Word-boundary Effects on Pitch Timing in Spanish

Miquel Simonet[1]
University of Illinois at Urbana-Champaign

1. Introduction

There is evidence from a number of languages that the position of the accented syllable with respect to the word in which it appears has an effect on the realization of the corresponding tonal event. Peaks tend to shift back (to the left) as word edges approach. Word-edge effects have been shown for English (Silverman and Pierrehumbert 1990), Catalan (Estebas 2003, Prieto *forthcoming*, Prieto and Estebas 2005), Serbian-Croatian (Godjevac 2000) and Spanish (de la Mota 2005, O'Rourke 2005, Prieto *et al.* 1995, Prieto and Estebas 2005, Simonet and Torreira 2005). For instance, in Prieto *et al.* (1995), a significant effect of lexical-stress configuration in a three-way comparison of Spanish sequences such as *número rápido*, *numero nervioso* and *numeró regular* was found. The data in Prieto *et al.* (1995) were collected from the performance of two speakers of Mexican Spanish.

In Iberian Romance, the role of word edges on pitch peak alignment is an area of some controversy. Recent research has focused on Castilian Spanish, as spoken in Madrid and Salamanca (Prieto and Estebas 2005, de la Mota 2005), and on Central Catalan (Estebas 2003, Prieto and Estebas 2005, Prieto *forthcoming*), as spoken in Barcelona. In this section, we will briefly review the findings for the two languages. Differences in the shape of pre-nuclear pitch accents between Castilian Spanish and Central Catalan are yet to be found: pre-nuclear pitch accents consist of a rise with the valley aligned with the stressed syllable's onset and with the peak in the post-stressed syllable (Hualde 2000). Thus, effects found for Central Catalan can hypothetically be extended to Castilian Spanish and vice versa.

Investigations on the modeling of Catalan pre-nuclear accents have found robust word-boundary effects (Estebas 2003). Estebas (2003) showed that F_0 peaks in Catalan pre-nuclear accents are consistently aligned with respect to the right edge of the word. In her data, peak delay increases as the distance from the stressed syllable to the word boundary increases, i.e. proparoxytones have a longer peak delay than paroxytones, and paroxytones have a longer peak delay than oxytones. This was established with the use of correlations between the duration of the word and a measure of peak delay. Consequently, Estebas (2003) proposes a labeling of Catalan pre-nuclear accents that consists of a monotonal pitch accent, L*, and a monotonal word-edge tone Hω: L*…Hω.

Prieto (*forthcoming*) and Prieto and Estebas (2005) revisited the effect of word boundaries on the location of F_0 peaks and the status of word-edge tones in Catalan pre-nuclear accents. These authors examined the production of minimal pairs that are distinguished only by the position of the word boundary. Instances of such potentially ambiguous utterances are as follows: *mirava talles* '(she) used to watch carvings' versus *mirà batalles* '(she) watched battles'. The results of the experiments showed that peaks were not strictly anchored at the edge of the word since in all sequences peaks were displaced to the post-accentual syllable. However, a consistent and significant word-position effect was found in that peak delay differed between words with lexical stress in the final syllable – *mirà* – and words with lexical stress in the penultimate syllable – *mirava*. Prieto (*forthcoming*) and Prieto and Estebas (2005) reject the existence of word-edge tones due to the fact that all pitch accents, including those in oxytones, show peak delay to the post-accentual syllable; however, they do not offer any explanation for the lexical-stress effects on pitch timing they still find.

---

[1] I gratefully acknowledge the helpful comments of José I. Hualde, Jennifer S. Cole, Pilar Prieto, Laura Colantoni and Barbara E. Bullock. I also thank Francisco Torreira and Eunice Díaz for technical assistance. Special thanks are due to the editors and to the audience of HLS, 2005. Evidently, all shortcomings are my own.

Prieto and Estebas (2005) and de la Mota (2005) investigated the effect of word boundaries on pitch peak timing in Castilian Spanish. Their data consisted of minimal pairs such as: *ve bovinos* '(she) sees bovines' versus *bebo vinos* 'I drink wines'. The results are identical to those in Prieto (*forthcoming*) and Prieto and Estebas (2005) for Catalan: word-edge effects are found but not to the point to present evidence for the existence of word-edge tones (Estebas 2003), since peaks do not consistently align with word edges and there is peak delay to the post-tonic in all pitch accents.

The data and their interpretation are, thus, far from being clear. First, while word-boundary effects have been shown for Central Catalan, there are differences in the degree of robustness found: while Estebas (2003) shows peaks to be consistently aligned with the word edge, Prieto (*forthcoming*), using different materials, shows them not to be. Prieto (*forthcoming*) suggests that the differences between Estebas's (2003) and Prieto's (*forthcoming*) data may be due to the fact that the former were in fact implementations of H- intermediate boundary tones, which unsurprisingly align to the word's offset, while the latter were simple rising pitch accents. Second, the question remains as to why word boundaries have an effect on pitch peak timing if it is not due to the existence of word-edge anchoring.

