. Secondly, it is also revealing the selection of variables such as type of verb and structural type of the subject, in two of the three works, which, in general, allowed to assess the influence of contexts with coordinated noun phrase and singular elements, with the relative *que* (“that”), and also with the unexpressed subject. The control of the type of verb enables to validate the particularity of copula constructions in EP.

As already stated here and in Brandão & Vieira (2012a), the status of group of factors related to the *Position* and *Semantic feature of the subject* does not contribute to the definition of the characteristics of EP, BP or any variety. Actually, it is an influence of universal nature¹⁴, which occurs in languages and linguistic varieties of absolutely distinct origins, well exemplified by Corbett (2000)¹⁵:

¹⁴ Several scholars have hypothesized universal variable principles, either in general (trans-linguistically) or in particular (in each language or language type). See, in this regard, Chambers (2004), Szmrecsányi & Kortmann (2009), Trudgill (2009).

¹⁵ We thank Professor Dante Lucchesi for suggesting the reading of Corbett (2000).

Table XXVII: Agreement with conjoined noun phrases
(Corbett 2000: 201)

	[+animate]		[-animate]	
	N	% PL	N	% PL
Subject-predicate				
Medieval Spanish	288	96	243	31
German	1,095	96	1,702	67
Russian	115	100	67	85
Serbo-Croat	21	100	35	91
Predicate-subject				
Medieval Spanish	318	69	239	6
German	379	93	925	40
Russian	89	84	114	28
Serbo-Croat	23	70	62	26

As for the group of *Verbal phonic saliency*, selected systematically only in the work of Monte (2012), it is understood that its supposed relevance is bound to an overlap with the control of two constructions: (i) structures with copula verb *ser* (“to be”), (ii) presence of post-verbal subject and/or structures with possible interpretations of topic and null subject expletive *isso* (“that”).¹⁶ Moreover, it is emphasized that the observation of the characteristics regarding the pronunciation of the personal number desinences, made in this article and substantially considered in Mota (this volume), reinforced the need to reconfigure the levels of phonic differentiation between singular and plural forms taking into account the phonetic-phonological features present in the European varieties, as it has been previously proposed.

5. Final remarks: the nature of verbal agreement rules in English varieties

After the extensive description of the 3rd person plural agreement in Brazilian, European and São Tomé varieties of Portuguese, we believe it is necessary to highlight the contributions of the investigation so that new scientific questions can be established.

Concerning the particular characteristics of European data that differentiate them from BP and STP, we can postulate relevant quantitative

¹⁶ In Monte (2009), the only three non-marked verbs involving high phonic salience, in the group of data with no occurrences of the verb *ser* (*to be*), are of the following types: post-posed subject, in two cases: (*começo os anos a passare*), and subject possibly interpreted as an expletive *isso* (*já tem se feito (algumas) greves mas nunca deu... (isso) nunca deu resultado...*).

and qualitative evidences to support the proposal that there are two patterns of agreement in Portuguese: a semi-categorical rule, typical of EP, and more evident in continental varieties used by all the population, educated or not; and another one, a variable agreement rule, typical of BP and STP, conditioned by specific linguistic and social restrictions. Linguistic restrictions make agreement and non-agreement variation applicable to a wide diversity of conditioned structures, and extra-linguistic variables are especially related to the level of education and the urban *vs.* rural opposition.

The semi-categorical pattern – or a categorical one, as we can observe in nominal agreement (Brandão, this volume) or in 1st person plural verbal agreement, with 100% of plural marks with the pronoun *nós* (“we”) (Vianna, 2009; Rubio, 2012; Marcotulio, Vianna & Lopes, this volume) – can also be found in educated persons’ speech in Brazil, to whom the absence of plural mark is a stereotype. In low educated people’s speech, especially from rural areas, the pattern of variable agreement rule is widely observed. Thus, according to Brandão & Vieira (20012b), it can be postulated to BP, and also to STP, the coexistence of a semi-categorical standard prestigious pattern of agreement *versus* a natural variable rule.

