Language Variation and Change, **13** (2001), 91–107. Printed in the U.S.A.
© 2001 Cambridge University Press 0954-3945/01 $9.50

Phrase-level parallelism effect on noun phrase number agreement

MARIA MARTA PEREIRA SCHERRE
Universidade Federal do Rio de Janeiro

ABSTRACT

Linguistic parallelism (i.e., the tendency of similar forms to occur together within a stretch of discourse) has been shown to be very strong in many linguistic phenomena and in a vast number of languages. Examining the role of phrase-level parallelism on noun phrase number agreement, this article demonstrates that Puerto Rican Spanish and Brazilian Portuguese exhibit more similarities than differences with regard to this constraint. A detailed analysis of Brazilian Portuguese data is presented, and the results are compared with those found by Poplack (1980a) for Puerto Rican Spanish. Guy's (1981a) and Labov's (1994) hypothesis about missing zeroes in Brazilian Portuguese is also discussed. In conclusion, it is claimed that the phrase-level parallelism effect on noun phrase number agreement is embedded in a universal principle of linguistic use: parallel processing.

According to Poplack (1980a:55), "Puerto Rican Spanish (s) is variably subject to two weakening processes, aspiration and deletion, so that a phrase such as *las cosas bonitas* 'the beautiful things' can also be realized [lah 'kosah bo'nitah] or [la 'kosa bo'nita]." Therefore, in Puerto Rican Spanish, a plural noun phrase can exhibit all overt plural markers, some overt and some zero plural markers, or all zero plural markers without losing the semantic plural information. In order to account for the Puerto Rican Spanish data, Poplack (1980a:63–64) claimed that it is necessary to consider preceding markers across linear position as well as other important constraints, such as grammatical category, following phonological segment, and stress. Poplack's own analysis took into account each token with respect to its position in the noun phrase and to the position of preceding plural markers; her results are presented in Table 1 (her Table 3.4B). Poplack (1980a:64; emphasis added) asserted that[1]

> When preceding S, ∅S, SS, and S∅ are treated as a single category, they appear to have the same disfavoring effect on deletion (between .40 and .43).

The research on which this article was based was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico. I am deeply grateful to Anthony J. Naro, who taught me how to be a researcher; I hope I have picked up a few of his ideas. I would like to thank Shana Poplack, who inspired me with her research on Puerto Rican Spanish, as well as Helena Gryner, who gave me an unpublished version of Poplack's (1980a) article in 1977. Finally, my thanks go to Elisabete Malvar, who read the final version of this article; her suggestions were very helpful.

The main ideas of this article were presented at NWAVE-XXVIII, Toronto, October 1999.

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 1. *Contribution of position of token in noun phrase string and presence of immediately preceding marker to the deletion of plural -s in Puerto Rican Spanish*

	Position of Token in String		
	3	2	1
Marker(s) preceding token			
∅∅_	.73		
S∅_	.68		
∅_		.52	
S_		.44	
∅S_, SS_	.40		
None possible			.24

Source: Poplack (1980a:63, Table 3.4B).

> Table 3.4B, on the other hand, shows that the differential effect on deletion is actually due to the presence of a marker in the slot immediately preceding the token: ∅S, SS, and S continue to disfavor deletion . . . , while ∅ and S∅ favor it. The most favorable context for marker deletion is precisely when the two preceding markers have already been deleted. *This is exactly the opposite effect from what was found by Guy and Braga for Brazilian Portuguese, a language that marks plurality on the NP and VP in much the same way as Spanish.*

However, even the most detailed analyses of noun phrase agreement with respect to preceding markers in Brazilian Portuguese have not allowed an adequate comparison with Poplack's (1980a) results for Puerto Rican Spanish. Therefore, I analyzed new data from Brazilian Portuguese, keeping such a comparison in mind. I codified each token, taking into account its position in the noun phrase as well as the type and position of preceding plural markers. The results of this analysis are presented here.[2]

In this article, I discuss the effect of preceding markers across linear position on noun phrase number agreement in Brazilian Portuguese, a language that exhibits a pattern of variable number agreement similar to that of Spanish (see, e.g., Braga, 1977; Braga & Scherre, 1976; Guy, 1981a, 1981b; Naro, 1981; Naro & Scherre, 1996; Scherre, 1978, 1981, 1988, 1998a, 1998b; Scherre & Naro, 1991, 1992, 1998). First, I point out the similarities between Puerto Rican Spanish and Brazilian Portuguese with respect to this constraint and present a detailed analysis of the Brazilian Portuguese data, highlighting the regularities. Second, I discuss the differences of preceding markers across linear position in both languages and reanalyze Guy's (1981a:172) and Labov's (1994:562–564) hypothesis about missing zeroes in Brazilian Portuguese. Finally, I claim that this constraint is further evidence of the phrase-level parallelism embedded in a universal principle of linguistic use: parallel processing (see Scherre, 1988, 1998b; Scherre & Naro, 1991).

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

Our sample consisted of 64 hours of speech with 64 speakers from Rio de Janeiro, recorded by the Programa de Estudos Sobre o Uso da Língua (PEUL) research project (cf. Silva & Scherre, 1996); the sample was stratified for gender, age, and years of schooling (excluding illiterates and university level). Examples of the noun phrases are presented in (1).[3] (In all examples, S indicates an overt plural marker, ∅ indicates a zero plural marker, and N represents a number.)

