Grammar and frequency effects in the acquisition of prosodic words in European Portuguese*

Marina Vigário+, Maria João Freitas!, and Sónia Frota!

+ILCH, Universidade do Minho, !DLGR, Universidade de Lisboa

Submitted to SPECIAL ISSUE ON PROSODIC WORDS, ed. by Katherine Demuth

Running head: Acquisition of prosodic words in EP

Corresponding author's full address:

Marina Vigário,
ILCH, Universidade do Minho,
Campus de Gualtar,
4710-057 Braga
Portugal
marina.vigario@mail.telepac.pt

* This paper develops the materials and ideas presented by the authors at the Second Lisbon Meeting on Language Acquisition (University of Lisbon, 2004). We are grateful to the audience of the Panel session on the Acquisition of Prosodic Word Structure in Comparative Romance, and especially to Katherine Demuth, Heather Goad, Conxita Lleó, Pilar Prieto, and Marilyn Vihman. We also thank Ana Lúcia Santos and Carla Soares for making the child-directed speech transcriptions available to us, and João Horta for her help with the figures. The paper has greatly benefited from comments by two *Language and Speech* reviewers. Special thanks are due to the children and their mothers, that have made this (and other) studies on the acquisition of European Portuguese possible (in particular, João and Antónia, Inês and Guilhermina, Marta and Cristina). This work is part of the research developed within the SILC project, held by ONSET-CEL (Faculdade de Letras, Universidade de Lisboa), in collaboration with Universidade do Minho.

Abstract

This paper investigates the acquisition of prosodic words in European Portuguese (EP) through analysis of grammatical and statistical properties of the target language and child speech. The analysis of grammatical properties shows that there are solid cues to the prosodic word (PW) in EP, and the presence of early word-based phonology in child speech shows that EP children are aware of these cues. It is thus hypothesized that grammatical properties could play a role in the development of the PW by promoting the early production of the different word shapes found in the language. The analysis of statistical properties of the input, namely word shape frequencies in adult speech and child-directed speech, shows that they constrain early word shapes in child speech in ways similar to recent reports on other languages: a fairly high frequency of monosyllabic shapes, and especially of monosyllabic CV shapes, in the input agrees with the production of subminimal words in child speech; a fairly high frequency of trisyllabic and larger shapes in the input (adult speech in particular) matches the early development of words larger than a binary foot. These patterns, together with the co-occurrence of truncation to subminimal shapes in the initial and later stages, as well as the presence of prosodic fillers regardless of word size, support the claim that early words in EP are not constrained by minimality or maximality requirements. The potential interaction of grammar and frequency effects in PW acquisition is discussed in the light of the present findings and comparable data available in the literature for English, French, Spanish and Catalan.

Key words: prosodic word, language development, frequency, phonology, European Portuguese

1. INTRODUCTION

Languages are known to differ in their phonological grammars and in the frequency of use of different phonological properties or units. Nevertheless, the effect of language-specific traits on the timing and course of prosodic word development is still largely unknown. Studying the extent to which children's early words are sensitive to the language-specific prosodic word phonology and language-specific frequency patterns involves combining three major areas of research: the grammar of the target system, the impact of frequency on language development, and the study of early phonological acquisition. The present paper addresses these three areas with the aim of providing a first contribution to the understanding of the acquisition of the prosodic word in European Portuguese (hereafter EP).

A growing body of literature has recently come to light showing various sorts of frequency effects in adult language (see, among many others, Bybee, 2000, 2001, the papers collected in Bybee & Hooper, 2001, Jurafsky, Bell & Girand, 2002, Moates, Bond & Stockmal, 2002, Pierrehumbert, 2002, and references therein). The relevance of statistical properties of the input has also been suggested to play an important role in language acquisition and development (Fikkert & Freitas, 1998, Roark & Demuth, 2000, Beckman & Edwards, 2000, Lleó, 2001a, 2001b, 2004, Lleó & Demuth, 1999, Demuth & Johnson, 2003, and Prieto, 2004, inter alia). Frequency can be seen from the perspective of the use of specific structures, patterns, categories, items, or the like. It is not always easy to disentangle frequency effects from grammar effects. For example, if the grammar of a particular language allows for complex codas, as in Dutch, it is expected that the frequency of syllables with more than one final consonant in this language is higher than in a language where their occurrence would result in the violation of a highly ranked constraint, such as in EP. Additionally, it can be hypothesized that phonological phenomena may play a role, regardless of their frequency of use, if they concur as a constellation of cues to signal a given prosodic structure or extract a particular phonological generalization (e.g Morgan, Shi & Allopenna, 1996). In this case, it would be the presence of cues pointing in the same direction that would be frequent just because the grammar offers a cluster of facts that highlight the same property of the language.

Cross-linguistic differences in prosodic word structure in the adult target system include the grouping of languages into those that show a constellation of phenomena cuing the prosodic word (PW) and those that exhibit weaker evidence for the prosodic word. Languages like Dutch, German, or English, where phonotactic restrictions, segmental phenomena and resyllabification are bound by the prosodic word, are among the former (Booij, 1995, 1999, for Dutch; Wiese, 1996, Hall, 1999, for German; Raffelsiefen, 1999, for English). Languages like French, Italian, Spanish, or Brazilian Portuguese belong to the latter group. In these languages, phonological facts that depend on the prosodic word are clearly less common (and more so in French than in the other Romance languages), and resyllabification spans a larger domain, the intonational phrase (Hannahs, 1995a, 1995b, for French, and Kleinhenz, 1997, for the contrast between Dutch and French; Peperkamp, 1997, for Italian and Spanish; Bisol, 2000, Schwindt, 2000, for Brazilian Portuguese). The properties of EP grammar seem to blur the apparently neat division between Romance and Germanic languages: on the one hand, EP phonology offers a rich array of evidence for the prosodic word (Vigário, 2003); on the other hand, resyllabification in EP is Romance-like (Frota, 2000). The properties of prosodic words in EP are described in section 2, as well as the predictions based on the assumption that grammatical properties of the target language could play a role in PW development.

Language-specific frequency distributions of prosodic word shapes in the input have been argued to constrain prosodic word development. A general trend found crosslinguistically shows an initial state, where children may produce monosyllabic monomoraic utterances, followed by early words which are both minimally and maximally a binary foot (see Demuth & Johnson, 2003, and references therein). However, there seems to be some variation, depending on the frequency of different prosodic word shapes in the input language. The frequency patterns may constrain the timing of emergence and the course of development of subminimal words and/or words with more than a binary foot (Demuth & Johnson, 2003, Lleó, 2001b, 2004, Prieto, 2004). For example, Lleó (2001b) shows that children acquiring Spanish exhibit early violations of maximality, when compared to German learning children. Demuth & Johnson (2003) show that subminimal words are found not only at the initial state but at later stages in the acquisition of French. Both cases are explained by the frequency distribution patterns in the input: a higher percentage of trisyllabic words with a pretonic syllable in Spanish relative to German, and the relative high frequency of subminimal words in French. The import of word shape frequency to children's developing phonology has not been dealt with in previous work on the acquisition of EP. One of the main goals of the present paper is to assess the frequency effects found in prosodic word acquisition. To address the issue, empirical databases of word shape frequencies in the ambient language (both in child-directed speech and adult speech) and in children's early productions have been used. The findings are reported in section 3, where the predictions based on the frequency patterns (section 3.1) are assessed through the analysis of child speech data (section

3.2). This analysis covers the word shapes found in early child speech, their course of development, their relation with the target word shapes, and the presence of prosodic fillers.

The impact of grammar and frequency effects on EP children's early PW structures is discussed in section 4. It is concluded that word shape frequencies in the ambient language constrain early word shapes in child speech by promoting the presence of subminimal shapes and the early development of words larger than a foot. In the case of EP, the grammatical constraints are argued to interact with the frequency effect in yielding the early production of the different prosodic word shapes found in the language.

2. THE PROSODIC WORD IN EUROPEAN PORTUGUESE

It is by now well-established that the phonology of EP offers a rich array of evidence for the prosodic word. Different kinds of phonological phenomena act as a constellation of cues that signal prosodic word structure. Section 2.1 below provides an illustrative summary of these facts, based on Vigário (2003). Given that there are solid cues to the PW in EP, it may be the case that EP children are aware of these cues, and that this is reflected in children's productions. In section 2.2 we present data showing that early child speech indeed exhibits facts pointing to the presence of word-based phonology, namely asymmetrical distributions of sounds within words and prosodic influences of position within the word in the realization of sounds. Based on the evidence available in the ambient language and in the child production facts, it is reasonable to assume that the grammatical properties of the language could play a role in the acquisition of the PW. We conclude this section stating the predictions derived from such assumption.

2.1. Evidence for the PW in the input grammar

Before describing the PW properties that characterize EP, more general information on the phonetics and phonology of EP (standard variety spoken in the region of Lisbon by educated speakers) is given.

EP is characterized by the presence of a stressed vowel system and an unstressed vowel system, as shown in Table 1. The latter system is a consequence of vowel reduction in unstressed positions, namely centralization and raising of [e , ɛ] to [ɨ], raising of [a] to [ɐ], and raising of [o , ɔ] to [u].[1] EP also has nasal vowels, which may appear in stressed and unstressed positions, and many diphthongs. Rising diphthongs also exist and are generally regarded as phonetic variants of vowel sequences of the [i]/[u] plus vowel. A further factor relevant to the issues under discussion in this paper is the absence of quantity contrasts in the sound inventory of the language.

INSERT TABLE 1 ABOUT HERE

EP syllable structure is rather simple, as only three different syllable types make up 80% of the most frequent syllables found in adult speech: CV, V and CVC (see Table 2). There are severe restrictions on final syllable consonants, which comprise only the alveolar liquids and palatal fricatives. Consonant clusters that appear in CCV syllables are restricted to sequences of obstruent plus alveolar liquid.

INSERT TABLE 2 ABOUT HERE

Word stress in EP may fall on the penultimate syllable, on the final syllable, or on the antepenultimate syllable, as shown in Table 3. Most accounts of word stress agree that morphology plays a major role in constraining the position of stress. The generalization that stress falls on the last vowel of the stem accounts for about 80% of the native nouns and adjectives. In verb forms, stress occurs on the penultimate syllable in 75% of cases (Mateus & Andrade, 2000). As most nouns and adjectives have a class marker, these two facts together make penultimate stress the dominant pattern. Another relevant property of the EP stress system is the absence of sensitivity to quantity, as shown by most accounts of word stress (see Mateus & Andrade, 2000, p.109-119 for a discussion of the (non)role of quantity in the EP stress system). The examples in Table 3 also illustrate why quantity has been argued not to play a role (e.g. [tuˈnɛɫ] and [ˈtunɛɫ];[tuˈfɐ̃w̃] and [ˈɔɾfɐ̃w̃]).

INSERT TABLE 3 ABOUT HERE

Prosodic words in EP may contain from one up to more than three syllables, and monosyllabic words with open syllables are also present in the lexicon, as well as in the list of the most frequent words used in the language, as shown in Table 4.