In the case of STP, we can observe that a variable rule pattern also operates, with strong preference for agreement, and, at the same time, data are clearly influenced by social conditioning. Thus, being educated in São Tomé and Príncipe clearly promotes the knowledge of Portuguese itself and, consequently, the use of agreement according to the European norm. Anyway, the presence of a variable rule in several structural contexts constitutes an evidence to support that the loss of inflectional marks are typically concerned with situations of intense linguistic contact.

Considering all the results presented, it is inevitable to conclude that the patterns of agreement in EP (and in the BP spoken by highly educated people) does not share the same qualitative or quantitative profile of spontaneous, vernacular BP. Social / extralinguistic restrictions (such as *education, geographic region* in terms of rural *vs.* urban opposition, *sex/gender*) and the presence or absence of non-marked forms in specific morph-syntactic structures (verbs with a high *phonic saliency* in singular/plural opposition; pre-verbal subject, without material between subject and verb) allow us to establish Brazilian own characteristics.

On the other hand, no evidence could be found in order to postulate the social value of agreement variation in EP. In other words, the contexts in which we face the supposed non-agreement cannot reveal anything regarding social status of the EP, except for insular varieties, where some differences have been observed. The relevant question is: What can general linguistic constraints – such as the position of the subject, semantic feature of the subject or even discursive parallelism – say about historical origin of Portuguese varieties, like BP and EP? As a matter of fact, these variables are concerned with syntactic, semantic and phonetic specific structures. The

range of factors acting in the three EP varieties seem to obey morpho-phonological constraints (external Sandhi) and syntactic-semantic generic types (which looks like a kind of “universal” constraints), which cannot identify each of these EP varieties (see Brandão & Vieira, 2012a).

Accordingly, the data presented in this study show no evidence to support the proposal that the expansion of European structures would have been in the original constitution of the BP. At least in contemporaneous data, it seems unjustifiable to suppose that from such specific contexts (such as, for example, the post-verbal subject, *to be* particular constructions, or the presence of the topicalized SN and null expletive subject) Brazilian variable patterns of agreement have emerged. Although phonic saliency may even operate in European data, it does not seem sustainable to propose that an occurrence like *nós sai* (“we go out”) came from an occurrence like “*eles canta*” (“they sing”). Appealing to analogy (as some authors do), in this case, seems to go back to the period when theories and scientific methods were insufficient to explain the linguistic data. According to Lucchesi (2012: 263), “invoking the analogy, so dear to Philology of the nineteenth century, in this case, is not only an anachronism, but a theoretical and epistemological foolishness, because, doing so, it is possible to explain anything (...)”.

Considering the Principle of “Uniformitarianism”, the different patterns of agreement support the thesis of different historical processes of formation. According to it, BP, given the extent of contexts without plural marks and the configuration of different social norms, similar to what occurs today in STP, clearly reflects the profile of a polarized linguistic contact, which opposed and still opposes the rural and urban areas.

Finally, we hope that the description of the significant amount of data collected under the Project *A comparative study of patterns of agreement in African, Brazilian and Europe varieties* may have helped to solve the problem of empirical validation of the existing interpretive proposals. Doing so, we hope it has metonymically contributed to the understanding of several issues related to the relationship between linguistic variation and the characterization of social dialects. Though the interpretation of the constitution in the history of languages and their varieties remains an exciting challenge to linguistic studies, it is expected that this article has provided relevant information to the scientific evidence of the hypotheses formulated.

References

- Andrade, E. (1990) Algumas particularidades do Português Falado no Funchal. *Actas do IX Encontro Nacional da Associação Portuguesa de Linguística (APL)*. Lisboa: Colibri, 17-29.
- Bazenga, A. (2012) Variation in subject-verb agreement in an insular dialect of European Portuguese. In *Non-dominating Varieties of pluricentric Languages*.

