A focus intonational morpheme in European Portuguese: Production and perception

Sónia Frota

Abstract

Production studies of the intonational signalling of focus in European Portuguese (EP) have shown that focus is expressed by a specific pitch accent type, thus revealing a systematic contrast between nuclear accents associated with different meanings. In declarative utterances, the contrast between the neutral/broad focus reading and the narrow/contrastive focus reading is essentially realized as an alignment difference: $H+L^*$ (neutral accent) and $H^*+L$ (focus accent). A pilot perceptual study using natural stimuli has shown that subjects are able to distinguish between members of neutral/focus minimal pairs and to match them to the appropriate production context. However, the perception of the contrast found in production has not yet been investigated in detail. The present paper revisits the production contrast and investigates its categorical nature using a multiple methodology approach that resorts to semantically motivated tasks. Several experiments tested whether differences in F0 peak and valley alignment would trigger a perceptual change from one meaning to the other, and whether the alignment differences pattern alike in stimuli based on a neutral and a focus sentence. In Experiment 1, stimuli were classified in a context-matching identification task. In Experiment 2, participants rated appropriateness of stimulus to context in a semantic scaling task. Finally, in Experiment 3, pairs of stimuli were discriminated in a context-matching discrimination task. The results of the three experiments provide converging evidence for the distinction between $H+L^*$ and $H^*+L$. Moreover, they support the claim that the neutral/focus accent distinction is primarily an alignment contrast phonologically encoded at the pitch accent level. These findings have implications for the understanding of the nature of intonational contrasts, and the discussion about the approaches and methods to define prosodic categories.

Keywords: focus, perception, pitch accent type
1. Introduction

Production studies of the intonational signalling of focus in European Portuguese (hereafter EP) have shown that focus is expressed by a specific pitch accent type, that contrasts with the neutral pitch accent, thus revealing a systematic difference between nuclear accents associated with different meanings (Frota 1997, 2000, 2002a, 2002b, in press; Vigário 1998; Fernandes 2007; Viana and Frota 2007). In declarative utterances, the contrast between the neutral/broad focus reading and the narrow/contrastive focus reading is essentially realized as an alignment difference: the neutral accent is characterized by a fall where the Low target aligns with the stressed syllable and is immediately preceded by a peak; the focus accent is characterized by a peak on the stressed syllable of the focalized word, immediately followed by a fall to the post-tonic syllable. Under the Autosegmental-Metrical theory of intonation, the two pitch accent categories have been respectively analysed as H+L* and H*L. Crucially, the two contours differ in the location of the peak and the fall relative to the nuclear syllable, as depicted in (1) (where the nuclear syllable is signalled by the box). The H*L accent is accompanied with focus-related lengthening.

(1) \[
\begin{array}{c}
\text{H+L*} \\
\hline
\end{array} \\
\begin{array}{c}
\text{H*L} \\
\hline
\end{array}
\]

A pilot perceptual study using natural stimuli has shown that subjects are able to distinguish between members of neutral/focus minimal pairs and to match them to the appropriate production context (Frota 2000: 323–350). However, the perception of the contrast found in production has not yet been investigated in detail.

The main goal of this paper is to examine whether the two pitch accents, defined in the literature as different tonal categories on the basis of production, are perceived in a categorical way. Specifically, the following two questions are addressed: Do differences in F0 alignment trigger a perceptual change from one meaning to the other? Do such differences pattern alike in stimuli based on a neutral sentence and a focus sentence?

A related goal of this paper is to contribute to the understanding of the nature of intonational contrasts, and to the discussion around the various approaches and methods to define prosodic categories. These two key issues have been repeatedly addressed and are under debate in much of the recent literature. The nature of intonation has long been established as being discrete and gradient, or in other words, as linguistic and paralinguistic (see Ladd 1996; Gussen-
hoven 1999, 2006). This dual nature of intonation makes it difficult to unravel discrete intonational differences from gradient intonational differences, thus raising the problem of how to define the phonological categories of intonation.

The body of research on alignment of fundamental frequency (F0) targets with respect to the segmental string has shown that timing differences are associated with semantic distinctions in a number of languages (e.g. Bruce 1977; Kohler 1987; D’Imperio and House 1997). However, it is also true that several studies examining alignment contrasts associated with intonational meaning distinctions in different languages have provided mixed results (e.g. Pierrehumbert and Steele 1989; Chen 2003; Dilley submitted; D’Imperio 2000; Niebuhr and Kohler 2004; Gili Fivela 2008). For example, the peak alignment contrast in the English rising accents conveying question/uncertainty versus incredulity has been reported to be discrete in Pierrehumbert and Steele’s (1989) study; but in Chen’s (2003) the same contrast is claimed not to be discrete.

Similar controversial findings have been reported for pitch scaling differences. While such differences have been associated with meaning distinctions in several languages, not all studies conclude in favour of the presence of discrete categories (e.g. Ladd and Morton 1997; Chen 2003; van Heuven and Kirsner 2004; Gili Fivela 2008; Savino and Grice 2011; Schneider and Möbius 2005; Vanrell 2007; Prieto, Torres-Tamarit, and Vanrell 2008; Schneider, Dogil, and Möbius 2009).

A comparison across studies of intonational contrasts reveals that the results obtained are sensitive to the methodologies used (the comparison between Ladd and Morton 1997, Pierrehumbert and Steele 1989, and Chen 2003 is a case in point). The implications of different experimental approaches have been discussed in recent work (Gussenhoven 2006; Rietveld and Chen 2006; Schneider et al. 2006; Schneider, Dogil, and Möbius 2009). From this discussion, as well as from the set of controversial findings offered in the literature, it seems clear that (1) the suitability of the traditional categorical perception (CP) concept to investigate intonational meaning must be questioned, and that (2) as intonation categories are semantically defined, the investigation of such categories calls for semantically motivated tasks or tasks where contextual information is taken into account. The traditional CP paradigm has been questioned not only in the study of intonation (Chen 2003; Niebuhr and Kohler 2004; Niebuhr 2007; Schneider, Dogil, and Möbius 2009), but also in the study of segmental contrasts (Schouten, Gerrits, and van Hessen 2003). One of the arguments against the CP in the study of prosodic contrasts is precisely the absence of a semantically motivated discrimination task, whereas the identification task has an explicit semantic dimension. The discrimination task, however, is simply an acoustic auditory task. Schneider, Dogil, and Möbius (2009) have demonstrated
that the presence of contextual information positively affects discrimination yielding a clear distinction between L% and H% in German, against previous results reported in the literature resorting to the traditional CP.

In the present paper, the EP neutral accent-focus accent distinction was used as a test case to investigate the production and perception of intonational contrasts. In Section 2, some background on the prosody of focus in EP is provided, including prominence and phrasing. In Section 3, the production contrast is revisited and a detailed description of the two pitch accents is provided, both in adult and child speech. Section 4 is devoted to the perception study of the intonational contrast. Multiple methodologies were used, always resorting to semantically motivated tasks: experiment 1 was a context-matching identification task; experiment 2 consisted of a semantic scaling task; and experiment 3 was a context-matching discrimination task. Finally, Section 5 concludes the paper with a discussion of the findings in the light of the argument for a focus intonational morpheme in EP, and of the implications for the study of intonation categories.