INSERT TABLE 4 ABOUT HERE

Let us now consider the properties that signal the prosodic word in EP. Evidence for the prosodic word in the adult target system comprises edge-related phenomena – signalling both the left and the right edges of PW –, word-bound phenomena – phenomena that select the PW as their domain, such as clipping and deletion under identity –, and prominence-related phenomena. The full set of tests available as diagnostics for the PW is described in Vigário (2003, p.157-163).

Evidence for prosodic word edges comes from phonotactic constraints and segmental processes. Prosodic words in EP cannot start with [ʎ] and [ɲ]. Other segments that can occur ω-internally but do not surface in ω-initial position are [ɾ] and [ɨ]: the former undergoes a strengthening process and is realized as [ʀ] or [ʁ] in this position; the later may not surface due to restrictions imposed by PW-initial position on vowel reduction, which is the only source of schwas in EP. The examples in (1) illustrate some of these constraints.

(1) *rato* [ˈʀa.tu] vs. *caro* [ˈka.ɾu] ‘mouse’/’car’

erguer *[ɨɾ.ˈgeɾ] vs. *perder* [pɨɾ.ˈdeɾ] ‘to raise’/to loose’

Vowels in PW-initial position are exempt from regular vowel reduction: specifically, non-high palatal vowels ([e, ɛ]) are not raised and centralized when unstressed (*[ɨ]), and non-high round vowels ([o, ɔ]) are not raised when unstressed (*[u]). This is shown in (2).

(2) *erguer* [eɾ.ˈgeɾ] vs. *roedor* [ru.ɨ.ˈdoɾ] ‘raise’/’rodent’

opinião [o.pi.ni.ˈɐ̃w̃] vs. *miolinho* [mi.u.ˈli.ɲu] ‘opinion’/ ‘soft part of bread-DIM’

Emphatic stress also marks the left-edge of the prosodic word. It consists of a high F0 level at the beginning of the PW, with the function of highlighting that particular word. Notice that this type of stress may occur in phrase internal position and thus behaves as a PW left-edge phenomenon and not a phrasal edge phenomenon.

There are also segmental phenomena that refer to the right edge of prosodic words (although this edge is less prominent than the left edge in the language): non-high stressless vowels in PW final syllables closed by sonorants are lowered (as in (3a)); stressless vowels at the edge of a prosodic word that is part of a morphological compound are lowered (see (3b)); non-high palatal vowels in PW final position are deleted (3c); and the semivocalization process that creates falling diphthongs also applies at the right-edge of prosodic words (3d).

(3) a. *líder* [ˈli.dɛɾ] vs. *pode* [ˈpɔ.dɨ / ˈpɔ.d] 'leader'/ 'can'

âmbar [ˈɐ̃.baɾ] vs. *acaba* [ɐ.ˈka.bɐ] 'amber'/ 'finish'

sénior [ˈsɛ.njɔɾ] vs. *livro* [ˈli.vɾu] 'senior'/ 'book'

b. *mono-acentual* [ˈmɔnɔ ɐsẽˈtwaɫ] vs. *monografia* [munugɾɐˈfiɐ]
'mono-accentual / monograph'

c. *passe* [ˈpas] vs. *passemos* [pɐˈsemuʃ]/ *passear* [pɐˈsjaɾ]
'to pass / pass-SUBJ-2pp, take a walk'

d. *rio* [ˈʀi.u / ˈʀiw] vs. graudinho [gɾɐ.u.ˈdi.ɲu] / *[gɾɐw.ˈdi.ɲu]
'river / big-DIM'

There are deletion processes that apply with reference to prosodic words, but not to morpho-syntactically defined units, as shown in (4): (4a) illustrates an instance of clipping, and (4b) a case of deletion under identity where the target sequences have the same morphological status but crucially differ prosodically (see Vigário 2003, p.250-255 for further details; parentheses signal prosodic word status).

(4) a. ~~tele~~*móvel* > *móvel* (tele) (móvel) > móvel [ˈtɛlɛ ˈmɔvɛɫ] ‘mobile’

telefonia > **fonia* (telefonia) > *fonia [tɨlɨfuˈniɐ] ‘radio’

b. *mono~~gamia~~ ou poligamia* (mono) (gamia) ou (poli) (gamia)
[ˈmɔnɔ gɐˈmiɐ] [ˈpɔli gɐˈmiɐ] ‘monogamy or polygamy’

**bio~~grafia~~ e discografia* (biografia) e (discografia)
[biugrɐˈfiɐ] [diʃkugrɐˈfiɐ] ‘biography and discography’

Several stress-related phenomena also cue the prosodic word. Word stress is perceptually salient in EP. Furthermore, there are many segmental processes that refer to the presence/absence of word-stress, both PW-internally (e.g. glide insertion to break hiatus, semivocalization of high vowels if followed by another vowel), and across PWs (e.g. PW-final vowel deletion if the following word starts with a vowel). Vowel reduction, which is a rather general process in EP, never affects stressed vowels, as shown by the minimal contrasts in (5).

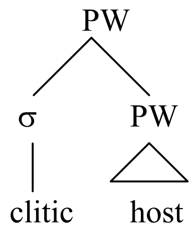
(5) *d<u>á</u>* [ˈda] vs. *da* [dɐ] ‘to give / of-the-FEM’

d<u>ê</u> [ˈde] vs. *de* [dɨ] ‘to give-SUBJ-3PSING / of’

A number of morphophonological processes further show the asymmetry between stressed and unstressed syllables: for example, the realization of plurals in nominal forms ending with fricative or lateral varies depending on the status of the word-final syllable relative to stress (e.g. *nariz* [nɐˈɾiʃ] / *narIzes* [nɐˈɾiziʃ] ‘nose / noses’ but *lápis* [ˈlapiʃ] */ lápis* [ˈlapiʃ] ‘pencil’ / pencils’; *funil* [fuˈniɫ] / *funis* [fuˈniʃ] ‘funnel / funnels’ but *fácil* [ˈfasiɫ] / *fáceis* [ˈfasɐjʃ] ‘easy / easy-PL’). All these phenomena that highlight the stressed syllable concur to signal the prosodic word because the presence of one stress implies the presence of a prosodic word.

We have seen that the adult target system displays a constellation of phenomena cuing the prosodic word. There are two other properties in the phonological grammar of EP that affect PW structure: the direction of cliticization and resyllabification. Adult speech displays around 30% of phonological clitics (Vigário, Martins & Frota, 2005). Most phonological clitics in EP are proclitics (as shown in Vigário 2003) and frequency data on the production of clitics in adult speech shows that 97% of clitics are proclitics (Vigário, Martins & Frota, 2005). According to the analysis proposed by Vigário (2003), which we adopt, proclitics are adjoined at the level of the prosodic word, and thus they may constitute an additional PW-initial syllable, as it is shown in (6).

(6) EP proclitics: adjunction to the following PW



Consequently, proclitics may contribute to word shape through the addition of one more syllable to the prosodic host. The result is the enlargement of the different prosodic word shapes, as there is no reason to suppose that clitic distribution is constrained by the size of the prosodic host. Resyllabification is probably the only property in the phonological grammar of EP that may weaken the otherwise neat picture for the PW. Resyllabification and related phonological processes (e.g. across word semivocalization or vowel deletion as in *músico africano* [ˈmu.si.kwɐ.fri.ˈkɐ.nu]/[ˈmu.si.kɐ.fri.ˈkɐ.nu] ‘African musician’, and vowel merger as in *aluna africana* [ɐ.ˈlu.na.fri.ˈkɐ.nɐ] ‘African student-FEM’) are not PW-bound, but span the intonational phrase (Frota, 2000). Like other Romance languages, but unlike Germanic languages, in EP resyllabification has the intonational phrase as its domain (e.g. Nespor & Vogel, 1986, Peperkamp, 1997, Vigário, 2003). This fact may yield a restructuring of the prosodic word, thus weakening the evidence for the PW in the input (see Vigário, 2003, p.35-37, for a review of possible analyses). Still, the clustering of numerous phenomena signalling the PW in EP, as described above, is an important property of the language that makes it closer to Germanic than to other Romance languages.

To sum up, EP phonology offers a robust set of phonological cues for the prosodic word. The availability of such strong evidence in the input may have consequences for prosodic word acquisition, as it may be the case that EP children are aware of (at least) some of these cues. Under this assumption, it is expected that early child speech will exhibit characteristics of word-based phonology, matching the target system.

2.2 Evidence for the PW in the acquisition data

A number of facts show that EP children are aware of cues for the prosodic word. In children's early production data, PW-edges are treated differently from word-internal positions. This is evidenced in the asymmetrical distribution of sounds within words and in the influence of position within the word in the realization of sounds.

Word-final coda fricatives have been shown to be mastered well before word-internal ones, as illustrated in (7) (Freitas, 1997, and Freitas, Miguel & Faria, 2001):

(7)	*festa*	/ˈfɛʃ.tɐ/ →	[ˈtɛtɐ]	(Inês: 1;9.19)	'party'
	gosta	/ˈgɔʃ.tɐ/→	[ˈgɔ]	(Inês: 1;8.2)	'(he/she) likes'

estas	/ˈɛʃtɐʃ/ →	[ˈɛtɐʃ]	(Inês: 1;10.29)	‘these-fem’
vs.				
bonecas	/bu.ˈnɛ.kɐʃ/→	[mɨˈɲɛkɐʃ]	(Inês: 1;9.19)	‘dolls’
bolos	/ˈbo.luʃ/→	[ˈboloʃ]	(Inês: 1;9.19)	‘cakes’
nariz	/nɐˈɾiʃ/ →	[ðɨˈɾiʃ]	(Inês: 1;10.29)	‘nose’

presta	/ˈpɾɛʃ.tɐ/→	[ˈpɛti]	(Marta: 1;5.17)	‘(it) is good’
gosta	/ˈgɔʃ.tɐ/→	[ˈkɔtɨ]	(Marta: 1;6.23)	‘(he/she) likes’
estes	/ˈeʃtɨʃ/ →	[ˈetiʃ]	(Marta: 1;5.17)	‘these-masc’
vs.				
bananas	/bɐ.ˈnɐ.nɐʃ/→	[ɐˈmɐnɐʃ]	(Marta: 1;4.8)	‘bananas’
maçãs	/mɐ.ˈsɐ̃ʃ/→	[mɨˈsɐ̃ʃ]	(Marta: 1;4.8)	‘apples’
nariz	/nɐ.ˈɾiʃ/→	[uˈdiʃ]	(Marta: 1;6.23)	‘nose’

A similar asymmetry is shown by the treatment of sequences of consonants that must be syllabified in different syllables (see Vigário & Falé, 1994, and Mateus and Andrade, 2000). Children avoid lexical items with these consonant clusters word-internally (there are only 0.19% of productions of items of this kind until age 3;7.24 – cf. Freitas 1997). However, they do select targets with such sequences word-initially, namely words starting with s+C clusters. The examples in (8) not only show the properties of the targets selected, but also children’s early acquisition of complex clusters at the left word-edge (Freitas, 1997, Freitas & Rodrigues, 2003).