Regarding the case of Spanish, some issues arise, most of which having to do with potential flaws in the experimental design of the two papers that have addressed this topic: Prieto and Estebas (2005) and de la Mota (2005). The data in both papers were collected exclusively with the use of minimal pairs. It is at least a possibility that the subjects in these experiments were explicitly aware of these contrasts and that they intended to produce the contrast as clearly as possible, thus exaggerating the effects. If this were true, the external validity of these studies would be seriously threatened since we would not be able to extend the findings to natural speech with enough confidence.

The issues that arise, (*i*) that Catalan evidence so far is contradictory and (*ii*) that Spanish evidence is unclear and, most importantly, has been collected with the use of minimal pairs, motivate further research on the topic. The present paper is a contribution to the study of word-edge effects on pitch timing.

2. The experiment

2.1. Research questions and hypotheses

The main goal of the present experiment is to examine the potential effect of word boundaries in the modeling of Spanish pre-nuclear accents, while controlling for several potentially confounding factors, including the potential effect of tonal context. Several research questions inform this paper. The main research questions are as follows: Does the distance between the stressed syllable and the corresponding word boundary (or the lexical-stress configuration of words) have a consistent effect on peak location? If so, how robust are these effects? Do peaks align consistently with respect to word edges? Is there evidence for word-edge tones in Castilian Spanish, as it has been suggested for Centreal Catalan (Estebas 2003)?

Our hypothesis is that word boundaries have some shaping effect on tonal gestures, considering de la Mota's (2005) Prieto and Estebas' (2005) findings. From our point of view, this would support a position in which peaks are associated in some loose way with right-hand word edges (Estebas 2003). The fact that all oxytonic and paroxytonic words have some degree of peak displacement towards the post-accentual syllable, as Prieto and Estebas (2005) and de la Mota (2005) show, does not strongly exclude, as Prieto and Estebas argue, the possibility of these tonal landmarks to be associated with word-edges. While strict word-edge alignment would provide the clearest evidence of association, other possibilities, including a weaker, more variable, alignment are also possible. It is not entirely evident to us how word-boundary effects, if found, could be explained without some reference to word edges in one way or another.

2.2. Methods

Six participants read a total of 1872 sentences. The subjects are in their late twenties or early thirties. Three of the subjects are male (1M, 2M, 3M) and three are female (4F, 5F, 6F). They are from similar social and educational backgrounds: None of them have college-level studies and all belong to what we could call the working class. The data were collected in the Mediterranean island of Majorca, Spain. Due to the language contact situation in Majorca, the subjects were selected after having answered a language background questionnaire (LBQ) based on the one used in Flege and MacKay (2004). The subjects have only passive knowledge of Majorcan Catalan, their L1 is Spanish, and they self-reportedly use Spanish in their daily lives over 95% of the times. While they were all born in the island, both their parents are from the centre-south of the Iberian Peninsula, from the provinces of Jaén, Granada or Albacete. We believe that we could classify our speakers as speakers of some broadly defined Castilian variety of Spanish. Thus, we believe our data to be perfectly comparable to that of Prieto and Estebas (2005), from Madrid, and de la Mota (2005), from Salamanca.

The sentences contained balanced instances of utterances where the number of intervening unstressed syllables between the target accent and the following stressed syllable was manipulated. The subjects read equal numbers of sequences where 0, 1, 2, 3 or 4 unstressed syllables appeared between the target pitch accent and the following one. In all these tonal-context situations, the position of the word boundary was moved as much as constraints on the lexicon would allow us. Thus, under most of the tonal-context situations, we tested for the placement of the peak in oxytones (stress in final syllable), paroxytones (stress in penultimate) and proparoxytones (stress in antepenultimate). This is so only in tonal-context situations where 2, 3 or 4 unstressed syllables intervened between the two accented syllables. Regarding utterances where only one unstressed syllable appeared in between the two accented ones, only a two-way comparison is possible: paroxytones versus oxytones.

All consonantal segments between and around the two relevant pitch accents, the target and the following one, were voiced. We tested pairs of words, but only the first word was measured: the following one was used to manipulate the context, both the accentual context and the position of the word boundary. All items were instances of a noun plus an adjective. All pre-nuclear accents occurred within the sentential predicate, not including the verb. We believe that this was necessary due to the following fact: Face (2002), in his corpus of 2677 utterances, found that almost all H- (high intermediate boundary tones) in Spanish declaratives occurred either after the subject or after the verb. Estebas (2003), Prieto (*forthcoming*) and Prieto and Estebas (2005) did not control for this fact in their corpora: their materials mostly consisted of combinations of verbs plus nouns. As also pointed out by Prieto (*forthcoming*) this could be a confounding factor, since some delayed peaks that could seem to be very much associated with the word edge, could in fact be instances of phrasal H- and not word-edge tones, or delayed pitch accents. Or they could in fact be a combination of both. Thus, in order to minimize the error rate, we decided to restrict the position of our tested pitch gestures to those that only sporadically may receive intermediate high boundary tones (H-).

Furthermore, 50% of the sentences had an added last phrase to the right of the two-word (N+A) sequence we were controlling for. This control balances the possibility for utterance-final accents or final lowering effects to introduce additional variation in peak alignment. In the other 50% of the sentences the adjective was indeed the last word in the sentence.

