

On the Morphosyntax of Augmentatives in Brazilian Portuguese

Sobre a morfossintaxe dos aumentativos em português brasileiro

Paula Roberta Gabbai Armelin*

ABSTRACT

This paper investigates augmentative formation with *-ão* and *-zão* in Brazilian Portuguese. The empirical guideline is divided into two main lines: (i) the interaction between augmentative and gender, and (ii) the (im)possibility of a non-compositional interpretation being attributed to the augmentative structure. Assuming a syntactic approach to word formation (HALLE; MARANTZ, 1993; BORER, 2003), this paper proposes that *-ão* augmentative structure may either show one or two gender heads. When two gender heads are present in the structure, a mechanism of gender agreement (CHOMSKY, 2000, 2001) is assumed to be operating within words. Nevertheless, when only one gender head is present, no gender agreement is possible and the resulting augmentative gender is a default masculine. Based on Borer (2013), this paper proposes that elements within the first functional head of the Extended Projection are able to trigger non-compositional reading. This is exactly the case of the *-ão* formative when only one gender head is present. The differences in the behavior between *-ão* and *-zão* are derived from the fact that the gender head that attaches to the augmentative *-zão* is not optional, and while *-ão* attaches below a number head, *-zão*, on the other hand, attaches above a number head. Finally, the impossibility that *-zão* augmentative constructions be non-compositionally interpreted is derived from the fact that there is too much syntactic material intervening between the augmentative head and the root.

Keywords: *augmentatives; gender; compositionality.*

RESUMO

Este artigo investiga a formação de aumentativos com *-ão* e *-zão* no português brasileiro. A diretriz empírica é dividida em duas linhas principais: (i) a interação entre aumentativo e gênero e (ii) a (im)possibilidade de que uma interpretação não-composicional seja atribuída à formação aumentativa. Assumindo uma abordagem sintática de formação de palavras (HALLE; MARANTZ, 1993; BORER, 2003), este artigo propõe que a estrutura sintática do aumentativo pode apresentar uma ou duas projeções de gênero. Quando dois núcleos de gênero estão presentes na estrutura, um mecanismo de *Agree* (CHOMSKY, 2000, 2001) é assumido no interior da palavra. No entanto, quando apenas um núcleo de gênero está presente na estrutura, nenhum mecanismo de concordância de gênero é possível e, como consequência, o gênero do aumentativo resultante é *default*. Baseado em Borer (2013), este artigo propõe que elementos dentro do primeiro segmento funcional da Projeção Estendida são capazes de desencadear leitura não-composicional. Este é exatamente o caso do morfema *-ão* quando apenas um núcleo de gênero está presente. As diferenças de comportamento entre *-ão* e *-zão* derivam do fato de que o núcleo de gênero que se liga a *-zão* não é opcional. Além disso, enquanto *-ão* se anexa abaixo de uma projeção de número, *-zão*, por outro lado, anexa-se acima dela. Finalmente, a impossibilidade de que interpretação não-composicional seja atribuída às construções com *-zão* é consequência da presença de material sintático intervindo entre o núcleo que aloja o aumentativo e a raiz.

Palavras-chave: *aumentativos; gênero; composicionalidade.*

* Universidade Federal de Juiz de Fora (UFJF)

This work is inserted within the scope of the studies that investigate word formation, and it aims to analyze the morphosyntactic structure of augmentatives in Brazilian Portuguese (BP), built with the formatives *-ão* and *-zão*. More specifically, the empirical guideline of this work can be divided into two main lines: one that addresses the interaction between augmentatives and the notion of grammatical gender, and another one that addresses the possibilities and impossibilities of a non-compositional interpretation being attributed to the augmentative structure.

Empirically, the interaction between augmentative and gender is evidenced by the fact that a feminine noun in BP may present two different corresponding *-ão* augmentative forms: a feminine augmentative form, and also a masculine augmentative form. Interestingly, however, this is not a possibility when the base noun is masculine. In this case, only the masculine

¹ I would like to thank the anonymous reviewers for their helpful suggestions and comments that greatly contributed to the improvement of the final version of this paper.

augmentative corresponding form is licensed, being the feminine augmentative ungrammatical. The augmentative formations with *-zão*, on other hand, necessarily show the very same gender as the non-augmentative corresponding form, being the base noun either a feminine or a masculine noun.

In order to investigate the relation between augmentative and gender, it is necessary to review the status of notions like gender and noun class in the grammar, and the formal representation attributed to them. Following Armelin (2014, 2015) it is proposed that gender and noun class occupy the very same position in the syntactic structure. This position is identified as a gender projection, which is part of the extended projection of the noun. The augmentative makers are, then, analyzed based on the relations they establish with the syntactic gender head. The underlying hypothesis is that the relation established between the augmentatives formatives and the gender head reveal important aspects of the structural position that hosts them.

It is also an empirical fact that the *-ão* augmentative formative may trigger non-compositional interpretation. When this is the case, the augmentative forms are ambiguous between a compositional and a non-compositional reading. Importantly, however, the non-compositional augmentative, even when related to a feminine base form, results in a masculine augmentative formation. Quite independently of the gender of the formation, *-zão* augmentative constructions, on the other hand, do not trigger non-compositional interpretation.

In order to investigate the interaction between the augmentative formatives and (non)compositional interpretation, I assume a localist view of grammar, in which the licensing of non-compositional meaning must be conditioned by local domains of syntactic structure. In this sense, the possibility and impossibility of non-compositional interpretation being attributed to augmentative formations is derived from the syntactic positions assigned to each of the formatives.

The proposal developed in this work is implemented within a syntactic approach to word formation (cf. HALLE; MARANTZ, 1993; BORER, 2003). The syntactic structure proposed for the *-ão* augmentative formation shows two gender heads: one that attaches to the root, and another one that attaches directly to the augmentative. When the feature specification of this internal gender head is preserved in augmentative formation, I propose that a mechanism of gender agreement is operating within words. Basically, the assumptions made here are in line with the system proposed in Chomsky (2000, 2001), in which a probe-goal relation is established between an unvalued feature and a valued one of the same type, given that the appropriate structural c-command relation between them is matched. The unvalued feature probes in its c-command domain for a valued version of the relevant feature and, as a result of a successful probing, the involved heads share the same value of that feature. Nevertheless, I also propose that the gender head that attaches directly to the augmentative is, in

fact, optional. This optionality is justified by the fact that the *-ão* augmentative can clearly present its own gender head, which may or may not agree with that of the base. Based on the proposal put forth in Borer (2013), I propose that elements within the first functional head of the Extended Projection are able to trigger non-compositional reading. This is exactly the case of the *-ão* augmentatives in which the internal (optional) gender head is absent.