2. Background: The prosody of focus in EP

Narrow and contrastive focus respectively refer to the pragmatic uses of focus as identification and correction: in the former, focus identifies the part of the answer that corresponds to the wh-part of a question; in the latter, focus leads to a corrective interpretation relative to a previous proposition (cf. Krifka 2007). Narrow and contrastive focus differ from broad focus in the domain or size of the focus constituent: in the broad focus case the whole sentence expresses new information. In the prosodic literature, broad focus is commonly equated with neutral or unmarked utterances (e.g. Gussenhoven 1984; Hayes and Lahiri 1991; Ladd 1996). In the present paper, the notation neutral (meaning broad focus) versus focus (meaning narrow or contrastive focus) is used for the sake of brevity. This section characterises the prosody of focus utterances in EP relative to neutral utterances, with respect to phrasing and prominence patterns. The description of these patterns informs the role in the neutral/focus distinction played by the pitch accent contrast under study.

Detailed studies of the production of focus utterances compared with their neutral counterparts have shown that the same prosodic phrasing patterns are obtained (Frota 2000, 2002b, 2002c, in press). Utterances produced as replies to wh-questions or utterances with contrastive focus show the same prosodic phrasing as neutral (i.e. broad focus) utterances.

Evidence for phrasing at the phonological phrase level (PhP) comes from stress clash data, where the lengthening of the first clashing syllable within PhP
applies regardless of the presence or absence of focus; similarly, no contrastive lengthening in the first clashing syllable is obtained across a PhP boundary, irrespective of focus. These patterns are illustrated respectively in (2aa′) and (2bb′) below (stressed syllables in capitals, focus in bold, lengthening of the 1st stressed syllable signalled by underlying).

(2) a. \[o\text{caFÉ} Luso\] PhP contém Neutral Stress Clash ‘the Lusitanian coffee contains’ [o\text{caFÉ} lusiTano] PhP contém Neutral No Clash ‘the Lusitanian coffee contains’ a’. \[o\text{caFÉ} Luso\] PhP contém Focus Stress Clash ‘the Lusitanian coffee contains’ [o\text{caFÉ} lusiTano] PhP contém Focus No Clash ‘the Lusitanian coffee contains’

b. \[o\text{caFÉ} \] PhP Luta . . . Neutral Stress Clash ‘the coffee disputes’ [o\text{caFÉ} \] PhP luTOU . . . Neutral No Clash ‘the coffee disputed’ b’. \[o\text{caFÉ} \] PhP Luta . . . Focus Stress Clash ‘the coffee disputes’ [o\text{caFÉ} \] PhP luTOU . . . Focus No Clash ‘the coffee disputed’

Additional evidence for the absence of changes in phrasing due to focus is provided by the constant alignment of the beginning and end of the pitch fall, regardless of the number of pretonic or post-tonic syllables, or the distance from a previous or to a following PhP boundary. For example, the pitch fall shows the same alignment relative to the stressed syllable in the phrases as angoŁAnas ‘the Angolans’, as arQUEÓłogas ‘the archeologists’, or uma maNHĀ angelical ‘an angelic morning’.

Evidence for phrasing at the intonational phrase (IP) level is offered by many sandhi rules that span the IP. In neutral utterances, as in utterances with narrow or contrastive focus, the conditions for these rules to apply are exactly the same. This situation clearly contrasts with the blocking effect produced by the presence of a topic phrase, as shown in (3) by the Fricative Voicing phenomenon (whereby a word-final fricative followed by a word-initial vowel is realized as [z] within the IP versus [J] at the juncture).

(3) a. \[a[z] \text{angolana[z] ofereceram especiaria[z] aos jornalista[J]}\] IP Neutral b. \[a[z] \text{angolana[z] ofereceram especiaria[z] aos jornalista[J]}\] IP Focus c. \[a[z] \text{angolana[z] ofereceram especiaria[J]}\] IP [aos jornalista[J]}\] IP Topic ‘the Angolan offered spices to-the journalists’

‘The Angolan women offered spices to the journalists.’
The identical phrasing of neutral and focus utterances, in contrast with topic utterances, is also shown by tonal facts: the pitch falls to the bottom of the speaker’s range before or after a topic, but not before or after a focus; there is always a pitch rise (or reset) after a topic, but never after a focus. These facts are depicted in (4).

(4)  a. Initial focus (black), initial topic (dotted)  
     b. Final focus (black), final topic (dotted)

In EP, duration is a cue to prominence, both at the word and phrase levels (Delgado Martins 1977; Andrade and Viana 1989; Frota 1991, 2000). Phrasal prominence is final in neutral utterances. In utterances with prosodic focus, the focused word is the head of the IP, as it bears the nuclear accent as well as the lengthening of the stressed syllable that characterizes IP-heads. Importantly, in the late focus cases the lengthening is larger than in the neutral counterpart (Frota 2000: 251–266). This durational difference between the IP-head in the neutral case and the IP-head in the focus case strongly suggests the presence of a specific focus-related lengthening as an extra cue to prosodic focus.

In short, the prosody of focus in EP is characterized by the absence of focus-induced prosodic phrasing effects and by the presence of stress and accent effects. The following section is devoted to a detailed description of the contrast between H+L*, the neutral accent, and H*+L, the focus accent, as it was found in production.

3. The production contrast revisited: H+L* and H*+L

The systematic occurrence of two nuclear pitch falls that differ in their alignment relative to the stressed syllable, as shown in (1) above, is certainly one of the major features of the neutral-focus distinction. The H+L* accent is the nucleus of broad focus (all new) declaratives, and also the nuclear accent of topic phrases; the H*+L accent is the nucleus of narrow/contrastive focus declaratives irrespective of early or late focus position.

The contrast between the two accents is illustrated below, both in single word and multiword utterances (Figures 1 and 2, respectively).
As is apparent from Figures 1–2, peak height does not differentiate the two pitch accents. Unlike the alignment distinction, which is robust within and across speakers, peak height was found to be much more variable and dependent on the speaker (Frota 2000: 237–240; 270–271).

The phonetic detail of the alignment of the H and L targets in the two accents is shown in Figure 3 (adapted from Frota 2002a). The H target in the neutral fall occurs before the onset of the nuclear syllable (only in topic phrases, which are preceded by an IP-boundary, the H target may be delayed) and the L target occurs before the offset of the nuclear vowel. By contrast, in the focus fall the H target aligns after the nuclear syllable onset and the L target aligns well after the offset of the nuclear vowel. There is thus a clear difference in the pattern of alignment relative to the segmental landmarks between the two accents, both for the H and the L targets.

