Available online at www.sciencedirect.com
ScienceDirect
Journal of Phonetics 35 (2007) 473–500
Journal of Phonetics
www.elsevier.com/locate/phonetics

The segmental anchoring hypothesis revisited: Syllable structure and speech rate effects on peak timing in Spanish

Pilar Prieto[a,*], Francisco Torreira[b]

[a] Department Filologia Catalana, Institució Catalana de la Recerca i Estudis Avançats (ICREA), Universitat Autònoma de Barcelona (UAB), Edifici B, 08193 Bellaterra, Barcelona, Spain
[b] Université Libre de Bruxelles/University of Illinois at Urbana-Champaign, USA

Received 20 February 2006; received in revised form 28 December 2006; accepted 16 January 2007

Abstract

This paper addresses the validity of the *segmental anchoring hypothesis* for tonal landmarks (henceforth, SAH) as described in recent work by (among others) Ladd, Faulkner, D., Faulkner, H., & Schepman [1999. Constant 'segmental' anchoring of f0 movements under changes in speech rate. *Journal of the Acoustical Society of America*, *106*, 1543–1554], Ladd [2003. Phonological conditioning of f0 target alignment. In: M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences*, Vol. 1, (pp. 249–252). Barcelona: Causal Productions; in press. Segmental anchoring of pitch movements: Autosegmental association or gestural coordination? *Italian Journal of Linguistics*, *18* (1)]. The alignment of LH* prenuclear peaks with segmental landmarks in controlled speech materials in Peninsular Spanish is analyzed as a function of syllable structure type (open, closed) of the accented syllable, segmental composition, and speaking rate. Contrary to the predictions of the SAH, alignment was affected by syllable structure and speech rate in significant and consistent ways. In: CV syllables the peak was located around the end of the accented vowel, and in CVC syllables around the beginning-mid part of the sonorant coda, but still far from the syllable boundary. With respect to the effects of rate, peaks were located earlier in the syllable as speech rate decreased.

The results suggest that the accent gestures under study are synchronized with the syllable unit. In general, the longer the syllable, the longer the rise time. Thus the fundamental idea of the anchoring hypothesis can be taken as still valid. On the other hand, the tonal alignment patterns reported here can be interpreted as the outcome of distinct modes of gestural coordination in syllable-initial vs. syllable-final position: gestures at syllable onsets appear to be more tightly coordinated than gestures at the end of syllables [Browman, C. P., & Goldstein, L.M. (1986). Towards an articulatory phonology. *Phonology Yearbook*, *3*, 219–252; Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*, 140–155; (1992). Articulatory Phonology: An overview. *Phonetica*, *49*, 155–180; Krakow (1999). Physiological organization of syllables: A review. *Journal of Phonetics*, *27*, 23–54; among others]. Intergestural timing can thus provide a unifying explanation for (1) the contrasting behavior between the precise synchronization of L valleys with the onset of the syllable and the more variable timing of the end of the f0 rise, and, more specifically, for (2) the right-hand tonal pressure effects and 'undershoot' patterns displayed by peaks at the ends of syllables and other prosodic domains.
© 2007 Elsevier Ltd. All rights reserved.

*Corresponding author. Tel.: +34 11 34 93 225 48 99; fax: +34 11 34 93 581 27 82.
E-mail address: pilar.prieto@uab.es (P. Prieto).

0095-4470/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wocn.2007.01.001

1. Introduction

Early work on tonal alignment has acknowledged the contrasting behavior between L valleys and H peaks in rising accents in a variety of languages such as Catalan, Dutch, English, Spanish, and Greek. While L turning points are consistently 'anchored' to the onset of the accented syllable in extremely consistent ways, H positions are found to be more variable (Caspers & van Heuven, 1993 for Dutch; Prieto, van Santen, & Hirschberg, 1995 for Spanish; Arvaniti, Ladd, & Mennen, 1998 for Greek; Ladd, 2003; Ladd, Faulkner, Faulkner, & Schepman, 1999 and Ladd, Mennen, & Schepman, 2000 for English; Xu, 1998 for Mandarin Chinese; Estebas-Vilaplana, 2000 for Catalan). Some of these studies have emphasized the role of syllable duration and time pressure from the right-hand prosodic context (i.e., the proximity of an upcoming accent or boundary tones) in determining the location of H peaks (Bruce, 1977; Caspers & van Heuven, 1993; Silverman & Pierrehumbert, 1990; Prieto et al., 1995). However, more recent work has suggested that when such right-hand prosodic effects are excluded (i.e., when the tonal features under investigation are not in the vicinity of pitch accents or boundary tones), the alignment of f0 peak targets is consistently governed by segmental anchoring and that strict alignment effects are pervasive under changes of syllabic/segmental structure and speech rate (Arvaniti et al., 1998 for Greek; Ladd et al., 1999 for English; Xu, 1998 for Chinese; see also Schepman, Lickley, & Ladd, 2006 for Dutch and Atterer & Ladd, 2004, for German).

The source of the segmental anchoring hypothesis (henceforth, SAH) was the study by Arvaniti et al. (1998). The original goal of this research was to study the phonetic manifestation of the LH prenuclear accent in Greek, a pitch accent similar to its Spanish counterpart in the variability displayed by its H tone. However, an unexpected and consistent stability of alignment was found when little or no tonal pressure was exerted on the pitch accent studied. This stability of alignment, which was interpreted as a sort of tonal 'anchoring', followed much simpler rules than those governing peak delay under tonal pressure. In sentences such as [tesera pa ranoma no mizmata vre ikan stin kato ʎi tu] 'Four illegal coins were found in his possession', the H target in the LH pitch accent associated with the test stressed syllable ['ra] was consistently aligned over the frontier between the postaccentual onset and rhyme ([n] and [o]).

The idea behind the SAH is that both the beginning and the end of a rising pitch accent are anchored to specific landmarks in the segmental structure, regardless of segmental or syllable structure composition, and regardless of speaking rate. In order to replicate these findings, Ladd and his collaborators conducted experiments on English (1999) and Dutch (2000). For example, Ladd et al. (1999) showed that the alignment of the valley and peak of an English rising pitch accent is unaffected by changes in segmental duration brought about by modifications of speech rate. The duration and slope of the accents become shorter and steeper as rate increases. Thus English rising prenuclear accents remain anchored to segmental landmarks regardless of speaking rate. Other studies have partially confirmed the predictions of the anchoring hypothesis (Ishihara, 2003 for Japanese; Igarashi, 2004 for Russian; Xu, 1998 for Mandarin Chinese), suggesting that segmental anchoring is a universal phenomenon. For example, Xu (1998) shows that the end of a rising contour in Mandarin Chinese is synchronized with the end of the syllable, regardless of speech rate and syllable composition. In line with this hypothesis, recent studies have also shown very subtle differences in alignment for dialectal varieties within a single language. In Atterer and Ladd (2004), speakers of northern and southern varieties of German align the L and the H targets in a LH accent at points that differed by a few ms. Moreover, in both varieties of German, L and H targets were aligned later than in English.

