

Patterns of plural agreement within the Noun Phrase

SÍLVIA F. BRANDÃO

Abstract

We focus on plural agreement within the Noun Phrase in urban varieties of Portuguese, in the light of Theory of Variation and Change assumptions (Weinreich, Labov, Herzog, 1968) and based on the speech of individuals with fundamental, medium and superior education level, distributed, yet, for three age groups and sex. We suggest that nominal agreement is a categorical rule in European Portuguese and a variable rule in Brazilian Portuguese and in São Tomé Portuguese, although these two varieties tend to adopt the canonical plural marking. The variable rule implementation is strongly determined by the linear and relative position of the constituent within the Noun Phrase and dependent on the action of external factors, among which education level, that seems congregate different aspects that characterize individuals from social point of view.

1. Introduction

Since we attempted to infer the characteristics of Brazilian Portuguese (BP) markedly different from European Portuguese (EP) not only in the lexical and phonological domains but also at the morphosyntactic level, the agreement proved to be one of the most stimulating topics of research. The differences observed gave rise to the formulation of several hypotheses, in order to deduce the motivations of the phenomena detected and thus contribute to a better understanding of the plurality of rules that have been outlined over time and that now typify BP.

Both nominal and verbal agreement have been the object of countless analyses of philologists, dialectologists and creolists, which were intensified

in the last 40 years, especially in light of the Theory of Variation and Change and of the different ramifications of Generative Theory. Those analyses have exposed most of the constraints governing the presence/absence of the plural number mark and have brought important elements to the debate about the origins of BP. Structural and extra-linguistic constraints operate in all Brazilian regional and social dialects, distancing or approximating them according to the greater or lesser input of the rules. This means that agreement, both nominal and verbal, corresponds to variable rules and that the contrasting [-agreement] vs. [+agreement] patterns found in spoken BP constitute a strong indication of the social status of the speakers.

As we shall see in the following sections, in BP, and also in African varieties of Portuguese, there are different ways of marking the plural value within the Noun Phrase (NP), by using the plural agreement marker, the morpheme /S/, in every inflectional constituent of the NP (the canonical form of marking plural), as in (1) below, or by using the mark only in one or in some of the inflectional constituents of the NP, as exemplified in (2)-(4) below¹:

- (1) TodoS oS meuS amigoS estrangeiroS gostam de vir ao Brasil
(All my foreign friends like to come to Brazil)
- (2) Vou mostrar [aS minhaS rede() nova()]
(I will show my new fishing nets)
- (3) Esse peixe anda n[aS parte() mais baixa()]
(this fish swims in the lower parts)
- (4) [AquelaS ondaS perigosa()]
(those dangerous waves).

The variable character of the rule and the low percentages of the presence of the plural marker in some communities formed basically by individuals who are illiterate or have a low education level (in particular, those who live inland or are Afro-descendants) contrast to what is observed in BP cultivated variety and in EP. Those facts have served as a basis for formulating different interpretations (creolisation/decreolisation, light creolisation, irregular linguistic transmission, Romance drift) about the history of the Portuguese language in Brazil.

In a work on the subject, Naro & Scherre (2007: 32) claim that it is “perfectly possible [...] that Portuguese language already had the embryo of the new more analytical system before leaving Europe”². According to the authors, this would result from “the secular drift of the Romance languages –

¹ More examples can be found in section 4.1.

² This and the other quotes from Brazilian and Portuguese authors were translated into English by the author of this paper.

and of the Indo-European languages in general – toward the morphological uniformity, with the survival of the more salient “irregular” forms”. To sustain their thesis, they searched for evidence in medieval and dialectal Portuguese works but the result was an extremely limited set of data. Nevertheless, the authors concluded that “Brazilian Portuguese has always been the Portuguese language” (ibid: 179), meaning that the Portuguese language has always been the same in Brazil and in Portugal.

Lucchesi (2003) formulated an alternative and opposite hypothesis: the historical process observed in BP would fit into the concept of “irregular linguistic transmission”, triggered by an intense inter-linguistic contact which is reflected, until today, in the variable patterns present in the popular and even in the cultivated speech:

Despite the diversity of BP structures that may be related to the contact between languages, there is a basic aspect of the irregular linguistic transmission process which is crucial for understanding the historical development of the Portuguese language in Brazil. The starting point of any process of irregular linguistic transmission triggered by contact between languages is the loss of inflectional morphology in the target language acquisition by speakers of other languages (p. 281).

Lucchesi's ideas are supported not only by studies on Brazilian Afro-descendant communities or on Brazilian communities where a low education level is prevalent, but also by the similarity of the agreement patterns observed in other areas to where the Portuguese language has been transplanted and where it is spoken today as L1 or L2. Such are the cases of São Tomé and Príncipe, Mozambique, Angola, and Cape Verde.

Until 2008, when we began to establish the *Agreement corpora* used for the studies presented in this volume, very little had been written on the nominal agreement patterns that prevail in the speech of individuals coming from different social strata and living in large urban areas. In those areas (characterised by diversity, interaction and social mobility par excellence), we observe in general (and also in the case of agreement) not only a greater awareness of behavioural and/or linguistic patterns socially stigmatised/prestigious, but also (and consequently) a tendency for neutralising such patterns in terms of frequency, whenever alternating forms are involved.

Until 2008, socially stratified *corpora* were not available either to EP and African varieties. Consequently, it was impossible to measure what was happening in the referred varieties among less and more educated individuals, for example. In fact, in the first variety the object of evaluation was the performance of individuals with a higher level of education, while in the African varieties it was just the opposite group who was taken into account.

The comments presented in the subsequent sections start from some hypotheses which have been outlined as the different *corpora* used in this study were being analysed and compared (Brandão, 2011; Brandão & Vieira,

2012a; Brandão & Vieira, 2012b). In what concerns the urban variety, the hypotheses are the following: (a) EP keeps the canonical nominal agreement rule in both the continental and insular environment, whereas the non-European varieties are characterised by having variable patterns in greater or lesser degree in the speech of all social strata; (b) structural constraints that determine the variation of nominal agreement patterns are the same in any of the non-European varieties, which is thought to be a reflection of the situation of Portuguese language contact with a great number of languages of different origins (this contact is still vigorous in African countries); (c) an irregular learning process is at the base of such diversity of patterns; (d) the main structural factor that implies variation – *linear and relative position of the constituents in the NP* – corresponds to a continuum; if considered only synchronically, this continuum constitutes a vernacular universal (Chambers, 2004: 128) that takes place only by the action of social factors in the speech of communities with a history of intense multilingual contact.

To discuss these hypotheses and expose the analyses on the nominal agreement in urban varieties, this paper is organised in five more sections. In section 2, we provide general information on the methodology used. In the third and fourth sections, we examine the nominal agreement as a categorical rule (data from EP), and as a variable rule (data from BP and from São Tomé Portuguese (STP)). We discuss the structural and extra-linguistic variables that showed up as the more salient ones for the implementation of the variable patterns; the results from previous studies on BP and on other African varieties have contributed in detecting those patterns. The fifth section presents some conclusions and, in the final section, we formulate new questions that can guide future studies on this subject.

2. Methodology: general issues

Data analyses were guided by the theoretical and methodological principles of the Theory of Variation and Change (Weinreich, Labov & Herzog 1968), as well as by the developments of Variationist Sociolinguistics (Labov 1972, 1994, 2001), regarding, among other assumptions, the notion of variable rule, which depends on the actuation of linguistic and social constraints. We considered also the proposal of Labov (2003) concerning the distribution of the linguistic rules into three classes, according to the frequency with which they operate: (a) categorical (100%); (b) semicategorical (95-99%); and (c) variable (5% to 95%).

The analyses presented in this paper are based on the spoken data described below:

(a) 18 informants for each sample, distributed by sex, three age groups (18-35 years, 36-55 years, 56-75 years) and three levels of education (fundamental: 5-8 years; medium: 9-11 years; high: over 11 years),

representative of EP (LIS-OEI, LIS-CAC; MAD-FNC samples, relating respectively to Oeiras and Cacém, in the Metropolitan Area of Lisbon, and Funchal, in Madeira Island), BP (RJ-COP and RJ-NIG samples, relating respectively to Copacabana and Nova Iguaçu, in the Metropolitan Region of Rio de Janeiro) and STP (sample ST(1), relating to São Tomé).

(b) 22 informants of São Tomé distributed by sex and age group and two levels of education: fundamental (5-8 years) and medium (9-11 years) – sample ST(2);

(c) 9 students of São Tomé from different levels of education – sample ST(3).

The samples which we analyse most extensively are those in (a) above, with 18 informants each. Some results from these samples are found in Table I. With regard to EP, the unit of analysis is the NP as a whole, in a total of 6.952 tokens; considering the very high productivity of the canonical agreement patterns, we present only the simple percentages. Regarding BP and STP, the unit of analysis corresponds to each inflectional constituent of the NP, in total 7493 and 5931 tokens, respectively.

Table I. Number of observed data by variety/sample

<u>EP</u>			<u>BP</u>		<u>STP</u>		
Number of NPs			Number of inflectional NP constituents		Number of inflectional NP constituents		
<u>LIS-OEI</u>	<u>LIS-CAC</u>	<u>MAD-FNC</u>	<u>RJ-COP</u>	<u>RJ-NIG</u>	<u>ST(1)</u>	ST(2)	ST(3)
2.312	2.449	2.191	3.716	3.777	2.923	2.375	633
4.761							
6.952			7.493		5.931		

In the quantitative analyses related to BP and STP, we used the Goldvarb-X software package, and systematically controlled eight structural variables: (1) phonic saliency, (2) toniticity and (3) number of syllables of the word in singular, (4) the word class, (5) linear and relative position of the constituent within the NP, (6) previous plural marks, (7) subsequent phonological context, (8) animacy of the NP nucleus. We also controlled three social variables: (1) age group, (2) sex, and (3) education level. In the analysis of an expanded number of adult informants from São Tomé having fundamental and medium education levels, we also considered the variable “frequency of use of a Creole”. In the analysis of the speech of the 9 students from São Tomé – ST(3) sample – we controlled the performance of each individual.

3. Nominal agreement as a categorical rule: EP

In the introduction, we mentioned the hypothesis of Naro & Scherre (2007) regarding the origin of the variation observed in BP, founded on the idea that Brazil presented the ideal conditions to manifest latent linguistic tendencies in the Portuguese language and, roughly speaking, in the Indo-European languages. This fact would imply the possibility of finding, throughout history and in popular speech, data that provides evidence for those tendencies. Some examples of the absence of plural marking in NPs were found through the “dig” (in the words of these researchers) effectuated in medieval works, dialectal monographs and in the inquiry note-books of the *Linguistic-Ethnographic Atlas of Portugal and Galiza-ALEPG*. Although numerically very modest and not always appropriately contextualised, those examples constituted evidence that the *Corpus Concordância*³ would allow testing in urban varieties, especially because, between the aforementioned Portuguese works, was the Mira's (1954), based on data from the “speech of simple and illiterate people in poor neighbourhoods of Lisbon”, as indicated in Naro & Scherre (2007: 191).

For the Metropolitan Area of Lisbon, represented by Oeiras and Cacém, the hypothesis that guided this study was that, regardless of the education level of the speaker and the evidence mentioned above, the agreement rule would categorically act. In the case of Funchal, on the contrary, we expected to find examples of the absence of plural marking, due to the following aspects:

(a) the variety spoken in Madeira Island, spatially discontinued regarding other varieties of Portuguese, presents, in a phonetic-phonological plan and in the words of Cruz (2003), “a set of phenomena of its own that are not found in continental dialects”. Among others:

The diphthongisation of the high stressed vowels /i/ in [ej] or [ij] and /u/ in [ew]: *navio* [nəv'ɐju] or [nəv'ijju] (...), *lua* [l'ɛwɐ] (...) the palatalization of the /l/, when preceded by [i] or [j], i.e. its realisation as [ʎ], as in (...) *vila* [v'ɛjʎɐ] or [v'ijʎɐ] and the vocalisation or semi-vocalisation of the final -s when followed by voiced consonant or voiceless fricative, as in [aj v'akeʃ]...(Cruz, 2003: 186-187).

