Word stress perception in European Portuguese

Susana Correia, Sónia Frota, Joseph Butler, Marina Vigário

Universidade de Lisboa, Portugal

correia.smd@gmail.com, sonia.frota@mail.telepac.pt, jbutler@fl.ul.pt, marina.vigario@mail.telepac.pt

Abstract

Previous research has reported stress "deafness" for languages with predictable stress, like French, contrary to languages with non-predictable stress, like Spanish ([1], [2], [3], [4]). The contrastive nature of stress appears to inhibit stress "deafness", but segmental and/or suprasegmental cues may also enhance stress discrimination ([5], [6]). In this study we carried out two experiments aiming to investigate stress perception in European Portuguese (EP), a language with non-predictable stress that utilizes duration and vowel reduction as main cues to stress. We used nonsense words that differed only in stress location, thus removing vowel reduction as a cue to stress. Experiment 1 was an ABX discrimination task ([1]). Experiment 2 was a sequence recall task ([2]). In both experiments, the stress contrast condition was compared with a phoneme control condition, in nuclear and post-nuclear position. Results of both experiments strongly suggest a stress "deafness" effect in EP. Despite its variable nature, word stress is hardly perceived by EP native-speakers in the absence of vowel reduction. These findings have implications for claims on prosodic-based cross-linguistic perception of word stress in the absence of vowel quality, and for stress "deafness" as a consequence of a predictable stress grammar.

Index Terms: perception, word stress, nuclear position, post-nuclear position, stress "deafness"

1. Introduction

Word stress is a prosodic dimension that varies across languages. In some languages, stress position may vary within the word and this variation is lexically contrastive (English, Spanish, EP). In other languages, stress position is fixed (French, Hungarian, Polish). Previous studies have shown that speakers of languages with non-predictable or variable stress are more proficient than speakers of languages with predictable or fixed stress at distinguishing nonsense words that vary only in stress position ([1], [2], [3], [4], [5]). The correlates of stress also vary cross-linguistically and languages differ in the way they signal word stress, by using various combinations of suprasegmental cues (duration, F0 and intensity) and sometimes also segmental cues (vowel quality). Spanish, a language with no phonological vowel reduction, uses suprasegmental cues, whereas Catalan and English which have vowel reduction, use a more diverse set of stress cues ([7], [8]). The co-variation between word stress and pitch accent also patterns differently across languages, depending on the domain for pitch accent distribution of the language ([9]). Pitch accent is a strong cue for word stress if there is a strict co-variation between stress and pitch accent.

In EP word stress is variable, as it may fall within the last three syllables of the Prosodic Word, and it is lexically contrastive (e.g. bamboo 'lax' ['bibul]/bamhu 'bamboo' [bi'bu]). Duration is the main cue for word stress ([10], [11]). In EP there is low co-variation between stress and pitch accent, due to a sparse pitch accent distribution: only 17% of Prosodic Words internal to the Intonational Phrase carry a pitch accent ([12]). Vowel reduction is a regular phenomenon in unstressed position (with few exceptions): /i, e, ə, a, u/ are realized as [i, ɨ, ɐ, u] in unstressed position. The vowel [ə] appears in stressed position before nasals and palatals.

EP is an interesting test case to examine word stress perception. On the one hand, EP patterns with Spanish, Catalan and English against French in having variable stress. Stress perception in EP is thus expected to pattern with Spanish, Catalan and English ([11] [2], [6], [8]), in particular stress "deafness" is not expected. On the other hand, EP patterns with English and Catalan against Spanish in the diversity of cues used to signal stress: whereas Spanish, Catalan and English use prosodic cues (namely, duration) to signal stress, EP, just like Catalan, might provide evidence for prosodic-based cross-linguistic perception of word stress, as prosodic cues (namely, duration) should be enough to signal stress contrasts [6]. Finally, EP differs from Spanish, Catalan and English in the low co-variation between stress and accent. EP may thus contribute with new data to the research on stress perception in accented (nuclear) and unaccented (post-nuclear) contexts.

