Revisiting “Stress Deafness” in European Portuguese – An ERP Study

Shuang Lu¹, Susana Correia², Rita Jerónimo³, Marina Vigário², Sônia Frota²

¹ School of Foreign Languages, Renmin University of China, China
² Centre of Linguistics, University of Lisbon, Portugal
³ Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal

Shuanglu2013@outlook.com, correia.smd@gmail.com, rita.jeronimo@iscte.pt, marina.vigario@mail.telepac.pt, sonia.frota@mail.telepac.pt

Abstract

Several behavioral studies have suggested that speakers of languages with variable stress (e.g., Spanish) are better than speakers of languages with fixed stress (e.g., French) at discriminating stress contrasts. European Portuguese (EP) is a language with variable stress, and the main cues for stress are duration and vowel reduction. However, when the vowel quality cue is absent, native speakers are not able to behaviorally discriminate nonsense words that differ only in stress pattern. Using a passive oddball paradigm, the present study recorded event-related potentials (ERPs) to investigate whether native speakers of EP can unintentionally discriminate CVCV pseudo-words with trochaic and iambic stress patterns in the absence of vowel reduction. The results showed that both the trochaic and iambic conditions yielded mismatch negativity (MMN) and late negativity. Moreover, the components in the iambic condition span over a larger temporal window than in the trochaic condition. These results suggest that native speakers of EP can discriminate stress patterns without vowel quality cues at the unintentional level. Furthermore, they are more sensitive to the iambic stress pattern than the trochaic one, which is at odds with their relative frequency in the language, but matches recent developmental findings in the acquisition of stress.

Index Terms: stress discrimination, ERPs, mismatch negativity, late negativity, European Portuguese

1. Introduction

Lexical stress refers to the prominent syllable in a word. Some languages (e.g., Finish, Polish and Turkish) have fixed stress, meaning that stress always falls on a particular position (e.g., the first, the penultimate or the final syllable). Other languages (e.g., English, Spanish, and German) have variable stress, meaning that the position of stress in a word is not predictable. In these languages there may be minimal pairs that only differ in stress pattern (e.g., insight /ɪnsaɪt/ vs. incite /ɪnsaɪt/ in English). Thus, the processing of word stress is particularly relevant in the use of such languages. Previous studies have shown that speakers of languages with variable stress are better than speakers of languages with fixed stress in distinguishing non-words that differ only in stress pattern (e.g., [1], [2], [3], [4] and [5]). Moreover, lexical stress is typically signaled by phonetic cues such as duration, F0, intensity and vowel quality [6]. Languages differ in the weighing of these phonetic cues and the absence of certain cues may influence listeners’ perception of stress. For instance, in English the primary cue for stress is relative pitch prominence (i.e. F0 contour), which outranks intensity, duration and vowel quality ([7] and [8]). However, in Catalan, syllable duration, spectral balance and vowel quality have been found to be the reliable acoustic correlates of stress differences ([9]).

European Portuguese (EP) is a language with variable stress, with penultimate stress being more frequent than final stress ([10]). Vowel reduction has been claimed to be the primary cue for the perception of stress in EP. Behavioral studies have shown that without the vowel quality cues, speakers of EP exhibited a stress “deafness” effect similar to that found for languages with fixed stress ([11]). Duration, which is the main prosodic cue of word stress in the absence of vowel reduction, is not sufficient for the processing of stress contrasts ([12] and [13]). Pitch is a low correlate of stress, due to the sparse distribution of pitch accents in EP ([14]).