(b) there are some syntactic phenomena common to BP, as summarised by Bazenga (2012: 330):

Some non-standard syntactic properties have been revealed by recent studies on dialectology field, as the work done by Carrilho (2010), based

³ Comparative study of agreement patterns in African, Brazilian and European varieties Project. See www.letas.ufri.br/concordancia.

on CORDIAL-SIN (*Syntax-oriented Corpus of Portuguese Dialects*⁴). In this study, non-standard constructions are in areal distribution, and they mainly occur in a restricted part of the EP area. It is the case of *existential impersonal* constructions with *ter* (*to have*), only attested in Madeira dialect, *aspectual constructions involving the gerund form* of a main verb (found also in the Azores and Central Southern dialects of EP), and constructions as *Pre-nominal possessives without article* (in Madeira but also attested in the Azores but with less frequency). It may be noticed too that these constructions are all attested in standard BP. Martins (2003, 2009) referred to the (...) sharing of some properties with BP variety and pointed to the possibility of a lack of the properties of the *Null-Subject parameter*, making this variety closer to BP;

(c) at different times in its history, the process of settlement of Madeira Island counted with the arrival of migrants from North and Central/South areas of Portugal, from other parts of Europe and also from Africa, which presupposes a situation of interlinguistic/interdialectal contact, which still persists, due to the natural touristic potential of the region.

The data analysis, however, demonstrated that the agreement rule operates categorically, in the terms of Labov (2003), in the speech of the three focused areas, as explained in Table II.

Table II. Distribution of NPs with and without number mark within urban varieties of EP

Sample	NPs			
	With number mark(s)		Without number mark(s)	
	Tokens	%	Tokens	%
Oeiras	2310/2312	99.9	2/2312	0.08
Cacém	2448/2449	99.9	1/2449	0.04
Funchal	2186/2191	99.7	5/2191	0.22

Table II shows that: (a) in the EP area, whether continental or insular, the canonical rule of agreement is categorical; only eight NPs from a total of 6.952 showed no number mark in all constituents, which equates to index 0.11%; (b) as the three samples include the same number of speakers in the three education levels (fundamental, medium and high), it is possible to conclude that the rule is categorical both in popular and in cultivated speech. As there are only eight cases of no number marking in all constituents, it is worth transcribing them and commenting on them, especially because the five data from Funchal could lead one to believe that, in this variety, the hypothesis formulated above operates, which does not seem to be true.

⁴ See <http://www.clul.ul.pt/en/research-teams/212-cordial-sin-syntax-oriented-corpus-of-portuguese-dialects>.

Table III. Tokens of NPs without number mark(s) within the three EP samples

Sample	Informant	Tokens
Oeiras	OEI-B2h	(1) com vários [vários tipo] de pessoas (with several [several types] of people)
	OEI-B2m	(2) e já há muita gente a gostar de [novela... portuguesas] eu pessoalmente não consigo. (and there are already a lot of people liking [Portuguese novel...] I personally cannot)
Cacém	CAC-C1h	(3) nada...ali nada tem uma pracetazinha...há uma ou [duas praceta] onde os miúdos se entretêm onde há menos carros pra jogar um bocadinho à bola ou andarem de bicicleta de resto não tem lá nada...aquela zona é mesmo mesmo mesmo sem nada. (nothing...there nothing has a ...very small square...there is one or [two small squares] where kids are entertained and where there are fewer cars to play a little ball or ride a bike the rest has nothing there...that area has really really really nothing)
Funchal	FNC-A1m	(4) queria que não faltasse nada a[o meu...filhes] [filhos] (I would like to lack nothing for [my...children] [children])
	FNC-B1m	(5) a maioria d[as pessoa] das mulheres era tra/ era em casa era a bordar (most of [the people] of women was/were at home was embroidering)
		(6) [as brincadeira]...era quando chegava os meus primos vinha à nossa casa brincar à à pilhagem. ([the jokes]...was when my cousins came to our house to play looting)
	FNC-A2m	(7) [outros conhecimento] outra maneira de ver as pessoas ([other knowledge] another way of seeing people)
FNC-A3h	(8) foi d[as primeiras vez] em que me apercebi o quão difícil é liderar um grupo (it was of [the first times] that I realised how difficult it is to lead a group)	

Save for (4) and (8) – presenting three elements – all the NPs correspond to two constituents, the type of structure that is prevalent in approximately 80% of the different sample data. Despite seven of them fall into the prototypical case of variable agreement (no mark in the nucleus), some of them seem to result from some failure in the discursive procedure, from hesitation, as in (2), (3) and (4). The data in (5) and (7), wherein the informant abandons the initially selected structure – *das pessoas* (‘of the people’), *outros conhecimento* (‘other knowledge’) – and rearranges the text with another structure – *das mulheres* (‘of women’), *outra maneira* (‘another way’) – thereby allowing us to deduce that the enrolled cases may comprise lapses as indicated by Labov (2003) when explaining the nature of a categorical rule as opposed to semi-categorical and variable rules.

We believe that the framework outlined in this study applies to EP. However, because strictly variationist sociolinguistic studies practically started in Portugal with the work undertaken by the *Projeto Concordância* team, it is essential to develop further research on this variety, focusing on less educated urban speakers and on inland areas, e.g., in the small villages that still retain traditional customs that singularise them.

4. Nominal agreement as a variable rule: BP and STP

Several studies on Portuguese-based Creoles and on Afro-descendant communities have contributed, for a long period, to the understanding of the agreement under BP, especially in order to justify variant patterns of agreement found in it. Nevertheless, the recovery of the ancient idea that essentially synchronic comparative studies between different national varieties of Portuguese could contribute to a better characterisation of the variable rule that operates on non-European varieties (whether spoken as L1 or used as L2) is recent.

As we try to demonstrate in this section, the variable rule seems to follow, today, the same structural and social constraints in the speech of all the communities outside Portugal, as suggested by the analyses of two Brazilian samples representative of the Metropolitan Region of Rio de Janeiro (RJ—COP and RJ-NIG) and of a sample of São Tomé – ST(1); the three samples are organised according to the common criteria of EP samples (cf. the general indexes shown in Table IV).

Table IV. Distribution of NP constituents with and without number mark(s) within urban varieties of BP and STP

NP inflectional constituents				
Sample	With number mark(s)		Without number mark(s)	
	Tokens	%	Tokens	%
Copacabana	3432/3716	92.4	284/3716	7.6
Nova Iguaçu	3439/3777	91.1	338/3777	8.9
São Tomé	2524/2612	93.4	173/2612	6.6

The average of number marking is high, similar in all communities; this result is surprising to some extent if we take into account the history of the formation of BP and STP, as well as the polarised socio-economic-cultural conditions that occur in the two countries. Such a result, however, masks significant differences, not between Brazilian and São Tomé varieties taken as a whole, but among groups of speakers depending on their level of education, an aspect also common to both varieties, as will be seen later.

Variationist analyses clearly demonstrate the constraints that govern the presence/absence of the number marking within the NP (cf. Table V) and therefore the presence of agreement variant patterns.

Table V. Operating variables for implementing the number marking by sample/variety

BP		STP
Copacabana	Nova Iguaçu	São Tomé
<u>EDUCATION LEVEL</u>	<u>EDUCATION LEVEL</u>	<u>EDUCATION LEVEL</u>
	<i>LINEAR RELATIVE POSITION OF THE NP CONSTITUENTS</i>	
<i>LINEAR AND RELATIVE POSITION OF THE CONSTITUENT WITHIN THE NP</i>	<u>EDUCATION LEVEL</u>	<i>LINEAR AND RELATIVE POSITION OF THE CONSTITUENT WITHIN THE NP</i>
<u>Age group</u>	<u>Age group</u>	<u>Nucleus animacy</u>
<u>Gender</u>	<i>Phonic salience</i>	<u>Gender</u>
<i>Phonic salience</i>	<u>Nucleus animacy</u>	Subsequent phonological context
Significance: .000	Significance: .001	Significance: .023
Input: .98	Input: .96	Input: .98

For what is exposed in Table V and in line with previous studies (cf. section 4.1 below), the variable rule operates depending on some major factors:

- (a) *the linear and relative position of the constituent within the NP and the education level of the speakers* are highly relevant variables, even if they alternate in order of importance depending on the samples; and
- (b) other social and structural variables play a supplemental role; among them stand, respectively, *age/gender* and *morphophonological process of plural formation* (here called *phonic salience*)/*nucleus animacy*.

4.1 Highly relevant variables

4.1.1. Linear and relative position of the constituent within the NP

The first variationist studies on nominal agreement in BP (Braga & Scherre, 1976; Braga, 1977; Scherre, 1978) showed the importance of the variable *constituent linear position*. It was possible to check that the first position is opposed to the others, and that it favours the implementation of the number mark. Guy (1981), in his thesis on spoken Portuguese in Rio de Janeiro, arrives at the same conclusion, setting, however, a correlation between position and grammatical class; he assumes that, in canonical terms, determiners, nouns and adjectives should occupy the first, second and third position within the NP, in this order. Such a correlation was challenged by Scherre (1988), who, while confirming the tendency towards a greater presence of marks in the first position, claimed that it could reach any constituent that occurred in the first position, as can be seen from the examples below, taken from samples of BP and STP:

- (5) a. todos os seus aluno (all his students)
- b. aquelas coisa toda (that whole things)
- (6) a. ter condições financeira boa (to have good financial conditions)
- b. ter boas ideia (to have good ideas)

The constituents *todo(s)* ('all/whole') and the adjective *boa(s)* ('good') hold, respectively, the position at the left and at right of the nucleus, as in (5)a and (6)a. The change in the position ultimately determines how the same elements behave, with regard to the agreement, depending on the position they occupy; therefore, they are less (cf. (5b)) or more (cf.(6b)) susceptible to receive the plurality mark.

Thus, posterior analyses began to consider a new variable, the *position of the constituent relative to the nucleus*, which, combined with *linear position*, would originate *linear and relative position of the constituent within the NP*, a variable that has not shown significance only in relation to the urban

Brazilian and São Tomé varieties, and also to other varieties – cf. Figueiredo (2010) on the community of Almojarife in São Tomé and Jon-And (2011) about Mozambique and Cape Verde.

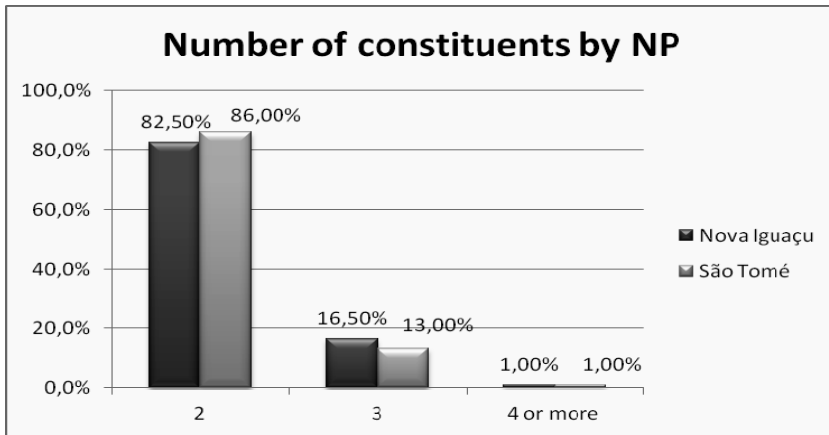
Table VI shows the indexes⁵ obtained in the analyses of the three samples, where we may observe the locus par excellence of the number marking: the pre-nuclear position, the first position, in a prototypical NP – Determiner(s) + Noun + (Modifier) – or the nuclear position, in non-prototypical NPs, where a noun or an adjective occurs at the beginning of the NP.