(1)	(SSS)	umaS pessoaS ricaS a people rich	'some rich people'
	(SS∅)	essaS estradaS nova∅ these roads new	'these new roads'
	(S∅∅)	aS porta∅ aberta∅ the doors open	'the open doors'
	(∅SS)	do∅ meuS paiS of+the my parents	'it belongs to my parents'
	(∅S∅)	na∅ minhaS coisa∅ in+the my things	'in my things'
	(SS)	meuS filhoS my children	'my children'
	(S∅)	unS amigo∅ a friends	'some friends'
	(NS)	dois anoS two years	'two years'
	(N∅)	sete ano∅ seven years	'seven years'
	(SSNS)	oS meuS quatro filhoS the my four children	'my four children'
	(NSS)	dois filhoS pequenoS two children little	'two little children'
	(NS∅)	três cavaloS lindo∅ three horses beautiful	'three beautiful horses'
	(N∅∅)	dois risco∅ verde∅ two lines green	'two green lines'

In the analysis presented here, each element of the noun phrase was treated as a separate datum for purposes of coding and statistical processing. A total of 13,225 nominal tokens were analyzed: 71% (9,382/13,225), including determiners, adjectives, and nouns, showed overt plural markers, and 29% (3,843/13,225) showed zero plural markers. The data were analyzed using VARBRUL, which calculated the relative weights (Pintzuk, 1988; Sankoff, 1988). Eight factor groups were considered in the analysis: five were linguistic (relative and linear position; phonic salience; preceding markers across linear position; stress; following phonological segment), and three were social (years of schooling, gender, and age). However, for purposes of this article, I only discuss the results of preceding markers across linear position—or, more precisely, the phrase-level parallelism effect.[4]

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 2. *Contribution of position of token in noun phrase string and presence of immediately preceding marker to the deletion of plural -s in Puerto Rican Spanish and Brazilian Portuguese*

	Position of Token in String			
	3		2	
	Puerto Rican Spanish	Brazilian Portuguese	Puerto Rican Spanish	Brazilian Portuguese
Marker(s) preceding token				
S∅_	.68	.95		
S_			.44	.48
∅S_/SS_	.40	.41		

SIMILARITIES AND REGULARITIES

To avoid inevitable misunderstandings when comparing analyses of different languages by different researchers at different times, I looked for categories that I believed to be maximally comparable. I codified each relevant token, noting its position in the noun phrase string as well as the type and position of preceding plural markers. In (2), I exemplify the factors that are directly comparable to those in Poplack's (1980a) analysis (see Table 1). (The analyzed nominal segment is shown in italics.)

(2)	Preceding Markers	Examples	
	S∅_	aS porta∅ *aberta*∅	'the open doors'
	S_	unS *amigo*∅	'some friends'
		aS *boa*s a*ções*	'the good actions'
	SS_	aS boaS *ações*	'the good actions'
	∅S_	do∅ meuS *pais*	'it belongs to my parents'

Similar to what Poplack found, the results for Brazilian Portuguese show that the S∅ preceding string leads to more zero plural markers than do the S and ∅S/SS preceding strings (see Table 2).[5] The S∅ string strongly favors following zero plural (.95), whereas the S (.48) and ∅S/SS (.41) strings disfavor following zero plural.

The Brazilian Portuguese results reveal that, for these factors, the effect of preceding markers across linear position is even more polarized. Because of the difference in probabilities between S∅_ and ∅S_/SS_ and between S_ and ∅S_/SS_ in Brazilian Portuguese and Puerto Rican Spanish, I reanalyzed ∅S_/SS_ as two factors in order to account for the behavior of ∅S_ and SS_ (see Table 3).[6]

As can be seen in Table 3, the difference between ∅S_ and SS_ factors is statistically significant at a p level of .05.[7] Hence, the non-mixed SS_ string (.38) disfavors zero plural markers more than does the mixed ∅S_ string (.62). In other words, similar preceding overt plural markers lead to overt plural markers more

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 3. *Contribution of position of token in noun phrase string and presence of immediately preceding marker to the deletion of plural -s in Brazilian Portuguese: Reanalysis of ∅S_ and SS_ strings as two factors*

	Position of Token in String	
	3	2
Marker(s) preceding token		
S∅_	.95	
S_		.48
∅S_	.62	
SS_	.38	

often than does an immediately preceding overt plural marker that is preceded by a zero plural.

Moreover, additional regularities occur in the Brazilian Portuguese data. The results presented here are based on a large run, with eleven factors in preceding markers across linear position, exemplified in (3). (The part of the string that represents the factor is shown in boldface.)

(3)	Factor Symbols of Preceding Markers	Examples	
a.	SS_	aS boaS *ações*	'the good actions'
b.	∅S_	do∅ meuS *pais*	'it belongs to my parents'
c.	NS_	**três** cavaloS *lindo*∅	'three beautiful horses'
d.	SN_	aS **duas** *perna*∅	'the two legs'
e.	S∅_	aS porta∅ *aberta*∅	'the open doors'
f.	N∅_	**dois** risco∅ *verde*∅	'two green lines'
g.	SN+S+prep_		
	higher plural semantic NP with overt plural marker	**milhare**S **de** *cabeças*	'thousands of heads'
	SN+prep+S_		
	higher plural semantic NP without overt plural marker followed by lower NP with overt plural preceding marker	um **grupo de** crianças*abandonadas*	'a group of children abandoned'
h.	SN+prep_		
	higher plural semantic NP without overt plural	uma **porção de** *carro*∅	'a lot of cars'
	SN+prep+∅_		
	higher plural semantic NP without overt plural marker followed by lower NP with zero plural preceding marker	uma **porção de** coisa∅ *interessante*∅	'a lot of interesting things'
i.	S_	unS *amigo*∅	'some friends'
		aS *boas ações*	'the good actions'
j.	N ending in *-s* consonant	**três** *capitulos*	'three chapters'
k.	N not ending in *-s* consonant	**nove** *anos*	'nine years'

Table 4 presents the results for these factors.[8]

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 4. *Results for rate of zero plural marker according to preceding markers across linear position in Brazilian Portuguese*