- Getting the Picture. In memory of Prof. Michael Clyne* (M. Rudolf, ed.), pp. 327-348. Wien et. al: Peter Lang Verlag.
- Bazenga, A. (2011) Concordância verbal e variantes de 3ª pessoa do plural em PE: Resultados preliminares de um estudo sociolinguístico com base numa amostra de Português Falado no Funchal. *XVI Congreso Internacional de la Asociación de Lingüística y Filología de la América Latina* (ALFAL). Alcalá de Henares. Madrid 7 de junho de 2011.
- Boléo, M. de P. (1974) *Estudos de linguística portuguesa e românica: dialectologia e história da língua*. Coimbra: Universidade de Coimbra.
- Bortoni-Ricardo, S. M. (2005) Um modelo para a análise sociolinguística do português brasileiro. In *Nós chegemu na escola, e agora?: Sociolinguística e Educação*, pp. 39-52. São Paulo: Parábola Editorial.
- Brandão, S. F. & S. R. Vieira (2012a) Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa: Revista de Linguística. (São José Rio Preto)* [online]. 56 (3), 1035-1064. <http://dx.doi.org/10.1590/S1981-57942012000300013>
- Brandão, S. F. & Vieira, S. R. (2012b) Concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Papia*, 22.1, 7-40.
- Brescancini, C. R. (2009) A redução de ditongos decrescentes seguidos por fricativa em coda no açoriano catarinense. In *Português do sul do Brasil : variação fonológica* (L. Bisol & G. Collischonn, orgs.), pp. 34-49. Porto Alegre: EDIPUCRS.
- Cabeleira, S. (2006) Atitudes e preconceitos linguísticos face a variedades dialectais. *XXI Encontro da Associação Portuguesa de Linguística*, pp. 245-260. Lisboa: APL.
- Cardeira, E. (2009) Revisitando a periodização do português: o português médio. *Domínios de Linguagem*, ano 3, 2, 103-115.
- Cardoso, A., Carrilho, E. & S. Pereira (2011) On verbal agreement variation in European Portuguese: syntactic conditions for the 3SG/3PL alternation. *Diacritica* 25.1, 135-157.
- Carrilho, E. & Pereira, S. (2011) Sobre a distribuição geográfica de construções sintáticas não-padrão em Português europeu. *Textos seleccionados do XXVI Encontro Nacional da Associação Portuguesa de Linguística*. CD-ROM. Lisboa: APL.
- Castro, I. (2006) *Introdução à História do Português*. Lisboa: Colibri.
- Chambers, J. K. (2004) Dynamic typology and vernacular universals. In *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective*. (B. Kortmann, ed.), pp. 127-145. Berlin / New York. Mouton de Gruyter.
- Corbett, G. G. (2000) *Number*. Cambridge: Cambridge University Press.
- Ferraz, M. de L. F. (1990) A Cidade do Funchal na 2ª metade do século XVIII – Freguesias urbanas. *Actas do II Colóquio Internacional de História da Madeira*. Vol. I, pp.265-284. Funchal: CEHA.
- Ferreira, C. S. (2009) Percepções dialectais e atitudes linguísticas. O método da Dialectologia perceptual e as suas potencialidades. *Textos Seleccionados. XXIV Encontro Nacional da Associação Portuguesa de Linguística*, pp. 251-263. Lisboa: APL.

