Journal of Phonetics 38 (2010) 687–705

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/phonetics

The relevance of prosodic structure in tonal articulation Edge effects at the prosodic word level in Catalan and Spanish

Pilar Prieto [a,*], Eva Estebas-Vilaplana [b], Maria del Mar Vanrell [c]

[a] Institució Catalana de Recerca i Estudis Avançats (ICREA)—Universitat Pompeu Fabra (UPF), Departament de Traducció i Ciències del Llenguatge, Campus de la Comunicació-Poblenou, C. Roc Boronat, 138, Room 53.710, 08018 Barcelona (Catalunya), Spain

[b] Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

[c] Universitat Pompeu Fabra (UPF), and Universitat Autònoma de Barcelona (UAB), Barcelona (Catalunya), Spain

ARTICLE INFO

Article history:
Received 14 July 2008
Received in revised form
7 October 2010
Accepted 7 October 2010

ABSTRACT

We conducted a production experiment with 1600 potentially ambiguous utterances distinguished by word boundary location in Catalan and Spanish (e.g., Cat. *mirà batalles* '(s)he looked at battles' vs. *mirava talles* 'I/(s)he used to look at carvings'; Span. *da balazos* '(s)he fires shots' vs. *daba lazos* 'I/(s)he gave ribbons'; stressed syllables are underlined). Results revealed strong effects of within-word position on H location. Peaks tended to be timed earlier with respect to the end of the syllable when their associated syllables occurred later in the word than when they occurred earlier in the word. These results confirmed previous findings for other languages (Silverman & Pierrehumbert, 1990 for English; Arvaniti, Ladd, & Mennen, 1998 for Greek; and Ishihara, 2006 for Japanese; and Godjevac, 2000 for Serbo-Croatian) and for Spanish and Catalan (Prieto, van Santen, & Hirschberg, 1995 for Spanish; de la Mota, 2005; Simonet, 2006, Simonet & Torreira, 2005 for Catalan). A set of perception experiments suggested that tonal alignment patterns influence listeners' judgments of word boundary location both in Catalan and in Spanish. Listeners were able to employ fine allophonic details of H tonal alignment due to within-word position to identify lexical items that are ambiguous for word-boundary position. The data is consistent with the view that prosodic structure plays an essential part in determining the temporal coordination of f_0 contours with segmental material.

1. Introduction

One of the controversial issues in intonation studies is the degree of stability of the alignment of f_0 (fundamental frequency) turning points with segmental material. While there is robust evidence in favor of the stability of L in rising accents, which are found to be consistently 'anchored' to the consonantal onset of the accented syllable in a variety of languages (Arvaniti et al., 1998 for Greek; Caspers and van Heuven, 1993 for Dutch; Estebas-Vilaplana, 2000 for Catalan; Ladd, Faulkner, Faulkner, & Schepman, 1999 and Ladd, Mennen, & Schepman, 2000 for English; Prieto et al., 1995 for Spanish; Xu, 1998 for Mandarin Chinese), the alignment patterns of H peaks are found to be more variable. First, work on tonal alignment in different languages has shown that peak timing tends to be influenced by syllable structure. For example, D'Imperio (2000) found that the peak was located closer to the vowel offset in closed syllables than in open syllables in Neapolitan Italian (see also D'Imperio, Petrone, & Nguyen, 2007). While in open syllables the peak was aligned with the end of the accented vowel, in closed syllables the peak was located within the coda consonant. The same pattern was found by Gili-Fivela and Savino (2003) for Pisa and Bari Italian, by Welby and Lœvenbruck (2005, 2006)[1] for the late rise in French, and by Prieto and Torreira (2007) for Peninsular Spanish. A similar effect of syllable structure on tonal timing has been reported in falling nuclear accents in Catalan (Prieto, 2009). The results indicated that while the beginning of the falling accent gesture (H) is tightly synchronized with the onset of the accented syllable, the end of the falling gesture (L) is more variable and it is affected by syllable structure: in general, while in open syllables the end of the fall is aligned roughly with the end of the accented syllable, in closed syllables it is aligned well before the coda consonant.[2]

[1] As one of the reviewers points out, Welby and Lœvenbruck's (2005, 2006) results show that one of the six speakers in this study (Speaker 6) aligned her peaks to the end of the vowel, regardless of syllable structure. Even though this study showed that peak alignment can change across speakers, the other five speakers consistently located the peak in the coda consonant for closed syllables.

[2] As the Associate Editor points out, one might think that the generalization is that the end of the fall aligns with the end of the accented vowel. Even though this generalization cannot be maintained if we interpret tonal alignment in strictly phonetic terms – L occurs before the end of the accented vowel in the case of falling accents (see Prieto, 2009) and well into the coda consonant in the case of rising accents (see for example, Prieto & Torreira 2007) – we could also argue that this tonal target is loosely 'associated' with the end of the vowel.

* Corresponding author. Tel.: +34 93 225 4899.
E-mail address: pilar.prieto@upf.edu (P. Prieto).

0095-4470/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2010.10.004

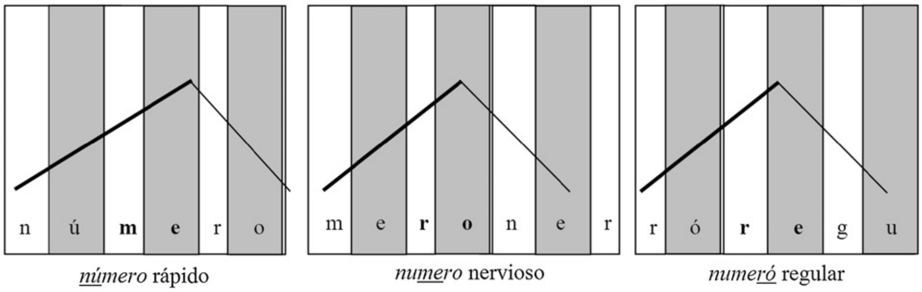

Fig. 1. Schematic representation of the difference in f_0 timing patterns in the three conditions, namely, *número rápido*, *numero nervioso*, and *numeró regular*. Double lines represent word boundaries and post-tonic syllables are marked in bold.

Second, acoustic work in a variety of languages has shown that H peaks are consistently affected by the position of the accented syllable with respect to the end of the word (for Greek, see Arvaniti et al., 1998[3]; for Lebanese Arabic, see Chahal, 2001, 2003; for Serbo-Croatian, see Godjevac, 2000; for Spanish, see de la Mota, 2005; Prieto et al., 1995; Simonet, 2006; Simonet & Torreira, 2005; for Japanese, Ishihara, 2006; for English, see Silverman & Pierrehumbert, 1990). That is, peaks tend to shift backwards as their associated syllables approach the end of the word: in other words, the time from the peak to the beginning of the accented syllable is longer in words with antepenultimate stress than in words with penultimate stress or words with final stress. In order to correct for the potentially confounding effects of stress clash (or time to the next accented syllable), Prieto et al. (1995) analyzed a subset of the data obtained from test syllables in different positions in the word (*número*, *numero*, *numeró*). This subset consisted of word sequences in which there was a difference of two unstressed syllables between one accented syllable and the next (e.g. *número rápido*, *numero nervioso*, *numeró regular*—throughout the paper, accented syllables are underlined). A significant effect of word position on different measures of peak alignment was found in all the comparisons (albeit stronger for one of the three speakers). The three diagrams in Fig. 1 show a schematic representation of the difference in f_0 timing patterns in the three conditions, namely, *número rápido*, *numero nervioso*, and *numeró regular*. Silverman and Pierrehumbert's (1990) model of f_0 peak location showed that two factors, rhyme duration and upcoming prosodic events, were the main source of peak location variation in English. Crucially, the dropping of the variable 'Word-Boundary' (while leaving the variable 'Stress Clash' as a main predictor) significantly worsened the fit of the linear model that was applied to the data.

With respect to duration, previous studies have reported small effects of word boundary location on syllable duration. While Lehiste (1960) found that English syllables in word-initial and word-final positions are longer than word-medial syllables, Turk and White (1999) found that it was not the case that final unstressed syllables were always longer in word-final position: word-final unstressed syllables were only longer than word-initial unstressed syllables when the word was contrastively accented. Similarly, in a study of 11 triads like *tune acquire*, *tuna choir* and *tune a choir*, Turk and Shattuck-Hufnagel (2000) investigated the durational patterns at either side of the word with the presence of a contrastive accent in English. Support was found for word-initial lengthening and for word-final lengthening in phrasally stressed contexts.

The goal of the present study is thus to contribute to our understanding of the temporal organization of the tonal gestures and their patterns of coordination with the segmental material. Specifically, we will investigate how H alignment with respect to the end of the accented syllable is affected by position of the syllable within the word to examine the effects of position-in-word both in Catalan and in Spanish and how it may contribute to the perceptual identification of word boundaries in these two languages. We will also investigate whether we find durational correlates in Catalan and in Spanish. The motivation behind choosing these two languages is twofold. First, even though previous production experiments (see the abovementioned references) have shown consistent effects of within-word position on peak placement for both languages, it will be valuable to compare the two languages in a single study that shares the same methodology and materials. Second, it will be important to compare production and perception behavior in a single study.

First, we carried out a production experiment in which 10 speakers of Central Catalan and Peninsular Spanish read 1600 minimal pair utterances which were only distinguished by word boundary location (e.g., Cat. *mirà batalles* [miˌɾa βəˈtaʎəs] '(s)he watched battles' vs. *mirava talles* [miˌɾaβə ˈtaʎəs] 'I/(s)he used to watch carvings', Span. *da balazos* [ˌda βaˈlaθos] '(s)he shoots' vs. *daba lazos* [ˌdaβa ˈlaθos] 'I/(s)he gave ribbons'). The results showed strong effects of within-word position on H location, whereby peaks tended to be timed earlier with respect to the end of the syllable as their associated syllables approached the end of the word. No systematic lengthening effects of vocalic and syllabic intervals were found when stressed syllables were either in word-final or word-initial positions.

Given these results, our hypothesis was that Catalan and Spanish listeners would be capable of using differences in peak alignment due to word structure for the identification of lexical minimal pairs of the type *mirà batalles* vs. *mirava talles*. Specifically, our hypothesis was that listeners would more easily perceive a word boundary after the accented syllables in those cases where the pitch peak was timed earlier with respect to the end of the syllable, since this is the typical tonal feature of words with final stress. In order to test this prediction, we carried out a controlled perception experiment that consisted of an identification task. A total of forty Catalan and Spanish listeners (20 for each language) were asked to identify stimuli of the type *mirà batalles* vs. *mirava talles* that were manipulated as to H tonal alignment (in five steps).

The paper is organized as follows. Section 2 presents the production experiment and discusses the effects of position-in-word of the accented syllable on peak placement and duration patterns in the two languages. Section 3 presents the identification experiment and discusses the results for the two languages. Finally, Section 4 discusses the relevance of the results of both production and perception experiments for our understanding of tonal coordination with the segmentals.

[3] For Greek, experimental results indicated that while H peak alignment is relatively stable in paroxytones and oxytones, in proparoxytones it exhibits greater variability and between-speaker variation (Arvaniti et al., 1998).

2. Experiment 1

The aim of Experiment 1 was to examine the timing of prenuclear peaks relative to the position of the accented syllable within the word in Catalan and Spanish. In both languages, prenuclear rising accents consist of a rise with the valley aligned with the onset of the stressed syllable and with the peak generally in the postaccentual syllable. As it is well-known, the placement of H peaks is strongly influenced by the upcoming prosodic context, such as the presence of subsequent tones (Estebas-Vilaplana, 2000 and Prieto, 2005 for Central Catalan; Prieto et al., 1995 for Spanish). Recent studies on the modeling of Catalan and Spanish prenuclear accents have found word-boundary effects on peak placement in Castilian Spanish, as spoken in Madrid (Estebas-Vilaplana, 2006; Estebas-Vilaplana & Prieto, 2007), Salamanca (de la Mota, 2005), and Majorca (Simonet & Torreira, 2005; Simonet, 2006), and on Central Catalan as spoken in Barcelona (Estebas-Vilaplana, 2003; Estebas-Vilaplana & Prieto, 2007; Prieto, 2006). With this experiment, our aim was to analyze instances of potentially ambiguous utterances such as *mirava talles* '(s)he used to watch carvings' versus *mirà batalles* '(s)he watched battles'.

2.1. Method

2.1.1. Experimental procedure

The prenuclear pitch accent used in this study is illustrated in Fig. 2. The test H peak corresponds to the first f_0 maximum and belongs to the target word. Prenuclear rises in Catalan and Spanish consist of an f_0 valley which presents a consistent alignment with the onset of the accented syllable, while the alignment of H is much more unstable and has prompted different phonological interpretations.[4] In the present study we will adopt the L+ > H* representation argued for in the recently developed Cat_ToBI (Prieto, Aguilar, Mascaró, Torres Tamarit, & Vanrell, 2009; Prieto, in press) and Sp_ToBI proposals (Estebas-Vilaplana & Prieto, 2008; Face & Prieto, 2007).[5] This phonological analysis enables us to capture the three-way alignment contrasts between L+H*, L*+H and L+ > H* attested in rising pitch accents for both Catalan and Spanish and which are explained in Prieto, D'Imperio and Gili-Fivela (2005). In this study the expected delayed f_0 peak in prenuclear accents justifies the choice for L+ > H* to describe these tonal movements.

The two examples in Fig. 2 illustrate the realization of the prenuclear pitch accent L+ > H* (a rise with a delayed peak) which is associated with the first stressed syllable of the sentence, both in Catalan (upper panel) and in Spanish (bottom panel). The sample utterances are the broad focus utterances *Li venen mandarines* [li ˌβɛnəm mən̪dəˈɾinəs] (Catalan) and *Le venden mandarinas* [le ˌβen̪den man̪daˈɾinas] (Spanish) 'They sell tangerines to him/her'. In both languages the tonal transcription of the utterance consists of one prenuclear accent L+ > H* followed by the nuclear accent !H* and by the boundary tone L%.

2.1.2. Materials

The data used in the two production tests consisted of 20 minimal pairs of potentially ambiguous sentences which had the same segmental and stress composition and were *only* distinguished by word boundary location. Both the Catalan and Spanish sentences included target words with two stress distributions: *oxytones* (words with stress on the final syllable), and *paroxytones* (words with stress on the penultimate syllable). For a complete list of utterances in both languages, see the Appendix. The database included two types of utterances, namely, utterances with either two or three prosodic words, in order to be able to verify whether the same effects would be obtained in longer utterances. Two pairs of utterances from the database are provided in (1) for Catalan and (2) for Spanish.

(1) Examples from Catalan

	Word-final stressed syllables (oxytones)		*Penultimate stressed syllables* (paroxytones)
a.	*comprà ventalls* [kumˌpɾa βən̪ˈtaʎʃ] '(s)he bought fans'	b.	*compraven talls* [kumˌpɾaβən̪ ˈtaʎʃ] 'they bought pieces'
a′.	*comprà ventalls de vim* [kumˌpɾa βən̪ˌtaʎʒ ðə ˈβim] '(s)he bought wicker fans'	b′.	*compraven talls de vim* [kumˌpɾaβən̪ ˌtaʎʒ ðə ˈβim] 'they bought wicker pieces'

(2) Examples from Spanish

	Word-final stressed syllables (oxytones)		*Penultimate stressed syllables* (paroxytones)
a.	*compraré mostazas* [kompɾaˌɾe mosˈtaθas] 'I'll buy mustards'	b.	*compraremos tazas* [kompɾaˌɾemos ˈtaθas] 'we'll buy cups'
a′.	*compraré mostazas alemanas* [kompɾaˌɾe mosˌtaθas alemˈanas] 'I'll buy German mustards'	b′.	*compraremos tazas alemanas* [kompɾaˌɾemosˌ taθas alemˈanas] 'we'll buy German cups'

The phonetic transcriptions show that these paired utterances have the same segmental and accentual composition, and are potentially ambiguous with respect to word boundary location. Potentially confounding effects of stress clash on f_0 peak location were neutralized, as the distance in syllables between stressed syllables was controlled (i.e., there was always either one or two intervening unstressed syllables between the two accents, depending on the target utterance).