Factors	Frequency		Relative Weight
	N	%	
SS_	132/439	30%	**.35**
∅S_	45/95	47%	.60
NS_	11/18	61%	.65
SN_	124/261	48%	.58
S∅_	181/193	94%	**.93**
N∅_	30/32	94%	**.93**
SN+S+prep_or SN+prep+S_	6/19	32%	**.31**
SN+prep_ or SN+prep+∅_	173/256	68%	**.69**
S_	2,258/4,812	47%	.50
Number ending in *-s*	372/872	43%	.43
Number not ending in *-s*	381/911	42%	.39
Total	3,713/7,907	47%	

First of all, recall that the mixed ∅S_ string favors zero plural (.62) more than does the non-mixed SS_ string (.38). Similar effects can be found in other mixed preceding strings, such as cardinal number and overt plural marker (NS_: **três** cavaloS *lindo*∅ 'three beautiful horses') or overt plural marker and cardinal number (SN_: aS **duas** *perna*∅ 'the two legs'). The results in Table 4 show that the three mixed categories (∅S_, NS_, SN_) exhibit intermediate effects (.60, .65, and .58, respectively). They are between the poles of SS_ (.35) and S∅_ (.93).

I also measured the effect of number and zero plural marker (N∅_: **dois** risco∅ *verde*∅ 'two green lines'). The result confirms the strong effect of an immediately preceding zero plural (S∅_: aS porta∅ *aberta*∅ 'the open doors'): both factors favor a following zero plural (.93 and .93). As can be seen, this effect is stronger than the effect of similar preceding overt markers (.35). In other words, if there are two preceding overt markers it is possible to have a following overt plural marker (generating strings SSS) or a zero plural marker (generating strings SS∅), although an SSS string is the preferred one in the analyzed sample. But an immediately preceding zero plural leads almost categorically to a following zero plural (generating strings S∅∅). The S∅S string almost never occurs.

Factors (g) and (h) focus on different kinds of embedded noun phrases. Two results emerge. First, higher plural semantic NP with overt plural marker (SN+S+prep_: **milhare**S **de** *cabeças* 'thousands of heads') and higher plural semantic NP without overt plural marker followed by lower NP with overt plural preceding marker (SN+prep+S_: um **grupo de** criançaS *abandonada*s 'a group of children abandoned') disfavor a following zero plural marker (relative weight of .31). Second, higher plural semantic NP without overt plural

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

(SN+prep_: uma **porção de** *carro*∅ 'a lot of cars') and higher plural semantic NP without overt plural marker followed by lower NP with zero plural preceding marker (SN+prep+∅_: uma **porção de** coisa∅ *interessante*∅ 'a lot of interesting things') favor a following zero plural marker (relative weight of .69).

Although the embedded noun phrases are syntactically different from the other noun phrases I analyzed, the results are undoubtedly consistent with those already presented: a preceding zero plural marker leads to more zero plural markers and a preceding overt plural marker leads to more overt plural markers. Therefore, in terms of internal opposition, it is possible to demonstrate that factor (g) works like the SS_ factor, disfavoring zero plural (.31), and factor (h) works like the S∅_ or the N∅_ factor, favoring zero plural (.69). I suggest that the results involving higher plural semantic NP allow us to analyze structures such as *uma porção de carro* 'a lot of cars' or *uma porção de coisa interessante* as a kind of ∅∅∅ or ∅∅∅∅ string.

The Brazilian Portuguese results presented so far not only corroborate Poplack's analysis, but also are even more revealing. I reemphasize that these striking results can be shown only if one analyzes preceding markers across linear position, as demonstrated by Poplack (1980a).

Factors (i), (j), and (k) analyze the effect of the preceding marker in the first position of a two-element noun phrase. In Brazilian Portuguese, some cardinal numbers end with a nonmorphemic *-s*. I separated data with a number in the first position that ends with an *-s* (**trê**S *capítulos* 'three chapters') from data with a number in the first position that does not end with an *-s* (**nove** *anos* 'nine years'). This enabled me to verify whether the nonmorphemic preceding *-s* would present an effect similar to that of the morphemic preceding *-s*.[9]

As can be seen in the results in Table 4, cardinal numbers ending in *-s* present a relative weight of .43, whereas cardinal numbers not ending in *-s* present a relative weight of .39. Thus, the difference between the two kinds of cardinal numbers is very small (.04), and, furthermore, it is not statistically significant.[10] The results show that overt markers that lead to more markers must have a semantic load. In the case analyzed in this article, it must have a plural semantic load. Although Labov (1994:555, 568) assigned a mechanical aspect to this kind of effect, I conclude that the effect of preceding markers in Brazilian Portuguese—a manifestation of linguistic parallelism on the phrase level—is not a mechanical effect. However, linguistic parallelism on the phrase level is undoubtedly further "evidence that morphological and syntactic variation is controlled by a tendency to preserve parallel structures" (Labov, 1994:550).

I also found that a preceding cardinal number slightly disfavors a following zero plural marker more than a preceding overt plural marker (.43 and .39 vs. .50). The differences between .43 and .50 (.07) and between .39 and .50 (.11) are statistically significant. With regard to linguistic parallelism, I consider that this fact is not adequately interpretable. If numbers as a whole do not exhibit overt plural markers, we would expect them to favor more zeroes. Future research could solve this unexpected effect.[11]

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 5. *Contribution of position of token in noun phrase string and presence of immediately preceding marker to the deletion of plural -s in Puerto Rican Spanish*

	Position of Token in String			
	3		2	
	Puerto Rican Spanish	Brazilian Portuguese	Puerto Rican Spanish	Brazilian Portuguese
Marker(s) preceding token				
∅∅_	.73	—		
∅_			.52	0%

Source: Poplack (1980a:63).