To our knowledge, no study has been conducted to examine the unintentional processing of stress by native speakers of EP. Previous research has suggested that perceptual discrimination may occur at the unintentional level, but not (yet) at the intentional/behavioral level ([15]). Using a passive oddball paradigm, the present study recorded participants’ event-related potentials (ERPs) to investigate whether speakers of EP can unintentionally discriminate CVCV pseudo-words that only differ in stress pattern (i.e. trochee vs. iamb) in the absence of vowel reduction. We focused on two ERP components: (1) the mismatch negativity (MMN), which is a negative wave elicited by the deviant stimuli in a sequence of frequently presented stimuli. The MMN peaks at about 100–250ms after change onset (may vary slightly according to different paradigms and type of deviant stimuli) and has a prominent frontal distribution [16]. (2) Late negativity, which is another negative wave that occurs around 350–600ms after the onset of deviant stimuli. This component has been associated with neural processes of auditory rule extraction (e.g., [17]). If native speakers of EP are able to discriminate stress in the absence of vowel reduction, they would show MMN and late negativity to both the trochaic and iambic conditions. Moreover, they may show asymmetric effects, due to the frequency asymmetry of the two stress patterns in EP.

2. Method

2.1. Participants

Twenty-four native speakers of European Portuguese (6 males and 18 females) were recruited in the present study. All participants were between the ages of 18 and 32 years old (M = 21.92, SD = 3.97), and were students at the University of
Lisbon. They were right-handed according to the Edinburgh Handedness inventory [18], and reported having normal vision and hearing. None of them had history of speech or neurological impairment. One additional participant was recruited, but was excluded from data analysis due to technical problems. All participants received either course credit or a voucher for their participation.

2.2. Stimuli

The disyllables [bubu] with either a trochaic or an iambic stress pattern were naturally produced by a female native speaker of EP. Each of the stress patterns was produced twice, resulting in four stimuli in total ([ˈbubu1], [ˈbubu2], [ˈbubu3], and [ˈbubu4]). The stimuli were pseudo-words in EP and were recorded at a sampling rate of 22050 Hz. The mean durations of the trochaic and iambic tokens are 872ms and 873ms respectively. Following [19], the first 100 millisecond of [ˈbubu1], [ˈbubu2], and [ˈbubu3] were replaced by the first 100 millisecond of [ˈbubu4], in order to control the acoustic onset differences. After the manipulation no pitch discontinuity was observed in any of the stimuli (see Figure 1). Three native EP speakers who did not participate in the ERP experiment judged all the stimuli as perceptually natural.

2.3. Procedure

Two types of blocks were created in a passive oddball paradigm: (1) Trochaic block: the iambic tokens were presented as standards, while the trochaic tokens served as deviants; and (2) Iambic block: the frequently occurring trochaic tokens were occasionally replaced by the deviant iambic tokens. Within each block each token of the deviants and standards were presented 50 times and 250 times respectively, resulting in 600 trials in total (50 × 2 tokens plus 250 × 2 tokens). The stimuli were presented in a pseudo-random order, with at least two standards preceding each deviant. We selected 100 clean standards (50 × 2 tokens) that were not immediately preceded or followed by any deviants in each block to compare with the same acoustic stimuli used as deviants in the other block. The offset-to-onset inter-stimulus interval randomly varied between 800 and 850ms to prevent participants’ automatic anticipation of stimulus onset. In order to avoid participant fatigue, each block was split equally into two sub-blocks, with each one lasting for about 8 minutes. The order of the four sub-blocks was counterbalanced across participants. Before the experimental blocks all participants received a practice block, in which each token of the two stress patterns was equally presented for 75 times. The practice block was excluded from data analysis.

During the experiment, participants were watching a muted movie (The Gold Rush by Charlie Chaplin) in a sound-attenuating booth while the stimuli were presented through a loudspeaker at a constant and comfortable hearing level. Participants were asked to ignore the sounds and focus on the movie. They were given comprehension questions regarding the movie after each block. Stimulus presentation was controlled by the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, [20]). The entire experiment took about 2 hours including preparation.

2.4. EEG recording and averaging

Continuous EEG was recorded from 29 Ag/AgCl scalp electrodes according to the international 10-20 system of electrode placement and was sampled at a rate of 500Hz. The electrodes were mounted in an elastic cap (Easy-Cap, Falk Minow, Herrchen-Breitbrunn, Germany) and a SynAmps2 amplifier (Compumedics NeuroScan, Victoria, Australia). The horizontal eye movements were recorded from electrodes at the outer canthus of each eye, and the vertical eye movements from electrodes placed above and below the right eye. Two additional electrodes were affixed at mastoid locations, and the ground electrode was placed on a cervical site. The EEG was referenced online to the left mastoid. During EEG recording electrical impedances were kept below 5 kΩ.

