

# THE MINIMAL WORD IN EUROPEAN PORTUGUESE: THE ORALIZATION OF ABBREVIATED FORMS

João VELOSO\*

- **ABSTRACT:** The phonological constraints that define which phonetic chains are eligible to become a word in a given language generally include a Minimality Condition (MC). MC imposes a minimal amount of phonological substance that must be included by any single word of the language. Traditionally, MC is measured in terms of syllable weight or syllabic extension of the word. It has been much debated whether MC corresponds to a truly universal constraint; as for Portuguese, previous studies have been conflicting between the acceptance and the refusal of its productivity in the phonology of the language. In this study, we will assess the functionality of MC for the building of Portuguese lexical entries by means of observing the oralization of abbreviations. This is assumed here as a word-productive process in Portuguese. Based upon an experimental study with a group of native speakers of Contemporary European Portuguese (CEP), we conclude that (i) MC is an operative constraint in the phonology of Portuguese, and (ii) its defining criterion is purely linear and segmental. According to our proposal, MC is respected in CEP when a phonetic chain contains a minimum of three segments, regardless of syllable weight and syllabic extension.
- **KEYWORDS:** Minimality Condition. Syllable Weight. Word. Wordiness Constraints. Well-formedness.

## Introduction

A satisfactory definition of the word, as it is well-known, is one of the main difficulties for linguists. A substantial part of such difficulty lies, first of all, on the lack of universal, objective, formal criteria that make it possible to identify this concept in an unambiguous way and to segment larger linguistic structures in units unanimously recognized as “words”. In addition, we lack further criteria allowing an intrinsic distinction and characterization, based solely on linguistic arguments, of the various types of meaning-bearer units, like affixes, clitics, clitic groups, among others. Different approaches, comprehending typologically different languages, have discussed this *paradox*: why an intuitive, basic unit such as the word proves to be so hard to define

---

\* FLUP/CLUP – Universidade do Porto. Faculdade de Letras/Centro de Linguística. Porto – Portugal. [jveloso@letras.up.pt](mailto:jveloso@letras.up.pt)

in purely technical terms? Examples of such reflections on this issue, some of which focusing on this very contradiction between the “easiness” of the self-intuitive notions of the word and the difficulties of a technical explicitation of this concept, may be found in Jones (1931), Krámský (1969), Juilland; Roceric (1972), Halle (1973), Aronoff (1976), Bisol (2000; 2004; 2007), Julien (2006), Rosa (2006), Veloso (2009; 2010; 2016), Haspelmath (2011; 2012a,b), Ferreira (2012), Villalva (2012), Mendes (2013), Ulrich (2013), Elordieta (2014), Ibarretxe-Antuñano; Mendivil-Giró (2014) and others.

Among the requirements commonly mentioned as a fundamental criterion for a satisfactory definition of the word, the *Minimality Condition* (MC) is found very often: in a given language, a sound chain will be a good candidate to be recognized as a word if and if only, together with other properties, it contains a minimum amount of phonological material. If below such limit, no units can be admitted as words of that language (see, for a review of these arguments, McCARTHY; PRINCE, 1995). MC is the central issue of the present study.

In some previous analyses of the same problem (see again, e.g., McCARTHY; PRINCE, 1995), it has been argued that MC is not a universal property: some languages could not include any minimality constraints in their phonologies, whereas cross-linguistic MC parametrization may vary importantly from grammar to grammar.

As far as Portuguese is concerned, two positions are assumed by the literature: on the one hand, some authors postulate the ineffectiveness of any minimality constraint in this language; oppositely, other analyses advocate MC as a mandatory constraint in the phonology of Portuguese:<sup>1</sup>

In languages like Portuguese (either in its European or Brazilian varieties), the occurrence of words like *pé* ‘foot’, *mi* ‘mi’ or *nu* ‘naked’ has led several authors – e.g., Bisol (2000), Vigário (2003) – to the conclusion that such constraint [Minimal Word Constraint] is inoperative in Portuguese. However, Vigário (2003: 159) notices that, in the *Português Fundamental* corpus [a corpus of the most frequent words of European Portuguese], including approximately 7.000 inflected forms, only 138 (lexical) monosyllabic words are found; within this subset, 28 only consist of open syllables. Such low values have motivated Booij

---

<sup>1</sup> Original quotation, in Portuguese: “Em línguas como o Português (na variedade brasileira ou europeia), a possibilidade de encontrarmos palavras como *pé*, *mi* ou *nu*, levou autores como Bisol (2000) e Vigário (2003) a considerar que tal restrição [Restrição de Palavra Mínima] não se encontra operativa nessa língua. Contudo, Vigário (2003: 159) não deixa de notar que, tendo em conta a lista do *Português Fundamental* [...] que inclui cerca de sete mil formas flexionadas, apenas 138 palavras (lexicais) são monossilábicas, e destas apenas 28 constituídas por sílaba aberta. Tais baixos valores conduzem Booij (2004) a contrapor que o Português é de facto sensível a restrições de minimalidade, mas que existe um reduzido número de palavras que a violam.

(.....)

Os resultados [do nosso próprio estudo] mostram que a proporção das palavras monomoraicas/monossilábicas em relação aos restantes formatos de palavra é muito maior do que a revelada [por estudos anteriores] [...]. O efectivo uso de formas monomoraicas/monossilábicas aponta, assim, para que *a palavra prosódica no Português (Europeu) não seja de facto sensível a restrições de tamanho mínimo.*” (VIGÁRIO; MARTINS; FROTA, 2005, p.903; our italics)

(2004) to argue that Portuguese, in fact, admits minimality constraints, even though a small amount of words can violate them.

(.....)

[Our own] results show that the proportion of monomoraic/monosyllabic words in relation to the other word formats is higher than [previously] demonstrated [...]. The effective usage of monomoraic/monosyllabic forms suggests, thus, that *the prosodic word in (European) Portuguese is not conditioned by minimal extent constraints.*”

(VIGÁRIO; MARTINS; FROTA, 2005, p.903; our translation, our italics)

In our study, it is our aim to gather evidence that could lead us to a better evaluation of two specific, inter-related questions: **is MC a functional constraint in Portuguese, and, if so, how is it parametrized in this language?**

Our analysis will be limited to European Portuguese (EP)<sup>2</sup> and follow a methodology previously adopted: the oralization of abbreviations as initialisms or acronyms.

Throughout the paper, the terms “abbreviation” and “abbreviated form” will be used to make reference to the process and to the result of a specific word-formation routine which converts special phonetic chains into “new” lexical words by means of one of the following ways<sup>3</sup> (which are not the only abbreviation procedures available in Portuguese):<sup>4</sup>

(i) *acronymy*: phonetization of a longer expression into a shorter one (=“acronym”) through the combination of morphologically unmotivated subsegments of the several words found in the original, larger expression; this gives rise to a phonetic chain similar to a phonological word ( $\omega$ ) of the language, since it respects its phonotactic rules and is in accordance with all phonological, morphological, syntactic and semantic conditions (stress, lexical class, thematic class, inflection, gender, meaning, etc.). Example: *Organização das Nações Unidas* ‘United Nations Organization’ → “*ONU*”=[ $\text{ɔ}^{\text{h}}\text{nu}$ ] <sub>$\omega$ -Noun\_Feminine</sub>;

(ii) *initialization*: phonetization of a longer expression into a shorter one (=“initialism”) through the spelling of the initial letters of each lexical word

---

<sup>2</sup> For this reason, only initialisms and acronyms that are current in the European variety of Portuguese will be considered here; accordingly, all phonetic transcriptions will take the standard norm of European Portuguese as the pattern.

<sup>3</sup> Bearing in mind that abbreviations are functionally equivalent to the longer expressions they replace and the fact that they absorb the essential grammatical properties of morphosyntactic words (see examples and analyses in the text), we assume them here as true entries of the speakers’ mental lexicon. In fact, most of them are largely shared by the speakers’ community and acquire a stable meaning (very often, they are even admitted by lexicographers as entries of current dictionaries). Their lexical representation, though, is inevitably subject to the natural variability that is found when individual lexicons are compared. For a discussion of this topic, see, among others. (CORREIA; LEMOS, 2005).

<sup>4</sup> See, for instance, and limiting the literature on this subject to the studies referring to European Portuguese exclusively, Correia e Lemos (2005), Villalva (2008) and Pereira (2013), who offer more detailed descriptions of abbreviation (and other non-concatenative) processes in Portuguese and discuss their role as word-formation – hence, as truly *morphological* – processes in this language.

contained in the original, longer expression, so that a prosodic word group (VIGÁRIO; FERNANDES-SVARTMAN, 2010) is formed, behaving as a sole morphosyntactic word though (one sole gender, one sole lexical class, etc.). Example: *Movimento Democrático das Mulheres* ‘Democratic Women’s Movement’ → “MDM”=[<sup>l</sup>εmĩ]<sub>o</sub>[<sup>l</sup>de]<sub>o</sub>[<sup>l</sup>εmĩ]<sub>o</sub>]<sub>PWG\_Noun\_Masculine</sub>.

In the second part of our study, an empirical exploration will be presented. We shall then observe the linguistic behaviour of a group of native speakers of EP when confronted with this kind of specific chains and asked to treat them as “words” of their language.

Subjects will be given a list of written representations<sup>5</sup>, introduced as abbreviated forms and embedded in syntactically invariant sentences. They will be asked to read that list. Therefore, we shall try to find cues giving us relevant information about the linguistic processing of this kind of stimuli and, simultaneously, about the obligatory conditions of the language’s words. More specifically, it is our purpose to check the eventual functionality of MC in the implicit phonological knowledge of EP native speakers/listeners. In brief, this way to investigate this specific issue coincides with Plénat’s (1993) methodology to study the same question with French speakers.

Based on these preliminary questions, our article will be divided as follows:

- in section 2, a review of some basic assumptions related to MC and other phonological constraints, either at a general level or focusing on Portuguese, will be made;
- in section 3, we shall discuss the relevance of the methodological option that will be followed in our empirical study – the oralization of abbreviations – as a valid means to evaluate MC and other phonological constraints as mandatory requirements for the wordiness in Portuguese;
- the empirical study, with data and subjects of Contemporary EP, will be presented in section 4;
- section 5 will present the final observations of the study.

## Phonological constraints and wordiness

As said in the Introduction, an exhaustive inventory of all linguistic criteria leading to a clear definition, identification and delimitation of words as linguistic units, even for inflectional languages like Portuguese, has challenged and motivated different interpretations.

Studies such as the aforementioned have attempted to identify the *wordiness conditions* (ARONOFF; FUDEMAN, 2005, p.36-38) which make it possible, at

---

<sup>5</sup> As explicitly admitted by authors such the researchers mentioned in note 4, in languages with an alphabetic, conventional system of writing – as it is the case of Portuguese – the abbreviation processes are regularly grounded on the written forms of the larger words or expressions that are the input for abbreviation.

least partially, to assign or to refuse the status of “word” to given phonetic chains in a given language.