2.1.3. Participants

Five speakers of Central Catalan (AG, MB, MR, PG, and PP –PP is the first author of the paper) and five speakers of Castilian Spanish (LA, MP, RA, RC, and SP) read the 40 ambiguous utterances four times (160 utterances per speaker × 5 speakers × 2 languages), for a total of 1600 utterances. The Catalan participants were all speakers of Central Catalan and the Spanish participants were from Northern and Central Castile. They were aged between 20 and 40, and were university teachers or students.

2.1.4. Recording procedure

The Catalan speakers were recorded by the first author and the Spanish speakers by the second author on professional equipment

[4] For Spanish, some studies have classified prenuclear rises with a peak in the posttonic syllable as H* or L+H* pitch accents with a peak delay (Prieto et al., 1995; Nibert, 2000). On the other hand, Sosa (1999), Face (2002) and Beckman, Díaz-Campos, McGory and Morgan (2002) have interpreted the rising contour as a bitonal accent, namely, L*+H, since the f0 maximum was found to be located after the accented syllable. Hualde (2002) finds that neither H* nor L*+H describe prenuclear rises in Spanish satisfactorily and proposes a pitch accent where both tones are associated with the stressed syllable (L+H)* because both target tones might be phonetically aligned with the stressed syllable. For Catalan, prenuclear rises have been described as instances of a low pitch accent (L*) followed by a word edge tone (H) since the f0 peak was consistently aligned with the end of the accented word irrespective of the number of postaccentual syllables (Estebas-Vilaplana, 2000, 2003). Possibly, the consistent alignment of H at the end of words reported in these two studies is actually triggered by the presence of an H- boundary tone which aligns with a major break between the subject and the predicate. In utterances with shorter subject/verb and object domains, such as the ones presented in this paper, there is no systematic alignment of H at the end of words and therefore a word edge tone category cannot be postulated to describe the data.

[5] For the *Cat_ToBI Training Materials*, see http://prosodia.upf.edu/cat_tobi/en/index.php.

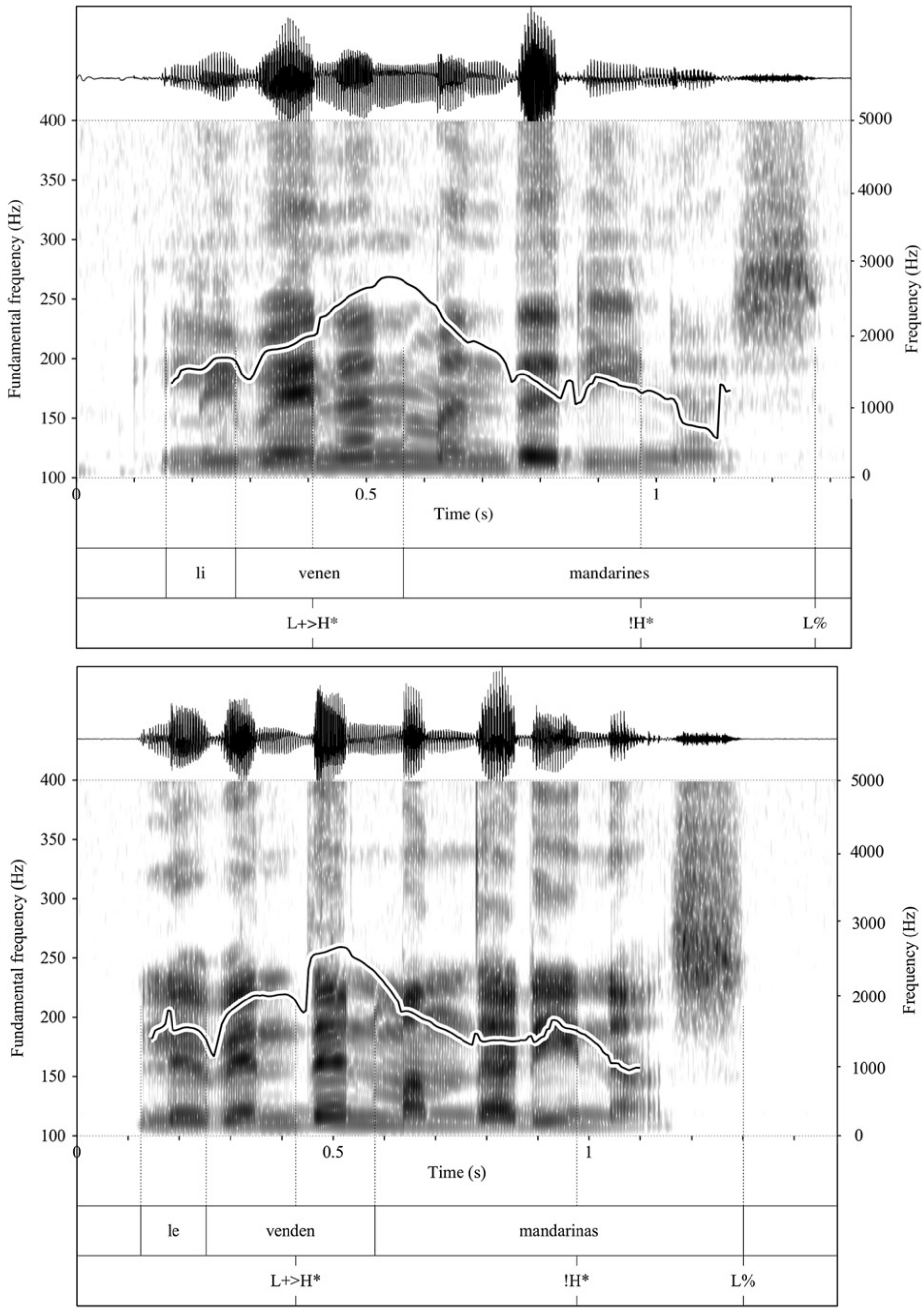

Fig. 2. Waveform and f_0 contour of the broad-focus statement *li venen mandarines* (Catalan, upper panel) and *le venden mandarinas* 'They sell tangerines to him/her' (Spanish, bottom panel). The vertical lines in the textgrid indicate word division.

in a sound-attenuated booth at the Universitat Autònoma de Barcelona (UAB) and the Universidad Nacional de Educación a Distancia (UNED, Madrid), respectively.

Participants were instructed to read the sentences naturally and at a normal rate of speech. They were asked to read the sentences without any pause or prosodic break within the utterance. The recording session was carefully monitored in order to guarantee that the speech was fluent and that no prosodic phrase breaks occurred between words. It was important that all utterances were pronounced in a single intonational phrase in order to avoid the presence of intermediate phrase boundaries (which are typically marked by a continuation rise in both languages). Thus, if speakers produced a prosodic phrase break within a sentence, they were instructed to repeat that sentence at the end. Similarly, there were a few cases of contrastive focus readings, which were repeated. Speakers were also asked to reproduce any sentence showing any type of disfluency or unwanted phrasing or contour. Each recording session lasted about half an hour.

2.1.5. Measurements

Sentences were analyzed by means of *Praat* (Boersma & Weenink, 2005; Wood, 2005). Measurements were made on simultaneous displays of the speech waveform, wide-band spectrogram and f_0 tracks. For each sentence, the following segmental and f_0 landmarks were manually placed in the two test syllables:

(1) Segmental landmarks:
- On the target accented syllable: beginning of onset and vowel.

- On the target postaccentual syllable: beginning of onset, vowel, and coda (whenever present); end of the postaccentual syllable, end of the onset of the following syllable.

f_0 landmarks

(2)
- L1, valley of the first pitch accent (in Hz).
- H1, peak of the first pitch accent (in Hz).

The two graphs in Fig. 3 show the *Praat* windows with the segmentation of the Spanish utterances *compraremos tazas* 'we'll buy cups' (upper panel) vs. *compraré mostazas* 'I'll buy mustards' (lower panel). The upper two boxes displayed in each graph show the speech waveform and a spectrogram with an overlapping f_0 track. The bottom boxes show the segmental landmarks with the segmental boundaries of the two test syllables and the f_0 landmarks.

In some cases, the identification of peaks and valleys was not a trivial task. For example, when the H points formed a plateau where no clear f_0 peak value emerged as the highest point, the standard endpoint in the plateau was selected. Potential microprosodic effects (such as the typical f_0 dip produced by voiced plosives) were visually disregarded while the pitch segmentation was being performed. This was not a problem in the detection of f_0 peaks because the majority of target consonants in this position were sonorant consonants. With regards to the location of segmental boundaries, segmentation procedures described in Peterson and Lehiste (1960) were followed. Importantly, around 90% of the boundaries in the database are instances of sonorant consonant +vowel or vowel+sonorant consonant combinations. The beginning or end of the target sonorant consonants [m, n, l, r] was

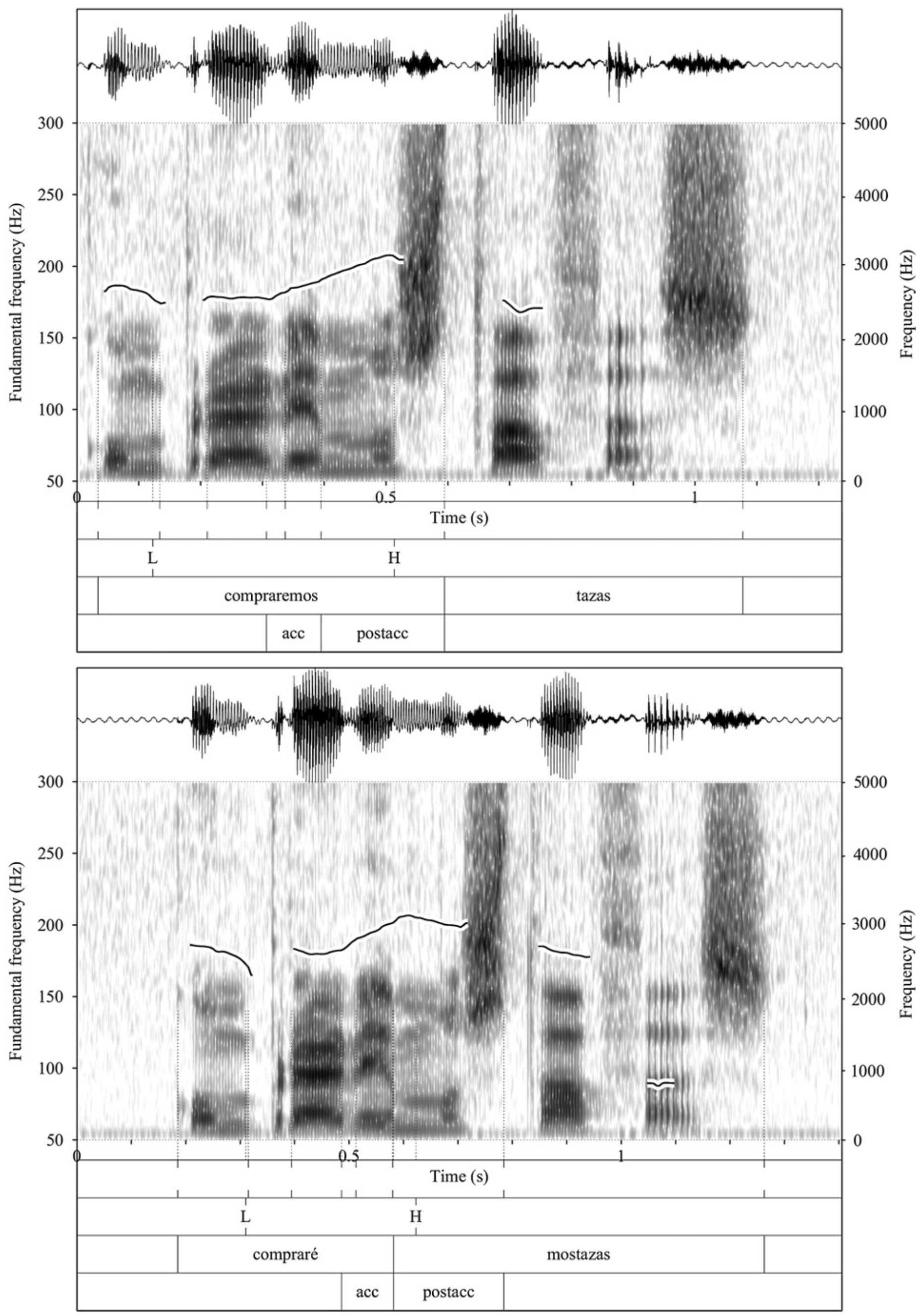

Fig. 3. Displays of speech waveform, spectrogram and f_0 track, and the segmental and pitch landmarks for the Spanish utterances *compraremos tazas* 'we'll buy cups' (upper panel) and *compraré mostazas* 'I'll buy mustards' (lower panel). The vertical lines indicate word division (upper textgrid) and the limits of the accented and post-accented syllables (lower textgrid).

identified as the start of the abrupt change from the steady-state period in the spectrogram to the onglide transition movement to the vowel (e.g. the sequence *mo* in both panels in Fig. 3). Generally, intensity changes (from nasal to vowel or from lateral to vowel, for example) could be used very reliably.

After segmentation, a *Praat* script automatically collected the duration and f_0 marks. We then transferred those into an SPSS file (SPSS for Windows, 2006) and the duration measures relevant for our study were calculated, namely, the alignment of the H tonal target relative to different segmental landmarks (syllable offset, word boundary, etc.), as well as duration measures such as the duration of the segments in the accented and postaccented syllables.

2.1.6. Statistical analyses

Two different measures of H location were used as dependent variables for statistical exploration, namely (i) peak delay from Accented Syllable Onset, or time from the peak to the beginning of the accented syllables, and (ii) peak delay from Accented Syllable End, or time from the peak to the end of the accented syllable. Moreover, two measures of duration were included in the analysis as dependent variables, namely (i) duration of the accented syllable, and (ii) duration of the postaccented syllable.

For each one of the four dependent variables, we ran a factorial repeated measures ANOVA (one for each language) including Within-Word Position as the main predictor, Repetition as an intra-subject factor and Subject as an inter-subject factor. To examine potential subject differences, post-hoc comparison tests with the Bonferroni correction were performed on the Subject factor.

An attempt was also made to capture the combined effects of the two factors, Within-Word Position (Paroxytonic, Oxytonic) and duration of accented syllable (DurAcSyl) on peak placement by means of linear regression, using the delay from the peak to the end of the syllable (Peak delay from Accented Syllable End) as the response measure.

2.2. Results

The results of the production test will be presented in three parts: (1) H delay (or time from the onset of the accented syllable to the f_0 peak) as a function of the duration of the accented syllables, in order to test the reported relationship between syllable duration and H delay (see refs in Section 2.2.1); (2) the effects of within-word position on H location, which will be characterized quantitatively as the delay from the beginning and/or the end of the accented syllable to the peak; and (3) the effects of within-word position on the syllable duration of the accented and postaccentual syllable.