However, contradictory evidence is found in the literature with respect to the predictions of the SAH. On the one hand, work on different languages in different languages has shown that the position of the peak changes consistently across syllable structures. For example, D'Imperio (2000) found that the peak was located closer to the vowel offset in closed syllables in Neapolitan Italian (see also Prieto, 2006b). While in open syllables the peak was aligned with the end of the accented vowel, in closed syllables the peak was somewhat retracted and located within the coda consonant. The same pattern is found by Gili-Fivela and Savino (2003) for Pisa and Bari Italian, Hellmuth (2005, 2006) for Egyptian Arabic, and Welby and Lœvenbruck (2005, 2006) for the late rise in French. Similarly, effects of segmental composition (onset and coda type) have been found on H alignment in English (van Santen & Hirschberg, 1994), Dutch (Rietveld & Gussenhoven, 1995), and French (Welby & Lœvenbruck, 2005, 2006). Finally, with respect to the effects of speaking rate on f0 peak alignment, several studies have found a significant effect of speech rate on peak alignment (Xu, 1988; Ishihara, 2006). Given the inconsistent evidence reported in the literature, some authors have proposed a weakened SAH. Welby and Lœvenbruck (2005, 2006) propose the notion of an "*anchorage*", that is, a temporal range within which an intonational turning point can anchor. For the peak of the French late rise, this *anchorage* stretches from just before the end of the vowel of the last full syllable of the accentual phrase to the end of the phrase.

One possible source of the contradictory results reported in the literature is that they come from studies that were not specifically designed to test the predictions of the SAH; in other words, in these studies the effects of syllable structure and speech rate on H alignment are described together (and sometimes confounded) with the influence of other factors. In order to clarify the situation, it is important to conduct studies designed to test the predictions of this hypothesis. The main goal of the present study is to examine the stability of alignment of prenuclear LH* peaks in Castilian Spanish, and the effects of the two factors that are predicted to have no effect on H alignment, namely: (i) syllable structure; and (ii) speech rate. Castilian Spanish is a particularly well-suited language to study the effects of syllable structure on f0 peaks, as sonorant codas like nasals, rhotics, and laterals are the most common codas in this language.

Two experiments were carried out based on alignment data from 3 speakers of Peninsular Spanish. Experiment 1 was specifically designed to examine the potential effects of syllabic structure and segmental composition on H alignment. The second experiment dealt primarily with the combined effects of syllable structure and speaking rate. If the predictions of the SAH are true and there are segmental anchors for H targets in rising accents, one would expect that the position of the peak should be unaffected by changes in syllable structure and segmental composition (Experiment 1), and by changes in speech rate (Experiment 2).

2. Experiment 1

The first experiment has the goal of analyzing the potential effects of syllable structure and coda type on H timing in prenuclear (non-sentence-final) accents in Spanish. The study addresses the validity of two alternative hypotheses:

(i) In Spanish, prenuclear peaks are consistently aligned with segmental or syllabic landmarks when external pressure effects (e.g. tonal clash) are not present. More precisely, in the absence of tonal pressure, changes in syllabic composition of the test syllable do not modify the positioning of the f0 peak with respect to its anchoring point. These are the predictions of the SAH.
(ii) In Spanish, prenuclear peaks are not anchored with reference to segmental landmarks. They show an invariant timing with respect to the preceding L valley. This is the prediction of the invariant rise hypothesis (IRH).

2.1. Methodology

The target H peak, corresponding to the prenuclear pitch accent and belonging to the target word *lánguido* 'languid', is illustrated in Fig. 1. Pilot recordings of several speakers had suggested that the first prenuclear peak of this contour, which is typical of reading style, was consistently produced towards the end of the stressed vowel (after CV in open syllables and after CVC in closed syllables), and not in the postaccentual syllable, as is usually the case in prenuclear rising accents in different Spanish dialects (Prieto et al., 1995; Face, 2002; Sosa, 1999; Face & Prieto, 2007). This accent is referred to with the label LH*, as opposed to L*H, which is the standard ToBI notation for the prenuclear accent with a peak in the postaccentual syllable. Crucially, this anchored peak is obtained when the first accent belongs to an utterance-initial phrase which contains two accents. Even though this type of prenuclear rise is not among the most common in Spanish, it was selected because it displayed less variability in alignment than the more usual rise with a displaced peak, where no obvious segmental boundary could be regarded as a segmental anchor. This target peak appeared therefore to be an optimal candidate for testing the anchoring hypothesis.

For both experiments, speech materials were designed in such a way as to control relevant structural and phonetic variables in the target accented syllable. After that, Spanish speakers read the materials in laboratory

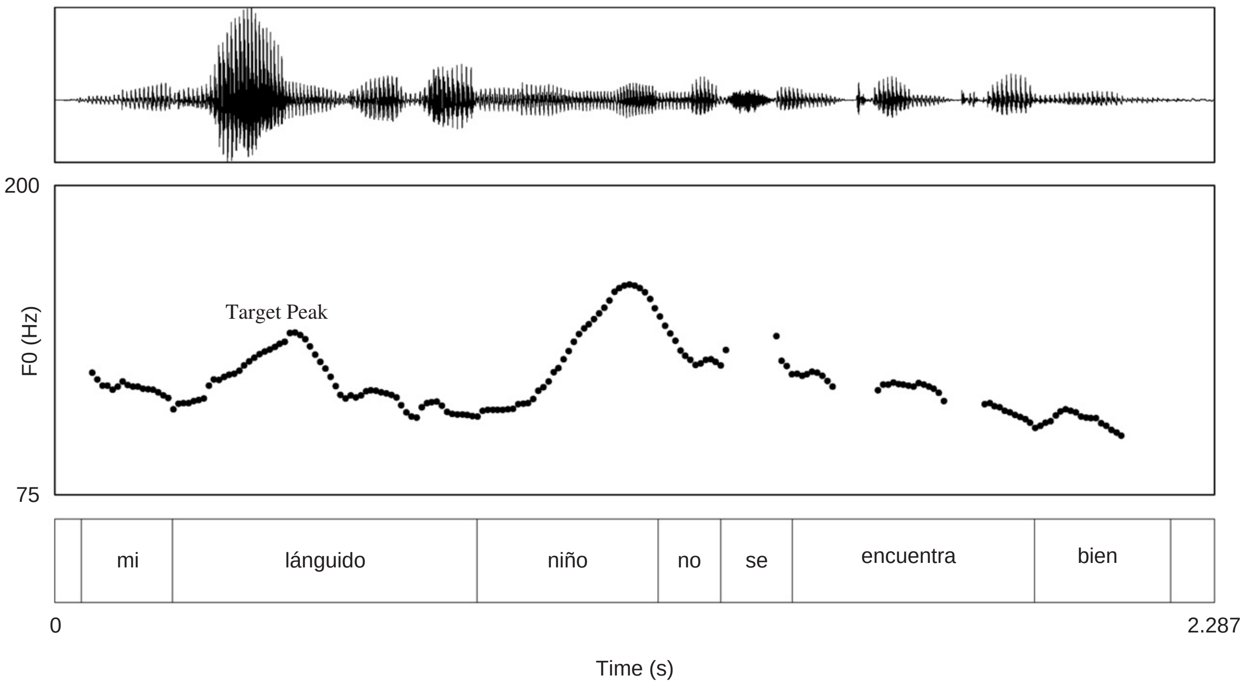

Fig. 1. Waveform and f0 contour of the utterance *Mi lánguido niño no se encuentra bien* 'My languid child does not feel well' produced by speaker SB. The test H peak corresponds to the first f0 maximum.

recording conditions and the relevant acoustic measurements were performed (namely, the alignment of specific f0 points in reference to segmental locations).

