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

Pilar Prieto[+], Eva Estebas-Vilaplana[↑] & Maria del Mar Vanrell[#]

[+]Institució Catalana de Recerca i Estudis Avançats (ICREA)
and Universitat Pompeu Fabra (UPF), Spain
[↑]Universidad Nacional de Educación a Distancia, UNED
[#] Universitat Autònoma de Barcelona (UAB)
and Universitat Pompeu Fabra (UPF), Spain

Corresponding author:
Pilar Prieto
ICREA and UPF
Departament de Traducció i Ciències del Llenguatge
Campus de la Comunicació-Poblenou
Universitat Pompeu Fabra
C. Roc Boronat, 138
Room 53.710
08018 Barcelona (Catalunya)
Tel.: +34 93 225 4899
Email: pilar.prieto@upf.edu

Running title: The relevance of prosodic structure in tonal articulation

Abstract

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) was undertaken. Results revealed strong (and statistically significant) effects of within-word position on H location, in such a way that peaks tended to be timed earlier when their associated syllables occurred later in the word than when they occurred earlier in the word, confirming previous results for other languages (Silverman & Pierrehumbert 1990 for English; Arvaniti et al. 1998 for Greek; and Ishihara 2006 for Japanese; Godjevac 2000 for Serbo-Croatian) and for Spanish and Catalan (Prieto et al. 1995 for Spanish; de la Mota 2005, Simonet 2006, Simonet & Torreira 2005 for Catalan). In the light of these results, the production experiment was followed up by a set of perception experiments with the goal of determining whether listeners are able to use H alignment information for lexical access. The results of these latter experiments supported the hypothesis that Catalan and Spanish listeners are 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 empirical evidence discussed in this paper contributes to our understanding of the temporal organization of tonal gestures and their patterns of coordination with segments. The data advocates a view in which prosodic structure plays an essential part in our understanding of the coordination of pitch gestures with the segments.

Keywords: tonal alignment, effects of within-word position, Catalan intonation, Spanish intonation

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 gestures. While there is robust evidence in favor of the stability of L in rising accents, which are found to be consistently 'anchored' to the onset of the accented syllable in a variety of languages (Caspers & van Heuven 1993 for Dutch; Prieto et al. 1995 for Spanish; Arvaniti et al. 1998 for Greek; Ladd et al. 1999 and Ladd et al. 2000 for English; Xu, 1998 for Mandarin Chinese; Estebas-Vilaplana 2000 for Catalan), 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 2007b). 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 & Lœvenbruck (2005, 2006)[1] for the late rise in French, and by Prieto & 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]

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 English, see Silverman & Pierrehumbert 1990; for Spanish, Prieto et al. 1995, de la Mota 2005, Simonet 2006, Simonet & Torreira 2005; for Greek, Arvaniti et al. 1998;[3] for Japanese, Ishihara 2006; for Serbo-Croatian, see Godjevac 2000; for Lebanese Arabic, see Chahal 2001, 2003). That is, peaks tend to shift backwards as their associated syllables approach the end of the word: in other words, the distance 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 distance to the next accented syllable), Prieto, van Santen & Hirschberg (1995) analyzed a subset of the data obtained from test syllables in different positions in the word (*<u>nú</u>mero*, *nu<u>me</u>ro*, *nume<u>ró</u>*). This subset consisted of word sequences in which there was a distance of two unstressed syllables

[1] As one of the reviewers points out, Welby and Lovenbruck's (2005, 2006) results show that for one of the 6 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.

[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 2008) 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.

[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, Ladd & Mennen 1998).

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 peak alignment was found in all the comparisons (albeit stronger for one of the three speakers). The three diagrams in Figure 1 show a schematic representation of the difference in f0 timing patterns in the three conditions, namely, *número rápido*, *numero nervioso*, and *numeró regular*. Crucially, in Silverman & Pierrehumbert's (1990) model of f0 peak location, the dropping of the variable 'Word-Boundary' (while leaving the variable 'Stress Clash' as a main predictor) significantly worsened the fit of the model. This behavior suggests the possibility 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.

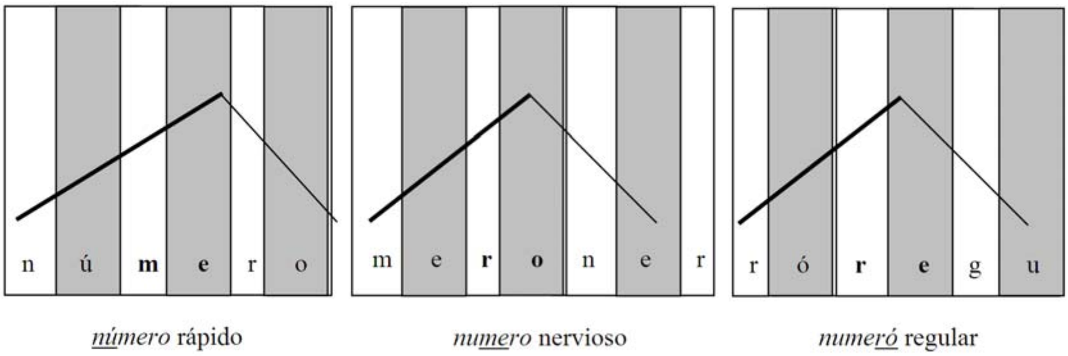

Figure 1. Schematic representation of the difference in f0 timing patterns in the three conditions, namely, *número rápido*, *numero nervioso*, and *numeró regular*.

A number of studies have found articulations of a larger magnitude in word-initial as compared to non-initial positions. For example, Fougeron and Keating (1997) found stronger lingual constrictions (i.e., with more extensive palatal contact) for consonants located at the beginning of prosodic domains (such as the phonological phrase, the intonational phrase, the utterance) and that, generally speaking, the higher the prosodic domain, the greater this strengthening. Yet even though there is well-documented evidence of the connection between prosodic structure and supraglottal gestures, very little research has been performed in the tonal domain. In this paper we test the view initially put forward in Ladd (2006) and Prieto & Torreira (2007) that prosodic domains are crucial in the coordination between f_0 gestures and units of speech production. This proposal is in line with recent studies which have demonstrated that there are a variety of prosodic influences on articulation and speech motor control, including effects of syllable structure and phrasal, emphatic, and lexical stress, and phrasing (for a review, see Byrd 1996, Byrd & Saltzman 1998, Krakow 1999).

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 prosodic structure. Specifically, we will investigate how H alignment with respect to the end of the accented syllable is affected by within-word position in Catalan and in Spanish and how it may contribute to the perceptual identification and discrimination of word boundaries in these two languages. The motivation behind choosing these two languages is that previous production experiments (see the abovementioned references) have shown consistent effects of within-word position on peak placement for both languages. First, we report the results of 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 identify more accented syllables as being followed by a word boundary when the pitch peak was timed earlier with respect to the end of the word, since this is the typical tonal feature of words with final stress. In order to test this prediction, a two-fold controlled perception experiment was carried out which included identification and discrimination tasks. A total of forty Catalan and Spanish listeners were asked to identify stimuli of the type *mirà batalles* vs. *mirava talles* that were manipulated as to H tonal alignment (in 5 steps). A total of 40 listeners (20 for each language) participated in the identification and discrimination tasks and each listened to a total of 226 stimuli.

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 set of perception experiments and discusses the results of the identification and discrimination parts. Finally, Section 4 discusses the relevance of the results of both production and perception experiments for our understanding of tonal alignment.

2. Experiment 1

The aim of Experiment 1 was to examine the effects of position-in-word of the accented syllable on H prenuclear peaks 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 2005a 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 Eastern 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. 1. 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 stable alignment, while the alignment of H is much more unstable and has prompted different phonological interpretations. Some studies have classified such rises as H* with a peak delay (Prieto, van Santen and Hirschberg 1995, and 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 f_0 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. Catalan prenuclear rises have also 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). It is possible that the consistent alignment of H at the end of words reported in these studies actually responds to the presence of a major break between the subject or the verb and the object, especially in those cases where the object was long. In utterances with shorter subject/verb and object domains, such as the ones presented in this paper, the systematic alignment of H at the end of words is not attested and therefore a word edge tone category does not describe the data properly.[4] In the present study we will adopt the L+H* representation argued for in the recently developed Cat_ToBI and Sp_ToBI proposals (Prieto, Aguilar, Mascaró & Vanrell 2009, Prieto in press) and Spanish (Face & Prieto 2007, Estebas-Vilaplana & Prieto 2008).[5] This phonological analysis enable us to capture the alignment contrasts attested in rising pitch accents for both Catalan and Spanish.

The two examples in Figure 2 illustrate a prenuclear rise with a delayed peak 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 <u>ven</u>en mandarines* [li ˌbɛnəm mən̪dəˈɾinəs] (Catalan) and *le <u>ven</u>den manda<u>ri</u>nas* [le ˌben̪den man̪daˈɾinas] (Spanish) 'They sell tangerines to him/her'. Following the proposal by Prieto, D'Imperio and Gili-Fivela (2005), the tonal transcription of the utterance would correspond to one prenuclear accent L+>H* followed by the nuclear accent !H* and by the boundary tone L%.

[4] The lack of systematic alignment between H and the end of the word was attested by inspecting the target utterances. The results in Figure 7 also show the values of distance from the H peak location relative to the right edge of the word (in ms) as a function of within-word position for both Catalan and Spanish. As the data show, no systematic alignment or distance to the end of the word is attested.

[5] For the Cat_ToBI Training Materials, see http://prosodia.uab.cat/cat_tobi/en/index.php.