Concerning the *-zão* augmentative formative, I propose that the presence of two independent gender heads is even clearer, since they can be phonologically identified in the output form. Differences in the behavior of these constructions point to a split between *-ão* on one side, and *-zão* on the other. I propose that these differences are derived from the fact that: (i) the gender head that attaches to the augmentative *-zão* is not optional, and (ii) while *-ão* attaches below a number head, *-zão*, on the other hand, attaches above a number head. Finally, the impossibility that the *-zão* augmentative constructions be non-compositionally interpreted is derived from the fact that there is too much syntactic material intervening between the augmentative head and the root.

This paper is organized as it follows: section 2 describes the empirical relation between augmentative and gender; section 3 discusses a formal approach to gender and noun class; section 4, implements a syntactic structure to represent augmentatives in BP; section 5 describes the empirical relation between augmentative and non-compositionality; section 6 presents the state-of-art regarding the formal proposals available in the literature to deal with meaning computation; section 7 revisits the syntactic structures proposed in the paper in order to derive the possibilities and impossibilities that non-compositional reading be licensed in each case; finally, section 8 closes the paper with the final considerations.

10

1. Augmentative and Gender

The augmentative *-ão* and *-zão* may, in some contexts, alternate in augmentative formation in BP, as showed in (1)².

- | | | | |
|--------|-----------------|----------------------|------------------------|
| (1) a. | carro ('car') | carrão ('big car') | carrozão ('big car') |
| b. | mapa ('map') | mapão ('big map') | mapazão ('big map') |
| c. | dente ('tooth') | dentão ('big tooth') | dentezão ('big tooth') |
| d. | anel ('ring') | anelão ('big ring') | anelzão ('big ring') |

² The alternation between between *-ão/-zão* is not licensed in all contexts: in the so-called non-thematic nominals, for example, the presence of the consonant /z/ is often obligatory. For a detailed account of the status of the *-z* consonant see Armelin (2015) and Bachrach and Wagner (2007).

Empirically, the interaction between augmentative and gender is evidenced by the fact that a feminine noun in BP may present two different corresponding *-ão* augmentative forms: a feminine augmentative form, and also a masculine augmentative form³:

- | | | | |
|--------|--|--|---|
| (2) a. | a chuva
det(f) rain(f)
'the rain' | a chuvona
det(f) rain-aug(f)
'the big rain' | o chuvão
det(m) rain-aug(m)
'the big rain' |
| b. | a tigela
det(f) bowl(f)
'the bowl' | a tigelona
det(f) bowl-aug(f)
'the big bowl' | o tigelão
det(m) bowl-aug(m)
'the big bowl' |

Interestingly, however, this is not a possibility when the base noun is masculine. In this case, only the masculine augmentative corresponding form is licensed, being the feminine augmentative ungrammatical.

- | | | | |
|--------|---|--|--|
| (3) a. | o livro
det(m) book(m)
'the book' | *a livrona
det(f) book-aug(f)
'the big book' | o livrão
det(m) book-aug(m)
'the big book' |
| b. | o carro
det(m) car(m)
'the car' | *a carrona
det(f) car-aug(f)
'the big car' | o carrão
det(m) car-aug(m)
'the big car' |

Concerning the augmentative headed by the *-z* consonant, the gender of the augmentative form is always identical to the gender of corresponding non-augmentative form. In this sense, a feminine noun may only have a feminine corresponding *-zão* augmentative, while a masculine noun may only form a masculine augmentative.

- | | | | |
|--------|--|--|-------------------------------------|
| (4) a. | a chuva
det(f) rain(f)
the rain' | a chuvazona
det(f) rain-aug(f)
'the big rain' | *o chuvazão
det(m) rain-aug(m) ' |
| b. | a tigela
det(f) bowl(f)
'the bowl' | a tigelazona
det(f) bowl-aug(f)
'the big bowl' | *o tigelazão
det(m) bowl-aug(m) |

³ An anonymous reviewer pointed out that *colherzão* (spoon-aug(m)) would be a possible exception to this generalization. It is important to notice, however, that *colher* (spoon) is a non-thematic noun.

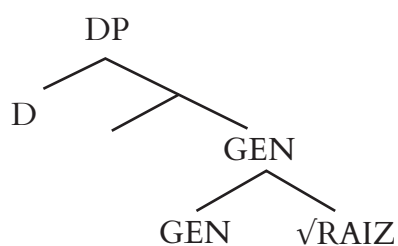
(5) a.	o livro	*a livrazona	o livrozão
	det(m) book(m)	det(f) book-aug(f)	det(m) book-aug(m)
	'the book'	'the big book'	'the big book'
b.	o carro	*a carrozona	o carrão
	det(m) car(m)	det(f) car-aug(f)	det(m) car-aug(m)
	'the car'	'the big car'	'the big car'

It is interesting to highlight two important contrastive empirical facts regarding the relation between augmentative and gender. The first fact is the contrast between (2) and (5), which shows that there are contexts in which *-ão* augmentative may present a gender value that differ from the gender of the non-augmentative form. This behavior, nevertheless, is not a possible one when it comes to the *-zão* augmentative formation. The second fact is the contrast between (2) and (3), which shows different possibilities of gender values for the augmentative, regarding the gender of the base noun.

2. Gender and Theme Vowel: a unified approach

In order to investigate the relation between the augmentatives and gender, it is necessary to review the status of notions like gender and noun class in the grammar, and the formal representation attributed to them. In this section, it is proposed that inflectional class is not a primitive in BP nominal system (see Armelin, 2014, 2015). More specifically, it is proposed that the split between gender and inflectional class is both unrevealing and misleading. In this sense, we unify both notions of gender and class by saying that they are phonological exponents of the gender head. The basic syntactic structure of our system is the following:

(6) The basic noun structure



The gender head is part of the extended projection (see GRIMSHAW, 1991) of the Noun and it is responsible by triggering agreement between the noun and its modifiers. We are assuming a theoretical view in which roots

do not have any internal grammatical value and are devoid of any syntactic feature (see BORER, 2013). This being so, roots cannot project or select for complements. In Bare Phrase Structures (see CHOMSKY, 1995) roots are at the same time minimal and maximal projections: they are not the projection of any head and they do not project further in the structure. The resulting label from the merge between the root and the Gender head has to be gender itself.