2.1.1. Materials

A corpus of 48 sentences was designed to exhibit the phonological variety needed to test our hypotheses. The test sentences all met the following criteria:

(i) The test word was generally an adjective followed by a noun or a noun followed by an adjective plus a verbal phrase (e.g., *Mi* ***lánguido*** *niño no se encuentra bien* 'My languid child does not feel well', *La* ***lámina*** *blanca está en la mesa* 'The white foil is on the table'). This typically ensured that, in normal reading, utterances were produced in two prosodic phrases.

(ii) Special attention was paid to maintaining the same number of potential accents across the carrier sentences. Utterances were produced in two intonational phrases containing two pitch accents each. The first phrase ended in a H^- phrase accent, a clear acoustic cue to the end of the prosodic phrase. The target peak was the first peak of the first prosodic phrase (see Fig. 1).

(iii) Two accentual patterns were chosen for the test words: paroxytone and proparoxytone, i.e. words with the stress on the penultimate and the antepenultimate syllables, respectively.

(iv) The test syllable was preceded by one unstressed syllable. Also, in order to exclude potential pressure effects from upcoming tonal targets, test syllables were followed by two unstressed syllables.[1] Following Arvaniti et al. (1998) for her study of Greek prenuclear accents, it is assumed that two unaccented postaccentual syllables provide enough prosodic space to allow for an anchored alignment.

[1]There are just two exceptions in the number of accents following the test syllable. One is the sentence containing the words '*nódulo central*', and the other one contains the words '*Tablones López*'. In these two utterances, there is one unstressed syllable following the target accent. As each of these exceptions belongs to one of the two word types in the experiment (paroxytones vs. proparoxytones), its potential effect is counterbalanced across the two data sets.

(v) The test syllables displayed various syllabic structures (open vs closed; complex vs simple onset) as well as different consonant types and vowels (high: /i/, /u/ vs mid/low: /a/, /o/). To minimize microprosodic perturbations, only sonorant consonants were used in coda position (nasals or laterals) and only voiced consonants were used in onset position.
(vi) In a small number of cases, nonce words were used in order to fill lexical gaps in our database. Nonce words are indicated with an asterisk in the Appendix A.

The total database counts were as follows: 6 segmental/syllabic structures $\times$ 4 vocalic rhymes $\times$ 2 accentual patterns (paroxytones/proparoxytones) $\times$ 3 repetitions = 144 utterances per speaker. The complete list of sentences in Experiment 1 can be found in Appendix A.1.

2.1.2. Subjects

Three female speakers of Castilian Spanish (ES, RG and SB) read the set of test sentences three times at a normal speaking rate, for a total of 432 utterances (3 speakers $\times$ 48 sentences $\times$ 3 repetitions). The speakers were from Northern Castile (RG) and Central Castile (ES and SB), aged between 20 and 40, and were university teachers or students. Their participation was voluntary and did not imply any kind of compensation.

2.1.3. Recording procedure

Speakers were recorded by the second author on professional equipment in a sound-attenuated booth in the Universidad Nacional de Educación a Distancia (UNED, Madrid). The utterances were initially recorded on a MiniDisc and were later digitized at a sampling rate of 44100 Hz.

Speakers were instructed to read the sentences naturally and at a normal rate of speech. Since several potential contours could be applied to the same text, readers were briefly instructed to produce the required contour before the recording session took place. A small number of random sentences from the corpus were used in this task. The sentences were presented in random order. The recordings were carefully monitored. After each set, speakers were asked to reproduce any sentence showing any type of disfluency or unwanted phrasing or contour. The materials for the two experiments were recorded in separate sessions, with a 1-day interval in between. Each session lasted about 4 h.

After recording, sentences were prosodically monitored to check that each sentence was produced in two prosodic phrases and that the test word was produced with a prenuclear rising accent followed by a nuclear accent on the following noun. Several utterances were found to be not usable for analysis and were recorded again in a different session. In most cases, the problem was that speakers produced a continuous f0 rise from the accented syllable up to the H$^-$ intermediate phrase boundary after the second content word, without a clear valley between the two rises.

2.1.4. Segmentation

Following previous research on f0 alignment, the following measurements in ms were taken. Note that segment labels (based on Arvaniti et al., 1998; Atterer & Ladd, 2004; Schepman et al., 2006, among others) represent the start of the segment and that the number stands for its corresponding syllable (accented = 0, postaccentual = 1):

(i) **c0** beginning of the initial consonant in the accented syllable;
(ii) **v0** beginning of the vowel in the accented syllable;
(iii) **k0** beginning of the coda in the accented syllable;
(iv) **c1** beginning of the initial consonant in the postaccentual syllable;
(v) **L** f0 valley location in LH accent;
(vi) **H** f0 peak location in LH accent.

Fig. 2 illustrates the labeling scheme used for open and closed syllables.

The acoustic measurements were made on simultaneous display of waveform, wide-band spectrograms and f0 tracks using Praat 4.2 (Boersma & Weenink, 2005). In general, H's were placed directly at f0 maxima for

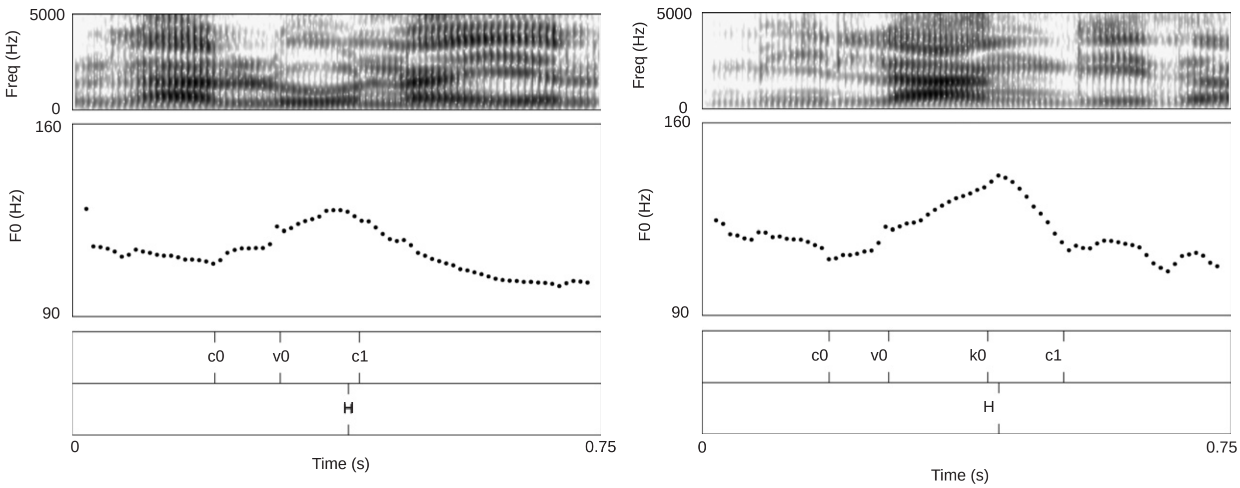

Fig. 2. Waveform display, f0 contour, and labeling scheme for two test words, one containing an open accented syllable (left) and the other a closed accented syllable (right), produced by speaker RG.

salient f0 peaks. However, in some cases the identification of peaks was not trivial, for example, when the H points formed a plateau where no clear f0 value emerged as the highest. Following research on f0 peak perception and tonal plateaux by D'Imperio (2000) and Knight (2003), in these cases, the H point was marked at the falling elbow, where it is estimated to be perceived by hearers. Microprosodic effects (such as the typical dip produced by nasal segments) were disregarded. With regards to the location of segmental boundaries across vowels and sonorants [m n l] (which constitute around 90% of the segmentation cases) standard segmentation procedures were followed (Peterson & Lehiste, 1960). The beginning or end of a sonorant consonant was identified at the start of the abrupt change from the steady-state period in the spectrogram to the onglide transition movement to the vowel. When the formant transitions were not abrupt enough, the criterion used was the expected change in amplitude displayed in the waveform.