In recent work on early intonational development in EP, it was found that the contrast between the neutral and the focus contour is present very early on in the second year of life, and its use is appropriate according to context (Frota and Vigário 2008). Although the alignment of the leading peak in the neutral accent is not consistently adult-like until 2:00, with a delayed peak relative to
the adult pattern, there is still a systematic difference in the alignment of the two accents. In child speech (like in adult speech), this difference is evident even in cases where the nuclear syllable is the last one in the utterance, as shown in Figure 4 (from Frota and Vigário 2008).

From previous work, it can be concluded that the neutral accent-focus accent distinction is quite robust in production and is acquired early in intonational development. Its major feature is an alignment difference (which is accompanied by focus-related lengthening, in adult speech). This

Figure 2. F0 contours of *As angolanas ofereceram especiarias aos jornaLIStas* ‘the Angolan women offered spices to the journalists’: top panel, neutral contour (in response to ‘What happened?’; bottom panel, focus contour (in response to ‘Whom did they give spices to?’)
Figure 3. The intervals H to nuclear syllable onset (HtoS0) and L to nuclear vowel offset (LtoV1) for the neutral fall (‘N’eutral and ‘T’opic) and the focus fall (‘F’ocus) in the jornalistas utterances (as in Figure 2 above). The circles highlight the neutral and focus cases. Data from the Jornalistas dataset.

Figure 4. The utterance Tatá ‘Tatá (the way the child calls itself)’ produced as a reply to Como te chamas? ‘What is your name?’ and E quem é esta (na foto)? ‘Who is this (in the photo)?’, at 1;09. The second question was uttered by the adult immediately after the child’s reply to the first question.
The perception of the H+L* / H*+L contrast

The main goal of this paper is to examine whether the two pitch accents, defined as different tonal categories on the basis of production, are perceived in a categorical way. Do native listeners distinguish between intended neutral utterances (i.e. with a broad focus interpretation) and intended focus utterances (i.e. with a narrow/contrastive focus interpretation)? Do differences in F0 alignment trigger a perceptual change from one meaning to the other? Do such differences pattern alike in stimuli based on a neutral sentence and a focus sentence? This section is devoted to the perception study of the neutral accent-focus accent distinction. First, we briefly report on a pilot study designed to address the first question listed above. Then, 3 experiments are described which specifically address the two remaining questions.

4.1. Pilot study

A set of sentences uttered to convey a neutral/broad focus reading, a narrow or contrastive focus reading (in which focus is early or late), and also a topic reading (in which the topic is initial or final), were used as natural stimuli in a context-matching task (Frota 2000: 326–331). The utterances chosen were representative of the features that were found to systematically characterize each of the readings. Ten test item pairs, occurring four times each, were presented to ten subjects. Subjects listened to pairs of identical utterances with the exception of their prosody (as those in Figure 2 above), and had to choose the felicitous one in the given context. The answer was forced choice.

The main results obtained were the following. The contrast between topic and neutral (that involves a prosodic phrasing distinction but not a pitch accent distinction – see Section 2 above) and the contrast between focus and neutral are equally well perceived (above 95%). The same holds for the contrast between neutral and early focus (that involves both a difference in phrasal prominence position and pitch accent type – Section 2) and the contrast between neutral and late focus (where the position of the IP head is the same). It could thus be concluded that the distinction between neutral utterances and focus utterances, as produced by speakers, is distinguished by native listeners.
4.2. Experiments and results

In order to test whether the distinction established in the pilot study is implemented in a gradual or categorical fashion, and specifically whether changes in F0 alignment alone trigger a categorical change from one meaning to the other, three experiments were run. As intonation categories are semantically defined, the investigation of such categories (in the present case the H+L* / H*+L distinction) calls for semantically motivated tasks or tasks where contextual information is taken into account (see Section 1). Thus, all three experiments were designed as semantically motivated tasks: Experiment 1 was a context-matching identification task; Experiment 2 was a semantic scaling task; Experiment 3 was a context-matching discrimination task.

4.2.1. General methodology: Stimuli and participants

The three experiments used the same stimuli. Four stimulus series were constructed based on natural productions. The four source utterances contained the two neutral-focus pairs shown in (5): the Jornalistas pair (taken from Frota’s 2000 data) and the Caramelos pair. Therefore, two of the series had a focus accent base (CF, JF) and the other two had a neutral accent (CN, JN). The inspection of the whole F0 contour of the utterances in each pair guaranteed a similar pitch accent distribution in the prenuclear stretch (namely, just a rising accent on the first prosodic word, as in Figure 2 above), and similar peak levels among accents.

(5) a. As angolanas ofereceram especiarias aos jornalistas Neutral (JN)
    As angolanas ofereceram especiarias aos jornalístas Focus (JF)
    ‘The Angolan women offered spices to the journalists.

b. A Maria ofereceu-me carameiros Neutral (CN)
    A Maria ofereceu-me carameiros Focus (CF)
    ‘Mary gave me candies.’

The use of two focus-neutral pairs, instead of just one, with varying lexical items and syntactic structure, each pair produced by a different speaker, ensured that the experimental materials were more varied than what is traditionally found in similar studies, thus diminishing the artificiality of the experimental tasks.

The distance between the peaks in the CN/CF (150 ms) and JN/JF (183 ms) source utterances was divided in equal steps of 30 ms, yielding six steps for
CN/CF, and seven steps for JN/JF. In the two neutral accent based series (CN, JN), peak and valley alignment were shifted forward in steps of 30 ms; in the two focus accent based series (CJ, JF), peak and valley alignment were shifted backwards in steps of 30 ms (as shown, respectively, in (6a) and (6b)). This manipulation was obtained in Praat (Boersma and Weenink 2008).

(6) a. H+L* 
   b. H*+L

The participants in the three experiments were 25 students and staff at colleges in Lisbon, all native speakers of EP (Lisbon variety), with normal hearing and 18 years of age or older. Each task started with a training section with natural stimuli (not used in the experiment itself). Participants that performed at chance level or below in the training were excluded from the analysis.

The experimental tasks were set up using the SuperLabPro software (version 4.5) from Cedrus, and were conducted in one session of about 70 minutes, in the Center of Linguistics of the University of Lisbon. The context-matching identification task preceded the semantic scaling task, which preceded the context-matching discrimination task. There was a short break between tasks.

4.2.2. Experiment 1: Context-matching identification task

The first experiment was a context-matching identification task, a version of the semantic difference tasks that have been used to examine the extent to which listeners assign different meanings to different contours (Gussenhoven 1999, 2006). Specifically, this task addresses the semantic appropriateness of the stimulus to a given context.

Two contexts (a neutral context and a focus context, labelled A or B) were presented on screen for 8000 ms, followed by the audio stimulus. Participants were asked to choose the appropriate context (A or B, forced choice) for the stimulus they heard, and to provide a quick answer. An example of the contexts given in a trial is reproduced in (7). The order of context presentation was

1. The fixed presentation order of the tasks followed similar designs in other studies, to the benefit of cross-study comparisons (e.g. Schneider, Dogil, and Möbius 2009; Prieto, Torres-Tamarit, and Vanrell 2008).
counter-balanced. Trials were randomized across participants, and stimuli were repeated 3 times.