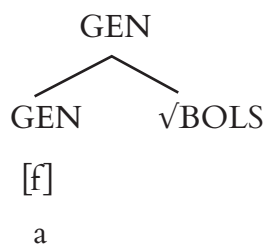
Returning to BP data, we have seen that the most general picture in the nominal system is the following: *-a* is the default feminine marker and *-o* is the default masculine marker. In this sense, most nouns ending in the vowel *-a* is feminine, while most nouns ending in the vowel *-o* is masculine. Let us suppose, then, that *-o* and *-a* are the default phonological exponents of the gender head. Assuming that the syntactic terminals are devoid of phonological content (Late Insertion), the system is very straightforward: if the gender head has a [feminine] feature in syntax, then phonological exponent *-a* is inserted, if the gender head has a [masculine] feature in syntax, then the phonological exponent *-o* is inserted.

(7) Default Vocabulary Items for the GEN head

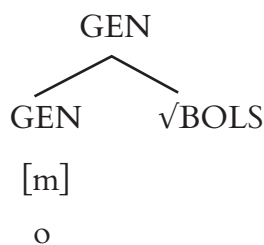
- a. $[f]_{\text{gender}} \leftrightarrow a$
- b. $[m]_{\text{gender}} \leftrightarrow o$

(8) Syntactic structure:

a. Output: bolsa ('purse')



b. Output: bolso ('pocket')



However, as we have seen earlier, the general pattern is not without exceptions. In fact, the existence of feminine nouns ended in *-o* or of masculine nouns ended in *-a* is one of the arguments for the traditional split between what is gender and what is theme vowel. These cases clearly deviate from the general pattern. To account for them, let us just emphasize the separation between what is syntactic (the agreement features) and what is phonological (the actual realization of the final vowel). We do that by assuming, as we did before, that the phonological content of syntactic heads is inserted late. It is clear that unexpected patterns, like masculine nouns in *-a* and feminine nouns in *-o* have to be listed somewhere. We propose, then, that the final

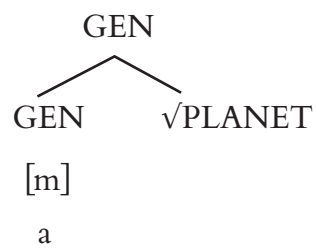
unexpected vowels are just a matter of phonology. This means that the non-default patterns are nothing more than “special” phonological exponents of the gender head. By special, we mean contextually specified with root information.

(9) Contextual specification in Vocabulary Items

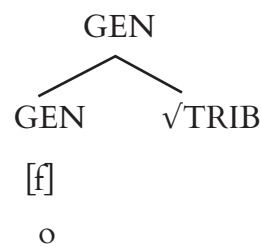
- a. $[m]_{\text{gender}} \leftrightarrow a/ \{ \sqrt{\text{PLANET}}; \sqrt{\text{MAP}}; \sqrt{\text{PROBLEM}}; \text{etc.} \}$
- b. $[f]_{\text{gender}} \leftrightarrow o/ \{ \sqrt{\text{TRIB}}; \sqrt{\text{LIBID}}; \sqrt{\text{VIRAG}}; \text{etc} \}$

(10) Syntactic structure:

a. Output: planeta (‘planet’)



b. Output: tribo (‘tribe’)



Vocabulary Items (VIs) compete for insertion into the terminal nodes generated by syntax. One of the important aspects of Vocabulary Insertion is that it respects the Subset Principle:

(11) The Subset Principle

The phonological exponent of a Vocabulary item is inserted into a morpheme in the terminal string if the item matches all or a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary item contains features not present in the morpheme. Where several Vocabulary items meet the conditions for insertion, the item matching the greatest number of features specified in the terminal morpheme must be chosen (HALLE, 1997, p. 428).

All things being equal, the exponent that realizes a Maximal Subset of the features on the relevant syntactic terminal node must win the competition. The consequence of this principle is that the existence of a more specific VI overrides the application of a more general one. Crucially, the Subset Principle

as stated does not allow us to conclude that the VIs in (9) are more specific than the VIs in (7), since both groups realizes the same number of features. However, contextual specification does have to be taken into account in order to compare for specificity. If that is true, it becomes clear that, in the context of the relevant roots, the VIs in (7) are more specific than the VIs in (9).

In this sense, in the context of the roots $\sqrt{\text{planet}}$ or $\sqrt{\text{map}}$, for example, the exponent $-a$, is the most specified one, while in the context of $\sqrt{\text{trib}}$ or $\sqrt{\text{libid}}$, on the other hand, the VI $-o$ is the winner of the competition. Crucially, in the absence of any relevant root for the contextual specification, the VIs in (9a) or (9b) lose the competition by virtue of being overspecified: they contain information that is not present in the terminal nodes. Consequently, the expected defaults are phonologically realized.

There is a valid generalization that has been noticed in the relation between gender and inflectional class in BP: whenever there is a mismatch between gender and class, the one that gets realized is class, and not gender. In a split approach for gender and inflection class, it is possible to account for that by saying that, somehow, class has hierarchical precedence over gender. If you have a masculine noun ended in $-a$, like *mapa*, for example, the final $-a$ is considered to be class and not gender, because class outranks gender. In our account, however, it is just the case that if there is a VI more specified for the relevant context, then, the default exponents cannot be inserted. This is no different from the following way to account for irregular plural formation in English:

(12) VIs specification for plural in English

- a. $[-\text{sing}]_{\text{num}} \leftrightarrow s$
- b. $[-\text{sing}]_{\text{num}} \leftrightarrow \emptyset / \{\sqrt{\text{SHEEP}}; \sqrt{\text{FISH}}; \text{etc}\}$
- c. $[-\text{sing}]_{\text{num}} \leftrightarrow \text{ren} / \{\sqrt{\text{CHILD}}; \sqrt{\text{BROTHER}}\}$

The VI in (12a) is the default one, while the ones in (b) and (c) have contextual specification about the roots to which they apply. In the context of $\sqrt{\text{SHEEP}}$, then, (12b) is the most specified, while in the context, of $\sqrt{\text{CHILD}}$, on the other hand, (12c) wins the competition. In the absence of contextually listed root, the default VI is going to be inserted. Crucially these are all different phonological realizations of the same syntactic head, that is, the number head. It is not at all necessary to stipulate beforehand that the relevant roots are marked with diacritics in order to ensure the right plural ending will match the right root.

