Production and perception of word-edge tones in Catalan and Spanish

Eva Estebas-Vilaplana and Pilar Prieto
UNED and ICREA-UAB

1. Introduction

The aim of this paper is to clarify the role of word-edge tones in prenuclear rising accents (LH) in Catalan and Spanish broad focus declaratives. The modelling of prenuclear rises in different languages within the Autosegmental-Metrical approach of intonational analysis has shown that whereas the F0 valley presents a stable alignment and scaling, the anchoring of H is much more unprecise and prompts different phonological interpretations. Spanish prenuclear accents have been classified differently in the literature: H* with a peak delay (Prieto et al. 1995), L*+H (Sosa 1999, Face 1999 and Beckman et al. 2002) or (L+H)* (Hualde 2002). In Catalan (Estebas-Vilaplana 2000), prenuclear rises have been described as instances of a low pitch accent (L*) followed by a word edge tone (H). In this study a production test and a perceptual test were designed for both Catalan and Spanish so as to check whether the anchoring of H to the end of the word is consistently produced by speakers to disambiguate the location of the word boundary and also whether this H alignment is a helpful perceptual cue to identify word boundaries.

2. Production test

The aim of the production test is to examine whether H anchoring is used to disambiguate sentences. The hypothesis is that if an H word edge tone is present the F0 peak should be aligned at the end of words regardless of the number of postaccentual syllables.

2.1. Experimental procedure

The data used in the two production tests consisted of 20 pairs of potentially ambiguous sentences which had the same segmental and stress composition and were *only* distinguished by word boundary location. Spanish sentences included words with three stress distributions: *oxytones*, *paroxytones* and *proparoxytone*, as for example, ***ve*** *bovinos* (s(he) sees cows) vs. ***be****bo vinos* (s(he) drinks wines) or ***su****be Melino* (Melino goes up) vs. ***sú****beme lino* (bring up the linen for me). In Catalan only oxytones and paroxytones were examined, as in *com****prà*** *ventalls* ((s)he bought fans) vs. *com****pra****ven talls* (they bought pieces). Three speakers for each language read the utterances four times (360 utterances for each language). Sentences were analysed by means of *Praat*. Measurements were made on simultaneous displays of speech wave, wide-band spectrogram and F0 tracks. For each sentence, the following segmental and pitch landmarks were manually placed in the two test syllables: 1) Segmental landmarks: a) beginning of onset and vowel on the target accented syllable, b) beginning of onset, vowel, and coda

(whenever present) on the target postaccentual syllable, c) end of the postaccentual syllable, d) end of the onset of the following syllable and e) phrase-final point. 2) Pitch landmarks: a) L1, valley of the first pitch accent (in Hz) and b) H1, peak of the first pitch accent (in Hz). Figure 1 shows the Praat windows with the segmentation of the Spanish utterances *ve bovinos* and *bebo vinos*. The first two boxes displayed in each graph show the speech waveform and a spectrogram with an overlapped F0 trace. The last two boxes show the segmental landmarks with the segmental boundaries of the two test syllables and the pitch landmarks.

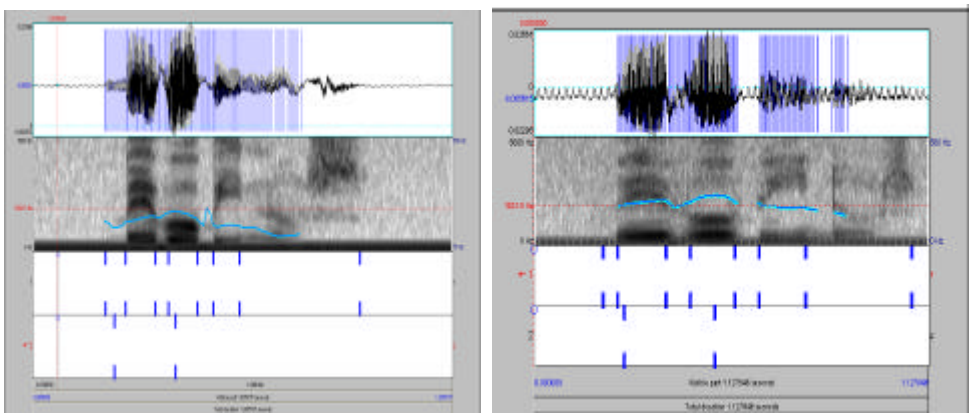

Figure 1. Diplays of speech waveform, spectrogram and F0 trace, and the segmental and pitch landmarks for the Spanish utterances *Ve bovinos* '(s)he sees cows' vs. *Bebo vinos* 'I drink wines'.