(8) *estrela* /ʃ.ˈtɾe.lɐ/ → [ʃˈtɛlɐ] (Marta: 2;1.19) 'star'

esticar /ʃ.ti. ˈkaɾ/ → [ʃ:tiˈkaj] (Marta: 2;2.17) 'to stretch'

escova /ʃ.ˈko.vɐ/ → [ʃˈkovɐ] (Raquel: 2;5.19) 'brush'

estrela /ʃ.ˈtɾe.lɐ/ → [ʃˈtewɐ] (Raquel: 2;5.19) 'star'

In addition, in child speech the left word-edge plays an important role with respect to other properties of the target system. Preliminary work on the acquisition of place of articulation (POA) in EP shows that, at early stages, children use the same place of articulation for the entire word, matching the behaviour attested for Dutch (Levelt, 1994). By the time they start to assign POA to a specific root node, they begin by mastering consonants' place features at the word-left periphery, using the unmarked *Coronal* for word-internal consonants, as shown in (9) (Costa & Freitas, 2003).

(9) Inês (1;8.2 – 1;9.19)

umbigo	/ũˈbigu/	→	[ˈbidu]
copo	/ˈkɔpu/	→	[ˈpatu]
tampa	/ˈtɐ̃pɐ/	→	[ˈpɐ̃tɐ]
quarto	/ˈkʷaɾtu/	→	[ˈkatu]
folha	/ˈfoʎɐ/	→	[ˈkuʎɐ]

Similarly, children's word-initial unstressed vowels exhibit a different behaviour from word-internal unstressed vowels, matching the target system: namely, vowel reduction does not apply word-initially as it regularly does word-internally: e.g. *e̲le̲fante* [eliˈfɐ̃ti] 'elephant' never shows the reduced vowel [ɨ] at the left-edge unlike within the word. Other facts showing the relevance of the left-word edge are the preservation of the word-initial syllable or the preservation of the POA features of the initial consonant in trisyllabic targets produced as disyllabic tokens, as it will be shown in section 3.2 below.

We take the facts just reported as a confirmation of our expectations based on the properties of the input grammar: word-based phonology, matching the target system, emerges early in EP children's speech. Given that the grammatical properties of EP provide a robust set of cues for the PW and that early child speech does exhibit the presence of word-based phonology, it seems reasonable to hypothesize that the grammatical properties of the input could play a role in PW acquisition. The prediction would be that such properties facilitate the early segmentation and production (along the lines of e.g. Cutler, 1996, Demuth 1996, or Peters & Strömqvist 1996) of the different word shapes found in the language. In particular, the clustering of phenomena signaling the PW in EP would make the phonological form of PWs salient thus helping learners to focus on it early and to eventually produce PW shapes matching the target system early in the acquisition process. In the next section, adult speech and child speech data will be inspected to determine the frequency patterns of prosodic word shapes. These patterns, once established, will allow the potential relevance of the grammatical properties to be compared with the potential impact of statistical properties on child speech. This issue will be developed in the discussion section, together with implications of our predictions for the acquisition of the PW in other languages.

3. PROSODIC WORD SHAPES: A FREQUENCY STUDY

In the previous section we have characterized the input from the perspective of the grammatical system and its phonological properties. In the present section we investigate statistical properties of the input, namely frequency distributions of prosodic word shapes in European Portuguese and their import for children's early word shapes.

One of the topics that has raised increased interest in recent literature is the effect of input frequency on the lower and upper bounds of early prosodic words. As mentioned in the introduction, it has been hypothesized that the word shape frequency patterns shown by the input constrain the emergence and development of subminimal words and/or words with more than a binary foot (e.g. Demuth & Johnson, 2003, Lleó, 2004, Prieto, 2004, this volume). In line with this research trend, we set out to assess the extent of word shape frequency effects in prosodic word acquisition in EP. To address the issue, empirical databases of word shape frequencies both in adult speech and in child-directed speech have been established, and the predictions based on the frequency patterns of the ambient language put forward (section 3.1). The input frequency patterns were then compared with those shown by children's early productions (section 3.2). We conclude this section with a summary of our findings (section 3.3).

3.1. Word shape frequencies in the input

The child ambient language includes both child-directed speech (CDS) and adult speech (AS). As word shape frequencies in these two sources of input have not been previously compared in the EP literature, we have inspected the distribution of word shapes in both. Further motivation to include AS and compare it with CDS relies on two sorts of reports found in the literature: (i) although some important differences have been described between the two types of speech, the impact of CDS specific properties on child speech is not always clear (e.g. Ratner, 1996, among many others); (ii) Frota and Vigário (1995) have compared child speech intonational properties with the intonational properties of AS and CDS in EP and concluded that child speech is closer to adult speech than to child directed speech.

3.1.1. Data. The AS data is part of the *Português Falado* corpus, a spontaneous speech corpus of Portuguese developed at the Centro de Linguística da Universidade de Lisboa, and available in CD-ROM. For the present paper, the European Portuguese data from the 90s was used. This database, with a total of 23.459 phonological tokens (both prosodic words and clitics), was taken from Vigário, Martins & Frota (2005). The CDS data, like the child speech data examined in section 3.2, is part of a larger *corpus* of longitudinal cross-sectional data from several monolingual European Portuguese children (database on L1 acquisition of EP, developed at the Laboratório de Psicolinguística da Universidade de Lisboa). For the present paper, the adult productions uttered in the first three recording sessions of two children (Inês and Marta) have been analysed. These sessions took place at ages 0;11.14, 1;0.25 and 1;1.30 for Inês, and ages 1;2, 1;3.8 and 1;4.08 for Marta. Only the first three sessions were considered because CDS data collection is presently work in progress and these were the data already transcribed when the present study was conducted.[2] This is certainly a limitation of our study that we must take into account in the interpretation of our findings. The CDS database contains 23.207 phonological tokens (both prosodic words and clitics).

Both the AS and the CDS data were analysed using to a new tool – FreP – that allows the automatic extraction of prosodic words and prosodic clitics from orthographic transcriptions of EP speech, and the computation of the distribution of words shapes in number of syllables (Vigário, Martins & Frota, 2005). The third author checked by hand the automatic calculations generated by FreP. Reliability was above 99.5%. In this paper, we will concentrate on the shape of prosodic words, that is all open class lexical items plus stressed adverbials/prepositions like *agora* [ɐ.ˈgɔ.ɾɐ] 'now', *após* [ɐ.ˈpɔʃ] 'after', and stressed pronouns like *nós* [ˈnɔʃ] 'us', *nosso* [ˈnɔ.ssu] 'ours' (for a comprehensive list of prosodic words in EP according to morphosyntactic class, see Vigário, 2003). The total number of PWs included in the analysis was 16.507 for AS and 17.454 for CDS.

3.1.2. Results. The first stage of our calculations was the division of phonological tokens into two categories: PWs and clitics. The two types of speech display a similar breakdown of phonological tokens into prosodic words and clitics (74.4% of PWs and 25.6% of clitics in CDS; 70.4% and 29.6% in AS). We then examined the frequency patterns of the different prosodic word shapes (clitics excluded): monosyllabic, disyllabic, trisyllabic, and larger shapes (with more than three syllables). Within the monosyllabic category, the monosyllabic CV shape was also computed. This shape comprises all monosyllabic PWs that end with an oral vowel. Although there is no clear evidence for quantity-based effects in the phonology of EP (see section 2), the monosyllabic CV category was included for comparison with the monomoraic PW shape discussed in other studies. The monosyllabic shape was thus divided into CV and 'other', the latter category including all monosyllabic PWs that do not end with an oral vowel.

The word shape frequency patterns found in the input are shown in Figure 1. The two types of speech show similarities and differences. The frequency of disyllabic prosodic words is alike (47.1% in CDS; 44.4% in AS). The frequency of monosyllabic non-CV words is very similar (21.3% in CDS; 21.2% in AS). However, the frequency distribution of other word shapes shows a contrasting pattern. In CDS, monosyllabic CV shapes (1:CV) prevail over trisyllabic and larger PWs. In AS, monosyllabic PWs and PWs larger than a binary foot (3 and >3) have a similar frequency (28.6% and 27%, respectively). Likewise, AS shows monosyllabic CV shapes and long words (>3) in similar proportions (7.4% and 8%, respectively). The two main differences between AS and CDS reside in the lower frequency of trisyllables and larger PWs in CDS compared to AS, and the much higher proportion of CV monosyllables in CDS compared to AS.[3]

INSERT FIGURE 1 ABOUT HERE

Assuming that the frequency patterns found in the EP input may constrain the shape of children's early words, the two types of speech make different frequency-based predictions. The CDS patterns, with a high frequency of subminimal PWs and a very low frequency of trisyllabic and larger PWs, predict that early child speech will show a high incidence of subminimal words while complying with maximality constraints on word-size, namely words larger than disyllables should be avoided or truncated in early speech and only acquired at later stages. By contrast, the AS patterns, with a comparable amount of monosyllables and words larger than disyllables, as well as a similar proportion of monosyllabic CV and long words, predict that early child speech will show **both** the presence of subminimal words **and** larger words, thus not complying with constraints on word-size. It is therefore expected that both subminimal shapes and shapes with more than a binary foot, trisyllabic words in particular, will appear early in EP children's speech.

3.2. The shape of early words

3.2.1. Data. The child speech data (CS), like the CDS data, is part of a larger corpus of longitudinal cross-sectional data from several monolingual EP children. For the purposes of this paper, we will mainly consider the productions of the 3 youngest children: João, aged 0;10.2 to 2;0.19; Inês, aged 0;11.14 to 1;10.29; and Marta, aged to 1;2.0 to 2;0.26. The children were videotaped once a month by the second author (during 1993 and 1994); the sessions took place at the children's home, with the presence of the mother and the researcher, and the child data were collected in child-adult conversations. The duration of the sessions ranges from 30 to 60 minutes. The database format used to analyse the children's data is the CHILDPHON wordbase, developed at the Max Plank Institut for Psycholinguistics – Nijmegen (see Freitas 1997 for further details). Like in section 3.1, tokens comprise all the prosodic word forms uttered by the children (clitics are thus excluded), and were selected according to the following criteria: monosyllabic forms and multisyllabic forms with a single prominent syllable; all forms with external sandhi effects (i.e. segmental phenomena that apply across words), like the illustrative cases given in (10), were not considered.