The EEG data were processed offline using NeuroScan 4.3 EDIT software (Compumedics NeuroScan, Victoria, Australia). Data were band-pass filtered from 0.1 to 30 Hz (24dB/oct; zero phase-shift). Eye blink artifacts were corrected using the ocular artifact reduction algorithm implemented in the Edit 4.3 software. The raw EEG data were then segmented into epochs of 1000ms, with a 100ms pre-stimulus baseline and 900ms after the onset of the stimulus. The epoched data were arithmetically re-referenced to the average of both mastoids. Trials exceeding ±80μV in any channel on the entire epoch were rejected. Finally, the ERPs were averaged separately for each stimulus type, electrode and participant. On average, 96 trials for each stimulus type were included in data analysis. The grand-averaged difference waves were generated for each stress pattern by subtracting the average responses to the clean standard stimuli from average responses to the corresponding deviant stimuli.

2.5. Data analysis

The percentage of accurate responses for the comprehension questions on the movie was calculated for each participant. The participants were divided into high accurate and low accurate groups according to their accuracy percentages. We assumed that the participants in the low accurate group may pay more attention to the auditory stimuli and less attention to the movie than the high accurate group, and this attention difference may influence the ERP effects.

Based on visual inspection of the raw ERPs, mean amplitudes within six consecutive time windows of 100 milliseconds were analyzed from 300 to 900 milliseconds after stimulus onset. The mean amplitudes were computed for four regions: left-frontal (LF) included the electrodes F7, F3, FT7 and FC3; right-frontal (RF) included the electrodes F4, F8, FC4 and FT8; left-posterior (LP) included the electrodes TP7, CP3, P7 and P3; and right-posterior (RP) included the electrodes CP4, TP8, P4 and P8.

Figure 1: Spectrograms of the trochaic and iambic stress patterns. Physical differences start at 100 milliseconds.
The mean amplitudes for each stress pattern and latency window were submitted to 2 × 2 × 2 × 2 repeated measures ANOVAs with Discrimination (deviant vs. standard), Hemisphere (left vs. right), and Anteriority (anterior vs. posterior) as within-subject factors and Group (high accurate vs. low accurate) as between-subject factors. All the p-values and the F-values were adjusted using the Greenhouse-Geisser correction and the post-hoc paired t-tests were adjusted using the Bonferroni correction for multiple comparisons.

3. Results

3.1. Comprehension questions on the movie

The high accurate group included participants who correctly answered at least 90% of the comprehension questions, while the low accurate group consisted of participants whose accuracy was below 90%. Table 1 shows the number of participants in each group and their mean accuracy percentages. Independent samples t-test revealed that the two groups significantly differed from each other in the accuracy percentages \( t(22) = 7.28, p < .001 \).

<table>
<thead>
<tr>
<th>Group</th>
<th>Nr. Of participants</th>
<th>Mean accuracy percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>High accurate</td>
<td>11</td>
<td>97% (3.40%)</td>
</tr>
<tr>
<td>Low accurate</td>
<td>13</td>
<td>84% (4.81%)</td>
</tr>
</tbody>
</table>

Table 1. Number of participants and mean accuracy percentages for the high accurate and low accurate groups. Standard deviations are in parentheses.

3.2. ERP data

Grand averages of the frontal electrodes (F3, Fz and F4), the central electrodes (C3, Cz and C4) and the parietal electrodes (P3, Pz and P4) for the whole group are presented in Figure 2a for the trochaic stress pattern and in Figure 2b for the iambic stress pattern. A MMN component was elicited for the deviant versus standard stimuli, with a prominent frontal distribution between 300 to 400 milliseconds for the trochaic stimulus, and between 300 to 500 milliseconds for the iambic stimulus. A late negativity component was also observed at the frontal and central electrodes between 500 to 700 milliseconds for the trochaic stimulus and between 500 to 900 milliseconds for the iambic stimulus. Figure 3 displays the grand-average difference waves (deviant minus standard) for the two stress patterns.