Table VI also allows us observing other structural features that recur in all samples. It is inferred by the number of data that the prototypical NPs are those that present the nucleus in the second position; therefore, the first pre-nuclear position is the more productive. It is a fact that the two constituent NPs predominate in the samples, the three or more constituent ones of the type Determiner + Noun + Modifier being uncommon.

The survey of the NPs characterised by the cancellation of the number mark in the NIG sample and in the ST(2) sample (data obtained from 22 informants from fundamental and middle education level) confirms this assertion, as Figure 1 shows.

⁵ In Table VI, the indexes are shown as follows: number of tokens; percentage; relative weight (RW).

Figure 1. Number of constituents by NP in NIG and ST(2) samples



Based on this ST sample, where 286 NPs were found with no plural mark in one or more of one constituent, a survey was conducted of the structural patterns observed therein; we consider that these patterns are the most commonly found in all samples.

Table VII describes the structures of the 246 NPs with two constituents, prototypical and non-prototypical, which are exemplified in (7) to (14) and in (15) to (17), respectively:

- | | |
|---|---|
| (7) ...com a idade de [oito ano]
(...with the age of [eight years old]) | (13) ...com [mais pessoa]
(...with [more people]) |
| (8) ... [umas coisinha] assim
(... [some little thing] like this) | (14) ... [diversos tipo] de pessoas
(... [various type] of people) |
| (9) N [aqueles tempo] era diferente
(On [those time] it was different) | (15) Não vivi [situações difícil]
(I didn't live [difficult situations]) |
| (10) Eles tem [seus salário]
(They have [their salary]) | (16) [Curso noturnos] , cursos de
alfabetização
([Night courses] , literacycourses) |
| (11) Fiquei [três ano] na vida militar
(I stayed [three years] in the
military) | (17) ... [pequenas obra] , [grandes obra]
(... [small works] , [great works]) |
| (12) Aqui tem [muitas cabra]
(Here there is [many goats]) | |

Table VII. Prototypical and non-prototypical two constituents NPs

TWO CONSTITUENTS NPs					
PROTOTYPICAL			NON-PROTOTYPICAL		
Pre-nucleus (Pos. 1)	Cancel	Nucleus (Pos. 2)	Cancel	Post-nucleus (Pos. 2)	Cancel
Def. article	3/51		48/51	Adj	11/12 91%
Indef. article	0/6		6/6		
Demonstrative	0/42		42/42		
Possessive	0/14	Noun	14	Nucleus (Pos. 2)	Cancel.
Numeral			79/79	Noun	3/3 100%
Quantifier/Indefinite	1/36		35/36		
<i>Mais</i> (more)	0/2		2/2		
<i>Diversos</i> (<i>several</i>)	0/1		1/1		
	4 2%		227/231 98%		

In the NPs we consider prototypical, in the first position, predominate, among the inflectional constituents, the articles (48 tokens) and the demonstratives (42 tokens), with a very similar degree of productivity; among the non inflectional, the numerals (79 tokens). In the last case, it is worthy of note that the nucleus becomes the element that does not present the mark (11/12), which is reversed only when the noun occupies the first position (1/12 mark cancellation).

Table VIII presents the structures inferred from the 31 NPs with three constituents with nucleus in the second and third positions exemplified, respectively, by (18) to (21) and by (22) to (30):

- | | |
|---|--|
| (18) ...aproveitar [os tempos livre]
(...enjoy [the free times]) | (24) São sempre [os mesmo programa]
(There are always [the same programme]) |
| (19) [As coisas toda] mudaram
([The things all] changed) | (25) [as outras igreja] tem estado...
([the others church] has been...) |
| (20) ...[essas coisa toda]...
(...[these whole thing]...) | (26) [umas pequena venda]...
([some small sale]...) |
| (21) [outras coisas pior]
([others worse thing]) | (27) em [todas as casa] temos vasilha
(in [all house] we have canister) |
| | (28) n[as duas coisa]
(in[the two thing]) |
| (22) [as minha ferramenta]
([my tool]) | (29) Pelo menos [umas três fábrica]
(At least [three factory]) |
| (23) [nos primeiro momento]...
([the first moment]...) | (30) ...com [esses dois trabalho]
(...with [these two work]) |

Table VIII. Three constituents NPs

NUCLEUS IN POSITION 2						NUCLEUS IN POSITION 3					
Pre-nucleus		Cancel	Nucleus	Cancel	Post nucleus	Pre-nucleus		Nucleus			
Pos. 1		Pos. 2		Pos. 3		Pos1	Cancel.	Pos. 2	Cancel.	Pos3	Cancel.
Def. article	1c	Noun	2a	Adj.	2a	Def. art.		Poss	1a	1a	3b
					2b						
			1a		Ordinal			2a	2a		
Demonst.			2a	Todo	2a			Numeral	2a		
			1a		1a			Mesmo	1a	1a	
				Adj.				(same)			
Quant/Ind.			2a		2a			Indef	3a		
(1)			(7)		11/11			Adj	1a	Subst	1a
								Ind. Art.		3a	
								Dem	Num	1a	
					Def. Art.		1a				
					Todo	1a	Dem.	1a			
					(all)	1b					
					(2)		(5)	19/20			

In the 11 tokens with the nucleus in the second position, only one has no marking in the pre-nuclear first position (in a definite article), seven in the nuclear second position and also all of them (11) in the post-nuclear position. In the case of nuclei in third position, the nucleus gathers the largest number of mark cancellations (19/20), followed by the pre-nuclear element in the second position.

There are, also, 6 tokens of three constituents NPs where the nucleus is in the first position, the cancellation falling on the inflectional post-nuclear constituent:

(31) [**peessoas mais velha**] faziam calulu ([**older people**] made calulu)

Table IX. Three constituents NPs within the STP sample

Nucleus		Post nucleus		
(Pos. 1)	Cancel	(Pos. 2)	(Pos. 3)	Cancel
Noun		Intensifier	Adjective	6

We describe below the three single NPs from four constituents recorded in the sample, two with the nucleus in second position, one in the fourth position.

	Pre-nucleus	Nucleus	Post-nucleus	
(32)	[Aquelas (Those (dem.)	lula skid (noun)	mesmo really (intens.)	grande] big) (adj).
(33)	[Umas (Some (ind. art.)	parte part (noun)	mais (intens.)	grossa] thicker) (adj.)
		Pre-nucleus	Nucleus	
(34)	[Os (The) (def. art.)	seus your (poss.)	quatro four (num.)	ano] year) (noun.)

This exhaustive description is aimed at demonstrating that, as explained in Table V above, referring to the *Linear and relative position of the constituent within the NP*, the behaviour of the constituents in two and three

constituents NPs, the most productive in all samples, suggests that when the agreement rule is variable, the number mark preferentially concentrates in the left *locus* of the NP. This fact allows us:

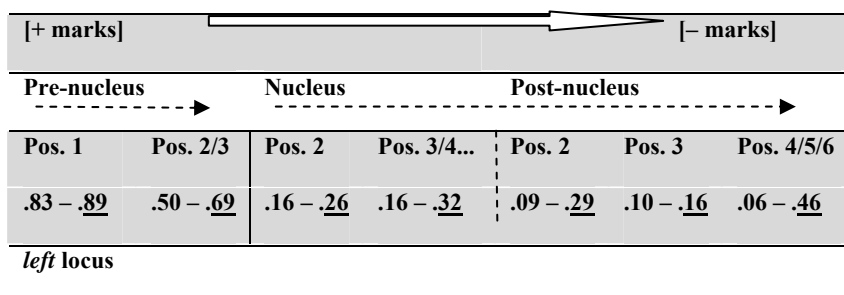
(i) on the one side, to indicate the positional saliency (either linear or relative) as one of the key elements for the presence/absence of marks in NP constituents; and

(ii) on the other, to question the role of the so-called principle of the *formal parallelism* (“marks leading to marks, zeros leading to zeros”) within the NP, advocated by some scholars (Scherre, 1988, for example).

Based on these results, it is proposed, as a hypothesis, the *continuum* outlined in Figure 2, which is believed valid for urban (and even non-urban) Portuguese varieties that the rule is not categorical. We consider what is here called prototypical NPs, that is, those comprised of pre-nuclear constituent(s), nuclear constituent and (post-nuclear constituent). At each position, we indicate the minimum and maximum relative weights (RW) of the plural mark presence obtained in the analyses presented in Table V.

Thus, the first linear position in the pre-nucleus tends to be more marked, which suggests this is the locus mark par excellence. From the nucleus in the second position, its presence will gradually decrease, whether considering the minors or major indexes obtained (highlighted by underlining). In the proposed scheme, a continuous vertical line that separates the pre-nucleus from the nucleus would represent the furthest space limit for marking (the left *locus*), while the dotted vertical line between the nucleus and the post-nucleus would indicate that these areas would not suffer, as the tendency not to mark, continuity solution. Note, as pointed out before, the displacement to the first position of a nucleus or a constituent that occupies post-nucleus position will compete for it to become more susceptible to receive number mark.

Figure 2. *Continuum* of plurality marking on prototypical NPs in BP and STP, based on WRs obtained within the different analyses carried out



The behaviour of the constituents for the presence of plural mark, as proposed above, appears to be an inherent tendency, a kind of universalising characteristic of Portuguese varieties, because, in greater or lesser extent, it is found in all the varieties where there is determined external conditions for

variation. As stated in another study (Brandão & Vieira, 2012a), structural constraints that operate in such varieties refer to components of a cognitive procedural nature that reveal themselves at the syntactic level (primarily) and also in sound and semantic levels as we shall see, but it would be external factors, whose performance must be interpreted in the light of the socio-history of each variety, that are the key elements for implementing such restrictions.

Studies on other varieties (Lopes, 2001; Andrade, 2003; Baxter, 2004; Figueiredo, 2008; Jon-And, 2010, 2011), even within the scope of BP (Scherre, 1988), indicate the pre-nuclear position adjacent to the nucleus (position 2 or 3) as the most subject to marks, which is why the variable, according to Jon-And (2010, 2011), has served as the basis for testing the hypothesis on the possible influence of African languages (transfer or substrate influence).

Also, according to Jon-And, in the cases of Mozambique and STP (the group of Tonga and the locality of Almoxarife), the hypothesis would be confirmed, given that these varieties have as substrate Bantu languages in which plural marking is made through suffix. However, the same explanation would not apply to Cape Verde, which has no Bantu languages as substrate. In the Creole substrate languages of Cape Verde, which belong to the Atlantic and Mande families, there is a varied way of indicating the plurality as seen in Table X, with being the marking at the right through suffix the most usual. However, in the Cape Verdean Creole, predominates marking in the determiner in the first position, although other constituents may also receive mark.

Table X. Plural marking in Cape Verdean Creole substrates

LANGUAGE	FAMILY	PLURAL MARKING
Wolof	Atlantic	The noun, typically, does not have plural mark. The noun is followed by a particle that defines the nominal class, with the function of definite article, and varies in form depending on whether the noun expresses singular or plural (Campbell, 1991: 1776).
Mandinka	Mande	The plural is marked only in specific contexts (Derive, 1990: 179). The plural is probably marked through suffix, as well as in languages related with Bambara and Dyula.
Bambara	Mande	The plural is marked in the noun through suffix — μ . The demonstratives have the same plural marks, but only when positioned after the noun (Campbell, 1991: 190). The plural is marked only in specific contexts (Derive, 1990: 179).