(7) a. A: *Ouvi dizer*:
   ‘I heard that . . .’
   B: *Foi aos artistas que as angolanas ofereceram especiarias?*
   ‘Was it to the artists that the Angolans gave spices?’
   Stimulus: *As angolanas ofereceram especiarias aos jornalistas.*

b. A: *A Maria ofereceu-te chocolates?*
   ‘Did Mary give you chocolates?’
   B: *O que aconteceu?*
   ‘What happened?’
   Stimulus: *A Maria ofereceu-me caramelos.*

The task was preceded by a training section consisting of 8 trials with feedback for correctness and item repetition if incorrect response. 20 participants performed above chance level in the training. A total of 1560 responses were thus obtained for analysis (4 stimuli series = 26 items x 3 repetitions x 20 subjects).

The results for the *Caramelos* series are shown in Figure 5. Both in the focus based (CF) and the neutral based continua (CN), the curves present a shift from one category to another in the range of stimulus 3 (CF) and stimuli 3 and 4 (CN). One-way ANOVAS with stimulus number as the factor were run on the data of each series. There was a significant effect of stimulus number on category identification: for CF, $F(5, 354) = 15.45$, $p < .001$; for CN, $F(5, 354) = 49.49$, $p < .001$. *Post hoc* Tukey HSD tests revealed that the only significant differences between pairs of adjacent stimuli involved stimulus 3 for CF and stimuli 3 and 4 for CN ($p < .05$), as shown in Table 1.

The results for the *Jornalistas* series also show curves with a shift from one category to another. This shift occurred, both for JF and JN, around stimulus 4 (Figure 6). One-way ANOVAS revealed a significant effect of stimulus number: for JF, $F(6, 413) = 20.99$, $p < .001$; for JN, $F(6, 413) = 23.99$, $p < .001$. *Post hoc* Tukey HSD tests confirmed that the only significant differences between pairs of adjacent stimuli involved stimulus 4 (Table 2).

It is clear that the shift did not occur at exactly the same point in the stimuli continua for both members of the *Caramelos* and the *Jornalistas* pairs. It should be recalled that only alignment was manipulated, and not duration, and that syllable duration was known to be longer for the focus accent (the actual difference in the base stimuli was 19 ms for both series). Further, the alignment manipulation took the actual alignment values of the base stimuli as the initial starting point to implement the alignment steps. For these two reasons,
Figure 5. Identification rates (expressed as proportion of Focus responses) as a function of stimulus step number. CF series on the top panel; CN series on the bottom panel.

Table 1. P-value of statistical comparisons of identification rates for adjacent pairs of stimuli, according to Tukey HSD test, in the Caramelos series

<table>
<thead>
<tr>
<th></th>
<th>1–2</th>
<th>2–3</th>
<th>3–4</th>
<th>4–5</th>
<th>5–6</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>.9999</td>
<td>.0324*</td>
<td>.0089*</td>
<td>.5989</td>
<td>.9987</td>
</tr>
<tr>
<td>CN</td>
<td>.9999</td>
<td>.0022*</td>
<td>.0008*</td>
<td>.0260*</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Figure 6. Identification rates (expressed as proportion of Focus responses) as a function of stimulus step number. JF series on the top panel; JN series on the bottom panel.

Table 2. P-value of statistical comparisons of identification rates for adjacent pairs of stimuli, according to Tukey HSD test, in the Jornalistas series

<table>
<thead>
<tr>
<th></th>
<th>1–2</th>
<th>2–3</th>
<th>3–4</th>
<th>4–5</th>
<th>5–6</th>
<th>6–7</th>
</tr>
</thead>
<tbody>
<tr>
<td>JF</td>
<td>.6441</td>
<td>1.000</td>
<td>.0007*</td>
<td>.9996</td>
<td>.0995</td>
<td>.9999</td>
</tr>
<tr>
<td>JN</td>
<td>1.000</td>
<td>.4849</td>
<td>.9806</td>
<td>.0002*</td>
<td>.4849</td>
<td>.9958</td>
</tr>
</tbody>
</table>
it was expected that the alignment details of a given stimulus step number in a focus based series may not be exactly equal to the alignment details of the same stimulus number in a neutral based series. In other words, base stimulus effects were expected. The inspection of the alignment patterns obtained in the two pairs of series shows that the shift is consistently located as a function of alignment (see Table 3): the key factor is the presence/absence of the peak in the stressed vowel and of the low within the post-stressed syllable.\(^2\) This result is coherent with the production data described in Section 3 above (see especially Figure 3).

**Table 3.** Alignment of H and L relative to the segmental string in the *Caramelos* and *Jornalistas* series: H alignment is respectively given relative to the ‘raME’ and ‘naLS’ stretches (1/3, 1/2 and 2/3 refer to location of the peak within the segment); L alignment is signalled as coincident with the post-stressed vowel (PSV), the post-stressed consonant (PSC), the transition between stressed and post-stressed syllables (TRAN), or the stressed vowel (SV). Steps in the focus condition go from focus to neutral; in the neutral condition the reverse applies. Shading marks the stimuli identified as having a focus interpretation.

<table>
<thead>
<tr>
<th>Steps</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>H</td>
<td>m (1/2)</td>
<td>m (1/2)</td>
<td>m (1/2)</td>
<td>m (1/2)</td>
<td>m (1/3)</td>
<td>m</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>PSV</td>
<td>PSV</td>
<td>PSV</td>
<td>PSC</td>
<td>TRAN</td>
<td>SV</td>
</tr>
<tr>
<td>CN</td>
<td>H</td>
<td>m (2/3)</td>
<td>m (2/3)</td>
<td>m (2/3)</td>
<td>m (2/3)</td>
<td>m (1/2)</td>
<td>e (2/3)</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>SV</td>
<td>SV</td>
<td>SV</td>
<td>TRAN</td>
<td>PSC</td>
<td>PSV</td>
</tr>
<tr>
<td>JF</td>
<td>H</td>
<td>i (1/2)</td>
<td>i (1/2)</td>
<td>i (1/2)</td>
<td>i (1/2)</td>
<td>i (1/2)</td>
<td>i</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>PSV</td>
<td>PSV</td>
<td>PSV</td>
<td>PSC</td>
<td>PSC</td>
<td>TRAN</td>
</tr>
<tr>
<td>JN</td>
<td>H</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>SV</td>
<td>SV</td>
<td>TRAN</td>
<td>PSC</td>
<td>PSC</td>
<td>PSV</td>
</tr>
</tbody>
</table>

Base stimulus effects were also found in the distribution of the proportion of focus responses. Although the proportion of focus responses was quite similar across the four series, the analysis of the distribution of participants responses after the identification boundaries revealed significant differences between the

\(^2\) There is a slight difference between the *Caramelos* and the *Jornalistas* series: in the former, peak alignment right at the beginning of the stressed vowel yield an ambiguous stimulus, whereas in the latter this alignment pattern was enough to trigger focus interpretation. Data from previous work had already shown that peak alignment in the neutral accent with nasal consonants tended to be later than with other sonorants (Frota 2000). The reasons behind this segmental effect on alignment are beyond the scope of the present paper.
neutral and the focus based series (Figure 7): in the latter, the proportion of focus responses after the boundary was higher than the proportion of neutral responses after the boundary in the former (p < .05). Again, this is as expected due to the presence of an additional cue to focus, besides alignment, in the focus based stimuli: namely, focus-related lengthening. It should be further noted that baseline effects were larger overall in the Jornalistas series than in the Carame-los series. Participants consistently reported that the Jornalistas sentences were more difficult. This point will be taken up later on in this section.