(10)	*boneca outra*	/buˈnɛkɐ ˈotɾɐ/	→	[mɨˈɲɛˈkotɐ]	Marta: 1;09.19
	senta aqui	/ˈsẽtɐ ɐˈki/	→	[ˈtetaˈki]	Marta: 1;10.29
	outra agora é	/ˈotɾɐ ɐˈgɔɾɐ ˈɛ/	→	[ˈotɐˈgɔˈlɛ]	Marta: 1;10.29

Word forms with prosodic fillers, i.e. an initial syllable added to the segmental material that realizes the target word regardless of the grammatical category of this word, are included in the tokens (on filler syllables, see Peters, 2001, inter alia). When relevant prosodic fillers are kept apart in the calculations, this will be made explicit. The CS data analysed for the present paper comprises a total of 4.073 tokens.

3.2.2. Results. The child speech frequency results presented in Table 5 are based on the speech of the two children for which we also have the child-directed speech data, that is Inês and Marta. The two children have produced a total of 3381 tokens. The motivation for considering only the data from Inês and Marta for Table 5 is two-fold: (i) first, as the comparison with the CDS data is one of our main goals, it seemed advisable to consider only the speech of the children to which the adult speech analyzed before was directed; (ii) second, João is clearly less fluent than the other children (his productions only amount to 17% of all the data) and, unlike Inês and Marta, monosyllabic forms always predominate in his speech, even in the last recording sessions, whereas trisyllabic targets are very rarely attempted (they represent 1.7% of all the attempted targets).

INSERT TABLE 5 ABOUT HERE

Table 5 shows that disyllabic shapes predominate in child speech, as expected, with a frequency value not far from the frequency found in adult speech. However, the crucial data for the frequency based predictions advanced in section 3.1 concerns the frequency patterns of monosyllables (monosyllabic CVs in particular) and of words larger than disyllables. The incidence of monosyllabic shapes, as well as that of monosyllabic CVs, in CS is lower than expected on the basis of the CDS data alone: 28% of monosyllables in CS against 43% in CDS; 16% of CV monosyllables in CS against 22% in CDS. The frequency patterns shown by larger words in CS go in a similar direction: CS has twice as much trisyllabic and larger shapes than CDS (25% against 10%). By contrast, the frequency patterns of monosyllables, trisyllables and longer words closely resemble those found in AS. CS deviates both from AS and CDS in the lower proportion of monosyllabic non-CV shapes (12% against 21% in AS and CDS).

A first conclusion can be drawn from the results just reported. AS is a better predictor of CS word shapes than CDS. This is confirmed by the correlation results in Table 6: CS is highly correlated with AS and the correlation is statistically significant. However, as mentioned in section 3.1.1, our CDS database is limited to the first three recording sessions and thus the present finding has to be interpreted with caution. It may be the case that CDS evolves during the period of child speech observed, becoming closer to AS. In any event, the quasi-perfect match between the frequency distributions of PW shapes in CS and AS is a robust finding. This finding naturally entails that the first prediction based on AS frequency patterns was borne out: early CS shows both the presence of subminimal words and words larger than a binary foot.

INSERT TABLE 6 ABOUT HERE

Examples of the word shapes found in CS (for the three children) are given in (11) to (14). Tables 7 to 9 provide complete information on the number of appropriately produced tokens by target word shapes over time, for the three children.

INSERT TABLES 7, 8, AND 9 ABOUT HERE

(11) Monosyllabic targets

a. monosyllabic CV targets

dá	/ˈda/	→	[ˈdɐ]	João: 1;02.01	‘give-imp’
é	/ˈɛ/	→	[ˈɛ]	João: 1;02.01	‘(it) is’
há	/ˈa/	→	[ˈa]	João: 1;04.05	‘there is’
dá	/ˈda/	→	[ˈda]	Inês: 1;0.25	‘give-imp’
pé	/ˈpɛ/	→	[ˈpɛ]	Inês: 1;0.25	‘foot’
má	/ˈma/	→	[ˈma]	Inês: 1;1.30	‘bad’
dá	/ˈda/	→	[ˈda]	Inês: 1;03.06	‘give-imp’
é	/ˈɛ/	→	[ˈɛ]	Marta: 1;02.0	‘(it) is’
pé	/ˈpɛ/	→	[ˈpe]	Marta: 1;02.0	‘foot’
dá	/ˈda/	→	[ˈda]	Marta: 1;03.08	‘give-imp’

b. monosyllabic CVC targets

quer	/ˈkɛɾ/	→	[ˈkɛ]	João: 1;0.12	‘(he/she) wants’
cão	/ˈkɐ̃w̃/	→	[ˈkɐ]	João: 1;06.18	‘dog’
cais	/ˈkajʃ/	→	[ˈka]	João: 1;09.11	‘to fall – 2nd sg’
cão	/ˈkɐ̃w̃/	→	[ˈpɐ̃w̃]	João: 1;09.11	‘dog’
cão	/ˈkɐ̃w̃/	→	[ˈkɐ̃w̃]	João: 1;11.13	‘dog’
cais	/ˈkajʃ/	→	[ˈkajʃ]	João: 1; 11.13	‘to fall – 2nd sg’
quer	/ˈkɛɾ/	→	[ˈkɛ]	Inês: 0;11.14	‘(he/she) wants’

mão	/ˈmɐ̃w̃/	→	[ˈmɐ]	Inês: 1;1.30	'hand'
cão	/ˈkɐ̃w̃/	→	[ˈka]	Inês: 1;1.30	'dog'
mais	/ˈmajʃ/	→	[ˈmɐ]	Inês: 1;06.06	'more'
quer	/ˈkɛɾ/	→	[ˈkɛj]	Inês: 1; 08.02	'(he/she) wants'
dois	/ˈdojʃ/	→	[ˈduʃ]/[ˈdojʃ]	Inês: 1;09.19	'two'
mais	/ˈmajʃ/	→	[ˈmaʃ:]/[ˈmaj]/[ˈmajʃ]	Inês: 1;09.19	'more'
quer	/ˈkɛɾ/	→	[ˈkɛ]	Marta: 1;02.0	'(he/she) wants'
cão	/ˈkɐ̃w̃/	→	[ˈkɔ]	Marta: 1;03.08	'dog'
quer	/ˈkɛɾ/	→	[ˈkɛ]	Marta: 1; 06.23	'(he/she) wants'
cais	/ˈkajʃ/	→	[ˈka]	Marta: 1;04.08	'to fall – 2nd sg'
cão	/ˈkɐ̃w̃/	→	[ˈkɐ̃w̃]	Marta: 1; 05.17	'dog'
quer	/ˈkɛɾ/	→	[ˈkɛj]	Marta: 1; 05.17	'(he/she) wants'

Monosyllabic words in CS, and monosyllabic CV shapes in particular, not only appear in large numbers, but they are also produced frequently even at later stages: from 1;08 onwards, monosyllabic CVs amount to 58% of all monosyllabic tokens for João, 59% for Inês, and 32% for Marta (see the examples in (11) above). Monosyllabic non-CV targets can be produced as CV (11b). Also important is the course of development of disyllabic targets (see the examples in (12) and (13)). Truncation to monosyllabic shapes is possible and common, even in the later stages. Illustrative examples of truncation to monosyllabic non-CV shapes can be found in (12). Examples of truncation to monosyllabic CV shapes in the initial and later stages are given in (13). Note that both iambic and trochaic targets can be truncated, the common feature being the preservation of the stressed syllable. These patterns are consistent across the three children's data. Although trochaic forms are more frequent in the input (see section 2.1), in both types of targets the unstressed vowel is typically reduced. The preservation of the stressed syllable is thus in line with the prominence-related cues to the PW in EP and the asymmetry between stressed and unstressed syllables that characterizes the language. We will return to this finding in section 4.

(12) Disyllabic targets (iambic and trochaic)

atirou	/ɐtiˈɾo/	→	[ˈtiw]	João: 1;05.13	'daddy'
fugiu	/fuˈʒiw/	→	[ˈʃiw]	Inês: 1;01.30	'(he/she) ran away'
balão	/bɐˈlɐ̃w̃/	→	[ˈlaw]	Marta: 1;03.08	'balloon'
caiu	/kɐˈiw/	→	[ˈkiw]	Marta: 1;03.08	'(it) fell'
água	/ˈagwɐ/	→	[ˈa:]	João: 1;03.11	'water'
praia	/ˈpɾajɐ/	→	[ˈpa]	João: 2;0.19	'beach'
água	/ˈagwɐ/	→	[ˈa]	Inês: 1;01.30	'water'
água	/ˈagwɐ/	→	[ˈaw]	Marta: 1;02.0	'water'
caixa	/ˈkajʃɐ/	→	[ˈkaʃ]	Marta: 1;03.08	'box'
avó	/ɐˈvɔ/	→	[ɐ.ˈdɐ]	João: 1;0.12	'grandmother'
mamã	/mɐˈmɐ̃/	→	[mɐ.ˈmɐ̃]	João: 1;05.13	'mummy'
mamã	/mɐˈmɐ̃/	→	[mɐ.ˈmɐ̃]	Inês: 0;11.14	'mummy'
Inês	/iˈneʃ/	→	[i. ˈni]	Inês: 1;01.30	'Inês (name)'
bebé	/bɛˈbɛ/	→	[βɐ.ˈbɛ]	Marta: 1;02.0	'baby'
maçãs	/mɐˈsɐ̃ʃ/	→	[mɨ.ˈʃaʃ]	Marta: 1;05.17	'apples'

tia	/ˈtiɐ/	→	[ˈti.a]	João: 1;08.13	‘aunt’
anda	/ˈɐ̃dɐ/	→	[ˈɐ.nɐ]	João: 1;09.11	‘come’
bolso	/ˈboɫsu/	→	[ˈbɔ.tɐ]	Inês: 1;08.02	‘pocket’
copo	/ˈkɔpu/	→	[ˈpa.ku]	Inês: 1;08.02	‘glass’
pato	/ˈpatu/	→	[ˈpa.tɨ]	Marta: 1;02.0	‘duck’
quadro	/ˈkwadɾu/	→	[ˈka.lu]	Marta: 1;07.18	‘painting’

(13) Truncation of disyllabic targets

a. disyllabic iambic targets

papá	/pɐˈpa/	→	[ˈpa]	João: 0;10.02	‘daddy’
titi	/tiˈti/	→	[ˈdi]	João: 1;03.04	‘aunt-dim’
papá	/pɐˈpa/	→	[ˈpa]	João: 1;06.18	‘(he/she) threw (it)’
aqui	/ɐˈki/	→	[ˈki]	João: 1;07.24	‘here’
avô	/ɐˈvo/	→	[ˈbo]	João: 1;11.13	‘grandfather’
papá	/pɐˈpa/	→	[ˈpa]	Inês: 0;11.14	‘daddy’
bebé	/bɛˈbɛ/	→	[ˈbi]	Inês: 1;01.30	‘baby’
chapéu	/ʃɐˈpɛw/	→	[ˈpɐ]	Inês: 1;01.30	‘hat’
Inês	/iˈneʃ/	→	[ˈne]	Inês: 1;03.06	‘Inês (name)’
aqui	/ɐˈki/	→	[ˈi]	Marta: 1;02.0	‘here’
fugiu	/fuˈʒiw/	→	[ˈʒu]	Marta: 1;03.08	‘(he/she) ran away’

b. disyllabic trochaic targets

papa	/ˈpapɐ/	→	[ˈpa]	João: 1;03.04	‘food’
pato	/ˈpatu/	→	[ˈpa]	João: 1;06.18	‘duck’

bola	/ˈbɔlɐ/	→	[ˈu]	João: 1;09.11	'ball'
toma	/ˈtɔmɐ/	→	[ˈtɔ]	João: 2;0.19	'take (it)'
toma	/ˈtɔmɐ/	→	[ˈtɔ]	Inês: 1;0.25	'take (it)'
manta	/ˈmɐ̃tɐ/	→	[ˈmɐ]	Inês: 1;03.06	'blanket'
barco	/ˈbaɾku/	→	[ˈba]	Inês: 1;03.06	'boat'
mostra	/ˈmɔʃtɾɐ/	→	[ˈmɔ]	Inês: 1;04.09	'show (me)'
meia	/ˈmɐjɐ/	→	[ˈme]	Inês: 1;05.11	'sock'
livro	/ˈlivɾu/	→	[ˈti]	Marta: 1;02.0	'book'
chama	/ˈʃɐmɐ/	→	[ˈʃɐ]	Marta: 1;02.0	'call (it)'

Both the frequency of monosyllabic shapes in CS and the possible reduction of non-CV monosyllables to CV shapes, as well as the possible truncation of disyllabic targets to monosyllabic CVs, show that early words in EP are not constrained by minimality requirements.