3.2.1. Trochee

The main effect of Discrimination was significant in the time windows of 300-400ms \( F(1, 22) = 17.41, p < .001 \) and 600-700ms \( F(1, 22) = 15.66, p = .001 \). In the time window of 300-400, there was a significant main effect of Anteriority \( F(1, 22) = 17.41, p < .001 \) and a significant interaction of Discrimination × Hemisphere × Anteriority \( F(1, 22) = 5.44, p = .029 \). Post-hoc analyses showed that the Discrimination effect was only significant in the left frontal region \( t(23) = 4.15, p < .001 \), suggesting that this effect can be considered as a MMN with a typical distribution. In the time window of 600-700ms, the main effects of Anteriority \( F(1, 22) = 12.11, p = .002 \) and Hemisphere \( F(1, 22) = 5.57, p = .028 \) were significant. Moreover, significant interactions of Discrimination × Anteriority \( F(1, 22) = 5.32, p = .031 \), Discrimination × Hemisphere × Anteriority \( F(1, 22) = 8.67, p = .007 \) and Discrimination × Hemisphere × Anteriority × Group \( F(1, 22) = 6.70, p = .017 \) were observed. Post-hoc analyses revealed that the Discrimination effect was only significant in the frontal region \( t(23) = 3.15, p = .004 \), but not in the parietal region \( t(23) = 1.36, p = .19 \).

3.2.2. Iamb

A significant main effect of Discrimination was observed from 300 to 900 milliseconds after the stimulus onset (300-400ms: \( F(1, 22) = 29.13, p < .001 \); 400-500ms: \( F(1, 22) = 23.24, p < .001 \); 500-600ms: \( F(1, 22) = 14.50, p = .001 \); 600-700ms: \( F(1, 22) = 44.70, p = .001 \); 700-800ms: \( F(1, 22) = 8.03, p = .01 \); 800-900ms: \( F(1, 22) = 5.84, p = .027 \)). The main effect of Anteriority was also significant in all the six time windows (300-400ms: \( F(1, 22) = 28.53, p < .001 \); 400-500ms: \( F(1, 22) = 36.61, p < .001 \); 500-600ms: \( F(1, 22) = 62.88, p < .001 \); 600-700ms: \( F(1, 22) = 89.51, p < .001 \); 700-800ms: \( F(1, 22) = 110.73, p < .001 \); 800-900ms: \( F(1, 22) = 88.34, p < .001 \)). For the time windows of 400-500ms, a main effect of Hemisphere was found \( F(1, 22) = 6.76, p = .016 \). There was a significant interaction of Discrimination × Anteriority for the time windows of 700-800ms \( F(1, 22) =
13.38, \( p = .001 \) and 800-900ms \([F(1, 22) = 8.95, p = .007]\). In both time windows, post-hoc analyses only yielded significant discrimination effects in the frontal region \([t(23) = 3.80, p = .001]\) and \([t(23) = 3.34, p = .003]\), but not in the parietal region. In the time window of 800-900ms, a significant main effect of Group was obtained \([F(1, 22) = 8.95, p = .007]\).

### a. Trochee

<table>
<thead>
<tr>
<th></th>
<th>300-400</th>
<th>400-500</th>
<th>500-600</th>
<th>600-700</th>
<th>700-800</th>
<th>800-900</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc</td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ante</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc x Hemi</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Disc x Ante</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemi x Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc x Hemi x Ante</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc x Hemi x Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ante x Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### b. Iamb

<table>
<thead>
<tr>
<th></th>
<th>300-400</th>
<th>400-500</th>
<th>500-600</th>
<th>600-700</th>
<th>700-800</th>
<th>800-900</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disc</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Ante</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Hemi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc x Ante</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemi x Ante</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemi x Ante x Group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Main effects and interactions in the six time windows of 100 milliseconds for a) trochaic stress pattern; and b) iambic stress pattern.*** \( p \leq .001 \), **\( p \leq .01 \), *\( p < .05 \).**