The analysis of reaction times (RTs) for the within-category stimuli and the across-category stimuli, as established by the identification results, showed that RTs are shorter within-category and longer across-category (although the difference does not reach significance – see Table 4). This tendency in RTs is taken as a confirmation of the identification results obtained (along the lines of Chen 2003).

In summary, the results from Experiment 1 show that listeners perceive the H+L* / H*+L contrast in a discrete manner, as a shift from one category to another was present in the four series. The location of the shift was shown to be consistently related to the alignment of the peak before or within the stressed vowel, and the alignment of the low before or within the post-stressed syllable.

**Figure 7.** Proportion of Focus responses in the four series (F-answers). Proportion of focus answers after the identification shift in the focus series, and proportion of neutral answers after the identification shift in the neutral series (*answers after shift)
Next, a semantic scaling task was carried out in Experiment 2, in order to determine whether a similar shift is obtained in a task that does not enforce a categorical answer.

4.2.3. Experiment 2: Semantic scaling task

The second experiment was a semantic scaling task using a seven-point Equal Appearing Interval scale for the rating of appropriateness of stimulus to context (on the use of unidimensional scaling in perceptual studies of intonational meaning, see Rietveld and Chen 2006). This kind of approach has been used in many studies of intonational contrasts to establish the gradience or discreteness of the implementation of a given meaning (e.g. Wichmann 1991; Caspers 2000; Rietveld et al. 2002; Gussenhoven 2006). Like Experiment 1, Experiment 2 also addressed the semantic appropriateness of the stimulus to a given context. However, a semantic scaling task will not only reveal differences in interpretation but also provide additional and arguably more robust evidence in favour (or against) the discrete nature of the neutral/focus contrast in EP.

In Experiment 2, participants heard the context plus the stimulus, like in natural conversation, and were asked to rate how appropriate the stimulus was to the given context on a scale of 7 points. Score 1 stood for ‘least appropriate’ and score 7 to ‘most appropriate’. The contexts participants listened to in Experiment 2 were the same ones used in Experiment 1. All the stimuli presented were preceded by both a neutral context (e.g. Ouvi dizer ‘I heard that’), and a focus context (e.g. A Maria ofereceu-te bolachas? ‘Did Mary give you cookies?’). Participants were instructed to answer as quickly as possible. Trials were randomized across participants, and stimuli were repeated 3 times.

<table>
<thead>
<tr>
<th>Mean RTs</th>
<th>Within</th>
<th>Across</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>3180</td>
<td>3562</td>
<td>0.546</td>
</tr>
<tr>
<td>CN</td>
<td>2552</td>
<td>3301</td>
<td>0.224</td>
</tr>
<tr>
<td>JF</td>
<td>3366</td>
<td>3556</td>
<td>0.797</td>
</tr>
<tr>
<td>JN</td>
<td>3231</td>
<td>4573</td>
<td>0.239</td>
</tr>
</tbody>
</table>

3. The 7-point scale was chosen considering the number of alignment steps in our data and the fact that 7 is also the most frequent used number, as reported in the literature (Rietveld and Chen 2006).
A training section consisting of 8 trials without feedback preceded the experimental task. 15 participants performed above chance level in the training. A total of 1170 responses were thus obtained for analysis (4 stimuli series = 26 items × 3 repetitions × 15 subjects).

A first analysis of the data looked for a possible effect of type of context on participants responses. A two-way ANOVA with context type (N, F) and stimulus type (C, J) as factors was run, showing a significant main effect of context on the rating with the F(ocus)-context triggering more appropriate responses, \( F(1, 1166) = 36.29, p < .001 \) (mean for F = 5.08, and for N = 4.18), a main effect of stimulus type, \( F(1, 1166) = 5.56, p < .05 \), and a significant interaction between context and stimulus type with the J-stimuli presenting a higher sensitivity to context type than the C-stimuli, \( F(1, 1166) = 12.89, p < .001 \) (means and standard deviations for J: \( F = 5.17 \) (2.26), \( N = 3.79 \) (2.54); means for C: \( F = 4.99 \) (2.42), \( N = 4.64 \) (2.53)). Further analysis for each series was thus run taking context type into account, confirming the F-context effect. Specifically, the ambiguous items, that is the items in the region of the identification shift established by Experiment 1, tend to be rated as appropriate to the F-context. This effect was not found in the N-context, where the rating curves were generally similar to the identification curves from Experiment 1. The stimulus type effect was also confirmed with the J-stimuli providing less clear results than the C-stimuli.

The results for the Caramelos series are shown in Figure 8. The two rating curves for CF and CN in the N-context closely match the identification curves from Exp. 1 (Figure 5), with stimulus 3 for CF and stimuli 3 and 4 for CN as the ambiguous items. The rating curves for the N-context and those for the F-context, however, do not show reflectional symmetry. Crucially, the ambiguous stimuli when preceded by the F-context were rated as appropriate to context and grouped with the non-ambiguous stimuli that precede (for CF) or follow (for CN) the category shift. Moreover, the right-hand side of the CF continuum shows the baseline effects for the focus-based stimuli already reported in Section 4.2.2. To further explore the differences in stimulus rating according to context, Post hoc Tukey HSD tests compared the rating of each stimulus number in the F- and N-contexts. For the CF continuum, they revealed a significant F/N difference for stimuli 1 and 2 (\( p < .05 \)); for the CN continuum, they revealed significant N/F differences for all steps, except for stimuli 3 and 4 (\( p < .01 \)). Overall, these results confirmed the general context effect, and

4. The rating task turned out to be the most difficult task, according to participants. Many of them reported that there were too many options (and too many keys) in the rating.
provided additional support for discreteness in the perception of the H+L* / H*+L contrast (once the effect of context is taken into account).