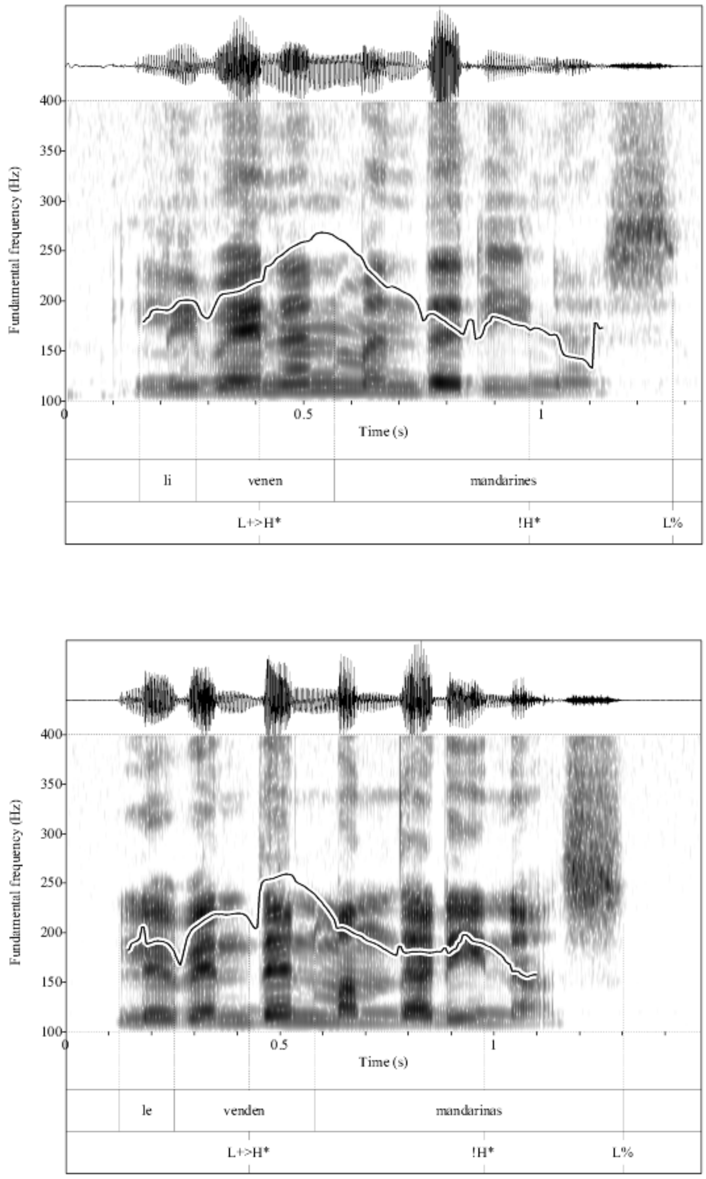

Figure 2. Waveform and f_0 contour of the broad-focus statement *li venen mandarines* (Catalan, upper panel) and *le venden mandarinas* 'They sell tangerines to him/her' (Spanish, bottom panel).

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.	*com<u>prà</u> ventalls* [kumˌpɾa βən̪ˈtaʎʃ] ‘(s)he bought fans’	b.	*com<u>pra</u>ven talls* [kumˌpɾaβən̪ ˈtaʎʃ] ‘they bought pieces’
a.	*com<u>prà</u> ventalls de vim* [kumˌpɾa βən̪ˌtaʎʒ ðə ˈβim] ‘(s)he bought wicker fans’	b.	*com<u>pra</u>ven 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.	*compra<u>ré</u> mostazas* [kompɾaˌɾemos ˈtaθas] ‘I’ll buy mustards’	b.	*compra<u>re</u>mos tazas* [kompɾaˌɾe mosˈtaθas] ‘we’ll buy cups’
a.	*compra<u>ré</u> mostazas alemanas* [kompɾaˌɾe mosˌtaθas alemˈanas] ‘I’ll buy German mustards’	b.	*compra<u>re</u>mos tazas alemanas* [kompɾaˌɾemos ˌtaθas alemˈanas] ‘we’ll buy German cups’

The phonetic transcription shows 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 between accents was controlled (i.e., there was always either one or two intervening unstressed syllables between the two accents).

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 x 5 speakers x 2 languages), for a total of 1600 utterances. The Catalan speakers were all speakers of Central Catalan and the Spanish speakers 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 Spanish speakers by the second author on professional equipment in a sound-attenuated booth at the Universitat Autònoma de Barcelona (UAB) and the Universidad Nacional de Educación a Distancia (UNED, Madrid), respectively.

Speakers were instructed to read the sentences naturally and at a normal rate of speech. Speakers were asked to read the sentences without any pause or prosodic break. 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 marked, for instance, by a continuation rise. Thus, if speakers produced a prosodic phrase break within a sentence, they were instructed to repeat that sentence at the end. Similarly, there were very few cases of contrastive focus readings, which were repeated. After each set, speakers were 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 analysed 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.

2) f_0 landmarks
- L1, valley of the first pitch accent (in Hz)
- H1, peak of the first pitch accent (in Hz)

The two graphs in Figure 3 show the *Praat* windows with the segmentation of the the Spanish utterances *compra<u>re</u>mos tazas* 'we'll buy cups' vs. *compra<u>ré</u> mostazas* 'I'll buy mustards'. 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.

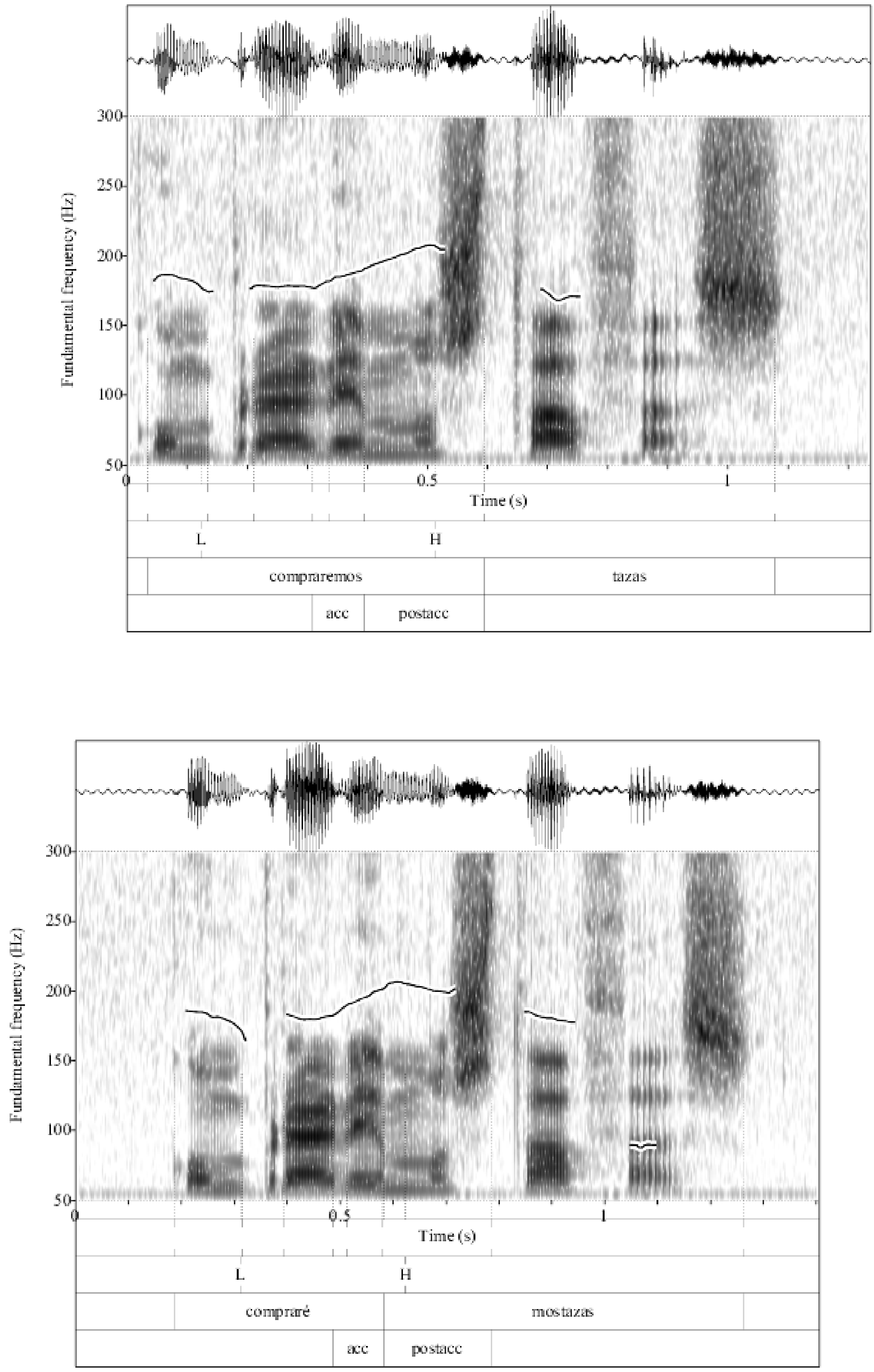

Figure 3. Displays of speech waveform, spectrogram and f_0 track, and the segmental and pitch landmarks for the Spanish utterances *compra<u>re</u>mos tazas* ‘we’ll buy cups’ vs. *compra<u>ré</u> mostazas* ‘I’ll buy mustards’.

In some cases, the identification of peaks and valleys was not a trivial task. For example, when the L or H points formed a plateau where no clear f_0 value emerged as the lowest or the highest, the endpoint in the plateau was selected (to illustrate this point, see the location of the H tone in Figure 3). Microprosodic effects (such as the typical f_0 dip produced by voiced plosives and nasal segments) were disregarded. With regards to the location of segmental boundaries separating vowels and sonorants [m, n, l, r], standard segmentation procedures using spectrograms were followed (Peterson & Lehiste 1960). The beginning or end of a sonorant consonant was identified as the start of the abrupt change from the steady-state period in the spectrogram to the onglide transition movement to the vowel (vg. the sequence *mo* in both panels in Figure 3). Similarly, intensity changes (from nasal to vowel, for example) were also used. Around 90% of the boundaries in the database are instances of sonorant consonant+vowel or vowel+sonorant consonant combinations.

After segmentation, a *Praat* script automatically collected the data points into an SPSS file (SPSS for Windows 2006) and the distance 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 postaccentual syllables.

2.1.6. Statistical analyses

Three different measures of H location were used for statistical exploration, namely (i) peak delay, or distance from the peak to the beginning of the accented syllables, (ii) HtoEndSyll, or distance from the peak to the end of the accented syllable; and (iii) HtoEndW, or distance from the peak to the end of the word.

In order to check for the general pattern of responses of the peak location variables (*peak delay*, *hdistvow*, *hdistwb)* and the duration variables *(duraccv*, *durposac*) and how they were affected by the independent variable WITHIN-WORD POSITION (*wb1)* — which had two possible values, word-final (oxytonic words) vs. word-medial syllables (paroxytonic words) — we first ran a Linear Mixed Model including subject as a random factor. These models have the following structure:

$$\mathbf{Y_i} = X_i\boldsymbol{\beta} + Z_i\mathbf{b_i} + \boldsymbol{\varepsilon_i}$$

where $\mathbf{Y_i}$ is the dependent variable (*peak delay*, *hdistvow...*), X_i is the independent variable WITHIN-WORD POSITION (*wb1)*, Z_i the matrix with the random effects and i $\boldsymbol{\varepsilon_i}$ the residuals. We also undertook subject-by-subject analyses in order to check whether the data showed any type of individual pattern of behavior or all of the subjects behaved as a homogeneous group. The level of significance was taken as $p < 0.05$.