- Galves, C. (2012) Concordância e origens do português brasileiro. In *Por amor à linguística: miscelânea de estudos linguísticos dedicados a Maria Denilda Moura* (A. P. Sedrins *et al.*, orgs), pp. 123-150. Maceió: UEDUFAL.
- Gandra, A. S. (2009) A concordância verbal no Português Europeu rural. In *Novos tons de Rosa... para Rosa Virgínia Mattos e Silva* (K. Oliveira *et al.*, orgs), pp. 142-161. Salvador: EDUFBA.
- Gonçalves, E. (1958) Os «Homens Bons» do Concelho do Funchal em 1471. *Das Artes e da História da Madeira*. 5. 28, 1-8, 76-82.
- Graciosa, D. (1991) *Concordância verbal na fala culta carioca*. Dissertação (Mestrado em Letras Vernáculas). Universidade Federal do Rio de Janeiro: Rio de Janeiro.
- Guy, G. R. (1981) *Linguistic variation in Brazilian Portuguese: aspects of phonology, syntax and language history*. Tese (Doutorado em Linguística). Pennsylvania: University of Pennsylvania.
- Hagemeijer, Tjerk (2009) As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* 1:1, 1-27.
- Labov, W. (2003) Some sociolinguistic principles. In *Sociolinguistics: the essential readings* (C. P. Paulston, & G. R. Tucker, eds.), pp. 235-250. Oxford: Blackwell.
- Labov, W. (1990) The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2, 205-254.
- Labov, W. (1972) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Lemle, M. & Naro, A. J. (1977) *Competências básicas do português*. Rio de Janeiro: Fundação Movimento Brasileiro de Alfabetização.
- Lindley Cintra, L. (2008) Os dialectos da ilha da Madeira no quadro dos dialectos galego-portugueses. In *Cultura madeirense – Temas e Problemas* (J. E. Franco, coord.), pp. 95-104. Funchal: DRAC.
- Lucchesi, D. (2012) A Teoria da Variação Linguística: um balanço crítico. *Estudos Linguísticos* 41 (2), 793-805.
- Lucchesi, D. (2009) História do Contato entre Línguas no Brasil. In *O Português Afro-Brasileiro* (D. Lucchesi, A. Baxter, & I. Ribeiro, orgs.), pp. 41-73. Salvador: Edufba.
- Lucchesi, D., Baxter, A. & Silva, J. A. A da (2009) A concordância verbal. In *O Português Afro-Brasileiro* (D. Lucchesi, A. Baxter, & I. Ribeiro, orgs.), pp. 331-371. Salvador: Edufba.
- Lucchesi, D., Baxter, A. & Ribeiro, I. (orgs.) (2009) *O Português Afro-Brasileiro*. Salvador: Edufba.
- Martins, A. M. (2012) Aparente variação na concordância sujeito-verbo no português europeu: ambiguidade quanto ao carácter singular ou plural do sujeito frásico. In *Rosae: linguística histórica, história das línguas e outras histórias* (T. Lobo *et al.*, eds), pp. 191-219. Salvador: EDUFBA.
- Mateus, M. H. M. & Andrade, E. (2000) *The Phonology of Portuguese*. Oxford: Oxford University Press.
- Mateus, M. H. M. & Andrade, E. (1998) The Syllable Structure In European Portuguese. *DELTA* 14 (1), 13-32.

- Monguilhott, I. (2009) *Estudo sincrônico e diacrônico da concordância verbal de 3ª pessoa do plural no PB e no PE*. Tese (Doutorado em Linguística). Florianópolis: Universidade Federal de Santa Catarina.
- Monguilhott, I. & Coelho, I. (2012) Sujeito: entre ordem e concorrência. *DIADORIM* 8, 307-328.
- Monte, A. (2012) *Concordância verbal e variação: um estudo descritivo-comparativo do Português Brasileiro e do Português Europeu*. Tese (Doutorado em Linguística e Língua Portuguesa). Araraquara/SP: UNESP.
- Mota, M. A. & Vieira, S. R. (2008) Contrastando variedades do português brasileiro e europeu: padrões de concordância sujeito-verbo. In *Língua portuguesa: identidade, difusão e variabilidade* (C. A. Gonçalves & M. L. L. Almeida, orgs.), pp. 111-137. Rio de Janeiro: AILP/UFRJ.
- Mota, M. A., Rodrigues, C. & Soalheiro, E. (2003) Padrões flexionais nos pretéritos fortes, em pe falado setentrional. In *Razões e Emoção. Miscelânea de estudos em homenagem a Maria Helena Mira Mateus* (I. Castro & I. Duarte, orgs.), pp. 129-156. Lisboa: Imprensa Nacional.
- Mota, M. A., Miguel, M. & Mendes, A. (2012) A concordância de P6 em português falado. Os traços pronominais e os traços de concordância. *Papia* 22.1, 161-188.
- Naro, A. (1981) The social and structural dimensions of a syntactic change. *Language*, 57 (1), 63-98.
- Naro, A. & Scherre, M. M. (2007) *Origens do português brasileiro*. São Paulo: Parábola.
- Nina, T. (1980) *Concordância nominal/verbal do analfabeto na Micro-Região Bragançana*. Dissertação (Mestrado em Letras). Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul.
- Pinto, M. L. R. & Rodrigues, T. M. F. (1993) Aspectos do Povoamento das Ilhas da Madeira e Porto Santo nos séculos XV e XVI. *Actas do III Colóquio Internacional de História da Madeira*, Vol. III, pp. 403-471. Funchal: CEHA.
- Rubio, C. F. (2012) *Padrões de concordância verbal e de alternância pronominal no Português Brasileiro e Europeu: estudo sociolinguístico comparativo*. Tese (Doutorado em estudos linguísticos). São José do Rio Preto/SP: UNESP.
- Rodrigues, C. (2012) Variantes não-standard e tipo de discurso: (des)encontro de resultados. In *Nada na linguagem lhe é estranho. Estudos em homenagem a Isabel Faria*, pp. 215-228. Porto: Edições Afrontamento.
- Scherre, M. M. (1988) *Reanálise da concordância nominal em Português*. Tese (Doutorado em Linguística). Rio de Janeiro: UFRJ.
- Scherre, M. M. & Naro, A. J. (2006) Mudança sem mudança: a concordância de número no português brasileiro. *SCRIPTA* 9 (18), 107-129.
- Scherre, M. M. & Naro, A. J. (1997) A concordância de número no Português do Brasil: um caso típico de variação inerente. In *Diversidade linguística no Brasil* (D. da Hora, ed.), pp. 93-114. João Pessoa: Ideia.
- Schwindt, L. C., Silva, T. B. & Quadros, E. S. (2012) O papel da morfologia na redução da nasalidade em ditongos átonos finais no português do sul do Brasil. In *Vogais além de Belo Horizonte* (S. H. Lee, org.), pp. 349-359. Belo Horizonte: Editora da UFMG. Disponível em: www.lettras.ufmg.br/site/E-livros/VogaisAlemdeBH2012.pdf.