In the same sense, it is not necessary to mark roots in BP just to ensure that they match the right phonological ending. It is very important to highlight that no relevant generalization is being lost by discarding class features as I, II and so on, once there is nothing but phonological exponence grouping nouns together.

3. A syntactic structure for the augmentative in BP

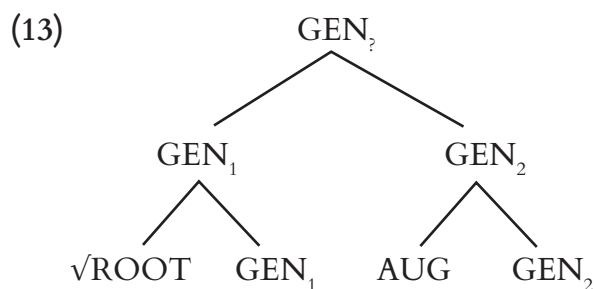
In this section we develop a syntactic structure for augmentatives in BP. More specifically, the augmentative makers are analyzed based on the relations they establish with the syntactic gender head proposed in the last section. The underlying hypothesis is that the relation established between the augmentative formatives and the gender head reveal important aspects of the structural position that hosts them. The main questions to be answered are the following:

- a) How to explain the cases in which the augmentative form and the corresponding non-augmentative form have the same gender specification?
- b) How to explain the cases in which the augmentative form and the corresponding non-augmentative form have different gender specifications?
- c) Why only feminine base nouns may show two different *-ão* augmentatives: a feminine one and a masculine one?
- d) Why the *-z* augmentative has to always preserve the gender of the base noun?

16

In order to deal with the question (a) above, we are going to assume that some sort of gender agreement mechanism is operating within words. We assume a view of Agree in which the very same feature is shared by two different heads. Basically, the assumptions made here are in line with the system proposed in Chomsky (2000, 2001), in which a probe-goal relation is established between an unvalued feature and a valued one of the same type, given that the appropriate structural c-command relation between them is matched. The unvalued feature probes in its c-command domain for a valued version of the relevant feature and, as a result of a successful probing the involved heads now share the same value of that feature.

If this rationale is on the right track, it is plausible to sketch a syntactic structure containing two gender heads: one that attaches to the root itself and one that attaches to the augmentative morpheme, as illustrated below:



It is important to note that the structure in (13) is able to account for some other empirical properties displayed by *-ão* augmentative formations. Regarding the interaction between the augmentative and derivational morphology, there seems to be an ordering restriction, as the augmentative may not appear inside derivational affixes.

- (14) Laranja ('orange') laranjão/ laranjona ('big orange')
 Laranjeira ('orange tree') laranjeirão/ laranjeirona ('big orange tree')
 *Laranjãozeiro/*Laranjeirona (intended meaning: a tree of big oranges)

The fact that derivational affixes appear inside, and not outside the augmentative marker is expected, since derivational affixes are, in general, very close to the root. These low derivational affixes are the ones that disrupt the relation between root and GEN, negotiating with the GEN head the feature and the phonological exponent this head should present.

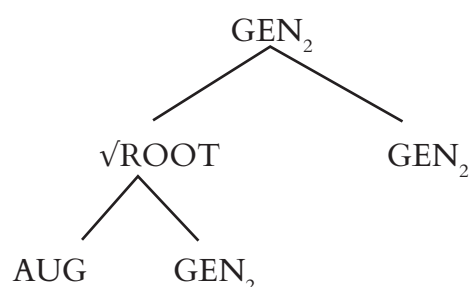
Also, the distance between the AUG head and the Root promptly captures the augmentative phonological properties. With the intervention of two GEN heads between them, it is expected that no allomorphic interaction could be licensed. Indeed there is no empirical evidence of phonological interaction between the augmentative and the root. More specifically, the augmentative fail to trigger phonological changes in the root even in contexts where this interaction is predicted to happen (see WETZELS, 1992; MORENO, 1977).

- (15) a. b[ɛ]r.ro b[e]r.rei.ro b[ɛ]r.rão
 'yell' 'yelling' 'big yell'
- b. p[ɔ].bre p[o].bre.za p[ɔ].brão
 'poor' 'poverty' 'very poor'

Differently from the derived forms of the second columns above, in the augmentative forms there is no [ɛ] → [e] or [ɔ] → [o] process.

Regarding the label that results from the merger between GEN₁ and GEN₂ in the structure (13), it seems plausible to assume that the internal gender head attached to the root is the one that carries a valued gender. The GEN₂

(17) Absence of GEN₁ in the augmentative structure



In (17), the merger between AUG and GEN₂ is prior to the merger between Root and GEN₂. In this context, the default [masculine] is the only feature GEN may assume. This readily explains why a feminine noun in BP may present masculine corresponding augmentative form, but not the other way around.

It remains to be discussed the relation between gender and -z augmentative. Before doing so, it is important to remark that the augmentative marker -zão is anteceded by a complete word, which includes the final vowel of the base form. It is then very plausible to say that the root and the -z augmentative have independent GEN heads, as (i) the vowel that follows the Root and the vowel that follows the -z form are both phonologically realized and (ii) the vowel that follows the -z form is completely independent from the Root, reflecting the gender patterns of the language and not specificities related to the Root itself.