2.2.1. H peak delay as a function of the duration of the accented syllable

A well-known effect found in a variety of languages is the positive correlation between vowel duration and peak delay, that is, when a vowel is lengthened the peak is correspondingly delayed with respect to the end of the accented syllable (Prieto et al., 1995 for Spanish; Silverman & Pierrehumbert, 1990 for English). The duration of the rise and the duration of the syllable can be correlated without being equal. Remember that peak delay has been measured as the time from the onset of the accented syllable to the f_0 peak, as in Silverman and Pierrehumbert (1990). Typically, the peaks under study are 'delayed' and are located in the postaccentual syllable. The two graphs in Fig. 4 plot H peak delay (in ms) as a function of the duration of the accented syllable for the 5 Catalan speakers (left graph) and for the 5 Spanish speakers (right graph). All data are plotted in the two graphs, which represent a total of 1600 data points. The graphs separate the data according to their position within the word: that is, accented syllables in word-final position, as in *compr<u>à</u> ventalls*, and accented syllables in penultimate position, as in *compr<u>a</u>ven talls*. First, the regression lines summarize the strong positive correlations found between the two variables (H delay and syllable duration) both in Catalan

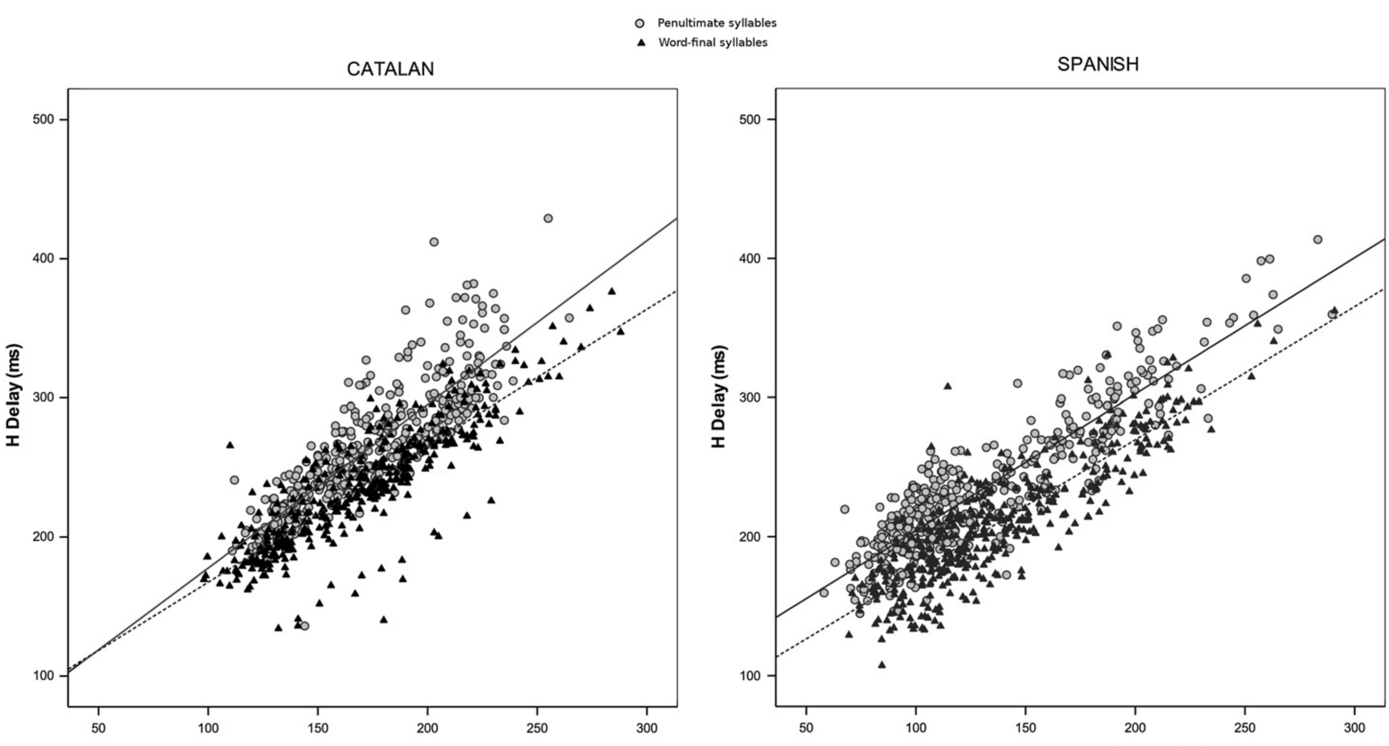

Fig. 4. Peak delay (or time from the onset of the accented syllable to the f_0 peak, in ms) as a function of the duration of the accented syllable (in ms) in the two within-word position conditions (word-final vs. penultimate position) in Catalan (left graph) and Spanish (right graph).

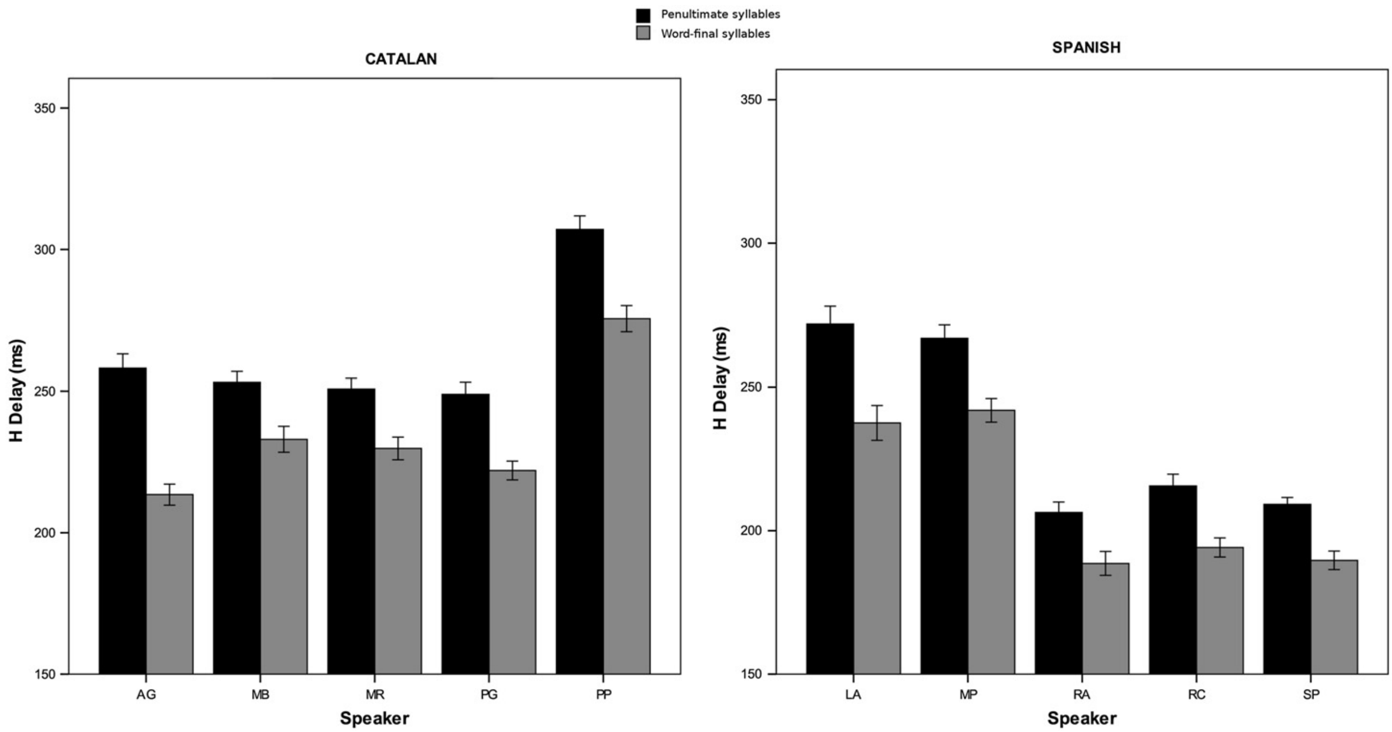

Fig. 5. Mean H peak delay (or time from the peak to the beginning of the accented syllables, in ms.), as a function of within-word position for all Catalan (left panel) and Spanish speakers (right panel). Light bars show mean H peak delay in word-final syllables and black bars show mean H peak delay in penultimate syllables. The vertical bars represent standard errors.

(correlation coefficients are $r\,(3)=.733$ for word-medial positions and $r\,(3)=.724$ for word-final positions, significant at $p<.0001$) and in Spanish (correlation coefficients are $r\,(3)=.808$ for word-medial positions and $r\,(3)=.755$ for word-final positions, significant at $p<.0001$). Second, the graph reveals a consistent difference in H delay depending on the two prosodic conditions, as the data are visually separated into two partially overlapping clouds: the H peak is located earlier in the syllable in word-final position (e.g., *comprà ventalls*; see black triangles in the graph) than in word-medial position (e.g., *compraven talls*; see grey circles in the graph). In addition, the graphs also suggest that the duration of the accented syllable (*x*-axis) is not significantly different across the two conditions, as values for both groups are evenly distributed along the *x*-axis. As we will see in Section 2.2.3, we did not find support for an effect of word-final lengthening in these data.

As shown in Fig. 4, these results reveal that in both languages there is a consistent tendency to have a smaller peak delay in words with final stress than in words with stress on the penultimate syllable (see the distribution of grey circles and black triangles in Fig. 4). This suggests that the location of the H is highly influenced by within-word position. To test whether within-word effect is statistically significant, specific analyses will be performed on a set of dependent variables in the next subsection.

2.2.2. *Effects of within-word position on H location*

In this section, peak location was characterized quantitatively in two different ways, namely, relative to both the beginning (peak delay) and the end of the target syllable. Following recent methodological arguments by Atterer and Ladd (2004) and Schepman, Ladd, and Lickley (2006), we included the latter measure of peak alignment, which takes as its main reference a nearby acoustic landmark (i.e., the end of the syllable) rather than a more distant one (i.e., the start of the syllable).[6] The results from the preceding subsection, which demonstrate a close relationship between the duration of the accented syllable and the peak delay, represent further evidence to select the end of the syllable as a reference point for H alignment for this specific pitch accent.

The two histograms in Fig. 5 plot mean peak delay measures (or mean time of the H tonal target relative to syllable onset) as a function of within-word position (word-final syllables, as in *comprà ventalls*, vs. penultimate syllables, as in *compraven talls*) for the ten speakers. The data reveal consistent effects of within-word position of the accented syllable on H delay: for all 10 speakers, peak delay is significantly shorter in word-final syllables than in penultimate syllables. The differences range from 27 to 45 ms, depending on the speaker.

This pattern is replicated by the data for H distance in time to the end of the accented syllable. The two histograms in Fig. 6 plot the mean time in ms from the location of the H tonal target to syllable offset in oxytonic and paroxytonic words for Catalan (left panel) and for Spanish (right panel). First, the positive measures reveal that *all* peaks are located in the postaccentual syllable. Taking the 0 value as the end of the accented syllable, the plots show that all f_0 peaks are displaced to the postaccentual syllable or syllables, since all peaks are located well beyond the 0 value. For the five Catalan

[6] Schepman et al. (2006) run the same peak alignment analyses on the same Dutch data using different quantitative definitions of alignment, including two with different segmental reference points (onset of the syllable and beginning of the postvocalic consonant) and one in which alignment is considered as a proportion of the syllable rhyme. They concluded from their results that "it is better to express alignment relative to a nearby acoustic landmark than a more distant one", since "the more distant the landmark, the greater the variance and the greater the likelihood of uninformative correlations." (Schepman et al., 2006: 22) The greater variance in the case of more distant acoustic landmarks (e.g., the onset of the syllable with respect to the peak) is probably due to intrinsic variations in segmental duration.

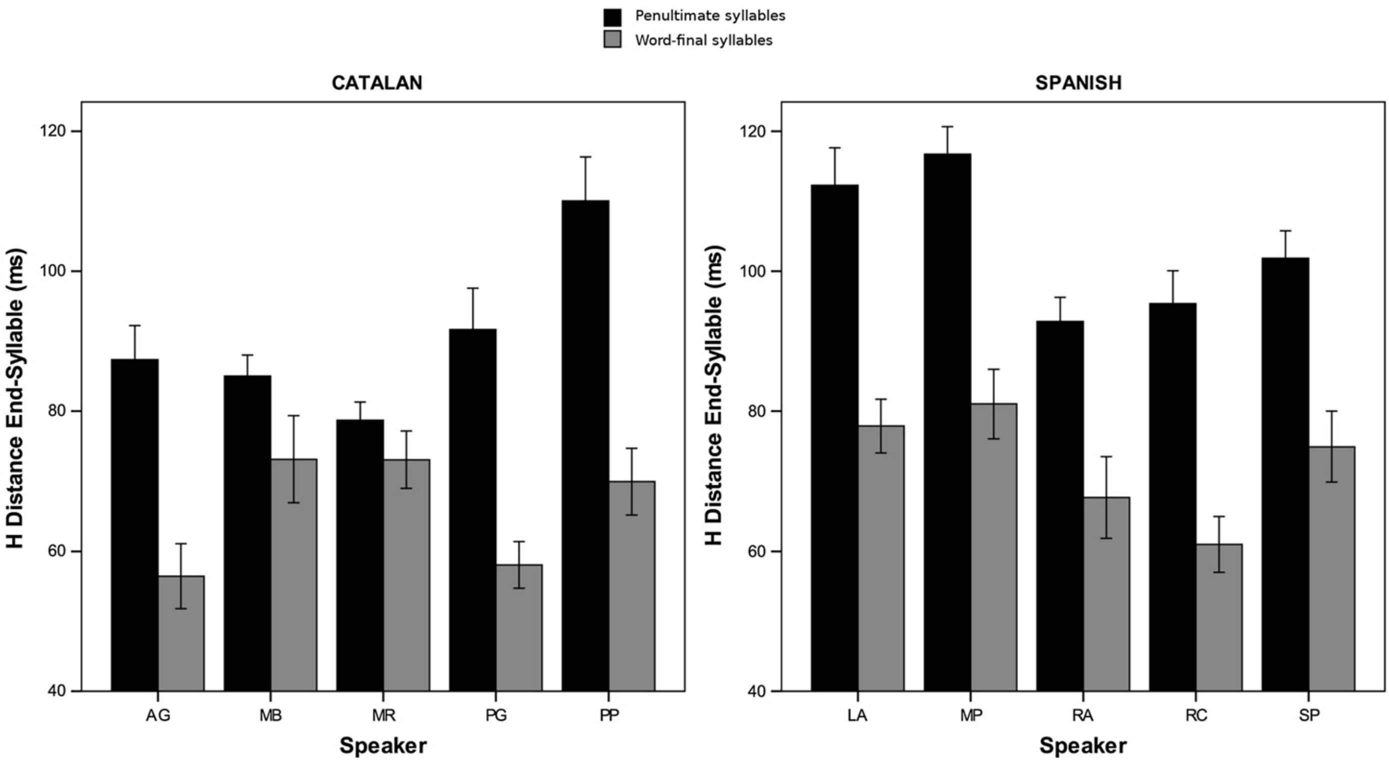

Fig. 6. Mean time from the H peak to the end of the accented syllable as a function of within-word position for five Catalan speakers (left panel) and for five Spanish speakers (right panel). Light bars show the means in the word-final condition and black bars show the mean in the word-medial condition. The vertical bars represent standard errors.

Table 1
One-way ANOVA summaries of the effects of Within-Word Position (Position) on two measures of H location, namely, Peak Delay from Accented Syllable Onset and Peak delay from Accented Syllable End (or time from the peak to the end of the accented syllable), for all Catalan and Spanish subjects.