2.2. Results

The results of the production test will be presented in three parts: 1) H delay (or temporal distance from the onset of the accented syllable to the f_0 peak) as a function of the duration of the accented syllables, 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 (Silverman & Pierrehumbert 1990 for English, Prieto, van Santen & Hirschberg 1995 for Spanish). The duration of the rise and the duration of the syllable can be correlated without being equal. Rembember that peak delay has been measured as the distance from the onset of the accented syllable to the f_0 peak, as in Silverman & Pierrehumbert (1990). Typically, the peaks under study are 'delayed' and are located in the postaccentual syllable. The two graphs in Figure 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 *comprà ventalls*, and accented syllables in penultimate position, as in *compraven talls*. First, the regression lines summarize the strong positive correlations found between the two variables (H delay and syllable duration) both in Catalan (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 gray 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.

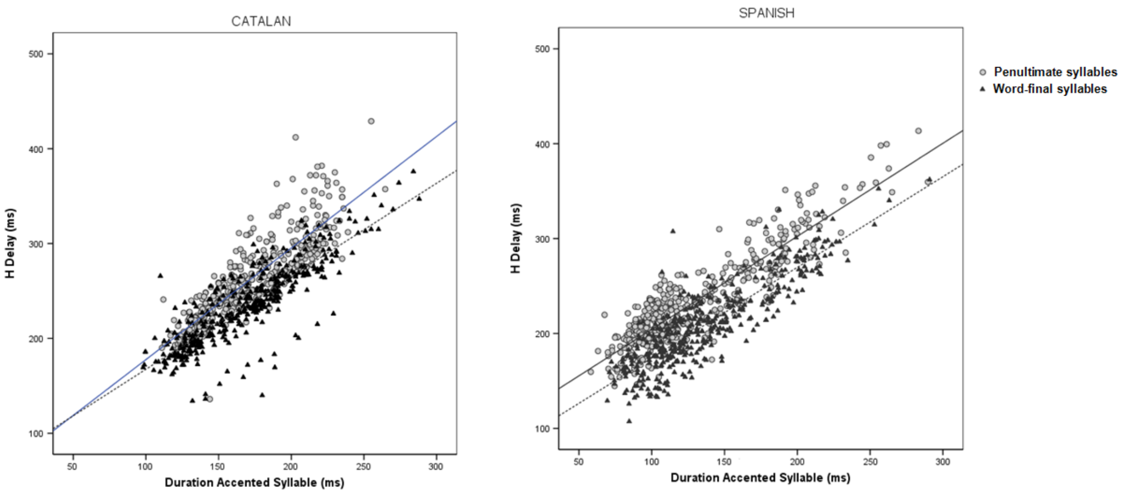

Figure 4. Peak delay (or distance from the onset of the accented syllable to the f_0 peak, in ms) as a function of the duration of the accented syllable (in ms) in the two within-

word position conditions (word-final vs. penultimate position) in Catalan (left graph) and Spanish (right graph).

As shown in Figure 4, these results reveal that in both languages there is a consistent difference in H delay depending on the prosodic condition, as the data for both Catalan and Spanish visibly cluster in two distinct groups. Hs in words with final stress are aligned earlier with respect to the end of the accented syllable than Hs in words with stress on the penultimate syllable. This suggests that the location of the H is highly influenced by within-word position.

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 et al. (2006), we decided to give more weight to a way of expressing 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). These authors demonstrate in their data that "the more distant the landmark, the greater the variance [of the distance in time between the target and the landmark], and the greater the likelihood of uninformative correlations." (Schepman et al. 2006:22)

First, the two histograms in Figure 5 plot mean peak delay measures (or mean distance of the H tonal target relative to syllable onset) as a function of within-word position (word-final syllables, as in *comp<u>rà</u> ventalls*, vs. penultimate syllables, as in *comp<u>ra</u>ven 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 ms to 45 ms, depending on the speaker.

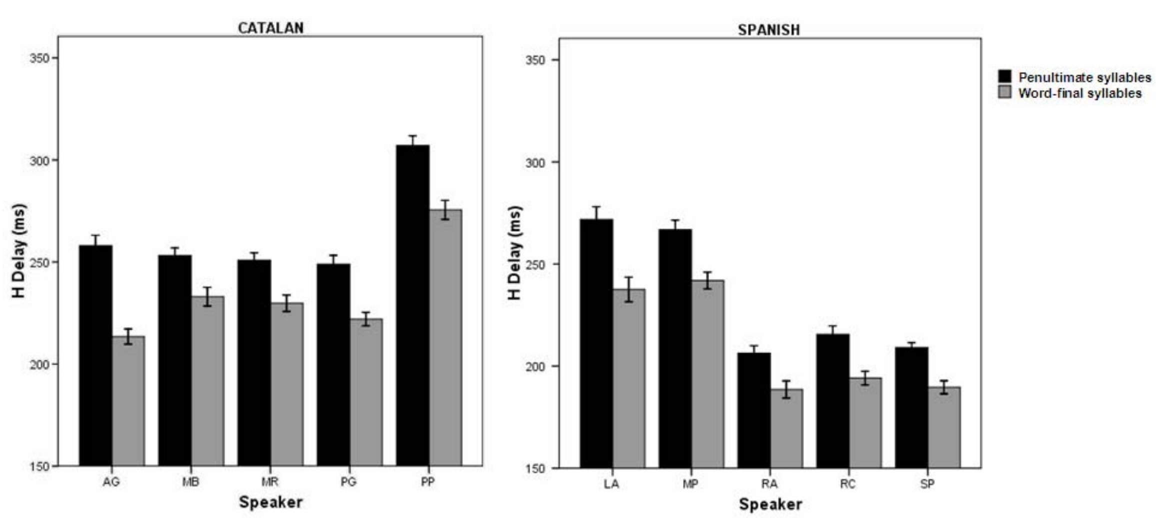

Figure 5. Mean H peak delay (or distance from the peak to the beginning of the accented syllables), as a function of within-word position for all Catalan (left panel) and Spanish speakers (right panel). Light bars show mean H peak delay in word-final

syllables and black bars show mean H peak delay in penultimate syllables. The vertical bars represent standard errors.

This pattern is replicated by the data for H distance to the end of the accented syllable. The two histograms in Figure 6 plot the mean distance 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 speakers and five Spanish speakers, peak distance 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 distance 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.[6]

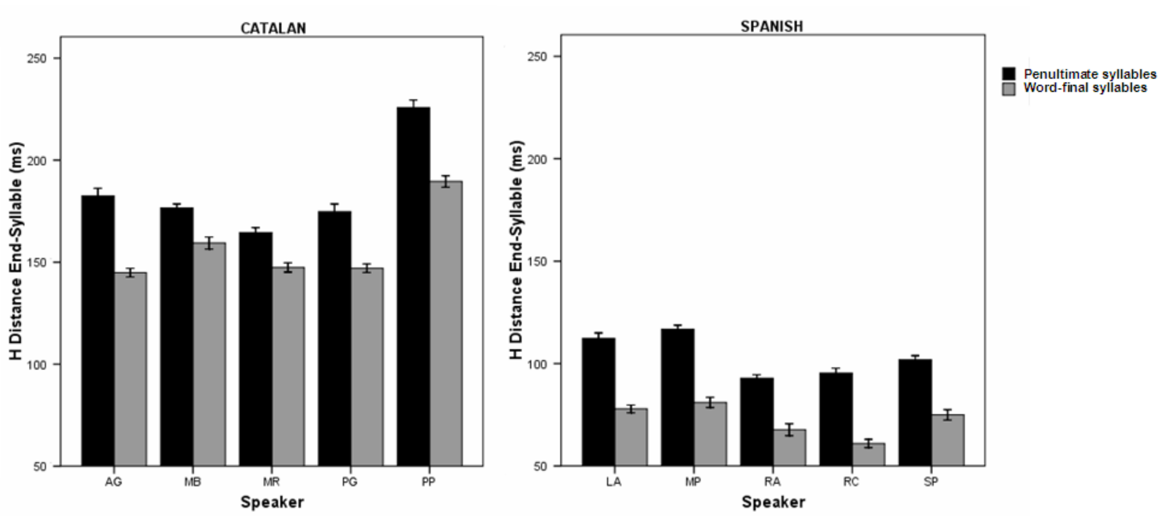

Figure 6. Mean distance from the H peak to the end of the accented syllable as a function of within-word position for 5 Catalan speakers (left panel) and for 5 Spanish speakers (right panel). Light bars show the means in the word-final condition and black bars show the mean in the word-medial condition. The vertical bars represent standard errors.

The differences observed in Figures 5 and 6 were statistically confirmed. For each variable (peak delay and peak distance to end of syllable), we ran two Linear Mixed Model analyses including subject as a random factor, one for Catalan and one for Spanish. For peak delay in the Catalan data, results indicate a main effect of Within-

[6] We note a small (and nonsignificant) 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.

Word Position (Position, at $p < 0.05$), an effect of Speaker (at $p < 0.05$) and an interaction between Within-Word Position and Speaker (at $p < 0.05$). Post-hoc analyses indicate that speaker PP is the only speaker that is significantly different from the rest; the reason is that the peak delay measures are larger for this speaker; yet crucially the effect of within-word position is the same. As for peak delay in the Spanish data, results also reveal a main effect of Within-Word Position (Position, at $p < 0.001$), an effect of Speaker (at $p < 0.001$) and no interaction between Within-Word Position and Speaker (at $p < 0.05$). Post-hoc analyses indicate that speakers LA and MP are significantly different from the rest. As for the HtoEndSyll analyses, both the Catalan and the Spanish data reveal a main effect of Within-Word Position (at $p < 0.05$ for Catalan and at $p < 0.001$ for Spanish). The variable Speaker is also significant in both languages (at $p < 0.05$), as well as the interaction between Within-Word Position and Speaker (at $p < 0.05$). Again, post-hoc analyses reveal several differences between pairs of speakers for both languages.

One-way ANOVAs were run separately for each speaker. Table 1 shows ANOVA summaries of the effects of Within-Word Position (Position) on two measures of H location, namely, Peak Delay and HtoEndSyll. Differences are statistically significant for all 10 speakers (two-tailed t-tests were significant at $p < 0.001$).