In summary, our materials consisted of a total of 1872 sentences (104 sentences * 6 speakers * 3 repetitions). Of the 104 sentences, each contained a distinct two-word sequence where the first word of the sequence is the target word. Examples can be observed in (1).

(1) a. *Juan visitó un jar**dín** maravillloso (y cuidado)*
oxytone target word followed by three intervening unstressed syllables
b. *Su padre disciplinaba con **ma**no vigorosa (y firme)*
paroxytone target word followed by three intervening unstressed syllables
c. *La catedral tiene una **bó**veda bonita (y grande)*
proparoxytone target word followed by three intervening unstressed syllables

The participants were asked to read the sentences as naturally as possible and to repeat any misread sentences. However, subjects were not coached. They were left with the decision of repeating any sentences they felt they misread. Recordings of the materials were made through a Shure SM10A head-worn dynamic microphone into a digital solid-state recorder, Marantz PMD 660. The speech signal was digitized at 44.1 kHz and 16-bit quantization.

The stimuli were presented to the participants through a slide presentation on a laptop screen. The presentation contained one sentence per slide and subjects controlled the pace of the presentation by pressing a key when they were ready for the next sentence.

The sentences were analyzed using the Praat signal-processing package (Boersma and Weenink 2005). Segmental landmarks were manually labeled based on time-synchronized displays of sound waves and wide-band spectrograms. Pitch turning points were labeled semi-automatically using Praat's built-in tools "Move cursor to maximum pitch", after selecting the area of the peak. For this, time-aligned F_0 tracks were used in addition to sound waves and wide-band spectrograms. The labeled landmarks were those in (2).

(2) a. The onset of the stressed syllable of the first word of the pair.
b. The offset of the stressed syllable of the same word.
c. The right-hand word-edge.
d. The F_0 peak corresponding to the F_0 rise of the target syllable.

Several time distances, which will be treated as dependent variables in the upcoming sections, were automatically derived from the landmarks in (2). The dependent variables will be individually explained in the upcoming sections.

3. Results

The conditions that have been included in the present study do not completely cross-classify. The three-way contrast between oxytones, paroxytones and proparoxytones is only available in sentences in which 2 or more intervening unstressed syllables appear between the two pitch accents. In the present article, we will be reporting exclusively on this subset of the dataset, in which the three-way comparison is possible. While proximity to the following pitch accent has been shown to affect peak delay, there are no significant differences in this respect in this subset. Thus, we will collapse in our analyses data from sentences in which 2, 3 or 4 unstressed syllables intervened between the target syllable and the following stressed syllable (see Arvaniti *et al.* 1998). We will be providing quantitative analyses of absolute and relative measures of peak delay and of location of pitch events with respect to the segmental string. Furthermore, we will provide correlation values between peak delay and the duration of the targeted word-stretch. The rest of the paper is based on a total of 1588 sentences, since a number of sentences had to be excluded due to noise in the sound files, pausing, disfluencies or other impediments.

3.1. Effects of lexical-stress configuration on the timing of pitch peaks

The absolute measures of peak delay in milliseconds were submitted to 6 independent one-way ANOVAs, one for each of the 6 speakers. Peak delay is calculated by subtracting the time value of the peak from the time value of the onset of the tonic syllable. The main factor had three levels of lexical-stress configuration or, in other words, distance of pitch accent to word boundary (oxytone, paroxytone and proparoxytone). For the statistical analyses, the R language and environment for statistical computing and graphics was used. For the six speakers, the results of the ANOVAs revealed a highly significant effect of the main factor [1M, $F(2, 204)=9.19$, $p<.001$; 2M, $F(2,184)=11.6$, $p<.001$; 3M, $F(2,177)=22.95$, $p<.001$; 4F, $F(2,192)=45.02$, $p<.001$; 5F, $F(2, 167)=9.74$, $p<.001$; 6F, $F(2,200)=45.41$, $p<.001$]. Post-hoc tests using Tukey HSD revealed four main grouping patterns: for two speakers (4F and 6F), the three levels were significantly different from each other; for two speakers (2M and 3M), proparoxytones were different from the other two, while paroxytones were not different from oxytones; for one speaker (5F), oxytones were different from the other two, while paroxytones were not different from proparoxytones; and, finally, for one speaker (1M) only the two extremes were different. A display of the results can be observed in Figure 1.