The labels were used for segment boundaries and f0 landmarks described above to derive dependent variables for expressing peak alignment. Our choices were based on past work and on our own hypotheses about the alignment behavior of Spanish prenuclear peaks. During preliminary visual inspection of the data, two segmental landmarks in particular—the end of the accented syllable and the end of the accented vowel—had emerged as possible anchor points for H peaks. Since open and closed syllables were included in our database, a special variable had to be created for marking the end of the stressed vowel across the two syllabic structures with the same label: EndV0. Therefore, the three alignment variables used in our analysis were the following, which all express time intervals in milliseconds:

C0toH	Distance (in ms) from the start of the accented syllable to the location of the H peak, or *peak delay*.
HtoEndV0	Distance from the H peak to the end of the accented vowel V0.
HtoEndSyll	Distance from the H peak to the end of the syllable (or HtoC1).

Finally, in addition to the three alignment measures, segment durations were also calculated, as control measures.

2.1.5. *Statistical analyses*

Three different measures of H location were used for statistical exploration, namely (i) C0toH, (ii) HtoEndV0, and (iii) HtoEndSyll. Note that for open syllables, HtoEndV0 and HtoEndSyll had the same values. Following recent methodological arguments by Schepman et al. (2006) and Atterer and Ladd (2004), our analyses were primarily based on the latter two ways of expressing peak alignment, namely, in reference to a nearby acoustic landmark (EndV0 or EndSyll), even though distant ones (C0) were also taken into account.

In order to check for individual effects, a series of univariate analyses of variance were performed separately for each speaker and for each independent variable, namely, Syllable Type, Stress Pattern, Coda Type, Onset Length and Vowel Type. Each variable had two possible values, as follows: Syllable Type (open vs. closed), Stress Pattern (proparoxytonic vs. paroxytonic words), Coda Type (nasal vs. lateral), Onset Length (complex vs. simple) and Vowel Type (high vs. low–mid, since low and mid vowels are significantly longer in duration than high vowels). After that, in order to model the combined effects of all factors and to observe the interactions between independent variables, a univariate linear model was run in which the independent factors were Syllable Type, Stress Pattern, Coda Type, Onset Length, Vowel Type, and Speaker.

2.2. Results

2.2.1. Effects of syllable duration

Previous studies of peak alignment have shown that onset/rhyme duration are key factors in the prediction of f0 peak location and that, typically, a high correlation between peak delay and duration of the stressed syllable is found (Silverman & Pierrehumbert, 1990 for English; Prieto et al., 1995 for Spanish; Arvaniti et al., 1998 for Greek). The three scatterplots in Fig. 3 show the peak delay measures (C0toH in ms) for the three speakers as a function of syllable duration in two syllable type conditions: open syllables (represented by a triangle) and closed syllables (represented by a circle). As shown by the three graphs in Fig. 3, the degree of correlation between the two variables is moderately high (R^2 coefficients range from 0.33 to 0.5, all significant below the 0.01 probability level). Even though in neither case correlation coefficients were comparable to those obtained for Greek in Arvaniti et al. (1998) or Prieto et al. (1995), it is clear that the fixed rise time hypothesis cannot be maintained for the Spanish data—for more details, see Torreira, 2004 and also Sections 3.2.5 and 4.1. Fig. 3 also reveals a potential effect of syllable structure. For speakers RG and SB, there are two distinct clouds corresponding to each syllable structure condition: while peaks in open syllables (triangles) are located close to the syllable boundary (solid line in the graph), peaks in closed syllables (circles) are retracted into the syllable. As might be expected, closed syllables are longer in duration than open syllables. However, this difference in duration across syllable structures does not seem to be correlated with an increase in C0toH. Even for speaker ES (left graph), who displays a more homogeneus alignment in the two syllable structure conditions, the correlation coefficient obtained for her utterances ($R^2 = 0.507$) is not high enough to support a strict alignment of f0 peaks at the end of accented syllables. In conclusion, the end of the syllable does not seem to be an anchoring point for our target f0 peaks. However, Fig. 3 leaves open the possibility that for speakers RG and SB f0 peaks might be aligned with the end of the accented vowel in both open and closed syllables, since f0 peaks are clearly retracted in closed syllables. This question is addressed in the following section.

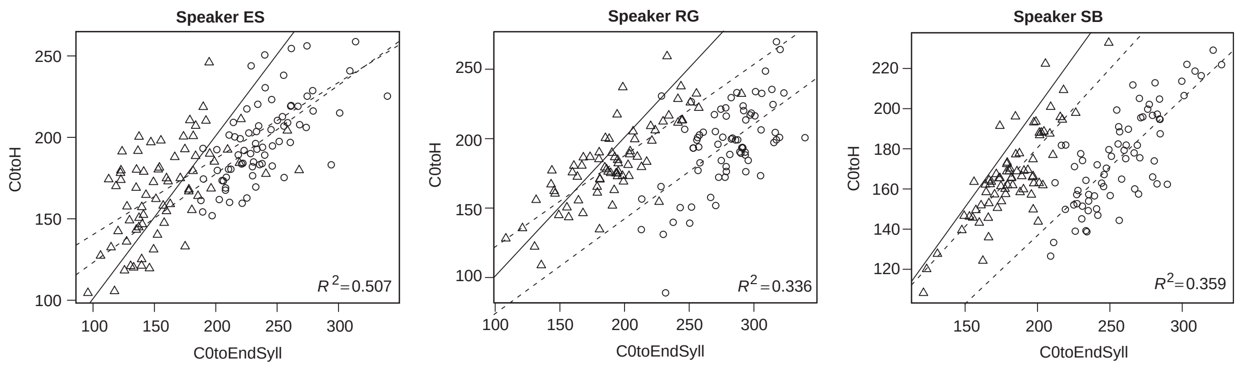

Fig. 3. Peak delay or C0toH (in ms) for the three speakers as a function of syllable duration in two syllable type conditions: open syllable (represented by a triangle) and closed syllables (represented by a circle). Solid lines represent the temporal position of syllable boundaries ($x = y$), while dashed lines are regression lines for each syllable type.

2.2.2. Effects of syllable structure

Fig. 3 seemed to support our preliminary observations that f0 peaks were aligned with the end of the stressed vowel for both open and closed syllables. Therefore syllable structure effects were examined in more detail. First, the effects of syllabic structure on vowel duration were analyzed, since it might be suspected that closed syllables would induce vowel shortening. Surprisingly, the small mean duration differences found were not significant for any speaker. Only speaker ES verged on significance ($p = 0.058$), with open syllables displaying longer vowels than closed syllables (average difference of 4 ms). After that, a comparison of H alignment patterns across the two syllabic structure conditions was undertaken.