The results for the Jornalistas series (Figure 9) also show the context effect and the baseline effect for the focus-based stimuli. Again, the two rating curves in the N-context were similar to the identification curves from Exp. 1
(Figure 6), unlike the rating curves in the F-context. In the former, the region between stimuli 4 and 5 was defined as ambiguous between the two readings, closely matching the results from the identification task. In the latter, stimuli in the ambiguous region were rated as appropriate to the F-context. Once more the right-hand side of the focus-based continuum (JF) shows strong baseline effects. Post hoc Tukey HSD tests compared the rating of each stimulus number in the F- and N-contexts and revealed significant differences for all steps in the JF continuum, except for stimulus 6 (p < .05). For the JN continuum,

**Figure 9.** Rating responses as a function of stimulus step number, by context type (F for the focus context – solid line, N for the neutral context – dotted line). JF series on the top panel; JN series on the bottom panel.
significant differences were only found for stimuli 1, 6 and 7 (p < .05), reflecting the absence of discreteness in the rating in the F-context condition. The results for the J-stimuli were thus less clear than those for the C-stimuli. This difference among series is in line with the interaction between context and stimulus type previously reported, that is especially due to the J-stimuli.

The inspection of reaction times (see Table 5) showed that the largest RT measures are for the stimuli judged not appropriate to the context: for appropriate M = 1267, SE = 72.97, for not appropriate M = 1623, SE = 83.67, \( t(7) = -4.31, p < .01, r = .85 \). Stimuli around the identification boundary (as defined in Exp. 1) showed larger RTs than appropriate stimuli especially in the neutral context condition. Still, the difference relative to the appropriate stimuli is globally significant (M = 1525, SE = 98.03; \( t(7) = -3.11, p < .05, r = .76 \). The RT results are thus taken as a confirmation of the rating results obtained.

Table 5. Experiment 2: Mean RTs for the appropriate stimuli, the not appropriate stimuli and the stimuli around the identification boundary (IB) as defined in Exp. 1, by stimulus series and context type

<table>
<thead>
<tr>
<th>Stimuli-Context</th>
<th>Appropriate</th>
<th>Not appropriate</th>
<th>IB (Exp. 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF – F</td>
<td>1405</td>
<td>1882</td>
<td>1391</td>
</tr>
<tr>
<td>CF – N</td>
<td>1527</td>
<td>2023</td>
<td>1805</td>
</tr>
<tr>
<td>CN – F</td>
<td>1055</td>
<td>1557</td>
<td>1298</td>
</tr>
<tr>
<td>CN – N</td>
<td>1086</td>
<td>1355</td>
<td>1630</td>
</tr>
<tr>
<td>JF – F</td>
<td>1014</td>
<td>1573</td>
<td>1116</td>
</tr>
<tr>
<td>JF – N</td>
<td>1498</td>
<td>1244</td>
<td>1863</td>
</tr>
<tr>
<td>JN – F</td>
<td>1365</td>
<td>1619</td>
<td>1327</td>
</tr>
<tr>
<td>JN – N</td>
<td>1184</td>
<td>1627</td>
<td>1767</td>
</tr>
</tbody>
</table>

The relationship between the rating results and the identification results from Experiment 1 was measured by means of the Spearman’s correlation coefficient. To control for the effect of focus context in the rating task, the rating data from the neutral context were used. Given the results previously described, it was predicted that as the neutral responses in the identification task increase, the rating as appropriate to the neutral context will increase. For each of the four series, the rating data were significantly correlated with the identification data: CF, \( r_s = .77, p \) (one-tailed) < .05; CN, \( r_s = .93, p \) (one-tailed) < .01; JF, \( r_s = .88, p \) (one-tailed) < .01; JN, \( r_s = .87, p \) (one-tailed) < .01.

Overall, it may be concluded that the results from Experiment 2 are consistent with those from Experiment 1, if the effect of context type in the rating task is taken into account.
4.2.4. Experiment 3: Context-matching discrimination task

The third and final experiment was a context-matching discrimination task. It innovates from traditional discrimination tasks in that it is semantically motivated, as the identification tasks usually are, and not purely an acoustic auditory task. Schneider, Dogil, and Möbius (2009) have successfully demonstrated that the presence of contextual information affects discrimination, yielding a clear distinction between L% and H% in German, a distinction that did not emerge when out-of-the-blue stimuli were used. In Experiment 3, the same kind of contextual information presentation was used as in Experiments 1 and 2, and thus the context given provides a background for stimuli comparison, or in other words the task is one of semantic discrimination.

A context (either of the neutral type or of the focus type) was presented on screen for 3000 ms, and was then followed by a pair of stimuli (with an Inter-Stimulus Interval of 200 ms). Participants had to decide which of the stimuli in the pair was appropriate to the context, by answering A, B, or I don’t know. They were explicitly instructed that the latter answer was to be used only if they could not decide between A and B. Stimuli within stimulus pairs involving different stimuli (the different pairs) were two steps apart, that is participants heard stimulus 1 paired with stimulus 3, stimulus 2 with 4, stimulus 3 with 5, etc. Half of the different pairs had an early-late peak order in stimuli presentation, and the other half a late-early peak order. Half of the pairs were preceded by a neutral context; the other half by a focus context. For the C-stimuli, there were 8 different pairs (4 from the focus continuum and 4 from the neutral continuum), and for the J-stimuli 10 different pairs (5+5). Six same pairs were included, 3 for C- and 3 for J-stimuli. Two extreme pairs were also added, one for Jornalistas and one for Caramelos. The extreme pairs consisted of stimuli from the opposite ends of the alignment continuum, and allowed a comparison with the results from the pilot experiment (Section 4.2.1). Trials were randomized across participants, and stimuli were repeated 3 times.

A training section consisting of 5 trials with feedback for correctness and item repetition if the response was incorrect preceded the experimental task. 17 participants performed above chance level in the training. A total of 1326 responses were thus obtained for analysis (26 pairs × 3 repetitions × 17 subjects).

A look at the results from the extreme pairs provides a first check at the semantic discrimination task. The mean proportion of different answers for these

---

5. It is known that step size may affect discrimination. The two-step size design was chosen along the lines of similar proposals in the literature (e.g. Niebuhr and Kohler 2004; Dilley submitted).
pairs was 0.94, a result similar to that obtained in the pilot experiment (with natural stimuli). The mean proportion of different answers for each of the different pairs was inspected for a possible effect of context type. Although overall the focus context triggered more different answers than the neutral context (respectively, M = 0.75, SE = 0.026 and M = 0.68, SE = 0.043), no significant effect of context type on discrimination was found, t(17) = 1.86, p > .05, r = .42.

For further and detailed data analysis, discrimination results are presented as d’ for each stimulus pair: the different responses to the different pairs were used as hit rate and the different responses to the same pairs as false alarms (Kaplan et al. 1978). A Wilcoxon Matched Pairs Test was carried out on d’ for the pairs in each of the four stimulus series (CF, CN, JF, JN), by context type. The results show no significant effect of context, confirming the general result on the mean proportion of different answers (Table 6).