Words larger than disyllabic shapes are not avoided in EP child speech either, as shown by their frequency distribution patterns (Table 5), by the course of word shape development (Tables 8 and 9), and by the illustrative cases of trisyllabic targets uttered as trisyllabic PWs given in (14).

(14) Trisyllabic targets

sa*pato*	/sɐˈpatu/	→	[ˈpa]	João: 1;07.24	'shoe'
sapato	/sɐˈpatu/	→	[ˈbɐ]	Inês: 1;05.11	'shoe'
cadeira	/kɐˈdɐjɾɐ/	→	[ˈka]	Marta: 1;03.08	'banana'
abelha	/ɐˈbɐʎɐ/	→	[ˈa.bɐ]	João: 1;08.13	'bee'

sapato	/sɐˈpatu/	→	[pɐ.ˈpɐ]	Inês: 1;04.09	‘shoe’
caracol	/kɐrɐˈkɔɫ/	→	[ˈto.ɐ]	Marta: 1;03.8	‘snail’
orelhas	/oˈrɐʎɐʃ/	→	[a.ˈge.ge]	Inês: 1;05.11	‘ears’
banana	/bɐˈnɐnɐ/	→	[ɐ.ˈmɐ.nɐ]	Inês: 1;05.11	‘banana’
leitinho	/lɐjˈtiɲu/	→	[ɨ.ˈti.ɲi]	Inês: 1;05.11	‘milk-dim’
macaco	/mɐˈkaku/	→	[ɐ.ˈka.ko]	Inês: 1;07.02	‘monkey’
sapato	/sɐˈpatu/	→	[pa.ˈta.tɐ]	Inês: 1;07.02	‘shoe’
ajuda	/ɐˈʒudɐ/	→	[ɐ.ˈdu.dɐ]	Inês: 1;08.02	‘help (me)’
piscina	/piʃˈsinɐ/	→	[pi. ˈti.nɐ]	Inês: 1;09.19	‘swimming pool’
chupeta	/ʃuˈpetɐ/	→	[ʃuj.ˈʃi.ʃi]	Marta: 1;02.0	pacifier’
sapato	/sɐˈpatu/	→	[ɤɐ.ˈpa.θɐ]	Marta: 1;02.0	‘shoe’
bolacha	/buˈlaʃɐ/	→	[ɨ.ʃa.ʃɐ]	Marta: 1;03.08	‘cookie’
banana	/bɐˈnɐnɐ/	→	[mɐ.ˈnɐ:.nɐ]	Marta: 1;03.08	‘banana’
chupeta	/ʃuˈpetɐ/	→	[ʃi.ˈʃi.ʃɨ]	Marta: 1;04.08	pacifier’
banana	/bɐˈnɐnɐ/	→	[mɐ.ˈmɐ.nɐ]	Marta: 1;04.08	‘banana’
caneta	/kɐˈnetɐ/	→	[kɨ.ˈle.lɐ]	Marta: 1;05.17	‘pen’
coelho	/kuˈɐʎu/	→	[kɨ.ˈke.ʎu]	Marta: 1;05.17	‘rabitt’
barulho	/bɐˈɾuʎu/	→	[bɨ.ˈlu.ʎu]	Marta: 1;11.10	‘noise’
gatinho	/gɐˈtiɲu/	→	[gɐ.ˈti.ɲɨ]	Marta: 2;00.26	‘cat-dim’

Figure 2 shows a detailed picture of the course of development of trisyllabic targets, for the two children (Inês and Marta) that have such targets in their vocabulary. The data in Figure 2 excludes word shapes with prosodic fillers, so that there remains no doubt as to the trisyllabic nature of the forms produced (filler syllables in EP are added to children's productions of target words independently of grammatical category, as it will be shown below).[4]

INSERT FIGURE 2 ABOUT HERE

As depicted in Figure 2, the first trisyllabic targets appear very early (1;01), trisyllabic word shapes for trisyllabic targets also emerge early in child speech (Marta:1;02; Inês:1;07), and are mastered soon in the path of phonological development (i.e. over 60% of accuracy at 1;05 for Marta and 1;09 for Inês). The second prediction based on the AS frequency patterns is thus borne out: word shapes with more than a binary foot do appear early in EP child speech (EP is thus similar to Spanish, as reported in Lleó, 2004).

Figure 2 also shows that there is, nevertheless, an initial stage when trisyllabic targets are truncated to dyllabic shapes. The interesting fact here is that besides deleting the first syllable, two other strategies are also available: Inês favours the deletion of the last syllable (and thus iambic truncated shapes predominate), whereas Marta deletes the initial syllable, preserving the place features of the PW-initial consonant (and thus trochaic truncated shapes predominate). These two strategies are illustrated in (15).

(15) Inês: CV'CV<u>CV</u> → CV'CV

boneca	/buˈnɛkɐ/ →	[meˈɲɛ]	(1;05.11)	'doll'
mamoca	/mɐˈmɔkɐ/ →	[mɐˈmɔ]	(1;05.11)	'little breast'

laranja	/lɐˈɾɐ̃ʒɐ/ →	[lɐˈlaː]	(1;05.11)	‘orange’
queijinho	/kɐjˈʒiɲu/ →	[kɛˈki]	(1;06.06)	‘cheese-dim’

Marta: CV'CVCV → 'CVCV

menino	/mɨˈninu/ →	[ˈmunu]	(1;03.08)	‘boy’
morangos	/muˈɾɐ̃guʃ/ →	[ˈmɐ̃guʃ]	(1;05.17)	‘strawberrys’
querido	/kɨˈɾidu/ →	[ˈkidu]	(1;05.17)	‘dear’
cenoura	/sɨˈnoɾɐ/ →	[ˈsoɾɐ]	(1;10.04)	‘carrot’

What the two strategies have in common is the preservation of the consonant (place features) from the left edge of the PW (besides the preservation of the stressed syllable), regardless of other factors such as the resulting stress pattern of the form produced. We will further comment on these findings in section 4.

Another issue relevant to the understanding of PW shapes in child speech is the presence/absence of prosodic fillers, that is an initial syllable added to the segmental material that realizes the target word, regardless of the grammatical category of the target word (fillers appear not only with nouns, but also with verbs, adverbs, or pronouns, as shown in (16) below). Prosodic fillers in EP occur from the first productions and **independently** of word size: they appear both with monosyllabic and disyllabic, thus yielding respectively disyllabic and trisyllabic word shapes. Illustrative examples are provided in (16).

(16)	*mãe*	/ˈmɐ̃j/ →	[ɨ.ˈmɐ]/[ˈmɐ]	João: 0;11.06	‘mother’
	pato	/ˈpatu/ →	[ɐ.ˈtɐ]/[ˈtɐ]	João: 0;11.06	‘duck’
	mãe	/ˈmɐ̃j/ →	[ɨ.ˈmɐ]/[ˈmɐ]	João: 0;11.06	‘mother’

pato	/ˈpatu/	→	[ɐ.ˈtɐ]/[ˈtɐ]	João: 0;11.06	‘duck’
não	/ˈnɐ̃w̃/	→	[ɨ.ˈnɐ]	João: 1; 10.08	‘no’
toma	/ˈtɔmɐ/	→	[ɐ.ˈdɔ]	Inês: 1;0.25	‘take (it)’
dá	/ˈda/	→	[ɐ.ˈda]/[ˈda]	Inês: 1;0.25	‘give (me)’
chupeta	/ʃuˈpetɐ/	→	[ɐ.ˈpi]/[ˈpi]	Inês: 1;01.30	‘pacifier’
põe	/ˈpõj̃/	→	[ɨ.ˈpo]	Inês: 1;04.09	‘put (it)’
mais	/ˈmajʃ/	→	[ɨ.ˈmɐ]	Inês: 1;05.11	‘more’
cão	/ˈkɐ̃w̃/	→	[ɐ.ˈkɐ̃w̃]/[ˈkɐ̃.u]	Marta: 1;02.0	‘dog’
tem	/ˈtɐ̃j̃/	→	[ɨ.ˈtɐ]	Marta: 1;02.0	‘(he/she) has’
outra	/ˈotrɐ/	→	[ɨ.ˈo.tɐ]	Marta: 1;04.08	‘other’
boca	/ˈbokɐ/	→	[ɔ.ˈbɔ:.tɐ]	Marta: 1;04.08	‘mouth’
mola	/ˈmɔlɐ/	→	[ɨ.mɔˈ.lɐ]/[ˈmɔ.lɐ]	Marta: 1;05.17	‘spring’
olha	/ˈɔ.ʎɐ/	→	[o. ˈɔ.ʎɐ]	Marta: 1;06.23	‘look’
fugiu	/fuˈʒiw/	→	[ɨ.gɨ.ˈʃiw]	Marta: 1;06.23	‘(he/she) ran away’
isto	/ˈiʃtu/	→	[ɐ.ˈi.tu]	Marta: 1;06.23	‘this’

Crucially, in EP prosodic fillers are not inserted to obtain a disyllabic Word shape, nor are they obligatory at a given stage as productions with and without fillers usually co-occur.

3.3. Summary of findings

The analysis of PW shape frequencies in the input has shown that they constrain early word shapes. Adult speech, in particular, provides a quasi-perfect match with the frequency distributions of PW shapes in CS: early CS shows both the presence of subminimal words and words larger than a binary foot. As predicted by the input frequency patterns, trisyllabic targets appear very early in child speech (Inês: 1;01; Marta: 1;02) and trisyllabic word shapes are produced quite early as well (Marta: 1;02; Inês:1;07). Monosyllabic word shapes, and monosyllabic CV shapes in particular, remain very frequent in child production until the later stages. Finally, the presence of prosodic fillers in early child speech was shown to be independent of word size. This set of findings strongly suggests that early words in EP are not constrained by minimality or maximality requirements.