In order to test the SAH, the variables HtoEndSyll (distance in ms from the H peak to the end of the syllable) and HtoEndV0 (distance in ms from the H peak to the end of the accented vowel) were singled out as the independent quantitative variables, following recent demonstrations by Schepman et al. (2006, pp. 22–23) that "the most appropriate quantitative variables for expressing f0/segmental alignment are those that define alignment as the time interval between the f0 target in question and a nearby segmental landmark; the more distant the landmark, the greater the variance". The two graphs in Fig. 4 plot mean HtoEndV0 and HtoEndSyll measures as a function of syllable structure for each of the speakers. Negative numbers indicate alignment before the relevant segmental frontier, while positive numbers indicate alignment after this landmark. The HtoEndSyll values (left graph) confirm the pattern illustrated in Figs. 2 and 3, namely that f0 peaks are retracted into the syllable in closed syllables. The HtoEndV0 values (right graph) reveal a more subtle effect: when measured with reference to the end of the stressed vowel, our hypothesized anchoring point, closed syllables exhibit a later alignment than open syllables. The mean duration difference of this effect, applied to all three speakers, was 18 ms.

The differences observed in Fig. 4 were statistically confirmed by one-way ANOVAs run separately for each speaker. Table 1 shows ANOVA summaries of the effects of syllable type (ST) on two measures of H location, namely, HtoEndSyll and HtoEndV0. The effects of ST were significant at $p < 0.001$ for both measures and for all speakers.

These data do not support the predictions of the segmental anchoring hypothesis, and reveal effects of syllable structure. Crucially, even though the relative alignment to adjacent segmental landmarks is different for each speaker, the identified syllable structure effects hold in every case. Speaker ES, for instance, tends to

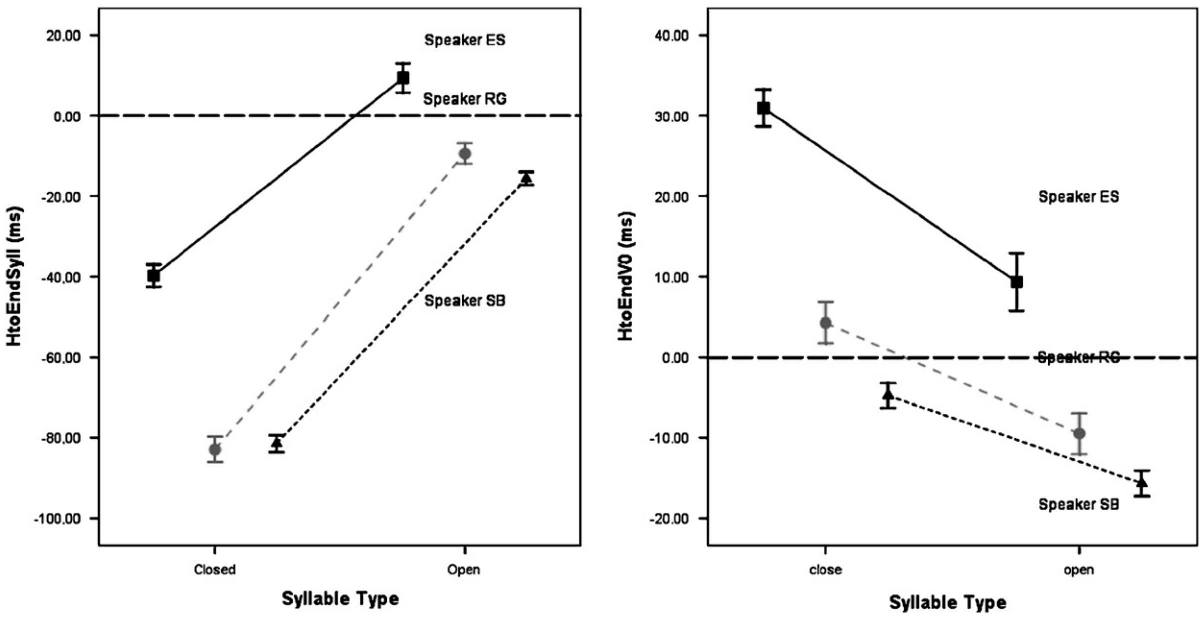

Fig. 4. Mean distance in ms from H to the end of the vowel (HtoEndV0) (right graph) and mean distance in ms from H to the end of the syllable (HtoEndSyll) (left graph) as a function of syllable type (open vs. closed) for the three speakers. The height of the bars represents standard errors.

Table 1
One-way ANOVA summaries of the effects of syllable type (ST) on two measures of H location, namely, HtoEndV0 and Ht0EndSyll

Speaker	HtoEndSyll/ST	HtoEndV0/ST
ES	$F(1,142) = 115.81$, $p < 0.001$	$F(1,142) = 28.555$, $p < 0.001$
RG	$F(1,142) = 335.64$, $p < 0.001$	$F(1,142) = 14.807$, $p < 0.001$
SB	$F(1,142) = 635.94$, $p < 0.001$	$F(1,142) = 24.356$, $p < 0.001$

Table 2
Mean onset/vowel duration values (in ms) across different conditions and results of two-tailed *t*-tests

Segment	Groups	Mean dur. (ms)	Diff. (ms)	*t*-test
C0	Complex onset	117.00	39.50	$t = 20.472$, df $= 395$
	Simple onset	77.50		$p < 0.001$
V0	Mid/low vowel (/o/,/a/)	92.06	8.08	$t = 7.859$, df $= 395$
	High vowel (/i/, /u/)	81.28		$p < 0.001$

place f0 peaks beyond the end of the accented vowel, both for open and closed syllables, whereas speaker SB tends to place f0 peaks within the accented vowel. For all three speakers, however, the distance from H to the end of the syllable (HtoEndSyll) is greater in closed than in open syllables. Similarly, the distance from H to the end of the vowel (HtoEndV0) is different in closed and open syllables for every speaker.

2.2.3. *Effects of segmental composition*

Previous evidence has shown that alignment of peak f0 targets seems to be affected by the duration and identity of intrasyllabic segments. For example, in an analysis of a large corpus of English utterances with nuclear rising pitch accents produced by a single speaker, van Santen and Hirschberg (1994) found that the peak time as measured from the syllable onset and vowel onset could be predicted from the durations of the onset and the vowel in the accented vowel plus constant values depending on coda class. By contrast, coda duration and onset/vowel class did not ameliorate the model significantly.

This section examines whether the difference between branching vs. non-branching onsets has any effect on H alignment, as well as the effects of intrinsic vowel duration or vowel type (high vowels /i,u/ vs. low and low–mid vowels /a, o/). First, complex onsets indeed were found to be longer than singletons and that low–mid vowels were longer than high vowels. Table 2 reports mean onset/vowel duration values in different conditions and related two-tailed *t*-tests.

A separate two-way ANOVA was run for each speaker with HtoEndV0 as a dependent variable and Onset Length and Vowel Type as independent factors. No speaker yielded or even approached significance for any factor, and no interaction was identified. Thus f0 peak alignment relative to the end of the accented vowel was not affected by the duration of onsets and vowels in the accented syllable. The apparent contrast between our results and previous results reported in the literature can be attributed to the different measures used to express peak alignment. As recently argued by Schepman et al. (2006), conclusions may be changed by the choice of dependent variables to express peak alignment. As they note, “in our present state of understanding, the most appropriate quantitative variables for expressing f0/segmental alignment are those that define alignment as the time interval between the f0 target in question and a nearby segmental landmark; the more distant the landmark, the greater the variance”.