2.2. Results

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

The graphs displayed in Figures 2 and 3 plot H peak delay (or distance between the onset of the accented syllable and the F0 peak in ms) as a function of the duration of the accented syllable for the three Catalan and the three Spanish speakers respectively. The regression lines summarize the strong positive correlations found between the two variables both in Catalan (correlation coef. 0,67-0,82) and in Spanish (correlation coef. 0,42-0,91).

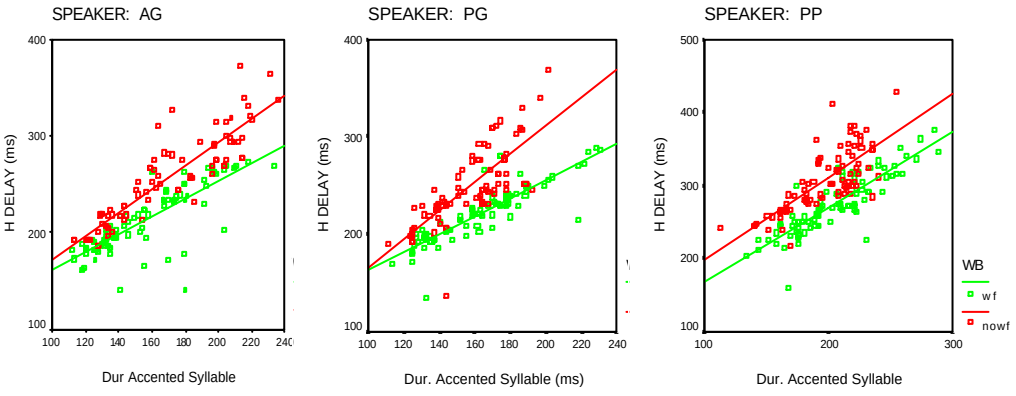

Figure 2 Peak delay as a function of the duration of the accented syllable in ms in the two stress conditions (word-final vs. word-medial stress) in Catalan for speakers AG, PG, and PP.

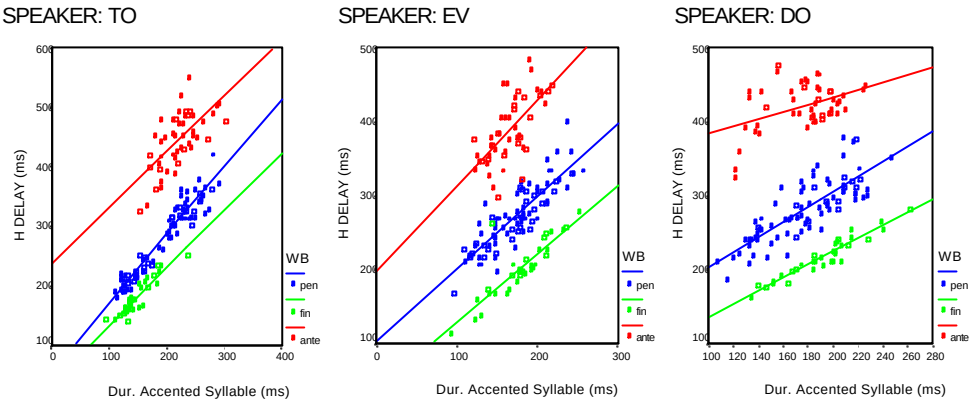

Figure 3. Peak delay as a function of the duration of the accented syllable in ms in three stress conditions (words with final, penultimate and antepenultimate stress) in Spanish for speakers TO, EV, and DO.

The results of Figures 2 and 3 also show that there is a consistent difference in H delay depending on the prosodic condition in both languages as the data are visually separated into two groups in Catalan and into three groups in Spanish. Hs in words with a final accent are less delayed than Hs in words with a penultimate accent. Similarly, in Spanish, Hs in words with a penultimate accent are less delayed than Hs in words with an antepenultimate accent. This seems to indicate that the location of the H is highly influenced by within-word position.