Wordiness conditions, in fact, are shared by the different grammar modules (namely, phonology, morphology and syntax). In the present study, however, as said before, we shall not deal with all these wordiness dimensions: we shall focus, deliberately and exclusively, on *Minimality Condition*.

### Minimality Condition: general formulation

MC is the phonological constraint that, interacting with others, imposing a minimum amount of phonological material so that a phonetic chain may be admitted as a word in a given language. Although its parametrization varies cross-linguistically, we assume, as a general formulation of MC potentially valid for a great number of languages of the world, a formulation like (1),<sup>6</sup> based on McCarthy and Prince (1995, p.321-322).<sup>7</sup>

(1). Generic formulation of the MINIMALITY CONDITION (*ap.* McCARTHY; PRINCE, 1995, p.321-322)

(1a). In **languages WITH quantitative distinctions**: the minimal word must contain at least one heavy syllable.

(1b). In **languages WITHOUT quantitative distinctions**: the minimal word must contain at least two syllables.

### The Minimality Condition in Portuguese

We shall now enter the discussion of the specific issue of the relevance of MC in EP.

Since Portuguese does not admit quantitative distinctions, we should expect, according to McCARTHY; PRINCE’s (1995) postulate referred to in (1) (see (1b)), that all words of Portuguese would correspond, at least, to two syllables.

Contemporary EP lexicon includes, though, a non-neglectable amount of monosyllabic words. In addition to clitics,<sup>8</sup> EP has many monosyllabic lexical words, as shown in Figure 1. Words in this list are categorized in *light* monosyllables (non-

---

<sup>6</sup> This formulation – taken here and among many other authors as generic – is not absolutely consensual. Vigário; Martins; Frota (2005), e.g., can be counted among the authors who discuss it very critically.

<sup>7</sup> “In *quantity-sensitive* languages, which distinguish syllable weight, the minimal word is bimoraic; in *quantity-insensitive* languages, all syllables are presumptively monomoraic, and so the minimal word is disyllabic.” (McCARTHY; PRINCE, 1995, p.321-322).

<sup>8</sup> The special status of clitics as real “words” has been largely discussed by previous literature, both with respect to Portuguese and to other languages as well. See, for example: Vigário (1998; 2003; 2007); Nespor (1999); Van Oostendorp (1999); Van Riemsdijk (Org., 1999); Vogel (1999); Ennaji (2000); Sadiqi (2000); Gerlach; Grijzenhout (2000); BISOL (2004); Nespor; Vogel (2007); Gori (2007); Ferreira (2012); Veloso (2012; 2013; 2016). Given its specificities and the limited scope of this study, we will not go any further into the analysis of clitics and their relation to MC in Portuguese.

branched rhyme) and *heavy* monosyllables (branched rhyme, according to the several prosodic and morphological conditions explicated in the right-foremost columns of Figure 1); the reasons for this categorization will be made more clear later on.

**Figure 1** – Examples of monosyllabic words in Portuguese<sup>9</sup>

<i>Light</i> monosyllables	<i>Heavy</i> monosyllables		
	Branched nucleus	Segmentally filled coda	Branched nucleus + Segmentally filled coda (lexical /S/)
<i>é</i> ‘is’	<i>rei</i> ‘king’	<i>três</i> ‘three’	<i>dois</i> ‘two’
<i>pé</i> ‘foot’	<i>pau</i> ‘stick’	<i>mal</i> ‘evil’	<i>pois</i> ‘then’
<i>dó</i> ‘sorrow’	<i>boi</i> ‘ox’	<i>mar</i> ‘sea’	
<i>ré</i> ‘re [musical note]’	<i>mau</i> ‘bad’	<i>cal</i> ‘lime’	
<i>mi</i> ‘mi [musical note]’	<i>sei</i> ‘(I) know’	<i>cor</i> ‘colour’	
<i>fã</i> ‘fa [musical note]’	<i>sou</i> ‘(I) am’	<i>faz</i> ‘(it) makes’	
<i>lá</i> ‘la [musical note]’	<i>teu</i> ‘your(s)’	<i>par</i> ‘even’	
<i>si</i> ‘ti [musical note]’	<i>céu</i> ‘sky’	<i>ter</i> ‘to have’	
<i>pá</i> ‘shovel’	<i>nau</i> ‘ship’	<i>sul</i> ‘south’	
<i>sé</i> ‘cathedral’	<i>lei</i> ‘law’	<i>sal</i> ‘salt’	
<i>pó</i> ‘dust’			
<i>há</i> ‘there is’			
<i>fé</i> ‘faith’			
<i>dá</i> ‘gives’			
<i>cá</i> ‘here’			
<i>tu</i> ‘you’			
<i>nu</i> ‘naked’			

**Source:** author’s elaboration.

At first sight, these examples could lead us to conclude that EP would belong to the set of languages for which MC, as formulated by McCarthy and Prince (1995) (see (1)), does not play any role as a phonotactic constraint or as a wordiness condition. In fact, according to such argument, if EP does not admit quantity-based vowel oppositions, its phonology should not allow any phonetic chains with less than two syllables. This is the main reason for several authors to refute the relevance of MC in Portuguese, as summarized in the aforementioned quotation from Vigário, Martins and Frota (2005).<sup>10</sup>

These conflicting, contradicting phonological interpretations feed the discussion that will be developed in the following sections of this paper.

<sup>9</sup> OBS.: 1. This table includes as “words” inflected forms of verbs and nouns. 2. In the examples of segmentally filled codas, we excluded all possible cases of inflected forms with a non-lexical final /S/. 3. VG diphthongs were assumed as branched nuclei. 4. In the present study, we deliberately exclude nuclei with phonetically nasalized vowels. 5. This table does not include forms traditionally described or classified as “clitics”.

<sup>10</sup> Other interpretations, referred to by the same study that is quoted here, see the scarcity of this kind of words – seen as *exceptional, marked* cases, then – as an argument to go on accepting MC as a wordiness condition in Portuguese. This is the explanation found in Booij (2004), as cited by Vigário, Martins and Frota (2005) (see quotation in Introduction).

## Syllable weight and wordiness in Portuguese

In our opinion, a complete evaluation of the effectiveness of MC in the phonology of Portuguese requires a deeper examination of dimensions that cannot be reduced to the minimum number of syllables per word. Thus, we shall pay attention, in the next paragraphs, to the specific question of syllable weight and its relation to wordiness in Portuguese, bearing in mind the connection between such property and MC, according to the formulation of this constraint as found in (1) and McCarthy and Prince's (1995) thoughts on the distinction between the language types under consideration there.<sup>11</sup>

The importance of syllable weight in Portuguese will be examined here based on several arguments related to two different phonological phenomena: stress-assignment and historical evidence attested in Medieval Portuguese (MP), that is to say, the cases of Latin etyma that almost invariably originated heavy monosyllables.

### Syllable weight and word stress

Traditionally, phonological descriptions of EP consider syllable weight as a neglectable factor in domains where it is accepted to operate quite systematically as far as other languages are concerned, as it is the case of stress-assignment. Studies such as Pereira's (1999), Roca's (1999), Mateus and D'Andrade (2000) and Mateus et al.'s (2003), among others, describe word stress in Portuguese as exclusively governed by morphological conditioning: nouns and verbs follow different stress patterns; within each word class, stress-assignment is sensitive to particular morphemes or morphological/morphosyntactic combinations. As we said before, this is the most consensual viewpoint regarding word-stress assignment among phonologists dealing with EP.

It is worthy to confront this interpretation, though, with some arguments provided by alternative explanations that do not exclude so categorically the interference of syllable weight on word-stress assignment. Brandão de Carvalho (1988; 1989; 2011), focusing on European Portuguese, as well as Wetzels (2007), considering the Brazilian varieties of the language, are among the authors who emphasize the inexistence of words stressed on the antepenultimate with a heavy syllable on the penultimate position, in addition to the statistical predominance of last-syllable-accented words with a heavy final syllable, as two key arguments in favour of the *phonological conditioning* of word stress in Portuguese and its sensitivity to syllable weight, mainly in the class of nouns and adjectives. This makes it a little more difficult to categorize Portuguese within the set of languages whose phonology does not admit the importance of syllable weight, a very pertinent criterion for the typological distinction proposed by McCarthy and Prince (1995) in (1). It should also be highlighted here, in accordance with Wetzels (2007), that quantity-based vocalic oppositions are not the sole evidence of syllable

---

<sup>11</sup> See quotation in note 7.

weight effectiveness in the phonology of any language, contrarily to what could be inferred from a quicker interpretation of McCarthy and Prince (1995).

### Syllable weight and heavy-monosyllable words in Medieval Portuguese

Our second argument in favour of the phonological importance of syllable weight in Portuguese stems from language diachrony. In Medieval Portuguese (MP), a non-neglectable number of words consisting of one sole heavy syllable is found. Tables 2 and 3 show us a few examples of such type of words. Quite interestingly, we could notice that many of such words descend from Latin monosyllabic etyma (see Table 2). In our view, however, a much greater interest comes from the observation of the cases found in Table 3, where the Portuguese heavy monosyllable (=monosyllabic word) is not explainable through etymology, as these words cannot be viewed as the historical inheritance of Latin monosyllables. Instead, they derive from other prosodic word formats; what is more, their formation in Medieval Portuguese preserves – in some cases, it even introduces – certain phonological structures that were generally deleted in regular historical changes from Latin to Portuguese. This is the case of heavy monosyllabic words which preserved final nasals (lost in Vulgar Latin itself) – e.g., *matre* > *mãi* ‘mother’, *fine* > *fim* ‘end’ – or an intervocalic lateral alveolar, as illustrated by examples such *sale* > *sal* ‘salt’, *sole* > *sol* ‘sun’). A plausible explanation for these exceptional cases of marked structure preservation could be the need for Portuguese to comply prosodically with MC prohibiting the existence of words containing less than one heavy syllable. This could allow us to postulate a MC constraint for MP such as found in (2); accordingly, we should accept that constraint (1a) would be available and mandatory at this early stage of Portuguese (the same postulate would allow us to reject the existence of quantity-based distinctions in Medieval Portuguese, too).