2.2.4. *Effects of stress pattern*

Studies of a variety of languages have shown that the position of the accented syllable within the word has a significant effect on the position of the peak in rising prenuclear accents, regardless of the measure of peak position used (see Silverman & Pierrehumbert, 1990 for English; Prieto et al., 1995; de la Mota, 2005; Estebas-Vilaplana &

Prieto, 2007; Simonet & Torreira, 2005; Simonet, 2006 for Spanish; Prieto 2006b for Catalan; Arvaniti et al., 1998 for Greek; Ishihara, 2006 for Japanese). This section checks whether the stress pattern of words or the prosodic distance to final word edges has any influence on the alignment of prenuclear peaks. A two-way ANOVA with HtoEndV0 as the dependent variable and Stress Pattern and Syllable Type as factors was run for each speaker. Syllable Type, which had yielded robust alignment effects, was included in order to check for possible interactions. Following recent findings that peaks are progressively retracted as the accented syllable approaches the end of the word, small differences were found based on Stress Pattern. However, significance was only reached by speaker SB [$F(1,142) = 4.27$, $p<0.05$]. As expected, proparoxytonic test words in this speaker's utterances exhibited a mean later alignment difference of 6 ms with respect to paroxytones. As for the two other speakers, who failed to reach significance, mean HtoEndV0 values also followed the expected trend. No interaction between Stress Pattern and Syllable Type was found for any of the speakers.

2.2.5. Effects of coda type

The two coda consonants present in our test words were coronal /l/ and /n/. First the durations of these two segments was measured in the production of our speakers. Only speaker RG displayed a significant duration difference ($p<0.001$), with /n/ being 16 ms longer than /l/. For the other two speakers, /l/ and /n/ were not significantly different in duration. Separate one-way ANOVAs were run with HtoEndV0 as the dependent variable and Coda Type as the main factor. Here statistical differences were found only for speaker SB [$F(1,70) = 6.47$, $p<0.05$], for whom /l/ and /n/ were not different in duration. The mean difference in H alignment for this speaker was 9 ms, with f0 peaks occurring later for /n/ (mean HtoEndV0 = 7 ms) than for /l/ (mean HtoEndV0 = −2 ms). An effect verging on significance was also found for speaker RG ($p = 0.066$), with a mean difference of 5 ms following the same trend as speaker SB. If it exists, the size of the coda type effect on H alignment is relatively small.

2.2.6. Univariate linear model

In order to evaluate the combined effects of the examined factors, two univariate linear models were run with Syllable Type, Stress Pattern, Onset Length, Vowel Type, and Speaker as independent factors. The dependent variables for each model were (1) HtoEndV0 and (2) HtoEndSyll. In our first model, only Syllable Type and Stress Pattern were statistically significant ($p<0.01$). No other factor was found to be significant, and no interaction was identified. Crucially, Speaker interacted with neither Syllable Type ($p = 0.867$) nor Stress Pattern ($p>0.1$), suggesting that within-speaker variability could not be attributed to interactions with any of these two factors. The results were relatively similar for the second model, with Syllable Type and Stress Pattern as the only significant factors ($p<0.05$ and 0.01, respectively) and no interactions.

In summary, the results of Experiment 1 reveal that f0 peaks in our target pitch accent are 'loosely' aligned around the final part of the accented syllable. For two of our three speakers, H occurred towards the end of the accented vowel, with peaks retracted into the coda consonant for closed syllables and near the syllable end in open syllables. This suggests that the end of the accented vowel might be an anchoring point for H peaks. However, a consistent effect of Syllable Type was found for all three speakers: in closed syllables, f0 peaks relative to the nearby segmental acoustic landmarks occurred later than in open syllables. A slight effect of Stress Pattern was also found, although it held for only one speaker. Finally, no effects on HtoEndV0 arose from the durations of accented onsets and vowels. Thus the results in Experiment 1 support the hypothesis that intonational targets are not strictly anchored to segmental boundaries, contradicting the claims of the segmental anchoring hypothesis. On the other hand, the fixed rise time hypothesis cannot account for the fact that the intrinsic durations of accented onsets and nuclei did not influence the alignment of H peaks relative to the end of the accented vowel (e.g., shorter CV sequences did not result in later peaks relative to the end of V).

3. Experiment 2

The purpose of Experiment 2 was to gather reliable and complementary data on the effects of syllable structure found in Experiment 1 and new data on the effects of speech rate on the alignment of prenuclear peaks in Spanish. In the materials for Experiment 2, some of the confounding factors present in Experiment 1 were controlled for, such as stress pattern of the word, onset length, and coda type. In order to avoid the effects of such factors, test words for Experiment 2 were all proparoxytonic, onsets were all non-branching, and almost all codas were nasal.

Based on the predictions of the SAH, one would expect that when prosodic pressure is excluded, f0 targets would be aligned with specific points in the segmental string regardless of syllable structure and speaking rate. However, given the results of Experiment 1, one expects to find at the very least a systematic effect of syllable structure.

3.1. Method and materials

The database consisted of a total of 16 test proparoxytonic words, all nouns or adjectives included in a structure of the type [N + AP] or [AP + N] (e.g., *Mi **lánguido** niño no se encuentra bien* 'My languid child does not feel well.', *La **lámina** blanca está en la mesa* 'The white foil is on the table'). A subset of the sentences consisted of the sentences used in Experiment 1. The target words' test syllables included a group of open and closed syllables. Factors such as position of the accent within the sentence and the word, distance to next accented syllable, and number of accents in the utterance were constant. The test words were exclusively proparoxytones and they were all followed by two unstressed syllables. A full list of the test sentences used in Experiment 2 can be found in Appendix A.2.

As in the case of Experiment 1, in normal reading, the test words were expected to bear a rising prenuclear accent with a peak aligned towards the end of the accented vowel (see Fig. 1). As Experiment 2 was performed 1 day after Experiment 1, speakers were acquainted with the expected intonational contour and thus were less often off-target than in Experiment 1. Yet occasionally sentences had to be repeated at the end of each reading block when they were not produced with the target pitch accent.

3.1.1. Subjects

The same speakers as in Experiment 1 were used in a different recording session 1 day after Experiment 1. The three speakers read the set of test sentences listed in Appendix A.2. For this experiment, they were told to read each target sentence aloud three times, first at a normal speech rate, then at a fast rate and finally at a slow rate. They thus read a total of 432 utterances (16 test sentences $\times$ 3 speaking rates $\times$ 3 speakers $\times$ 3 repetitions).

3.1.2. Analysis procedure

The analysis procedure and segmentation of the materials were the same as used in Experiment 1 (see Section 2.1.4). The only difference is that two timing points (utterance-initial and utterance-final values) were added in order to monitor whether the self-paced rate of speech was reflected in the actual duration of the utterances.

3.1.3. Statistical analyses

The data were analyzed as in Experiment 1. First, to analyze the individual effects, separate one-way ANOVAs were performed for each speaker and for each independent variable, namely, Syllable Type and Rate. The variable Syllable Type had two levels (open vs. closed) while the variable Rate had three (fast, normal, or slow). Second, in order to model the joint effects of all factors and to observe the interactions across independent variables, an univariate linear model was run taking Syllable Type, Rate, and Speaker as independent factors.