In this study we investigated the perception of word stress by Portuguese-native speakers, in the absence of vowel reduction, both in nuclear and post-nuclear position.

2. Experiment 1

Experiment 1 was an ABX discrimination task, following the same experimental procedure as in Dupoux et al. 1997's experiment 1. This would allow us to compare our findings with previous results for French and Spanish reported in [1]. Unlike [1], we tested both nuclear and post-nuclear positions.

2.1. Method

We tested the perception of disyllabic and trisyllabic nonsense words that varied only in stress location, uttered in two conditions: nuclear position (citation form) and post-nuclear position (within the carrier sentence “A MARIA comeu” [target word] com manteiga” “MARY ate [target word] with butter’, with narrow focus on 'Mary'). Fifteen pairs of nonsense words with penultimate and final stress and eleven trisyllabic words with antepenultimate, penultimate and final stress were constructed. Example nonsense words were ['mipu]/[mi'pu] and [drɪˈmitu]/[dɹɪˈmitu]. A phonemic contrast was also tested as a control condition. The stimuli in the control condition varied consonants or vowels (e.g. ['dɛtu]/[dɛtu], ['siu]/[ˈsiu],) while keeping the same stress pattern. As in the control condition stimuli were uttered in nuclear and post-
nuclear position. Three EP native-speakers, one male and two female, produced the stimuli.

Acoustic analysis of the stimuli showed that duration was a cue to stress, both in nuclear (NP) and post-nuclear (PF) position (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>Stressed syllables</th>
<th>Unstressed syllables</th>
<th>Sig. difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>M=251; SD=63</td>
<td>M=156; SD=56</td>
<td><em>p = .03</em></td>
</tr>
<tr>
<td>PF</td>
<td>M=169; SD=42</td>
<td>M=130; SD=36</td>
<td><em>p = .000</em></td>
</tr>
</tbody>
</table>

Table 1. Experiment 1: Mean duration and standard deviation of stressed and unstressed syllables, in nuclear position and in post-nuclear position

A pitch fall in the stressed syllable (due to the H+L* nuclear accent) was an additional cue in nuclear position, as shown in Figure 1.

Figure 1. Intonational contour in nuclear vs post-nuclear position for the nonsense word [pɐ'musi].

Participants listened to two words contrasting in adjacent stress positions (AB), such as [ˈdɐmitu]/[dɐˈmitu] or [dɐˈmitu]/[dɐˈmitu], or showing a segmental contrast. The third word, X, should be equivalent to either A or B. The A and B words were always uttered by two female speakers and X by the male speaker. The participants’ task was to respond whether X corresponded to A or B. The experiment included 200 trials (140 trials for the stress contrast and 60 trials for the phoneme contrast).

Thirty-two standard EP native-speakers participated in the experiment (16 in the nuclear condition and 16 the post-nuclear condition). The order of presentation of the stress contrast and the segmental contrast was counterbalanced across subjects. Participants’ responses and reaction times were recorded with SuperLab Pro v. 4.5. ANOVAs were run for the two dependent variables: error rate and reaction times.

2.2. Results

Table 2 presents the results for the error rates in the stress vs phoneme contrast, in NP and PF.

<table>
<thead>
<tr>
<th></th>
<th>Stress contrast</th>
<th>Phoneme contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>21%</td>
<td>5%</td>
</tr>
<tr>
<td>PF</td>
<td>36%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 2. Mean error rates (Exp. 1)

The results show that the error rates were significantly higher in the stress contrast condition than in the phoneme condition, both in nuclear (F1(1,14) = 71.07, p < .001, η² = .84; F2(1,98) = 23.8, p < .001, η² = .2) and in post-nuclear position ((F1(1,14) = 108.88, p < .001, η² = .89; F2(1,98) = 52, p < .001, η² = .35 - see Figure 2). In addition, an analysis of position (nuclear/post-nuclear) as a between-subject factor, showed that PF generated significantly more errors overall (F1(1,28) = 8, p < .01, η² = .22; F2(1,196) = 15.54, p < .001, η² = .07). No significant interaction was found between stimuli type (stress vs phoneme contrast) and position (nuclear vs post-nuclear position) (F1(1,30) = 1.77, p = .19, η² = .06; F2(1,196) = 2.12, p = .16, η² = .01).