Also, it is possible to find morphophonological changes triggered by the presence of a plural marker linearly before -zão (see MENUZZI, 1993; LEE 1997), as can be seen in (18) below⁴:

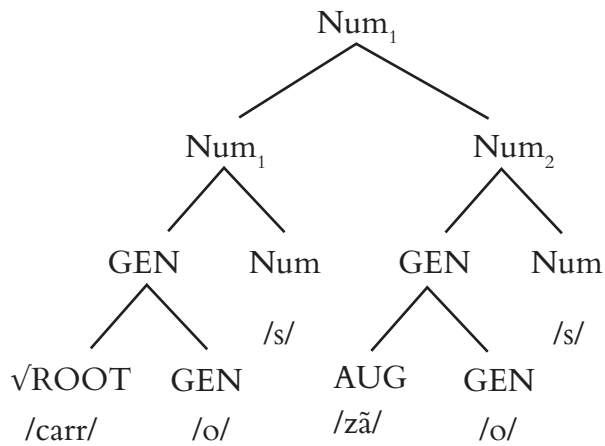
(18) -zão and morphophonological changes triggered by plural

Singular	Plural	Augmentative Plural
pastel ('pastry')	pastéis	pasteizões
jornal ('newspaper')	jornais	jornaizões
avião ('airplane')	aviões	aviõeões
pão ('bread')	Pães	pãezões

In order to represent the -z augmentatives we propose a structure that contains two gender heads, but also two different number heads, one related to the Extended Projection above the root and another one related to the functional structure above the -z augmentative, as shown below:

⁴ It is also possible that the morphophonological changes triggered by plural to be absent from augmentative formations, even in the presence of the consonant /z/, as in pastelzões, jornalzões, aviãozões, pãozões, for example. For accounts of these data see Bacharach and Wagner, 2007; Armelin, 2015).

(19) Anexação de -zão



output: carrozões ('big cars')

This analysis is supported by the fact that: (i) the final vowel completing the non-augmentative form is phonologically realized; (ii) gender agreement between *-zinh/-zã* formation and the base nouns is obligatory; the phonological realization of the vowel completing *-zão* is independent of the root that participates in the formation; (iv) the morphophonological changes triggered by the plural are found linearly before the augmentative marker, and (v) there is number agreement between *-zão* formation and the base noun.

20

5. Augmentative and (non)Compositionality: empirical facts

The *-ão* augmentative formative may trigger non-compositional interpretation. When this is the case, the augmentative forms are ambiguous between a compositional and a non-compositional reading. The *-z* augmentative counterpart, on the other hand, may only be compositionally interpreted.

- (20) a. sapato ('shoe') sapatão ('big shoe') or ('lesbian')
- b. sapato ('shoe') sapatozão ('big shoe')
- c. macaco ('monkey') macacão ('big monkey') or ('overalls')
- d. macaco ('monkey') macacozão ('big monkey')

Importantly, however, the non-compositional augmentative is masculine, regardless of the gender of the non-augmentative corresponding form.

- (21) a. sala salão compositional : big room
- room(f) room-aug(m) non-compositional : saloon

b. sacola	sacolão	compositional: big bag
bag(f)	bag-aug(m)	non-compositional: vegetable store

In this sense, if the augmentative formation in (21a-b) were to be replaced by its correspondent feminine counterpart – which shows the same gender as the non-augmentative form – the resulting formation would only be interpreted compositionally⁵:

(22) a. sala	salona	compositional: big room
room(f)	room-aug(f)	non-compositional: -----
b. sacola	sacolona	compositional: big bag
bag(f)	bag-aug(f)	non-compositional: ---

Quite independently of the gender of the formation, *-zão* augmentative constructions, on the other hand, do not trigger non-compositional interpretation.

(23) a. sala	salazona	compositional: big room
room(f)	room-aug(f)	non-compositional: ----
b. sacola	sacolazona	compositional: big bag
bag(f)	bag-aug(f)	non-compositional: ---

The nouns *sala* and *sacola* are feminine in BP. In the above data, it is possible to see that the masculine augmentatives in (21a-b) may have both a compositional and a non-compositional reading. On the other hand, for the feminine augmentatives in (22a-b) and (23a-b) the non-compositional interpretation is no longer available.

Another important point to be noticed is that non-compositional augmentatives may be formed from derived forms. This can be evidenced by the data below, in which a suffix intervenes between the augmentative and the root.

⁵ An anonymous reviewer noticed the existence of a pair in which only the feminine augmentative form would be non-compositionally interpreted: *bichão/bichona*. The first member of the pair may only be interpreted as a ‘big bug’, and this would be the compositional interpretation. On the other hand, the second member of the pair may only be interpreted as ‘someone who is really gay’, and this would be the non-compositional interpretation. It is important to notice however, that, in this case, the so-called non-compositional interpretation is not to be attributed to the augmentative morpheme, since the non-augmentative correspondent form *bicha* already means ‘gay’ in BP. Interestingly, the non-augmentative form may be used either with a feminine determiner (*uma bicha*) or with a masculine determiner (*um bicha*).

- | | |
|---|---|
| (24) a. bat-id-ão
to beat-part-aug
'a kind of music' | b. brasil-eir-ão
Brazil-suff-aug
'Brazilian football championship' |
| c. cint-ur-ão
cincture-suff-aug
'belt given to the winner of a fight' | d. paul-ist-ão
São Paulo-suff-aug
'Football championship of
São Paulo State' |
| e. mens-al-ão
month-suff-aug
'a corruption schmeme' | f. min-eir-ão
Minas-suff-aug
'the name of a stadium in Minas
Gerais' |

In the above formations, it is exactly the augmentative morpheme that is triggering the non-compositional interpretation. This is so because the relevant non-compositional interpretation is not available for the corresponding forms in the absence of the augmentative formative.

6. Compositionality: the state of the art

In this section, we present some proposals developed in the literature to deal with meaning computation. More specifically, three accounts are brought into discussion: Marantz (2001, 2007) and Arad (2003), Marantz (2013), and Borer (2013a, 2013b, 2014).

6.1 The Marantz/ Arad Hypothesis

Based on Chomsky (2001), Marantz (2001) proposes that phase heads should be identified within word formation. The merger of a phase head triggers the transfer of a chunk of syntactic structure to both phonological and conceptual interfaces. Consequently, the transferred material gets inaccessible to further operations. More specifically, Marantz (2001) proposes that the phase heads within words should be equated to the categorizing heads. The idea underlining this reasoning is that a root has no category on its own, and, as such, it is not a legible element for the interpretative component. Once categorizers are merged in the structure, the formed object can be sent to the interfaces.

This idea is further explored in Arad (2003), giving rise to what has been known as the Marantz/Arad hypothesis. Simply put, the idea is that the idiomatic syntactic domain applies at the first categorizing node. All further syntactic derivation must involve the meaning negotiated in the merger between root and the first categorizer. This is so, precisely because the categorizing head is

a phase, and subsequent derivation must include the interpretation provided by the phase head, which has been fixed on an earlier cycle of Spell-out.