3.2. Results

3.2.1. Speaking rate differences

In order to test whether the self-imposed rate contrast during the reading of the recording materials was reflected by the data, mean utterance duration (in ms) was calculated for each speaker at each speaking rate. Fig. 5 plots average utterance length (in ms) as a function of rate (fast, normal, and slow) for all three speakers. In general, all speakers successfully self-paced their speaking rate. The mean utterance-length

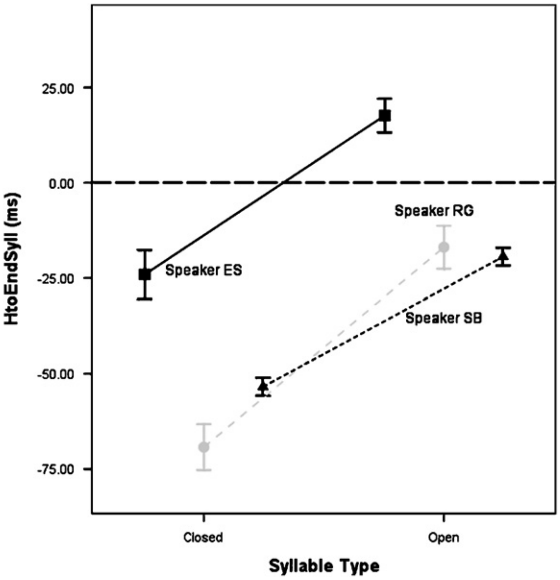

Fig. 5. Mean utterance length (in ms) as a function of speech rate (fast, normal, and slow), for all three speakers. The height of the bars represents standard errors.

difference across conditions was 238.5 ms for speaker ES, 768.2 ms for speaker RG, and 211.4 ms for speaker SB. As expected, a one-way ANOVA revealed a significant main effect of rate on utterance length for the 3 speakers [$F(2,93) = 19.976$, $p<0.0001$ (Speaker ES); $F(2,93) = 161.753$, $p<0.0001$ (Speaker RG); $F(2,93) = 15.857$, $p<0.0001$ (Speaker SB)]. Post-hoc tests with Bonferroni adjustments revealed that all pairwise comparisons were significant at $p<0.05$, except for slow-normal comparisons for speaker SB ($p = 0.246$). This result will be important in the discussion of the speech rate effects (see Section 3.2.3).

3.2.2. *Effects of speaking rate*

This section examines the effects of rate on H alignment patterns. Previous reports in the literature yield contradictory findings. While some studies report that speech rate affects the timing of peak placement of H peaks (Xu, 1998; some speakers in Ladd et al., 1999), others report no effects of speech rate on peak alignment (Steele, 1986; Silverman & Pierrehumbert, 1990; Ladd et al., 1999). The latter studies showed that changes in speech rate only had a proportional effect on the alignment of H peaks. The longer a syllable was as a result of a slower speaking rate, the farther the f0 peak aligned on a proportional basis, and vice versa.

The two graphs in Fig. 6 plot the mean distance from the peak to the end of the accented syllable (HtoEndSyll, left graph) and to the end of the accented vowel (HtoEndV0, right graph) as a function of speech rate (fast, normal, and slow) for the three speakers (see different line types). Although the two figures show moderate individual differences in overall alignment across speakers (again, speaker ES aligns peaks later), the effects of speaking rate on H alignment patterns are consistent across subjects. In general, peaks are located earlier in the syllable as speech rate decreases.

Table 3 shows the results of a set of one-way ANOVAs of the effects of Rate on two measures of H peak location, namely, HtoEndSyll and HtoEndV0, run separately for each speaker. The main effect of Rate was significant at $p<0.001$ for both measures and for all speakers. Post hoc analyses reveal that all pairwise comparisons were significant at $p<0.05$ except for one case, namely, pairwise comparisons between normal-slow for speaker SB. Crucially, this comparison is the one which did not reflect a significant difference in utterance length (see Section 3.2.1 above).

Thus the results in this section demonstrate that H alignment is sensitive to speech rate changes in each of the three conditions (fast, normal, and slow) for the three speakers, regardless of the measure of H alignment used. Similarly, Xu (1999) showed that peak delay in Mandarin Chinese occurred regularly in both high and rising tones; at slow rates, peak delay continued to occur regularly in rising tones but only rarely in H; at fast

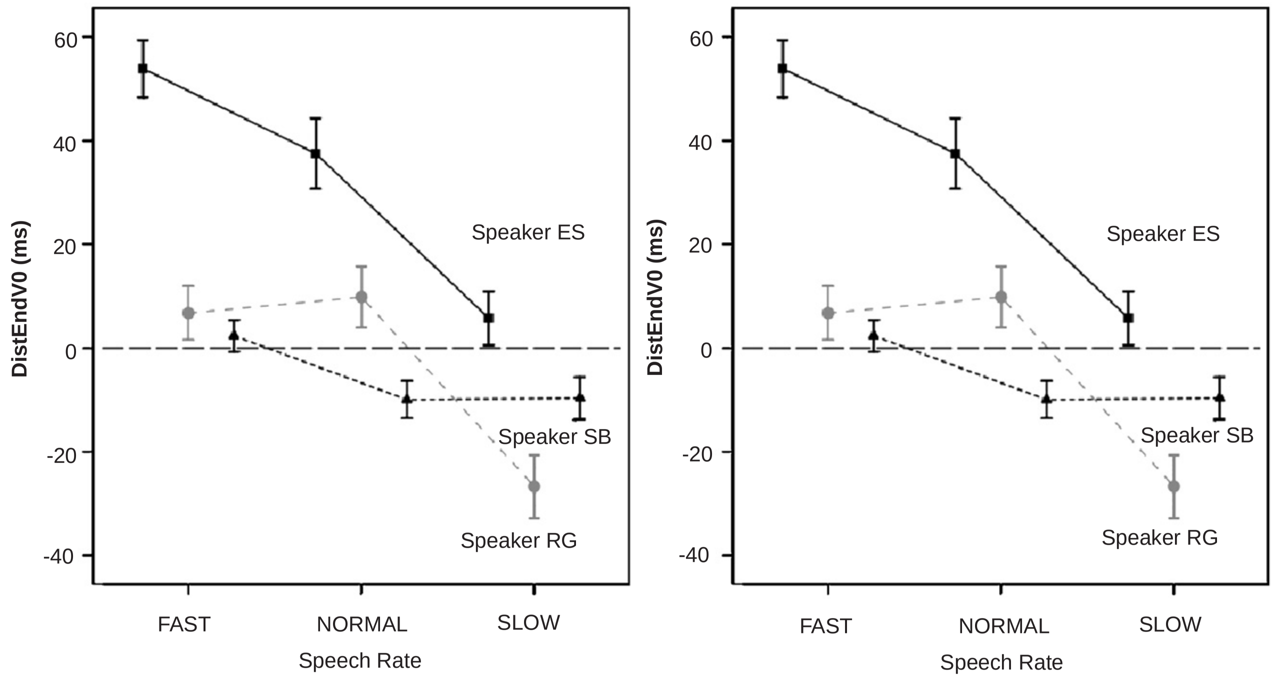

Fig. 6. Mean distance in ms from the peak to the end of the accented syllable (left graph) and to the end of the accented vowel V0 (right graph) as a function of speech rate (fast, normal, and slow) for the three speakers. The height of the bars represents standard errors and the dashed lines represent the temporal position of the end of the syllable and the end of the vowel, respectively.