**Figure 2** – Medieval Portuguese words corresponding to heavy monosyllables descendant from Latin heavy monosyllables too; their preservation could be explained as a way of satisfying Minimality Condition |MP: { $\omega \geq \sigma_H$ }|

Lat. <i>sum</i> > MP <i>sam</i> ‘(I) am’
Lat. <i>cum</i> > MP <i>com</i> ‘with’
Lat. <i>non</i> > MP <i>nom</i> ‘no’
Lat. <i>sic</i> > MP <i>sim</i> ‘yes’

**Source:** author’s elaboration.



**Figure 3** – Medieval Portuguese words corresponding to heavy monosyllables that do not descend from Latin heavy monosyllables; their formation as such is probably due to the need of satisfying Minimality Condition [MP:  $\{\omega \geq \sigma_H\}$ ]

Lat. <i>patre</i> > MP <i>pae</i> ‘father’
Lat. <i>matre</i> > MP <i>māi</i> ‘mother’
Lat. <i>fine</i> > MP <i>fim</i> ‘end’
Lat. <i>bene</i> > MP <i>bem</i> ‘well’
Lat. <i>sale</i> > MP <i>sal</i> ‘salt’
Lat. <i>sole</i> > MP <i>sol</i> ‘sun’

**Source:** author’s elaboration.

### Other phonotactic constraints related to wordiness requirements in Portuguese

In previous sections, CM and syllable weight were given particular emphasis in our discussion of wordiness conditions of EP, in view of the special relation of such constraints to the central questions of this study.

Nonetheless, as many other studies on the same subject have highlighted (see, a.o., ELORDIETA, 2014, VELOSO, 2016), other phonological constraints apply at the word-level. These will be the focus of the following sections of our paper, emphasizing those that are more relevant for our main purpose.

In (3), we give a short summary of the main phonological constraints that, in languages like Portuguese, a morphological word has to comply with, mandatorily:

(3). Phonological constraints mandatorily observed by any phonetic chain admissible as *morphological words* in Portuguese.

(3a). One and one only primary stress, in full compliance with the stress-assignment prosodic rules (with the notable exception of Clitics).

(3b). Strict accordance with the phonotactic rules of the language and with the Structure Preservation Principle (KIPARSKY, 1985).

As for the phonological constraint found in (3a), let us recall that stress-assignment is one of the major phonological processes that take the word as a domain of application; therefore, it contributes for the identification/delimitation of this linguistic unit AND for the distinction between *phonological words* (bearing one, and not more than one only, primary stress), *clitic words* (inherently unstressed) and prosodically more complex, longer units, such as the *clitic group* (comprehending more than one single morphological word, but bearing just one single primary stress, according to Vigário (1998; 2003; 2007; 2010), Bisol (2000; 2004; 2007), Veloso (2012; 2013) and others, as far as Portuguese is concerned) or the *prosodic word group* (grouping different phonological words and clitics in one single, complex, longer prosodic unit, according, for instance, to VIGÁRIO; FERNANDES-SVARTMAN, 2010).

In relation to the phonotactic rules mentioned in (3b), we highlight here the fact that candidates to wordiness in EP should also obey a specific set of rules governing the selection and combination of segments. Such rules correspond, basically, to the Syllabification Algorithm of Mateus and D'Andrade (2000, p.60-64). In short, these rules specify that Portuguese words be in accordance with the basic principles found in (4).

(4). Major phonotactic constraints of contemporary EP (mainly, MATEUS; D'ANDRADE, 2000, p.60-64)

(4a). *The syllable nucleus is primarily filled by vocoids.* (Exceptionally, “empty nuclei” are admitted, for structures not respecting (4b).)

(4b). *Branched onsets are strictly governed by Sonority Principle and Dissimilarity Condition.*

(4c). *Segmental and autosegmental filling of the coda is highly restricted in this language.* (Word-finally, coda restrictions can be somehow alleviated – VELOSO, 2009; 2010): inexistence of branched codas; coda consonants admitted in coda position limited to the subset of /r l ([ʃ]) S/.

### **The importance of abbreviations' oralization for the assessment of Minimality Condition**

In the following paragraphs, it is our aim to make it explicit why abbreviations' oralization was chosen as a way to assess the importance of MC as a wordiness condition in Portuguese. As said in the Introduction, the same methodological choice was followed by Plénat (1993) in his study about French; Plénat's (1993) methodology is the main basis for the empirical study that will be presented in section 4.

We come back, at this point of our study (see Table 4), to part of the information that was given before regarding the two main possibilities of abbreviations' oralization in Portuguese – initialization and acronymization. Current examples of EP will be also given.

The cases under A in Table 4 disallow the acronymization modality, since all of them contain gross violations of the major phonotactic rules of Portuguese like the ones that are given in (4): for example, if these potential phonetic chains do not seem to include vocoid nuclei, their prosodization as *acronyms*<sup>12</sup> becomes impossible; this makes it irrelevant to check any application of MC in their eventual oralization. For this reason, cases like those of column A, Table 4, are excluded from further analysis.

Examples under B and C of Table 4 correspond to a distinct case: both situations respect the constraint of vocoid-filled nuclei, and this makes their production as

---

<sup>12</sup> See the arguments that are summarized in note 3 for the acceptance of abbreviated forms as particular cases of phonological words.

acronyms possible. The question that arises from a careful comparison between both situations is as follows: if, due to the respect of that constraint, both types of candidates can be oralized as acronyms, how can we explain that this does not occur with the examples of A as well as with the examples of B? In fact, the examples under C are commonly produced as acronyms; notwithstanding, the examples under B – which could also be produced as acronyms due to the presence of potential vocalic nuclei – are mainly produced as initialisms. Therefore, *the phonotactic constraint that requires vocalic nuclei seems to be a necessary – yet not sufficient – condition for acronymization*. This observation renders research on the phonological conditions of wordiness in Modern EP more necessary. Previous research has suggested that a considerable multiplicity of phonological and prosodic variables may be implied in the so-called “marginal” processes of word formation (see, a. o., the work on Brazilian Portuguese by ARAÚJO, 2002; CHACON, 2004; TENANI, 2004; GONÇALVES; LIMA; THAMI DA SILVA, 2005; ABREU; ROSA, 2006; LIMA, 2008; ABREU, 2009; LIMA; THAMI DA SILVA; GONÇALVES, 2009; CUNHA, 2012, as well as recent research regarding EP – FERREIRA, 2012; PEREIRA, 2013). Evidence has been gathered from data of language acquisition, children’s early written production, abbreviation procedures, hypocoristic formation and other modalities of truncation. In this study, attention will be paid only to abbreviation procedures as a way to assess MC in Portuguese.

It should be borne in mind that, in the cases of column C of Table 4, C2 represents a specific subset of examples which, in spite of respecting the constraint of vocalic nuclei (see (4a)), do not agree with other important phonotactic constraints of EP, since they show codas potentially filled by obstruents or by sonorants unallowed in coda position. These are the cases, e.g., of FENPROE, [fɛn'prɔf], and REN, ['rɛn]). The formation of this particular type of acronyms has been taken into consideration by previous studies (see, e.g., PEREIRA, 2013) and might be explained, for instance, by the frequent phonetic deletion of final unstressed vowels in contemporary EP. This originates, at the phonetic level, the high frequency of occurrence of this type of “irregular” codas. Such result could be interpreted as a manifestation of their representation in the speakers’ phonological knowledge as “regular” syllables” and viewed as a first hint of their possible on-going phonologization in the current stage of the language. Just as in the cases under A in Table 4, these examples will not be considered further on in the present study, since they raise specific questions that do need the prior understanding of the more general, central topics that constitute the central research of our study.

**Figure 4** – Initialization and acronymization of current abbreviated forms in Contemporary European Portuguese<sup>13</sup>

CURRENT ABBREVIATIONS (CEP)		
Unadmitted as words of the language, in view of their violation of the mandatoriness of a vowel-filled syllabic nucleus	Admitted as words: full respect of the phonotactic constraints of European Portuguese	
<b>A: INITIALISMS</b>	<b>B: INITIALISMS</b>	<b>C: ACRONYMS</b>
FCT (=Fundação para a Ciência e a Tecnologia, ‘Foundation for Science and Technology’) : [ˈɛfiˈseˈte] ; *[fkt] CGTP (=Confederação Geral dos Trabalhadores Portugueses, ‘General Confederation of the Portuguese Workers’) : [ˈseˈzeˈteˈpe] ; *[kgtp]	BE (=Bloco de Esquerda, ‘Leftist Party’) : [ˈbeˈɛ] SU (=Serviço de Urgência, ‘Emergency Unit’) : [ˈesiˈu] (ou [ˈɛˈsjɔ])	<b>C1: Without violation of any phonotactic rules</b> BES (=Banco Espírito Santo, ‘Espírito Santo Bank’) : [ˈbɛʃ] NATO (=North Atlantic Treaty Organization) : [ˈnatu] <b>C2: Coda filled by an unadmitted sonorant or obstruent in this prosodic position</b> FENPROF (=Federação Nacional dos Professores, ‘National Federation of Teachers’) : [fɛnˈprɔf] REN (=Rede Elétrica Nacional, ‘National Electric Network’) : [ˈrɛn]

Source: author’s elaboration.

On the basis of these assumptions, we hypothesized that speakers, when faced with the task of reading letter sequences presented as abbreviations, could follow one of two possible paths:<sup>14</sup>

- **Option 1.** The speaker recognizes the chain as a current abbreviation of his/her language (as it the case of the abbreviations found in Table 4 for native speakers of CEP). In a way, these abbreviations correspond to lexicalized words; thus, they are processed as any other word pertaining to the speaker’s lexical inventory (although, as mentioned before, individual lexical repertoires vary considerably). Therefore, the main processing way activated by the speaker in the oralization of these stimuli is their **lexical knowledge** (which is conceived of here as the set of linguistic information related to all entries of the individual’s “mental lexicon”, which includes indications about the phonetic form of such entries);

<sup>13</sup> OBS.: 1. All acronyms of this table are often used in EP. So, they are assumed as potential lexical entries of the individual lexicons of a great number of speakers of the language. 2. Phonetic transcriptions (in this table and in all passages of our text) assume that each initialism contains more than one prominence; since the prosodic status of these units is not the main focus of our study, we do not make any distinction, within each initialism, between primary- and secondary-stress bearing syllables. 3. Each acronym is assumed as one sole prosodic word (bearing, thus, just one main stress).

<sup>14</sup> See Abreu (2009, p.48 ss.) for a broader, more detailed discussion on the various psycholinguistic mechanism that are activated by the processing of written abbreviations.

- **Option 2.** In the case of an unknown abbreviation, the speaker must activate other resources of his/her implicit linguistic knowledge, namely his/her phonological knowledge (PhK). Excluding the cases of column A, Table 4 – which would never be oralized as acronyms, for the reasons that were explained above –, we assume those abbreviations could be read either as acronyms OR as initialisms as the result of the activation of constraints such as MC, plausibly one of the main responsible factors, in the speakers' PhK, for the wordiness evaluation of any given phonetic chain.

