

Towards a non-linear account of plural marking in the ‘Caipira’ dialect of Brazilian Portuguese

Eduardo Rivail Ribeiro¹
University of Chicago

1 Introduction.

Standard Brazilian Portuguese (henceforth SP), as other Romance languages, presents gender and plural agreement between the head of a noun phrase and its determiners and modifiers. In addition, the verb agrees with the subject both in person and number. In colloquial Brazilian Portuguese, however, there is a strong tendency towards morphological simplification, both through the elimination of plural markers in the noun phrase and the reduction of the inflectional paradigm of verbs. This is shown by examples (1b) and (2b) below, which contrast sharply with their Standard Brazilian Portuguese versions given in (1a) and (2a):

- (1) a. *A-s menina-s bonita-s chegaram.*
the-PL girl-PL beautiful-PL arrive.past.3P
‘The beautiful girls arrived.’
b. *A-s menina bonita chegou.*
the-PL girl beautiful arrive.past.3S
‘The beautiful girls arrived.’
- (2) a. *Que menina-s bonita-s!*
whatgirl-PL beautiful-PL
‘What beautiful girls!’
b. *Que menina-s bonita!*
whatgirl-PL beautiful
‘What beautiful girls!’

Thus, as examples (1b) and (2b) show, plural tends to be marked only in the first elements of the NP in colloquial varieties of Brazilian Portuguese. Although this tendency is fairly common among speakers of dialects from different regions, it is probably in the ‘Caipira’ Portuguese (CP)² that it takes its

¹ I would like to thank Mônica Veloso Borges, from the Universidade Federal de Goiás (Brazil), for her help in confirming my native speaker’s intuitions, both with her field notes and her own native speaker’s intuitions, and Neha Dave, for editing support. This paper has also benefited from discussions with Jerry Sadock and Salikoko Mufwene. Any remaining shortcomings are, naturally, my entire responsibility. I would also like to thank the Wenner-Gren Foundation for Anthropological Research and the Brazilian National Scientific Development Council (CNPq, Grant 200018/98-1), for their financial support.

² The term ‘Caipira’ is generally used to refer to rural populations of the states of Goiás, Minas Gerais, Mato Grosso do Sul, São Paulo, and Paraná. Historically, these regions were colonized by *bandeirantes* (‘explorers’) from São Paulo (as in contrast with the coastal cities, founded by the Portuguese). Although the general characteristics of Caipira populations in the different states

most extreme form. In this dialect, the plural marker *-s* always attaches to the first element of the NP, regardless the part of speech to which this element belongs—an atypical behavior for an affix. Thus, as shown by examples (3) and (4) below, even the exclamative or interrogative ‘pronoun’ *que* and the interjections *ô* and *ê*, which are invariable words in Standard Portuguese and in the other non-standard dialects, take the plural marker in the Caipira dialect:³

may vary to a considerable degree, there are a number of shared cultural and linguistic features that justify the unifying use of the term. Among these features is the retroflex pronunciation of syllable-final *r* (generally referred to as ‘Caipira *r*’), which in the other dialects is pronounced as a glottal fricative [h] or an alveolar flap [ɾ]:

SP *carta* [ˈkahta] ‘letter’

CP [ˈkaɾta]

Another phonological peculiarity is the lack of the palatal liquid [ʎ], which becomes a yod in the Caipira pronunciation:

SP *palha* [ˈpaʎa] ‘straw’

CP *paia* [ˈpaia]

Furthermore, Caipira Portuguese lacks the so-called *proparoxítonas* words (that is, words whose stress falls on the antepenultimate syllable), which are systematically shortened:

SP *árvore* [ˈahvori] ‘tree’

CP *arve* [ˈaɾvi]

SP *córrego* [ˈkõhegõ] ‘stream’

CP *corgo* [ˈkõɾgõ]

The discussion in this paper is based on my own native dialect, spoken in Goiás, although data from different dialects will also be used when relevant for the present discussion. Additional data were obtained from Mônica Veloso Borges (personal communication) and Villefort’s (1985) monograph on the Caipira dialect as spoken in Morrinhos, Goiás. Villefort’s consultants were all selected among illiterate residents of the rural areas of the municipality, and therefore are less subject to the influence of Standard Portuguese than my other sources (Ms. Borges and myself, both from the region of Goiânia). Caipira Portuguese is commonly considered a less prestigious form of Portuguese, generally associated with uneducated, “outback” rural people. That is why Amadeu Amaral, considering particularly the variety spoken in São Paulo, predicted, in 1920, that this dialect would soon disappear, as a result of the democratization of the access to public education, the urbanization, and the arrival of European and Middle-Eastern immigrants. Instead, as a result of the formidable process of urbanization in most Brazilian regions in the last century, Caipira Portuguese is nowadays spoken in rather large cities, such as Goiânia (Goiás), São José do Rio Preto (São Paulo), and Uberlândia (Minas Gerais), among others.