Dyula	Mande	The plural is marked through several different suffixes. The plural is marked only in specific contexts (Derive, 1990: 180-86).
Fida	Atlantic	For regular nouns, the suffix <i>-o</i> of the singular is replaced by the suffix <i>-be</i> in plural. For irregular nouns the initial consonant also changes from singular to plural (Arrendsdorff, 1913: 23-25).
Balanta	Atlantic	For animated nouns, the plural is marked through the prefix <i>be-</i> . For unanimated nouns, the plural is indicated by the free morpheme <i>mòne</i> , positioned after the noun (Quintino, 1951: 3-4).

Given that (1) the plural mark in the determiner is not a linguistic universal, as can be seen, for example, in English, where the mark appears only in the nucleus, regardless of the position it occupies (The houses/The (big) houses, Our (old) friends); (2) in EP, we observe the marking of all inflectable constituents of the NP; and (3) languages that served as a substrate for the Portuguese varieties and Cape Verdean Creole have diversified plurality marking rules in NP, the trend outlined in Figure 2 must be considered, as seems to be, in the syntactic level, one of the general patterns of non-European varieties of Portuguese.

4.1.2 Education Level

The variable described in the previous section enabled a clear perception of the configuration of the variable patterns of agreement, these patterns that stated in EP itself, in the cases classified in item 3 as discourse processing failures. Such cases, however, do not appear to be subjected to the performance of the variable *education level*, as occurs in non-European varieties, which plays a role as relevant as the constituent position, as seen in Table IV: in the variationist analyses regarding Copacabana and São Tomé samples, the variable *education level* was selected in first place and in Nova Iguacu in second.

Apparently, this variable would best represent the action of the external factors for the implementation or not of the number marking by its complexity and due to the fact that, directly or indirectly, it allows to evaluate the productivity of the phenomenon. Despite admitting that, in the extralinguistic level, the comparisons must be relativized and that this variable, in a given manner, reflects the action of others related to it, as, for example, more or less contact with the media, higher or lower social-

-economic level, more or less exposition to cultural goods, we have been noticing that, in BP, and in African varieties, it has served as a parameter to understand how different agreement patterns become effective.

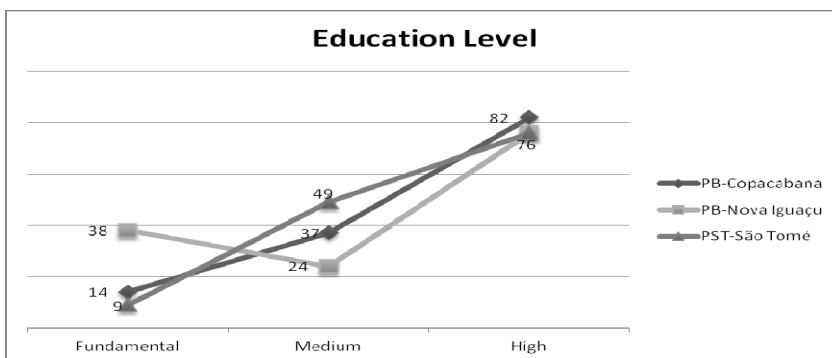
As can be seen from Table XI, complemented by Figure 3, regardless of the high percentage indexes of plural mark on all education levels (from 76.5% to 98.8%), the R.Ws show that there is between them convergent and divergent characteristics.

Table XI. Actuation of education level for implementing the number mark within the NP, in BP (Copacabana and Nova Iguaçu) and STP

Education Level	BP						STP		
	Copacabana			Nova Iguaçu			Oco	%	R.W.
	Oco	%	R.W.	Oco	%	P.R.			
5 to 8 years (Fundamental)	781/970	80.5	.14	1074/1175	91.4	.38	406/531	76.5	.09
9 to 11 years (Medium)	1127/1203	93.7	.37	820/1013	80.9	.24	908/942	96.4	.49
12 to 15 years (High)	1524/1543	98.8	.82	1545/1589	97.2	.76	1125/1139	98.8	.76

The varieties converge pursuant to the performance of more educated persons (R.W. between .76 and .82) compared with the least ones (R.W. between .09 and .38), but differ in other aspects (according to Figure 3).

Figure 3. Actuation of education level for implementing the number mark within the NP in BP (Copacabana and Nova Iguauçu) and also in STP in R.Ws



The actuation of the variable in Copacabana and São Tomé is compatible with the hypothesis that agreement rule application probability increases gradually from the lowest to the highest level of education respectively: .14 \mapsto .37 \mapsto .82 and .09 \mapsto .49 \mapsto .76. In Nova Iguauçu, however, individuals at fundamental level (.38) outweigh those in the medium education level (.24) regarding the use of the rule.

In both varieties, there is a polarised socio-linguistic situation, even though when we compare the two BP representative varieties, the one that depicts major polarisation is Copacabana, which will be subject to some considerations in section 6 of this work.

It allows us to evaluate the importance of this variable in other non-European urban and non-urban data shown in Table XII, in which we used and reorganised some information presented in table 21 from Jon-And (2011: 140), and added data related to the analyses we are developing in this study. It also shows results from researches of Almeida (1997) and Figueiredo (2010). In Table XII, the bold global marked number indexes are highlighted, sometimes followed by indexes related to certain education levels, according to the parameters defined in each study.

Table XII. Plural marking frequency within different studies on NP number agreement

Variety		Location	Education Level of the speakers	Plural Marking
BRAZIL	URBAN	Nova Iguaçu (Brandão, 2011)	5-15 years (all data)	<u>91.1%</u>
			5-8 years	91.4%
			9-11 years	80.9%
			12-15 years	97.2%
		Copacabana (Brandão, unpublished)	5-15 years (all data)	<u>97%</u>
			5-8 years	80.5%
			9-11 years	93.7%
			12-15 years	98.8%
		Rio de Janeiro (Scherre, 1988)	1-11 years (all data)	<u>72%</u>
			1-4 years	62%
5-8 years	72%			
Rio de Janeiro (Campos et al., 1993)	College	<u>98%</u>		
NON-URBAN	Rio de Janeiro (Almeida, 1997)	0-4 years (all data)	<u>44%</u>	
		0 years	42%	
		1-4 years	46%	

SÃO TOMÉ	URBAN	(Brandão, 2011) Portuguese L1	5-15 years (all data)	<u>93.4%</u>
			5-8 years	76.5%
			9-11 years	96.4%
			12-15 years	98.8%
	NON-URBAN	Almoxarife (Figueiredo, 2008)	0-8 years (all data)	<u>65 %</u>
Almoxarife (Figueiredo, 2010)			0-8 years (all data)	<u>51%</u>
			Illiterate	31%
	Primary level frequency	60%		
	Completion of the primary level	50%		
	Post-primary level frequency	71%		
	The Tongas (Baxter 2004, 2009)	0-5 years (all data)	<u>55%</u>	
MOZAMBIQUE		(Jon-And, 2011) L2	3-7 years (all data)	<u>88%</u>
CAPE VERDE		(Jon-And, 2011) L2	0-8 years (all data)	<u>82%</u>
			0 years	69%
			3-4 years	78%

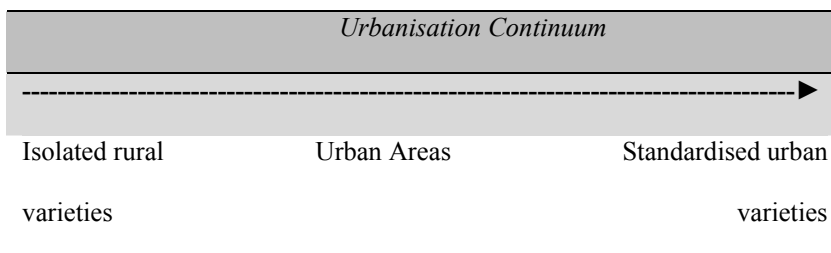
Observing the overall percentage (underlined), it appears that both in BP (72% to 98%) and in STP (93.4%), urban varieties tend to present a higher plural marking in NP than non urban ones (BP: 44%; STP: 51% to 65%). Although Jon-And's samples refer to Portuguese speakers as L2, individuals who contributed to its composition were born and live in Maputo, in Mozambique's case, and in Mindelo, one of the two major urban centres of Cape Verde, which justifies the high rates of agreement (respectively 88% and 82%).

Regarding the partial rates of Table XII, it is possible to note that roughly less education or illiteracy implies less marks in the NP. We emphasize also that, as observed in Almojarife (Figueiredo, 2010), individuals with low education in the schooling phase (60%) seem to be more susceptible to the agreement canonical patterns than those who no longer attend school (50%). It is possible to deduce that the implementation of the number mark in this type of community depends on constant exposure to EP, which is the standard, and is present in written texts and in textbooks.⁶

Moreover, the panel study (Labov, 1994) carried out by Naro & Scherre (2003a: 47-48) comparing the performance of 16 individuals in the 1980s and 2000 (Rio de Janeiro) shows that there was, during this period, at the individual level, in general, a major increase in the use of the agreement and that individuals who increased their education level tended to bypass those that remained at the same level.

Bortoni-Ricardo (2004) proposed that we consider the variable phenomena in BP according to three parameters (urbanisation, literacy-orality and stylistic monitoring). The urbanisation continuum (according to Figure 4) also applies fully to STP.

Figure 4. Urbanisation Continuum regarding the BP variation analysis, according to Bortoni-Ricardo (2004)

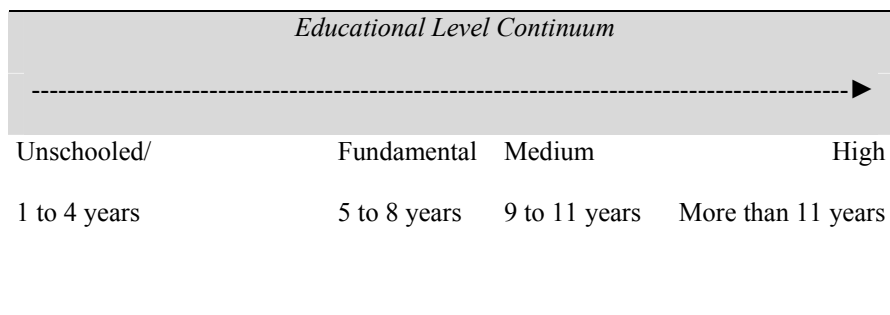


It should be highlighted, however, that such continuum overlaps the education level continuum proposed by Brandão (2011) below (see Figure 5)

⁶ It should be kept in mind, however, that many of the teachers who work in São Tomé are natives of the region and do not necessarily dominate EP standards.

and above mentioned, as in non-urban communities in Brazil (and also in São Tomé) there are few individuals with more than eight years of schooling.

Figure 5. Education Level *Continuum* according to Brandão (2013)



To better explain the aforesaid superposition, Figure 6 was organized. In it we consider together the two mentioned *continua* applied to the different non-European varieties, using the percentage indexes exposed in Table XII.

Figure 6. Intersection of the *urbanisation continuum* (Bortoni-Ricardo, 2004) and the *education level continuum* (Brandão, 2013) applied to the nominal plural agreement (AR = Rurban Areas)

		Plural Nominal Agreement					+marks	
		- marks		RA	Urban varieties			
Urbanisation	Rural varieties	Rural varieties			Urban varieties			
Continua	Education level	Illiterates/ 1-4 years	5-8 Years		Illiterates/ 1-4 years	5-8 years	9-11 years	More than 11 years
		RJ-42%/46%			RJ-62%	NI-91.4%	NI-80.9%	NI-97.2%
		AL 31%	ST-AL 71%			CP-80.5%	CP-93.7%	CP-98.8%
		TO 55%	ST-60%/50%			RJ-72%		RJ-98%
						ST-76.5%	ST-96.4%	ST-98.8%
					CV 69%/78%			
						MO-69%/78%		

RJ = city of Rio de Janeiro, as a whole; CP = Copacabana, district of the city of Rio de Janeiro; NI = Nova Iguaçu, Municipality of the Metropolitan Region of Rio de Janeiro; ST = São Tomé, as a whole; AL = Almoarifé in São Tomé; TO = Tonga community in São Tomé; CV = Cape Verde.