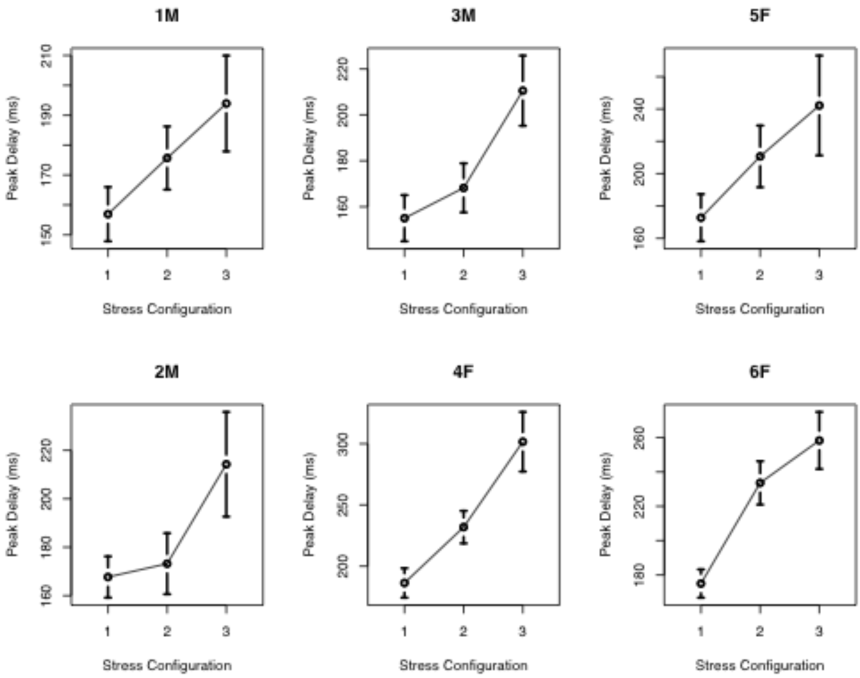

Figure 1. Error-bar plots (95% confidence interval) of peak delay (ms) as a function of lexical-stress configuration of words: 1 (oxytones), 2 (paroxytones) and 3 (proparoxytones) for 6 Spanish speakers.

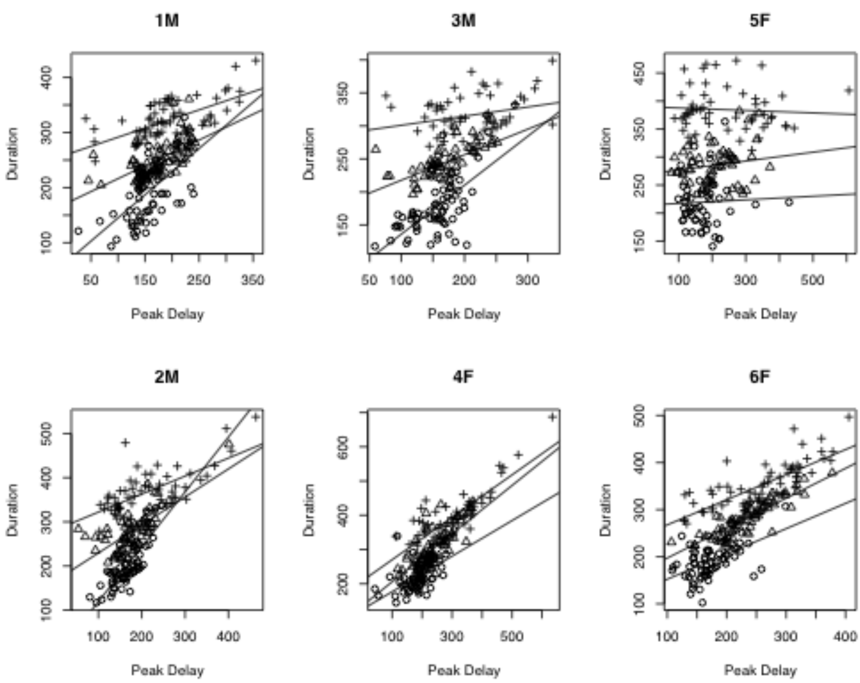

Figure 2. Correlation plots of peak delay and the duration of the word stretch. Circles correspond to oxytones, triangles, to paroxytones and crosses, to proparoxytones. Independent regression lines for each stress configuration.

In order to compare our results to those of Estebas (2003) for Catalan, we also performed correlations between the absolute measures of peak delay and the duration of the word-stretches, the distance from the onset of the stressed syllable to the word-edge. Estebas's results showed a very low correlation between these two measures when all word items (words belonging to the different lexical-stress configurations) were collapsed, R^2 ranging from 0.15 to 0.35. However, when correlations were performed differently on the three different groups, R^2 values raised to a range between 0.7 and 0.89, which is an acceptable correlation.

The results of the correlations can be observed in Figure 2. The R^2 values range from an acceptable $R^2 = 0.83$ to an unacceptable $R^2 = 0.28$ [1M, $R^2 = 0.55$; 2M, $R^2 = 0.62$; 3M, $R^2 = 0.59$; 4F, $R^2 = 0.83$; 5F, $R^2 = 0.28$; 6F, $R^2 = 0.79$]. Independent correlation tests were performed on the data separated by lexical-stress group, but no significant improvement was found. Interestingly, it can be observed that the three groups seem to cluster differently. The three clusters on the y-axis are not surprising since we are comparing a group with one syllable (oxytones) with a group with two (paroxytones) and a group with three (proparoxytones). However, the clustering on the x-axis, which can be seen for some of the speakers, reveal that distance to the word edge plays a main role on peak placement. In any case, correlations are not as high as those in Estebas (2003). In Figure 2, regression lines are shown for each lexical-stress configuration independently.