2.2.2. Effects of within-word position on H location

The graphs displayed in Figures 4 and 5 plot the mean H peak delay or distance in ms between the H peak relative to the beginning of the accented syllable in words with final (fin) and medial (med) stress for Catalan (Figure 4) and in words with fin, penultimate (pen) and antepenultimate (ante) stress for Spanish (Figure 5). The bars represent standard errors.

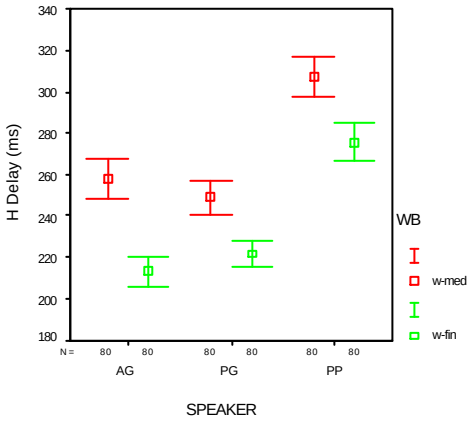

Figure 4. Mean H peak delay as a function of within-word position for all Catalan speakers in medial (w-med) and final (w-fin) position.

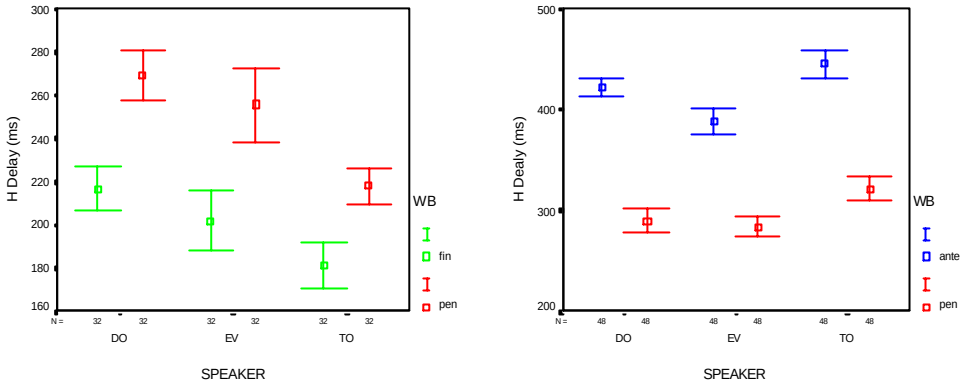

Figure 5. Mean H peak delay as a function of within-word position for all Spanish speakers: fin vs pen (graph on the left) and pen vs ante (graph on the right).

For both languages, the data reveal consistent effects of within-word position on H delay. For the three Catalan speakers, peak delay is significantly shorter ($p<0,001$) in word-final syllables (w-fin) than in word-medial syllables (w-med). For the Spanish data, peaks are less displaced in words with a final accent than in words with an accent on the penultimate (left graph) or antepenultimate syllables (right graph). Thus, F0 peaks are more displaced in paroxytones than in oytones and even more displaced in proparoxytones than in paroxytones.

The results in Figures 6 and 7 show the distance between the location of the H relative to the end of the accented syllable in words with fin and med stress for Catalan (Figure 6) and in words with fin, pen and ante stress for Spanish (Figure 7). Taking the 0 value as the end of the accented syllable, the plots show that all F0 peaks are displaced to the postaccentual syllable/s, since all peaks are located beyond the 0 value. For the three Catalan speakers, peak distance to the end of the syllable is significantly shorter ($p<0.0001$) in word-final position than in word-medial position. The same is true for the Spanish data (left graph). The results also show that in Spanish words with antepenultimate stress F0 peaks are even more displaced than in words with penultimate stress (right graph). These results reveal that there is no strict anchoring of the H at the end of the word.