DIFFERENCES

The first difference between Brazilian Portuguese and Puerto Rican Spanish is the effect of two preceding zero plurals (∅∅_) on the following element. In Puerto Rican Spanish, a preceding zero plural strongly favors another zero plural (relative weight of .73), generating a ∅∅∅ plural string (see Table 5). As Poplack (1980a:64) noted, "the most favorable context for marker deletion is precisely when the two preceding markers have already been deleted." She claimed that "almost 1,000 examples of NP strings containing two or three elements, with no morphological marker at all" (p. 63) were found in the analyzed corpus.

As far as I know, a noun phrase with a referential plural meaning, such as *la*∅ *cosa*∅ *bonita*∅ 'the beautiful things', does not occur in Brazilian Portuguese. However, it is interesting to focus on structures such as those shown in (4).

(4) a. *menina bonita não chora*
'all beautiful girls never cry' or 'beautiful girl never cries'
b. *eu gosto muito de com arroz com feijão*
'I like rice and beans so much'
c. *eu comprei muito livro caro*
'I bought many expensive books'
d. *tinha roupa suja espalhada no quarto*
'there were dirty clothes scattered all over the room'
e. *a coisa tá feia*
'a lot of things are ugly' or 'the situation is so ugly'

These structures indicate that, in Brazilian Portuguese, it is possible to express the notion of plurality by generic singular clauses or by other syntactically singular words in the clause, such as the noun *feijão* 'beans', the modifier *muito* 'many', 'the adjective *espalhada* 'scattered', or the noun phrase *a coisa* 'a lot of things' or 'the situation'. However, these cases are not always seen as examples of plural *-s* deletion. Pereira (1984) interpreted them as not having overt plural

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

markers in their abstract representation. Thus, I suggest that, with respect to the effect of the preceding ∅∅ string, the difference between the two languages is apparent. In my view, Poplack and I have analyzed different sorts of data (see Poplack, 1980a:58–59, for details). Certainly an analysis that considers this kind of data in Brazilian Portuguese would find results similar to those found for Puerto Rican Spanish.

If the independent variable were changed from plural number agreement to the notion of the plural, ∅∅∅ strings (*muito livro caro* 'many expensive books'; *roupa suja espalhada* 'dirty clothing scattered') would be found in Brazilian Portuguese. These ∅∅∅ are supported zeroes, as Labov (1994:589) termed them.

> . . . . An element of linguistic structure is *supported* if it is accompanied in the stream of speech by other sources of the information that it carries.
>
> This support may consist of morphological, lexical, syntactic, pragmatic, or cultural information, just as in the analyses of Poplack (1980, 1981). A structural element is *supported* in a particular sentence if it is accompanied by any of this information; otherwise, it is *unsupported*. The term *unsupported zero* indicates the total absence of an inflection or any other mark of plural.

Similar to Puerto Rican Spanish (Poplack, 1980a:65), the ∅∅S string is "virtually nonexistent" in Brazilian Portuguese. Even the S∅S string is infrequent (12/727 = 1.7%). The order of preference is SSS (307/727 = 42%), S∅∅ (181/727 = 25%), and SS∅ (132/727 = 18%) strings. Although not frequent, there are also ∅SS (50/727 = 6.9%) and ∅S∅ (45/727 = 6.2%) strings, although the ∅∅S string is virtually nonexistent (for details about preferred structures, see Scherre, 1988:189–193).

In Brazilian Portuguese, there are other supported zeroes (∅∅∅ or ∅∅ strings) that have not been analyzed yet. They can carry specific plural meaning, expressing inherent possession in singular or plural syntactic clauses, such as those shown in (5).

(5) a. *Vou fazer a minha unha hoje / vou fazer a*S *minha*S *unha*S *hoje*
'I will have my nails done today'.

b. *Paula, por favor, lava a mão, o pé escova o dente / Paula, por favor, lava a*S *mão*S, *o*S *pé*S *e escova o*S *dente*S
Paula, please wash your hand/hands, your foot/feet, and clean your teeth'.

In (5a), it is implied that all ten finger nails will be done and perhaps all ten toe nails as well. In (5b), it could mean that only one hand or one foot will be washed, but it is clear that all teeth will be cleaned.

The only supported zeroes that I analyzed quantitatively were those that were embedded (*uma porção de coisa*∅ *interessante*∅ 'a lot of interesting things'). This structure could be interpreted as ∅∅, preceded by a ∅∅ string in the higher semantic noun phrase (*uma porção*). Although the ∅∅+prep+∅∅ string is slightly preferred, SS strings preceded by this type of ∅∅ string are also natural and are, indeed, prescribed by grammarians (*um grupo de criança*S *abandonada*S 'a group of children abandoned'). Furthermore, S∅ strings

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 6. *Apparent functional effects in plural marking of Brazilian Portuguese noun phrases: Effect of the preceding marker in the first position of noun phrase on the second element of the noun phrase*

	∅	-*s*	Total	% Deletion
Guy's data (Labov, 1994:562, Table 19.9)				
-*s* inflection preceding (S_)	2,046	683	2,729	75
No -*s* preceding (∅_)	7	63	70	10
Our data				
-*s* inflection preceding (S_)	2,258	2,254	4,812	47
No -*s* preceding (∅_)	0	126	126	0

preceded by this type of ∅∅ string can also occur (*um monte de filho*S *pequeno*∅ 'a lot of small children'). But ∅S strings preceded by this type of ∅∅ string are improbable in vernacular Brazilian Portuguese (**uma porção de coisa*∅ *interessante*S 'a lot of interesting things'). The ∅∅+prep+∅∅ string is permitted because there are singular zeroes in the higher semantic noun phrases. The ∅∅+prep+SS and ∅∅+prep+S∅ strings are permitted because the semantic singular noun phrases are in the higher position. The ∅∅+prep+∅S string is virtually nonexistent because the immediately preceding zero is a zero plural. The examples given here reflect the preference for (∅)∅∅∅ but not (∅)∅∅S strings, as Poplack (1980a:65) claimed.