4. Discussion: Grammar and frequency effects

From the results reported in the previous section, it seems clear that the statistical properties of the input contribute to explain prosodic word acquisition in EP in ways similar to recent reports on other languages (e.g. Demuth & Johnson, 2003, Lleó, 2001b, 2004, Prieto 2004). However, this conclusion does not necessarily entail that the grammatical properties of the input described in section 2 are given no role. In section 2, we established that EP phonology offers a robust set of cues for the prosodic word, and the presence of early word based phonology in child speech has shown that EP children are aware of these cues. The hypothesis was entertained that the grammatical properties of the input could constrain PW acquisition. In EP, both the grammar and frequency effects concur to promote the early production of the different word shapes found in the language.

The availability of strong evidence in the input grammar for the PW may contribute to strengthen the frequency effects. As PW-edges are relatively well-delimited (certainly better delimited than in other Romance languages), a tendency is expected to faithfully reproduce word edges even in early child speech. Trisyllabic targets offer an interesting example of a possible interaction between grammar and frequency effects. We have seen that trisyllabic words are acquired early by EP children. There is, nevertheless, an initial stage involving truncation of the syllables of the target word shape. Two of the strategies that children display – deletion of the last syllable or deletion of the initial syllable but keeping the place features of the PW initial consonant (see section 3.2) – involve the preservation of the PW left edge, which is the most prominent word edge in EP grammar, i.e. the one that is signaled by a stronger set of phonological cues (see section 2.1). Thus, EP children not only produce trisyllabic targets early, but also show the relevance of the PW left edge in their initial truncated forms.

Disyllabic targets also provide a case of a possible interaction between grammatical cues to PW structure and word shape frequency. We have seen that truncation to monosyllabic shapes is possible and frequent even at the later stages, showing the absence of a minimality constraint on word size. Truncation affects both iambic and trochaic targets and typically involves the preservation of the stressed syllable. This truncation pattern is in line with the prominence-related cues to the PW in EP, namely the salience of word stress, and the asymmetry between stressed and unstressed vowels, the latter being subject to reduction both in pre- and post-tonic position (see section 2.1).

Another instance of a grammatical property that may enhance frequency patterns is the main direction of cliticization in EP. As described in section 2.1, most clitics in EP are proclitics, yielding an enlargement of the different prosodic word shapes in the language. Thus, proclitics increment the frequency of disyllables, trisyllables and larger words, adding to the evidence available against a maximality constraint on word-size in EP and possibly helping to promote the early emergence of trisyllabic words.

A third grammatical property that may affect the way prosodic words are acquired in EP is resyllabification across words and other processes spanning the intonational phrase. Unlike all the other grammatical properties mentioned in this paper, resyllabification weakens the evidence for the prosodic word in EP and makes it less strong than the evidence for the prosodic word generally found in Germanic languages (see section 1).

Since both grammar and frequency concur to yield the same kind of effects on EP prosodic word acquisition, it is important to examine potential grammatical and frequency interactions in languages different from EP. Although the data available is still scarce, the literature provides us with same relevant cases. We will take the emergence of shapes larger than a binary foot as a test case, and consider languages as different as English, Spanish and Catalan.

As a Germanic language, grammatical cues for the PW in English are stronger than in EP. Spanish and Catalan, by contrast, are known to display weaker cues for the PW than EP (see section 1). Considering grammar alone, one would expect the production of the different word shapes to come earlier in English and in EP than in Spanish and Catalan. However, the data reported in the literature shows that shapes larger than a binary foot emerge early in Spanish, like in EP, and later in English and Catalan (Demuth & Johnson, 2003, Prieto, 2004, Lleó, 2004). The frequency of word shapes larger than a binary foot in these languages' input places Spanish and EP first with similar frequencies (around 30% in Spanish and 27% in EP), followed by Catalan with 15% and finally English with around 5% (Roak & Demuth 2000, Prieto

2004). Based on frequency alone, it would be expected that words larger than a binary foot should emerge early in Spanish and EP, later in Catalan, and even later in English. Although they do appear early in both EP and Spanish, they seem to emerge equally late in Catalan and English despite the considerable frequency difference (Lleó & Demuth, 1999, Prieto, 2004). This result, unexplained on a mere frequency basis, would follow straightforwardly if a grammar and frequency interaction is assumed: although the frequency is considerably higher in Catalan, the cues provided by the grammar are stronger in English.

As already pointed out, there is little data available with which to fully assess the role that grammar effects, frequency effects and their interactions may play in prosodic word development. It is hoped that the findings reported in this paper will contribute to our understanding of the factors that may influence the acquisition of prosodic words.

5. CONCLUSION

This paper investigates the acquisition of prosodic words in European Portuguese through the analysis of grammatical and statistical properties of the target language and child speech. It was shown that the adult target system offers a constellation of properties cuing the prosodic word, namely edge-related, word-bound, and prominence phenomena. The presence of early word-based phonology in child speech suggests that children are aware of these cues. The hypothesis was put forward that the presence of such strong evidence could promote the early development of the different shapes of the prosodic word found in the ambient language.

The observation of the statistical properties of the input (both in adult speech and child-directed speech) and of early child speech has lead us to conclude that they do influence the shape of early words and thus play an important part in the understanding of prosodic word acquisition in EP. Word shape frequency patterns in adult speech were found to be a better predictor of word shapes in child speech than the frequency patterns that characterize child-directed speech. Early child speech shows both the presence of subminimal words and words larger than a binary foot, and in proportions similar to those found in adult speech. As expected on the basis of the input frequency patterns, EP early word shapes (like French) are not constrained by minimality effects (like those reported for English – cf. Demuth & Johnson, 2003). Thus, input information seems to promote the low ranking of the minimal word constraint early in the course of development, matching the target system (see Bisol, 2000 and Vigário, 2003, on the absence of minimal word requirements in Portuguese). Likewise, early word shapes in EP are not constrained by maximality effects (like in Spanish, but unlike in Catalan or English – Lleó, 2004, Prieto, 2004, Roark & Demuth, 2000). Input information seems to enforce the low ranking of the maximality constraint in EP, thus promoting the early emergence of trisyllabic prosodic words.

Given the grammatical properties of the EP target system and the word shape frequency properties that characterize the EP input, it was argued that possible grammar effects on prosodic word acquisition will strengthen the frequency-based effects: rather well-delimited and salient prosodic words are expected to favor early prosodic word development, enhancing the trends shown by the frequency patterns. The assessment of the cross-linguistic differences predicted on the basis of the proposed interaction between grammar and frequency effects requires the expansion of the available databases on the acquisition of the prosodic word to which the present Volume makes an important contribution.

FOOTNOTES

[1] [ɐ] may also appear in stressed position when followed by a nasal segment (as a result of an independent raising process) – in which case it becomes nasal ([ɐ̃]) if the nasal segment is tautosyllabic – or when followed by a palatal segment (as a consequence of the centralization of an underlying /e/ in this context: cf. Mateus, 1975, Vigário, 2003: ch.3).

[2] The CDS data was taken from transcriptions made by Ana Lúcia Santos and Carla Soares, respectively for Inês and Marta. These transcriptions are part of a larger data collection using the CHILDES format (Santos, in progress, Soares, in progress).

[3] We believe that these differences are due both to a type issue and a token issue: the choice of short lexical items and the avoidance of longer words, together with the repetition of the same CV monosyllables (like *é* '(it) is', *dá* 'give (me)', *já* 'already', *lá* 'there', or *tu* 'you').

[4] The exclusion of productions with fillers and respective targets is behind the differences found between Figure 2 and Tables 8-9: in the latter, as stated in section 3.2.1, prosodic fillers were included in the tokens, and thus trisyllabic targets produced with a filler (i.e. tokens with 4 syllables) count as not matching the target word; in the former, all such cases with filler syllables were excluded from the calculations.

REFERENCES

BACELAR, M. F., MARQUES, M. L. & SEGURA DA CRUZ, M. L. (1987). *Português Fundamental. Vol. II. Métodos e Documentos*, *Tomo 2: Inquérito de Disponibilidade*. Lisboa: INIC/CLUL.

BECKMAN, M., & EDWARDS, J. (2000). Lexical frequency effects on young children's imitative productions. In M. B. Broe and J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V. Acquisition and the Lexicon* (pp. 250-268). Cambridge: Cambridge University Press.

BISOL, L. (2000). O Clítico e o seu Status Prosódico. *Revista de Estudos de Linguagem UFMG*. Belo Horizonte, 9(1): 5-30.

BOOIJ, G. (1995). *The Phonology of Dutch*. Oxford: Clarendon Press.

BOOIJ, G. (1999). The Role of the Prosodic Word in Phonotactic Generalizations. In T. A. Hall and U. Kleinhenz (Eds.), *Studies on the Phonological Word* (pp. 47-72). Amsterdam: John Benjamins.

BYBEE, J. (2000). Lexicalization of sound change and alternating environments. In M. B. Broe and J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V. Acquisition and the Lexicon* (pp. 250-268). Cambridge: Cambridge University Press.

BYBEE, J. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

BYBEE, J., & HOPPER, P. (2001) (Eds.) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

COSTA, T. & FREITAS, M. J. (2003). Obstruções na aquisição. A propósito da natureza dos estádios iniciais de desenvolvimento segmental. In A. Mendes and T. Freitas (Eds.), *Actas do XVIII Encontro Nacional da Associação Portuguesa de Linguística* (pp. 275-292). Lisboa: APL

CUTLER, A. (1996). Prosody and the Word Boundary Problem. In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (pp. 87-99). Mahwah, NJ: Laurence Erlbaum Associates.

DEMUTH, K. (1996). The Prosodic Structure of Early Words. In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (pp. 171-184). Mahwah, NJ: Laurence Erlbaum Associates.

DEMUTH, K., & JOHNSON, M. (2003). Truncation to subminimal words in Early French. *Canadian Journal of Linguistics*, 48, 211-241.

FIKKERT, P., & FREITAS, M. J. (1998). Acquisition of syllable structure constraints: Evidence from Dutch and Portuguese. In *Proceedings of the GALA'97 Conference on Language Acquisition* (pp. 217-222). Edinburgh: University of Edinburgh.

FREITAS, M. J. (1997). *Aquisição da Estrutura Silábico do Português Europeu.* PhD dissertation, University of Lisbon.

FREITAS, M. J., MIGUEL, M., & FARIA, I. (2001). Interaction between Prosody and Morphosyntax: plurals within codas in the acquisition of European Portuguese. In J. Weissenborn and B. Höhle (Eds.), *Approaches to Bootstrapping. Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (pp. 45-58). Amsterdam/Philadelphia: John Benjamins.