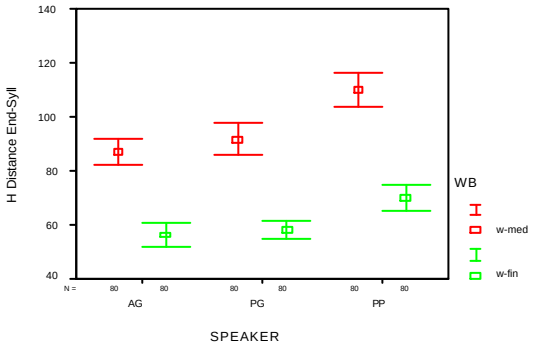

Figure 6. Mean distance in ms between the H peak relative to the end of the accented syllable as a function of within-word position for all Catalan speakers in med and fin positions.

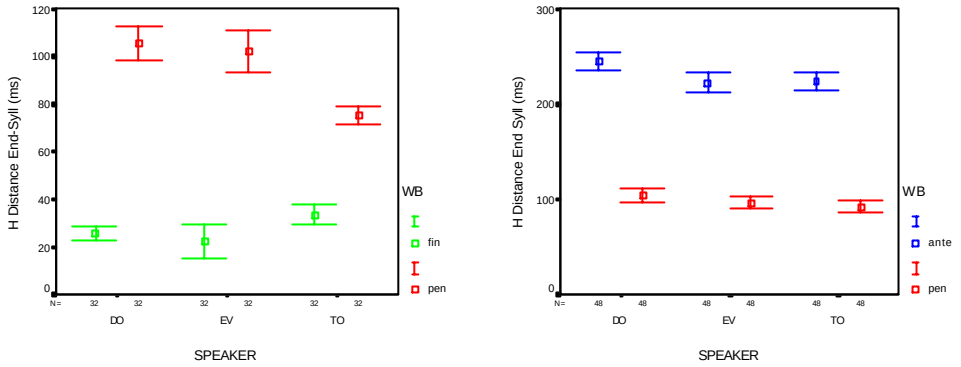

Figure 7. Mean distance in ms between the H peak relative to the end of the accented syllable as a function of within-word position for all Spanish speakers: fin vs pen (graph on the left) and pen vs ante (graph on the right).

Separate ANOVAs were done for the two measures of H alignment (relative to either the beginning and the end of the syllable) for the three Catalan and Spanish speakers. Both for Catalan and Spanish, the ANOVAs revealed a significant effect of word boundary both for the measures of peak delay and peak distance to the offset of the syllable. For Catalan, the results of peak delay are: Speaker AG: $F(1, 158) = 50{,}088$; $p<0{,}001$; Speaker PG: $F(1, 158) = 24{,}574$; $p<0{,}001$; Speaker PP: $F(1, 158) = 22{,}434$; $p<0{,}001$). The results of peak distance to the offset of the syllable are: Speaker AG: $F(1, 158) = 84{,}230$; $p<0{,}001$; Speaker PG: $F(1, 158) = 98{,}253$; $p<0{,}001$; Speaker PP: $F(1, 158) = 103{,}399$; $p < 0{,}001$. For Spanish, the results of peak delay are: Speaker TO: $F(2, 159) = 269{,}505$; $p<0{,}001$; Speaker EV: $F(2, 159) = 204{,}951$; $p<0{,}001$; Speaker DO: $F(2, 159) = 348{,}952$; $p<0{,}001$). The results of peak distance to the offset of the syllable are: Speaker TO: $F(2, 159) = 726{,}453$; $p<0{,}001$; Speaker EV: $F(2, 159) = 629{,}163$; $p<0{,}001$; Speaker DO: $F(2, 159) = 780{,}654$; $p < 0{,}001$. Thus all speakers in both languages show that peak location is affected by within-word position in a way that the peak systematically shifts backwards as the end of the word is closer.

Finally, the alignment of the H tonal target relative to the end of the word was also examined. If H was a word edge tone then it would be expected to be consistently anchored at the right edge of the word. The graphs in Figures 8 and 9 plot the distance in ms between the H and the end of the word in words with fin and med stress for Catalan (Figure 8) and in words with fin, pen and ante stress for Spanish (Figure 9). Taking the 0 value as the end of the word, the results show that in words with a final accent, the H is placed after the end of the word both in Catalan and in Spanish (left graph). Alternatively, in words with a penultimate accent, the H is located before the end of the word. Similarly, in the Spanish data, peaks in words with an antepenultimate accent are also located before the end of the word (right graph). These results confirm the idea that there are no strict word-anchoring effects in any of the two languages.