³ I suggest that not only *que*, but also the exclamative and vocative interjections in such examples are part of the NP, occupying the syntactic position of a determiner, which explains the fact that they cannot co-occur with articles or demonstratives. Notice that the plural marker can only attach to the first element of the NP. It does not attach to prepositions, for example:

de que-s menina cês tão falando? ‘about which girls are you (PL) talking?’

**de-s que menina cês tão falando?*

- (3) a. *Que-s*⁴ *menina* *bonita!*
 what-PL girl beautiful
 ‘What beautiful girls!’
 b. *Que-s* *menina?*
 what-PL girl
 ‘Which girls?’
- (4) a. *Ô-s,* *menina* *bonita!*
 VOC-PL girl beautiful
 ‘Hey, beautiful girls!’
 b. *Ê-s* *trem tentado,* *sô!*
 EXCL-PL thing annoying sir
 ‘Man, what annoying things!’

Some of these linguistic peculiarities of the Caipira dialect were described as early as 1920. Amaral (1920, 52), focusing particularly on the variety spoken in the interior of São Paulo, describes the process of ‘s-dropping’ in the following terms:

Como sinal de pluralidade, desaparece: *os pau, os nó, os ermão, os papé, as frô(r), os urubú*. (...) Exceptuam-se os determinativos, que conservam o *s*: *uⁿas, arguⁿas, certos, muitos, êstes, duas, suas, minhas*, etc. assim como o pronome *êles, elas*. Quando pronominalizados, porém, os determinativos podem perder o *s*: *Estas carta não são as MINHA*.

[As a sign of plurality, [the -s] disappears: *os pau* ‘the woods,’ *os nó* ‘the knots,’ *os ermão* ‘the brothers,’ *os papé* ‘the papers,’ *as frô(r)* ‘the flowers,’ *os urubú* ‘the vultures.’ (...) The exceptions are the determiners, which preserve the *s*: *uⁿas* ‘a (plural),’ *arguⁿas* ‘a (plural), some,’ *certos* ‘certain (plural),’ *muitos* ‘many (plural),’ *êstes* ‘these,’ *duas* ‘two,’ *suas* ‘yours (plural),’ *minhas* ‘my (plural),’ etc. as well as the pronoun *êles, elas* ‘they.’ When pronominalized, however, the determiners may lose the *s*: *Estas carta não são as MINHA* ‘these letters are not mine.’]

This process cannot be explained in purely phonological terms, since the final fricative /s/ is generally preserved in cases in which it is not a plural marker:

⁴ I have noticed the occurrence of the plural marker with the interrogative and exclamative *que* and the vocative and exclamative interjections only in the variety spoken in Goiás (and part of Minas Gerais). However, Amaral (1920, 53) mentions at least one example which suggests that *que-s* must have occurred also in the variety of São Paulo. It is the case of the exclamative idiom *há que zano!* ‘what a long time!’ (SP *há que anos!*).

lápiz ‘pencil’, *nós* ‘we’, *paz* [pajs] ‘peace’, *ônibus* ‘bus’, etc.⁵ Apparently, it cannot be explained in purely morphological terms either, since the same word may or may not present the plural marker depending on its syntactic position. Thus, as Amaral observes, possessive pronouns take the plural marker when occurring in the beginning of an NP (3a), but not when in phrase-final position (3b):

- (5) a. [A-s] *minha-s* *camisa* *é* *azul*.
 the.FEM-PL my.FEM-PL shirt is blue
 ‘My shirts are blue.’
- b. *A-s* *camisa* *azul* *é* *a-s* *minha*.
 the.fem-PL shirt blue is the.FEM-PL my
 ‘The blue shirts are mine.’

In sum, the plural marker seems to appear only in the leftmost elements of the noun phrase—generally, the determiners. This suggests that the plural marker in the Caipira dialect behaves as a clitic—that is, a phrasal affix, whose position is syntactically, rather than morphologically determined. This would explain, among other things, the fact that it attaches even to words that are invariable in other dialects of Brazilian Portuguese (and probably in other Romance languages as well), such as interjections and the exclamative and interrogative *que*.