FREITAS, M. J. & RODRIGUES, C. (2003). On the nature of sC-clusters in European Portuguese. *Journal of Portuguese Linguistics,* 2 (2), 55-86 (Special Issue on Portuguese Phonology, Guest-edited by L. Wetzels).

FROTA, S. (2000). *Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation*. New York: Garland Publishing.

FROTA, S. & VIGÁRIO, M. (1995). The intonation of one European Portuguese Infant: a first approach. In I. H. Faria and M. J. Freitas (Eds.), *Studies on the Acquisition of Portuguese* (pp.17-34). Lisboa: APL/Colibri.

HALL, T. A. (1999). Phonotactics and the Prosodic Structure of German Function Words. In T. A. Hall and U. Kleihenz (Eds.), *Studies on the Phonological Word* (pp. 99-131). Amsterdam: John Benjamins.

HANNAHS, S. (1995a). *Prosodic Structure and French Morphophonology*. Tübingen: Max Niemeyer Verlag.

HANNAHS, S. (1995b). Glide formation, prefixation, and the phonological word in French. In J. Amastae et al. (Eds.), *Contemporary Research in Romance Linguistics* (pp. 13-24). Amsterdam/Philadelphia: John Benjamins Publishers.

JURAFSKY, D., BELL, A., & GIRAND, C. (2002). The role of lemma in form variation. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology 7* (pp. 3-34). Berlin/New York: Mouton de Gruyter.

KLEINHENZ, U. (1997). Domain Typology at the Phonology-Syntax Interface. In G. Matos et al. (Eds.), *Interfaces in Linguistic Theory* (pp. 201-220). Lisboa: APL/Colibri.

LEVELT, C. (1994) *On the Acquisition of Place*. Dordrecht: HIL.

LLEÓ, C. (2001a) The Transition From Prenominal Fillers to Articles in Spanish and German. In M. Almgren, A, Barreña, M.-J. Ezeizabarrena, I. Idiazabal, and B. MacWhinney (Eds.), *Research on Child Language Acquisition. Proceedings of the 8th Conference of the International Association for the Study of Cild Language* (pp. 713-737). Somerville: Cascadilla Press.

LLEÓ, C. (2001b). The Interface of Phonology and Syntax. The emergence of the article in the early acquisition of Spanish and German. In J. Weissenborn and B.

Höhle (Eds.), *Approaches to Bootstrapping. Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (pp. 23-44). Amsterdam/Philadelphia: John Benjamins.

LLEÓ, C. (2004). The Acquisition of prosodic word structures in Spanish. Paper given at the *Second Lisbon Meeting on Language Acquisition – with special reference to Romance Languages*, Lisbon.

LLEÓ, C., & DEMUTH, K. (1999). Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and Romance languages. In A. Greenhill, H Littlefield, and C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development* (pp. 407-418). Somerville, MA: Cascadilla Press.

MATEUS, M. H. (1975). *Aspectos da Fonologia do Português*. Lisboa. INIC.

MATEUS, M. H., & ANDRADE, E. (2000). *The Phonology of Portuguese*. Oxford: Oxford University Press.

MOATES, D. R., BOND, Z. S., & STOCKMAL, V. (2002). Phoneme frequency in spoken word reconstruction. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology 7* (pp. 141-169). Berlin/New York: Mouton de Gruyter.

MORGAN, J. L., SHI, R., & ALLOPENA, P. (1996). Perceptual Bases of Rudimentary Grammatical Categories: Toward a Broader Conceptualization of Bootstrapping. In J.

L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (pp. 263-283). Mahwah, NJ: Lawrence Erlbaum Associates.

NESPOR, M., & VOGEL, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

PEPERKAMP, S. (1997). *Prosodic Words*. HIL Dissertations 34. The Hague: Holland Academic Graphics.

PETERS, A. M. (2001). Filler Syllables in Early Language Development. In M. Almgren, A, Barreña, M.-J. Ezeizabarrena, I. Idiazabal, and B. MacWhinney (Eds.), *Research on Child Language Acquisition. Proceedings of the 8th Conference of the International Association for the Study of Cild Language* (pp. 697-700). Somerville: Cascadilla Press.

PETERS, A. M., & STRÖMQVIST, S. (1996). The Role of Prosody in the Acquisition of Grammatical Morphemes. In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (pp. 215-232). Mahwah, NJ: Laurence Erlbaum Associates.

PIERREHUMBERT, J. B. (2002). Word-specific phonetics. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology 7* (pp. 101-139). Berlin/New York: Mouton de Gruyter.

PRIETO, P. (2004). Early prosodic word acquisition in Catalan. Paper given at the *Second Lisbon Meeting on Language Acquisition – with special reference to Romance Languages*, Lisbon.

RAFFELSIEFEN, R. (1999). Diagnostics for Prosodic Words Revisited: The Case of Historically Prefixed Words in English. In T. A. Hall and U. Kleinhenz (Eds.), *Studies on the Phonological Word* (pp. 133-201). Amsterdam: John Benjamins.

RATNER, N. B. (1996). From "Signal to Syntax": But What Is the Nature of the Signal? In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping From Speech to Grammar in Early Acquisition* (pp. 135-150). Mahwah, NJ: Lawrence Erlbaum Associates.

ROARK, B., & DEMUTH, K. (2000). Prosodic constraints and the learner's environment: A corpus study. In S. K. Howell, S. A. Fish, and T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 597-608). Somerville, MA: Cascadilla Press.

SANTOS, A. L. (in progress). Verbal answers to yes/no questions. A syntax study based on longitudinal corpora from Portuguese children. PhD dissertation, University of Lisbon.

SCHWINDT, L. C. (2000). *O prefixo no Português Brasileiro: análise morfofonológica*. PhD dissertation, Pontifícia Universidade Católica do Rio Grande do Sul.

SOARES, C. (in progress). *La périphérie gauche dans le processus de l'acquisition du Portugais Européen : étude longitudinale de la production spontanée de 3 enfants monolingues de 1;2.0 à 4;5.19 ans*. PhD dissertation, Université de Paris 8.

VIANA, M. C, TRANCOSO, I. M., SILVA, F. M., MARQUES, G., d'ANDRADE, E., & OLIVEIRA, L. C. (1996). Sobre a pronúncia de nomes próprios, siglas e acrónimos em Português Europeu. In I. Duarte and I. Leiria (Eds.), *Actas do Congresso Internacional Sobre o Português* (pp. 481-519). Lisboa: APL e Edições Colibri.

VIGÁRIO, M. (2003). *The Prosodic Word in European Portuguese*. Berlin/New York: Mouton de Gruyter.

VIGÁRIO, M. & FALÉ, I. (1994). A Sílaba no Português Fundamental: uma descrição e algumas considerações de ordem teórica. In *Actas do IX Encontro da Associação Portuguesa de Linguística* (pp. 465-477). Lisboa: APL/Colibri.

VIGÁRIO, M., MARTINS, F., & FROTA, S. (2005). Frequências no Português Europeu: A ferramenta FreP. In I. Duarte and I. Leiria (eds.), *Actas do XX Encontro da Associação Portuguesa de Linguística* (pp.897-908). Lisboa: APL/Colibri.

WIESE, R. (1996). *The Phonology of German*. Oxford: Clarendon Press.

TABLES

Table 1. Sound inventory of European Portuguese

Vowels	Stressed system			Unstressed system		
	Front	Central	Back	Front	Central	Back
High	i		u	i	ɨ	u
Mid	e		o		ɐ	
Low	ɛ	a	ɔ			

Nasal Vowels				Oral and Nasal Diphthongs		
	Front	Central	Back			
High	ĩ		ũ	iw		uj
						ũj̃
Mid	ẽ	ɐ̃	õ	ew	(ej >) ɐj	oj (ow>o)
					(ẽj̃>) ɐ̃j̃ ɐ̃w̃	õj̃
Low				ɛw	aj aw	ɔj

Consonants	Bilabial/ Labiodental	Dental/ Alveolar	Prepalatal/ Palatal	Velar	Uvular
Plosives	p b	t d		k g	
Fricatives	f v	s z	ʃ ʒ		
Nasals	m	n	ɲ		
Laterals		l, ɫ	ʎ		
Vibrants		ɾ			ʀ

Table 2. EP syllables. Frequency data from Viana, Trancoso, Silva, Marques, Andrade & Oliveira (1996). 'N' stands for the nasal feature that yields nasal vowels and diphthongs.

Syllable Structure						
Most Frequent Syllable Types		Possible Onsets			Examples	
		C			pa.to	'duck'
CV	53.00%	Obstruent+Alveolar Liquid			pra.to	'dish'
V	15.50%				glo.bo	'globe'
CVC	12.41%				li.vre	'free'
VC	3.34%				flor	'flower'
CCV	3.01%					
(C)VGC	<1.40%	Possible Rhymes			Examples	
(C)VCC	<0.01%					
(C)VGCC	0.00%	V			ca.ó.ti.co	'chaotic'
		V	N		a.tum	'tuna'
		V	l		al.to	'tall'
		V	r		ar.te	'art'
		V		s	as.tu.to	'astute'
		V	N	s	mons.tro	'monster'
		V	r	s	pers.pi.caz	'acute'
		V	l	s	sols.tí.cio	'solstice'
		VG			mau	'bad'
		VG	N		põe	'(he) puts'
		VG		s	pais	'parents'
		VG	N	s	pões	'(you) put'

Table 3. EP word stress patterns

Stress Patterns					
Final			Penult		
peru	[pɨˈɾu]	‘turkey’	pero	[ˈpeɾu]	‘apple’
dominó	[dɔmiˈnɔ]	‘domino’	domino	[duˈminu]	‘dominate’
alugar	[ɐluˈgaɾ]	‘to rent’	açúcar	[ɐˈsukaɾ]	‘sugar’
tonel	[tuˈnɛɫ]	‘large cask’	túnel	[ˈtunɛɫ]	‘tunnel’
tufão	[tuˈfɐ̃w̃]	‘tornado’	órfão	[ˈɔɾfɐ̃w̃]	‘orphan’
Antepenult					
áspero	[ˈaʃpɨɾu]	‘rough’			
indómito	[ĩˈdɔmitu]	‘untamed’			
víveres	[ˈvivɨɾɨʃ]	‘supplies’			
vírgula	[ˈviɾgulɐ]	‘comma’			
lâmpada	[ˈlɐ̃pɐdɐ]	‘lamp’			

Table 4. EP prosodic word shapes. Examples extracted from the list of EP most frequent words in Bacelar, Marques & Segura da Cruz (1987).