	Peak Delay from Accented Syllable Onset/Position	Peak delay from Accented Syllable End/Position
Catalan		
AG	$F(1,120)=15.29$, $p<0.001$	$F(1,120)=31.45$, $p<0.001$
MB	$F(1,120)=3.59$, $p<0.05$	$F(1,120)=10.07$, $p<0.05$
MR	$F(1,120)=4.29$, $p<0.05$	$F(1,120)=9.11$, $p<0.05$
PG	$F(1,120)=8.40$, $p<0.05$	$F(1,120)=14.93$, $p<0.05$
PP	$F(1,120)=7.47$, $p<0.05$	$F(1,120)=24.83$, $p<0.001$
Spanish		
LA	$F(1,120)=4.51$, $p<0.05$	$F(1,120)=37.96$, $p<0.001$
MP	$F(1,120)=4.80$, $p<0.05$	$F(1,120)=46.39$, $p<0.001$
RA	$F(1,120)=4.16$, $p<0.05$	$F(1,120)=40.08$, $p<0.001$
RC	$F(1,120)=7.34$, $p<0.05$	$F(1,120)=55.72$, $p<0.001$
SP	$F(1,120)=8.32$, $p<0.05$	$F(1,120)=24.53$, p <0.001

speakers and five Spanish speakers, peak time to the end of the syllable is significantly shorter in word-final position than in penultimate position (at $p<0.05$; see Table 1). The mean time of the f_0 peak to the right edge of the syllable is 61 ms for oxytonic words and 96 ms for paroxytonic words, that is, there is a mean difference of 35 ms.[7]

The alignment patterns observed in Figs. 5 and 6 were statistically confirmed. For each variable (peak delay and peak distance in time to end of syllable), we ran a factorial repeated measures ANOVA (one for each language) including Within-Word Position as the main predictor, Repetition as an intra-subject factor and Subject as an inter-subject factor.

For peak delay in the Catalan data, results indicate a main effect of Within-Word Position ($F(1,1)=116.57$, at $p<0.0001$), no effect of Repetition ($F(1,3)=1.83$, at $p=0.147$), and no interaction between Within-Word Position and Repetition ($F(1,3)=0.30$, at $p=0.993$). There was a significant interaction effect between Within-Word Position and Subject ($F(1,4)=3.385$, at $p<0.05$) and Repetition and Subject ($F(1,12)=0.35$, at $p<0.0001$). In the post-hoc subject two-pair comparisons, subject 5 (PP) is the only speaker that is significantly different from the rest; the reason for this is that her peak delay measures were larger for this speaker; yet crucially the presence and direction of the effect of within-word position is the same as in the other speakers. As for peak delay in the Spanish data, results also reveal a main effect of Within-Word Position ($F(1,1)=97.68$, at $p<0.0001$), no effect of Repetition ($F(1,3)=0.386$, at $p=0.764$), and a significant interaction between Within-Word Position and Repetition ($F(1,3)=3.406$, at $p<0.05$). There were no significant interactions between Within-Word Position and Subject ($F(1,4)=1.506$, at $p=0.207$) and Repetition and Subject ($F(1,12)=1.413$, at $p=0.159$).

The statistical results for the Accented Syllable End (or time from the H peak to the end of the accented syllable) reveal very similar patterns to those found for Peak Delay. For Catalan, results indicate a main effect of Within-Word Position ($F(1,1)=161.84$, at $p<0.0001$), no effect of Repetition ($F(1,3)=0.647$, at $p=0.587$), and no interaction between Within-Word Position and Repetition ($F(1,3)=0.460$, at $p=0.711$). There was a significant interaction effect between Within-Word Position and Subject ($F(1,4)=12.046$, at $p<0.0001$) and Repetition and Subject ($F(1,12)=2.213$, at $p<0.05$). In this case, subject two-pair comparisons analyses

[7] We note a small difference in the peak location results for Catalan and Spanish: the H is located slightly later in Spanish than in Catalan. As one anonymous reviewer points out, this might be due to the fact that the Spanish materials contained some cases where the target accent syllable was word- and phrase-initial (e.g., *bebo vinos* 'I drink wines'); whereas in the Catalan materials the target accented syllable is never word-initial (e.g., *compraven talls* 'they bought pieces'). We could speculate that word-initial accented syllables might have induced a later peak in the Spanish materials.

reveal several differences between pairs of speakers for both languages. For Spanish, results also reveal a main effect of Within-Word Position ($F(1,1)=271.57$, at $p<0.0001$), a significant effect of Repetition ($F(1,3)=3.246$, at $p<0.05$), and no significant interaction between Within-Word Position and Repetition ($F(1,3)=1.479$, at $p=0.225$). There were no significant interactions between Within-Word Position and Subject ($F(1,4)=1.293$, at $p=0.278$) and Repetition and Subject ($F(1,12)=0.885$, at $p=0.563$).

In order to evaluate the effects of Within-Word Position for each speaker, one-way ANOVAs were run separately for each speaker. Remember the significant interaction between Subject*Repetition and Subject*Within-Word Position found for Catalan in the factorial repeated measures ANOVAs. Table 1 shows ANOVA summaries of the effects of Within-Word Position on two measures of H location, namely, Peak Delay from Accented Syllable Onset and Peak delay from Accented Syllable End (or time from the peak to the end of the accented syllable). Differences are statistically significant for all 10 speakers (all effects are reported as significant at $p<0.05$).

Finally, we analyzed the alignment of the H tonal target relative to the end of the word. As mentioned above, prenuclear rising accents in Catalan have been analyzed as a combination of a L* tone plus a H word tone (Estebas-Vilaplana, 2000, 2003). If the hypothesis about the existence of a word-edge tone H in Catalan is correct, then the f_0 peak should be aligned systematically with the right edge of the word regardless of the number of posttonic syllables in the sequence. Thus, in *comprà ventalls* the f_0 peak would be located around the end of the syllable *prà* whereas in *compraven talls* the f_0 peak would be located towards the end of the postaccentual syllable.

The two histograms in Fig. 7 plot the mean values of the time from the H peak to the right edge of the word as a function of within-word position for the five Catalan speakers (left panel) and the five Spanish speakers (right panel). The horizontal line (at value 0 in the *y*-axis) graphically indicates the position of the word boundary. All ten speakers show a consistent trend: while peaks in word-final accented syllables are located after the end of the word (a mean of 61 ms), peaks in word-medial accented syllables are located before the end of the word (a mean of −47 ms). Thus the data show that the H tone is not anchored at the end of the prosodic word. Taking the 0 value as the end of the word, the results show that in words with a final accent, the H occurs after the end of the word in both Catalan and Spanish. Alternatively, in words with a penultimate accent, the H is located before the end of the word.

A factorial ANOVA was applied to this dependent variable (i.e., distance in time from the peak to the right edge of the word), including Within-Word Position as the main predictor, and Subject as a random factor. The analyses were run separately for each language. For both languages, results revealed a main effect of Within-Word Position on H location with respect to Word Boundaries (for Catalan $F(1,1)=529.41$, at $p<0.0001$, and for Spanish $F(1,1)=719.526$, at $p<0.0001$), no effects of Speaker (for Catalan $F(1,4)=0.470$, at $p<0.759$, and for Spanish $F(1,4)=0.855$, at $p<0.559$), and finally a significant interaction between Within-Word Position and Speaker (for Catalan $F(1,4)=5.934$, at $p<0.0001$, and for Spanish $F(1,4)=7.382$, at $p<0.0001$). Post-hoc pair comparisons revealed that the interaction is due to the different magnitude of the effect in different speakers. Table 2 shows the ANOVA summaries of the effects of Within-Word Position (Position) on this measure of H location, namely H location in time with respect to Word Boundaries, for every speaker. Effects are statistically significant for the 10 speakers (two-tailed *t*-tests were all significant at $p<0.001$).

Summarizing, the ten speakers show statistically significant effects of within-word position on the location of the f_0 peak relative to both the beginning and the end of the syllable. While the presence of an adjacent word boundary triggers a relatively earlier alignment of f_0 peak with respect to the end of the word, a late word

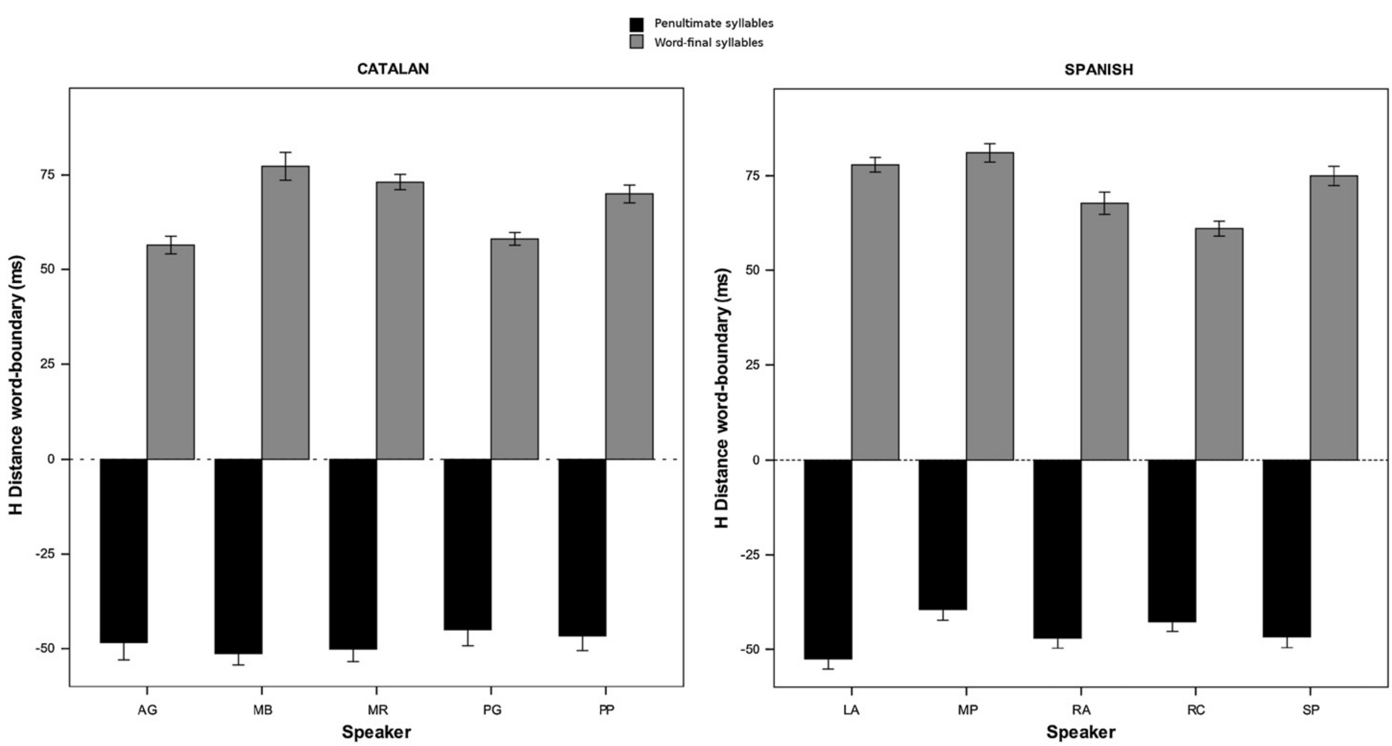

Fig. 7. Mean values of time from the H peak location relative to the right edge of the word (in ms) as a function of within-word position for the five Catalan speakers (left panel) and the five Spanish speakers (right panel). Light bars show the means in the word-final condition and black bars show the mean in the word-medial condition. The horizontal line (at value 0 in the *y*-axis) graphically indicates the position of the word boundary. The vertical bars represent standard errors.

Table 2
One-way ANOVA summaries of the effects of Within-Word Position (Position) on H location with respect to Word-Boundary.

	H Location with respect to Word Boundaries
Catalan	
AG	$F(1,120)=122.02$, $p<0.001$
MB	$F(1,120)=225.11$, $p<0.001$
MR	$F(1,120)=280.14$, $p<0.001$
PG	$F(1,120)=149.57$, $p<0.001$
PP	$F(1,120)=219.91$, $p<0.001$
Spanish	
LA	$F(1,120)=503.03$, $p<0.001$
MP	$F(1,120)=324.90$, $p<0.001$
RA	$F(1,120)=341.11$, $p<0.001$
RC	$F(1,120)=349.51$, $p<0.001$
SP	$F(1,120)=297.78$, $p<0.001$

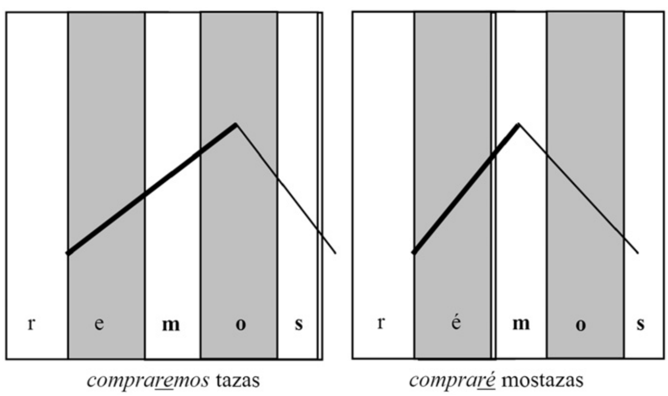

Fig. 8. Schematic diagram of the f_0 peak location with respect to the end of the syllable and the end of the word in paroxytones vs. oxytones. The double bars represent the location of the word boundary. Post-tonic syllables are marked in bold. Left panel: paroxytonic words and right panel: oxytonic words.

boundary leads to a later alignment of H. Fig. 8 shows a schematic diagram of the difference in f_0 location patterns with respect to the end of the syllable in two basic conditions, *compraremos tazas*, and *compraré mostazas*. In each graph, the solid vertical lines represent the syllable boundaries, and the thick dotted lines represent the location of the word boundary. The mean f_0 curve shows that the time from the peak to the end of the accented syllable is larger in paroxytonic than in oxytonic words:

In summary, the production data analyzed in this section do not support the idea that the rise on the target word is the implementation of a word-edge tone H that aligns with the right edge of the word.[8]

2.2.3. *Effects of within-word position on syllable duration*

We are interested in describing the duration patterns found in word-initial and word-final positions in our data, as they might be interacting with the peak alignment patterns. In this section, we analyze (1) the duration of the accented syllable as a function of within-word position (it is expected that accented syllables before a word-boundary will be longer than accented syllables in medial position), and (2) the duration of the postaccentual syllable as a function of within-word position (it is expected that postaccentual syllables before a word-boundary will be longer than non-accented syllables in medial position).[9] Beckman and Edwards (1990: 176) note that in English "phrase-final lengthening occurs at intonational-phrase boundaries, and is a large effect that is highly consistent across speakers and rates. Word-final lengthening occurs at some other constituent boundaries, and is a much smaller effect that is not consistently evident across speakers and rates." Similarly, recent studies on the durational and articulatory effects of word boundaries, show that articulation of a constituent-initial consonant tends to be stronger than articulation of the same consonant constituent-medially, and that this strengthening increases for segments which are initial in increasingly high-level prosodic constituents (for a review, see Byrd and Saltzman, 1998; Fougeron and Keating, 1997, among others).

The two histograms in Fig. 9 show the mean values of duration of the accented syllable in words with final and medial stress for Catalan (left panel) and Spanish (right panel) and in words with final and penultimate stress. The histograms show no consistent effects of word edges on syllabic duration. Word-final accented syllables are slightly longer than word-medial accented syllables for four Spanish speakers and for two Catalan speakers. However, these differences were only significant for two Spanish speakers according to the results of the one-way ANOVAs presented in Table 3 (at $p<0.004$). This is indicating that durational differences are optional correlates of the presence of word boundaries.

For each variable (DurAcSyl and DurPostAcSyl), we ran a factorial repeated measures ANOVA (one for each language) including Within-Word Position as the main predicting factor, Repetition as an intra-subject factor and Subject as an inter-subject factor.

For duration of the accented syllable (DurAcSyl) in the Catalan data, results indicate a main effect of Within-Word Position ($F(1,1)=10.036$, at $p<0.05$), an effect of Repetition ($F(1,3)=4.245$, at $p<0.05$), and no interaction between Within-Word Position and Repetition ($F(1,3)=0.121$, at $p=0.948$). There was a significant interaction effect between Within-Word Position and Subject ($F(1,4)=10.692$, at $p<0.0001$) and Repetition and Subject ($F(1,12)=5.109$, at $p<0.0001$). Post-hoc subject two-pair comparisons reveal that there are differences between the duration patterns produced by the five subjects (see also Fig. 9). As for duration of the accented syllable in the Spanish data, results also reveal a main effect of Within-Word Position ($F(1,1)=17.645$, at $p<0.0001$), no effect of Repetition ($F(1,3)=2.076$, at $p=0.109$), and no significant interaction between Within-Word Position and Repetition ($F(1,3)=1.734$, at $p=0.165$). There were no significant interactions between Within-Word Position and Subject ($F(1,4)=1.397$, at $p=0.241$) and Repetition and Subject ($F(1,12)=1.531$, at $p=0.112$).