Table 3 shows the results for the reaction times in the stress vs phoneme contrast, in NP and PF.

<table>
<thead>
<tr>
<th></th>
<th>Stress contrast</th>
<th>Phoneme contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>403ms</td>
<td>321ms</td>
</tr>
<tr>
<td>PF</td>
<td>546ms</td>
<td>545ms</td>
</tr>
</tbody>
</table>

Table 3. Mean reaction times (Exp. 1)

Reaction times were significantly slower in the stress contrast condition than in the control condition, in nuclear position (82 ms, F1(1,14) = 4.24, p = .06, η² = .23; F2(1,98) = 11.73, p < .01, η² = .11). No significant differences in reaction times were found between stress and phoneme contrast in post-nuclear position (F1(1,14) < 1 ; F2(1,98) = 3.27, p = .07, η² = .03). Furthermore, no significant interaction was found between stimuli type and position (F1(1,30) < 1; F2(1,196) < 1).

2.3. Discussion

Native EP subjects performed an ABX discrimination task based on stress contrasts, in nuclear (accented) and post-nuclear (unaccented) positions. Nonsense words contrasting minimally in word stress were used, and nonsense words contrasting only in one phoneme acted as a control condition. Subjects made significantly more errors in the stress contrast than in the phoneme contrast. Moreover, they showed faster reaction times in the phoneme contrast, in nuclear position. The post-nuclear position made discrimination more difficult overall, and not only in the case of the stress contrast.

A comparison with the results in [1], on French and Spanish, shows that the mean error rate in EP in the stress
condition is similar to the stress error rate reported for French (21% in EP vs 19% in French), whereas the error rate in EP in the phoneme control condition is similar to the stress error rate reported for Spanish (5% in EP vs. 4% in Spanish). The results from Experiment 1 strongly suggest that EP subjects attend less to suprasegmental cues to stress (pitch and/or duration) than Spanish or Catalan speakers ([8], [6]), and, in the absence of vowel reduction, show a similar stress “deafness” effect to that reported for languages with predictable stress in their phonological grammar.

3. Experiment 2

In Experiment 2, we used a sequence recall task, following the same procedure used in [2] (specifically, in experiment 4), but with stimuli phonetically similar to the ones used in [4]. Our aim was to test the perception of word stress in EP using a method that has been described to be more robust than the ABX discrimination task. A confirmation of the findings from Experiment 1 would demonstrate that EP speakers show stress “deafness” in the absence of vowel reduction.

3.1. Method

Using a sequence recall task ([2]), we tested EP subjects’ stress perception in the absence of vowel reduction. Two disyllabic (CVCV) minimal pairs consisting of nonsense words were constructed: [ˈmunɐ]/[ˈmunɐ], with a phonemic contrast, and [ˈnumi]/[ˈnumi], with a stress contrast. The 4 nonsense words were uttered in nuclear (citation form) and post-focus position (within the carrier sentence “A MARIA comeu [target word] com manteiga” ‘MARY ate [target word] with butter’, with narrow focus on ‘Mary’). Two EP native speakers, one male and one female, produced 10 tokens of each nonsense word, 6 of which were used as stimuli (three per speaker). The inclusion of multiple tokens uttered by two different talkers added to the phonetic variability of the stimuli.

As in Experiment 1, acoustic analysis of the stimuli showed that duration was a cue to stress, both in nuclear and in post-focus position (Table 4).

