Prosodic Complexity and Word Segmentation

Catarina Afonso, Dina Alves & Mª João Freitas
University of Lisbon, CLUL www.clul.ul.pt

Introduction

1. Goal
To study potential contrasts in the processing of different word shapes when performing a segmentation task in European Portuguese (EP).

2. Word shapes in European Portuguese
i) Word length frequencies in EP: both in adult speech and in child directed speech disyllabic words are more frequent than trisyllabic ones (Vigário, Freitas & Frota, 2006).

<table>
<thead>
<tr>
<th></th>
<th>CVCV</th>
<th>CVCCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult speech</td>
<td>44%</td>
<td>19%</td>
</tr>
<tr>
<td>Children directed speech</td>
<td>47%</td>
<td>8%</td>
</tr>
</tbody>
</table>

   ii) EP exhibits 3 stress patterns:
   - Word-final stress café ['kafɛ] “coffee”
   - Penultimate stress bebida ['bebide] “drink”
   - Antepenultimate stress chavêna ['xavɛnə] “cup”

   The penultimate stress pattern is clearly the most frequent one in EP (Vigário, 2000). Word length frequencies in EP: both in adult speech and in child directed speech dissyllabic words are more frequent than trisyllabic ones (Vigário, Freitas & Frota, 2006).

3. Hypotheses
Assuming that frequency is a relevant aspect to determine the (un)marked nature of linguistic structures, we will consider that:

   • A dissyllabic word shape matches the unmarked word length in EP;
   • The penultimate stress pattern is the unmarked one in EP (Mateus & Andrade, 2000).

   Following the mainstream in the literature we will assume that unmarked structures are easier to process than marked ones. Therefore, in EP:

   H1: It is easier to segment disyllabic words than trisyllabic ones.

   H2: It is easier to segment CV CV than CVCV ones.

Method

1. Subjects
   • 80 children kindergarten (40 female and 40 male);
   • Age 50 to 77 months old (average = 64 months);
   • No history of hearing or language deficits or disorders;
   • Four subgroups based on the results of the segmentation task:
     - Group 1: 26 children (average age of 62.5 months, s.d. = 7.3);
     - Group 2: 14 children (average age of 63.2 months, s.d. = 6.6);
     - Group 3: 22 children (average age of 64.9 months, s.d. = 6.9);
     - Group 4: 18 children (average age of 68.5, s.d. = 6.8).

2. Stimuli
   • 8 disyllabic words with simple Onset faca ['faʃa] “knife”
   • 15 trisyllabic words:
     - 7 with stress in the antepenultimate syllable médico ['mediku] “doctor”
     - 8 with stress in the penultimate syllable cavalho ['kɐvɐʎu] “horse”

3. Procedure
   a. Children individually tested in a quiet room at school;
   b. Children instructed to look at a picture in the computer screen and to simultaneously listen to the word registered in a sound file (same procedure for each stimulus);
   c. Segmentation by pressing a computer key as quickly as possible;
   d. Stimuli presented randomly and reaction times (RT) registered through the computer keyboard;
   e. Experimental procedure developed in E-Prime1.0 and results analyzed with SPSS 11.5.

Results

1. Segmentation task - word length
   - Reaction time
     - Disyllabic words: $x = 2635$ msec.
     - Trisyllabic words: $x = 4294$ msec.
   - Both the number of correct segmentations and the RT values confirm the unmarked nature of disyllabic words.

   ![](https://www.clul.ul.pt)

2. Segmentation task – word stress
   - Reaction time
     - CV CV: $x = 1904$ msec.
     - CV CCV: $x = 2132$ msec.
   - Although the number of correct segmentations confirmed the unmarked nature of CV CV, the RT scores do not exhibit the same result.

   ![](https://www.clul.ul.pt)

(i) Results for the segmentation task with disyllabic and trisyllabic words support Hypothesis 1: segmenting disyllabic words in EP is easier than segmenting trisyllabic ones. This is empirically supported by the increased number of incorrect segmentations for trisyllabic words and the higher RT values associated to the latter word shape. Literature in EP shows that disyllabic forms in early stages correspond to the most used word shape in production (see Vigário, Freitas & Frota, 2006). Our results for phonological awareness are consistent with the information available for phonological development and for the target system in EP (Mateus & Andrade, 2000; Vigário, Freitas & Frota, 2006; Vigário, Martins & Frota, 2004).

(ii) Hypothesis 2 (segmenting CV CV CV words is easier than segmenting CV CV CV) words was partially confirmed by data: children had a higher number of correct segmentation when stimuli had a CV CV CV shape, thus confirming the unmarked nature of penultimate stress in EP (Mateus & Andrade, 2000; Vigário, Martins & Frota, 2004). However RT scores show the opposite behavior: CV CV CV exhibit higher RT values than CV CV CV.

(iii) Results show that the word length plays a role when performing a word segmentation task (disyllabic words are easier to process than trisyllabic words); the role of word stress is less clear; further research on this matter is required.

(iv) For word length the results match the predictions in the literature; a similar pattern is not attested for word stress:

   ![](https://www.clul.ul.pt)

References