Catalan	**Peak Delay / Position** *(peak delay)*	**HtoEndSyll / Position** *(hdistvow)*
AG	$F(1,120)=15.29, p < 0.001$	$F(1,120)=31.45, p < 0.001$
MB	$F(1,120)=3.59, p = 0.0604$	$F(1,120)=10.07, p = 0.0019$
MR	$F(1, 120)=4.29, p = 0.0406$	$F(1,120)=9.11, p = 0.0031$
PG	$F(1,120)=8.40, p = 0.0045$	$F(1,120)=14.93, p = 0.0002$
PP	$F(1,120)=7.47, p = 0.0072$	$F(1,120)=24.83, p < 0.001$
Spanish		
LA	$F(1,120)=4.51, p = 0.0357$	$F(1,120)=37.96, p < 0.001$
MP	$F(1,120)=4.80, p = 0.0304$	$F(1,120)=46.39, p < 0.001$
RA	$F(1,120)=4.16, p = 0.0435$	$F(1,120)=40.08, p < 0.001$
RC	$F(1,120)=7.34, p = 0.0077$	$F(1,120)=55.72, p < 0.001$
SP	$F(1,120)=8.32, p = 0. 0046$	$F(1,120)=24.53, p < 0.001$

Table 1. One-way ANOVA summaries of the effects of Within-Word Position (Position) on two measures of H location, namely, Peak Delay and HtoEndSyll.

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 an 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 Figure 7 plot the mean values of the distance 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 no strict word-edge anchoring of the H tone is present in our data. 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.

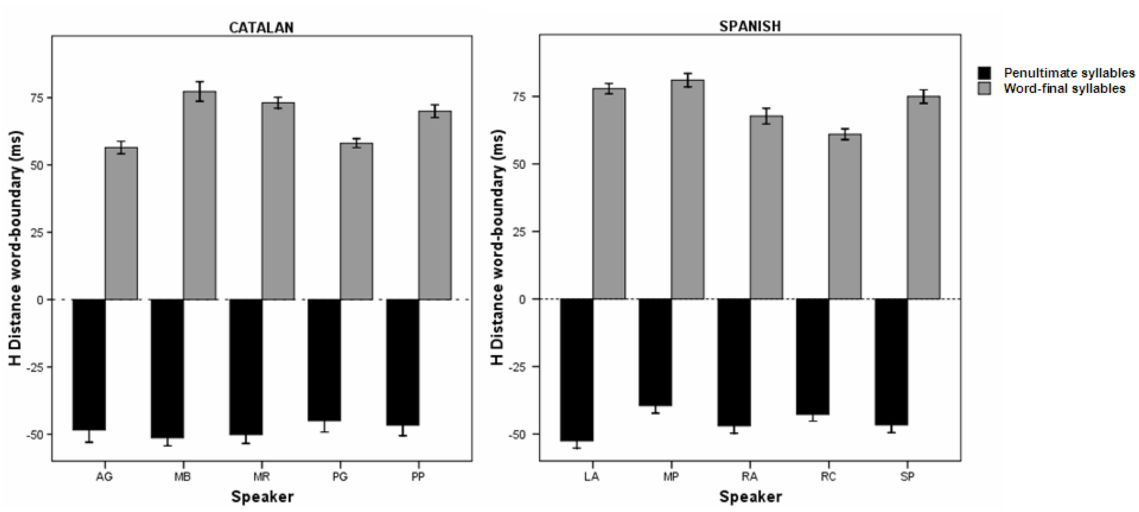

Figure 7. Mean values of distance from the H peak location relative to the right edge of the word (in ms) as a function of within-word position for the five Catalan speakers (left panel) and the five Spanish speakers (right panel). Light bars show the means in the word-final condition and black bars 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.

A linear mixed model was applied to the data including subject as a random factor, one for Catalan and one for Spanish. For both languages, results indicate a main effect of Within-Word Position on H Distance to Word Boundaries (*hdistwb*) (at $p < 0.001$), no effects of Speaker, and finally a significant interaction between Within-Word Position and Speaker (at $p < 0.001$). The interaction is probably 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 Distance to Word Boundaries (*hdistwb*), for every speaker. Effects are statistically significant for the 10 speakers (two-tailed t-tests were significant at $p < 0.001$).

Catalan	**HtoWB** *(hdistwb)*
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, $p < 0.001$

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

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 boundary leads to a later alignment of H. Figure 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 distance from the peak to the end of the accented syllable is larger in paroxytonic than in oxytonic words:

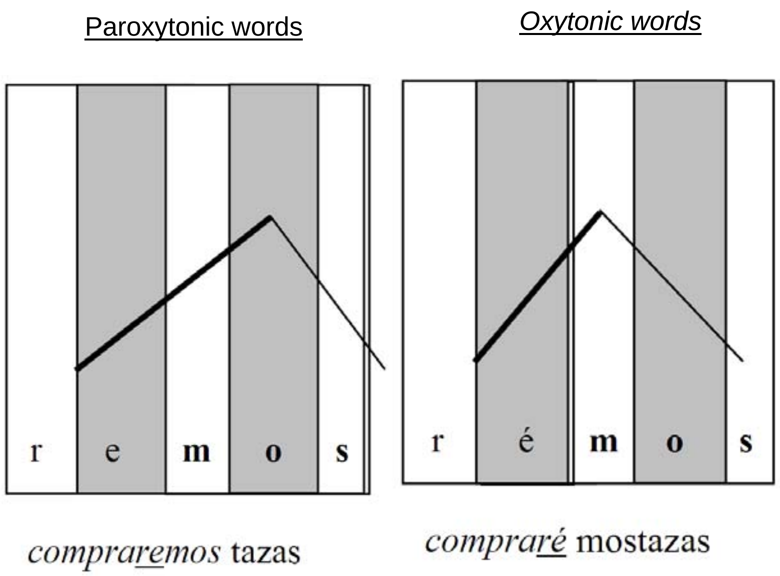

Figure 8. Schematic diagram of the f_0 peak location with respect to the end of the syllable and the end of the word in paroxytones vs. oxytones. The thick dotted lines represent the location of the word boundary.

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.[7]

2.2.3. Effects of within-word position on syllable duration

Previous studies have reported partially contradictory results on the 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 & 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, but not for word-final lengthening.

In this section, we analyse 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).[8] 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 Fougeron and Keating 1997, Byrd and Saltzman 1998, among others).

The two histograms in Figure 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 results show no clear durational effects of word edges. 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

[7] 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.

[8] Upon Alice Turk'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.

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.

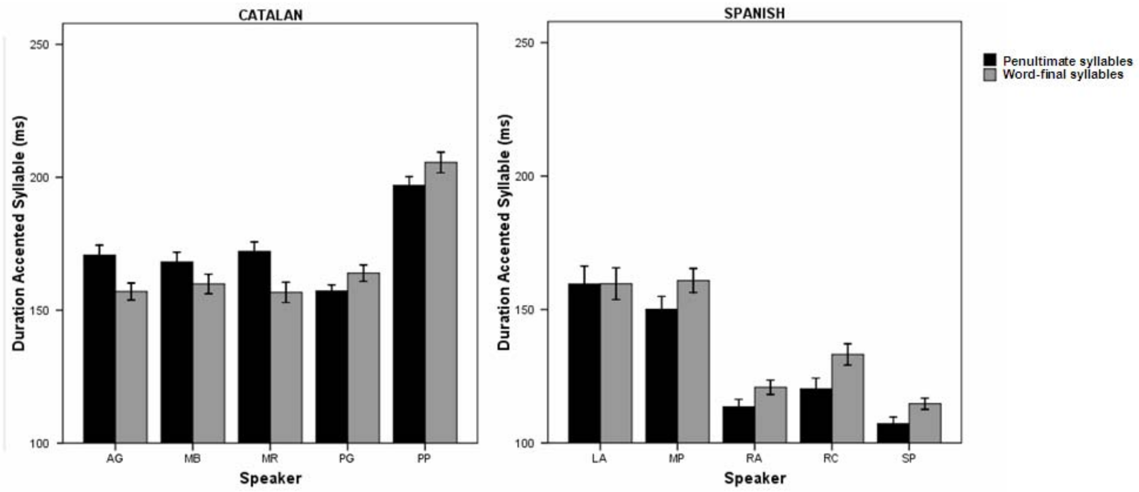

Figure 9. Mean values of duration of the accented syllable (in ms) as a function of within-word position for all Catalan (left panel) and Spanish speakers (right panel) in word-final syllables (black solid bars) and penultimate syllables (grey bars). The vertical bars represent standard errors.

The two histograms of Figure 10 show the mean values of duration of the postaccentual syllable in words with final and penultimate stress for Catalan and for Spanish. Again, the results show no clear duration effects. Postaccentual syllables in non-word-final position are slightly longer than postaccentual syllables in word-final position for some Spanish and speakers, even though there are exceptions (e.g., Catalan speakers PP and PG and Spanish speakers LA, MP, and RC). However, none of the speakers showed statistically significant differences (at $p > 0.05$) 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.

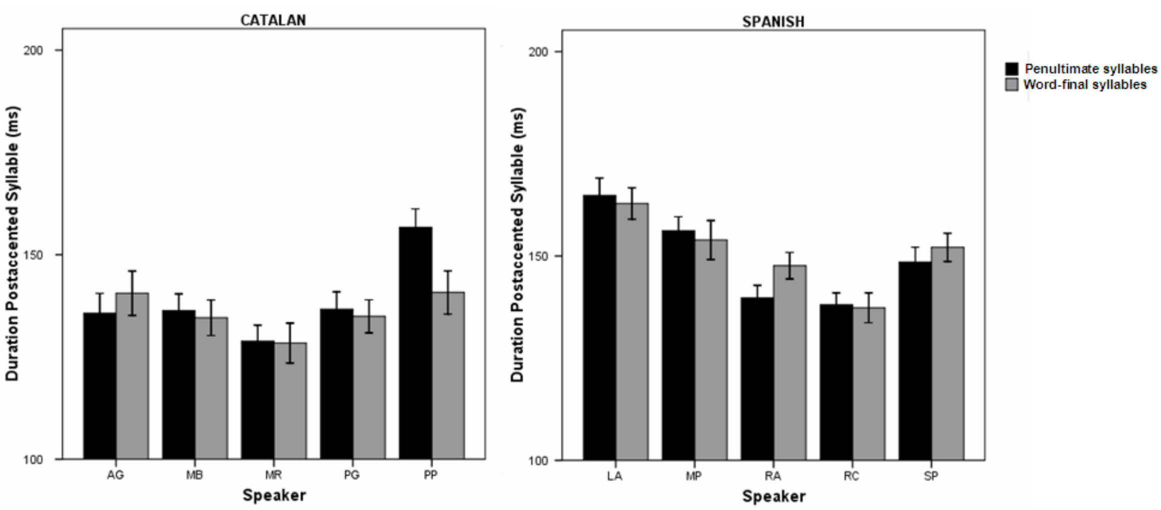

Figure 10. Mean values of duration of the postaccentual syllable (in ms) as a function of within-word position for all Catalan (left panel) and Spanish speakers (right panel) in

word-final syllables (black solid bars) and penultimate syllables (grey bars). The vertical bars represent standard errors.