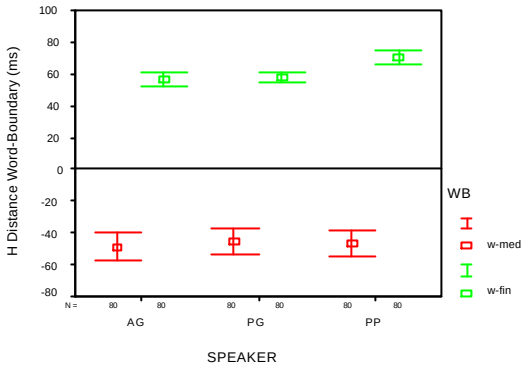

Figure 8. Mean values of distance between H peak location relative to the right edge of the word (in ms) as a function of within-word position for all Catalan speakers in med and fin positions.

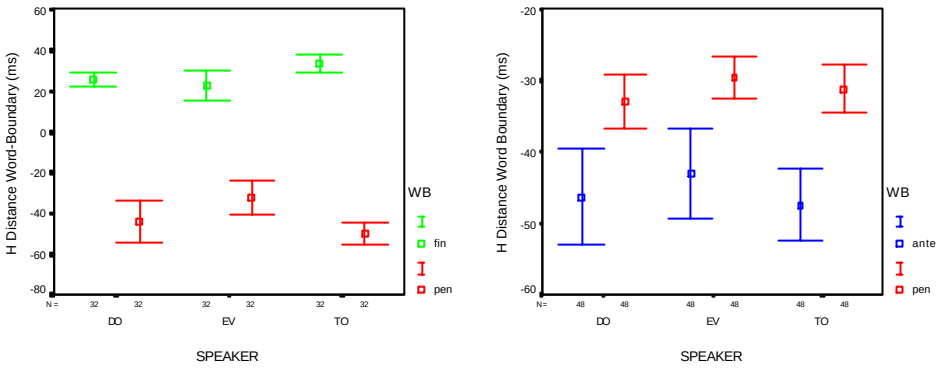

Figure 9. Mean values of distance between H peak location relative to the right edge of the word (in ms) as a function of within-word position for all Catalan speakers: fin vs pen (left graph) and pen vs ante (right graph).

2.2.3. Effects of within-word position on syllable duration

The results displayed in Figures 10 and 11 show the mean values of duration of the accented syllable in words with fin and med stress for Catalan (Figure 10) and in words with fin, pen and ante stress for Spanish (Figure 11). It is expected that accented syllables before a word-boundary will be longer than accented syllables in medial position (Lehiste 1960, Turk and White 1999).

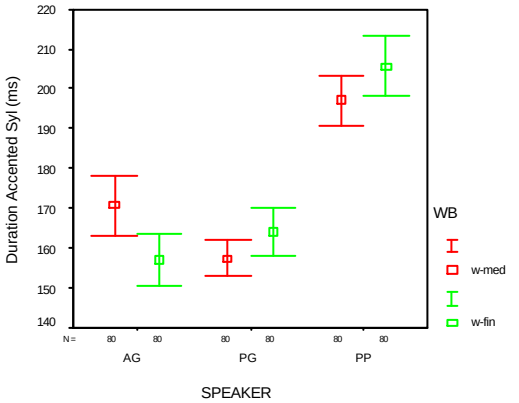

Figure 10. Mean values of duration of the accented syllable as a function of within-word position for all Catalan speakers in med and fin positions.