(25) The Marantz/Arad Hypothesis

Locality constraint on the interpretation of roots: roots are assigned an interpretation in the environment of the first category-assigning head with which they are merged. Once this interpretation is assigned, it is carried along throughout the derivation. (ARAD, 2003, p.747)

Idiosyncratic meaning (and phonology) is, then, restricted to the inner domain, that is, the domain below the first categorizing head. The phase is a closed domain, and elements that merge above it cannot have access to what is inside it. Consequently, elements merging above the inner domain cannot alter the interpretation (or the phonology) previously established in the first phase.

This hypothesis has faced many counterexamples (see Anagnostopoulou and Samioti (2014), concerning adjectival participles and verbal adjectives in Greek; Borer (2013a), concerning constructs N-N in Hebrew and, also, Lemle (2013), concerning derived forms in BP) pointing to the fact that the first categorizing head constitutes an excessively restricted domain for non-compositional interpretation to be licensed. In this sense, in Marantz (2013) some of the counterexamples are further explored and the resulting proposal is that the boundary of non-compositional interpretation is still the phase, but based on Embick (2010), the author proposes that the limits for contextual allomorphy are the same ones that restricts contextual allomorphy.

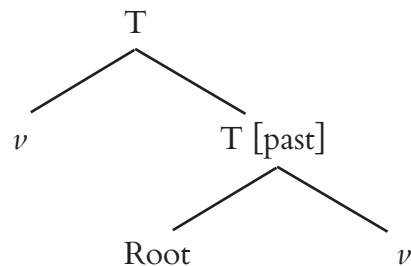
6.2 Marantz (2013): a theory of contextual allomorphy

The central idea developed in Marantz (2013) is that there is, in the semantic side, an equivalent to what is known as contextual allomorphy in the phonological side. Contextual allomorphy is at stake when the competing forms realize the same set of features of a syntactic node, and the choice between them is determined exclusively by the context. In the semantic side, this is equivalent to say that there are multiple meanings competing to be realized in a certain terminal node, let's say the root, and the winning interpretation is chosen based on the syntactic context that is local to the relevant node.

In order to pursue this idea, a theory of contextual allomorphy is in need. Exploring the hypothesis that the restrictions imposed on contextual allomorphy, and on contextual allomorphy are the very same, Marantz (2013) follows the main lines of the system proposed in Embick (2010), which investigates, exclusively on the PF side, the locality relations that may hold for allomorphic interactions to be possible.

Embick (2010) argues that allomorphy should be computed by local domains. In the definition of what the allomorphic domain should be like, the author advocates in favor of the idea that adjacency is the relation that licenses allomorphic interactions. However, if the Categorization Assumption holds (see EMBICK; NOYER, 2007), adjacency cannot be the whole story. See, for example, the formation of past verbs in English.

(26) Allomorphic interaction in the English past tense



(EMBICK, 2010, p. 14)

In the above representation, the root has to merge with the categorizer little *v* before Tense enters the structure, since the Categorization Assumption forces this step. The issue is that the spell-out of the T head is clearly determined by the root: it is, for example, \emptyset in the context of the roots $\sqrt{\text{hit}}$ and $\sqrt{\text{cut}}$, it can also be $-\text{t}$ in the context of the root $\sqrt{\text{bend}}$, contrasting with the default $-\text{ed}$.

24 In order to deal with that, Embick (2010) proposes two conditions in the computation of allomorphic domains: (i) the invisibility of phonological zeros for adjacency, and (ii) an asymmetry between categorizing heads and non-categorizing heads in the outer domain. The idea in (i) is roughly that phonological zeros do not count as interveners. In the structure in (26), for example, because little *v* is phonological empty, it does not block the adjacency between the root and the T head. The idea in (ii), by its turn, is summarized below:

(27) Asymmetry between phase heads and non-phase heads

- a. ... α] x] Z]
 - Generalization: Noncyclic Z may show contextual allomorphy determined by α , as long as x is not overt.
- b. ... α] x] y]
 - Generalization: Cyclic y may not show contextual allomorphy determined by α , even if x is not overt.

(EMBICK, 2010, p, 16)

In the above generalization, x and y are phase heads, whereas Z is not. In (27a), if x is not overtly spelled-out, the form of Z can be determined by α ,

precisely because Z is not a phase head. On the other hand, in (27b) there is no possible allomorphic interaction between α and y , regardless of the phonological status of x , and this is so because y is also a phase head.

The system proposed in Marantz (2013) to deal with contextual allosemy is parallel to Embick's (2010) system sketched above. This means that two elements have to be adjacent to each other in order to influence each other's interpretation they. Since we are dealing with interpretation, the adjacency at stake has to be a semantic one. In this sense, if there is an intervener, in order for this intervener not to preclude non-compositional interpretation, it has to be semantically null. As a consequence, the first phase head is not the only domain for non-compositional interpretation anymore, since this head could be phonologically overt but semantically null, not counting as an intervener.

Besides the idea that semantically null elements do not count for adjacency in the interpretative branch, Marantz (2013) also claims that it is possible that configurations like the one presented in (26) trigger contextual allosemy. Concretely speaking, this means that a higher head may condition the meaning of a lower head over a (semantically null) intervener, only if the relevant higher head is not a phase. In the structures bellow, x and y are phase heads, while Z is not. Consider the hypothetical syntactic structure:

(28) Hypothetical Syntactic Structure: phase heads vs. non-phase heads



Let's imagine that the first categorizing head x is semantically null and does not intervene on the semantic adjacency between the root and the higher node. In the Marantz (2013) system, it is possible for Z, a non-cyclic head, to condition the allosemy of the root node (28a), but it is not possible for y , a cyclic head, to condition the semantics of the root (28b).

6.3 Borer (2013a, 2013b, 2014): Structuring Content

Borer (2013a, 2013b, 2014) shares with Marantz (2001) and Arad (2003) the basic intuition that meaning, or Content, as the author calls it, must be calculated locally, and, hence, defined by a well delimited syntactic domain. However, contra the Marantz/Arad hypothesis, the proposal put forth by the

author is that the domain of Content is not defined by the first categorizing head (that is little n, little v or little a, for example), but by the lowest segment of an Extended Projection.