#### 3.2.3. Difference wave

In order to directly compare the differences between the trochaic and iambic conditions, we further performed six \( 2 \times 2 \times 2 \times 2 \) repeated measures ANOVAs with Stress (Trochee vs. Iamb), Hemisphere (left vs. right), and Anteriority (anterior vs. posterior) as within-subject factors and Group (high accurate vs. low accurate) as between-subject factors on the difference waves for the six time windows. The results yielded significant main effect of Stress in the time windows of 400-500ms \([F(1, 22) = 10.84, p = .003]\) and 800-900ms \([F(1, 22) = 7.17, p = .014]\), and marginal effect in the time windows of 500-600ms \([F(1, 22) = 3.68, p = .068]\), with the negativity being more prominent in the iambic condition than in the trochaic condition. Moreover, a significant main effect of Group was observed in the time windows of 700-800ms \([F(1, 22) = 5.55, p = .028]\) and 800-900ms \([F(1, 22) = 5.06, p = .035]\), with the low accurate group showing larger negativities than the high accurate group.

### 4. Discussion

In the present study, we recorded native EP speakers’ event-related potentials (ERPs) to examine whether they can unintentionally discriminate CVCC nonsense words with trochaic and iambic stress patterns in the absence of vowel quality cues. The results showed that both the trochaic and iambic conditions yielded mismatch negativity (MMN), and late negativity, indicating that native speakers of EP are able to discriminate the two stress patterns without vowel reduction at the unintentional level. This result is inconsistent with previous behavioral studies ([11]), which demonstrated a stress “deafness” effect in the EP speakers when the vowel reduction cue was removed. This suggests that listeners perceived the difference between the two types of stimuli using some acoustics-based strategies, but these cues are not enough on a meta-linguistic level ([11] and [21]). Hence, EP listeners failed in the behavioral perceptual tasks because these cues are not meaningful or sufficient to match their phonological representations of stress (unlike the vowel reduction cue). Our result is, however, consistent with [19], which used the same paradigm and showed that the native speakers of German (a language with variable stress) can unintentionally discriminate CVCC non-words with trochaic and iambic stress patterns. In [19], only one token of each stress pattern was used, resulting in a fine-grained discrimination situation. In the present study, we included two tokens of each stress pattern and thus provided some evidence that the participants are able to group non-words with different stress types together on the basis of some higher level category representations.

Unexpectedly, the present study showed that the MMN and late negativity components in the iambic condition span over a larger temporal window than in the trochaic condition, indicating that native speakers of EP may be more sensitive to the iambic than the trochaic stress pattern. These results conflict with the frequency distribution of the stress patterns in EP and previous literature on other languages. For example, [22] employed ERP measures and revealed that native speakers of Russian are more sensitive to the trochaic stress pattern, which faithfully represents the frequency asymmetries of the stress patterns in the language. Contrary to [22], it could be suggested that EP listeners appeared to be more sensitive to the less common iambic pattern. However, this explanation cannot account for the similar results we got using other methodologies. We are currently running a follow-up behavioral study on the same participants as in the current study, using an ABX paradigm. Preliminary results replicated the stress deafness effect found in [11]. Nevertheless, results also showed that participants had more accurate and faster discrimination when X is an iambic stimulus ([23]). In addition, a recent study on native EP infants’ perception of stress also showed that 5-6 month old EP-learning infants prefer the iambic to the trochaic stress pattern ([24]). Taken together, these results in adult and infant studies seem to suggest that EP speakers are more sensitive to iambic stress.

### 5. Conclusions

Using the ERP measures, the present study demonstrated that native speakers of EP can unintentionally discriminate CVCC pseudo-words with trochaic and iambic stress patterns in the absence of vowel quality cues. These results argue against stress “deafness” in EP at the unintentional level, and suggest the need of a multi-methodological approach to stress processing.

### 6. Acknowledgements

The authors would like to thank two anonymous reviewers for their insightful comments. We also thank Simão Cortês for his assistance in coordinating with the participants and running the experiment. This project was supported by grant EXCL/MHC-LIN/0688/2012 from Fundação para a Ciência e a Tecnologia, Portugal.
7. References