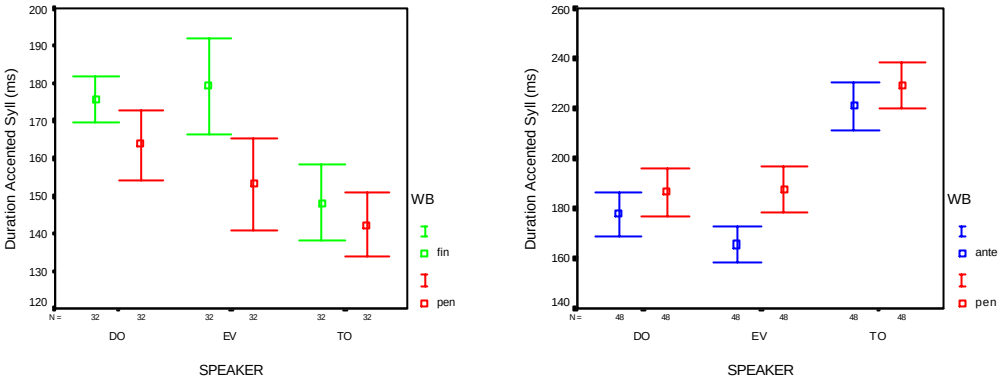

Figure 11. Mean values of duration of the accented syllable (in ms) as a function of within-word position for all Spanish speakers: fin vs pen (graph on the left) and pen vs ante (graph on the right).

The results show no clear duration effects. Word-final accented syllables are slightly longer for all speakers in both languages except for AG (Catalan). However, these differences were only significant (p=0.004) for speaker EV (Spanish). This seems to indicate that duration differences as a cue to indicate word boundary location are optional.

The results displayed in Figure 12 and 13 show the mean values of duration of the postaccentual syllable in words with fin and med stress for Catalan (Figure 12) and in words with fin, pen and ante stress for Spanish (Figure 13). It is expected that postaccentual syllables before a word-boundary will be longer than non-accented syllables in medial position. However, the results show no clear duration effects. Postaccentual syllables in non-word-final position are slightly longer than postaccentual syllables in word-final position except for speakers AG (Catalan) and TO (Spanish). These differences are not statistically significant for any of the speakers (at p>0,05) except for speaker PP (Catalan). Thus, the hypothesis that the postaccentual syllable will be longer before a word boundary is not confirmed by these data.

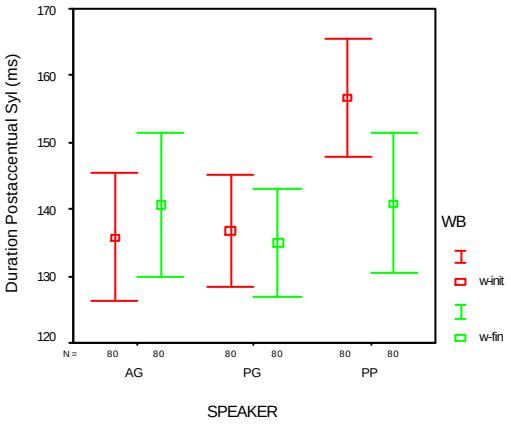

Figure 12. Mean values of duration of the postaccentual syllable (in ms) as a function of within-word position for all Catalan speakers speakers in medial (w-med) and final (w-fin) position.

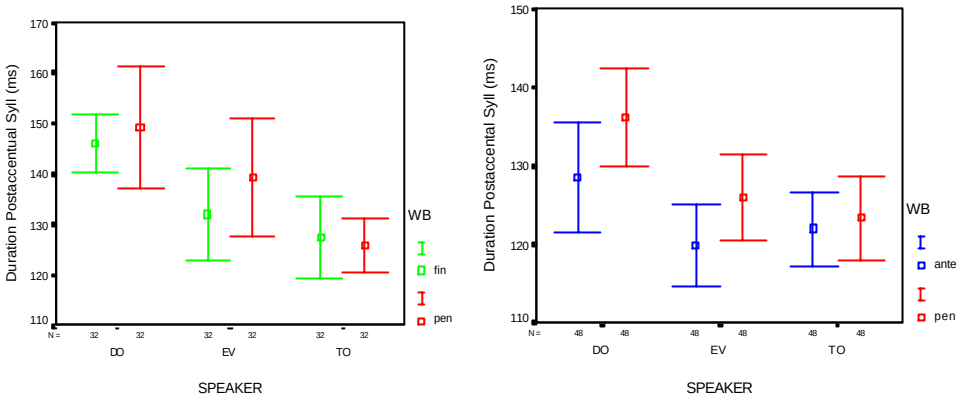

Figure 13. Mean values of duration of the postaccentual syllable (in ms) as a function of within-word position for all Spanish speakers: fin vs pen (graph on the left) and pen vs ante (graph on the right).