The system for Content derivation is developed under the framework of the Exo-skeletal model (see Borer, 2005a, 2005b, and 2013b for details), henceforth the XS. Broadly speaking, in the XS model, the building blocks of syntax are roots and functors. Roots are devoid of any formal properties, while functors name syntactic or semantic functions that are constant across their instantiations. The functors are divided into two different categories: the Categorical functors (C-functors) and the Semantic functors (S-functors).

The C-functors divide the categorial space: at the same time, they project a lexical node and define the category of its complement. To make it concrete, the suffix *-ation* in English, for example, would be a C-functor that projects a noun and defines its complement as equivalent to a verb.

The S-functors, on the other hand, are implicated in the valuing of functional nodes and are typically assumed to be linked to Extended Projections. In this sense, D (determiner) is a semantically valued node (by the S-functor THE, for example) in the extended projection of the noun, and T is a semantically valued node (by the S-functor PAST, for example) in the extended projection of the verb. The segments in the Extended Projection collectively define the category of their complement.

In the XS model, the component responsible by meaning assignment is called Encyclopedia. In order to do that, the Encyclopedia searches post-syntactic representations delimited by qualified brackets. Borer's (2013a, 2013b, 2014) proposal is that non-Compositional Content emerges as the output of a single encyclopedic search. The workings of the system also have two important points: (i) Content is optional, which means that it is possible that an encyclopedic search returns no meaning at all, and (ii) segments of the Extended Projection are barriers for encyclopedic searches, in the sense that when they enter the derivation, Contentless sequences cannot be rescued. The articulation of the system is illustrated by the data below:

(29) Encyclopedic searches in the XS model

Base	Content	Derived Form	Content
a. sturd, flim	No listed Content	sturdy, flimsy	STURD, FLIMSY
b. bulge, boss	BULGE, BOSS	bulgy, bossy	No listed Content
c. blood, flake	BLOOD, FLAKE	bloody, flakey	BLOODY, FLAKEY

(Adapted from BORER, 2014, p. 86)

Using the terms of the XS model, the formative *-(s)y* is a C-functor that projects an adjective and defines its complement as a noun. This being so,

the bases on the first column, although co-extensive with the correspondent roots, are noun equivalent. In (29a) the encyclopedic search targeting the base returns no meaning, since the noun equivalent forms *sturd* and *flim* are not matched with any listed Content. The correspondent derived forms *sturdy* and *flimsy*, however, are successfully matched with Content in the Encyclopedia. Importantly, the relevant interpretation cannot be a compositional one, since the bases have no Content on their own. The emergence of the non-compositional meaning is explained by the fact that meaning is computed on the bases of the en-search that returns meaning for the whole derived form.

An opposite case is illustrated by (29b): Content is matched with the base form, but not with the derived form. Crucially, the fact that the derived form is Contentless does not preclude it from participating in further syntactic derivation. As a result of the previous Content matched with the base, compositionality is a consequence, and the derived forms may only have whatever Content emerges from composing the listed Content of the base with a (Contentless) C-functor.

The cases in (29c) are more complex, since the base noun equivalent form and the derived form are both matched with listed Content in the Encyclopedia. As emphasized by Borer (2014) the derived forms can have a compositional and a non-compositional meaning. In the non-compositional interpretation *bloody* means ‘damned’, while *flakey* means ‘flippant’, ‘odd’. The compositional interpretation is derived in the following way: the encyclopedic search that targets the base forms *blood* and *flake* are matched with listed meaning and, consequently this meaning is carried throughout the derivation. With respect to non-compositional meaning, once Content is optional, the idea is that no Content was matched with the bases *blood* and *flake*, but Content was matched with the derivative as a whole in the next encyclopedic search.

Summarizing the guidelines for Content assignment in Borer’s system are the following: (a) Encyclopedic searches targets bracketed syntactic structures; (b) Content is optional and (c) nodes in the Extended Projection are barriers for encyclopedic search.

7. Deriving (Non)Compositionality in Augmentative Structures

Crucially, assuming a derivational theory of word formation does not entail dispensing with the idea that non-compositional meaning needs to be listed. The question to be answered is, then, at which point in the syntactic derivation meaning can be accessed. In this context, syntactic approaches to word formation have focused on defining a syntactic well-defined domain within which non-compositional interpretation could be licensed.

The interaction between augmentative markers and derivational morphology seems to be showing that the syntactic limit for non-compositionality set out in the Marantz (2001, 2007)/Arad (2003) hypothesis is too narrow. As showed in the description section, non-compositional interpretation may be triggered by augmentative formatives that are attached to derived forms (see (24a-f)). The possibility that the augmentative marker licenses non-compositional interpretation after a derivational morpheme is not expected if, as proposed in Marantz (2001, 2007)/Arad (2003), non-compositional interpretation is restricted to the domain below the first categorizing head.

On the other hand, from the perspective developed in Marantz (2013), non-compositional meaning is related to the notion of semantic adjacency. More specifically, the presence of intervening phonological material is tolerated, provided such material is semantically empty. Thus, a possible analysis for the non-compositional augmentatives that trigger non-compositional meaning regardless the intervention of the derivational morpheme is that the intervening element is actually semantically null. However, in some cases, this position does not seem simple to sustain. Note that it is difficult to say that *-eiro* in *brasileirão*, for example, is not semantically computed, since the interpretation of the whole formation includes the interpretation of *brasileiro* ('Brazilian'): *brasileirão* is actually a sports league in which the participating teams are necessarily Brazilian. In this sense, it is not always possible to say that the intervening material is ignored.

Given the empirical inadequacy of both the Marantz (2001, 2007)/Arad (2003) hypothesis and Marantz (2013), we now reevaluate the syntactic structures proposed in the previous sections with the aim of opening perspectives that may help defining the structural domain within which non-compositional interpretation can be licensed. In order to do so, we assume with Borer (2013a, 2013b, 2014) the idea that segments of Extended Projection constitute barriers to the meaning attribution. The syntactic structures proposed to derive the empirical pattern found in BP augmentatives are the following:

- a) A syntactic structure containing two gender heads: augmentatives that agree in gender with the base noun (see (16)).
- b) A syntactic structure in which the internal gender head is absent: augmentatives that do not agree in gender with the base noun, and are always masculine by default (see (17)).
- c) A syntactic structure containing both two gender heads and two number heads: augmentatives that agree in gender and number with the base noun, and preserve phonological changes triggered by the plural (see (19)).