Option 2 admits that the competing oralization procedures (acronymization vs. initialization) are not selected at random. Following Plénat's (1993) proposal for French, it is admitted that minimality is the factor that determines the choice of either strategy. This will be the basic assumption for our empirical research, even though other phonological variables are not excluded.<sup>15</sup>

This assumption seems to be specially enhanced if we compare, for instance, one of the pairs found in Table 4. Indeed, if we compare the current oralization of "BE" ( $\sigma_L$ ) as an initialism ( $=[b'e'\epsilon]$ ) and the oralization of an abbreviation which differs from this one in just one segment and in syllable weight – "BES" ( $\sigma_H$ ), oralized as an acronym ( $=[b'e\zeta]$ ) –, we could infer that syllable weight is, indeed, the key factor eliciting the Portuguese speakers' preference for the acronymization (i.e., for the oralization of abbreviations as phonological words: these will emerge only with abbreviations with a syllable weight equaling  $\sigma_H$ ). This inference would lead us to accept constraint (2) (postulated for Medieval Portuguese) as still operational in CEP, too, somehow contradicting McCarthy and Prince (1995)'s constraint (1), which limits the importance of syllable weight to quantity-sensitive languages.

All these questions will be assessed by the empirical study that will be presented in the next section.

## Empirical study

### Preliminary remarks

In order to understand the linguistic processing of phonetic chains candidate to be treated as words and to gather evidence regarding the role of MC in Portuguese, we developed the empirical study that will be presented in this section. A group of monolingual speakers of CEP was asked to read a list of 88 abbreviations specifically prepared for this experiment. As far as we could know, these abbreviations were not previously existent – they were assumed thus as unavailable to the subjects' linguistic knowledge. We will refer to them as "*pseudoabbreviations*". The use previously

---

<sup>15</sup> Other methodologies could have been followed, of course. Abreu and Rosa (2006) e Abreu (2009), e.g., in their studies of the linguistic processing of this kind of stimuli by native speakers of Brazilian Portuguese, take into consideration data obtained from grammaticality judgments and from latency times of stimuli reading. Such studies conclude that variables such as frequency and familiarity also play a role on the processing of written abbreviations.

unattested stimuli was chosen in order to cancel the interference of any level of lexical knowledge (that could comprehend any information regarding the phonological representation of such forms, acceptable as lexical entries, comparable to the ones which are found in Table 4). That is to say, it was our intention to make the subjects follow mandatorily the processing path previewed by Option 2 (see above), i.e., we wanted to activate the subjects' *phonological knowledge* as their only means of processing these stimuli. Only the isolation of this processing strategy could lead us to clear information about the role of MC and other phonological constraints on the phonological representation of wordiness conditions among Portuguese speakers.

## Hypotheses

The three following hypotheses were formulated.

(5). Experimental Hypotheses:

**Hypothesis A:** Monosyllables corresponding to a heavy syllable (and respecting the language's phonotactic rules), when presented as written abbreviations, will be oralized as ACRONYMS.

**Hypothesis B:** Monosyllables corresponding to a light syllable, when presented as written abbreviations, will be oralized as INITIALISMS.

**Hypothesis C:** Disyllables and trisyllables, when presented as written abbreviations, will be oralized as ACRONYMS.

Hypotheses A and B, related to the processing of monosyllabic stimuli only, are motivated by our previous analysis of data shown in Table 4 (see section 3).

As for Hypothesis C, dealing with the processing of stimuli with 2 or more syllables, its motivation is found in de McCarthy and Prince's (1995) Minimality Condition (see (1b)).

## Method

### Participants

Subjects were 12 adult (8 women, 4 men), monolingual, native speakers of CEP, mostly speakers of the Northern dialects of this language. All of them are undergraduate students at the University of Porto. Their participation was free. None was previously informed of the specific purposes or questions of the study. Participants' age ranged from 16 to 60 years old (mean=38 years; SD=16,06 years).

## Linguistic material and procedure

### Linguistic material

Pseudoabbreviations were embedded in a set of 130 sentences bearing a syntactic and lexical invariant structure: the only spot that varied was where abbreviations were inserted. In 88 sentences, such abbreviations corresponded to the aforementioned pseudoabbreviations. The remaining 42 sentences included abbreviations assumed as common in everyday language for a native speaker of European Portuguese. In all sentences, grammatical gender was attributed to the abbreviated form: in the case of current abbreviations, the current grammatical gender was maintained in the corpus. As abbreviations and pseudoabbreviations were always embedded as the nucleus of a noun phrase, gender agreement was overtly given by the definite article (“masculine”/“feminine”) determining the abbreviation.

The structure of the carrier sentence of our stimuli is given in (6).

(6). Structure of the carrier sentences used as stimuli

O Presidente d(x) [Z] ABREVIACÃO/PSEUDOABREVIACÃO demitiu-se.

The\_Masc\_Sg President of(x) [ Z ]  
ABBREVIATION/PSEUDOABBREVIATION has quit.

(x=definite article, singular, masculine or feminine)

[Z]=word or expression that, in certain sentences, was introduced for the sake of semantic plausibility)<sup>16</sup>

The complete list that subjects were asked to read is found in (7). The order in which stimuli appear in this list corresponds to the order of presentation of all sentences to each participant; it was randomly obtained during the preparation of the linguistic material.

Including abbreviations *and* pseudoabbreviations in the stimuli list was meant to create a distractor effect so that participants were induced to oralize each abbreviation/pseudoabbreviation either as an initialism or as an acronym according to their own lexical knowledge (in the case of abbreviations) or to their implicit phonological knowledge (in the case of pseudoabbreviations). That is to say, we aimed at eliciting more than one single path in the processing of the stimuli, assuming that participants would choose the most adequate way of processing each stimulus according to its linguistic status.

Thus, each participant was asked to read a set of 130 sentences; 88 sentences contained written chains corresponding to pseudoabbreviations. 42 sentences contained abbreviations which are current among speakers of Portuguese – assumed as probably

---

<sup>16</sup> Example: “*O Presidente da Comissão da CREL demitiu-se.*”

previously known by most of the participants, then<sup>17</sup>. These current abbreviations – containing cases of sequences in accordance with the phonotactic rules of the language and cases that violate such rules – comprehend 23 current initialisms and 19 current acronyms.

As for the 88 pseudoabbreviations found in the corpus, all respected the phonotactic rules of Portuguese as formulated in (3) and (4). That is to say, all pseudoabbreviations of the stimuli list – on which our analysis will be exclusively focused – correspond to the cases under B and C1 of Table 4. The more debatable cases of A and C2 of Table 4 were then excluded from the present study.

In our analysis, we will pay attention to the results regarding the oralization of pseudoabbreviations only. In (7), such pseudoabbreviations are found in the underlined written sentences; it should be borne in mind, however, that the materials given to all participants during the experiment did not in any manner emphasise or differentiate pseudoabbreviations from supposedly real abbreviations. All instructions and materials were absolutely clueless regarding this matter.

---

<sup>17</sup> In fact, it is rather difficult to determine which current abbreviations do or do not belong to the specific individual lexicon of an individual speaker, due to the aforementioned variability among individual lexicons. Similarly, we cannot exclude categorically that some of the pseudoabbreviations that were coined on purpose for our study did not belong to the individual lexicon of some of our participants: for instance, they could correspond to the current designation of small, local companies or institutions eventually familiar to some or few of our subjects. For the list of (7), our own linguistic intuition and knowledge was strongly taken into consideration. In addition, for each (real) abbreviation that was assumed as current in Portuguese, we looked for at least one occurrence in any printed materials, in the internet and/or in general dictionaries of the language. The pseudoabbreviations were explicitly forged on the basis of the phonological variables that we wanted to test; their inoccurrence in the same sources (general press, internet, dictionaries) was also controlled.



**Figure 5** – List of stimuli presented to the participants<sup>18</sup>

1 - O Presidente da RDIS demitiu-se.	41 - O Presidente da ONU demitiu-se.
2 - <u>O Presidente do RU demitiu-se.</u>	42 - O Presidente da RTP demitiu-se.
3 - <u>O Presidente do Instituto EUPA demitiu-se.</u>	43 - O Presidente do PS demitiu-se.
4 - <u>O Presidente da UISMA demitiu-se.</u>	44 - <u>O Presidente do SOUGA demitiu-se.</u>
5 - O Presidente do BPN demitiu-se.	45 - O Presidente da AIP demitiu-se.
6 - O Presidente da SPA demitiu-se.	46 - <u>O Presidente do GA demitiu-se.</u>
7 - <u>O Presidente da FAS demitiu-se.</u>	47 - <u>O Presidente do GREL demitiu-se.</u>
8 - O Presidente do PSD demitiu-se.	48 - <u>O Presidente do LAIRA demitiu-se.</u>
9 - O Presidente da UEFA demitiu-se.	49 - O Presidente da NATO demitiu-se.
10 - O Presidente do SITAVA demitiu-se.	50 - O Presidente da CCVM demitiu-se.
11 - O Presidente da Comissão da VCI demitiu-se.	51 - <u>O Presidente do PRICA demitiu-se.</u>
12 - <u>O Presidente da FAIMA demitiu-se.</u>	52 - O Presidente da Comissão do SIGARRA demitiu-se.
13 - <u>O Presidente do BLO demitiu-se.</u>	53 - <u>O Presidente do FLEU demitiu-se.</u>
14 - O Presidente do SINDEPO demitiu-se.	54 - <u>O Presidente da CO demitiu-se.</u>
15 - <u>O Presidente do CLAUS demitiu-se.</u>	55 - <u>O Presidente do CLARCO demitiu-se.</u>
16 - O Presidente da UGT demitiu-se.	56 - <u>O Presidente da ARCI demitiu-se.</u>
17 - <u>O Presidente da AU demitiu-se.</u>	57 - <u>O Presidente da FE demitiu-se.</u>
18 - <u>O Presidente do CRAITA demitiu-se.</u>	58 - <u>O Presidente da FIBA demitiu-se.</u>
19 - O Presidente da Direcção da CREL demitiu-se.	59 - <u>O Presidente do COIS demitiu-se.</u>
20 - <u>O Presidente do GATE demitiu-se.</u>	60 - <u>O Presidente da UNIVA demitiu-se.</u>
21 - O Presidente da FAP demitiu-se.	61 - <u>O Presidente do CLA demitiu-se.</u>
22 - <u>O Presidente da FLEUDO demitiu-se.</u>	62 - O Presidente da APE demitiu-se.
23 - O Presidente da FNE demitiu-se.	63 - <u>O Presidente da BLURMA demitiu-se.</u>
24 - <u>O Presidente do EUS demitiu-se.</u>	64 - <u>O Presidente da AUPE demitiu-se.</u>
25 - O Presidente da CGTP demitiu-se.	65 - O Presidente do GROISMI demitiu-se.
26 - <u>O Presidente da CLAUSBA demitiu-se.</u>	66 - O Presidente da GIU demitiu-se.
27 - O Presidente da CIP demitiu-se.	67 - O Presidente da ECOFIN demitiu-se.
28 - O Presidente do BE demitiu-se.	68 - <u>O Presidente da UIS demitiu-se.</u>
29 - <u>O Presidente do BLUR demitiu-se.</u>	69 - <u>O Presidente do GREMA demitiu-se.</u>
30 - O Presidente do CDUP demitiu-se.	70 - <u>O Presidente da AUSPA demitiu-se.</u>
31 - <u>O Presidente do GRALCO demitiu-se.</u>	71 - <u>O Presidente da AICA demitiu-se.</u>
32 - <u>O Presidente do CEUNA demitiu-se.</u>	72 - O Presidente da Fiscalização do IVA demitiu-se.
33 - <u>O Presidente da CIL demitiu-se.</u>	73 - <u>O Presidente do BLOLI demitiu-se.</u>
34 - O Presidente da EI demitiu-se.	74 - <u>O Presidente da FAUSPA demitiu-se.</u>
35 - O Presidente da ULMI demitiu-se.	75 - <u>O Presidente do IR demitiu-se.</u>
36 - O Presidente da CLAUPA demitiu-se.	76 - <u>O Presidente do SEI demitiu-se.</u>
37 - <u>O Presidente da EISTI demitiu-se.</u>	77 - <u>O Presidente da ASPO demitiu-se.</u>
38 - O Presidente da CDU demitiu-se.	78 - O Presidente do BES demitiu-se.
39 - <u>O Presidente da CRAI demitiu-se.</u>	79 - <u>O Presidente da FEU demitiu-se.</u>
40 - O Presidente da FEUP demitiu-se.	80 - <u>O Presidente do FRISPE demitiu-se.</u>
	81 - O Presidente da UP demitiu-se.