Separate ANOVAs were performed for the two dependent variables (namely, the duration of the accented syllable and the duration of the postaccentual syllable) for all speakers. In Catalan, for speaker PG, the ANOVAs revealed that none of the variables showed a significant effect of word boundary, $F(1, 158)=3,054$; $p=0,082$ (accented syllable), and $F(1, 158)=0,088$; $p=0,767$ (postaccentual syllable). For speaker AG, the ANOVA for the duration of the accented syllable showed a significant effect of word boundary, $F(1, 158)=7,631$; $p<0,005$, but not for the duration of the postaccentual syllable, $F(1, 158)=0,450$; $p=0,503$. Conversely, for speaker PP, the effect of word boundary was only significant for the duration of the postaccentual syllable ($F(1, 158)= 5,221$; $p=0,024$), bot not for the duration of the accented syllable ($F(1, 158)=2,903$; $p=0,024$). In Spanish, speakers EV and DO showed no significant effect of word boundary for none of the variables. Speaker EV: $F(2, 159)=1,791$; $p=0,170$ (accented syllable), and $F(2, 159)=3,996$; $p=0,02$ (postaccentual syllable); Speaker DO: $F(2, 159)=0,007$; $p=0,993$ (accented syllable), and $F(2, 159)=4,321$; $p=0,015$ (postaccentual syllable). For speaker TO, the ANOVA for the duration of the accented syllable showed a significant effect of word boundary, $F(2, 159)=27,773$; $p<0,001$, but not for the duration of the postaccentual syllable, $F(2, 159)=0,939$; $p=0,393$. Syllable duration measures thus indicate that syllable lengthening is not systematic neither in word-final nor in word-initial syllables.

3. Perception tests

The main goal of the perception tests was to check whether H alignment differences could be used to identify word boundaries between otherwise identical sentences.

3.1. Identification test

The materials used in the identification test consisted of 10 ambiguous utterances from the production test. 12 Catalan and 20 Spanish listeners heard the ambiguous utterances a maximum of three times and had to

identify the sentence in a two-choice task. The results of the identification test are presented in Figure 14 for Catalan and 15 for Spanish which the frequency of correct responses for all listeners for each sentence type: Type A includes oxytonic words, Type B paroxytonic words and type C proparoxytonic words. The results show a low rate of identification for the Catalan sentences (35 to 66% of correct responses) and variable rate of identification for Spanish sentences (25% to 95%). Words with an antepenultimate or penultimate accent tend to be better identified than words with final stress. These results indicate that the clear H displacement towards the right edge of the word in paroxytones and proparoxytones helps the hearer to a better identification. In oxytones, H placement is more ambiguous and hence contributes to a more difficult identification.

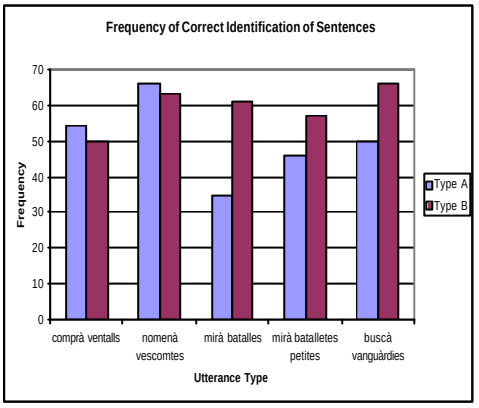

Figure 14. Frequency of correct identification of the 10 ambiguous Catalan sentences.

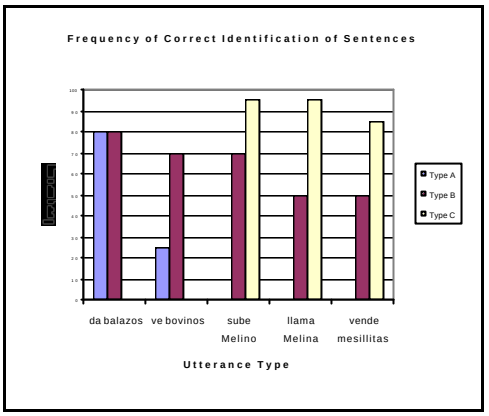

Figure 15. Frequency of correct identification of the 10 ambiguous Spanish sentences.

3.2. Discrimination test

The discrimination test investigates whether listeners could identify sentences in a better way when heard in pairs. The stimuli used were the same 10 sentences as before but presented in pairs. The same listeners as for the identification test heard the pairs of sentences for a maximum of three times. Subjects had to decide in which order sentences were presented.