The statistical results for the duration of the postaccentual syllable (DurPostAcSyl) analyses revealed inconsistent patterns of results. For Catalan, results indicate a main effect of Within-Word Position ($F(1,1)=4.253$, at $p<0.05$), no effect of Repetition ($F(1,3)=1.862$, at $p=0.142$), and no interaction between Within-Word Position and Repetition ($F(1,3)=1.853$, at $p=0.143$). There was a significant interaction effect between Within-Word Position and Subject ($F(1,4)=6.295$, at $p<0.0001$) and Repetition and Subject ($F(1,12)=2.731$, at $p<0.05$). For Spanish, results revealed no main effect of Within-Word Position ($F(1,1)=0.621$, at $p=0.433$), no significant effect of Repetition ($F(1,3)=0.470$, at $p=0.704$), and no significant interaction between Within-Word Position and Repetition

[8] We entertain the possibility that the strict alignment effects between H peaks and the end of the word found by Estebas-Vilaplana (2000, 2003) might have been conditioned by the presence of a phrase break signaled by a phrase accent after subjects and verbs when the object is long enough. We do not discard the possibility, though, that this might be an optional phenomenon that is not attested in our data.

[9] Upon the associate editor's request, we performed the analysis on the word-initial consonant, and the main results were comparable to the syllabic results, that is, no significant effects were found across conditions. Yet the lack of statistically significant effects could have been due to consonant type variation. In our data, the variation in the consonant type that appears in word-initial position might have been responsible for the high magnitude in the variability measures.

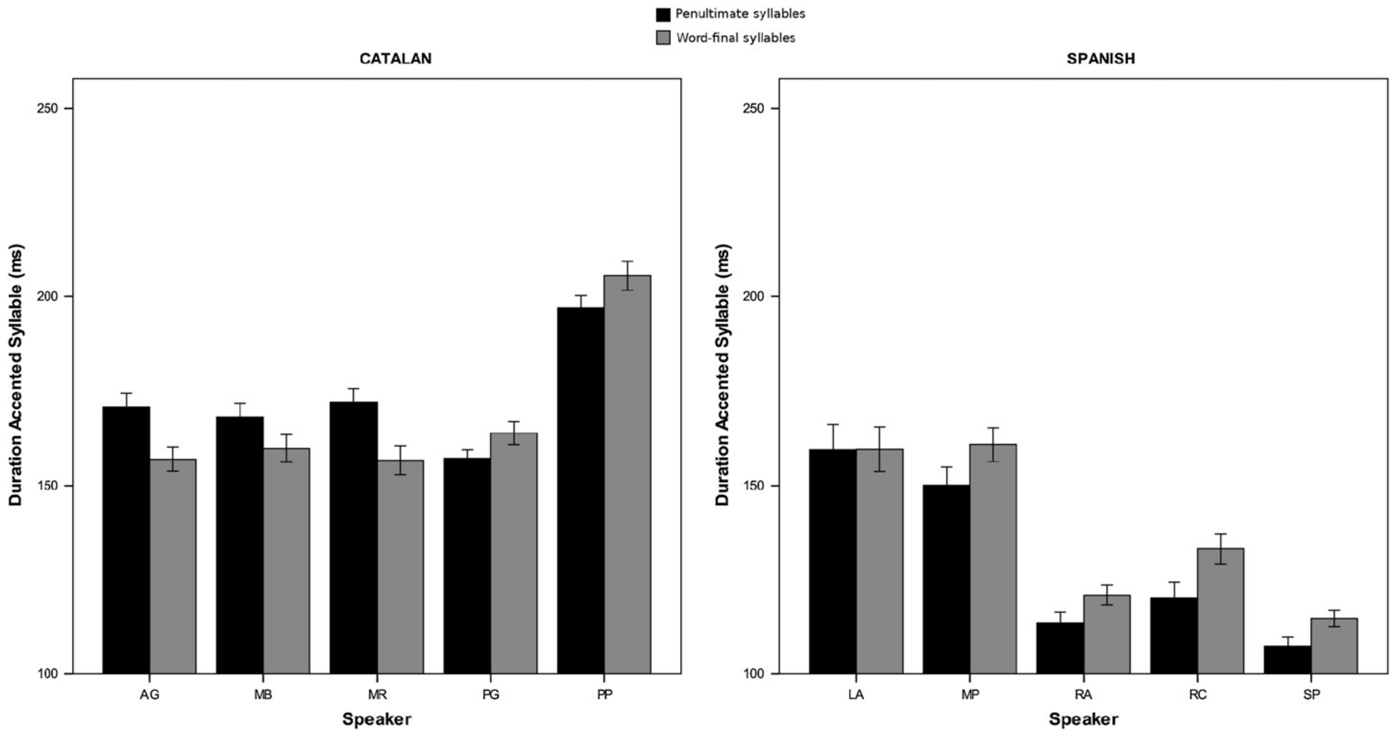

Fig. 9. Mean values of duration of the accented syllable (in ms) as a function of within-word position for all Catalan (left panel) and Spanish speakers (right panel) in word-final syllables (grey bars) and penultimate syllables (black solid bars). The vertical bars represent standard errors.

Table 3
One-way ANOVA summaries of the effects of Within-Word Position (Position) on the duration of accented syllables and post-accentual syllables.

	DurAcSyl	DurPostAcSyl
Catalan		
AG	$F(1,120)=2.30$, $p=0.1318$	$F(1,120)=0.11$, $p=0.7354$
MB	$F(1,120)=2.62$, $p=0.1079$	$F(1,120)=0.02$, $p=0.8791$
MR	$F(1,120)=7.05$, $p=0.0090$	$F(1,120)=0.00$, $p=0.9703$
PG	$F(1,120)=2.02$, $p=0.1574$	$F(1,120)=0.02$, $p=0.8810$
PP	$F(1,120)=0.93$, $p=0.3367$	$F(1,120)=1.19$, $p=0.2783$
Spanish		
LA	$F(1,120)=0.38$, $p=0.5395$	$F(1,120)=0.03$, $p=0.8632$
MP	$F(1,120)=0.62$, $p=0.4332$	$F(1,120)=0.05$, $p=0.8180$
RA	$F(1,120)=3.69$, $p=0.0573$	$F(1,120)=0.98$, $p=0.3246$
RC	$F(1,120)=12.46$, $p=0.0006$	$F(1,120)=0.01$, $p=0.9281$
SP	$F(1,120)=4.04$, $p=0.0468$	$F(1,120)=0.14$, $p=0.7084$

($F(1,3)=2.646$, at $p=0.054$). There were no significant interactions between Within-Word Position and Subject ($F(1,4)=1.420$, at $p=0.233$) and Repetition and Subject ($F(1,12)=1.410$, at $p=0.160$). Again, subject two-pair comparison analyses reveal great differences between pairs of speakers for both languages.

The two histograms of Fig. 10 show the mean values of duration of the postaccentual syllable in words with final and penultimate stress for Catalan, the language that showed a significant interaction between Within-Word Position and Subject. Again, the histogram reveals no consistent durational patterns in the data. Postaccentual syllables in non-word-final position are slightly longer than postaccentual syllables in word-final position for some speakers, with the exception of Catalan speakers PP and PG. Moreover, none of the speakers showed statistically significant differences (at $p>0.05$) across both conditions except for speakers PP (Catalan) and RA (Spanish) (see Table 3). Thus, the hypothesis that the postaccentual syllable will be longer before a word boundary is not confirmed by these data.

In order to evaluate the effects of Within-Word Position for each speaker, Table 3 shows ANOVA summaries of the effects of Within-Word Position (Position) on the duration of accented and postaccentual syllables, for each speaker. For the duration of accented syllables, effects are only statistically significant for one Catalan speakers (MR) and for one Spanish speaker (RC, at $p<0.01$). As for the duration of postaccentual syllables, none of the speakers showed significant results.

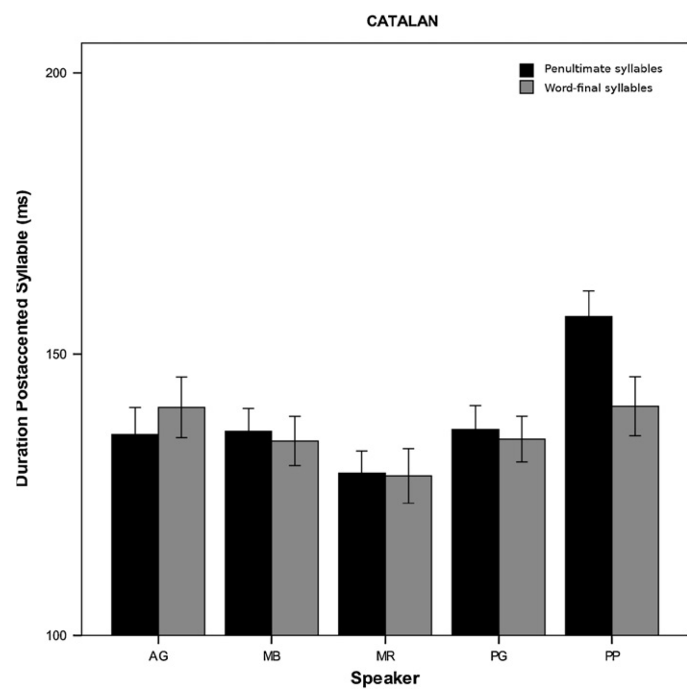

Fig. 10. Mean values of duration of the postaccentual syllable (in ms) as a function of within-word position for all Catalan speakers in word-final syllables (grey solid bars) and penultimate syllables (black solid bars). The vertical bars represent standard errors.

2.2.4. Linear models of Peak Placement

In order to capture the combined effects of the two factors, Within-Word Position (Paroxytonic, Oxytonic) and duration of accented syllable (DurAcSyl) on peak placement, a linear regression model was applied to the data using the delay from the peak to the end of the syllable (Peak delay from Accented Syllable End) as the response measure. In essence, our goal in this section is to test the relevance of the wihin-word position factor together with accented syllable duration in the prediction of H alignment. Table 4 shows the results of the linear mixed effects model. First, the main factors (Within-Word Position and the duration of accented syllable, DurAcSyl) were statistically significant (at $p<0.0001$) in both languages. In Catalan, the response variable Peak delay from Accented Syllable End estimate takes the value of 78 ms. As can be seen in the first column, when the factor within-word position equals 1 (that is, when the word is paroxytonic) the variable Time-to-End-Syllable increases by 25 ms in Catalan and by 30 ms in Spanish, the mean value of Peak delay from Accented Syllable End is 101 ms. Finally, the effects of syllable duration are contradictory: while in Catalan when the syllabic duration increases Peak delay from Accented Syllable End also increases, in Spanish the situation is the reverse (by 33 percent). The estimate term works in absolute terms for the factor within-word position and in percentage terms for the duration of the accented syllable.

Summarizing, the mixed model analyses for the two languages clearly support the prediction that H alignment is affected by the position of the accented syllable within the word, that is, by the Word Boundary condition. This behavior confirms the hypothesis that the end of the word (and not only the presence of upcoming accents or boundary tones) is acting as a kind of prosodic boundary exerting prosodic pressure on H tonal targets. In the production experiment, the other potentially confounding factors, that is the upcoming presence of boundary tones and pitch accents, have been controlled for and thus the only possible reason for this behavior is the presence of a word edge.

To summarize the results of this section, the results of the production experiment have revealed a set of production regularities in the alignment of H prenuclear peaks in both languages, namely, that H placement is sensitive to the prosodic word domain including the target accented syllable. Catalan and Spanish speakers show statistically significant effects of within-word position on the location of the f_0 peak relative to both the beginning and end of the syllable. While the presence of an adjacent word boundary triggers a relatively earlier alignment of f_0 peak, a late word boundary leads to a later alignment of H. This evidence supports other acoustic studies of a variety of languages, which have reported the same tendency (Arvaniti et al., 1998 for Greek; Chahal 2001, 2003 for Lebanese Arabic; Prieto et al., 1995 for Spanish; Silverman & Pierrehumbert, 1990 for English).

Table 4
Results of the Linear Mixed Effects Model that uses Time-to-End-Syllable as the response variable and stress pattern (*Oxytonic, Paroxytonic*) and duration of the accented syllable (*DurAcSyl*) as predictive factors.

	Estimate	Error	Df	T Value	Signif
CATALAN					
Intercept	78.5242	5.7356	4	13.69	$p<0.0002$
Paroxytonic=1	24.7803	5.7356	793	15.85	$p<0.0001$
Oxytonic=0	0				
DurAcSyl	0.8647	0.04663	793	18.55	$p<0.0001$
SPANISH					
Intercept	101.29	6.1609	4	16.44	$p<0.0001$
Paroxytonic=1	29.5373	1.3676	791	21.60	$p<0.0001$
Oxytonic=0	0				
DurAcSyl	−0.3297	0.03758	791	−8.77	$p<0.0001$

3. Perception experiments

Based on the findings of the production experiment that the alignment of the f_0 peak is significantly affected by the position of the accented syllable within the word both in Spanish and in Catalan, we then proposed to test the perceptual effects of the peak alignment on word boundary identification. The perception experiments were a set of identification tasks (see Section 3.1). The purpose was to verify whether f_0 peak alignment changes would help in lexical identification tasks of otherwise identical utterances (e.g., *buscà vanguàrdies* '(s)he looked for newspapers' and *buscaven guàrdies* 'they were looking for guards').

We hypothesize that if these tonal alignment patterns are used as perceptual cues by listeners, helping them to identify the position of the word boundary, the identification functions will be affected. At this point, we do not make any assumption as to whether other prosodic cues such as duration might be relevant for word segmentation in Catalan and Spanish. In this experiment, we simply test whether a systematic change in peak alignment affects the identification responses given by listeners.

3.1. Method

The general method for this set of experiments involved altering the alignment of the H peak artificially with respect to its position to the end of the syllable, and then using a series of identification to test the effect that these alterations produce on the listeners' word identification tasks. Given the results of previous pilot tests, our hypothesis is that the contrast between the two stimuli is not categorical in nature, but rather of a continuous type. Thus we were expecting to find a small but significant effect of stimulus type in the identification tasks.

3.1.1. Stimuli

The following four utterances were selected from the corpus of sentences read in Experiment 1, two produced by a female Catalan speaker (MB) and two produced by a female Spanish speaker (SP).

Stimuli	Stimulus 1	Stimulus 2
Catalan	*buscà vanguàrdies*	bus*caven guàrdies*
Spanish	*compraré mostazas*	compra*remos tazas*

For the selection procedure, the main criterion to select the sentences for the perception test was that the target accent in the two utterances was representative of the average values of f_0 peak location and duration of the target vowels obtained in the production test. Also, in order to avoid confounding our results with durational and f_0 peak scaling (or f_0 height) cues, we closely checked that the chosen target stimuli showed minimal differences in either the syllable duration of the two target syllables (Catalan: ca: 166 ms in *buscaven* vs. 158 ms in *buscà##ven*: 211 ms in *buscaven* vs. 209 ms in *buscà*; Spanish: *ré*: 66 ms in *compraré* vs. 60 ms in *compraremos*; *mos*: 204 ms in *compraré* vs. 196 ms in *compraremos*) or the scaling of the peak of the two target accents (Catalan: 239 Hz in target word *buscaven* vs. 240 Hz in target word *buscà*; Spanish: 206 Hz in target word *compraré* vs. 206 Hz in target word *compraremos*).