Concerning the syntactic structure in (a), the AUG head is not local to the root due to the presence of the internal functional projection of Gen_1 , responsible by caring the valued gender feature for the Agree process to be licensed. In this position, it is predictable that AUG may not trigger non-compositional interpretation, and that's exactly the empirical pattern found in the language. On the other hand, if the structure in (b) is considered, AUG is inside the first segment of Extended Projection that categorizes the root. This being so, the prediction is that AUG may trigger non-compositional interpretation. Besides that, given the impossibility of gender agreement, it is also predictable that the gender of non-compositional augmentatives is always masculine. Finally, the syntactic structure in (c) contains even more syntactic functional material intervening between the AUG head and the root, since it attaches after a number head. This explains why the *-z* augmentative may only be compositionally interpreted.

8. Final Considerations

This paper investigates augmentative formation with *-ão* and *-zão* in Brazilian Portuguese, focusing specially on the relation these formatives establish with gender and compositionality. Assuming a syntactic approach to word formation, the different empirical patterns were captured by the presence of either one or two gender heads in the augmentative structure. When two gender heads are present in the structure, a mechanism of gender Agree (Chomsky, 2000, 2001) is assumed to be operating within words. Also, given the presence of intervening syntactic material, non-compositional interpretation is not licensed. Nevertheless, when only one gender head is present, no gender agreement is possible and the resulting augmentative gender is a default masculine. Finally, this is the local configuration that licenses non-compositional interpretation.

References

ARMELIN, Paula Roberta Gabbai. *Classifying Nominals in Brazilian Portuguese*. In: VESELOVSKÁ, L.; JANEBOVÁ, M. (Eds.) *Complex Visible Out There: Language, Use and Linguistic Structure*. Olomouc Modern Languages Series, v. 4, 2014. p. 67–82.

_____. *A Relação entre Gênero e Morfologia Avaliativa nos Nominais do Português Brasileiro: Uma abordagem sintática da formação de palavras*. Tese (Doutorado) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.

ANAGNOSTOPOULOU, Elena; SAMIOTI, Yota. *Domains within words and their meanings: A case study*. In: ALEXIADOU, A.; BORER, H.; SCHÄFER, F. (Eds.). *The Syntax of Roots and the Roots of Syntax*. Oxford: Oxford University Press, 2014. p. 81–111.

ARAD, Maya. *Locality Constraints on the Interpretation of Roots: The Case of Hebrew Denominal Verbs*. *Natural Language and Linguistics Theory*, v. 21, p. 737–778, 2003.

BACHRACH, Asaf; WAGNER, Michael. *Syntactically Driven Cyclicity vs. Output- Output Correspondence: The Case of Adjunction in Diminutive Morphology*. *U. Penn Working Papers in Linguistics*, v. 10, n. 1, 2007.

BORER, Hagit. *Taking Form: Structuring Sense, Vol. II*. Oxford: Oxford University Press, 2013a.

_____. *The syntactic domain of Content*. In: BECKER, M.; GRINSTEAD, J., and ROTHMAN, J. (Eds.). *Generative Linguistics and Acquisition: Studies in Honor of Nina Hyams*. Amsterdam: John Benjamins, 2013b. p. 205–48.

_____. *Derived nominal and the domain of Content*. *Lingua*, v.141, p. 71–96, 2014.

CHOMSKY, Noam. *Bare phrase structure*. In: WEBELHUTH, G. (Ed.). *Government and Binding Theory and the Minimalist Program*. Cambridge, MA: Basil Blackwell, 1995. p. 383–440.

_____. *Minimalist Inquiries: The Framework*. In: MARTIN, R.; MICHAELS, D.; URIAGEREKA, J. (Eds.). *Step by Step: Essays on Minimalist Syntax in*

Honor to Howard Lasnik. Cambridge, MA: MIT Press, 2000. p. 89-156.

_____. *Derivation by Phase*. In: KENSTOWICZ, M. (ed.) *Ken Hale: A life in language*. Cambridge, MA: MIT Press, 2001. p. 1-52.

EMBICK, David. *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press, 2010.

EMBICK, David; ROLF, Noyer. *Movement operations after syntax*. *Linguistic Inquiry*, v.32, n.4, p.555-595, 2001.

_____. *Distributed Morphology and the Syntax/Morphology Interface*. In: RAMCHAND, Gilliam; REISS, Charles (Eds.). *The Oxford handbook of Linguistics Interfaces*. Oxford: Oxford University Press, 2007. p. 298-324.

GRIMSHAW, Jane. *Extended Projections*. MS, Brandeis University, 1991.

HALLE, Morris. *Distributed morphology: Impoverishment and fission*. MIT Working Papers in Linguistics 30: papers at the interface, edited by Benjamin Bruening, Yoonjung Kang and Martha McGinnis. MITWPL, Cambridge p. 425-449, 1997.

HALLE, Morris; MARANTZ, Alec. *Distributed Morphology and the pieces of inflection*. In: HALE, K.; KEYSER, S. J. (Eds.). *The view from Building 20*. Cambridge, MA: MIT Press, 1993. p. 111-176.

LEE, Seung-Hwa. *Morfologia e Fonologia Lexical do Português do Brasil*. Tese de Doutorado, Unicamp – IEL, Campinas, 1995.

LEMLE, Miriam. *Aspectos arbitrários e composicional na morfologia derivacional*. Apresentação no II Colóquio Brasileiro de Morfologia. UFRJ, 2013.

MARANTZ, Alec. *Words and Things*. Manuscrito. NYU, 2001.

_____. *Phases and words*. Manuscrito. NYU, 2007.

_____. *Locality Domains for Contextual Allomorphy across the Interfaces*. In: MATUSHANSKY, O.; MARANTZ, A. (Eds.). *Distributed Morphology Today: Morphemes for Morris Halle*. Cambridge, MA: MIT Press, 2013. p. 95-116.

MORENO, Cláudio. *Morfologia nominal do português: um estudo de fonologia lexical*. Tese de Doutorado – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 1997.

WETZELS, W. Leo. *Mid Vowels Neutralization in Brazilian Portuguese*. *Cadernos Estudos Linguísticos*, v. 23, Campinas, p. 19-55, 1992.

Submetido em: 05/03/2017

Aceito em: 06/06/2017