The second difference between the two languages is the effect of a preceding zero (∅_), which categorically disfavors a zero plural marker in Brazilian Portuguese. This difference seems to be real. In the PEUL data, 126 cases had a zero plural marker in the first position and an overt plural marker in the second position of the noun phrase. In Brazilian Portuguese, then, a preceding zero in the first position categorically favors a following overt plural marker—an effect that could be interpreted as functional. Hence, the categorical effect of ∅_ seems to contradict the results already presented. Recall that, in Puerto Rican Spanish, this factor contributes to deletion with a relative weight of .52, whereas SS_ and ∅S_ contribute to deletion with a relative weight of .40 (cf. the results in Tables 2 and 5).

Discussing the functional effect of no preceding -*s* in the first position of the noun phrase (∅_ in my analysis), Labov (1994:561–568) claimed that it is an apparent functional effect. Following Guy's hypothesis (1981a:172), he asserted that this kind of context could involve unsupported zero: "These are the crucial cases noted by Guy (1981) in the discussion of Rio de Janeiro Portuguese . . . : 58 zeroes that had escaped the notice of the analyst" (Labov, 1994:589).

Comparing the most relevant part of Labov's (1994:562, Table 19.9) to the data that I analyzed, an interesting distribution may be seen (see Table 6).[12] The categorical effect of no preceding -*s* emphasizes the functional effect of a preceding zero marker in the first position. As I mentioned before, this result seems to contradict all the results previously presented: that is, the tendency of a preceding zero plural to favor a following zero plural.

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

TABLE 7. *Effect of the preceding marker in the first position of noun phrase on the second element of the noun phrase: Three-element noun phrases*

	∅	*-s*	Total	% Deletion
Three-element noun phrases				
No *-s* preceding (∅_)	0	95	95	0
-s inflection preceding (S_)	4	58	62	6
Other noun phrases[a]				
-s inflection preceding (S_)	2,254	2,196	4,750	47

[a]See totals of preceding *-s* in Table 6: zero plural variant, 2,258; overt plural variant, 2,254; Total, 4,812; frequency of zero variant, 47%.

As did Guy (1981a), Labov (1994) believed that some missing zero plurals without morphological support would be interpreted as singular, especially in two-element noun phrases. Using the frequency of zeroes in the first position, Labov (1994:564) recalculated the supposed missing zeroes of the second position and reaffirmed that there is an apparent functional effect.

I agree with Guy (1981a) and Labov (1994) that there is an apparent functional effect on the behavior of the no preceding *-s* factor (or ∅_). However, I am not sure if this fact is due to missing zero plurals. I verified the following facts about noun phrases with zero plural marker in the first position in the PEUL data. First, 75% (95/126) of the tokens with zero plural marker in the first position were indeed in three-element noun phrases, predominately with the definite article + possessive pronoun + noun configuration (92/95 = 97%): for instance, *na casa do*∅ *meu*S *pai*S/*pai*∅ 'in my parent's houses'; *gosto de ter a*∅ *minha*S *coisa*S/*coisa*∅ 'I like to have my own things'; *você bate no*∅ *teus filho*S/*filho*∅? 'Do you beat your children?'; *cê num paga o*∅ *teus imposto*S/*imposto*∅? 'Don't you pay your taxes?'. Second, 25% (31/126) of the unmarked tokens were in two-element noun phrases without any specific syntactic configuration: for instance, *aquela*∅ *azeitona*S 'those olives'; *temo filho*∅ *casado*S 'we have married children'.

Only the two-element noun phrase is a candidate for nonmorphological support if the second position occurs with a zero plural (i.e., only this kind of noun phrase is a candidate for unsupported zeroes). However, the two-element noun phrase does not seem to present a clear, regular pattern. Only noun + adjective (*filho*∅ *casado*S 'married children') and adjective + noun (*grande*∅ *viagen*S 'big trips') configurations have some regular influence on this sort of variation. Future studies should address this question.

As I have demonstrated, three-element noun phrases are very regular. Besides, in this kind of data, the categorical plural marker in the second position is surely not related to the zero plural marker in the first position of the noun phrase. I carefully examined all three-element noun phrases whose syntactic configuration does not have a zero plural in the first position, and I found the distribution presented in Table 7 (see also Scherre, 1988:181).

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

As displayed in Table 7, three-element noun phrases with a definite article + possessive pronoun + noun configuration show a high presence of markers, almost independently of the zero plural marker in the first position. Even if there is an overt plural marker in the first position, the frequency of zero plural in the second position is only 6%, a number closer to 0% than to 47% (preceding *-s* in other noun phrases). Scherre (1988, 1998a) examined another constraint—relative and linear position—that explains the high probability of overt plural markers in the second position in three-element noun phrases. The results from relative and linear position clearly demonstrate that elements to the left of the head exhibit more overt plural markers than do elements to the right of the head, almost independently of their grammatical class and their linear position from the left (Scherre, 1988, 1998a, in press). Moreover, zero plural marker in the first position may also occur in the written language, as the following dialogue transcription from a comic magazine illustrates:[13] *Tem tanta terra embaixo da*∅ *tua*S *unha*∅ *que dava pra começá a reforma agrária* 'there is so much soil under your nails that it would be possible to start the agrarian reform'.

The special behavior of the definite article + possessive pronoun + noun structure could be related to the contraction of the article with the preposition in the prepositional phrase (cf. Silva, 1982): for example, *do* (*de* + *o*) 'of the', *no* (*em* + *o*) 'in the', *pro* (*para* + *o*) 'for the'. Silva (1982) believed that the contraction effaces analysis and perception of the article as a distinct category. But yet it is not clear. These data occur with the prepositional phrase almost 75% of the time, but similar data for the overt plural marker in the first position (*na casa do*S *meu*S *pai*S/*pai*∅ 'in my parent's houses') occur with the prepositional phrase close to 65%. Besides, there is a zero plural marker with non-contracted prepositional phrases (*para o*∅ *meu*S *filho*S/*filho*∅ 'to my children') or without prepositional phrases (*gosto de ter a*∅ *minha*S *coisa*S/*coisa*∅ 'I like to have my own things').