<sup>18</sup> OBS.: In this list, the underlined sentences correspond to those where pseudoabbreviations were embedded. In the materials given to the participants, no formal distinction was made between these sentences and those that supposedly carried current abbreviations.

82 - <u>O Presidente da FREIS demitiu-se.</u>	106 - <u>O Presidente da TERPE demitiu-se.</u>
83 - <u>O Presidente do PRI demitiu-se.</u>	107 - <u>O Presidente do GALVE demitiu-se.</u>
84 - <u>O Presidente da OITI demitiu-se.</u>	108 - <u>O Presidente do LITE demitiu-se.</u>
85 - <u>O Presidente do SIR demitiu-se.</u>	109 - <u>O Presidente do PRUI demitiu-se.</u>
86 - <u>O Presidente do REISTA demitiu-se.</u>	110 - <u>O Presidente do GOISPA demitiu-se.</u>
87 - <u>O Presidente da ESPE demitiu-se.</u>	111 - <u>O Presidente do PRUILE demitiu-se.</u>
88 - <u>O Presidente da FLEUSTA demitiu-se.</u>	112 - <u>O Presidente da FREISPO demitiu-se.</u>
89 - <u>O Presidente da GNR demitiu-se.</u>	113 - <u>O Presidente da Comissão da IVG demitiu-se.</u>
90 - <u>O Presidente da OL demitiu-se.</u>	114 - <u>O Presidente da AS demitiu-se.</u>
91 - <u>O Presidente do CESP A demitiu-se.</u>	115 - <u>O Presidente da FUIS demitiu-se.</u>
92 - <u>O Presidente da FIFA demitiu-se.</u>	116 - <u>O Presidente do GEUS demitiu-se.</u>
93 - <u>O Presidente da OU demitiu-se.</u>	117 - <u>O Presidente da AUS demitiu-se.</u>
94 - <u>O Presidente da PAC demitiu-se.</u>	118 - <u>O Presidente da FRIS demitiu-se.</u>
95 - <u>O Presidente do PRU demitiu-se.</u>	119 - <u>O Presidente do CLADA demitiu-se.</u>
96 - <u>O Presidente do Instituto da CRIL demitiu-se.</u>	120 - <u>O Presidente do NEFA demitiu-se.</u>
97 - <u>O Presidente da DAUS demitiu-se.</u>	121 - <u>O Presidente do SASU demitiu-se.</u>
98 - <u>O Presidente da AL demitiu-se.</u>	122 - <u>O Presidente do OISCI demitiu-se.</u>
99 - <u>O Presidente da AI demitiu-se.</u>	123 - <u>O Presidente da RAI demitiu-se.</u>
100 - <u>O Presidente da CLAU demitiu-se.</u>	124 - <u>O Presidente da SORPI demitiu-se.</u>
101 - <u>O Presidente do Gabinete do IRS demitiu-se.</u>	125 - <u>O Presidente da PT demitiu-se.</u>
102 - <u>O Presidente do OIS demitiu-se.</u>	126 - <u>O Presidente da PLAIS demitiu-se.</u>
103 - <u>O Presidente da URSS demitiu-se.</u>	127 - <u>O Presidente da FER demitiu-se.</u>
104 - <u>O Presidente da FAUSTE demitiu-se.</u>	128 - <u>O Presidente do GROUS demitiu-se.</u>
105 - <u>O Presidente da FENPROF demitiu-se.</u>	129 - <u>O Presidente da PSP demitiu-se.</u>
	130 - <u>O Presidente do PLAR demitiu-se.</u>

Source: author's elaboration.

The 88 pseudoabbreviations contained in the corpus are prosodically controlled: monosyllables for syllable weight; polysyllables for syllable type of initial syllable and syllabic extension (number of syllables). As for syllable type, all unmarked syllable formats of Portuguese were included, either as the only syllable of monosyllabic pseudoabbreviations or as the initial syllable of polysyllabic pseudoabbreviations. In these positions, nasal vowels were excluded. As for syllabic extension, two sets of pseudoabbreviations were created, so that we could obtain results with monosyllables and results with di- and trisyllables, in order to test the validity of (1b) for EP. All stimuli whose written form could potentially elicit the production of one single vocalic nucleus ( $\cong$  one single written vowel) were assumed here as monosyllabic; stimuli whose written forms contained more than one potential vocalic nucleus ( $\cong$  two or more written vowels) were assumed as polysyllabic.

Therefore, we aimed at gathering information on the roles of *syllable weight* (see Hypotheses A and B) and *syllabic extension* (see Hypothesis C) on the definition of wordiness conditions in Portuguese.

Crossing these 3 variables throughout our corpus, the 88 pseudoabbreviations can be categorized as it is done in Table 5.

**Figure 6** – Categorization of the 88 pseudoabbreviations of the corpus according to the variables SYLLABLE WEIGHT, SYLLABLE TYPE and SYLLABIC EXTENSION

SYLLABLE WEIGHT (monosyllabic pseudoabbreviations and initial syllable of polysyllabic abbreviations)	SYLLABLE TYPE <sup>19</sup> (monosyllabic pseudoabbreviations and initial syllable of polysyllabic abbreviations)	SYLLABIC EXTENSION
Heavy syllables: 20 Light syllables: 68	V: 2 CV: 9 VC: 6 VG: 8 VS: 1 VGS: 7 CVC: 6 CVG: 7 CVS: 2 CVGS: 8 CCV: 9 CCVC: 6 CCVS: 2 CCVG: 7 CCVGS: 8	Monosyllabic: 40 Di-/Trisyllabic: 48

Source: author's elaboration.

## Method

Each subject was tested individually in a silent room at the Faculty of Arts of the University of Porto (Portugal).

All sentences were presented in written form and displayed on A4 sheets, printed out in Times New Roman, 14 pt. type, double-spaced. Each abbreviation was capitalized, without any signs separating the constituting letters. Sentences were sequentially numbered, from 1 to 130, with Arabic numerals.

Each subject was asked to read each sentence as naturally as s/he could. All were previously informed that all sentences had the same structure; this structure was explicitly explained beforehand. In the pre-test session, each participant was also told that sentences would vary in just one point of the word-chain, after "Presidente de" ('president of'). They were then informed that, after this sequence, they would be shown an "abbreviation". The instruction given at this point was as follows: if they recognized such abbreviation, they should read it according to the usual form they pronounced it. They were also told that many abbreviations could be unknown, since they designated small companies or institutions which were not very well-known. This precision was made in order to prepare the participants for the oralization of pseudoabbreviations.

<sup>19</sup> (\*): C=Consonant (contoid); V=Vowel (vocoid); G=Glide; S=Coronal Fricative (corresponding to written <S>).

They were asked to read such abbreviated forms so that they sounded “more natural” to them. In the beginning of each session, participants were given a brief explanation of the two possible ways of “reading” an abbreviated form (either as an initialism or as an acronym, without using this explicit terminology). Participants were allowed to make pauses during their reading sessions whenever they felt like it. Each individual session lasted about 15 minutes, approximately. Participants’ productions of abbreviated forms were immediately transcribed in IPA by the experimenter as the reading session was going on.

## Results

Table 1 displays the results for each syllable type, both with monosyllabic and polysyllabic pseudoabbreviations.

Each cell shows the mean number (and standard deviation) of the oralization of the 88 pseudoabbreviations, either as initialisms or as acronyms. Values were obtained from the 1056 recorded productions (=88 pseudoabbreviations X 12 participants). Both types of response (initialisms vs. acronyms) are statistically compared separately for monosyllabic and di-/trisyllabic stimuli. Results are restricted to the cases for which it was possible to obtain a statistical comparison on the basis of a T-Student test. As for the variables *syllable type* and *syllable weight* of disyllables and trisyllables, results regard first syllable only.

**Table 1** – Mean (and SD) and statistical comparison (T-Student Test) of the participants’ responses for each syllable type. Mono-, di- and trisyllabic pseudoabbreviations. For di- and trisyllabic pseudoabbreviations, analysis takes into consideration responses regarding the first syllable only. Variables considered: *syllable type*, *syllable weight*<sup>20</sup>

	MONOSYLLABLES			DISYLLABLES/TRISYLLABLES		
	<b>Initialism:</b> Mean (Standard Deviation)	<b>Acronym:</b> Mean (Standard Deviation)	T-Student	<b>Initialism:</b> Mean (Standard Deviation)	<b>Acronym:</b> Mean (Standard Deviation)	T-Student
<b>LIGHT SYLLABLES</b>						
CV	2,67 (1,56)	1,33 (1,56)	t(11)=1,483, n.s.	0,17 (0,39)	3,83 (0,39)	<b>t(11)=-16,316,</b> <b>p&lt;0,000</b> **
CCV	1,33 (1,37)	2,58 (1,50)	t(11)=-1,512, n.s.	0,08 (0,29)	3,92 (0,29)	<b>t(11)=-23,000,</b> <b>p&lt;0,000</b> **

<sup>20</sup> Significance levels: \* p<0,05; \*\* p<0,005

OBS.: Results not allowing a statistical comparison based on a T-Student Test were excluded.