In Figure 6, it can be seen that the lowest rates of plural marking refer to illiterate or low level of education, both in rural and in urban varieties.

In rural areas, there are factors that predispose to the preservation of linguistic patterns deemed as socially less prestigious by society: in these areas, the most isolated communities are concentrated (as Almojarife and the Tonga's group) and the vast majority of the population consists of individuals with low education level, as the schools available are limited, in general, to elementary education (four or eight years).

It is believed that the high level of canonical agreement in the speech of individuals with a low education level in urban areas is due to the fact that they are in constant interaction with speakers of high education level and the non canonical agreement standards have, in these communities, strong entangling nature. In complex communities, individuals tend to adopt patterns of prestige due to increased contact with the media, the ease of access to education and the requirements for the achievement of better jobs, factors that promote social mobility.

In Brazil, the vast majority of the population is concentrated in urban areas, contrary to what occurred in the 1950s. As can be seen from Table XIII, the process was abrupt, in some regions, such as the Midwest and South, and gradual in the Southeast, North and Northeast, although in the latter two this process is in a less advanced stage, which gives an idea of the transformation that took place in Brazil, which, as characterised by Cunha (1972) "*was a vast rural country during four centuries*".

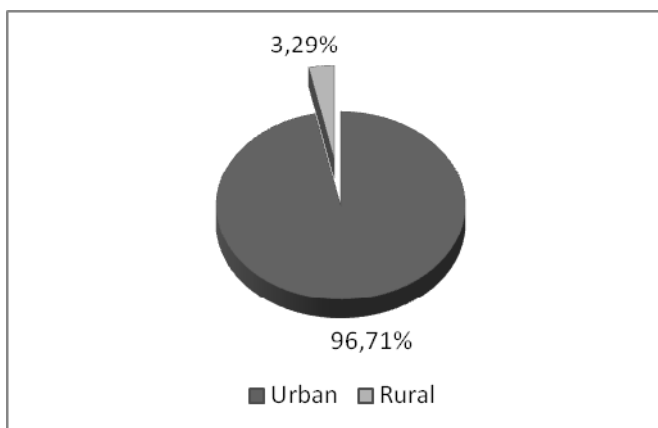
Table XIII. Brazil: Urbanisation percentage rates per region

	1940	1950	1960	1970	1980	1990	2000
North	27.75	31.49	37.80	45.13	51.65	57.84	69.9
Northeast	23.42%	26.40	34.24	41.81	50.46	60.64	69.1
Southeast	39.42	47.55	57.36	72.68	82.81	88.01	90.5
South	27.73	29.50	37.58	44.27	62.41	74.12	80.9
Midwest	21.52	24.38	35.02	35.02	67.79	79.16	86.7
Brazil	31.24	36.16	45.08	55.92	67.59	75.47	81.2

Source: IBGE

The State of Rio de Janeiro is located in the Southeast, the most urbanised region of the country, where rates attain almost categoricity, as depicted in Figure 7.

Figure 7. State of Rio de Janeiro: population distribution in accordance with urban and rural areas



Source: IBGE

In this context, the Metropolitan Region of Rio de Janeiro, which comprises 19 municipalities, has, according to the 2010 census (IBGE), a total population of 11.835.708 inhabitants, 99.5% of them can be found in urban areas and only 0.5% in rural areas.

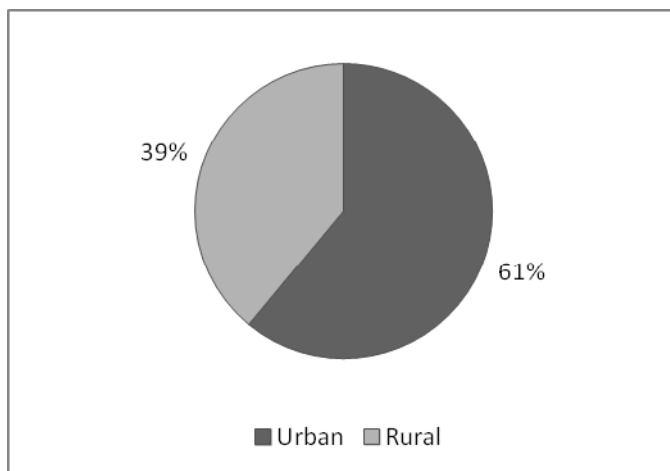
In São Tomé – and this seems to be a global trend – the situation seems to be the same: the majority of the population, although not as striking as in Brazil, is concentrated in urban areas, as shown in Figure 8⁷.

Afonso (2008: 6) complements this information:

It is estimated that the population of São Tomé and Príncipe Island reckons currently, according to the statistics, on one hundred and seventy (170) thousand individuals. About 47% of the population is young and predominantly urban. The deterioration of economic indicators, especially in rural areas, where poverty levels are extreme, has caused an acute exodus and the city of São Tomé is the main location of the occurrence of this phenomenon.

⁷ Data obtained in <http://www.alea.pt/html/lusofonos/html/fichluso.asp?pais=7>

Figure 8. São Tomé: population distribution in accordance with urban and rural areas



Source: INE, Estatísticas da CPLP 2012

Afonso (op. cit.: 13) also alludes to the fact that greater or lesser use of prestige standards depends on the socio-cultural level of the users of the language. It means, ultimately, considering the education level of the individuals:

The Portuguese used in São Tomé and Príncipe officially follows the standard European Portuguese, but in reality coexists with different sublanguages (registers) of Portuguese, which remains faithful to the rule, others with greater or lesser degree of displacement. Obviously, the degree of deviation varies with the environment (rural/urban) and socio-cultural level of the users.

Despite the distance separating the African and Brazilian communities and the different socio-history aspects of each of them, there are linguistic and social aspects that brings them together instead of distancing them, and the cities convenes spaces in which the dialectal differences seem to neutralise and obey common principles.

4.2. Complementary variables

4.2.1. Age and sex

As shown in Table V, when the performance of individuals at all education levels is taken into consideration, among social factors, age appears as the second salient variable for the application of the agreement rule in BP (both in Copacabana as in Nova Iguaçu), while sex is shown relevant in the Copacabana and São Tomé samples (ST(1)). In the last

community, however, age is relevant when considering only individuals of the elementary education level.

In Figures 9 and 10 below, one can observe the R.Ws achieved for the variable age in the cases mentioned.

Figure 9. Actuation of age for the application of number mark within the NP, in R.Ws, in Copacabana and Nova Iguaçu (all education levels)

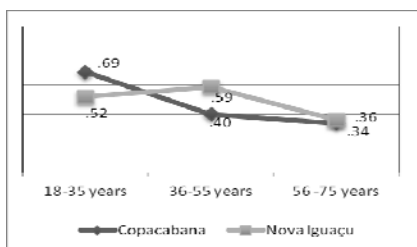
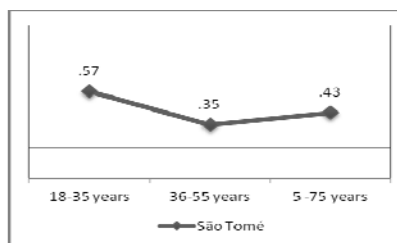


Figure 10. Actuation of age for the application of number mark within the NP, in R.Ws, in PST (speakers of elementary and medium education levels)



Even though Copacabana and Nova Iguaçu can be considered as locations belonging to the same speech community (the Metropolitan Region of Rio de Janeiro), which is supported by studies on other linguistic variables, such as /S/ and /R/ in the syllable coda (according to, among others, Lima, 2006; Brandão, 2012c; Brandão, 2012d), individuals behave differently according to their age groups. In Copacabana, it is clear that the trend of the adoption of the agreement rule by the younger individuals (R.W. .69), with distancing indexes from middle-aged and older, respectively, .29 and .33 points, is quite significant. In Nova Iguaçu, older individuals are, as in Copacabana, the most prone to marking cancellation (pr .36), whereas, among younger, middle-aged individuals (.59) stand out on the use of the agreement mark, with indexes quite close to the younger individuals (.52).

In the aforesaid analysis of São Tomé, based on the speech of 22 subjects (Sample ST (2)), there is a curve that is different from the others: even though younger individuals use more number marks (.57), the middle-aged individuals (.35) tend to use them on a smaller scale than the older (.43).

Despite the differences, generalisations can be drawn:

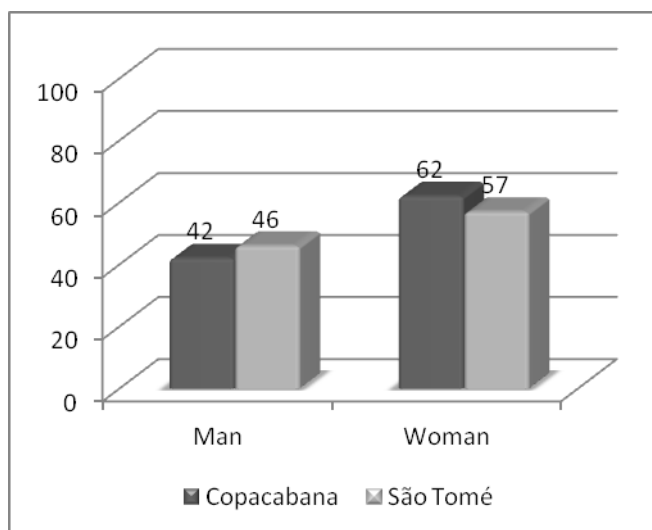
(a) younger and older individuals have similar behaviours in all communities, more likely (R.W. above .50) and less favourable (R.W. below .50) for the implementation of the number mark.

(b) the middle-aged individuals have fluctuated behaviour according to the community: they can employ the use of the rule with more productivity than younger individuals (as in Nova Iguaçu, R.W. .59) or they may tend not

to apply it, sometimes surpassing, in this sense, the older (as in São Tomé, R.W. 35) hence, overcoming them (as in Copacabana, R.W. 40)

The variable *sex* is relevant only in Copacabana and São Tomé (both in the sample with all education levels as in the elementary and medium levels). Given that the rates showed no significant quantitative differences, the rates obtained from the first sample can be seen in Figure 11.

Figure 11. Actuation of *sex* for the application of number mark within the NP, in R.Ws, in Copacabana and São Tomé (all education levels)



The results meet the initial expectations: women, when it comes to linguistic variables subject to social evaluation, tend to use more prestigious variants. Even in Nova Iguaçu, in which sex was not relevant to the application of the rule, the percentage rates indicate the same direction, even with a small percentage margin: women with 92% and men with 90% agreement rate.

4.2.2. *Phonic salience and nucleus animacy*

From a structural point of view, besides *linear and relative position of the constituent within the NP*, *phonic salience* (selected for Copacabana and Nova Iguaçu) and *nucleus animacy* are relevant in the samples of Nova Iguaçu and São Tomé (see Figures 12 and 13).

Figure 12. Actuation of the *Phonic Saliency* for the application of number mark within the NP, in R.Ws, in Copacabana and Nova Iguaçu (all education levels)

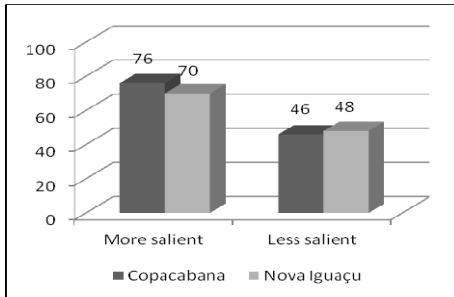
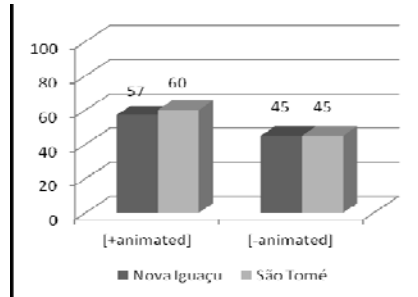


Figure 13. Actuation of the Nucleus Animacy for the application of number mark within the NP, in R.Ws in Copacabana and São Tomé (all education levels)



The two variables have been shown significant in many studies, not only in PB but also in other varieties.