Furthermore, the ∅S_ pattern bears no clear relation to the following nasal in the possessive of the first person *meu*/*nosso* 'my/our', as Guy (1981a:176) supposed. As I have shown, these types of clauses occur with the possessive pronoun *teu* 'your', whose first phone [t] is a non-nasal consonant, even in informal written language. In addition, 36% (11/31) of the first nominal constituent in two-element noun phrases occur in front of vowels, 36% (11/31) in front of voiced consonants (3 of which are nasals), and 29% (9/31) in front of voiceless consonants. This distribution does not present a clear phonological pattern, especially if one considers the effect of following vowels (in which case less deletion is expected).

Now, I am sure that the reason why "noun phrase agreement never skips over any words" is not because of the "phonetic process of S-deletion," as Guy (1981a:176) claimed. The 14 three-element noun phrases (S∅S or N∅S) do not present a clear phonological pattern. The nominal elements with zero plural marker show the following distribution: 50% (7/14) occur before voiced consonants (5 of which are nasals), but 43% (6/14) occur before voiceless consonants, and 7% (1/14) occur before a vowel.

The major regularities with regard to these cases involve intervening segments and phonic salience in the plural/singular opposition. Six cases (36%) present an intervening segment, such as *assim* 'like this', *bem* 'very', or *muito* 'so much',

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

between ∅S markers (*as coisa∅ muito uniforme*S 'things that are very similar'). Six cases without intervening segments that have an overt plural marker (67%) are more salient, such as *carnaval/carnavais* 'carnival/carnivals' (*esses dois último∅ carnavai*S 'these two last carnivals'). It is well known that this sort of opposition favors an overt plural marker in all phenomena of number agreement in Brazilian Portuguese (see, e.g., Naro, 1981:74–79, 94–95; Naro & Scherre, 1996:7; Scherre, 1989, in press).[14]

To summarize, I have demonstrated the following points. First, the difference between Brazilian Portuguese and Puerto Rican Spanish with regard to two preceding zeroes is more supposed than real. I suggested that there are differences in the kind of analyzed data and not in the type of observed effect. And more than that, I showed that some of the analyzed data are supported zeroes, in which the support element is either lexical or cultural information (Labov, 1994:560–661, 589). Second, the difference between Brazilian Portuguese and Puerto Rican Spanish with regard to a preceding zero in the first position in the noun phrase is more real than supposed. In this case, there are data specific to Brazilian Portuguese. I found that 75% (95/126) of the tokens are in fact three-element noun phrases, with a regular syntactic configuration, and undoubtedly they do not represent any functional effect. Therefore, 25% (31/126) of the data remain for future analysis. These data are truly two-element noun phrases, and they could participate in the area where the probable unsupported missing zeroes are located. In Labov's terms, this type of data is crucial to postulating models of acquisition of variable constraints or models of linguistic stability and change.

CONCLUSION

With respect to preceding markers across linear position, I have pointed out that the similarities between Brazilian Portuguese and Puerto Rican Spanish are stronger than the differences. The supposed differences between the two languages "that mark plurality on the NP and VP in much the same way" (Poplack, 1980a:64) are less than the early research led us to believe.[15] Earlier studies on number agreement in Brazilian Portuguese (e.g., Braga, 1977; Braga & Scherre, 1976) analyzed the effect of preceding markers in a very simplified way. Hence, they did not uncover the major regularities presented by Poplack (1980a:63–64). Even Scherre (1978) and Guy (1981a), who analyzed the variable in more detail, did not demonstrate the regularities that I have shown. I also demonstrated that the supposed functional effect of a preceding zero plural marker in the first position of a noun phrase on the second element is apparent. However, this fact is not completely related to the missing zeroes, as supposed by Guy (1981b) and Labov (1994).

Finally, I claim that the constraint of preceding markers across linear position on noun phrase number agreement is a parallelism effect at phrase level, undoubtedly another manifestation of linguistic parallelism.[16] This effect is naturally embedded in a universal principle of linguistic use: parallel processing (see Scherre & Naro, 1991). Surely parallel processing reflects one of several ways in which

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

the human mind works. My analysis of the parallelism effect on phrase level demonstrates that parallel processing operates more efficiently when preceding markers are more similar. Specifically, if the immediately preceding markers are zeroes, this principle increases in efficiency.[17] As Labov claimed (1994:547–568), this effect does not reveal "tendencies to preserve information." Moreover, in view of the facts presented in this article, this tendency is clearly not mechanical. Hence, the real nature of parallel processing remains to be completely understood as well as its consequences for the organization of discourse (see Scherre, 1998b; Tannen, 1989) and for the mechanism of linguistic change (see Labov, 1994:567–605; Naro, 1996).

NOTES

1. Factors ∅S, S, SS, and S∅ mean that the analyzed token is preceded, respectively, by zero plural marker and overt plural marker; one overt plural marker; two overt plural markers; overt plural marker and zero plural marker.

2. In Brazilian Portuguese, preceding markers were analyzed in a very simplified way by Braga and Scherre (1976) and Braga (1977). More detailed analyses were developed by Scherre (1978) and Guy (1981a). In my doctoral dissertation, completed in 1988, I analyzed new Brazilian Portuguese data. At that time, I discussed in detail the interaction between grammatical category and preceding markers on the one hand (Scherre, 1998a) and between phonic salience and stress on the other hand (Scherre, 1989). Recently (Scherre, in press), I presented a partial synthesis of these constraints.