HEAVY SYLLABLES						
VC	3,58 (0,90)	0,42 (0,90)	<b>t(11)=-6,092,</b> <b>p&lt;0,000</b> **	0,75 (0,75)	3,25 (0,75)	<b>t(11)=-5,745,</b> <b>p&lt;0,000</b> **
VG	3 (1,54)	1 (1,54)	<b>t(11)=2,253,</b> <b>p&lt;0,05</b> *	1,25 (1,36)	2,75 (1,36)	t(11)=-1,915, n.s.
VGS	2,33 (1,37)	1,58 (1,37)	t(11)=0,950, n.s.	1 (1,28)	2,92 (1,24)	<b>t(11)=-2,653,</b> <b>p&lt;0,05</b> *
CVC	0,83 (0,83)	3,17 (0,83)	<b>t(11)=-4,841,</b> <b>p&lt;0,000</b> **	0,17 (0,58)	3,83 (0,58)	<b>t(11)=-11,000,</b> <b>p&lt;0,000</b> **
CVG	0,5 (0,67)	3,42 (0,67)	<b>t(11)=-7,705,</b> <b>p&lt;0,000</b> **	0,25 (0,87)	3,75 (0,87)	<b>t(11)=-7,000,</b> <b>p&lt;0,000</b> **
CVGS	0,75 (0,62)	3,25 (0,62)	<b>t(11)=-6,966,</b> <b>p&lt;0,000</b> **	0,17 (0,58)	3,83 (0,58)	<b>t(11)=-11,000,</b> <b>p&lt;0,000</b> **
CCVC	0,25 (0,45)	3,75 (0,45)	<b>t(11)=-13,404,</b> <b>p&lt;0,000</b> **	0,42 (1,16)	3,58 (1,16)	<b>t(11)=-4,710,</b> <b>p&lt;0,005</b> **
CCVG	0,33 (0,65)	3,67 (0,65)	<b>t(11)=-8,864,</b> <b>p&lt;0,000</b> **	0,33 (1,15)	3,58 (1,16)	<b>t(11)=-4,892,</b> <b>p&lt;0,000</b> **
CCVGS	0,25 (0,45)	3,75 (0,45)	<b>t(11)=-13,404,</b> <b>p&lt;0,000</b> **	0,25 (0,87)	3,75 (0,87)	<b>t(11)=-7,000,</b> <b>p&lt;0,000</b> **

Source: author's elaboration.

## Discussion

### Monosyllabic stimuli

Hypotheses A and B made us expect that light syllables (CV and CCV) would be preferably oralized as initialisms, whereas heavy syllables were expected to be oralized mainly as acronyms.

Results as shown in Table 1, however, do not seem to be completely conclusive as far as this initial assumption is concerned:

- CV monosyllables ( $=\sigma_L$ ) were, in fact, mainly oralized as initialisms, but the difference between this type of response and acronymization was not statistically significant;

- by the other hand, CCV monosyllables (light syllables, too) were very often oralized as acronyms, contrarily to what was previewed by Hypothesis B, although the difference between both responses was not significant;

- finally, it should be noticed that the only cases where a comparison between both types of response (initialism vs. acronym) is not significant are found in light syllables and with VGS-syllables: with other syllable types ( $=\sigma_H$ ), results always show a clear statistical preference towards either of response.

Such preference, for monosyllabic stimuli only, is as follows:

-with (heavy) syllables with the structure VC, VG and VGS (for the former, without a statistically significant difference between the two types of response), *initialisms* are statistically preferred over *acronyms* (contrarily to Hypothesis A);

-with syllables CVC, CVG, CVGS, CCVC, CCVG and CCVGS (also heavy syllables), *acronyms* are statistically predominant; contrarily to results obtained with VC-, VG- and VGS-syllables, these results could be interpreted as a confirmation of Hypothesis A.

Finally, it should be borne in mind that, within monosyllabic stimuli, the syllable type that reaches the higher score of oralizations as acronyms is VC, a *heavy syllable*; let us remind here that, according to Hypothesis A, acronymization was more expectable than initialization in these cases.

## Disyllabic and trisyllabic stimuli

As far as di- and trisyllabic stimuli are concerned, the predominant strategy that is followed by most participants (see Table 1) is acronymization. This seems to confirm Hypothesis C. For each syllable type, and just looking at the first syllable of each stimulus ( $\sigma_1$ ), the mean number of acronymizations is always statistically higher than the number of initializations (with one single exception: the disyllable where  $\{\sigma_1=VG\}$ , which attains the highest scores of acronymization, although the difference between this kind of oralization and initialism-type is not statistically significant).

\*\*\*\*\*

In our view, these results need further examination: more data, from a larger number of participants, would give us a clearer picture of the several, interacting variables.

Notwithstanding, a few general comments can be drawn. First, on the one hand, we could conclude that syllable weight effect is not entirely clear for determining which candidates are better or worse to be accepted as words in European Portuguese: heavy monosyllables are very often oralized as acronyms (see the results for CV and VG monosyllables, for instance), and this is not absolutely clear for a definite confirmation or refutation of Hypothesis A. On the other hand, light monosyllables were not systematically produced as initialisms, contrarily to the initial assumptions of

Hypothesis B: with CV stimuli, indeed, initialisms are predominant, but this is not the case for CCV (mostly produced as acronyms). In these stimuli, comparisons between both types of response were not statistically significant either.

If we confront these results with our initial hypotheses, the only hypothesis that seems minimally confirmed is Hypothesis C: when processing stimuli with two or more syllables, participants show a clear preference towards acronymization in detriment of initialisms.

This preliminary observation – not allowing a clear refutation or confirmation of Hypotheses A and C, but supporting, at least in a limited way, the acceptance of Hypothesis C – could be seen as empiric evidence in favour of McCarthy and Prince’s (1995) proposal as found in (1b): in languages without quantitative oppositions (as it is the case of CEP), Minimality Condition obeys the word’s syllabic extension, not syllable weight.

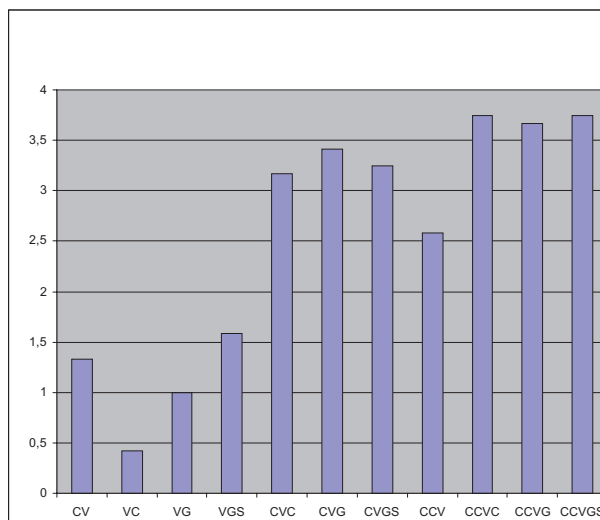
However, a deeper analysis of this rough conclusion brings us a few more supplementary arguments to be taken into consideration for us to understand it better how monosyllabic stimuli are processed by CEP native speakers. These arguments challenge in some ways the state of the art as well as the experimental hypotheses that were explored earlier in this paper. The most striking finding that such analysis provides us with has to do with the stimuli extension, not in terms of syllables, but in terms of segments.

Indeed, there seems to be a highly interesting, suggestive regularity regarding the possible effect of a Minimality Condition of any potential word in European Portuguese – *be that minimality be measured, though, in number of segments, not in syllable weight or number of syllables*. In our view, such regularity could explain, quite satisfactorily, why monosyllabic pseudoabbreviations are systematically oralized according to one single type of production. *Native speakers of CEP process phonetic chains with three or more segments preferably as acronyms, and process phonetic chains with less than three segments preferably as initialisms – regardless of such stimuli being mono- or polysyllabic, and regardless of syllable weight, too*. If this interpretation proves true, this is the explanation why stimuli like, e.g., “BLO” and “PRU”, are preferably oralized as acronyms ([blo], [pru]) and seldom as initialisms. This would lead us to postulate a **linear** MC constraint in CEP, contrary to Hypothesis B. The same explanation would apply to the results obtained with stimuli such as “AU” and “EI”. Although they correspond to heavy monosyllables, they were mainly oralized as initialisms ([a’u], [e’i]), not as acronyms, contrarily to Hypothesis A.

Figure 7, with results of the 40 monosyllabic pseudoabbreviations only, illustrates how the score of acronymizations increases as the stimuli extension measured in number of segments also gets higher: stimuli with just two segments (CV, VC, VG) show an average number of acronymizations also lesser than 1,5. Stimuli with 3 or more segments show higher scores of acronymization, even when the only case of light syllable within this set is considered: CCV, in spite of showing a lesser number of acronymizations than other stimuli with 3 or more segments, was processed as

an acronym more often than light or heavy monosyllables (formed by 2 segments only) were.

**Figure 7** – Mean number of oralizations of *monosyllabic* pseudoabbreviations as acronyms



Source: author's elaboration.

We could then postulate a Minimality Condition, which would apply specifically in CEP, clearly grounded on the number of segments of the word, regardless of syllable weight. This constraint is formulated in (8).

- (8). Minimality Condition in Contemporary European Portuguese  
 CEP: {WMinExtSg  $\geq$  3}  
 WMinExtSg=Word Minimal Extension, measured in number of segments

This constraint, in our view, would also explain results such as the following:

1 - Disyllables and trisyllables are produced, quite regularly, as acronyms, not as initialisms (as previewed by Hypothesis C and admitted by McCarthy and Prince (1995) Minimality Condition found in (1b));

2 - Monosyllables with just 2 segments are produced as initialisms, not acronyms, regardless of syllable weight (see Table 1 and Figure 7);

3 - Monosyllables with 3 or more segments are mainly produced as acronyms, not as initialisms, also regardless of syllable weight (see again Table 1 and Figure 7).

We admit that a formulation like (8) does not explain the presence of Portuguese words that are formed just by a light monosyllable (see examples at 1<sup>st</sup> column of Table 1). In fact, it remains unanswered how a constraint such (8) is violated by this set of



words of the language, even though it consists of a not very large set. We cannot ignore, first, that the answer for this question, very crucial for the phonological description of Portuguese, cannot be categorically given by a study like this one and on the sole basis of the oralization of abbreviated forms. Notwithstanding, we will try to gather some possible explanations for this interrogation.

First of all, it should be reminded that these words form – as it is explicitly acknowledged by several authors who were previously mentioned – a very small subset of the Portuguese lexicon. In an oversimplified manner, they could be presented as *exceptions* to MC in Portuguese, in line with interpretations such as Vigário's (2003) and Booij's (2004), for instance.

Additionally, it is interesting to notice here that most of these words descend from Latin or medieval disyllables (words that, originally, agreed with (2) or (8), thus), as the result of heterosyllabic vowel contraction, as exemplified by *mala*>*maa*>*má* 'bad\_FEM', *fede*>*fee*>*fê* 'faith', *pala*>*paa*>*pá* 'shovel', *pede*>*pee*>*pé* 'foot', *nudu*>*nuu*>*nu* 'naked', *dolu*>*doo*>*dó* 'sorrow, mourning', *solu*>*soo*>*só* 'only' (see MACHADO, 1977, for these etymologies).