The *phonic saliency* variable control aimed to test the hypothesis of Naro & Scherre (1977) according to which less marked forms, in other words, those in which the difference between the singular and the plural is only due to the presence of the morphemes of number – as in *filho/filhos* (child/children) – are more prone to loss of the marks than the forms as *[o]vo/[ɔ]vos* (egg/eggs).

In view of the results obtained, the seven factors initially defined were distributed in only two: (a) less salient, items ending in oral or nasal vowel – *filho/filhos* (son/sons), *irmã/irmãs* (sister-sisters); (b) more salient, items ending in –R, S, L e –ÃO and dual marks: *cantor/cantores* (singer/singers); *mês/meses* (month/months); *atual/atuais* (current/current); *cartão/cartões* (card/cards); *novo/n[ɔ]vos* (new/new).

As can be attested, the more marked words from the phonic point of view tend to retain the number marks, significantly in both communities (the R.Ws in Copacabana, for example, are 30 points higher than the less marked).

Regarding the semantic animacy features, it is evident in samples of Copacabana and Nova Iguaçu, that the nucleus [+ animated] (R.W. ranging from .46 to .48) disfavours non marking, contrary to what occurs with the feature [-animated], which clearly favours it (R.W. ranging from .70 to .76).

4.2.3. *Frequency of use of a Creole and individual performance*

As already mentioned, education level, in complex societies such as urban ones, may be related to other variables such as age and gender, as discussed here. The STP analysis in which only fundamental and medium educational level individuals were confronted – sample ST(2) – and the one in which we

focused on the speech of students aged 10 to 19 years – sample ST(3) – also account for this complexity.

In the analysis with sample ST(2), already referred to and incorporated with the speech data of 22 individuals, the variable *frequency of use of a Creole* was the third most relevant to the non-application of the plural mark, surpassed only by the variables related to the position of constituents and education level. Although all individuals who contributed to the formation of the sample have declared themselves Portuguese speakers such as L1, we took into consideration the STP multilingual society character in which there is three creoles and other minority languages. There are individuals who (a) only use Portuguese and (i) not have fluency or (ii) have fluency in a Creole; (b) use Portuguese and one or more Creoles with (i) less or (ii) higher frequency (in the market, during leisure moments with friends). Furthermore, the use of Creoles have a stigmatising character, as many informants mentioned the fact that their parents forced them to talk, even at home, only Portuguese so that they could master the rules of prestige and be able to take advantage of the few opportunities for socioeconomic improvement.

Due to this situation, we sought to verify whether individuals who supposedly make more frequent use of a Creole (usually, the forro) than Portuguese would be more liable to avoid the usage of the plural mark.

As indicated by Table XIV, individuals who claim to communicate preferably in Creole (high frequency) are the most likely not to implement the number mark (R.W. .25). With these individuals, there is a contrast among those who never or poorly expressed themselves in Creole (R.W. .59) and those who use it only occasionally (R.W. .48).

Table XIV. Actuation of frequency of use of a Creole for the application of number mark within the NP – sample ST(2)

Frequency of use of a Creole	Oco	Perc.	P.R.
Zero /Low	964/1039	92.8%	.59
Medium	977/1117	87.5%	.48
High	129/219	58.9%	.25

However, the cross *frequency of use of a Creole* and *education level* (fundamental and medium) reveals that the first of these variables only seems to take effect among fundamental level individuals (see Table XV).

Table XV. Crossing of frequency of use of a Creole and Education level – sample ST(2).
Application: number mark presence within the NP

Frequency of use of a Creole	Education Level				Partial total mark	
	Fundamental (5 to 8 years)		Medium (9 to 11 years)			
	Apl.	%	Apl.	%	Apl.	%
Zero/Low	224/964	81%	740/964	97%	964/1039	92.8%
Average	325/977	75%	652/977	96%	977/1117	87.5%
High	129/219	58.9%	----	-----	129/219	58.9%
Total	678/930 – 73%		1392/1445 – 96%		2070/2375 – 87%	

It should be verified that:

(i) none of the 11 informants with medium education level said that they used Creole frequently;

(ii) in this group of individuals, the application of the number mark is quite significant: 97% and 96%, respectively, among those who use it with zero/low (7 subjects) and medium (4 subjects) frequencies;

(iii) among the 11 informants of fundamental education level, gradually increasing, the higher the frequency of use of a Creole, the lower the tendency of plural mark in NP. In this group, two uses it on a reduced scale, six use a medium manner and three individuals use the Creole on a large scale.

The analysis of the speech of students from São Tomé, even if not developed in the classic standards, because it was based on the speech of only nine informants and there was no filling of all social cells, did not indicate, as expected, the grade the student was as the more relevant social variable, but instead the *individual performance*. In Table XVI⁸ below, the results are summarised.

⁸ The abbreviations in the table correspond to information about the informants: ST (Sao Tome); E (student); 1 to 9 (student identification); E (among 10 and 14 years) or F (among 15 and 19 years); 6 (6th class), 8 (8th and 9th class) and D (10th and 11th class).

Table XVI. Actuation of individual performance for the presence of plural mark within the NP, based on a sample of speech of students from São Tomé

<i>Individual Performance</i>							
Informant	Token	%	R.W.	Informant	Token	%	R.W.
ST-E1-E6m	11/19	57.9%	.09	ST-E6-FDh	80/88	90.9%	.52
ST-E2-E6m	28/28	100%	(1.00)	ST-E7-FDh	35/38	92.1%	.62
ST-E3-F8h	62/62	100%	(1.00)	ST-E8-FDh	80/83	96.4%	.74
ST-E4-F8m	261/261	100%	(1.00)	ST-E9-FDm	91/91	100%	(1.00)
ST-E5-F8m	26/26	65.4%	.09				
Input: .05				Significance: .000			

The results suggest that the plural nominal agreement has, in São Tomé society, a strong socio-economic-cultural component. While, in the speech of four students, the rule is categorical (two students from the sixth grade, one from the eighth and another from the eleventh), in the speech of another five it is variable, in a major or a lesser degree (one from the sixth grade, one from the eighth and three from the eleventh). This framework is certainly associated with the aspects related to the family environment, the major or minor exposure to culture, the languages spoken in the region and the type of school attended. It is worth noting the comments of two of the students who apply categorically to the rule: one student stated that his father provided all means for his intellectual development, another that his parents preferred that he should study at the Portuguese School because they thought that school, unlike public institutions, functioned in a smoother manner and, in principle, would provide a better quality of teaching.

It is noteworthy that, in the speech of students who applied the variable agreement rule, the only other variable selected for the marking or not of the plural is the linear and relative position of the constituents within the NP, as already discussed.

5. Preliminary Conclusions (still)

Due to the complex nature of the plural nominal agreement in the universe of Portuguese speakers and whatever little that is yet known of its application in the European variety, this section will present the possible conclusions, based on the samples considered, and reserves the new ideas and hypotheses for section 6, some of them arising from empirical observations to be tested.

This article focused on the plural agreement within the NP as categorical rule and as variable rule. It was demonstrated that the canonical rule of agreement, which requires the presence of the mark in all inflectional NP constituents, has full application in European urban varieties (continental or not) currently and therefore is categorical. Even in Funchal, with a settlement that was made with the assistance of persons of different origins, there is, in relation to nominal agreement, uniformity which assumes a highly consolidated standard.

For the implementation of variable agreement rule, variables that co-operate here were distributed in two categories: relevant and complementary. It is believed that the non-European varieties, from a structural point of view, follow the same general pattern that, applied with greater or lesser frequency, implies the decrease of marks from the left margin of the NP.

We consider social-cultural factors crucial for the application of this standard, including the greater or lesser knowledge of individuals according to the standards deemed as preferential and according to more interaction with other individuals who use and value them. Accordingly, in urban areas, where the majority of the population is now concentrated, the education level, that reflects the exposure time of the individual not only to socially valued patterns, but also to critical structures seen as deviant (stigmatised), emerges as the main external variable that “licenses” the diversified indexes of NP marks presence.

In non-European varieties, the high overall percentage of agreement registered in urban areas, conceal, however, situations that are quite diverse, evidencing the existence of different performances, as suggested by the continua proposed here, of plural marking in prototypical NPs (Figure 2) and urbanisation/education level (Figure 6).

The results of the analyses of the commented varieties suggest, in summary:

- (a) in Portuguese, there are two general trends regarding the presence of marks in the NP constituents: full marking and variable marking, whose implementation although strongly determined by the linear and relative position of the constituent within the NP and also dependent on the action of factors that act on the nucleus (animacy, phonic salience), it is subject to the action of external factors, among which the education level that seems congregate different aspects that characterise individuals from a social point of view;
- (b) hence, as indicated by Brandão & Vieira (2012b):
 - (1) there seems to be, in percentage terms, a general tendency to adopt the rule of agreement in the NP, even if with different status in the different urban areas considered here;
 - (2) in Brazilian and STP even though preference is given to the agreement – preference compliant to the communities where the access to school, to a greater or lesser extent, is provided to individuals – the rule is variable, stressing that the STP has higher overall rates of plural marking than BP, which is the variety where the continuum of education level has clearly a scaling profile;
 - (3) structural variables in the syntactic level (*linear and relative position of the NP constituents*), semantic level (*animacy*) and phonetic level (*phonic salience*) act as constraints in the achievement of the plural marks, duly attested in BP and STP, according to the following trends:
 - (i) in syntactic structures used to indicate the plural by means of inflection, the left area is the most prone to the mark presence;
 - (ii) the plural mark is concentrated in the pre-nuclear constituents, when it comes to the prototypical NPs, and in the first constituent in the case of non-prototypical;
 - (iii) the feature [+ animate] or [-animated] of the NP nucleus determines respectively the presence or absence of the number mark;
 - (iv) greater or lesser degree of phonic salience conditions determines, inversely, higher or lower index of the plurality mark;
 - (v) the similarity of non-marking patterns in the different varieties suggests the existence of a trend of universalising character, inherent to the Portuguese varieties emerging from situations of language contact;
- (c) the different plural marking percentages opposing the urban and non-urban areas varieties does not constitute, in essence, diatopic but diastratic variation considering not only the compact population concentration in urban areas, but also different conditions of access to education and cultural patterns in opposition to rural areas in

- which the heavily contingent of the population is of low education and where there are, in some cases, even illiterate;
- (d) considering only percentage indexes, according to Labov (2003), we can draw the following picture regarding the varieties targeted in this study (cf. Table XVII)

Table XVII. Plural agreement within the NP: rule status in three varieties of Portuguese

Plural agreement within the NP	
Rule Status	Variety/Sub-variety
Categorical	European (EP)
Semi-categorical	São Tomé (STP)/cultivated Brazilian (BP)/cultivated
Variable	São Tomé (STP)/ popular Brazilian (BP)/popular

- (e) whether in São Tomé or Rio de Janeiro, the sociolinguistic polarisation is based on different levels of agreement that oppose more educated individuals to those less educated;
- (f) with the principle of uniformitarianism in mind which was postulated by Labov (1972) and envisages the performance of the same variation processes and changes in different stages in the history of a language or, in new terms, at different stages in the history of the varieties of a language that are exposed to common development situations (in this case, for example, inter-linguistic contact), we believe that performing analyses on African varieties, yet emerging, in multilingual contact condition, may greatly help understanding the process of BP constitution.

6. Some (few) insights

Much has been discussed, based on the nominal and verbal agreement, on the origins of the Portuguese in Brazil, as noted in the introduction to this article. The data that is available about BP and now about STP, in contrast to those that are released about Portuguese L2 on Cape Verde and Angola (JON-AND, op. cit) suggest that the massive inter-linguistic contact that would have determined the irregular learning of Portuguese would be the reason for the variable rule just described.