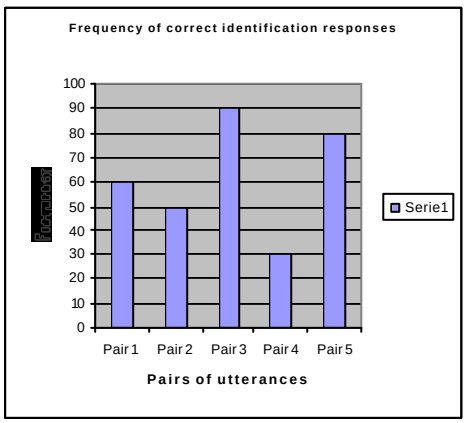

Figure 16. Frequency of correct identification responses for 5 pairs of Catalan ambiguous sentences.

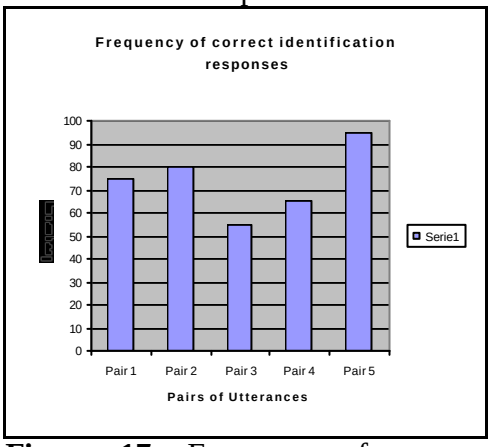

Figure 17. Frequency of correct identification responses for 5 pairs of Spanish ambiguous sentences.

The results of the discrimination test are presented in Figure 16 for Catalan and 17 for Spanish. The graph plots the frequency of correct identification responses for each pair. The results show a variable rate of identification for both Catalan (30% to 90% of correct responses) and Spanish (55% to 95%). Overall, the results show some improvement on the identification rates when sentences are heard in pairs.

4. Conclusion

The results obtained in the production test clearly show that the F0 peak is aligned after the accented syllable in all kinds of words (oxytones, paroxytones and proparoxytones) indicating that there is no strict anchoring of H at the end of the word. Thus, the idea to analyse Catalan and Spanish prenuclear rises with a final H word-edge tone is ruled out by the evidence of the data. Alternatively, the data showed a clear effect of within-word position on the alignment of H, since the F0 peak is more retracted in oytones than in paroxytones and proparoxytones. The results of the perception tests show that listeners do seem to rely on H alignment to distinguish sentences. The close location of H after the word boundary in oxytones causes some difficulties in identifying single sentences. However, some improvement on the identification rates is observed when sentences appear in pairs. This indicates that H alignment differences due to word position are a strong cue for word-boundary identification.

Bibliography

BECKMAN, M.; M. DÍAZ-CAMPOS, J. T. MCGORY and T. A: MORGAN: «Intonation accross Spanish, in the Tones and Break Indeces Framework», *Probus*, 14, 2002, pp. 9-36.

ESTEBAS-VILAPLANA, Eva: *The use and realisation of accentual focus in Central Catalan with a comparison to English*, University College London, PhD dissertation, 2000.

FACE, Timothy: «A Phonological analysis of rising pitch in Castilian Spanish», *Hispanic Linguistics*, 11, 1999.

HUALDE, J. Ignacio: «Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestinois», in C. Wiltshire and J. Camps (eds.): *Romance phonology and variation*, Amsterdam, J. Benjamins, 2002, pp. 101-115.

LEHISTE, Ilse: «An acoustic-phonetic study of internal open juncture», *Phonetica*, 5, 1960.

PRIETO, P.; J. VAN SANTEN and J. HIRSCHBERGH: «Tonal alignment patterns in Spanish», *Journal of Phonetics*, 23, 1995, pp. 429-451.

SOSA, J. Manuel: *La entonación del español: su estructura fónica, variabilidad y dialectología*, Madrid, Cátedra, 1999.

TURK, A. E. and L. WHITE: «Structural influences on accentual lengthening in English», *Journal of Phonetics*, 27, 1999, pp. 171-206.