For each variable (DurAcSyll and DurPostAcSyll), we ran two Linear Mixed Model analyses including subject as a random factor, one for Catalan and one for Spanish. For DurAcSyll in the Catalan data, results indicate no significant effect of Within-Word Position (Position, at $p = 0.453$), an effect of Speaker (at $p < 0.05$) and an interaction between Within-Word Position and Speaker (at $p < 0.001$). Post-hoc analyses indicate that almost all pairs of speakers are significantly different from the rest. As for DurAcSyll in the Spanish data, results reveal a significant main effect of Within-Word Position (Position, at $p < 0.05$), an effect of Speaker (at $p < 0.05$) and no interaction between Within-Word Position and Speaker (at $p = 0.148$). Post-hoc analyses also indicate that differences are statistically significant between the majority of speaker pairs. As for the DurPostAcSyll analyses, both the Catalan and the Spanish data reveal a no significant effects of Within-Word Position (at $p = 0.434$ for Catalan and at $p = 0.545$ for Spanish) and no significant interactions between Within-Word Position and Speaker (at $p = 0.589$ for Catalan and at $p = 0.227$ for Spanish). The variable Speaker is only significant in Spanish (at $p < 0.05$).

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.

Catalan	**DurAcSyll**	**DurPostAcSyll**
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	$F(1,120)=4.04$, $p = 0.0468$	$F(1,120)=0.14$, $p = 0.7084$

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

2.2.4. Linear models of Peak Placement

An attempt was made to capture the combined effects of the two factors, stress pattern (*wb*) and duration of accented syllable (*duraccsyl*) on peak placement by means of linear regression, using the delay from the peak to the end of the syllable (*hdistvow*) as the response measure. In essence, our goal in this section is to test the relevance of the wihin-word position factor in the prediction of H alignment. Table 4 shows the results of the linear mixed model. First, the main dependent variables (*wb* and *duraccsyl*) were statistically significant (at $p < 0.0001$) in both languages. In Catalan, the response variable *hdistvow* estimate takes the value of 78 ms. As can be seen in the first column, when *wb* equals 1 (that is, when the word is paroxytonic) the variable Distance-to-End-Syllable increases by 25 ms in Catalan and by 30 ms in Spanish, the mean value of *hdistvow* is 101 ms. Finally, the effects of syllable duration are contradictory: while in Catalan when the syllabic duration increases *hdistvow* also increases, in Spanish the situation is the reverse (by 0.33 percent).

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

Table 4. Results of the Linear Mixed Model that uses Distance-to-End-Syllable (*hdistvow*) as the response variable and stress pattern (*wb*) and duration of the accented syllable (*duraccsyl*) as predictive factors.

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 accenst, 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 (Silverman and Pierrehumbert 1990 for English, Prieto, van Santen and Hirschberg 1995 for Spanish, Arvaniti, Ladd and Mennen 1998 for Greek, Chahal 2001, 2003 for Lebanese Arabic).

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 alignent on word boundary identification. The perception experiments were designed following the classical Categorical Perception based paradigm, that is, by including complementary identification and discrimination 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 two tonal features are used as perceptual cues by listeners, helping them to identify the position of the word boundary, both the identification and the discrimination functions will be affected. At this point, we do not make any assumption as to what the other segmental and prosodic cues for word segmentation in Catalan and Spanish might be. We simply propose to test whether a systematic change in 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, and then using a series of identification and discrimination tasks (the classical Categorical Perception paradigm) to test the effect that these alterations produce on the listeners' word identification and discrimination 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 task and a non-significant effect in the discrimination task.

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*	*buscaven guàrdies*
Spanish	*compraré mostazas*	*compraremos 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 f0 peak location and duration of the target vowels obtained in the production test. Also, in order to avoid confounding our results with durational and f0 peak scaling cues, the test pair of stimuli had 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 perceptual 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 PSOLA (Pitch Synchronous Overlap and Add) resynthesis routine available in the Praat speech analysis and resynthesis software (Boersma and 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 was created in which the *alignment of the f_0 peak* was varied systematically in 5 equidistant steps in the continuum, 5 for each stimuli (for a total of 10 stimuli per language). Figure 11 shows the 5-step alignment continuum created from the two source utterances (v1 and g1, indicated in the graph through the use of the solid line). Note that the original stimuli, v1 and g1, were named after the actual Catalan object examples buscà vanguàrdies and buscaven guàrdies. Thus v1 corresponds to the example with a word break after the target syllalbe and g1 to the example with a word break one syllable after the target syllable. The peak manipulation was done by delaying the peak by 22 ms in the 'v' cases and by retracting it by 22 ms in the 'g' cases. 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.

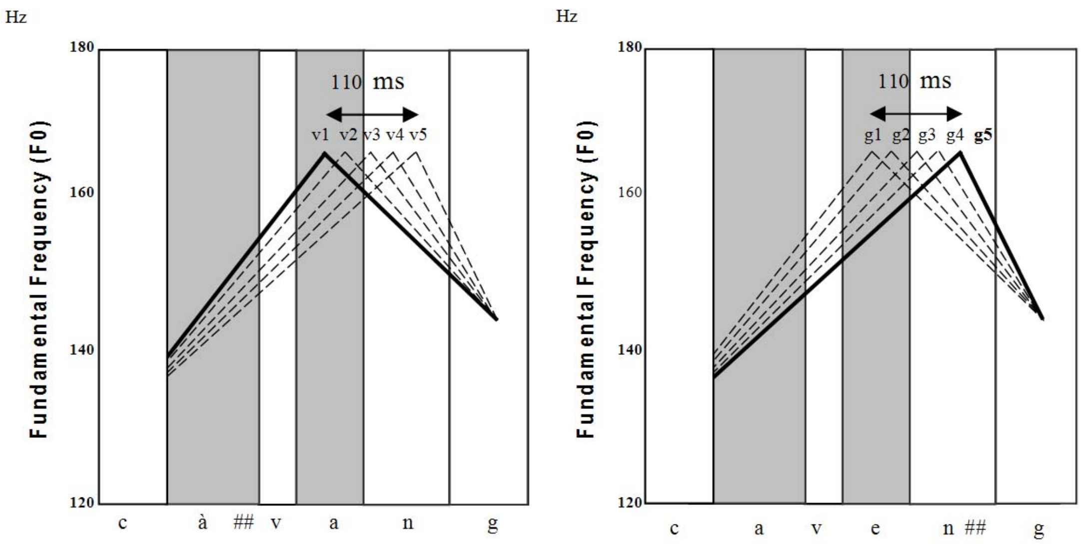

Figure 11. Five-step H alignment continuum diagrams for the *guàrdies/tazas* base stimulus (that is, g1, g2, g3, g4, g5; left panel) and for the *vanguàrdies/mostazas* base stimulus (v1, v2, v3, v4, v5; right panel). The solid line represents the original pair of utterances.

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: v1, v2, v3, v4, v5) and for the *guàrdies/tazas* base stimulus (right columns: g1, g2, g3, g4, g5). As shown by the table, the peak position in the base stimuli (v1 and g5, coded using bold letters) is different in both languages: while the peak in Catalan v1 is located 53 ms after the end of the syllable, in Spanish it is located 20 ms after; in the case of g1, 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. NB: Note that the initials used to refer to the stimuli are general throughout ('t'/'v') and are based on the Catalan example *guàrdies*/*vanguàrdies*. Thus g1, g2,… refers to the stimuli derived from the *guàrdies/tazas* base stimulus, and v1, v2… to the steps generated with the *vanguàrdies/mostazas* base stimulus.

CATALAN Steps	Base stimulus: *buscà vanguàrdies*	CATALAN Steps	Base stimulus: *buscaven guàrdies*
v1	**53 ms after end syllable**	g1	53 ms after end syllable
v2	75 ms (+22 ms)	g2	75 ms (+22 ms)
v3	97 ms (+22 ms)	g3	97 ms (+22 ms)
v4	119 ms (+22 ms)	g4	119 ms (+22 ms)
v5	141 ms (+22 ms)	**g5**	**141 ms (+22 ms)**
SPANISH Steps	Base stimulus: *compraré mostazas*	SPANISH Steps	Base stimulus: *compraremos tazas*
m1	**20 ms after end syllable**	t1	20 ms after end syllabñe
m2	42 ms (+22 ms)	t2	42 ms (+22 ms)
m3	64 ms (+22 ms)	t3	64 ms (+22 ms)
m4	88 ms (+22 ms)	t4	88 ms (+22 ms)
m5	110 ms (+22 ms)	**t5**	**110 ms (+22 ms)**

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

3.1.2. Tasks and experimental procedures

The experiment was set up by means of the PERCEVAL software developed in the Laboratoire Parole et Langage, Aix-en-Provence (André et al. 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 response frequencies and the Reaction Time (RT) were automatically recorded in PERCEVAL.

Following the classical paradigm of categorical perception (Liberman et al. 1957), one discrimination and one identification test were created for the sets of stimuli. In the first experimental session, participants judged the stimuli generated from the peak alignment continuum. Participants were seated in front of a computer and given different sets of instructions for the identification and discrimination experiments. First, they were told that they would hear a series of similar utterances either presented individually (i.e., the identification tests) or in pairs (the discrimination tests). 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 avoid order effects, the order of the sessions was counterbalanced. Half the participants started with the second session (discrimination task), and the other half heard the first session first (identification task). Subjects were allowed to rest between sessions. 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 45 minutes.

Identification tasks

For the identification tests, 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 depressing the "G" key (for the object name GUÀRDIES) or the "V" key (for **V**ANGUÀRDIES). For the Spanish version, participants indicated their choice by depressing either the "T" key (for **T**AZAS) or the "M" key (for **M**OSTAZAS).

Listeners performed the identification task in which the 10 stimulus group with each of the alignment steps was played 5 times in 5 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.