A Wilcoxon Matched Pairs Test was also carried out on d’ for the pairs of the four stimulus series, by order of presentation (early-late peak, late-early peak – see Table 7). One of the series showed a significant effect of order of presentation (JF) and another showed a borderline result (CN). Overall, there was a tendency for better discrimination in pairs of stimuli with early-late peak order (d’ means for early-late and late-early were respectively 2.15 and 1.99). This effect seems similar to that reported for low-high pitch order relative high-low

| Table 6. Experiment 3: Wilcoxon Test results by context type, for each stimulus series |
|---------------------------------|-----|-----|
| Stimulus series | T | z  | p  |
| CF | 3 | 0.730 | 0.465 |
| CN | 0 | 1.603 | 0.109 |
| JF | 3.5 | 1.079 | 0.281 |
| JN | 4.5 | 0.809 | 0.418 |

| Table 7. Experiment 3: Wilcoxon Test results by order of presentation, for each stimulus series |
|---------------------------------|-----|-----|
| Stimulus series | T | z  | p  |
| CF | 3 | 0.730 | 0.465 |
| CN | 0 | 1.826 | 0.067 |
| JF | 0 | 2.023 | 0.043* |
| JN | 5 | 0.674 | 0.500 |
order in many studies (e.g. Ladd and Morton 1997; Remijsen and van Heuven 1999; Vanrell 2007; Prieto, Torres-Tamarit, and Vanrell 2008; Savino and Grice 2011). The interpretation of early as low and late as high finds support in House’s (1990) model of the perception of falling tonal contours.

Figure 10 shows the discrimination results (d’) for the Caramulos series. In the CF continuum, a discrimination peak was found at 2–4, that is precisely across stimulus 3, the stimulus identified and rated as ambiguous in the two previous experiments. The discrimination curve further shows lower discrimination for the pairs involving the ambiguous stimulus. Lowest discrimination

![Figure 10](image_url)

**Figure 10.** Sensitivity (d’) to difference between stimulus pairs. CF series on the top panel; CN series on the bottom panel.
was reached at the end of the continuum, between two stimuli within the same category, again matching the results from Experiments 1 and 2. A Wilcoxon Matched Pairs Test carried out on $d'$ for the adjacent pairs showed that 2–4 responses are significantly different from 1–3 and 4–6 from 3–5, whereas no significant difference is found between 3–5 and 2–4 (respectively, $T = 0$, $z = -3.07$, $p < .01$; $T = 0$, $z = -2.72$, $p < .01$; $T = 3$, $z = -1$, $p > .05$).

In the CN continuum, there was a discrimination peak region across the ambiguity region as defined in the two previous experiments (stimuli 3–4). Lowest discrimination was again reached at the end of the continuum. Results of the Wilcoxon Test showed that responses to 1–3 and 2–4, as well as responses to 2–4 and 3–5 were not significantly different, whereas 4–6 differed from 3–5 (respectively, $T = 0$, $z = -1.41$, $p > .05$; $T = 6.50$, $z = -1.29$, $p > .05$; $T = 5.50$, $z = -2.08$, $p < .05$). The lower $d'$ values in the discrimination peak region are consistent with the fact that the pairs tested involved one of the ambiguous stimulus. These results thus confirm previous results on the discreteness in the perception of the H+L* / H**+L contrast in the Caramilos series.

The results for the Jornalistas series show a less clear pattern (Figure 11), in line with previous results for the J-stimulus type (namely, larger baseline effects and more difficulties in identification and in the rating).

In the JF continuum, a discrimination peak was found at 3–5, matching the region of stimuli 4 and 5, which was defined as ambiguous between the two readings in the rating experiment, and also in accord with the identification results. Lowest discrimination is reached at the beginning of the continuum, between two stimuli within the same category. However, at the right-end of the continuum, discrimination raises again. This pattern is consistent with the strong baseline effects previously found for focus-based stimuli in the Jornalistas series, that is some interaction of the duration due to focus with the alignment pattern may underlie this result. The Wilcoxon Test results showed that none of the contrasts between adjacent pairs reaches significance, with the exception of that between 4–6 and 5–7 (for 1–3 and 2–4, $T = 4.67$, $z = -1.71$, $p > .05$; for 2–4 and 3–5, $T = 4.50$, $z = -1.27$, $p > .05$; for 3–5 and 4–6, $T = 7.0$, $z = -1.75$, $p > .05$; for 4–6 and 5–7, $T = 0$, $z = -2.26$, $p < .05$).

In the JN continuum, there was also a discrimination peak region matching the identification boundary defined in Experiment 1 (that is, across 4–5). Discrimination was lower at the beginning and end of the continuum, again matching previous results. The Wilcoxon Test results showed that the contrast between 5–7 and 4–6 is significant, while the differences between the other adjacent pairs did not reach significance (for 1–3 and 2–4, $T = 5.50$, $z = -0.22$, $p > .05$; for 2–4 and 3–5, $T = 5.12$, $z = -0.73$, $p > .05$; for 3–5 and 4–6,
A focus intonational morpheme in European Portuguese

Figure 11. Sensitivity (d') to difference between stimulus pairs. JF series on the top panel; JN series on the bottom panel.

\[ T = 5.25, z = -0.18, p > .05; \text{ for } 4–6 \text{ and } 5–7, T = 1, z = -2.68, p < .01 \]. However, sensitivity is overall lower, and the sensitivity range was smaller, in line with the less clear results for the J-stimuli previously reported. A comparison between the false alarm rates in the same pairs for the J and C-stimuli further confirms the finding that the J-stimuli were more difficult to judge (J = 0.43, C = 0.29).

In short, the results from semantic discrimination largely support previous results by showing higher discrimination across the identification boundaries defined in Experiment 1 and across the stimuli rated as ambiguous in Experiment 2.
4.2.5. Summary of main findings: The three experiments

In Section 3 it was shown that EP speakers produce two pitch accent types that are systematically related to different meanings: the neutral H+L\(^*\) accent, and the focus H\(^*\)+L accent. The three experiments just reported have provided converging evidence that EP speakers perceive the two pitch accent types in a discrete way, and that differences in F0 alignment are enough to trigger a perceptual change between neutral and focus interpretation.

The results from the three experiments also show an effect of the focus accent base that leads to more focus responses at the end of the continuum than the neutral accent base leads to neutral responses. It seems thus clear that the duration cue to focus helps to promote focus responses in focus-based continua. However, the absence of the duration cue has not a symmetrical effect in preventing focus responses in the neutral-based continua. This asymmetry supports the crucial role alignment has as the major cue to the neutral accent-focus accent distinction.

5. Discussion and conclusion

The main goal of this paper was to examine whether the neutral H+L\(^*\) accent and the focus H\(^*\)+L accent, defined as different tonal categories on the basis of production, were perceived in a categorical way. Specifically, the following two questions were addressed: Do differences in F0 alignment trigger a perceptual change from one meaning to the other?; Do such differences pattern alike in stimuli based on a neutral sentence and a focus sentence?