Therefore, this paper investigates the hypothesis that the plural marker *-s* in Caipira Portuguese acquired a certain syntactic independence, becoming a clitic—a process we may term ‘autolexicalization,’ after Sadock’s Autolexical Syntax theory (1991). In this framework, the various components of grammar—such as morphology, syntax and semantics, among others—are organized as autonomous subgrammars called modules. Since each of these modules is an autonomous grammar, “the organization of an expression in one module need not correspond to its organization in another” (Sadock *op. cit.*, 1). Thus, the usually problematic properties of clitics, for example, encounter a straightforward explanation in this framework: while clitics are affixes in the morphological component (thus combining with words), they combine with phrases in the syntactic component. As I intend to show, such an approach seems to provide a rather coherent account of the distribution of the plural marker in Caipira Portuguese.⁶

⁵ Further arguments against a ‘purely phonological’ analysis of the lack of plural marking come from consonant-final stems and some words ending in the nasal diphthong *-ão*, whose plural forms present additional modifications besides the mere suffixation of *-s*. Thus, while the plural forms of the words *anel* ‘ring’ and *leão* ‘lion’ in SP—*anéis* and *leões*, respectively—involve changes in the stem, their counterparts in Caipira Portuguese do not present any modification whatsoever (*os anel* ‘the rings’, *os leão* ‘the lions’). If the absence of plural marking in colloquial Brazilian Portuguese were the result of an *s*-deletion rule, one would have forms such as **os anéi* and **os leõe*, which are unattested.

⁶ The distribution of the plural marker would thus reflect the overwhelming tendency towards left-attachment also displayed by the pronominal clitics in most varieties of Brazilian Portuguese,

2. A second-position clitic?

The hypothesis to be investigated here is that the NP plural marker *-s* in the Caipira dialect of Brazilian Portuguese presents a behavior which resembles in several respects that of second-position clitics (Sadock 1991, 67). An example of second-position clitics is found in Amharic, “which has clitic determiners that agree with the head noun in gender but attach as suffixes to the first word of the NP that they determine” (Sadock *op. cit.*, 68):

- (6) a. *bet-u*
 house(M)-DET/M
 b. *tillik?u bet*
 big-DET/M house(M)
 c. *tillik?u k?iyy bet*
 big-det/M red house(M)

There is a noticeable parallel between the distribution of the clitic determiner in Amharic and that of the plural marker in Caipira Portuguese. Both always attach to the first element of the noun phrase, no matter what its part-of-speech is. One of the differences is that the plural marker in Caipira Portuguese may optionally spread to all pre-nominal elements. Thus, in the example (7) below, both the demonstrative pronoun *aquela* ‘that’ and the possessive *minha* ‘my’ take the plural marker:

- (7) *toda-s aquela-s minha-s roupa véia*
 all-PL that.FEM-PL my.FEM-PL clothes old.FEM
 ‘all those old clothes of mine’

Notice that possessives may also occur after the head noun. In this case, they do not inflect:

- (8) *toda-s aquela-s roupa véia minha*
 all-PL that-PL clothes old.FEM my.FEM
 ‘all those old clothes of mine’

where procliticization (*ele me deu* ‘he gave me’) is in fact the norm, encliticization (*ele deu-me* ‘he gave me’) being generally found only in written language or in formal, educated speech. This tendency towards left-marking is also revealed by the preference for ‘analytical,’ periphrastic future-tense constructions (*eu vou cantar* ‘I will sing’, *eu ia cantar* ‘I would sing’) over the ‘synthetic,’ morphological future-tense construction (*eu cantarei, eu cantaria*). The reanalysis of the plural marker as a second-position clitic (or its ‘autolexicalization’) may probably be seen as a strategy to resolve a ‘positional paradox’ presented by the language, where a general tendency towards inflectional suffixation co-exists with a general preference for leftward cliticization.

Preposed adjectives, although common in Standard Portuguese, are rare in Caipira Portuguese. For the cases in which they occur, they tend to follow the expected pattern—that is, they occur with the plural marker:⁷

- (9) *N-o-s* *meu-s* *bon-s* *tempo de* *menino...*
 in-the-PL my.MASC-PL good-PL time of boy
 ‘In my good times of childhood...’

Notice that the ‘adjective’ *outro* ‘other,’ which generally occurs preposed (10a), may lose its plural marker when occurring in noun phrase-final position (10b). Notice, however, that it takes the plural marker when not preceded by a determiner, as in (10c):

- (10) a. *Eu trouxe* *o-s* *outro-s* *trem.*
 I brought the-PL other-PL thing
 ‘I brought the other things.’
 b. *Eu trouxe* *o-s* *outro.*
 I brought the-PL other
 ‘I brought the others.’
 c. *Eu trouxe* *outro-s.*
 I brought other-pl
 ‘I brought others.’