Discrimination tasks

In the discrimination tests, listeners had to decide if the two stimuli they heard were identical or dissimilar. Listeners were told to respond according to whether they heard two similar utterances (in which case they were to press the "I" key, for Cat. IGUALS-Span. IGUALES 'equal') or two different sentences (in which case they were to press the "D" key, for Cat. DIFERENTS/ Span. DIFERENTES 'different'). We used a 300 milliseconds interval between the stimuli.

Each stimulus in the continuum was paired with either itself or a dissimilar stimulus, in either an "early-late" or "late early" order. Three or four steps separated the dissimilar stimulus pairs, as Table 6 shows. Each stimulus set was presented in random order. For this task, the 18 pairs of stimuli were played 5 times in 5 different blocks, for a total of 90 stimuli.

CATALAN	*guàrdies* base stimulus	*vanguàrdies* base stimulus	*Total pairs stimuli*

	1-1, 3-3, 5-5, 1-4, 4-1, 2-5, 5-2, 1-5, 5-1	1-1, 3-3, 5-5, 1-4, 4-1, 2-5, 5-2, 1-5, 5-1	18 stimuli x 5 repetitions=90 stimuli
SPANISH	*tazas* base stimulus	*mostazas* base stimulus	
	1-1, 3-3, 5-5, 1-4, 4-1, 2-5, 5-2, 1-5, 5-1	1-1, 3-3, 5-5, 1-4, 4-1, 2-5, 5-2, 1-5, 5-1	18 stimuli x 5 repetitions=90 stimuli

Table 6. Pairs of stimuli for the discrimination task.

Altogether, each listener heard a total of 140 trials in the experiment (6 training items + 50 identification trials + 90 discrimination trials).

3.1.3. Participants

20 native speakers of Central Catalan and 20 native speakers of Peninsular Spanish participated in the experiment. The Catalan participants were first and second year undergraduates from the Catalan Philology programme 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.

For the discrimination results, two main analyses were performed using a Wilcoxon signed-rank test. First, dissimilar stimulus pairs (that is, those pairs separated by 2, 3, or 4 steps) were compared with identical pairs in order to see whether the former yielded more ‘different’ judgments than the latter. A second Wilcoxon signed-rank test was used to uncover a potential order-of-presentation effect, that is, to determine whether the presentation of the stimuli in “early-late” or “late early” order affected the ratio of “different” judgments.

3.2. Results

3.2.1. Effects of H alignment: Identification test

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 Figure 12. The two graphs in this figure plot the identification rate for the two continua created from the "vanguàrdies/mostazas" (in black) and "guàrdies/tazas" (in grey) base stimuli. The identification rate is defined as the number of "vanguàrdies/mostazas" responses (in vanguàrdies/mostazas-based stimuli) or "guàrdies/tazas" responses (in guàrdies/tazas-based stimuli) over the total responses. The x-axis represents the five steps of H alignment manipulations, either from the *guàrdies/tazas* base stimulus (t1, t2, t3, t4, t5) or the *vanguàrdies/mostazas* base stimulus (v1, v2, v3, v4, v5). Each data point represents a total of 100 judgments (20 listeners x 5 repetitions).

Results from both languages show a tendency towards an S-shaped curve, that is, in all panels we see a shift from *guàrdies* to *vanguàrdies* (and from *tazas* to *mostazas*) judgments as a function of H alignment: at late locations of H within the continuum (g5, v5) *guàrdies/tazas* responses are more dominant, while at the opposite end of the continuum (g1, v1) *vanguàrdies/mostazas* responses are more dominant within a group. Second, though an effect can be seen in both languages, no categorical boundary shift due to the H alignment manipulation is present, as ratio differences range from 0,20 to 0,40 at the two ends of the curve. Third, the identification curves are different depending on the base stimulus, meaning that there are 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.

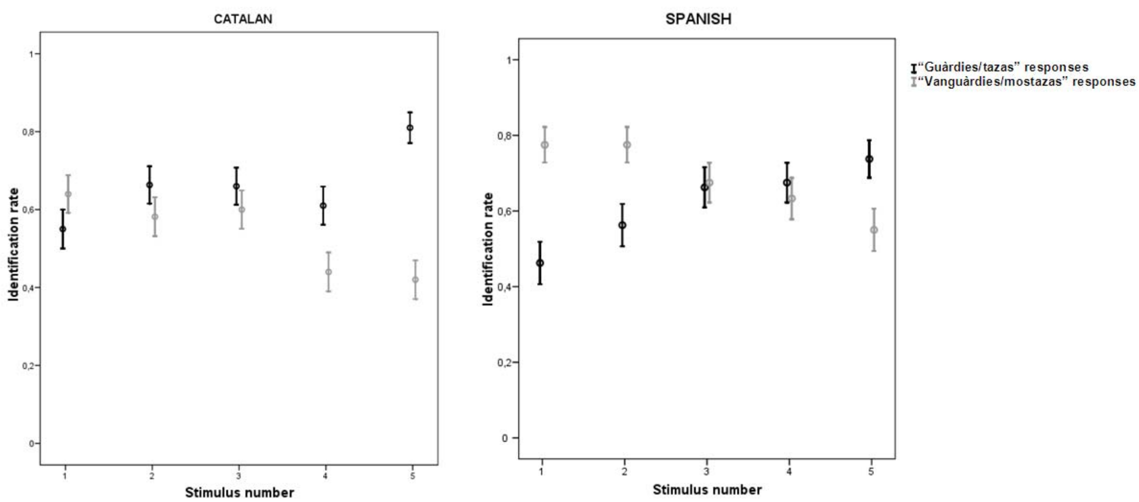

Figure 12. Identification curves of the alignment stimuli in Catalan (left panel) and in Spanish (right panel). The vertical bars represent standard errors.

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 7 shows the results. As expected, the main effect of H alignment is significant both for Spanish and Catalan (both for the *guàrdies/tazas* and the *vanguàrdies/mostazas* stimuli set) —the significance level was set to 0.05, asterisks mark significant effects.

	Catalan	Spanish
guàrdies/tazas base stimulus	$\chi^2(4)$= 29.024*	$\chi^2(4)$= 16.622*
vanguàrdies/*mostazas* base stimulus	$\chi^2(4)$= 27.351*	$\chi^2(4)$= 15.373*

Table 7. Results of the Friedman non-parametric test on the overall effects of H alignment steps on word-boundary identification.

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 8, 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.

CATALAN	1-2	2-3	3-4	4-5	1-3	1-4	1-5
guàrdies	z = -2.200*	z = -.229, ns	z = -.962, ns	z = -4.082*	z = -2.043*	z = -1.061, ns	z = -4.596*
vanguàrdies	z = -.928, ns	z = -.164, ns	z = -2.828*	z = -.447, ns	z = -.707, ns	z = -4.082*	z = -4.017*

SPANISH	1-2	2-3	3-4	4-5	1-3	1-4	1-5
tazas	z = -1.298, ns	z = -1.234, ns	z = -.192, ns	z = -1.043, ns	z = -2.530*	z = -2.592*	z = -3.317*
mostazas	z = .000, ns	z = -1.569, ns	z = -.577, ns	z = -1.183, ns	z = -1.569, ns	z = -2.117*	z = -3.087*

Table 8. Results of the Wilcoxon matched pairs signed rank test between pairs of alignment steps.

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 *guàrdies/tazas* responses, since a late peak location is a good cue to paroxytonic word types. Yet in no way does this effect 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.

3.2.2. Effects of H alignment: Discrimination test

Figures 13 and 14 illustrate the results of the discrimination experiments as the ratio of 'different' responses for each pair pooled over the 20 Catalan listeners (upper graphs) and 20 Spanish listeners (bottom graphs). Figure 13 shows the rate of false alarms (number of "different" responses to identical pairs) and "early-late" and "late-early" hits ("different" responses to different pairs) separated by 3 steps corresponding to the continuum created from the “guàrdies/tazas” base stimulus (left) and “vanguàrdies/mostazas” base stimulus (right). Figure 14 shows the rate of “different” responses to the “early-late” and “late-early” hits separated by 4 steps corresponding to the continuum created from the “guàrdies/tazas” base stimulus (left) and “vanguàrdies/mostazas” base stimulus (right). It is clear from the graphs that there is a tendency for the series dissimilar pairs (that is, pairs separated by 3 or 4 steps) to yield more ‘different’ judgments than identical ones. In sum, pairings of identical stimuli were less frequently judged to be different than those of dissimilar ones, in both series and for all languages. Yet, in contrast with typical categorical perception graphs, no clear discrimination maximum was found.

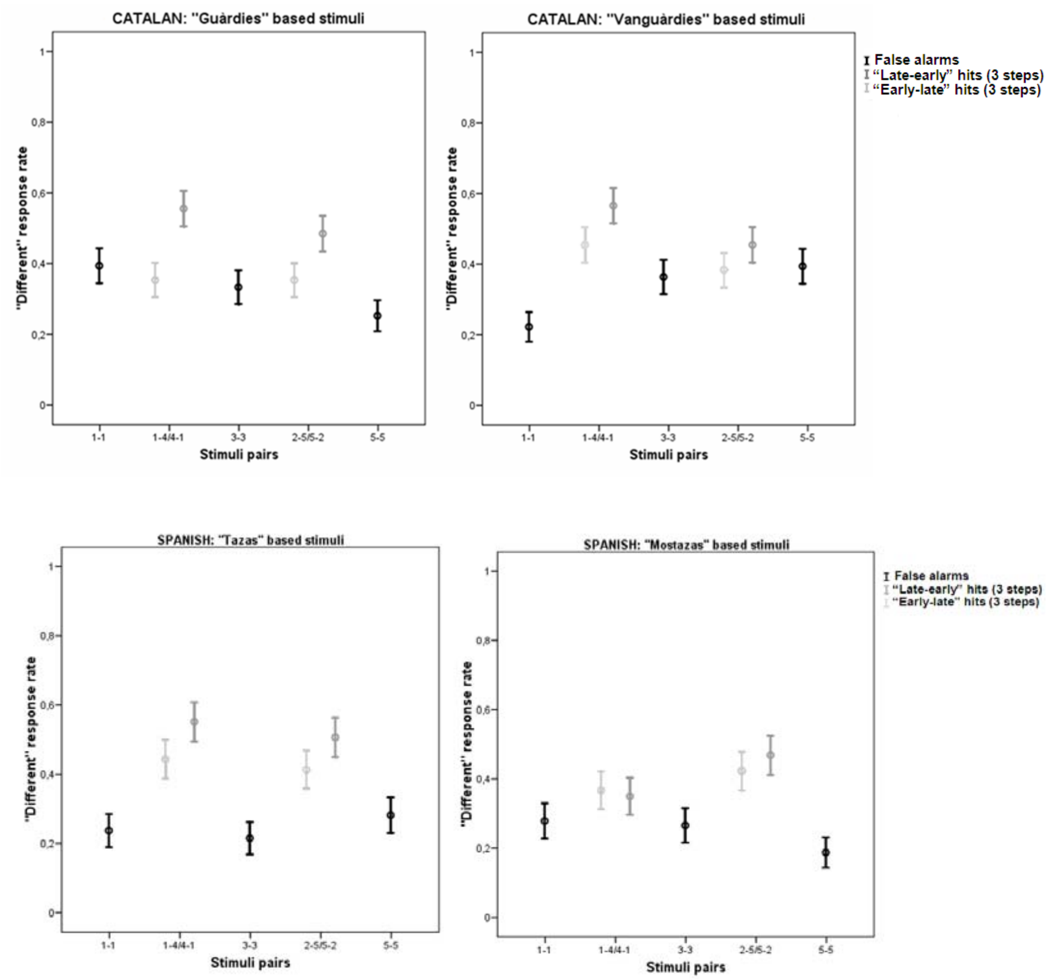

Figure 13. Discrimination functions for the false alarm and “early-late” and “late-early” hits separated by 3 steps. The right panels show the results with the *guàrdies/tazas* base stimulus set and the left panels the results with the *vanguàrdies/mostazas* base stimulus set. The vertical bars represent standard errors.