3. The examples presented here are extracted from the discourse of the speaker, recorded as sociolinguistic interviews. Noun phrases are presented in isolation because examples within longer parts of speech are not relevant to this article. In Scherre and Naro (1992), there are parts of speech relatively longer as examples for noun phrase agreement, subject/predicate adjective agreement, and subject/verb agreement in Brazilian Portuguese, selected from the same data base.

4. Aiming at a better comparison with Poplack's (1980a) results, I submitted the data to three specific runs, testing alternative analyses for preceding markers across position. The first run involved three factors (Table 2), the second run involved four factors (Table 3), and the third run involved eleven factors—a complete run (Table 4). Differences among the results of the three runs were very small and meaningless. The other seven constraints were together in the three runs. Although the total number of tokens analyzed was 13,225, only 7,907 were used for the analysis presented in Table 4; 5,192 tokens in the first position of a noun phrase were analyzed as "doesn't apply" to preceding markers across linear position, and 126 tokens presented a categorical effect (see Pintzuk, 1988:3). In this article, the relative weights are shown in relation to the plural zero variant to facilitate the comparison between both analyses.

5. The results in Table 2 were obtained in a run containing eight factors groups, all of which were selected by VARBRUL using the 0.05 threshold:

number of cells: 1,084
corrected mean: .26
factor groups, in order of selection:
1. relative and linear position
2. phonic salience
3. years of schooling
4. phrase-level parallelism
5. gender
6. stress
7. age
8. following phonological environment

6. The results in Table 3 were obtained in a run containing eight factor groups, all of which were selected by VARBRUL using the 0.05 threshold:

number of cells: 1,116
corrected mean: 26
factor groups, in order of selection:

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

1. relative and linear position
2. phonic salience
3. years of schooling
4. phrase-level parallelism
5. gender
6. stress
7. age
8. following phonological environment

7. For details about tests of significance between two factors that belong to the same constraint, see Naro (1981:73) and Sankoff (1988:992).

8. The results in Table 4 were obtained in a run containing eight factors groups, all of which were selected by VARBRUL using the 0.05 threshold:

number of cells: 1,707
corrected mean: .30
factor groups, in order of selection:

1. relative and linear position
2. years of schooling
3. phonic salience
4. gender
5. following phonological environment
6. phrase-level parallelism
7. stress
8. age

9. Cardinal numbers ending in *-s* include *dois* 'two', *três* 'three', *seis* 'six', *dez* 'tem', some combinations (e.g., *dezesseis* 'sixteen', *vinte e seis* 'twenty six'), and all those that indicate more than one hundred (e.g., *duzentos* 'two hundred').

Cardinal numbers not ending in *-s* include *quatro* 'four', *cinco* 'five', *sete* 'seven', *oito* 'eight', *nove* 'nine', their combinations (e.g., *dezenove* 'nineteen', *vinte e nove* 'twenty nine'), and *onze* 'eleven', *doze* 'twelve', *treze* 'thirteen', *quatorze* 'fourteen', *quinze* 'fifteen'. Numbers that indicate more than ten and numbers that indicate a hundred and thousand also do not end in *-s*. For purposes of coding, I always considered the end of the last number.

10. I also tested all pairs of the eleven factors in Table 4 for statistical significance. Only one pair does not show a clear statistical difference. The difference between SS_ (.35) and S_ (.50) is sometimes statistically significant and sometimes not. It depends on the insertion of data, such as *coisas assim lindas* 'things that beautiful', that seem to have a special behavior. I am carrying out an analysis that takes the whole noun phrase as the unit of analysis in order to understand some facts that are still not well understood (see Scherre, 1988a:178–182). A final remark must be made: in its origin, the final *-s* in numbers such as *duzentos* 'two hundred' represents morphemic value. At one point in my research, I separated them as a specific factor. There were 40 tokens; 24 (60%) exhibited plural zero variant. The final relative weight was .50, the same as that assigned to S_. I did all tests of statistical significance, and all the relevant comparisons did not exhibit meaningful differences. Considering that there is no clear perception of the morphemic value of final *-s* in numbers such as *duzentos* 'two hundred', I provisionally decided to leave these 40 tokens together with the factor number ending in *-s* until new facts emerged.

11. To understand the effect of preceding cardinal numbers, it would be interesting to compare noun phrases with cardinal numbers, noun phrases embedded in a higher semantic noun phrase, and noun phrases with determiners with semantic plural load, such as *muitos meninos* 'many boys', *diversos meninos* 'diverse boys', *vários meninos* 'various boys' etc.

12. The factors labeled by Labov (1994:562–563) as "one-element noun phrases" and "second element of two-element noun phrases" are equivalent to "first position of noun phrases" and "second position of noun phrases," respectively (see Guy, 1981a:168).

13. I wish to express my gratitude to my friend Ana Zilles, who kindly sent me a copy of this magazine.

14. Heloísa Salles and Gregory Guy suggested that I look for an explanation for these facts in the generative framework. I am currently pursuing this line of study.

15. Concerning the comparison between Puerto Rican Spanish and Brazilian Portuguese, two other dependent variables show major linguistic patterns on noun phrase and subject/verb phrase agreement. They are relative position and phonic salience (see Labov, 1994:560; Poplack, 1980b:380–381). Anthony J. Naro and I are currently comparing these factor groups.

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

16. Linguistic parallelism is a constraint that works on several levels: discourse, clause, phrase, and word level. For details on discourse level, see Sankoff and Laberge (1978), Schiffrin (1981), Weiner and Labov (1983), Scherre and Naro (1991), Mollica (1991), Labov (1994). For details on clause level, see Lefebvre (1981) and Scherre and Naro (1991). The present work deals with phrase-level parallelism. For details on word level, see Scherre et al. (1998). For details about a large number of studies focusing on linguistic parallelism in Brazilian Portuguese, see Scherre (1988:378–425, 1998b).
17. Similar behavior of zero plural markers can be seen for subject/verb agreement and subject/predicate adjective agreement in Scherre and Naro (1991).