The determiner-like adjectives *mesmo* ‘the same’ and *próprio* ‘proper, self’ follow the same pattern illustrated in (9) and (10) above. However, the ‘spreading’ of plural marking to other pre-nominal elements in cases such as (7)-(10a) above is clearly optional, as shown by a number of examples presented by Villefort (*op. cit.*: 36-38), in which plural marking is restricted to the first word of the NP: *un-s pouco tempo* ‘a short (period of) time’, *o-s outro irmão* ‘the other brothers’, *essa-s nossa história* ‘these stories of ours’. Thus, the data from this dialect strongly support the analysis of the plural marker as a second-position clitic in Caipira Portuguese.⁸

The examples we have considered until now suggest that plural marking is restricted to the prenominal position within the noun phrase. This tendency also occurs in other varieties of colloquial Portuguese, as described by Scherre (1988). The question, then, is what happens when the head is the first (or the only)

⁷ The same tendency is found in the few examples of preposed adjectives provided by Villefort: *nova-s idéia* ‘new ideas’, *boa-s coisa* ‘good things’, *nova-s administração* ‘new managements’ (Villefort, *op. cit.*: 48, 99).

⁸ The ‘spreading’ of plural marking to other pre-nominal elements of the NP constitutes a case to be further investigated. The analysis in this paper will focus on the data presented by Villefort (*op. cit.*), based on the assumption that they are less influenced by Standard Portuguese (see Section 3).

element in a noun phrase. Would it take the plural marker? That is clearly the case with pronouns, such as *(o)cê* ‘2nd person’ and *ele* ‘3rd person’:

- (11) *Ele-s* *chegou* *cedo.*
 3rd.MASC-PL arrive.3S early
 ‘They (masc.) arrived early.’

Noun-initial plural NPs are quite rare in colloquial varieties of Brazilian Portuguese. In Caipira Portuguese, plural NPs must generally occur with a determiner. ‘Generic,’ ‘universal’ NPs are generally singular, such as in (12) and (13a). In order to take the plural marker, they must occur with a determiner (13b):

- (12) *Home é* *tudo* *igual.*
 man is all equal
 ‘Men are all the same.’
- (13) a. *Brasileiro* *gosta de* *futebol.*
 Brazilian like.3S of soccer
 ‘Brazilians like soccer.’
- b. *O-s* *brasileiro* *gosta de* *futebol.*
 the-PL Brazilian like.3S of soccer
 ‘[The] Brazilians like soccer.’

One of the few circumstances in which plural nouns occur in initial position is in idioms and proverbs, such as in (14) below. In this case, the noun takes the plural marker, but the adjective does not:

- (14) *Água-s* *passada* *não* *roda* *moinho.*
 water-PL past not move.3S mill
 ‘Past waters do not move a watermill.’

Another circumstance in which plural nouns may occur is in advertisements, such as (15) and (16) below. My intuition is that, when ‘forced’ to pluralize a noun, in a more formal, ‘semi-learned’ speech register, the speaker tends to do it according to a pattern that favors the marking of the first element only. Constructions such as the ones below are commonly seen in hand-made signs in stores and informal commercial establishments:

- (15) [*Temos*] *roupas-s* *fina* *de* *marca-s* *famosa.*
 have.2p clothes-PL fine of brand-PL famous
 ‘[We have] fine clothes of famous brands’

- (16) Vende-se *doce-s* *caseiro*.
 sell.3s-REFL candy-PL home-made
 ‘Home-made candies sold [here].’

3. A note on variation.

As already thoroughly discussed (Scherre 1988; Scherre & Naro 1991, 1992; etc.), there is a great deal of variation in plural marking in Brazilian Portuguese. A very interesting aspect of such variation is what Scherre and Naro call “the serial effect,” a general principle according to which “birds of a feather flock together” (Scherre & Naro 1991): “marking leads to more marking and lack of marking leads to more lack of marking” (Scherre & Naro 1992, 1-2). This phenomenon is also noticed among speakers of Caipira Portuguese.