If we accept Brandão de Carvalho's (1988; 2011) explanations, higher sonority degrees – rendering certain vowels immune to vowel reduction (see CEP examples such as, e. g., *pegada* 'footprint' ([pe'gadɐ], \*[pi'gadɐ]) – are another inheritance, in present-day EP, of such phonologically *unlost* syllable weight.

Therefore, we cannot exclude completely that such monosyllables preserve, in their underlying forms, an abstract structure that maintains their status as heavy syllables: for example, *pé* 'foot' =/pɛ/<sub>σH</sub>, *má* 'bad\_FEM' =/ma/<sub>σH</sub>, their syllable heaviness ( $\sigma_H$ ) being kept by higher degrees of vowel sonority, according to Brandão de Carvalho (2011, p.59). In some cases, we can even admit the presence, at the maximally abstract level of phonological representation, of 3 phonological segments (*nu* 'naked' =/nuu/, *dó* 'sorrow, mourning' =/dɔɔ/). This even becomes segmentally very transparent in morphological derivations of these stems, in cases like *desnudo* 'naked [poetic]', *dolorido* 'consumed by sorrow', etc. (accordingly, then, with (8)). Such explanation would render these words more compatible with the various MC versions that were analyzed in this study, cancelling the need for *ad hoc* explanations strongly based on the exceptionality of those forms as the only explanation for their insertion in the Portuguese lexicon (even if we cannot ignore that these cases do correspond to very particular cases and further investigation would be necessary to reach more robust conclusions about them).

## Concluding remarks

Based on our data and our discussion, and not ignoring the limitations of our study, we think it possible to draw some final observations that can shed some light on the initial questions and hypotheses of our research.

The main conclusion is that it seems indeed possible *to advocate the pertinence of a Minimality Condition in Contemporary European Portuguese*, mandatorily obeyed by the gross majority of lexical entries of this language.

Contrarily to the general understanding about the MC parametrization in Portuguese and other languages, though, CEP's MC does not depend on syllable weight or syllabic extension: according to our data, CEP's MC seems to be peculiarly conditioned by one single parameter, different and independent from those – *the number of segments of the word*. We propose, in this paper, that phonetic chains with 3 or more segments are good candidates to wordiness in CEP, regardless of their syllable weight and syllabic extension. The participants' behavior in the experimental task they were confronted with showed a consistent tendency, statistically probable, confirming this proposal. In several moments of our paper, we acknowledged that this MC formulation does not comply with most previous versions of minimality for Portuguese and other languages, precisely because of the apparent cancelling of syllable weight and/or syllabic extension, which are generally assumed as the most important variables related to wordiness and minimality (McCARTHY; PRINCE, 1995 – see (1)).

A second conclusion that can be drawn from our study is that MC, as any other phonological property of a given language, is subject to historic variation. The MC that is proposed here for CEP ( $\{W_{MinExtSg} \geq 3\}$  - see (8)) is not the same as the MC that was eventually productive in Medieval Portuguese ( $\{w \geq \sigma_H\}$  - see (2)), much more sensitive to syllable weight.

We emphasize here that the apparent lack of importance of syllable weight as a crucial criterion of MC in CEP does not equate its elimination from the set of phonological constraints of this language. The conclusion that, *at the current stage of the language*, syllable weight does not play a decisive role for the parametrization of MC has to do with data that show that monosyllables with at least 3 segments, be they light (with a branched onset but an unbranched rhyme, as in “BLO”, “PRI” or “CLA”) or heavy (with a branched rhyme, as in “EUS” and “OIS”), are preferably oralized as acronyms, contrasting with monosyllables with just 2 segments, also regardless of their syllabic weight, which are preferably oralized as initialisms. At other levels – namely as for non-verb stress-assignment, for instance – arguments in favour of syllable weight as a productive phonological property seem to be very relevant, as we have briefly seen in 2.2.1 and according to the descriptions of Brandão de Carvalho (1988; 1989; 2011) or Wetzels (2007). This specific subject, however, was not addressed as a central question of this study.

Four questions remain unanswered in this study, in need of future research:

- the processing of pseudoabbreviations phonologically similar to clitics;
- the specific case of words corresponding to light monosyllables formed by 2 segments only (underneath, thus, the MC in terms of syllable weight, syllabic extension or number of segments);

- the oralization of phonetic chains that respect important phonotactic constraints of the language but violate others, namely the prohibition of segments in principle unadmitted in coda position (e.g.: FENPROF, RENN; see examples C2 in column C, Table 4);

- the oralization of pseudoabbreviations potentially ending in a nasal segment, that were deliberately excluded from the scope of our study, as previously explained.

VELOSO, J. Palavra mínima em português europeu: a oralização de abreviações. *Alfa*, São Paulo, v.61, n.1, p.131-163, 2017.

- *RESUMO: Entre as restrições fonológicas a que qualquer cadeia fonética está obrigatoriamente sujeita para poder ser aceite como uma palavra da língua conta-se a Condição de Minimalidade (CM), que impõe uma quantidade mínima de material fonológico que deve ser contido por qualquer palavra. A CM costuma ser medida em termos de peso silábico ou de extensão silábica. Sendo discutível se se trata de uma restrição verdadeiramente universal, estudos anteriores relativos ao português têm-se dividido quanto à sua operacionalidade na fonologia desta língua. Neste estudo, avaliaremos a funcionalidade da CM na constituição do léxico do português através da oralização de abreviações, assumida como um processo produtivo em português. A partir de um estudo empírico realizado com um grupo de falantes nativos do português europeu contemporâneo (PEC), propomos (i) que a CM seja uma condição fonológica operacional nesta língua e (ii) que o critério a que ela obedece em PEC seja um critério de ordem puramente linear e segmental. De acordo com a proposta aqui apresentada, respeitam a CM em PEC cadeias com três ou mais segmentos, independentemente do peso silábico ou da extensão silábica.*
- *PALAVRAS-CHAVE: Condição de Minimalidade. Peso silábico. Palavra. Restrições de Palavridade. Boa Formação.*

## REFERENCES

ABREU, K. N. M. **Um estudo sobre as siglas do português do Brasil**. 143 f. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2009.

ABREU, K.; ROSA, M. C. Isso é uma palavra? *Revista da ABRALIN*, Rio de Janeiro, v.5, n.1-2, p.113-130, 2006.

ARAÚJO, G. Truncamento e Reduplicação no Português Brasileiro. *Revista de Estudos da Linguagem*, Belo Horizonte, v.12, p.61-90, 2002.

ARONOFF, M. **Word Formation in Generative Grammar**. Cambridge, MA: The MIT Press, 1976.

- ARONOFF, M.; FUDEMAN, K. **What is Morphology?** Malden, MA: Blackwell, 2005.
- BISOL, L. O Clítico e o seu Status Prosódico. **Revista de Estudos da Linguagem**, Belo Horizonte, v.9, n.1, p.5-30, 2000.
- \_\_\_\_\_. Mattoso Câmara Jr. e a palavra prosódica. **DELTA**, São Paulo, v.20, p.59-70, 2004.
- BISOL, L. A palavra fonológica pós-lexical. In: GUIMARÃES, E.; MOLLICA, M. C. (Org.). **A Palavra. Forma e Sentido**. Campinas: Pontes, 2007. p.13-22.
- BOOIJ, G. 2004. The morphology-phonology interface in European Portuguese. Review article of M. Vigário, The Prosodic Word in European Portuguese. **Journal of Portuguese Linguistics**, Lisboa, v.3, n.1, p.175-182, 2004.
- BRANDÃO DE CARVALHO, J. Réduction vocalique, quantité et accentuation: pour une explication structurale de la divergence entre portugais lusitanien et portugais brésilien. **Boletim de Filologia**, Lisboa, v.XXXII, p.5-26, 1988.
- \_\_\_\_\_. Phonological conditions on Portuguese clitic placement: on syntactic evidence for stress and rhythmical patterns. **Linguistics**, Berlin, v.27, p.405-436, 1989.
- BRANDÃO DE CARVALHO, J. Contrastive hierarchies, privative features, and Portuguese vowels. **Linguística**, Porto, v.6, n.1, p.51-66, 2011.
- CHACON, L. Constituintes prosódicos e letramento em segmentações prosódicas não-convencionais. **Letras de Hoje**, Porto Alegre, v.39, n.3, p.223-232, 2004.
- CORREIA, M.; LEMOS, L. S. P. **Inovação lexical em português**. Lisboa: Colibri/ Associação de Professores de Português, 2005.
- CUNHA, A. P. As segmentações não convencionais da escrita inicial: um estudo sobre o troqueu silábico e a sua relação com o ritmo linguístico do PB e do PE. **Linguística**, Porto, v.7, n.1, p.45-63, 2012.
- ELORDIETA, G. The Word in Phonology. In: IBARRETXE-ANTUÑANO, I.; MENDÍVIL-GIRÓ, J.-L. (Org.). **To be or not to be a Word. New Reflections on the Definition of Word**. Newcastle upon Tyne: Cambridge Scholars Publishing, 2014. p.6-65.
- ENNAJI, M. Aspects of cliticization in Arabic. In: ENNAJI, M. (Org.). **La variation linguistique: des faits aux théories**. Fès: Université Sidi Mohamed Ben Abdellah/ Faculté des lettres et des sciences humaines, 2000. p.98-118.
- FERREIRA, M. L. S. **Contributos para uma definição de palavra fonológica**. 211 f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade do Porto, Porto, 2012.

GERLACH, B.; GRIJZENHOUT, J. (Org.). **Clitics in phonology**: morphology and syntax. Amsterdam: John Benjamins, 2000.

GONÇALVES, C. A. V.; LIMA, B. C.; THAMI DA SILVA, H. Hipocorização no português brasileiro: moldes, circunscrições e correspondência em dois padrões de formação. In: DUARTE, I.; LEIRIA, I. (Org.). **Actas do XX Encontro Nacional da Associação Portuguesa de Linguística**. Lisboa: Associação Portuguesa de Linguística, 2005. p.605-614.

GORI, B. **La grammatica dei clitics portoghesi**. Firenze: Firenze University Press, 2007.

HALLE, M. Prolegomena to a theory of word formation. **Linguistic Inquiry**, Cambridge MA, v.4, n.1, p.3-16, 1973.

HASPELMATH, M. The indeterminacy of word segmentation and the nature of morphology and syntax. **Folia Linguistica**, Berlin, v.45, n.1, p.31-80, 2011.

\_\_\_\_\_. How to compare major word-classes across the world's languages. In: GRAF, T. et al. (Org.). **Theories of everything**: in honor of Edward Keenan. UCLA Working Papers in Linguistics 17. Los Angeles CA: UCLA, 2012a. p.109-130.