Important word-boundary effects have been found for 6 speakers of Spanish. Even though post-hoc tests revealed differences among the speakers in the paired differences, the effect of lexical-stress configuration on pitch peak timing was highly significant for all of them, with $p<.001$. The question remains as to whether these effects are due to the existence of H word-edge tones in Spanish. Recall that Prieto (*forthcoming*) and Prieto and Estebas (2005) argued against a strong interpretation of the word-edge theory mainly due to the fact that, in their dataset, pitch peaks pertaining to oxytones were also displaced to the post-tonic syllable, as also were those that belonged to paroxytones.

3.2. Alignment of pitch peaks with respect to the offset of stressed syllables

In order to address this last question and to investigate the status of word-boundary tones in Spanish, we measured the distance in milliseconds from the offset of the stressed syllable to the position of the peak. This is done by subtracting the time value of the pitch peak from the time value of the offset of the tonic syllable. Considering Prieto's (*forthcoming*) results for Catalan and Prieto and Estebas's (2005) and de la Mota's (2005) for Spanish, we would expect that all peaks were displaced to the post-accentual syllable, even those that belong to oxytones. If strict alignment with word boundaries occurs, however, only peaks in paroxytones and proparoxytones will be displaced to the post-accented syllable, while those in oxytones will be aligned with respect to the stressed syllable, which is also the word boundary. Figure 3 is a display of the results of this measurement. Note that 0 in the y-axis stands for the offset of the stressed syllable.

The data were submitted to 6 independent one-way ANOVAs, one for each of the 6 speakers. The main factor is, again, lexical-stress configuration (3 levels). The results show that peaks are not aligned with respect to the stressed syllable's offset, since the measures are significantly different as a function of the main factor for all subjects [1M, $F(2,2040)=97.44$, $p<.001$; 2M, $F(2,184)=75.19$, $p<.001$; 3M, $F(2,177)=108.46$, $p<.001$; 4F, $F(2,192)=188.92$, p.<001; 5F, $F(2,167)=38.46$, $p<.001$; 6F, $F(2,200)=204.6$, $p<.001$]. As it can be seen, the peaks are more displaced in paroxytones than in oxytones and, even more, in proparoxytones. Note also that the average peak alignment for oxytones falls within the bounds of the stressed syllable for all 6 speakers. This is different from previous results cited above (Prieto and Estebas 2005, de la Mota 2005), in which peaks were displaced to the post-tonic also in oxytones. This may be due to the fact that all our oxytones included a coda consonant, while those in previous works ended in vowels, it may also be due to the position of our pitch accents with respect to the sentences.

The results presented so far seem to point towards a position in which pitch peaks are aligned with respect to the word-boundary and not to the syllable's offset. First, ANOVAs with peak delay as the dependent variable and lexical-stress configuration as the independent variable reached significance. Second, regarding the alignment of peaks with respect to the offset of the stressed syllable, peaks in oxytones appear in its vicinity (which is also the word edge for them), while peaks in the other two configurations are quite displaced to the right. However, it has yet to be shown that peaks in paroxytones and proparoxytones are actually aligned with respect to their word-boundaries. In the following section we aim at investigating this.

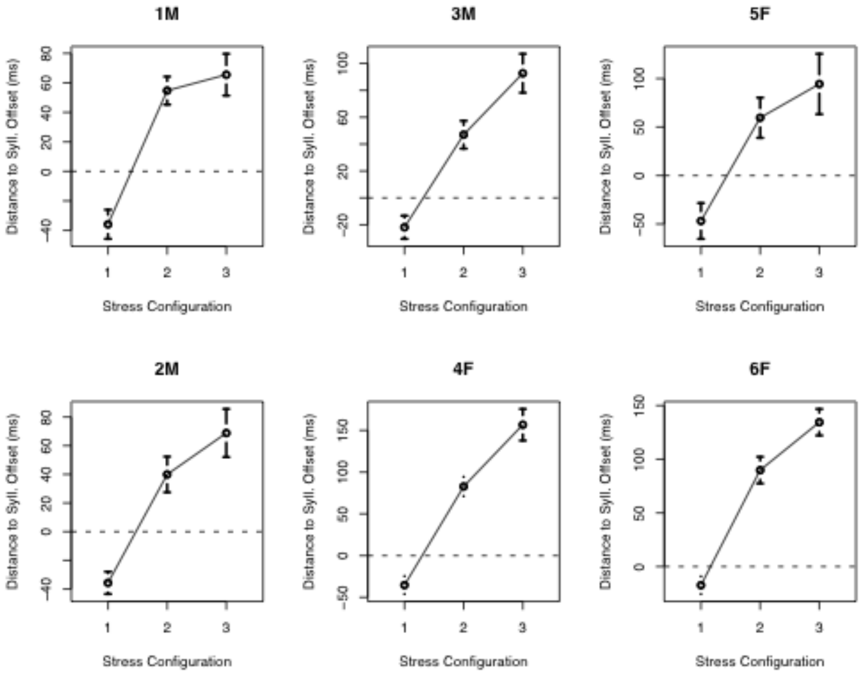

Figure 3. Error-bar plots (95% confidence interval) of distance of peak to the offset of stressed syllable (ms) as a function of lexical-stress configuration of words: 1 (oxytones), 2 (paroxytones) and 3 (proparoxytones) for 6 Spanish speakers. Horizontal dashed lines represent the offset of the stressed syllable.