These four sentences constituted the basis for the manipulations for utterances used in the perceptual experiments. One multi-step continuum from each sentence was created, by varying the alignment of the f_0 peaks. The stimuli for the perception experiments were obtained by acoustically manipulating this intonation variable artificially using the Pitch Synchronous Overlap and Add (PSOLA) resynthesis routine available in the Praat speech analysis and resynthesis software (Boersma & Weenink, 2005; Wood, 2005), with the goal of testing the effect of these changes on the listeners' judgments on the location of the word boundary.

From the original pair of utterances per language, a set of stimuli were created in which the alignment of the f_0 peak was varied in five equidistant steps in the continuum, five for each stimuli (for a total of ten stimuli per language). The closest landmark was the end of the syllable. Fig. 11 shows the five-step alignment continuum created from the two source utterances, indicated in the graph through the use of the solid line and the initials 'ewb' and 'lwb' ('early word boundary' and 'late word boundary', respectively). Note that the acronym 'ewb' (early word boundary) corresponds to the examples with a word break after the accented syllable (Cat. *buscà vanguàrdies* and Sp. *compraré mostazas*) and 'lwb' (late word boundary) corresponds to the examples with a word break one syllable after the target syllable (Cat. *buscaven guàrdies* and Sp. *compraremos tazas*). The peak manipulation was performed by delaying the peak by 22 ms in the 'ewb' cases and by retracting it by 22 ms in the 'lwb' cases (for a total of 88 ms). After this, a linear interpolation line was created between the onset and the f_0 peak. Note that this modification of alignment slightly affects the *slope* of the rising pitch movements. Duration and scaling properties of the original source utterance were not modified (see above).

The position of the peaks with respect to the end of the syllable were not the same in the Catalan and Spanish original base stimuli. Table 5 shows a summary of the manipulations in alignment made in the two languages, both for the *vanguàrdies/mostazas* base stimulus (left columns: ewb1, ewb2, ewb3, ewb4, ewb5) and for the *guàrdies/tazas* base stimulus (right columns: lwb1, lwb2, lwb3, lwb4, lwb5). As shown by the table, the peak position in the base stimuli (ewb1 and lwb5, coded using bold letters) is different in both languages: while the peak in Catalan ewb1 is located 53 ms after the end of the syllable, in Spanish it is located 20 ms after; in the case of lwb5, while in Catalan the peak occurs 141 ms after the end of the syllable, in Spanish occurs 110 ms after the end of the syllable. This was due to the fact that we preserved the peak timing in the original stimuli.

3.1.2. Experimental procedure

The identification task was set up by means of the PERCEVAL software developed in the Laboratoire Parole et Langage, Aix-en-

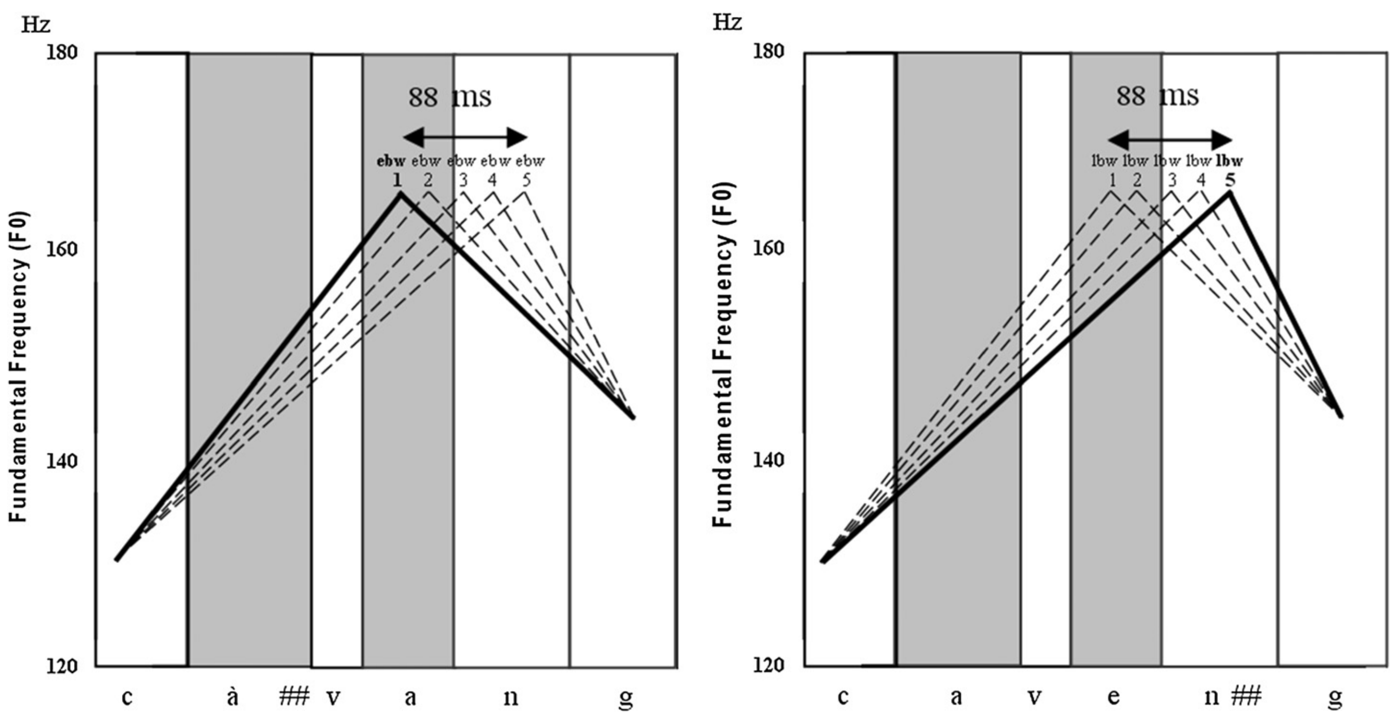

Fig. 11. Five-step H alignment continuum diagrams for the *vanguàrdies/mostazas* base stimulus (ewb1, ewb2, ewb3, ewb4, ewb5; left panel) and for the *guàrdies/tazas* base stimulus (that is, lwb1, lwb2, lwb3, lwb4, lwb5; right panel).

Table 5
Summary of H alignment steps. The original base (natural) stimuli are coded using bold letters.

CATALAN Steps	Base stimulus: *buscà vanguàrdies*	CATALAN Steps	Base stimulus: *buscaven guàrdies*
ewb1	**53 ms after end syllable**	lwb1	53 ms after end syllable
ewb2	75 ms (+22 ms)	lwb2	75 ms (+22 ms)
ewb3	97 ms (+22 ms)	lwb3	97 ms (+22 ms)
ewb4	119 ms (+22 ms)	lwb4	119 ms (+22 ms)
ewb5	141 ms (+22 ms)	**lwb5**	141 **ms (+22 ms)**
SPANISH Steps	Base stimulus: *compraré mostazas*	SPANISH Steps	Base stimulus: *compraremos tazas*
ewb1	**20 ms after end syllable**	lwb1	20 ms after end syllabñe
ewb2	42 ms (+22 ms)	lwb2	42 ms (+22 ms)
ewb3	64 ms (+22 ms)	lwb3	64 ms (+22 ms)
ewb4	88 ms (+22 ms)	lwb4	88 ms (+22 ms)
ewb5	110 ms (+22 ms)	**lwb5**	110 **ms (+22 ms)**

Provence (André, Ghio, Cavé, & Teston, 1995–2003). The stimuli were played back on laptop or desktop computers, using high quality headphones. The experimental sessions for the Catalan experiments were conducted in quiet rooms at the Universitat Autònoma de Barcelona, and the sessions for the Spanish experiments in quiet rooms at the Universidad Nacional de Educación a Distancia. The experiment was set up in such a way that the next stimulus was presented only after a response was given. Data of the responses and the Reaction Time (RT) were automatically recorded in PERCEVAL.

In the identification task, participants had to judge the stimuli generated from the peak alignment continuum and decide which of the two possible sentences they were listening to (i.e. *buscà vanguàrdies* vs. *buscàven guàrdies* for Catalan and *compraré mostazas* vs. *compraremos tazas* for Spanish). Participants were seated in front of a computer and given a set of instructions. They were told that they would hear a series of very similar utterances that would be presented individually. They were also asked to keep their hands on the keyboard and respond as fast as possible once they were sure of their response (but not before the end of the utterance), as we were also interested in analyzing response time.

To test the adequacy of the participants' performance and their sensitivity to the experimental method, an initial training test was performed prior to the experiment proper using six tokens of original utterances. The whole experiment lasted approximately 20 min.

For the identification test, listeners were seated in front of the computer and instructed to perform a two-alternative forced choice task. Specifically, participants heard one stimulus at a time and had to identify each step of the continuum as either *compraven guàrdies* or *comprà vanguàrdies* for Catalan and as either *compraremos tazas* or *compraré mostazas* for Spanish. Participants doing the test in Catalan were told to listen to each of the target utterances carefully, and indicate their choice by pressing the "G" key (for the object name ***G**UÀRDIES*) or the "V" key (for ***V**ANGUÀRDIES*). For the Spanish version, participants indicated their choice by pressing either the "T" key (for ***T**AZAS*) or the "M" key (for ***M**OSTAZAS*).

Listeners performed the identification task in which the ten stimulus group with each of the alignment steps was played five times in five different blocks, for a total of 50 stimuli. The start of each block presentation was preceded by a visual message on the screen. All the stimuli were automatically randomized by PERCEVAL.

3.1.3. Participants

Twenty native speakers of Central Catalan and twenty native speakers of Peninsular Spanish participated in the experiment. The Catalan participants were first and second year undergraduates from the Catalan Philology program at the Universitat Autònoma de Barcelona. The Spanish listeners were undergraduate and graduate students from the Spanish Philology program at the UNED. They all reported having normal hearing and reading skills.

3.1.4. Statistical analyses

Because the data from the perception experiments did not meet the distributional requirements essential for parametric testing (i.e., the data were not continuous and were thus not normally distributed), a set of non-parametric tests was used. First, the Friedman non-parametric test (a multiple testing technique similar to parametric ANOVA) was performed on the Catalan and Spanish identification data to test the overall effects of Alignment Step on the identification curve. The data were always separated by language and the two base stimuli in each language. After that, a Wilcoxon signed-rank test was used to compare between groups of Alignment Steps.

3.2. Results

3.2.1. Effects of H alignment

If details of H alignment are used as a perceptual cue in word identification tasks, we would expect that shifting the pitch peak later in time will progressively change the percept from *buscà vanguàrdies*/*compraré mostazas* to *buscaven guàrdies*/*compraremos tazas*, since later alignment is characteristic of paroxytones. The results of the identification experiments for the H alignment stimuli are presented in Fig. 12. The two graphs in this figure plot the identification rate for the two continua created from the "early word boundary" (in grey) and "late word boundary" (in black) base stimuli. The identification rate is defined as the number of "early word boundary" responses (in early word boundary-based stimuli) or "late word boundary" responses (in late word boundary-based stimuli) over the total responses. The *x*-axis represents the five steps of H alignment manipulations, either from the *late word boundary* base stimulus (lwb1, lwb2, lwb3, lwb4, lwb5) or the *early word boundary* base stimulus (ewb1, ewb2, ewb3, ewb4, ewb5). Each data point represents a total of 100 judgments (20 listeners $\times$ 5 repetitions).

Results from both languages show a shallow S-shaped curve, that is, in all panels we see a very gradual shift from *late word boundary* to *early word boundary* judgments as a function of H alignment steps: at late locations of H within the continuum (lwb5, ewb5) *late word boundary* responses are slightly more dominant, while at the opposite end of the continuum (lwb1, ewb1) *early word boundary* responses are more dominant. Second, though an effect can be seen in both languages, no categorical boundary shift due to the H alignment manipulation is present: ratio differences range from 0.20 to 0.40 at the two ends of the four curves and two of them do not cross the 50% identification boundary. Third, the identification curves are different depending on the base stimulus (see the patterns of responses in black and in grey), meaning that there might be other cues in the stimulus that lead to a given response. A Wilcoxon signed-rank test was used to compare the two identification functions from the different base stimuli. The difference between the two was found to be statistically significant at $p < 0.0001$. Finally, the identification functions show a difference between the behavior of Catalan and Spanish that the response shift from the left to right ends of the continua is more clearcut in the results for Spanish.

First, a Friedman non-parametric test—a test which uses a multiple testing technique similar to parametric ANOVA—was applied separately to Catalan and Spanish identification data, again separating the data for the two base stimuli. The test was used to evaluate the overall effects of H alignment steps (independent variable) on word-boundary identification (0 or 1, response variable). Table 6 shows the results. As expected, the main effect of H alignment is significant both for Spanish and Catalan (both for the *late word boundary* and the *early word boundary* stimuli set)—the significance level was set to 0.05, asterisks mark significant effects.

A second statistical analysis, namely the Wilcoxon matched pairs signed rank test, was performed in order to see whether there were significant differences between pairs of alignment steps. As can be seen in Table 7, this analysis revealed that almost all comparisons between groups separated by just one step (1–2, 2–3, etc.) were non-significant for both languages, except for three Catalan comparisons (asterisks mark significant effects). When stimulus pairs with a step difference greater than 1 were compared, more significant comparisons were found.

The results from the identification task reveal an overall significant effect of tonal alignment on word-boundary identification when steps are sufficiently large, thus supporting our main hypothesis that this cue is employed by Catalan and Spanish listeners in word identification tasks. As hypothesized, the later the alignment, the higher the ratio of *late word boundary* responses, since a late peak location is a good cue to paroxytonic word types. Yet in no way does this effect

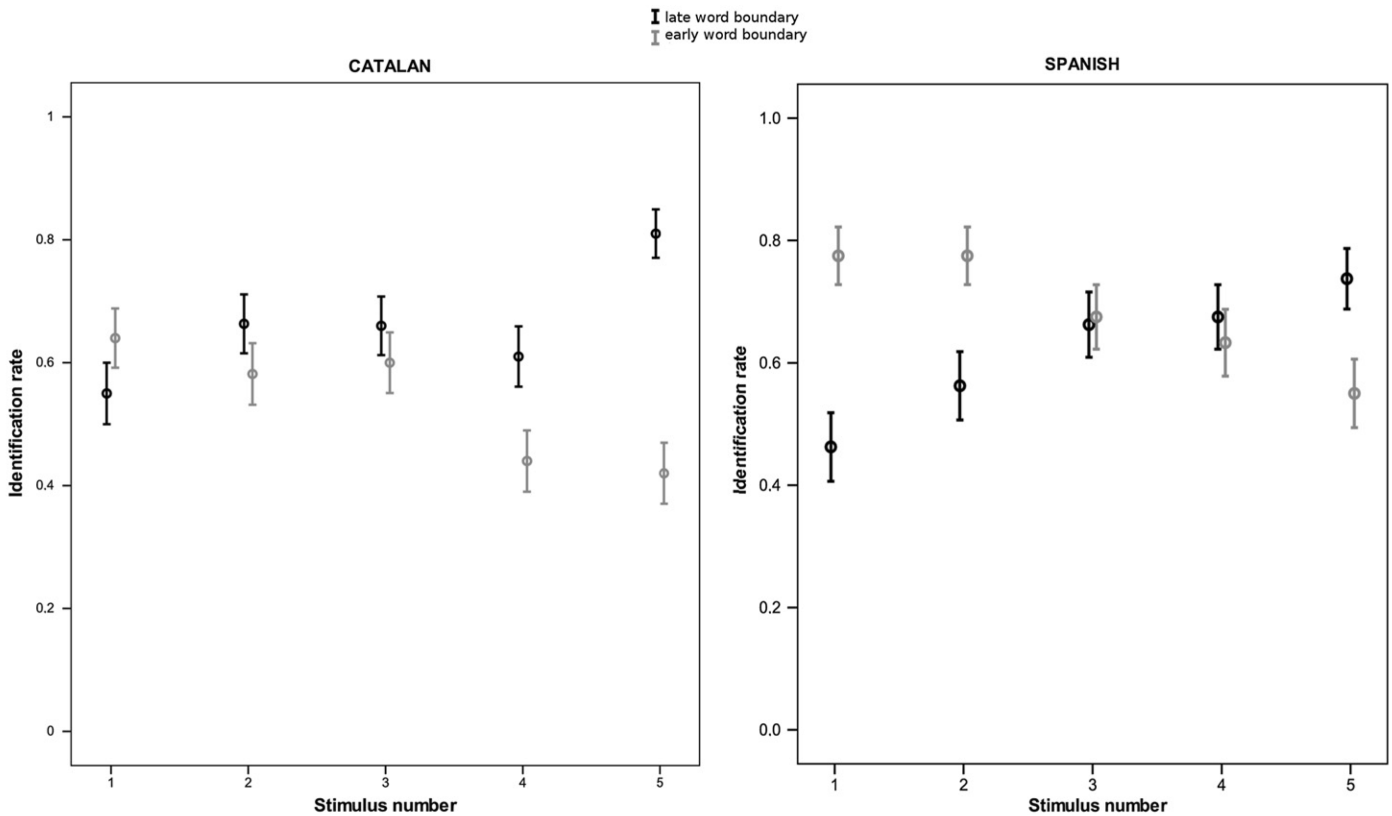

Fig. 12. Identification responses in Catalan (left panel) and in Spanish (right panel) as a function of alignment steps, for the two stimulus bases (''early word boundary'' in grey and ''late word boundary'' in black) and steps, for all listeners. The vertical bars represent standard errors.