Taking into account that the languages/varieties change throughout the course of time (considering BP, whose real-time analyses of short duration have shown, for example, the increase in the agreement rate depending on the actuation of social/situational factors), draws attention to the apparent monolithic character of the nominal agreement rule in EP that is used, as it seems, according to the canonical standard in different regions (here, the mainland and Madeira Island were taken into consideration), for speakers of all education levels.

In the grammars of the sixteenth century and in those of the following seventeenth and eighteenth centuries, there is no reference regarding the alteration of patterns of agreement, even though there are several remarks about phonetic-phonological aspects in terms of correction and/or change. It is known that the European variety has undergone changes in the eighteenth century, but apparently these did not affect the inflectional system, but the consonant and vowel systems, among which the most significant was the neutralisation of the mid and high vowels in the pre-stressed position, as attested by Teyssier (1997) and Castro (1991). It should be brought to mind that Peres & Moia (1995) indicate the agreement as one of the “critical areas” of Portuguese, not listing variable cases in the NP structure.

Given that Naro & Scherre (2007) have observed some examples of variable nominal agreement in medieval works and in dialectal works, even though we may find further restrictions to the data presented by them (c.f. the relevant critics made by Galves, 2012), it would be important to undertake an intensive and systematic sociolinguistic study which could provide the real dimension of the hypotheses scrutinised by them. A real-time long-term study based on the medieval works and also on those developed in subsequent centuries, with contextualised data collection, would enable the verification of variable agreement patterns evidence in other stages of EP. On the other hand, it would also be important to verify the variable performance in the speech of more conservative areas of Portugal, where there are small communities still relatively isolated. In those areas, it would be possible to find cases, although rare, according to some researchers, which can be classified as agreement absence⁹.

Among the factors that could be listed as motivators of canonical agreement regarding EP, the tendency to maintain the coda of the syllable (in opposition to the universal tendency to CV syllable) would be a possibility. In BP, in which the cancel and vocalisation of the consonants are usual

⁹ Maria Antónia Mota, CLUL researcher, in personal information, noted that during the conduction of the work with her students on Atlas data, she found virtually no cases of absence of agreement in the NP, and these rare cases are associated with phenomena void of hesitation or reformulation of speech. We take this opportunity to acknowledge some observations regarding EP and BP, incorporated in the final part of this article.

processes, even when these constitute morphological marks (the /R/, of the infinitive, that is hardly produced in some dialects, the /S/ nominal number marker, which may not occur in some NP constituents), the number mark that, in the speech of some speakers, generally occurs only in the pre-nuclear constituent seems to spread throughout the NP: [the [black cat]]SN +plural.

EP does not tend to suppress consonants in this context. To Mota (2012)¹⁰, the issue lies in the phonology/morphology interface: in EP, which is already prone to the maintenance of the coda segments, a phoneme with morphological value would have extra weight in the maintenance of the canonical nominal agreement. Besides that, studies conducted in the acquisition area demonstrate that after acquiring the CV structure, the first CVC syllabic type acquired by Portuguese children is comprised by a fricative, as stated by Almeida (2007: 74-75), when synthesising, in her work, research results on the topic:

The study of the syllabic structure by Portuguese monolingual children demonstrated that there is an order in the appearance of the segments in the end of syllable (Freitas, 1997). In the first place, appears the fricative coda in the final position of the words, later this emerges in an internal position of the word: first in stressed syllable, and then in non stressed syllable. After that, the liquid consonants can be seen, at first at the end of the word and later in the internal position (Freitas, 1997). [...] This author [Correia, 2004] demonstrates, with additional empirical evidence, that the fricatives in the final syllable appear long before the liquid consonant in this position. Regarding the liquid consonant form of the final syllable, it was attested that the final vibrant consonant is the first liquid consonant to be stabilised at the end of the syllable, whereas the stabilisation of the same occurs after the first syllable and then in unaccented syllable. The lateral consonant is then stabilised in all contexts. The last liquid consonant parameter is a vibrant consonant in the word internal position, and its stabilisation occurs initially in stressed syllable and later in the non stressed syllable. It is evident, therefore, that two linguistic facts which relate to the development of segments appear in the final syllable: the position in the word and the position of the accent.

Regarding BP fricative coda acquisition, Mezommo (2004: s/p.) observes that:

[...] it was possible to identify the age of the appearance of the plural marker (at 1:8) in VC rhyme and the early appearance of the lexical coda regarding the morphological coda (at 1:4). [...] Even though the plural

¹⁰ Paper presented at the Sociolinguistics Symposium 19, "Subject-Verb agreement 3rd forms in Portuguese. The significance of phonological and phonetic phenomena: a comparative study of agreement patterns in African, Brazilian and European varieties".

marker appeared after the lexical coda, the same is acquired beforehand, at 1;9. The lexical coda, on the other hand, has stabilised its use from 2;6. The domain process of /S/ in lexical and morphological codas shows pathways and uses of different repair strategies, inferring that the child treats these two types of coda in a unique manner from the start of its production until its stabilisation. Besides these differences, different results in the acquisition of lexical and morphological coda were also observed regarding grammatical categories. However, it is believed that a greater number of data is necessary to evaluate the variable role of word class in the /S/ final coda acquisition.

Based on this work, it is believed that the advance of the majority domain phonemes in final codas mainly determined by linguistic factors such as the tone, as these do not possess morphological information and the stressed variant favours their productions. In the case of the final coda /S/, the syllable appears to have a secondary role as, unlike other phonemes in coda, the /S/ termination is not favoured by the stressed syllable, but the non stressed syllable (according to the results of statistical analysis that are shown in Table 3). In this case, the morphological factor plays a key role, “leveraging” the acquisition of final /S/ codas.

[...] Despite appearing after the lexical coda, the morphological coda seems to help in the early stabilisation of the final coda in relation to the medial. These results corroborate partially the study [on EP acquisition] performed by Freitas, Miguel and Faria (2001)¹¹ which shows that morphosyntactic information assists the prosodic structure.

For a better understanding of the acquisition of the /S/ as a morphological mark, considering, on the one hand, both in BP and EP, that /S/ is the first coda that appears in the child's speech, on the other, that, regardless, the agreement rule has different standards in the two varieties, therefore it is timely to consider, in the case of BP, the observations made by Capellari & Zilles (2002). In this work, there are not just reports on the results of the previous research conducted by Lamprecht about the NP marking in the speech of bilingual children (Portuguese and German) and monolingual (in German and Portuguese) between 2.3 and 5.6 years but also it is developed a punctual study on the speech of a child between 4 and 8 years old.

Capellari & Zilles summarise some of the findings of the research conducted by Lamprecht (1997):

Lamprecht (1997: 115) establishes the following conclusions. The first stresses emphasis on the *input* role. Bilingual children do not have full

¹¹ The study to which Mezommo refers is FREITAS, M. J.; MIGUEL, M., Faria, I. Interaction between Prosody and Morphosyntax: plurals within codas in the acquisition of EP. In: B. HOEHLE, B.; WEISSENBORN, J. (eds) *Approaches to Bootstrapping. Phonological, Lexical, Syntactic and Neurological Aspects of Early language Acquisition*. Amsterdam: John Benjamin Publishers, 2001. vol. 2. pp. 45-58.

control of the system of rules of the German plural because they are less exposed to this system, because it is their second language. The same applies to monolingual in Portuguese. Those exposed to the most educated variant present a more intense use of the plural marking rule in speech. The second conclusion goes towards explaining why bilingual children make more standard plural in Portuguese than monolinguals. According to Lamprecht (1997: 116), *bilingual children seem to pay more attention to word endings than monolingual speakers of Portuguese because in German such attention is vital*. That is, in the termination of German words, it is very important not only for marking of the plural, but also for case declinations, with a frequent standard equivalent to WORD = ROOT + SUFFIX. However, Brazilian children are not faced with this structure as, in Portuguese, there is no case marking and the plural marking rule is variable and governed by syntactic rules related to the NP position, among other factors making optional the use of the suffix (Cappelari & Zilles, 2002: 197-198).

Regarding the punctual study conducted by Cappelari & Zilles (2002: 215-216), the main conclusions can be summarised as follows:

What we see therefore is a behaviour similar to that of adults. [...] Finally, it seems important to mention one last point. The influence that the school may have to encourage the use of standard forms of the language is not clearly depicted in this work. The speech of Carmela suggests that she implements two types of grammar, and this seems to be related to the discursive contexts. Thus, even if, in the interactions, there was no formality, her speech was based on a written text, standard forms that were deemed as relevant.

As attested, it all depends, ultimately, on the *input* received by the child during the acquisition of the language, and it is worthwhile to mention, in this sense, two more passages written by Lamprecht (apud Cappelari & Zilles: 195; 197):

Due to the input to which they are exposed, children appear to believe that the plural marking is not necessary when the plural is clear in the determiner or by means of the use of a numeral. In summary, children are aware about the system but seem unaware that the marking is required (p. 195).

Regarding the *input*, the author argues that their influence was evident in her research. Lamprecht highlights the performance of two children, designated as markers, and attests that these children come from families in which the language used is very close to the cultivated standards (p. 197).

In summary, in EP, the input for the agreement is robust, regardless of the education level of the individual. In BP, where the rule is variable, the input

is differentiated depending on the socio-economic-cultural class to which the child's family belongs: major exposure to formal education and cultural goods, in particular, will increase or reinforce the use of prestigious patterns in urban communities. Although the STP reality is far more complex than BP, one could say that the scenario resembles this variety.

Therefore, it would be interesting to conduct studies on agreement acquisition in which EP and BP or EP and other non-European varieties of Portuguese are contrasted.

Returning to the specific case of EP, it is worth mentioning, subliminal aspects of identity nature, as can be seen in informants' speech such as OEI-B-1-h, which demonstrated a great concern to defend Portuguese language, which is a reason of pride:

Portuguese individuals are attaining to the language that we all know is rich and is one of the major values that we have and we do not want to give in to the influence of the external side and we would say that we are confronting you because ... eh man! because I am not sure how many millions you are...and I am not even reckoning the Africans as it is like this () right now they have an official language that is theirs ... Portuguese is the official language but they have their own language ... they just talk to each other and no one ... no one will understand... now we have to respect their culture ... just like that ... if you are developing a language () over the centuries we have to respect that ... this is how I – I do not see any agreement between the Americans and the British and Americans also have a language that is evolving in a direction and the language spoken by the British is going towards another line and then that is a matter of a universal language ...¹²

When asked about the changes they observe in EP, Portuguese informants mention diatopic “accent” differences (“it is possible to perceive the accents in Alentejo ... further north is another Alentejo ... I know that Algarve have another accent”)¹³ or, regarding the generational diversity, refer to the “slang” (“in the full issue of the slang ... the language spoken by the youth is

¹² “o português está-se a agarrar muito à língua sabemos que é rica que é um dos grandes valores que temos e não queremos deixar-nos influenciar mais pelo lado de fora e diria que estamos muito em confronto com vocês porque... eh pá!... porque vocês são não sei quantos milhões... já nem contando com os africanos porque é assim o () neste momento têm uma língua praticamente deles é... o oficial é português mas têm uma língua deles... praticamente eles falam entre si e ninguém... ninguém lhes percebe... agora a cultura deles temos que respeitar... isso... se vocês estão a desenvolver uma língua () com o passar dos séculos temos que respeitar isso... isto é como eu – não vejo acordo entre os americanos e os ingleses e a língua dos americanos também está a evoluir num sentido e a dos ingleses noutro depois há aquela questão de uma língua universal...?”