3.3. Alignment of pitch peaks with respect to word edges

If greater peak delay in paroxytones and proparoxytones is due to the alignment of peaks with respect to word edges, we would expect that a measure of the time distance between the peak and the word boundary would be similar for all three stress configurations. This measure is calculated by subtracting the time value of the pitch peak from the time value of the word edge. Just like it has been shown for the oxytones in our dataset, we would expect that peaks in paroxytones and proparoxytones would also fall in the vicinity of the word edge, and generally within their bounds. We would expect so in, at least, a strong interpretation of the word-edge-tone hypothesis (Estebas 2003).

In order to investigate this, we measured the distance in milliseconds from the pitch peak to the word boundary. Recall that, only for oxytones, this is the same measure as the one previously shown: the distance between the pitch peak and the offset of the stressed syllable. For the other two stress configurations, these are two very different measures. Figure 4 is a display of the data regarding this measure. In Figure 4, 0 in the y-axis stands for the location of the word boundary. For oxytones, it can be seen that peaks fall within the bounds of the word-stretch, and close to the word edge. However, the relevant fact is that in both paroxytones and proparoxytones peaks tend to fall farther to the left as the distance from the onset of the stressed syllable to the word boundary increases. In other words, even though peaks move further to the right (peak delay) when the distance between the stressed syllable's onset and the word edge increases, they do not seem to go all the way down to the word offset.

The data were submitted to 6 independent one-way ANOVAs, one for each subject, to make sure that the displays reflect statistically significant differences. The dependent variable was distance between the peak and the word edge in milliseconds, and the main factor was, again, lexical stress configuration (3 levels). The results of the ANOVAs show that these measurements are statistically significant for all 6 speakers [1M, $F(2,204)=77.4$, $p<.001$; 2M, $F(2,184)=110.32$, $p<.001$; 3M, $F(2,177)=60.43$, $p<.001$; 4F, $F(2,192)=26.16$, $p<.001$; 5F, $F(2,167)=15.96$, $p<.001$; 6F, $F(2,200)=81.3$, $p<.001$]. Post-hoc tests using Tukey HSD revealed three main patterns: For two speakers (1M, 6F), all three groups are significantly different from each other; for three subjects (2M, 3M, 4F), oxytones are significantly different from the other two, while paroxytones are not different from proparoxytones; and for one speaker (5F), only the two extremes are significantly different from each other.

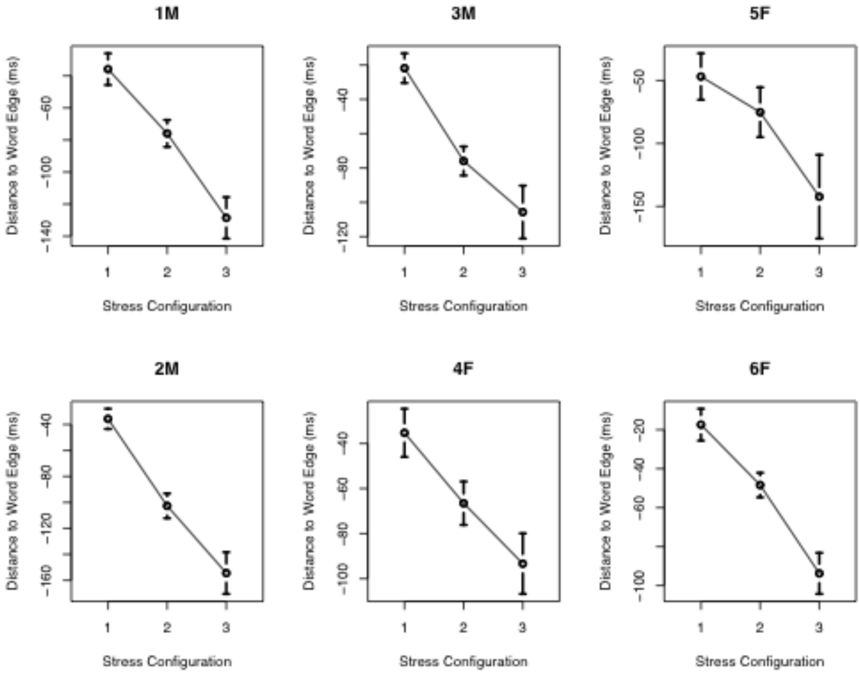

Figure 4. Error-bar plots (95% confidence interval) of distance of peak to the word edge (ms) as a function of lexical-stress configuration of words: 1 (oxytones), 2 (paroxytones) and 3 (proparoxytones) for 6 Spanish speakers.