It was found that alignment differences trigger a perceptual change between neutral and focus interpretation. The results from the context-matching identification experiment show a shift from one category to another. The shift was consistently related to the alignment of the peak before versus within the stressed vowel, and the alignment of the low before versus within the post-stressed syllable. The shorter RTs for within-category stimuli and longer for across-category stimuli supported the identification results. Converging evidence for the two categories was obtained in the semantic scaling task (especially when stimuli were preceded by a neutral context). RTs confirmed the rating results, with larger RTs for the utterances judged not appropriate to context and smaller RTs for the utterances appropriate to the semantic context; RTs for stimuli at the boundary shift were also larger than those for stimuli judged as appropriate to context. The results from the context-matching discrimination task were coherent with the identification and the rating data, by showing
higher discrimination across the boundary shift. Such a set of consistent data is not usually reported in studies on the perception intonational contrasts, especially in those using traditional discrimination experiments (e.g. Chen 2003; Schneider and Lintfert 2003; Niebuhr and Kohler 2004; Gili Fivela 2008; Prieto, Torres-Tamarit, and Vanrell 2008; Savino and Grice 2011).

Besides the consistent role played by alignment, it was also found that stimuli based on a neutral sentence and stimuli based on a focus sentence do not pattern exactly alike. The focus accent base led to more focus responses than the neutral accent base led to neutral responses, both in the identification and the rating tasks. Given the descriptions in the literature of the realization of the focus accent, which mention the presence of a duration cue to focus (see Section 2), this was an expected outcome. The duration cue to focus helped to promote focus responses in focus-based continua, and some interaction of the duration cue with the alignment pattern might have affected some of the discrimination results as well. However, the absence of the duration cue did not have a symmetrical effect in preventing focus responses in the neutral-based continua thus leading to more neutral responses. This asymmetry confirms the crucial role of alignment as the major cue to the neutral accent-focus accent distinction, as argued in production studies (see Section 3), although the exact role of duration as a complementary cue to focus perception remains to be examined. Overall, these findings support the claim that the neutral/focus accent distinction is essentially and primarily an alignment contrast phonologically encoded at the pitch accent level by means of the two H+L* and H*+L categories.

The findings reported in this paper have implications for the field of intonational phonology. They strengthen the view that some of the timing variation present in intonation may be phonologized in some languages, resulting in discrete contrasts that are meaning-related (as suggested in Gussenhoven 1999). This, in turn, lends support to the Autosegmental-Metrical approach, where pitch accents may be distinguished on the basis of F0 alignment (Pierrehumbert 1980, 2000; Beckman and Pierrehumbert 1986; Ladd 1996, 2000, among others), as well as to theories of intonational meaning that interpret tonal events as intonational morphemes (e.g. Gussenhoven 1984; Pierrehumbert and Hirschberg 1990; Hayes and Lahiri 1991). For the intonational system of EP, the present findings make a strong argument in favour of previous proposals where the bitonal nuclear tone carries the focus marker (neutral/broad focus H+L* versus narrow/contrastive focus H*+L, see Frota 2002b). As H+L* has been shown to be the default accent in the language, and thus the unmarked member of the distinction (Frota 2000, in press; Fernandes 2007), H*+L stands as a prosodic focus morpheme in the EP intonational system.
A related goal of this paper was to contribute to the understanding of the nature of intonational contrasts, and to the discussion around the various approaches and methods to define prosodic categories.

The categorical nature of the neutral accent-focus accent distinction was investigated using a multiple methodology approach that resorts to semantically motivated tasks. Multiple methodology or the application of various paradigms within the study of the same intonational contrast is an as yet not very common but promising approach (Gili-Fivela 2008; Prieto, Torres-Tamarit, and Vanrell 2008; Schneider, Dogil, and Möbius 2009; Dilley submitted). Only by comparing different methods and their (in)ability to provide convergent results, may the understanding of methodological contingencies be improved. On the other hand, studies of intonational contrasts, like the one in this paper, examine categories that are semantically defined. Therefore, the investigation of such categories calls for semantically motivated tasks or tasks where contextual information is taken into account. Recent studies that have considered the semantic dimension of intonation in their experimental set up have proven more successful than previous studies where such a dimension was not always present (e.g. Chen 2003; Schneider, Dogil, and Möbius 2009). The convergent results obtained in the present study may well be an outcome of the exclusive use of semantic tasks. On the methodological side, the findings from the present study thus argue for the use of semantic tasks to study intonational meaning, including a semantically motivated discrimination task.

Another systematic finding across experiments may hinge on methodological issues of a different nature: the Jornalistas stimuli was more difficult to judge and provided less clear results than the Caramelos stimuli. At least two reasons may underlie this difference. The Jornalistas utterances were much longer than the Caramelos ones: 4 prosodic words, 20 syllables and about 2,500s vs. 3 prosodic words, 13 syllables and about 1,600s. They were also syntactically more complex, having two objects instead of one. It may thus be the case that the processing load was much higher for the Jornalistas series. Moreover, in the Jornalistas utterances, and unlike in the Caramelos ones, there is an unvoicing stretch between the stressed and the poststressed vowels that may have favoured less clear results. However, this difference between Jornalistas and shorter/simpler utterances had not emerged in the pilot study with natural stimuli. Further experiments with utterances with different length, complexity, and syllable structures are required to assess any potential effects of these factors on the perception of intonational contrasts.

Last but not least, it is already known that the neutral accent-focus accent distinction is produced early in the intonational development of EP, but per-
ception studies of such contrast with very young children are non-existent. To fully understand the intonational contrast calls for research on the acquisition and development of the neutral and the focus accents in both production and perception.

Acknowledgments

I would like to thank the audiences of the Workshop on Prosody and Meaning (Barcelona) and of the Linguistics Seminar, Lund University, where different versions of this paper were presented. I’m deeply grateful to Aoju Chen, Barbara Gili Fivela, Carlos Gussenhoven, Pilar Prieto, and Vicent van Heuven for their helpful comments and discussion on the experimental set up of the different tasks. Marisa Cruz and Cátia Severino have helped to run the experiments. I also thank all the participants for their patience and generosity with their time. Finally, thanks are due to the reviewers of this paper and the editors of this volume. This research was partially funded by the projects PTDC/LIN/66202/2006 and PTDC/CLE-LIN/108722/2008, FCT – Fundação para a Ciência e a Tecnologia/MCTES (PIDDAC), Portugal.

References


Dilley, Laura. submitted. The role of F0 alignment in distinguishing categories in American English intonation.


Frota, Sónia. 1991. Para a prosódia da frase: Quantificador, advérbio e marcação prosódica (somente alguns tópicos em foco) [On the prosody of the sentence: Quantifiers, adverbs and prosodic marking (some topics under focus)]. MA dissertation, University of Lisbon.


Prieto, Pilar, Francesc Torres-Tamarit and Maria del Mar Varejell. 2008. The role of tonal scaling in encoding intonational contrasts in Catalan. Talk given at the Third Conference on Tone and Intonation (TIE3), Lisbon.