¹³ “consegue-se perceber os sotaques no Alentejo... mais a norte é outro alentejano... sei que o algarvio tem outro sotaque”.

full of slangs ... totally ... that is where we are completely under that influence”)¹⁴. None of them referred to the morphosyntactic variation, even though the informant that made these affirmations presented here, like others, at some point, expressed it as such: “defendo aquilo que a gente tem, mas não, mas temos que fazer um acordo para o manter, para que o brasileiro fale português, para que a gente sejamos não sei quantos milhões a falar português daqui a uns anos”: a clear example of variation¹⁵.

Given that the linguistic expression is a factor of individual/community identity, it would, however, be worthy to verify in the urban variety of Rio de Janeiro specifically, whether certain variants, despite the negative evaluation of the society as a whole, are being used, also subliminally, as a mark that grants the sense of belonging of a group. It can be observed, especially in the Copacabana sample, that the five individuals that presented high levels of number mark cancellation, in contrast with other informants (six of them categorically applied the mark), live in so-called communities, in which they feel fully integrated and to which they aggregate a high value: a man and a woman from the first age, in the first age group, pertaining to the fundamental education level; a man and a woman in the second age group, mid-level and a man of the third age group, also of mid-level. We could, therefore, formulate the assumption that in social communities with high social contrast such as Copacabana, the absence or presence of number marking (nominal or verbal) may be functioning as a mark of the group identity.

There is, however, a long way to go, although much has been said and known about the nominal agreement in the Brazilian variety: works on evaluation, attitudes and practice communities, in the line of Eckert (2005, 2008), are yet to be accomplished. “Diving the third wave” of the studies on variation would enable us to observe “the meanings that motivate particular variable performances” (Eckert, 2005: 30) and also to deepen this theme in EP and other non-European varieties. It is time to set new directions for the research of this stimulating and inexhaustible subject.

¹⁴ “no calão totalmente... da questão de falar dos jovens está totalmente calão... totalmente... aí é que nós somos muito influenciáveis”.

¹⁵ I defend what we have but we have to make a pact to ensure that Brazilians will speak Portuguese so that there can be millions of us to speak Portuguese some years from now.”

References

- Afonso, B. C. (2008) *A problemática do bilinguismo e ensino da língua portuguesa em S. Tomé e Príncipe*. Dissertação (Mestrado em Ensino do Português como Língua Segunda e Estrangeira). Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Lisboa.
- Almeida, E. M. (1997) *A variação da concordância nominal num dialecto rural*. Dissertação de Mestrado em Letras Vernáculas, Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro.
- Almeida, L. B. (2007) A aquisição das codas mediais por uma criança bilingue em português e francês. *Actas do XII Encontro Nacional da Associação Portuguesa de Linguística*, Lisboa, APL, 73-85.
- Andrade, P. R. D. (2003) *Um fragmento da constituição sócio-histórica do português do Brasil: variação na concordância nominal de número em um dialecto Afro-Brasileiro*. Dissertação (Mestrado em Linguística Histórica). Salvador: Instituto de Letras, Universidade Federal da Bahia.
- 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* (R. Muhr, editor), pp. 327-348. Wien: Peter Lang Verlag.
- Baxter, A. N. (2004) The development of variable NP plural agreement in a re-structured African variety of Portuguese. In *Creoles, contact, and language change* (G. Escure, & A. Schwegler, editors), pp. 97-126. Amsterdam/Philadelphia: John Benjamins.
- Baxter, A. (2009) A concordância de número. In *O português afro-brasileiro* (D. Lucchesi; A. Baxter & I. Ribeiro, eds.), pp. 269-293. Salvador: EDUFBA.
- Bortoni-Ricardo, S. M. (2004) *Educação em língua materna: a sócio-linguística na sala de aula*. São Paulo: Parábola Editorial.
- Braga, M. L. (1977) *A concordância de número no sintagma nominal no Triângulo Mineiro*. Dissertação (Mestrado em Língua Portuguesa) – Pontifícia Universidade Católica, Rio de Janeiro.
- Braga, M. L & Scherre, M. M. P. (1976) A concordância de número no SN na área urbana do Rio de Janeiro. In *Anais do Primeiro Encontro Nacional de Linguística*, pp. 464-477. Rio de Janeiro: PUC.
- Brandão, S. F. (2011) Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, 15(1), 164-178.
- Brandão, S. F. & Vieira, S. R. (2012a) A concordância nominal e verbal no Português do Brasil e no Português de São Tomé: uma abordagem sociolinguística. *Papia: Revista Brasileira de Estudos Crioulos e Similares*, 22(2), 7-41.
- Brandão, S. F. & Vieira, S. R. (2012b) Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa, Revista de Linguística*, (56)3, 1035-1064.
- Brandão, S. F. & Silva, C. B. (2012c) Róticos em coda silábica na fala de Nova Iguaçu-RJ com base em diferentes tipos de *corpus*. *Revista do GELNE*, (14) nº especial, 21-40.

- Brandão, S. F. (2012d) A variável (S) na fala do Estado do Rio de Janeiro. In *Múltiplos olhares sobre a diversidade linguística: uma homenagem a Vanderci de Andrade Aguilera* (F. Altino, editor), pp. 230-250. Londrina: Midiograf.
- Brandão, S. F. (2013) Réalité sociolinguistique brésilienne et géolinguistique pluridimensionnelle. In *Current approaches to limits and areas in Dialectology*. (E. Carrilho; C. Magro; X. Álvarez, ed.), pp. 3-26. Cambridge: Cambridge Schollars Publishing.
- Campos, O. G. L. S. et al. (1993) Flexão nominal: indicação de pluralidade no sintagma nominal. In *Gramática do português falado*. (R. Ilari, ed.), vol. 2, pp. 111-134, 2.^a ed. Campinas – SP: Editora da UNICAMP.
- Capellari, E. T. C. & Zilles, A. M. S. (2002) A marcação de plural na linguagem infantil – estudo longitudinal. *Revista da ABRALIN*, 1(1), 185-218.
- Castro, I. (1991) *Curso de história da língua portuguesa*. Lisboa: Universidade Aberta.
- Chambers, J. K. (2004) Dynamic typology and vernacular universals. In *Dialectology meets Typology. Dialect Grammar from a Cross-Linguistic Perspective* (B. Kortmann, editor), pp. 127-145. Berlin/New York: Mouton de Gruyter.
- Cruz, L. S. (2003) Variação dialectal no território português: conexões com o português do Brasil. In *Análise contrastiva de variedades do português: primeiros estudos* (S. F. Brandão & M. A. Mota, eds.), pp. 181-196. Rio de Janeiro: In-Fólio.
- Eckert, P. (2005) Variation, convention and social meaning. Paper presented at the Annual Meeting of the Linguistic Society of America. Oakland CA., Jan. 7.
- Eckert, P. (2008) Variation and the indexical field. *Journal of Sociolinguistics* (12)4, 453–476.
- Figueiredo, C. F. G. (2008) A concordância variável no sintagma nominal plural do português reestruturado de Almojarife. *Papia: Revista Brasileira de Estudos Crioulos e Similares* 18, 23-43.
- Figueiredo, C. F. G. (2010) *A concordância plural variável no sintagma nominal do português reestruturado da comunidade de Almojarife, São Tomé (Desenvolvimento das regras de concordância variáveis no processo de transmissão-aquisição geracional)*. Tese (Doutorado em Linguística). Faculdade de Ciências Sociais e Humanas, Departamento de Português, Universidade de Macau, Macau.
- 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; A. T. Castilho; M. A. Sibaldo & R. B. Lima, eds.), pp. 123-149. Maceió: EdUFAL.
- Guy, G. R. (1981) Linguistic variation in brazilian portuguese: aspects of the phonology, syntax, and language history. PhD Dissertation. Philadelphia: University of Pennsylvania.
- Jon-And, A. (2010) Concordância variável de número no SN no português L2 de Moçambique – algumas explicações sociais e linguísticas. *Revista de Crioulos e Base Lexical Portuguesa e Espanhola*, 2, 28-50.
- Jon-And, A. (2011) *Variação, contato e mudança linguística em Moçambique e Cabo Verde. A concordância variável de número em sintagmas nominais do português*. Tese (Doutorado). Department of Spanish, Portuguese and Latin American Studies, Stockholm University.

- Labov, W. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. (1994) *Principles of linguistic change: internal factors*. Oxford: Blackwell Publishers, vol. 1.
- Labov, W. (2001) *Principles of linguistic change: social factors*. Oxford: Blackwell Publishers, vol. 2.
- Labov, W. (2003) Some sociolinguistic principles. In *Sociolinguistics: the essential readings* (C. B. Paulston & G. R. Tucker, eds.), pp. 234-250. Cambridge: Blackwell Publishing.
- Lemle, M. & Naro, A. J. (1977) *Competências básicas do português*. Relatório final de pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro de Alfabetização (MOBRAL) e Fundação Ford. Rio de Janeiro: Fundação Movimento Brasileiro de Alfabetização.
- Lima, L. G. (2006) *Atlas Fonético do entorno da Baía de Guanabara-AFeBG*. Dissertação (Mestrado em Letras Vernáculas). Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2v.
- Lopes, N. (2001) *Concordância nominal, contexto linguístico e sociedade*. Tese (Doutorado). Salvador: Instituto de Letras, Universidade Federal da Bahia.
- Lucchesi, D. (2003) O conceito de transmissão linguística irregular e o processo de formação do Português do Brasil. In *Português brasileiro: contato linguístico, heterogeneidade e história* (C. Roncarati, & J. Abraçado, editors), pp. 272-284. Rio de Janeiro: 7 Letras.
- Mezommo, C. L. (2004) Interação entre a aquisição da fonologia e da marca de plural no português brasileiro: o domínio de /s/ nas codas lexical e morfológica. *Anais do VI Encontro do Círculo de Estudos Linguísticos do Sul – CELSUL, Florianópolis: CELSUL*.
- Mira, M. H. F. G. (1954) *Algumas contribuições para um estudo da fonética, morfologia, sintaxe e léxico da linguagem popular de Lisboa*. Monografia (Licenciatura em Filologia Românica), Faculdade de Letras, Universidade de Lisboa.
- Naro, A. J. & Scherre, M. M. P. (2003a) O conceito de transmissão linguística irregular e as origens estruturais do português brasileiro: um tema em debate. In *Português brasileiro: contato linguístico, heterogeneidade e história* (C. Roncarati, & J. Abraçado, editors), pp. 285-302. Rio de Janeiro: 7 Letras.
- Naro, A. J. & Scherre, M. M. P. (2003b) Estabilidade e mudança linguística em tempo real: a concordância de número. In *Mudança linguística em tempo real* (M. C. Paiva; M. E. L. Duarte, eds.), pp. 47-62. Rio de Janeiro: FAPERJ/Contra capa.
- Naro, A. J. & Scherre, M. M. P. (2007) *Origens do português brasileiro*. São Paulo: Parábola Editorial.
- Peres, J. A. & Mória, T. (1995) *Áreas críticas da língua portuguesa*. Lisboa: Caminho.
- Scherre, M. M. P. (1978) *A regra de concordância de número no sintagma nominal em português*. Dissertação (Mestrado em Letras). Rio de Janeiro: Departamento de Letras e Artes, Pontifícia Universidade Católica.
- Scherre, M. M. P. (1988) *Reanálise da concordância de número em português*. 1988. 2 v. Tese (Doutorado em Linguística). Faculdade de Letras, Universidade Federal do Rio de Janeiro.
- Teyssier, P. (1997) *História da língua portuguesa*. Tradução Celso Cunha. São Paulo: Martins Fontes.

Weinreich, U.; Labov, W. & Herzog, M. (1968) Empirical foundations for the theory of linguistic change. In *Directions for historical linguistics* (W. Lehmann & Y. Malkiel, editors), pp. 97-195. Austin: University of Texas Press.

Sílvia F. Brandão

Faculdade de Letras (Departamento/Programa de Pós-graduação em Letras Vernáculas)
Universidade Federal do Rio de Janeiro
silvia.brandao@terra.com.br