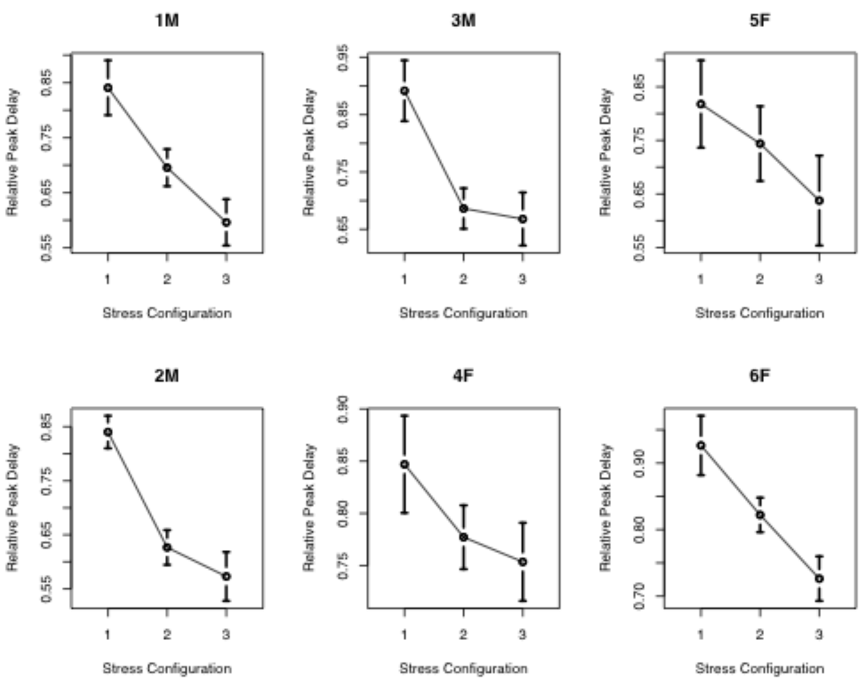

Figure 5. Error-bar plots (95% confidence interval) of relative peak delay as a function of lexical-stress configuration of words: 1 (oxytones), 2 (paroxytones) and 3 (proparoxytones) for 6 Spanish speakers.

These data seem to present evidence against a strong interpretation of the word-edge-tone hypothesis since peaks do not consistently align with respect to the word boundary. However, one could argue that, even though they may not reach the end of the word or word edge in their displacement to the right, it is still possible that peaks are consistently placed *with respect to* a certain predetermined proportion or percentage of the word-stretch duration. The absolute values we have shown so far could be obscuring the possibility for peaks to be consistently located at, say, 60%, 75% or 80% of the word-stretch duration. In order to capture this possibility in quantitative terms, we used a relative measure of peak delay that takes into account the duration of the word stretch (Myers 2003). This measurement results from dividing the distance value between the onset of the tonic and the pitch peak by the distance values between the onset of the tonic and the word edge.

Since ratio values are not well suited for statistical analyses, ANOVAs were not possible for these measurements. However, by eyeballing the means of the ratio values in Figure 5, it can be seen that the picture clearly resembles that in Figure 4, where absolute values are shown. In other words, peaks do not seem to be located at a given proportion or percentage of the word-stretch (the distance between the stressed syllable's onset and the word edge). Although the differences between oxytones, paroxytones and proparoxytones are not as large as for the absolute values shown in Figure 4, they seem to be large enough. Oxytones seem to have their pitch peaks at about 85% to 95% of the word-stretch, while paroxytones seem to have theirs at 65% to 75%, and proparoxytones, at 60% or 70%.

In summary, even though word-boundary effects on pitch peak timing have been found, it cannot be claimed that these effects fall from the fact that all peaks align with respect to word edges. For instance, peaks in proparoxytones are maximally displaced to the right. However, in these, peak delay does not trigger peaks to go all the way down to the word edge.

4. Discussion and conclusions

This study of pitch timing in Spanish has arrived at the following conclusions for a group of 6 speakers of the variety spoken in the western Mediterranean island of Majorca, Spain: (*i*) Peak delay varies as a function of lexical-stress configuration; (*ii*) Peaks in oxytones fall within the bounds of the tonic syllable, thus contrasting with previous results such as those in Prieto and Estebas (2005) and de la Mota (2005); (*iii*) The distance from the peak to the word edge also increases as a function of lexical-stress configuration: That is, while peak delay is greatest for proparoxytones, distance to the word edge is also greatest for these; (*iv*) A measure of relative peak delay shows that peaks are not consistently located at a precise proportion of the duration of the word (or, better, the time distance between the stressed syllable onset and the word edge). Due mainly to inter-speaker differences, the conclusions we just outlined deserve a deeper explanation.

Regarding the effects of lexical-stress configuration on peak timing, it seems that the most robust pattern, the one that holds across speakers, is that oxytones and proparoxytones are always significantly different. Pitch peaks in oxytones are quite consistently aligned with respect to the word edge and usually fall within the bounds of the tonic syllable. In proparoxytones, peaks are most displaced to the right but they do not go as far to the right as to the word boundary. Paroxytones, however, show a more variable behavior. With respect to peak delay (section 3.1), four patterns have been found: oxytones < paroxytones < proparoxytones (4F, 6F); oxytones < paroxytones = proparoxytones (5F); oxytones = paroxytones < proparoxytones (2M, 3M); oxytones < proparoxytones (1M). The results for the time distance between peaks and word edges (section 3.3) show similar patterns, although across different clusters of speakers: oxytones < paroxytones < proparoxytones (1M, 6F); oxytones < paroxytones = proparoxytones (2M, 3M, 4F); oxytones < proparoxytones (5F). Paroxytones, thus, seem to pattern either on their own or with one of the other two configurations. Above all, however, the ANOVAs show that the position of pitch accents with respect to the word has a highly significant effect on both peak delay and the distance between the peak and the word boundary. The effects are gradient and vary as a function of speaker, but they are still significant.