\_\_\_\_\_. Escaping ethnocentrism in the study of word-class universals. **Theoretical Linguistics**, Berlin, v.38, n.1-2, p.91-102, 2012b.

IBARRETXE-ANTUÑANO, I.; MENDÍVIL-GIRÓ, J.-L. Introduction. In: IBARRETXE-ANTUÑANO, I.; MENDÍVIL-GIRÓ, J.-L. (Org.). **To be or not to be a Word. New Reflections on the Definition of Word**. Newcastle upon Tyne: Cambridge Scholars Publishing, 2014. p.1-5.

JONES, D. The 'word' as a phonetic entity. *Le Maître Phonétique*, Bourg-La-Reine, v.36, p.60-65, 1931. Reproduzido in: JONES, W. E.; LAVER, J. (Org.). **Phonetics in Linguistics. A Book of Readings**. London: Longman, 1973. p.154-158.

JUILLAND, A.; ROCERIC, A. **The linguistic concept of word**: Analytic bibliography. The Hague: Mouton, 1972. Cit. by HASPELMATH, M. The indeterminacy of word segmentation and the nature of morphology and syntax. **Folia Linguistica**, Berlin, v.45, n.1, p.31-80, 2011.

JULIEN, M. Word. In: BROWN, K. et al. (Org.). **Encyclopedia of Language and Linguistics**. 2. ed. Amsterdam: Elsevier, 2006. v.13, p.617-624.

KIPARSKY, P. Some Consequences of Lexical Phonology. **Phonology Yearbook**, Cambridge, v.2, p.82-136, 1985.

KRÁMSKÝ, J. **The Word as a Linguistic Entity**. The Hague: Mouton, 1969.

LAUSBERG, H. **Linguística Românica**. 2. ed. Tradução de Marion Ehrhardt e Maria Luísa Schemann. Lisboa: Fundação Calouste Gulbenkian, 1981 [1963].

LIMA, B. C. **A formação de ‘Dedé’ e ‘Malu’**: uma análise otimalista de dois padrões de hipocorização. 67 f. Dissertação (Mestrado em Letras Vernáculas) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

LIMA, B. C.; THAMI DA SILVA, H.; GONÇALVES, C. A. V. A hipocorização no português do Brasil: análise por padrões gerais de formação. In: GONÇALVES, C. A. V. et al. (Org.). **Otimalidade em foco**: morfologia e fonologia do português. Rio de Janeiro: Publit, 2009. p.115-151.

MACHADO, J. P. **Dicionário Etimológico da Língua Portuguesa**. 5 vols. Lisboa: Livros Horizonte, 1977.

MATEUS, M. H. M.; BRITO, A. M.; DUARTE, I.; FARIA, I. H.; FROTA, S.; MATOS, G.; OLIVEIRA, F.; VIGÁRIO, M.; VILLALVA, A. **Gramática da Língua Portuguesa**. 5. ed. Lisboa: Caminho, 2003.

MATEUS, M. H.; D’ANDRADE, E. **The Phonology of Portuguese**. Oxford: Oxford University Press, 2000.

MCCARTHY, J. J.; PRINCE, A. S. Prosodic Morphology. In: GOLDSMITH, J. A. (Org.). **The Handbook of Phonological Theory**. Cambridge: Blackwell, 1995. p.318-366.

MENDES, A. Processos de gramaticalização. In: PAIVA RAPOSO, E. B. et al. (Org.). **Gramática do Português**. Lisboa: Fundação Calouste Gulbenkian, 2013. v.I, p.249-293.

NESPOR, M. The phonology of clitic groups. In: VAN RIEMSDIJK, H. (Org.). **Clitics in the languages of Europe**. Berlin: Mouton de Gruyter, 1999. p.865-887.

NESPOR, M.; VOGEL, I. **Prosodic Phonology. With a new foreword**. Berlin: Mouton De Gruyter, 2007.

NUNES, J. J. **Compêndio de Gramática Histórica Portuguesa. Fonética e Morfologia**. 5. ed. Lisboa: Livraria Clássica, 1956.

PEREIRA, M. I. P. **O acento de palavra em português**. Uma análise métrica. 286 f. Dissertação (Doutoramento em Linguística) – Faculdade de Letras da Universidade de Coimbra, Coimbra, 1999.

\_\_\_\_\_. Processos de construção não concatenativa. In: RIO-TORTO, G. et al. **Gramática Derivacional do Português**. Coimbra: Imprensa da Universidade de Coimbra, 2013. p.463-491.

PLÉNAT, M. Observations sur le mot minimal français. L’oralisation des sigles. In: LAKS, B.; PLÉNAT, M. (Org.). **De natura sonorum. Essais de phonologie**. Saint-Denis: Presses Universitaires de Vincennes, 1993. p.143-172.

ROCA, I. M. Stress in the Romance languages. In: VAN DER HULST, H. (Org.). **Word Prosodic Systems in the Languages of Europe**. Berlin: Mouton De Gruyter, 1999. p.659-811.

ROSA, M. C. **Introdução à Morfologia**. São Paulo: Contexto, 2006.

SADIQI, F. Issues in Berber cliticization. In: ENNAJI, M. (Org.). **La variation linguistique: des faits aux théories**. Fès: Université Sidi Mohamed Ben Abdellah/ Faculté des lettres et des sciences humaines, 2000. p.119-151.

SILVA, R. V. M. **O Português Arcaico. Uma Aproximação**. 2 vols. Lisboa: Imprensa Nacional – Casa da Moeda, 2008.

TENANI, L. Segmentações não-convencionais e teorias fonológicas. **Letras de Hoje**, Porto Alegre, v.39, n.3, p.233-244, 2004.

ULRICH, C. W. **A noção de palavridade na concepção de falantes do português brasileiro**. 101 f. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.

VAN OOSTENDORP, M. Italian s-voicing and the structure of the phonological word. In: HANNAHS, S. J.; DAVENPORT, M. (Org.). **Issues in phonological structure**. Amsterdam: John Benjamins, 1999. p.195-212.

VAN RIEMSDIJK, H. (Org.). **Clitics in the languages of Europe**. Berlin: Mouton de Gruyter, 1999.

VELOSO, J. Découpage de continuums phonétiques en mots: Critères formels vs. indices substantiels. In: CROUZET, O.; TIFRIT, A.; ANGOUJARD, J.-P. (Org.). **Actes des 6.èmes Journées d'Etudes Linguistiques**. Nantes: Université de Nantes, 2009. p.85-90.

\_\_\_\_\_. Rimes/VGNS/en position finale de mot en portugais: Une contrainte «sensible au mot». In: ILIESCU, M.; SILLER-RUNGGALDIER, H. M.; DANLER, P. (Org.). **Actes du XXVe Congrès International de Linguistique et de Philologie Romanes**. Berlin: De Gruyter, 2010. v.II, p.231-240.

\_\_\_\_\_. Unidades acentuais proparoxítonas e grupos clíticos em Português. In: COSTA, A.; DUARTE, I. (Org.). **Nada na linguagem lhe é estranho**. Homenagem. Isabel Hub Faria. Porto: Afrontamento, 2012. p.471-483.

\_\_\_\_\_. Prosodisation des complexes « mot phonologique+clitique(s) » : groupes accentuels proparoxytons et groupes clitiques en portugais. In: TIFRIT, A. (Org.). **Phonologie, morphologie, syntaxe**. Mélanges offerts à Jean-Pierre Angoujard. Rennes: Presses Universitaires de Rennes, 2013. p.233-244.

\_\_\_\_\_. ***Verba Manent. A palavra como unidade pertinente para a descrição linguística do português e de outras línguas flexionais.*** Porto Alegre: EDIPUCRS, 2016.

VIGÁRIO, M. Elisão da vogal não-recuada final e a palavra prosódica no português europeu. In: XIII ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA. **Actas.** Lisboa: APL/Colibri, 1998. p.359-376.

\_\_\_\_\_. **The Prosodic Word in European Portuguese.** Berlin: Mouton de Gruyter, 2003.

\_\_\_\_\_. O lugar do Grupo Clítico e da Palavra Prosódica Composta na hierarquia prosódica: uma nova proposta. In: XXII ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA. **Textos Seleccionados.** Lisboa: APL/Colibri, 2007. p.673-688.

\_\_\_\_\_. Prosodic structure between the Prosodic Word and the Phonological Phrase: recursive nodes or an independent domain? **The Linguistic Review**, Berlin, v.27, n.4, p.485-530, 2010.

VIGÁRIO, M.; FERNANDES-SVARTMAN, F. A atribuição tonal em compostos no Português do Brasil. In: BRITO, A. M.; SILVA, F.; VELOSO, J.; FIÉIS, A. (Org.). **XXV Encontro Nacional da Associação Portuguesa de Linguística.** Textos Seleccionados. Porto: Associação Portuguesa de Linguística, 2010. p.769-786.

VIGÁRIO, M.; MARTINS, F.; FROTA, S. Frequências no Português Europeu: a ferramenta FreP. In: DUARTE, I.; LEIRIA, I. (Org.). **Actas do XX Encontro Nacional da Associação Portuguesa de Linguística.** Lisboa: Associação Portuguesa de Linguística, 2005. p.897-908.

VILLALVA, A. **Morfologia do Português.** Lisboa: Universidade Aberta, 2008.

\_\_\_\_\_. Palavras, que as há. **Revista de Estudos da Linguagem**, Belo Horizonte, v.20, n.2, p.125-139, 2012.

VOGEL, I. Subminimal constituents in prosodic phonology. In: HANNAHS, S. J.; DAVENPORT, M. (Org.). **Issues in phonological structure.** Amsterdam: John Benjamins, 1999. p.249-267.

WETZELS, L. Primary Stress in Brazilian Portuguese and the Quantity Parameter. **Journal of Portuguese Linguistics**, Lisboa, v.5/6, p.9-58, 2007.

WILLIAMS, E. B. **Do Latim ao Português. Fonologia e Morfologia Históricas da Língua Portuguesa.** Tradução de A. Houaiss. 3. ed. Rio de Janeiro: Tempo Brasileiro, 1975 [1938].



LIST OF MAIN ABBREVIATIONS AND CONVENTIONS  
FOLLOWED THROUGHOUT THE TEXT

<p>C = Consonant/Contoid          CEP = Contemporary European Portuguese          EP = European Portuguese          G = Glide          Lat. = Latin          MC = Minimality Condition          MP = Medieval Portuguese          PhK = Phonological Knowledge          PWG = Prosodic Word Group          SD = Standard-Deviation</p>	<p>S = Underspecified Coronal Fricative          V = Vowel/Vocoid          WMinExtSg=Word Minimal Extension,          measured in number of segments  <math>\sigma_1</math> = Initial syllable of a word  <math>\sigma_H</math> = Heavy Syllable  <math>\sigma_L</math> = Light Syllable  <math>\omega</math> = Phonological/Prosodic Word</p>
--	--

Received in september 2015

Approved in november 2015