REFERENCES

Braga, Maria Luiza. (1977). *A concordância de número no sintagma nominal no triângulo mineiro*. Master's dissertation, Rio de Janeiro, Pontifícia Universidade Católica, Departamento de Letras e Artes.

Braga, Maria Luiza, & Scherre, Maria Marta Pereira (1976). A concordância de número no SN na área urbana do Rio de Janeiro. In *Encontro nacional de lingüistica*. Rio de Janeiro: Pontifícia Universidade Católica. 464–477.

Guy, Gregory Riordan. (1981a). *Linguistic variation in Brazilian Portuguese: Aspects of the phonology, syntax, and language history*. Doctoral dissertation, University of Pennsylvania.

_____ (1981b). Variability in American dialects of Spanish and Portuguese. In D. Sankoff & H. J. Cedergren (eds.), *Variation omnibus*. Edmonton: Linguistic Research. 85–93.

Lefebvre, Claire. (1981). Variation in plural marker: The case of Cuzco quechua. In D. Sankoff & H. J. Cedergren (eds.), *Variation omnibus*. Edmonton: Linguistic Research. 73–84.

Labov, William. (1994). The functional character of change. In *Principles of linguistic change: Internal factors*. Cambridge: Blackwell. 545–599.

Mollica, Maria Cecília de Magalhães. (1991). Processing and morpho-semantic effects in complementation in Brazilian Portuguese. *Language Variation and Change* 3:265–274.

Naro, Anthony. (1981). The social and structural dimensions of a syntactic language change. *Language* 57:63–98.

_____ (1996). *The functional nature of variation.* Paper presented at NWAVE-XXV, Las Vegas.

Naro, Anthony J., & Scherre, Maria Marta Pereira. (1996). Disfluencies in the analysis of speech data. *Language Variation and Change* 8:1–12.

Pereira, Maria Ângela Botelho. (1984). *Gênero e número em português. Estudos das relações forma-sentido na gramática*. Master's dissertation, Universidade Federal do Rio de Janeiro.

Pintzuk, Susan. (1988). VARBRUL programs. Unpublished manuscript.

Poplack, Shana. (1980a). The notion of the plural in Puerto Rican Spanish: Competing constraints on /s/ deletion. In Labov William (ed.), *Locating language in time and space*. Philadelphia: University of Pennsylvania. 55–67.

_____ (1980b). Deletion and disambiguation in Puerto Rican Spanish. *Language* 56:371–385.

Sankoff, David, & Laberge, Suzanne. (1978). Statistical dependence among successive occurrences of a variation in discourse. In D. Sankoff (ed.), *Linguistic variation: Models and methods*. New York: Academic Press. 119–126.

Sankoff, David. (1988). Variable rules. In U. Ammon, N. Dittmar, & K. Mattheier (eds.), *Sociolinguistics: An international handbook of the science of language and society*. Berlin: Walter de Gruyter. 984–998.

Scherre, Maria Marta Pereira. (1978). *A regra de concordância de número no sintagma nominal em português*. Master's dissertation, Rio de Janeiro, Pontifícia Universidade Católica, Departamento de Letras e Artes.

_____ (1981). Variation de la regle d'accord du nombre dans le syntagme nominal en Portugais. In D. Sankoff & H. J. Cedergren (eds.), *Variation omnibus*. Edmonton: Linguistic Research. 125–133.

_____ (1988). *Reanálise da concordância nominal em português*. Doctoral dissertation, Universidade Federal do Rio de Janeiro.

_____ (1989). Sobre a saliência fônica na concordância nominal em português. In F. Tarallo (ed.), *Fotografias sociolingüisticas*. São Paulo, Pontes. 301–332.

_____ (1998a). Variação da concordância nominal no português do Brasil: Influência das variáveis posição, classe gramatical e marcas precedentes. In S. Große & K. Zimmermann (eds.), *"Substandard" e mundança no português do Brasil.* Frankfurt am Main: TFM.

_____ (1998b). Paralelismo lingüístico. *Estudos da Linguagem* 7:29–59.

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press

_____ (in press). *Major linguistic patterns in noun phrase agreement in Brazilian Portuguese*. Paper presented at Atas do IX Convegno Internazionale di Linguisti, Instituto Di Glottologia, Universitá Degli Studi, Milano.

Scherre, Maria Marta Pereira, & Naro, Anthony Julius. (1991). Marking in discourse: Birds of a feather. *Language Variation and Change* 3:23–32.

_____ (1992). The serial effect on internal and external variables. *Language Variation and Change* 4:1–13.

_____ (1998). Sobre a concordância de número no português falado do Brasil. In R. Giovanni (Ed.), *Dialettologia, geolinguistica, sociolinguistica*. Tübingen: Max Niemeyr Verlag. 509–523.

Scherre, Maria Marta Pereira et al. (1998). Phonic parallelism: evidence from the imperative in Brazilian Portuguese. In C. Paradis et al. (eds.), *Papers in sociolinguistiocs*. Québec: Nota Bene. 63–72.

Schiffrin, Deborah. (1981). Tense variation in narrative. *Language* 57:5–62.

Silva, Giselle Machline de Oliveira. (1982). *Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro*. Doctoral dissertation. Universidade Federal do Rio de Janeiro.

Silva, Giselle Machline de Oliveira, & Scherre, Maria Marta Pereira (eds.). (1996). *Padrões sociolingüísticos: Análise de fenômenos variáveis do português falado no Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro.

Tannen, Deborah. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Weiner, E., Judith, & Labov, William. (1983). Constraints on the agentless passive. *Journal of Linguistics* 19:29–58.

https://doi.org/10.1017/S0954394501131042 Published online by Cambridge University Press