Our results seem to both support and contradict a proposal in which peaks are aligned with respect to word edges. We cannot accept a strong interpretation of the word-edge-tone hypothesis due to the fact that, while word boundaries exert a strong effect on pitch timing, peaks do not move all the way down to the word edge – except in oxytones. At the same time, however, the word-boundary effects we have shown in the present article would remain completely unexplained without any sort of reference to the lexical-stress configuration of words.

The results may suggest the existence of two competing attractors for peak alignment: the stressed syllable's offset and the word edge. For oxytones, the two competing attractors are located in exactly the same position and, thus, exert a strong influence on peak timing. This strong attraction would be responsible for blocking the delay of peaks to the post-tonic syllable in the oxytones in our corpus. For paroxytones and proparoxytones, the two competing attractors are found in different time positions. This could be the reason why peaks are located halfway between stressed syllable offsets and word boundaries. However, we are not claiming that the middle position between the two time points is an optimal position that acts as the main target or attractor. We are claiming that speakers may reach a compromise position in which they decide to place peaks in a broad area that falls within the two attractors. The two competing attractors may be in fact secondary associations (Prieto *et al.* 2005).

For findings like those in Prieto and Estebas (2005), in which peaks are displaced to the post-tonic syllable even in oxytones, a different explanation is needed. It is certainly possible that the previous valley exerts a strong influence on the peak and that this causes the peak to be displaced to the right, in order to allow for an optimal articulation of the rise. The lack of coda consonants in de la Mota's (2005) and Prieto and Estebas's (2005) data may have provoked a reorganization of the rising gesture that triggers peak delay, since the time range between the onset and the offset of the gesture (which may coincide with the syllable's onset and offset) may have been too compressed. Laryngeal gestures may exert an influence on each other that is stronger than that exerted by supralaryngeal gestures.

The extent to which our findings can be extended to other Iberian Romance varieties constitutes an open empirical question. An empirically-grounded theoretical interpretation for such results is yet to be found. We believe that our experiment has addressed a topic of Romance intonation that still deserves more investigation. The present study may have provided more questions than answers.

References

Arvaniti, Amalia, Robert Ladd and Ineke Mennen (1998) Stability of tonal alignment: the case of Greek prenuclear accents. In *Journal of Phonetics* 26: 3-25.

Boersma, Paul and David Weenink (2005) Praat: a system for doing phonetics by computer. [Computer program].

de la Mota, Carme (2005) Alignment, word boundaries and speech rate in Castilian Spanish. Poster presented at PaPI-2. June 20-21. Barcelona, Spain.

Estebas, Eva (2003) The modeling of pre-nuclear accents in Central Catalan declaratives. In *Catalan Journal of Linguistics* 2: 97-114.

Face, Timothy (2002) Intonational marking of contrastive focus in Madrid Spanish. Munich: Lincom Europa.

Flege, James E. and Ian MacKay (2004) Perceiving vowels in a second language. In *Studies in Second Language Acquisition*, 26: 1-34.

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

Hualde, José Ignacio (2002). Intonation in Spanish and the other Ibero-Romance languages: overview and status quaestionis. In Caroline Wiltshire and Joaquim Camps (eds.) *Romance phonology and variation*. Amsterdam: John Benjamins. 101-116.

Myers, Scott (2003) F_0 timing in Kinyarwanda. In *Phonetica* 60: 71-97.

O'Rourke, Erin (2005) Anchoring the H: correlates of peak alignment in Peruvian Spanish intonation. Paper presented at the 9th Hispanic Linguistics Symposium. November 11-13. PennState University.

Prieto, Pilar (*forthcoming*) Word-edge tones in Catalan. In *Italian Journal of Linguistics*. 18.1.

Prieto, Pilar, Mariapaola d'Imperio and Barbara Gili-Fivela (2005) Pitch accent alignment in Romance: Primary and secondary associations with metrical structure. In *Language and Speech*, vol. 48, 4: 356-396.

Prieto, Pilar and Eva Estebas (2005) The role of word-edge tones in Catalan and Spanish. Poster presented at PaPI-2. June 20-21. Barcelona, Spain.

Prieto, Pilar, Jan Van Santen and Julia Hirschberg (1995) Tonal alignment patterns in Spanish. In *Journal of Phonetics* 23: 429-451.

Silverman, Kim and Janet Pierrehumbert (1990) The timing of pre-nuclear high accents in English. In Kingston, John and Mary Beckman (eds.) *Papers in Laboratory Phonology I*. Cambridge University Press: 72-106.

Simonet, Miquel and Francisco Torreira (2005) Word-boundary effects on F_0 timing in Spanish: spontaneous vs. laboratory speech. 2nd Sp-ToBI Workshop. June 22. Barcelona, Spain.

Selected Proceedings of the 9th Hispanic Linguistics Symposium

edited by Nuria Sagarra and Almeida Jacqueline Toribio

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

ISBN 1-57473-413-X library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information