<table>
<thead>
<tr>
<th>Stressed syllables</th>
<th>Unstressed syllables</th>
<th>Sig. difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP M=233; SD=28</td>
<td>M=166; SD=29</td>
<td>p=.000</td>
</tr>
<tr>
<td>PF M=151; SD=22</td>
<td>M=129; SD=16</td>
<td>p=.000</td>
</tr>
</tbody>
</table>

Table 4. Experiment 2: Mean duration and standard deviation of stressed and unstressed syllables, in nuclear position and in post-nuclear position

Like in Experiment 1, a pitch fall (H=L*) was a further cue to stress in nuclear position only, as illustrated in Figure 3.

The experiment was divided into two parts. In the first part, participants were tested for phoneme contrast. In the second part, they were tested for stress contrast. Each part was subdivided into two phases: the training phase and the test phase. In the training phase, participants had to associate the two nonsense words to the keys [1] and [2]. By pressing each one of these keys, they listened to the two words as many times as they wanted. Before the test phase, there was a warm-up set of trials, consisting in 4 sequences of two of the newly learned words, presented with feedback. In the test phase, participants listened to twenty sequences composed of 5 tokens each, followed by the word ‘OK’. After listening to the word ‘OK’, participants should recall the order in which the two words had appeared in the 5-token sequence (e.g., [ˈnumi]-[ˈnumi]-[ˈnumi]-[ˈnumi]-[ˈnumi]). Only responses that were a 100% correct transcription of the 5-word sequence were coded as correct; all the others were coded as incorrect. Responses that were 100% incorrect were coded as reversals. As in [2], participants with more reversals than correct responses in either the phonemic or the stress contrast condition were rejected.

24 EP native speakers participated in the experiment (12 in the nuclear condition and 12 in the post-nuclear condition). Additional speakers were tested and excluded from the results due to the presence of too many reversals: 2 in the NP condition (1 had too many for the phoneme contrast and 1 for the stress contrast) and 10 in the PF condition (8 had too many reversals for the phoneme contrast and 2 for the stress contrast). Two other speakers were excluded for responding before the word ‘OK’. Participants’ responses were recorded with SuperLab Pro v. 4.5. The data were subjected to an ANOVA with position (nuclear vs post-nuclear) as a between-subject factor and type of contrast (stress vs phoneme) as a within-subject factor.

3.2. Results

Error rates as a function of type of contrast are shown in Table 5.

<table>
<thead>
<tr>
<th></th>
<th>Stress contrast</th>
<th>Phoneme contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>78%</td>
<td>50%</td>
</tr>
<tr>
<td>PF</td>
<td>92%</td>
<td>48%</td>
</tr>
</tbody>
</table>

Table 5. Mean error rate (Exp. 2)

As in Experiment 1, there was a significant effect of type of contrast (F(1,22) = 66.93, p < .001, η² = .75), with more errors in the stress than in the phoneme contrast, and an almost significant interaction between contrast and position (F(1,22) = 3.76, p = .065, η² = .15), due to the fact that stress, but not phoneme, showed more errors in the PF position (see Figure 4). There was no main effect of position (F(1,22) < 1). An ANOVA carried out for stress and phoneme separately between the two positions showed no difference between NP and PF for the phoneme contrast (F(1,23) < 1), but a significant difference between NP and PF for the stress contrast (F(1,23) = 10.01, p < .01).
In Experiment 2, participants made significantly more errors in the stress contrast condition than in the phoneme contrast condition, both in nuclear and in post-nuclear position. Contrary to the results found in Experiment 1, where the stimuli in post-nuclear position were overall more difficult to discriminate, in Experiment 2, the stress contrast (but not the phoneme contrast) was more difficult to discriminate.

### 3.3. Discussion

In Experiment 2 we investigated stress perception in EP by means of a sequence recall task. Participants made significantly more errors in the stress contrast than in the phoneme contrast, both in nuclear and in post-nuclear position. Furthermore, perception of the stress contrast, unlike that of the phoneme contrast, decreased substantially in post-nuclear position.