Table 6
Results of the Friedman non-parametric test on the overall effects of H alignment steps on word-boundary identification.

	Catalan	Spanish
late word boundary base stimulus	$\chi^2(4)=29.024^*$	$\chi^2(4)=16.622^*$
early word boundary base stimulus	$\chi^2(4)=27.351^*$	$\chi^2(4)=15.373^*$

points to the classical concept in categorical perception: none of the identification functions show a sharp transition between two separate categories, but a gradual effect towards a given answer, indicating that the listener's responses were systematic in exploiting this cue for word identification.

To summarize, results from the identification task support our main hypothesis that f_0 peak alignment differences help Catalan and Spanish listeners in word identification tasks. On the other hand, it is also clear that alignment f_0 differences are not able to yield a shift in category in listeners' responses.

4. Discussion

We had good reason to expect that our production experiments would reveal substantial effects of prosodic word structure on H location, given that parallel findings had been reported for Catalan and Spanish and for other languages. Acoustic work in a variety of languages has shown that *within-word position* has a robust effect on the position of the peak of rising prenuclear accents, revealing that the details of the alignment seem to depend not only on whether the accented syllable is open or closed, but also on its place within a larger prosodic domain (see Arvaniti et al., 1998 for Greek; Chahal 2001, 2003 for Lebanese Arabic; Estebas-Vilaplana & Prieto, 2007; de la Mota, 2005; Prieto et al., 1995; Simonet, 2006; Simonet & Torreira, 2005, for Spanish; Ishihara, 2003, 2006 for Japanese; Prieto, 2006 for Catalan; Silverman & Pierrehumbert, 1990 for English). In all these languages, peaks tend to be timed earlier as their associated syllables approach the end of the word: in other words, peak delay is longer in words with antepenultimate stress than in words with penultimate stress, which in turn have longer peak delay than words with final stress. For example, in Japanese the alignment of the f_0 peak of the lexical accent in Japanese is progressively slightly earlier, relative to the accented syllable, the later the accented syllable is located in the word. When the lexical accent is on the first syllable in a CV.CV.CV.CV word, the peak is aligned with the *beginning of the vowel* in the following syllable; when the accent is on the second syllable, the peak is aligned during the onset consonant of the following syllable; when the accent is on the third syllable, the peak is aligned at the beginning of the onset consonant of the following syllable. Again, the details of the alignment seem to depend not only on the structure of the accented syllable, but also on its place in a larger prosodic domain (Ishihara, 2006). In Spanish, Prieto et al. (1995) found a significant effect of word position on peak delay in all comparisons of the words *número*, *numero*, *numeró* when the time to the next accented syllable was constant (*número rápido*, *numero nervioso*, *numeró regular*). Similarly, in Silverman and Pierrehumbert's (1990) model of f_0 peak location (their dependent variable was peak delay), the dropping of the variable word boundary (while leaving the variable stress clash as a main predictor) significantly worsened the fit of the model.

Importantly, all this evidence suggests that prosodic structure domains act as domains of articulatory organization. This view has initially been put forward in recent papers by Ladd (2006) and Prieto and Torreira (2007). Ladd (2006) asked ''[do] right-context

Table 7
Results of the Wilcoxon matched pairs signed rank test between pairs of alignment steps.

CATALAN	1–2	2–3	3–4	4–5	1–3	1–4	1–5
late word boundary	$z=-2.200^*$	$z=-.229$, ns	$z=-.962$, ns	$z=-4.082^*$	$z=-2.043^*$	$z=-1.061$, ns	$z=-4.596^*$
early word boundary	$z=-.928$, ns	$z=-.164$, ns	$z=-2.828^*$	$z=-.447$, ns	$z=-.707$, ns	$z=-4.082^*$	$z=-4.017^*$
SPANISH							
late word boundary	$z=-1.298$, ns	$z=-1.234$, ns	$z=-.192$, ns	$z=-1.043$, ns	$z=-2.530^*$	$z=-2.592^*$	$z=-3.317^*$
early word boundary	$z=.000$, ns	$z=-1.569$, ns	$z=-.577$, ns	$z=-1.183$, ns	$z=-1.569$, ns	$z=-2.117^*$	$z=-3.087^*$

effects operate at the level of the foot, the (prosodic) word, or some larger prosodic unit like the intonation phrase?" The data reported here show that at least the prosodic word plays an essential part in our understanding of the coordination of pitch gestures with the segments.[10] The data reported in this article are consistent with the hypothesis that prosodic structure not only serves to modulate the dynamics of supraglottal gestures, but also of glottal gestures. Thus this structure is necessary for understanding the coordination of f_0 gestures with supraglottal gestures.

The effect of upcoming prosodic word edges on H alignment is comparable to a certain extent to that exerted by upcoming syllable boundaries in different languages (see D'Imperio, 2000; D'Imperio, Petrone, et al., 2007 for Neapolitan Italian; Gili-Fivela and Savino, 2003 for Bari and Pisa Italian; Hellmuth, 2007 for Egyptian Arabic[11]; Ladd et al., 2000 for Dutch; Welby & Lœvenbruck, 2005, 2006 for French[12]). In a recent experiment about the effects of syllable structure on H alignment in Spanish, Prieto and Torreira (2007) found that the peak occurred around the end of the stressed vowel in the word containing open accented syllables, whereas it tended to occur within the coda consonant in the word containing a closed accented syllable. This indeed seems to suggest the possibility that the prosodic units such as the syllable and the prosodic word (and not only the presence of upcoming accents or boundary tones) influence the coordination of f_0 events with segmental material. Chahal (2001, 2003) for Lebanese Arabic found that accented syllables preceding the right edge of intonational phrase, intermediate phrase, and prosodic word boundaries display progressively earlier peak alignment. The higher the level of the boundary, the earlier the peak. Accordingly, peak alignment in Lebanese Arabic is argued to constitute a phonetic correlate for prosodic constituency in the language. Similarly, our data demonstrate that the right edge of a prosodic word domain exerts edge effects on tonal articulation, that is, it might be causing production changes that might be similar to those that occur at the edges of other prosodic domains like syllables. Yet we have to be cautious in establishing neat parallelisms, as the French results show some variation in the patterns of alignment and the Egyptian Arabic results seem to show an effect of syllable weight on peak alignment (see footnote 11).

Importantly, results from the identification tasks support the view that differences in f_0 peak alignment help Catalan and Spanish listeners in word identification tasks. These f_0 peak alignment differences can be considered as secondary acoustic cues that help listeners in the lexical identification decisions, together with cues such as consonant strengthening (see Byrd & Saltzman, 1998; Fougeron & Keating, 1997, among others), glottalization (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996), VOT (Jun, 1995), and acoustic final lengthening (e.g., Ladd & Campbell, 1991; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992).

5. Conclusion

The experimental evidence presented in this paper with minimal pair utterances such as *comprà vanguàrdies* vs. *compraven guàrdies* demonstrates that when tonal pressure effects are controlled for (in our materials there is always either one or two intervening unstressed syllables between the two accents), within-word position continues to exert consistent effects on H alignment in prenuclear peaks in Catalan and Spanish. Statistically robust effects have been found for the 10 speakers and for the two dependent variables under study, namely, the position of the peak relative to the beginning and the end of the accented vowel, thus confirming recent investigations in other languages (for Greek, Arvaniti et al., 1998; for Spanish, Estebas-Vilaplana & Prieto, 2007; de la Mota, 2005; Simonet, 2006; Simonet & Torreira, 2005; for Serbo-Croatian, Godjevac, 2000; for Japanese; Ishihara, 2006; for Catalan, Prieto, 2006; for English, Silverman & Pierrehumbert, 1990). Moreover, perceptual experiments in the two languages support the hypothesis that fine phonetic details of H tonal alignment are employed by Catalan and Spanish listeners in offline word identification tasks. Averaged classification results on the identification tasks performed by 40 listeners is summarized through a curve that shows a decrease from *late word boundary* to *early word boundary* judgments as a function of H location.

The results of our experiments clearly show that the prosodic word domain has a significant shifting effect on f_0 peak location, and, moreover, these alignment patterns are used by listeners in word-identification tasks. The empirical evidence discussed in this article demonstrates that prosodic structure should play an essential part of our understanding of the coordination of pitch gestures with the prosodic structure and advocates for a view defended by other work that prosodic structure is manifested in details of articulation.

Acknowledgments

Parts of this study were presented at the ESF International Conference on Tone and Intonation (Santorini, September 2004),

[10] Even though at this point we do not have any empirical evidence as to whether the relevant unit is the morphological word or the prosodic word, we have the hypothesis that the relevant domain of analysis is the prosodic word level. In Catalan, words followed by enclitics which constitute prosodic word units (e.g., dóna-li 'give it to him/her') seem to exert the same kind of retracting effect on peaks.

[11] As one of our reviewers points out, the case of Egyptian Arabic is a bit more complex. In Egyptian Arabic the peak is aligned with the end of the accented vowel in heavy open syllables (CVV). Yet in light open syllables (CV) the peak is found in the postaccentual consonant. Thus in Egyptian Arabic the retraction pattern that mirrors Spanish/Catalan is found only by comparing CV–CVC pairs; if CVV–CVC pairs are compared there is no retraction and the 'upcoming syllable effect' disappears. The Egyptian Arabic facts could be argued to be evidence of variability of H alignment that is potentially consistent with a gestural account.

[12] Yet as pointed out before, the French results reported by Welby and Lœvenbruck's (2005, 2006) are not as clearcut. For one of the six speakers in this study (Speaker 6), the peak was located in the coda consonant for closed syllables; this speaker aligned her peaks to the end of the vowel, regardless of syllable structure.

the 2nd Phonetics and Phonology in Iberia (Bellaterra, June 2005), the Xth Conference on Laboratory Phonology (Paris, June 2006) and at talks at the Laboratoire Parole et Langage (Aix-en-Provence, April 2005) and Institut de la Communication Parlée (Grenoble, November 2005). We are grateful to the audience at these conferences, and especially to C. de la Mota, G. Elordieta, C. Gussenhoven, J.I. Hualde, D.R. Ladd, H. Lœvenbruck, J. Kingston, C. Petrone, M. Simonet, F. Torreira, P. Welby for very useful feedback. We are grateful to friends and colleagues for their participation in the production and perception experiments. We would like to thank the Associate Editor and the three reviewers (D.R. Ladd, P. Welby, and an anonymous reviewer) for their very thorough reviews which have helped improved the final quality of the paper. Finally, we would like to thank F. Torreira and P. Welby, for their help with the Praat scripts, C. André, M. D'Imperio and C. Petrone for their support with the use of Perceval, and the SEA (Servei d'Estadística de la Universitat Autònoma de Barcelona) for their help with the statistical analysis of the data. This research has been funded by projects FFI2009-07648/FILO and CONSOLIDER-INGENIO 2010 Programme CSD2007-00012 (both awarded by the Ministerio de Ciencia e Innovación) and by project 2009 SGR 701 (awarded by the Generalitat de Catalunya).

Appendix

Materials used in Experiment 1. Note that the items are presented in pairs (a and b): (a) refers to oxytonic words (early word boundary location) and (b) paroxytonic words (late boundary location) (Table A1).

Table A1

CATALAN			
1a.	*comprà ventalls* [kumˌpɾa βən̪ˈtaʎʃ] '(s)he bought fans'	1b.	*compraven talls* [kumˌpɾaβən̪ ˈtaʎʃ] 'they bought pieces'
2a.	*comprà ventalls de vim* [kumˌpɾa βən̪ˌtaʎʒ ðə ˈβim] '(s)he bought wicker fans'	2b.	*compraven talls de vim* [kumˌpɾaβən̪ ˌtaʎʒ ðə ˈβim] 'they bought wicker pieces'
3a.	*mirà batalles* [miˌɾa βəˈtaʎəs] '(s)he looked at battles'	3b.	*mirava talles* [miˌɾa βəˈtaʎəs] '(s)he used to look at carvings'
4a.	*mirà batalles grans* [miˌɾa βəˈtaʎəz ˈɣɾans] '(s)he looked at great battles'	4b.	*mirava talles grans* [miˌɾaβə ˈtaʎəz ˈɣɾans] '(s)he used to look at great carvings'
5a.	*buscà vanguàrdies* [busˌka βəŋˈɡwaɾðjəs] '(s)he looked for newspapers'	5b.	*buscaven guàrdies* [busˌkaβəŋ ˈɡwaɾðjəs] 'they were looking for guards'
6a.	*buscà vanguàrdies a la tarda* [busˌka βəŋˈɡwaɾðjəz ə lə ˈtaɾðə] '(s)he looked for newspapers' in the afternoon'	6b.	*buscaven guàrdies a la tarda* [busˌkaβəŋ ˈɡwaɾðjəz ə lə ˈtaɾðə] 'they were looking for guards' in the afternoon'
7a.	*dibuixà vessants* [diβuˌʃa βəˈsans] '(s)he drew depressions'	7b.	*dibuixava sants* [diβuˌʃaβə ˈsans] '(s)he was drawing saints'
8a.	*dibuixà vessants de fusta* [diβuˌʃa βəˌsanz ðə ˈfustə] '(s)he drew wooden depressions'	8b.	*dibuixava sants de fusta* [diβuˌʃaβə ˌsanz ðə ˈfustə] '(s)he was drawing wooden saints'
9a.	*nomenà vescomtes* [numəˌna βəsˈkomtəs] '(s)he appointed viscounts'	9b.	*nomenaves comtes* [numəˌnaβəs ˈkomtəs] 'you appointed counts'
10a.	*nomenà vescomtes al matí* [numəˌna βəsˌkomtəz əl məˌti] '(s)he appointed viscounts in the morning'	10b.	*nomenaves comtes al matí* [numəˌnaβəs ˌkomtəz əl məˈti] 'you appointed counts in the morning'
11a.	*està badant* [əsˌta βəˈðan] '(s)he was gaping'	11b.	*estava dant* [əsˌtaβə ˈðan] '(s)he was giving'
12a.	*està badant molt* [əsˌta βəˌðam ˈmol] '(s)he was gaping a lot'	12b.	*estava dant molt* [əsˌta βəˌðam ˈmol] '(s)he was giving a lot'
13a.	comprà ventallets [kumˌpɾa βən̪təˈ ʎɛt͡s] '(s)he bought little fans'	13b.	compraven tallets [kumˌpɾaβən̪ təˈʎɛt͡s] 'they bought little pieces'
14a.	*comprà ventallets de vim* [kumˌpɾa βən̪təˌʎɛd͡z ðə ˈβim] '(s)he bought little wicker fans'	14b.	*compraven tallets de vim* [kumˌpɾaβən̪ təˌʎɛd͡z ðə ˈβim] 'they bought little wicker pieces'
15a.	*mirà batalletes* [miˌɾa βətəˈʎɛtəs] '(s)he looked at little battles'	15b.	*mirava talletes* [miˌɾa βə təˈʎɛtəs] '(s)he used to look at little carvings'
16a.	*mirà batalletes petites* [miˌɾa βətəˌʎɛtəs pəˈtitəs] '(s)he looked at little battles'	16b.	*mirava talletes petites* [miˌɾaβə təˌʎɛtəs pəˈtitəs] '(s)he used to look at little carvings'
17a.	*dibuixà vessantets* [diβuˌʃa βəsən̪ˈtɛt͡s] '(s)he drew little depressions'	17b.	*dibuixava santets* [diβuˌʃaβə sən̪ˈtɛt͡s] '(s)he drew little saints'
18a.	*dibuixà vessantets de fusta* [diβuˌʃa βəsən̪ˌtɛd͡z ðə ˈfustə] '(s)he drew little wooden depressions'	18b.	*dibuixava santets de fusta* [diβuˌʃaβə san̪ˌtɛd͡z ðə ˈfustə] '(s)he drew little wooden saints'