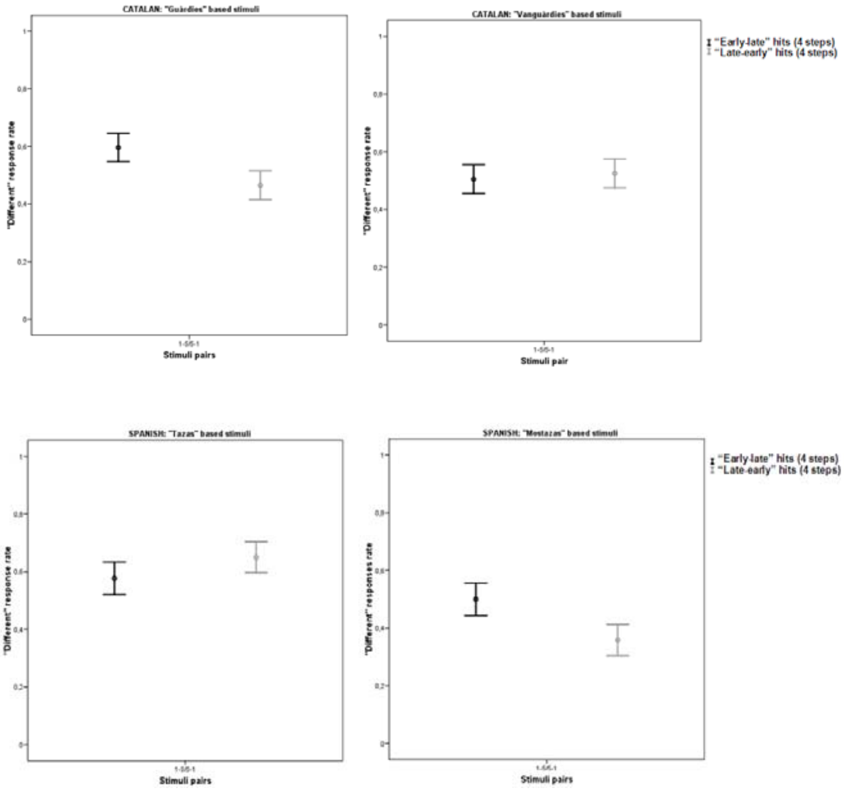

Figure 14. Discrimination functions for the “early-late” and “late-early” hits (separated by 4 steps) in Catalan (upper graphs) and in Spanish (bottom graphs). The right panels show the results with the *guàrdies*/*tazas* base stimulus set and the left panels the results with the *vanguàrdies*/*mostazas* base stimulus set. The vertical bars represent standard errors.

The next statistical test was intended to examine whether there were any significant differences in the frequency of ‘different’ judgments between identical (0 steps, false alarms) and dissimilar (3 or 4 steps, hits) stimulus pairs in the alignment continuum. This test was a Wilcoxon matched pairs signed rank test. As can be seen in Table 9, all the comparisons are significant in Catalan (cf. *guàrdies* vs. *vanguàrdies* stimuli) and some of them in Spanish (cf. *tazas* vs. *mostazas* stimuli; see asterisks). The analysis shows that both in Catalan and Spanish in all cases the dissimilar pairs (3 or 4 steps) were judged as ‘different’ significantly more often (p < 0.05) than the identical pairs (0 steps). Finally, even though the order-of-presentation effect is significant only in Catalan (see 4 steps “early-late” vs 4 steps “late-early”), this same tendency can be found in both languages.

	FA – “Late-early” Hits (3 steps)	“Late-early” H (3 steps) – “Early-late” H (3 steps)	“Early-late H” (3 steps) – “Late-early” H (4 steps)	“Late-early” H (4 steps) – “Early-late” H (4 steps)	FA – “Early-late” H (4 steps)	FA – “Late-early” H (4 steps)	FA – “Early-late” H (4 steps)
guàrdies	z = -6.708*	z = -5.745*	z = -4.899*	z = -3.606*	z = -3.464*	z = -5.831*	z = -4.583*
tazas	z = -2.224*	z = -.236, ns	z = -1.897, ns	z = -1.134, ns	z = -.250, ns	z = -1.461, ns	z = -2.475*

	FA – “Early-late” Hits	“Early-late” H (3steps) –	“Late-early” H (3 steps) –	“Early-late” H (4 steps)-	FA – “Late-early” H (4	FA – “Early-late”	FA – “Late-early” H (4

	(3 steps)	"Late-early" H (3 steps)	"Early-late" H (4 steps)	"Late-early" H (4 steps)	steps)	H (4 steps)	steps)
vanguàrdies	z = -5.000*	z = -4.243*	z = -2.449*	z = -1.414*	z = -6.557*	z = -5.292*	z = -5.477*
mostazas	z = -4.610*	z = -1.723, ns	z = -2.832*	z = -.898, ns	z = -3.263*	z = -4.041*	z = -3.286*

Table 9. Results of the Wilcoxon matched pairs signed rank test between pairs of discrimination steps.

To summarize, results from the identification and the discrimination tasks support our main hypothesis that f_0 peak alignment differences help Catalan and Spanish listeners in word identification tasks. Yet alignment differences do not trigger categorical perception effects, as the identification graphs show very low slope values and a discrimination peak is indeed absent from the discrimination function.

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 the structure of the accented syllable, but also on its place within a larger prosodic domain (see Silverman and Pierrehumbert 1990 for English, Prieto, van Santen and Hirschberg 1995, de la Mota 2005, Estebas-Vilaplana and Prieto 2007, Simonet and Torreira 2005, Simonet 2006 for Spanish, Prieto 2006 for Catalan, Arvaniti, Ladd and Mennen 1998 for Greek, and Ishihara 2003, 2006 for Japanese, Chahal 2001, 2003 for Lebanese Arabic). 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 the structure of the accented syllable, but its place in a larger prosodic domain (Ishihara 2006). In Spanish, Prieto, van Santen and Hirschberg (1995) found a significant effect of word position on peak delay in all comparisons of the words *número*, *numero*, *numeró* when the distance 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, 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 & Torreira (2007). Ladd (2006) asked "[do] right-context 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. 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 Ladd et al. 2000 for Dutch, D'Imperio 2000 and D'Imperio, Petrone and Nguyen 2007b for Neapolitan Italian, Gili-Fivela and Savino 2003 for Bari and Pisa Italian, Welby and Lœvenbruck 2005, 2006 for French[9], and Hellmuth 2007 for Egyptian Arabic[10]). 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) are acting as domains of intergestural coordination. 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 causes changes in articulatory gestural dynamics 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 parallellisms, 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 6).

Recent research in speech production has provided clear evidence that segmental articulation cannot be understood independently of prosodic structure. Several studies have shown that the syllable domain and position in the word influence the durational and spatial domains of a gesture's articulation (for a review, see Byrd 1996 and Krakow 1999, Byrd et al. 2000). A number of studies have found articulations of a larger magnitude in word-initial as compared to non-initial positions. For example, Fougeron and Keating (1997) reported that lingual constrictions for initial consonants are strengthened (have more extensive palatal contact) at the beginning of prosodic domains (phonological phrase, intonational phrase, utterance) and that, generally speaking, the

---

[9] Yet as pointed out before, the French results reported by Welby and Lovenbruck's (2005, 2006) are not as clearcut. For one of the 6 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.

[10] 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.

higher the prosodic domain, the greater the strengthening. Coordination of gestures at the ends of domains is expected to be off-phase. Byrd and Saltzman (1998) have proposed a theory of prosodic gestures that attempts to phase together segmental and suprasegmental structures in terms of their temporal and coordination properties within the task dynamics framework. They conceive boundary-adjacent lengthening as a local slowing of the gestures in the immediate vicinity of sufficiently strong prosodic boundaries at multiple levels. In their study, all participants appeared to use a diminished gestural stiffness as the main source of lengthening before prosodic boundaries. Yet even though there is well-documented evidence of this in supraglottal gestures, very little research has been done in the tonal domain.

The current formulation of the task dynamics model of speech production originally proposed by Browman and Goldstein (1986) and developed over the last two decades (Browman and Goldstein 1988, 1995a, 1995b, 2000, Byrd 1996, Byrd et al. 2000, Goldstein, Byrd and Saltzman 2006, Goldstein and Fowler 2003 among many others) provides explicit mechanisms to account for speech timing and gestural coordination patterns. Thus it offers the tools to adequately account for the alignment patterns reported by the literature on tonal alignment. Following the proposal by Byrd et al. (2000), several gestural phenomena will occur at prosodic domain edges, that is, their activity is governed by prosodic constituency. They refer to such prosodic boundary units as "π-gestures". As they point out, a given "π-gesture" affects independently and directly the following gestural attributes: “(1) the values of gestural parameters such as stiffness or target position for all tract-variable gestures with which it is concurrently active; or (2) the clock rate (i.e., local speaking rate) such that a stronger "π-gesture" yields more slowing of the clock rate than a weaker one.” (Byrd et al. 2000: 84).