Considering particularly the case of Caipira Portuguese, I suspect that such variation is motivated by the competition of two dialects—on the one hand, the stigmatized Caipira Portuguese, and, on the other, a more neutral, less socially-marked dialect, such as the Standard Portuguese prevalent in the media and taught at school, where a great deal of effort is devoted to teaching the ‘correct’ plural forms. Although it is hard to demonstrate that variations such as the ones mentioned above are a matter of ‘code-switching,’ given the great similarity between both ‘codes,’ some pieces of evidence seem to point in the direction of a similar phenomenon involving the potential choice between two linguistic systems or sub-systems. One of them is the fact that standard plural forms tend to increase according to the degree of formal education of the speaker—that is, among speakers with a higher degree of ‘diglossia’ in Caipira and Standard Portuguese. That is why Villefort (*op. cit.*), whose consultants were all illiterate inhabitants of rural areas, does not cite one single example displaying full number agreement, both inside the NP and between a plural subject and its predicate.

Another piece of evidence is the fact that the use of a more standard agreement pattern seems to come hand-in-hand with the choice of less dialectal phonological processes or lexical selection. As Head (1989: 646-47) demonstrates, the use of the Caipira *r*—a strongly stigmatized characteristic of this dialect—tends to decrease among college-educated speakers, but is less subject to variation among speakers with only elementary education, a factor that apparently also underlie the distribution of standard plural forms. Furthermore, although a speaker of CP can use a strongly dialectal form such as *meus fi* ‘my sons’, a less socially-marked colloquial form *meus filho*, or the Standard Portuguese form *meus filhos*, I never came across a mixed form such as **meus fis*, which combines a strongly dialectal lexical form, *fi*, with a strongly standard inflectional pattern. This suggests that the ‘birds of a feather’ phenomenon can be seen as a matter of register or dialectal coherence.

Thus, the possible explanation I suggest for the variation in plural marking in Caipira Portuguese implies the notion of *non-monolithic grammar*, such as proposed by Mufwene (1992) and adopted by Labov (1998) under the name of

coexistent systems, in his analysis of variation in the tense-aspect system of African-American Vernacular English. Without appealing to the notion of code-switching, this approach suggests that variation may be a result of the competition between two or more systems coexisting in the same language, each containing its own set of rules that may eventually overlap. Among the factors that favor the recognition of co-existent systems is what Labov calls *the segregation of variants*: “[T]he variants of linguistic variables are not evenly distributed across texts or situations, but concentrated in long runs of the same value, so that extended stretches of speech show one value rather than the other.” This is exactly what seems to be at play in the “birds of a feather” phenomenon in Brazilian Portuguese.

4. Final remarks.

The examples analyzed here seem to suggest that the plural marker in Caipira Portuguese occurs preferentially in second-position in the NP. Since the first element in a noun phrase is generally the determiner, the plural marker will mostly attach to determiners. However, the plural marker may also occur with nouns, in the few circumstances in which these occur phrase-initially, as well as pronouns. Examples such as (11) and (14)-(15) demonstrate that the distribution of the plural marker cannot be explained in terms of constraints against its occurrence after the head of an NP. Furthermore, the occurrence of the plural marker phrase-finally, such as in example (10c) and (11), shows that there are no constraints against its occurrence in the rightmost position of the NP.

The examples presented above show that plural marking is obligatory in (and usually restricted to) the first word of an NP, regardless of its part-of-speech. Thus, a hypothesis strongly suggested by the data is that the plural marker *-s* is associated with the second position within the NP, a behavior quite similar to the one of second-position clitics. The adoption of an autolexical perspective provides a straightforward explanation for the peculiar distribution of the plural marker in Caipira Portuguese, which would be a result of its conflicting morphological and syntactic properties. As the second-position clitics *u* in Amharic and *que* in Latin (Sadock 1991, 70), the plural marker *-s* in Caipira Portuguese would be a phrase-initial suffix, such as suggested in the lexical entry in (17). As the autolexical representations in (18) demonstrate, the plural marker can be seen as a proclitic element which associates to the whole NP in the syntax, but attaches as a suffix to the first word of the NP in the morphology:

- (17) a. *-s* ‘plural marker’
 syntax = [_{NP}[__N̄]]
 morphology = [_wW_____]

- Sadock, Jerrold. 1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations*. Chicago: University of Chicago Press.
- Scherre, Maria M. Pereira. 1988. Reanálise da concordância nominal em português. Doctorate thesis. Rio de Janeiro: PUC.
- Scherre, Maria M. Pereira and Naro, Anthony. 1991. Marking in discourse: "Birds of a feather." *Language Variation and Change* 3:23-32.
- Scherre, Maria M. Pereira and Naro, Anthony. 1992. The serial effect on internal and external variables. *Language Variation and Change* 4:1-13.
- Villefort, Mariza T. Costa. 1985. *Aspectos Sintáticos do Dialeto Caipira na Região de Morrinhos*. Goiânia: Editora da UCG.