Table A1 (*continued*)

CATALAN			
19a.	*nomenà vescomtets* [numəˌna βəskumˈtɛt͡s] ‘(s)he appointed little viscounts’	19b.	*nomenaves comtets* [numəˌnaβəs kumˈtɛt͡s] ‘you appointed little counts’
20a.	*nomenà vescomtets al matí* [numəˌna βəskumˌtɛd͡z əl məˌti] ‘(s)he appointed little viscounts in the morning’	20b.	*nomenaves comtets al matí* [numəˌnaβəs kumˌtɛd͡z əl məˌti] ‘you appointed little counts in the morning’
SPANISH			
1a.	*ve bovinos* [ˌbe βoˈβinos] ‘(s)he sees cows’	1b.	*bebo vinos* [ˌbeβo ˈβinos] ‘I drink wines’
2a.	*ve bovinos negros* [ˌbe βoˈβinoz ˈneɣɾos] ‘(s)he sees black cows’	2b.	*bebo vinos negros* [ˌbeβo ˈβinoz ˈneɣɾos] ‘I drink red wines’
3a.	*da balazos* [ˌda βaˈlaθos] ‘(s)he shoots’	3b.	*daba lazos* [ˌdaβa ˈlaθos] ‘(s)he gave knots’
4a.	*da balazos muy fuertes* [ˌda βaˈlaθos muʝ ˈfwertes] ‘(s)he shoots very strongly’	4b.	*daba lazos muy fuertes* [ˌdaβa ˈlaθos muʝ ˈfwertes] ‘(s)he gave very strong knots’
5a.	*compraré mostazas* [kompɾaˈɾe mosˈtaθas] ‘I’ll buy mustards’	5b.	*compraremos tazas* [kompɾaˈɾemosˈtaθas] ‘we’ll buy cups’
6a.	*compraré mostazas alemanas* [kompɾaˈɾe mosˈtaθas ale manas] ‘I’ll buy German mustards’	6b.	*compraremos tazas alemanas* [kompɾaˈɾe mos ˈtaθas ale ˈmanas] ‘we’ll buy German cups’
7a.	*venderé moscaretas* [ben̪deˈɾe moskaˈɾetas] ‘I’ll sell birds’	7b.	*venderemos caretas* [ben̪deˈɾemos kaˈɾetas] ‘we’ll sell masks’
8a.	*venderé moscaretas grandes* [ben̪deˈɾe moskaˈɾetaz ˈɣɾan̪des] ‘I’ll sell big birds’	8b.	*venderemos caretas grandes* [ben̪deˈɾemos kaˈɾetaz ˈɣɾan̪des] ‘we’ll sell big masks’
9a.	*da menudos* [ˈda meˈnuðos] ‘(s)he gives offal’	9b.	*dame nudos* [ˈdame ˈnuðos] ‘give me knots’
10a.	*da menudos enormes* [ˈda meˈnuðos eˈnoɾmes] ‘(s)he gives big offal’	10b.	*dame nudos enormes* [ˈdame ˈnuðos eˈnoɾmes] ‘give me big knots’
11a.	*ve bovinitos* [ˌbe βoβiˈnitos] ‘(s)he sees cows.dim’	11b.	*bebo vinitos* [ˌbeβo βiˈnitos] ‘I drink wines.dim’
12a.	*ve bovinitos negros* [ˌbe βoβiˈnitoz ˈneɣɾos] ‘(s)he sees black cows.dim’	12b.	*bebo vinitos negros* [ˌbe βoβiˈnitoz ˈneɣɾos] ‘I drink red wines.dim’
13a.	*da balazotes* [ˌda βalaˈθotes] ‘(s)he shoots’	13b.	*daba lazotes* [ˌdaβa laˈθotes] ‘(s)he gave knots.aug’
14a.	*da balazotes muy fuertes* [ˌda βalaˈθotes muj ˈfwertes] ‘(s)he shoots very strongly’	14b.	*daba lazotes muy fuertes* [ˌdaβa laˈθotes muj ˈfwertes] ‘(s)he gave very strong knots.aug’
15a.	*compraré mostacitas* [kompɾaˈɾe mostaˈθitas] ‘I’ll buy mustards.dim’	15b.	*compraremos tacitas* [kompɾaˈɾemos taˈθitas] ‘we’ll buy cups.dim’
16a.	*compraré mostacitas alemanas* [kompɾaˈɾe mostaˈθitas aleˈmanas] ‘I’ll buy German mustards.dim’	16b.	*compraremos tacitas alemanas* [kompɾaˈɾemos taˈθitas aleˈmanas] ‘we’ll buy German cups.dim’
17a.	*venderé moscaretitas* [ben̪deˈɾe moskaɾeˈtitas] ‘I’ll sell birds.dim’	17b.	*venderemos caretitas* [ben̪deˈɾemos kaɾeˈtitas] ‘we’ll sell masks.dim’
18a.	*venderé moscaretitas grandes* [ben̪deˈɾe moskaɾeˈtitaz ˈɣɾan̪des] ‘I’ll sell big birds.dim’	18b.	*venderemos caretitas grandes* [ben̪deˈɾemos kaɾeˈtitaz ˈɣɾan̪des] ‘we’ll sell big masks.dim’
19a.	*da menuditos* [ˈda menuˈðitos] ‘(s)he gives offal.dim’	19b.	*dame nuditos* [ˈdame nuˈðitos] ‘give me knots.dim’
20a.	*da menuditos enormes* [ˈda menuˈðitos eˈnoɾmes] ‘(s)he gives big offal.dim’	20b.	*dame nuditos enormes* [ˈdame nuˈðitos eˈnoɾmes] ‘give me big knots.dim’

References

André, C., Ghio, A., Cavé, C., & Teston, B. (1995–2003). PERCEVAL: PERCeption EVALuation Auditive & Visuelle, Université d'Aix-en-Provence. Available on-line from 〈http://www.lpl.univ-aix.fr/~lpl/dev/perceval〉.

Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics*, *26*, 3–25.

Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of segmental anchoring of f_0: Evidence from German. *Journal of Phonetics*, *32*, 177–197.

Beckman, M. E., Díaz-Campos, M., McGory, J. T., & Morgan, T. A. (2002). Intonation across Spanish, in the tones and break indices framework. *Probus*, *14*, 9–36.

Beckman, M. E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 152–178). Cambridge: Cambridge University Press.

Boersma, P., & Weenink, D. (2005). Praat: Doing phonetics by computer (Version 4.3.01) [Computer program], University of Amsterdam. Available on-line from 〈http://www.praat.org/〉.

Byrd, D., & Saltzman, E. L. (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics*, *26*, 173–199.

Caspers, J., & van Heuven, V. (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica*, *50*, 161–171.

Chahal, D. (2001). *Modeling the intonation of Lebanese Arabic using the autosegmental-metrical framework: a comparison with English*. Doctoral dissertation, University of Melbourne.

Chahal, D. (2003). Phonetic cues to prominence in Arabic. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 2067–2070). Barcelona: Causal Productions.

de la Mota, C. (2005). Alignment, word boundaries and speech rate in Castilian Spanish. Paper presented at *the II Phonetics and Phonology in Iberia, PaPI 2005*, Bellaterra, June 20–21.

Dilley, L., Shattuck-Hufnagel, S., & Ostendorf, M. (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics*, *24*, 423–444.

D'Imperio, M. (2000). *The Role of perception in defining tonal targets and their alignment*. Doctoral dissertation, The Ohio State University.

D'Imperio, M., Petrone, C., & Nguyen, N. (2007). Effects of tonal alignment on lexical identification in Italian. In C. Gussenhoven, & T. Riad (Eds.), *Tones and Tunes*, Vol. 2 (pp. 79–106). The Hague: Mouton de Gruyter.

Estebas-Vilaplana, E. (2000). *The use and realisation of accentual focus in central catalan*. Doctoral dissertation, University College London. Published in 2009. Muenchen: Lincom Europa.

Estebas-Vilaplana, E. (2003). Catalan pre-nuclear accents: Evidence for word-edge tones. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 1779–1782). Barcelona: Causal Productions.

Estebas-Vilaplana, E. (2006). Word-edge tones in Spanish prenuclear accents. *Estudios de Fonética Experimental*, *15*, 11–42.

Estebas-Vilaplana, E., & Prieto, P. (2007). Production and perception of word-edge tones in Catalan and Spanish. In M. González González, E. Fernández Rei, & B. González Rei (Eds.), *Actas del III Congreso de Fonética Experimental, Santiago de Compostela* (pp. 279–290). Xunta de Galicia: Universidad de Santiago de Compostela.

Estebas-Vilaplana, E., & Prieto, P. (2008). La notación prosódica en español. Una revisión del Sp_ToBI. *Estudios de Fonética Experimental*, *17*, 263–283.

Face, T. (2002). *Intonational marking of contrastive focus in Madrid Spanish*. Muenchen: Lincom Europa.

Face, T., & Prieto, P. (2007). Rising accents in Castilian Spanish: A revision of SpToBI. *Journal of Portuguese Linguistics*, special issue on Prosody of Iberian Languages (ed. by G. Elordieta and M. Vigário) *6.1*, 117–146.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, 3728–3740.

Gili-Fivela, B., & Savino, M. (2003). Segments, syllables and tonal alignment: A study on two varieties of Italian. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 2933–2936). Barcelona: Causal Productions.

Godjevac, S. (2000). An autosegmental/metrical analysis of Serbo-Croatian intonation. *Ohio State University Working Papers in Linguistics*, *54*, 79–142.

Hellmuth, S. (2007). The foot as the domain of tonal alignment of intonational pitch accents. In *Proceedings of XVIth international congress of phonetic sciences*, Saarbrücken (pp. 969–972).

Hualde, J. I. (2002). Intonation in Spanish and the other Ibero-Romance languages: Overview and status quaestinois. In C. Wiltshire, & J. Camps (Eds.), *Romance phonology and variation* (pp. 101–115). Amsterdam: John Benjamins.

Ishihara, T. (2003). A phonological effect on tonal alignment in Tokyo Japanese. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences* (pp. 615–618). Barcelona: Causal Productions.

Ishihara, T. (2006). *Tonal alignment in Tokyo Japanese*. Doctoral dissertation, University of Edinburgh.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell, & A. Arvaniti (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 235–253). Cambridge: Cambridge University Press.

Ladd, D. R. (2006). Segmental anchoring of pitch movements: autosegmental association or gestural coordination? *Italian Journal of Linguistics*, *18*(1) (special issue on Autosegmental-metrical approaches to intonation in Europe: Tonal targets and anchors, ed. by M. D'Imperio), 19–38.

Ladd, D. R. & Campbell, N. (1991). Theories of prosodic structure: evidence from syllable duration. In: *Proceedings of the 12th international congress of phonetic sciences*, Aix-en-Provence, France (Vol. 2, pp. 290–293).

Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant 'segmental' anchoring of f_0 movements under changes in speech rate. *Journal of the Acoustical Society of America*, *106*, 1543–1554.

Ladd, D. R., Mennen, I., & Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America*, *107*, 2685–2696.

Lehiste, I. (1960). An acoustic-phonetic study of internal open juncture. *Phonetica*, *5*(Suppl.), 5–54.

Nibert, H. (2000). *Phonetic and phonological evidence for intermediate phrasing in Spanish intonation*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, *32*(6), 693–703.

Prieto, P. (2005). Stability effects in tonal clash contexts in Catalan. *Journal of Phonetics*, *33*(2), 215–242.

Prieto, P. (2006). Word-edge tones in Catalan. *Italian Journal of Linguistics*, *18*(1) (special issue on Autosegmental-metrical approaches to intonation in Europe: tonal targets and anchors, ed. by M. D'Imperio), 39–71.

Prieto, P. (2009). Tonal alignment patterns in Catalan nuclear falls. *Lingua*, *119* (special issue on The phonology and phonetics of tone from a typological perspective, ed. by F. Kügler, L. Downing, and S. Zerbian), 865–880.

Prieto, P. (in press). The intonational phonology of catalan. In Sun-Ah Jun (Ed.), *Prosodic typology*, Vol. 2. Oxford: Oxford University Press.

Prieto, P., Aguilar, L., Mascaró, I., Torres Tamarit, F. J., & Vanrell, M. M. (2009). L'etiquetatge prosòdic Cat_ToBI. *Estudios de Fonètica Experimental*, *18*, 287–309.

Prieto, P., D'Imperio, M., & Gili-Fivela, B. (2005). Pitch accent alignment in Romance: Primary and secondary associations with metrical structure. *Language and Speech*, *48*(4), 359–396.

Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, *23*, 429–451.

Prieto, P., & Torreira, F. (2007). The segmental anchoring hypothesis revisited. Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics*, *35.4*, 473–500.

Schepman, A., Ladd, D. R., & Lickley, R. (2006). Effects of vowel length and "right context" on the alignment of Dutch nuclear accents. *Journal of Phonetics*, *34*, 1–28.

Silverman, K., & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In J. Kingston, & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 72–106). Cambridge: Cambridge University Press.

Simonet, M. (2006). Word-boundary effects on pitch timing in Spanish. In N. Sagarra, & A. J. Toribio (Eds.), *Selected proceedings of the 9th hispanic linguistics symposium* (pp. 103–112). Somerville, MA: Cascadilla Proceedings Project.

Simonet, M., & Torreira, F. (2005). Alignment, word boundaries and speech rate in Castilian Spanish. Paper presented at *the II Workshop on Spanish ToBI*, Bellaterra, June 22.

Sosa, J. M. (1999). *La entonación del español: su estructura fónica, variabilidad y dialectología*. Madrid: Cátedra.

SPSS for Windows (2006). Rel. 15.0.1. 2006. Chicago: SPSS Inc.

Turk, A. E., & Shattuck-Hufnagel, S. (2000). Word-boundary-related durational patterns in English. *Journal of Phonetics*, *28*, 397–440.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, *27*, 171–206.

Welby, P., & Lœvenbruck, H. (2005). Segmental anchorage and the French late rise. In *Proceedings of the INTERSPEECH 2005*, Lisbon, Portugal, September 4–8.

Welby, P., & Lœvenbruck, H. (2006). Anchored down in anchorage: Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics*, *18(1)* (special issue on Autosegmental-metrical approaches to intonation in Europe: tonal targets and anchors, ed. by M. D'Imperio), 73–124.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*(3), 1707–1717.

Wood, S. (2005). Praat for beginners [Manual]. Available on-line from 〈http://www.ling.lu.se/persons/Sidney/praate/〉.

Xu, Y. (1998). Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica*, *55*, 179–203.