In this paper we hypothesize that there is crosslinguistic evidence that the prosodic word is exerting right edge effects. Extrapolating our results, we claim that general properties of intergestural timing could provide a unifying explanation for (i) the contrasting behavior between the precise synchronization of L valleys with the onset of the syllable and the more variable timing of H peaks, and, more specifically, (ii) the right-hand tonal pressure effects and ‘undershoot’ patterns displayed by peaks at the ends of syllables and other prosodic domains.

The data reported in this article also has some implications for the so-called ‘segmental anchoring hypothesis’ (Ladd et al. 1999, Arvaniti, Ladd, & Mennen 1998, Ladd 2004, 2006, Xu 1998). As it is well-known, recent work on tonal alignment has emphasized that the temporal location of tonal targets relative to the segmental string are governed by principles of synchrony and stability. A general version of the segmental anchoring hypothesis claims that the beginning and end of a pitch gesture are anchored to specific landmarks in the segmental structure when no tonal pressure is exerted on them. Our data show that when right-hand tonal effects are controlled the right edge of a prosodic word domain exerts edge effects on tonal alignment, and thus the segmental anchoring hypothesis cannot account for those word-edge effects.

Importantly, results from the identification and the discrimination tasks support the view that such 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 Fougeron and Keating 1997, Byrd and Saltzman

1998, among others), glottalization (Dilley, Shattuck-Hufnagel and Ostendorf, 1996), VOT (Jun, 1995), and acoustic final lengthening (e.g., Wightman, Shattuck-Hufnagel, Ostendorf and Price, 1992; Ladd and Campbell, 1991).

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 distance from the peak relative to the beginning and the end of the accented vowel, thus confirming recent investigations in other languages (for English, Silverman and Pierrehumbert 1990; for Greek, Arvaniti et al. 1998; for Japanese, Ishihara 2006; for Serbo-Croatian, Serbian-Croatian see Godjevac 2000; for Catalan, Prieto and Estebas-Vilaplana 2005, Prieto 2006; for Spanish, Estebas-Vilaplana and Prieto 2007, de la Mota 2005, Simonet and Torreira 2005, Simonet 2006). 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 *guàrdies* to *vanguàrdies* 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 segments 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), the *2nd Phonetics and Phonology in the 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 editor of this paper, Alice Turk, and the three reviewers (D.R. Ladd, P. Welby, and an anonymous reviewer) for their very thorough reviews. 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 grant 2005SGR-00753, awarded by the Generalitat de Catalunya, and by grants HUM2006-01758/FILO "Estructura prosódica y adquisición de la prosodia en catalán y español" and CONSOLIDER-INGENIO 2010 "Bilingüismo y Neurociencia Cognitiva CSD2007-00012" awarded by the Spanish Ministerio de Educación y Ciencia.

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 accross Spanish, in the Tones and Break Indeces 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.

Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in English and Japanese. *Phonology Yearbook*, *3*, 255-310.

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/

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219-252.

Browman, C. P., & Goldstein, L. M. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*, 140-155.

Browman, C. P., & Goldstein, L. M. (1995a). Dynamics and articulatory phonology. In R. Port & T. van Gelder (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition* (pp. 175-193). Cambridge, MA: MIT Press.

Browman, C. P., & Goldstein, L. M. (1995b). Gestural syllable position effects in American English. In F. Bell-Berti & R. Lawrence (Eds.), *Producing Speech: Contemporary Issues* (pp. 19-33). New York: American Institute of Physics.

Browman, C. P., & Goldstein, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Bulletin de la Communication Parlée*, *5*, 25-34.

Byrd, D. (1996). A phase window framework for articulatory timing. *Phonology*, *13*, 139-169.

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

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. B. Broe, & J. B. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 70-87). Cambridge: Cambridge University Press.

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 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. (2002). Language-specific and universal constraints on tonal alignment: the nature of targets and "anchors". *Proceedings of the 1st International Conference on Speech Prosody* (pp. 101-106).

D'Imperio, M., Espesser, R., Lœvenbruck, H., Menezes, C., Nguyen, N., & Welby, P. (2007). Are tones aligned with articulatory events? Evidence from Italian and French. In J. Cole & J. Hualde (Eds.), *Papers in Laboratory Phonology 9: Change in Phonology* (pp. 577-608). The Hague: Mouton de Gruyter.

D'Imperio, M., Petrone, C., & Nguyen, N. (2007b). 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.

Goldstein, L. M., & Fowler, C. (2003). Articulatory phonology: a phonology for public language use. In A. S. Meyer & N. O. Schiller (Eds.), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* (pp. 159-207). New York: Mouton de Gruyter.

Goldstein, L., Byrd, D., & Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. Arbib (Ed.), *From action to language: the mirror neuron system* (pp. 215-249). Cambridge, UK: Cambridge University Press.

Hellmuth, S. (2007). The foot as the domain of tonal alignment of intonational pitch accents. In *Proceedings of XVIth International Congress of Phonetic Sciences* (pp. 969-972), Saarbrücken.

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.

Krakow, R. A. (1999). Physiological organization of syllables: a review. *Journal of Phonetics*, *27*, 23-54.

Ladd, D. R. (2004). Segmental anchoring of pitch movements: autosegmental phonology or speech production? In H. Quene, & V. J. van Heuven (Eds.), *Speech and Language: Studies for Sieb G. Noteboom* (pp. 123-131). Utrecht: LOT.

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, *Proceedings of 12th International Congress of Phonetic Sciences* 2: 290-293, Aix-en-Provence, France.

Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant 'segmental' anchoring of f0 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* (supplement), 5-54.

Liberman, A., Harris, K., Hoffman, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358-368.

Mücke, D., Grice, M., Becker, J., Hermes, A., & Baumann, S. (2006). Articulatory and acoustic correlates of prenuclear and nuclear accents. In R. Hoffmann & H. Mixdorff (Eds.). *Proceedings of Speech Prosody 2006* (pp. 297–300). Dresden: TUDpress Verlag der Wissenschaften GmbH.

Mücke, D., Grice, M., Becker, J., Hermes, A. (in press). Sources of variation in tonal alignment: evidence from acoustic and kinematic data. *Journal of Phonetics*.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology.* Dordrecht: Foris.

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

Prieto, P. (2005b). En torno a la asociación tonal en el modelo métrico-autosegmental. Puntos controvertidos en su aplicación al catalán. *Revista Internacional de Lingüística Iberoamericana, 6*, 9-28.

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. (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. *Prosodic Typology, 2*, ed. by Sun-Ah Jun. Oxford University Press: Oxford.

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

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., Mücke, D., Becker, J., Grice, M. (2007). Coordination Patterns between pitch movements and oral gestures in Catalan. In J. Trouvain and W. J. Barry (Eds.) *Proceedings of the XVIth International Congress of Phonetic Sciences*. Saarbrücken, Germany, 6-10 August 2007. (pp. 989-992). Pirrot GmbH: Dudweiler.

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 II Workshop on Spanish ToBI, Bellaterra, June 22.

Smith, R. (2004). *The role of fine phonetic detail in word segmentation.* Doctoral dissertation. University of Cambridge.

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.

Torreira, F. (2004). *Alignment between prenuclear H* tones and the segmental tier in Spanish: the anchoring hypothesis.* Undergraduate thesis, Université Libre de Bruxelles.

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

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

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

Welby, P., & Loevenbruck, 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., and 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.

Xu, Y., & Liu, F. (2006). Tonal alignment, syllable structure and coarticulation: Toward an integrated model. *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), 125–159.

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).

CATALAN

1a. *comprà ventalls*
[kumˌpɾa βən̺ˈtaʎʃ]
'(s)he bought fans'

1b. *compraven talls*
[kumˌpɾaβən̺ ˈtaʎʃ]
'they bought pieces'

2a. *comprà ventalls de vim*
[kumˌpɾa βən̺ˌtaʎʒ ðə ˈβim]
'(s)he bought wicker fans'

2b. *compraven talls de vim*
[kumˌpɾaβən̺ ˌtaʎʒ ðə ˈβim]
'they bought wicker pieces'

3a. *mirà batalles*
[miˌɾa βəˈtaʎəs]
'(s)he looked at battles'

3b. *mirava talles*
[miˌɾaβə ˈtaʎəs]
'(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 βəŋˈgwarðjəs]
'(s)he looked for newspapers'

5b. *buscaven guàrdies*
[busˌkaβəŋ ˈgwarðjəs]
'they were looking for guards'

6a. *buscà vanguàrdies a la tarda*
[busˌka βəŋˈgwarðjəz ə lə ˈtarðə]
'(s)he looked for newspapers'
in the afternoon'

6b. *buscaven guàrdies a la tarda*
[busˌkaβəŋ ˈgwarðjəz ə lə ˈtarðə]
'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 məˈti]
'(s)he appointed viscounts

10b. *nomenaves comtes al matí*
[numəˌnaβəs ˌkomtəz əl məˈti]
'you appointed counts

in the morning’

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 βən̪təˈʎɛt͡s]
‘(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*
[miˌɾaβə 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 san̪ˈtɛt͡s]
‘(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’

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 məˈti]
‘(s)he appointed little viscounts
in the morning’

20b. *nomenaves comtets al matí*
[numəˌnaβəs kumˌtɛd͡z əl məˈti]
‘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 muj ˈfwertes]
‘(s)he shoots very strongly’

4b. *daba lazos muy fuertes*
[ˌdaβa ˈlaθos muj ˈfwertes]
‘(s)he gave very strong knots’

5a. *compraré mostazas*
[kompɾaˈɾe mosˈtaθas]
‘I’ll buy mustards’

5b. *compraremos tazas*
[kompɾaˈɾemos ˈtaθas]
‘we’ll buy cups’

6a. *compraré mostazas alemanas*
[kompɾaˈɾe mosˈtaθas aleˈmanas]
‘I’ll buy German mustards’

6b. *compraremos tazas alemanas*
[kompɾaˈɾemos ˈtaθas aleˈmanas]
‘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]

12b. *bebo vinitos negros*
[ˌbeβo βiˈnitoz ˈneɣɾos]

‘(s)he sees black cows.dim’

‘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*
[kompɾaˈɾemostaˈθitas]
‘we’ll buy cups.dim’

16a. *compraré mostacitas alemanas*
[kompɾaˈɾe mostaˈθitas aleˈmanas]
‘I’ll buy German mustards.dim’

16b. *compraremos tacitas alemanas*
[kompɾaˈɾemostaˈθitas 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*
[ˈdamenuˈðitos eˈnoɾmes]
‘(s)he gives big offal.dim’

20b. *dame nuditos enormes*
[ˈdame nuˈðitos eˈnoɾmes]
‘give me big knots.dim’