EP word-shapes				
Monosyllables				
	CV	pé	[ˈpɛ]	‘foot’
		dá	[ˈda]	‘(he) gives’
		má	[ˈma]	‘bad (fem)’
	Other	mar	[ˈmaɾ]	‘sea’
		pau	[ˈpaw]	‘stick’
		mal	[ˈmaɫ]	‘badly’
		mão	[ˈmɐ̃w̃]	‘hand’
Disyllables		casa	[ˈka.zɐ]	‘house’
		certo	[ˈsɛɾ.tu]	‘right’
Trisyllables		colega	[ku.ˈlɛ.gɐ]	‘colleague’
		exemplo	[e.ˈzẽ.plu]	‘example’
≥ 4 Syllables		informação	[ĩ.fuɾ.mɐ.ˈsɐ̃w̃]	‘information’
		trabalhadoras	[tɾɐ.bɐ.ʎɐ.ˈdo.ɾɐʃ]	‘workers (fem)’

Table 5. Percentage of word shapes (tokens) produced by Inês and Marta as a function of the number of syllables per word

Word shapes	1:CV	1:other	2	3	3+
CS	15.96	12.08	46.53	20.88	4.58

Table 6. Correlation results for the three sets of data: CS, AS, and CDS. Pearson Product-Moment Correlation. An * signals the results statistically significant at $p < .05$.

Correlation matrix	CS	AS	CDS
CS	_______	,99*	,88
AS	,99*	_______	,91
CDS	,88	,91	_______

Table 7. Number (percent) production of target word shapes for João

Target shapes	0;10.2	0;11.6	1;0.12	1;0.27	1;2.1	1;3.4	1;3.11	1;4.5	1;5.13	1;6.18	1;7.24	1;8.13	1;9.11	1;10.8	1;11.13	2;0.19
Monosyllabic																
CV	---	1/1 (100)	---	---	5/5 (100)	7/7 (100)	---	4/4 (100)	2/2 (100)	---	4/4 (100)	5/5 (100)	17/17 (100)	9/9 (100)	15/15 (100)	24/24 (100)
other	---		0/1 (0)	1/1 (100)	0/1 (0)	---	---	---	1/1 (100)	8/15 (53)	3/4 (75)	6/8 (75)	11/21 (52)	8/10 (80)	31/36 (86)	17/28 (61)
Disyllabic																
‘σ σ	---	2/6 (33)	---	0/5 (0)	0/6 (0)	1/7 (14)	0/2 (0)	1/4 (25)	2/5 (40)	2/7 (29)	0/3 (0)	5/6 (83)	14/30 (47)	---	21/56 (38)	14/32 (44)
σ ‘σ	6/14 (43)	5/14 (36)	10/20 (50)	---	6/7 (86)	10/17 (59)	5/8 (67)	3/7 (43)	19/29 (66)	36/57 (63)	3/16 (19)	0/9 (0)	14/22 (64)	4/8 (50)	22/33 (67)	19/33 (58)
Trisyllabic	---	---	---	---	---	---	---	---	---	---	0/3 (0)	0/2 (0)	0/2 (0)	---	---	0/3 (0)
Total	14	21	21	6	19	31	10	15	37	79	30	30	92	27	140	120

Table 8. Number (percent) production of target word shapes for Inês

Target shapes	0;11.14	1;0.25	1;1.30	1;3.6	1;4.9	1;5.11	1;6.6	1;7.2	1;8.2	1;9.19	1;10.29
Monosyllabic											
CV	1/1 (100)	8/8 (100)	2/2 (100)	11/12 (92)	5/5 (100)	7/7 (100)	5/5 (100)	7/7 (100)	9/9 (100)	12/13 (92)	11/11 (100)
other	1/2 (50)	1/2 (100)	1/3 (33)	2/16 (13)	2/10 (20)	1/12 (8)	3/11 (27)	4/11 (36)	9/18 (50)	19/24 (79)	31/42 (74)
Disyllabic											
‘σ σ	1/2 (50)	0/3 (0)	0/6 (0)	0/18 (0)	1/18 (6)	1/64 (2)	2/58 (4)	4/36 (11)	38/80 (48)	15/92 (16)	115/134 (86)
σ ‘σ	6/10 (60)	7/11 (64)	3/19 (16)	3/10 (30)	5/16 (31)	12/22 (55)	31/52 (60)	29/51 (57)	28/58 (48)	20/34 (59)	40/71 (56)
Trisyllabic	0/1 (0)	0/1 (0)	0/6 (0)	0/3 (0)	0/6 (0)	5/26 (19)	4/28 (14)	15/35 (43)	23/42 (55)	24/38 (63)	44/61 (72)
> 3	---	---	---	---	---	0/2 (0)	0/2 (0)	0/4 (0)	1/3 (33)	1/8 (13)	9/23 (39)
Total	16	25	36	59	55	133	156	144	210	209	342

Table 9. Number (percent) production of target word shapes for Marta

Target shapes	1;2	1;3.8	1;4.8	1;5.17	1;6.23	1;7.18	1;8.18	1;10.4	1;11.10	2;0.26
Monosyllabic										
CV	6/6 (100)	7/8 (88)	6/6 (100)	2/2 (100)	6/6 (100)	6/6 (100)	6/6 (100)	12/12 (100)	12/12 (100)	12/13 (92)
other	19/26 (73)	11/16 (69)	16/21 (76)	19/22 (86)	13/18 (72)	12/14 (86)	14/17 (82)	22/27 (81)	21/26 (81)	23/37 (62)
Disyllabic										
‘σ σ	19/29 (66)	27/58 (47)	30/76 (39)	56/96 (58)	47/72 (65)	68/91 (75)	73/103 (71)	46/77 (60)	77/146 (53)	37/65 (75)
σ ‘σ	17/39 (44)	22/46 (48)	20/30 (67)	9/25 (36)	17/32 (53)	14/22 (64)	8/33 (24)	23/50 (46)	22/34 (65)	27/44 (61)
Trisyllabic	4/21 (19)	2/21 (10)	8/15 (53)	19/44 (43)	14/38 (55)	27/51 (53)	24/47 (51)	29/52 (56)	41/81 (51)	24/49 (49)
> 3	---	---	---	3/10 (30)	2/10 (20)	2/7 (29)	4/18 (22)	13/23 (57)	7/21 (33)	7/19 (37)
Total	121	149	148	199	176	191	224	241	320	227

FIGURE TITLES

Figure 1. Percentage of word shapes (tokens) in child-directed speech (CDS) and adult-directed speech (AS) as a function of the number of syllables per word.

Figure 2. Production of trisyllabic targets in EP child speech as a function of the number of syllables per uttered word shape (1, 2 and 3).

FIGURE 1

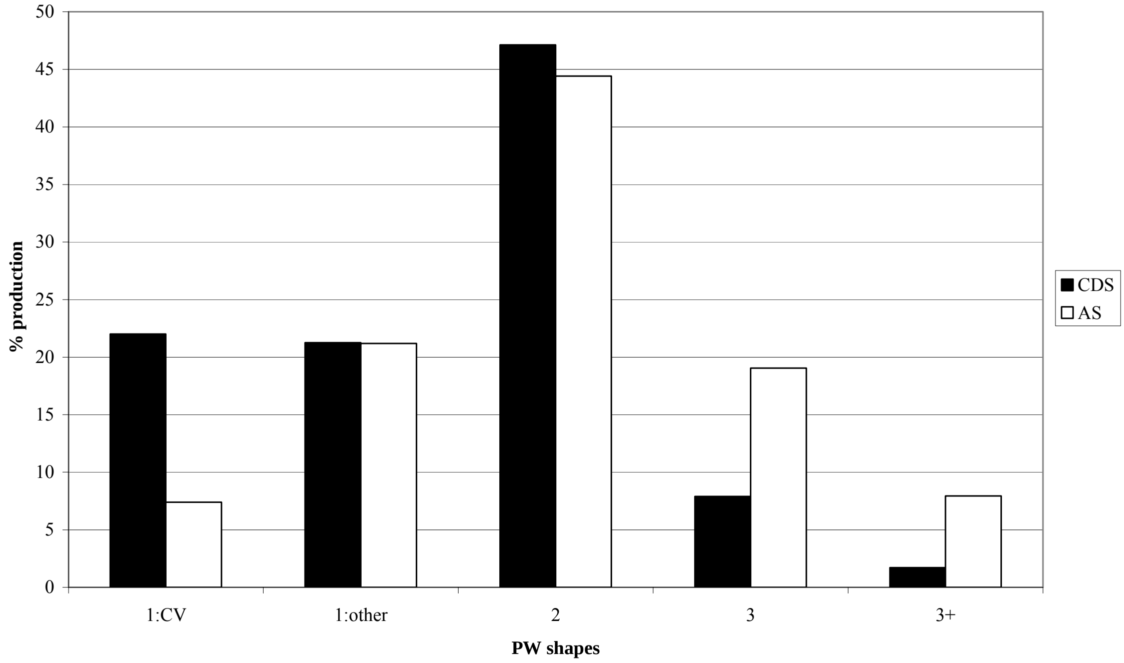
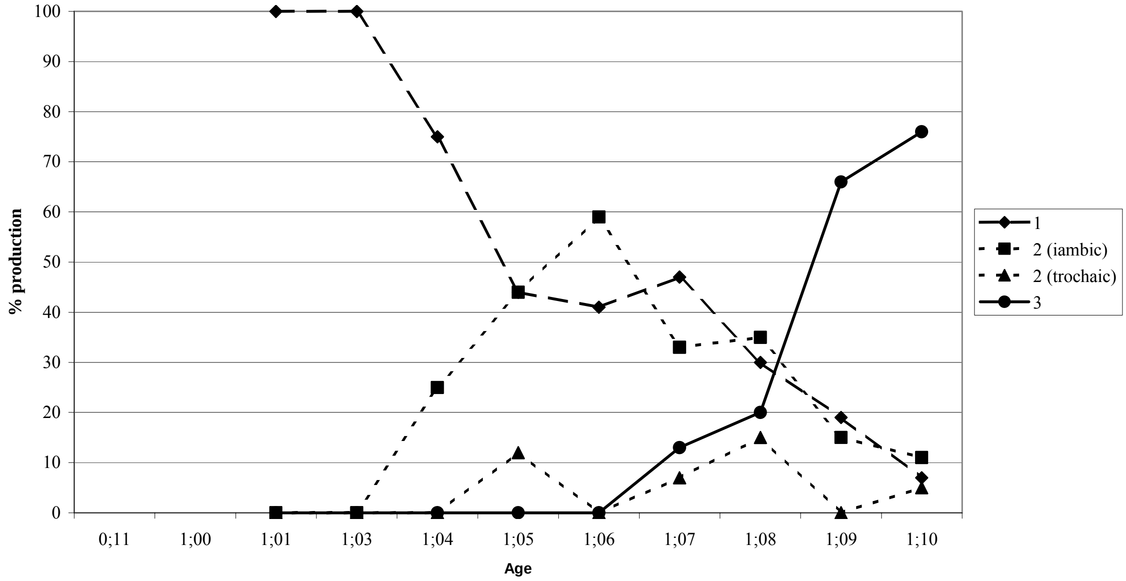

50
45
40
35
30
25
20
15
10
5
0
% production
CDS
AS
1:CV
1:other
2
3
3+
PW shapes

FIGURE 2

Inês



Marta