- Segura da Cruz, M. L. & Saramago, J. (1999) Açores e Madeira: autonomia e coesão dialectais. In *Lindley Cintra. Homenagem ao Homem, ao Mestre e ao Cidadão* (I. H. Faria, ed.), pp. 706-738. Lisboa: Cosmo.
- Szmrecsanyi, B. & Kortmann, B. (2009) Between simplification and complexification: Non-standard varieties of English around the world. In *Language Complexity as Evolving Variable* (G. Sampson, D. Gil & P. Trudgill, eds.), pp. 64-79. Oxford: OUP.
- Trudgill, P. (2009) Linguistic Universals and Vernacular data. In *Vernacular Universals and Language Contacts* (M. Filppula, J. Klemola & H. Pitkänen, eds.), pp. 323-348. Taylor & Francis Routledge.
- Varejão, F. (2006) *Variação em estruturas de concordância verbal e em estratégias de relativização no Português Europeu Popular*. Tese (Doutorado em Letras Vernáculas). Rio de Janeiro: Universidade Federal do Rio de Janeiro.
- Vianna, J. (2011) *Semelhanças e diferenças na implementação de A gente em variedades do português*. Tese (Doutorado em Letras Vernáculas). Rio de Janeiro: Universidade Federal do Rio de Janeiro.
- Vieira, A. (2001) As migrações e os descobrimentos portugueses. Séculos XV e XVI. In *Imigração e Emigração nas ilhas*, pp. 27-62. Funchal: CEHA.
- Vieira, A. (1997) Escravos. com e sem açúcar na Madeira. Disponível em: <http://www.madeira-edu.pt/Portals/31/CEHA/bdigital/hm-esc-1.pdf>
- Vieira, S. R (2011) O estatuto da regra variável e o fenómeno da concordância verbal em variedades do português. *Documentos para el XVI Congreso Internacional de la ALFAL*. CDROM. Alcalá de Henares: ALFAL.
- Vieira, S. R (1995) *Concordância verbal: variação em dialetos populares do Norte Fluminense*. Dissertação (Mestrado em Letras Vernáculas). Rio de Janeiro: Universidade Federal do Rio de Janeiro.
- Weinreich, U., Labov, W. & M. Herzog (1968) Empirical foundations for theory of linguistic change. In *Directions for historical linguistics* (W. Lehmann & Y. Malkiel, orgs.), pp. 97-195. Austin: University of Texas Press.

Silvia Rodrigues Vieira
Faculdade de Letras (Departamento/ Programa
de Pós-graduação em Letras Vernáculas) da
Universidade Federal do Rio de Janeiro
(Federal University of Rio de Janeiro/FAPERJ)
silviavieira@hotmail.com

Aline Bazenga
University of Madeira
alinebazenga@gmail.com