These results replicate the findings from Experiment 1. In the absence of vowel reduction, EP speakers show a stress “deafness” effect similar to the one reported for French [2]. Therefore, suprasegmental cues (pitch accent and/or duration) are not used to perceive stress, unlike in Catalan [6]. Also, Experiment 2 showed that in the unaccented context (PF), stress is even harder to perceive, whereas that is not the case for the phoneme contrast. This suggests that the nuclear pitch accent may nevertheless function as a residual weak cue to stress in EP.

### 4. General discussion

In this paper we investigated word stress perception in EP. Given an unusual combination between variable stress, vowel reduction, duration as the main acoustic cue to stress, and low co-variation between stress and pitch accent, EP is an interesting language to examine word stress perception in the absence of vowel quality cues. The variable nature of stress in the language predicts the nonexistence of a stress “deafness” effect ([1], [2], [3], [4]). EP is thus expected to pattern with Spanish and not French. The fact that stress is cued by suprasegmental information, besides vowel quality, as in Catalan, would predict EP data to support recent claims of prosodic-based perception of word stress in the absence of vowel quality cues ([1]).

Using two different paradigms – an ABX discrimination task and a sequence recall task – we demonstrated that EP speakers show a stress “deafness” effect similar to that found in speakers of languages with predictable stress, such as French. In the absence of vowel reduction, and in accented contexts (nuclear position), EP speakers have difficulties in perceiving stress contrasts. In Tables 6 and 7 we compare the error rates found for French and Spanish in [1], [2] and [4] with those found for EP, in a comparable condition (the accented context) and using similar methodologies.

<table>
<thead>
<tr>
<th></th>
<th>French</th>
<th>Spanish</th>
<th>EP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. 1</td>
<td>19%</td>
<td>4%</td>
<td>21%</td>
</tr>
<tr>
<td>Exp. 2</td>
<td>89%, 78%</td>
<td>39%, 48%</td>
<td>78%</td>
</tr>
</tbody>
</table>

Table 6. Error rates for stress contrast (nuclear position only)

<table>
<thead>
<tr>
<th></th>
<th>French</th>
<th>Spanish</th>
<th>EP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. 1</td>
<td>3%</td>
<td>6%</td>
<td>5%</td>
</tr>
<tr>
<td>Exp. 2</td>
<td>50%, 34%</td>
<td>56%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Table 7. Error rate for phoneme contrast (nuclear position only)

Clearly, EP patterns with French, not Spanish. These results bring evidence for a stress “deafness” effect in speakers of a language with non-predictable stress, contra previous findings ([1], [2], [3], [4]).

The EP results, both in accented and unaccented positions, also show that suprasegmental properties alone are not enough for stress perception in a language that uses both suprasegmental and vowel quality cues. Therefore, the present findings do not support the claims of prosodic based cross-linguistic perception of stress in the absence of vowel quality cues.

### 5. Conclusions

Our findings demonstrate that, in the absence of vowel quality cues to word stress, a stress “deafness” effect may occur in a language with non-predictable stress that combines both suprasegmental and segmental information to signal word stress. This raises a number of issues in stress perception and in the acquisition of word stress. Under which conditions may languages with contrastive stress show stress “deafness”? How and when does stress “deafness” arise during the acquisition of such languages? These are questions to be pursued in future research.

### 6. Acknowledgements

The first author is a post-doc research fellow funded by the Fundação para a Ciência e a Tecnologia (grant SFRH/BPD/66529/2009). We acknowledge financial support from the funding program Incentivo/LIN/UI0214/2013 and PEst-OE/LIN/UI0214/2013. This research was developed within the DEPE project (PTDC/CLE-LIN/108722/2008). The third author is a post-doc research fellow of the DEPE project. We would like to thank Cátia Severino for helping in setting up Experiment 1.
7. References