Table 3
One-way ANOVA summaries of the effects of speech rate (SR) on two measures of H location, namely, HtoEndSyll and HtoEndV0

Speaker	HtoEndSyll/RATE	HtoEndV0/RATE
ES	$F(2,93) = 21.912$, $p<0.001$	$F(2,93) = 17.358$, $p<0.001$
RG	$F(2,93) = 25.419$, $p<0.001$	$F(2,93) = 12.518$, $p<0.001$
SB	$F(2,93) = 6.928$, $p<0.002$	$F(2,93) = 3.857$, $p<0.001$

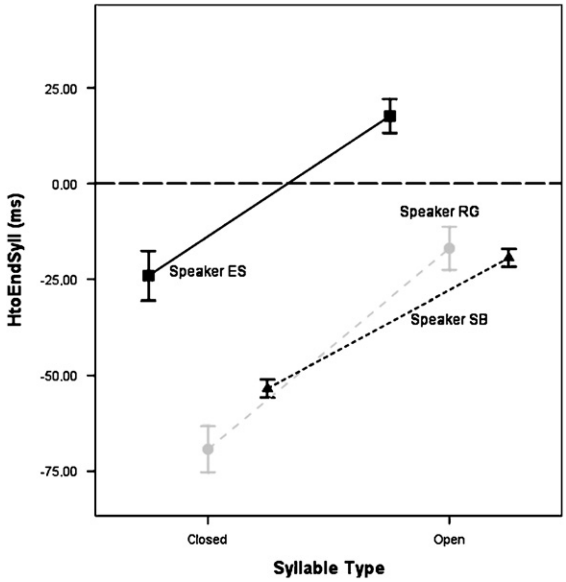

Fig. 7. Mean distance (in ms) from the peak to the end of the accented syllable as a function of syllable type (open vs. closed) for the three speakers. The height of the bars represents standard errors and the dashed horizontal line represents the temporal position of the syllable boundary.

rate, peak delay occurred not only regularly in *R* and *h*, but also frequently in H. Peak-alignment analyses revealed that peak delay tended to occur whenever f0 rose sharply near the syllable offset. The results were interpreted as indicating that peak delay is due to an articulatory constraint that limits how fast the larynx can reverse the direction of pitch movement.

3.2.3. *Effects of syllable structure*

Alignment of the f0 peak (H) was characterized quantitatively in two different ways, namely, relative to the end of the accented syllable (HtoEndSyll), and relative to the end of the accented vowel (HtoEndV0). Fig. 7 plots the mean distance (in ms) from the peak to the end of the accented syllable as a function of syllable type (open, closed) for the three speakers. The dashed lines represent the temporal position of the syllable boundary. Two main results stand out in this graph. First, even though H alignment exhibits a certain amount of inter-speaker phonetic variability (speakers showed variation in their mean alignment points, and speaker ES clearly aligns peaks later than the other two speakers), it is clear from the graph that H peaks are not anchored at the end of the syllable. Second, the graph reveals a consistent qualitative difference in alignment brought about by syllable structure, namely, peaks are more retracted into the syllable in closed syllables than in open syllables (mean difference across conditions is 41.79 ms for speaker ES, 52.36 ms for speaker RG, and 34.02 ms for speaker SB).

As with the findings of Experiment 1, the results illustrated in Fig. 6 leave open the possibility that the alignment difference could be interpreted as a reflection of aligning the f0 peak with the end of the accented

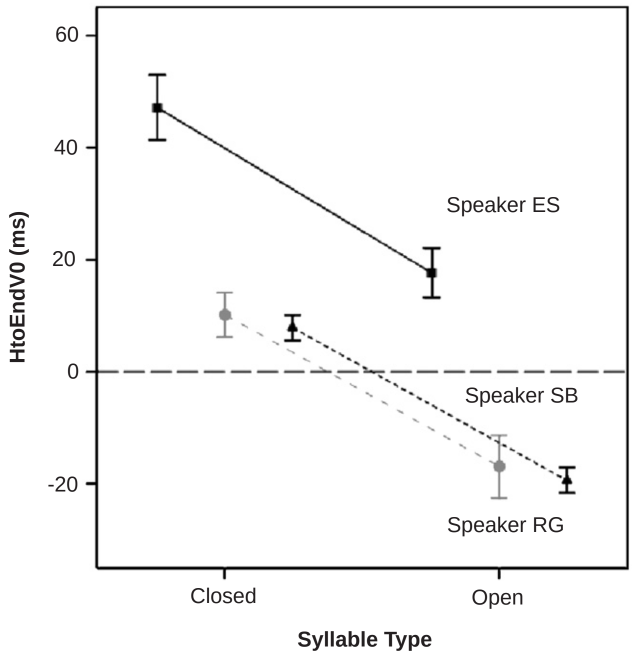

Fig. 8. Mean distance in ms from the peak to the end of the accented vowel V0 as a function of syllable type (open vs. closed) for the three speakers. The height of the bars represents standard errors and the dashed lines represent the temporal position of the end of V0.

Table 4
One-way ANOVA summaries of the effects of syllable type (ST) on two measures of H location, namely, HtoEndSyll and HtoEndV0

Speaker	HtoEndSyll/ST	HtoEndV0/ST
ES	$F(1,94) = 28.343$, $p<0.001$	$F(1,94) = 16.360$, $p<0.001$
RG	$F(1,94) = 40.689$, $p<0.001$	$F(1,94) = 15.535$, $p<0.001$
SB	$F(1,94) = 106.352$, $p<0.001$	$F(1,94) = 71.184$, $p<0.001$

vowel. Whereas in an open syllable the end of the syllable corresponds to the end of the vowel, that is not the case in a closed syllable. Fig. 8 plots the mean distance in ms from the peak to the end of the accented vowel V0 as a function of syllable type (open vs. closed) for the three speakers. This graph clearly shows no evidence of H anchoring to the end of the vowel. Clearly, H peaks are positioned within the coda consonant in closed syllables and around the end of the vowel/syllable in open syllables.

The tendencies observed in Figs. 6 and 7 are statistically confirmed by a set of one-way ANOVAs run separately for each speaker. Table 4 shows the results of one-way ANOVAs of the effects of Syllable Type (ST) on two measures of H peak location, namely, HtoEndSyll and HtoEndV0. All the effects of ST were significant at $p<0.001$ for both measures and for all speakers.

In sum, as in Experiment 1, results from Experiment 2 show that the alignment of the H prenuclear peak is consistently affected by syllable structure: in general, closed syllables trigger earlier peak alignment within the syllable. Interestingly, similar findings have been reported by Ladd et al. (2000) for Dutch, D'Imperio (2000) and Prieto (2006b) for Neapolitan Italian, Gili-Fivela and Savino (2003) for Pisa and Bari Italian, Welby and Lœvenbruck (2005, 2006) for French, and Hellmuth (2005, 2006) for Egyptian Arabic.

3.2.4. A univariate linear model

This section describes the effects of speech rate on H alignment patterns in our data and its interaction with the effects of syllable structure. Our interest is to determine whether the effects of syllable structure on f0 peak timing are maintained across speaking rates and, conversely, whether speech rate effects are consistent across syllable structure types.

The three graphs in Fig. 9 plot the mean distance in ms from the H peak to the end of the syllable across syllable structure (solid line = open syllables; dashed line = closed syllables) and speaking rate conditions, for the three speakers (ES = upper left graph; RG = upper right graph; SB = bottom graph). Clearly, the effects of syllable structure on peak alignment are consistent across speaking rates for all three speakers: the graphs show that peaks are aligned significantly earlier with respect to the syllable boundary in closed syllables than in open syllables, across the three speech rate conditions. Conversely, the effects of speech rate are consistent across syllable types for all three speakers: at fast tempo, peaks are aligned later than at normal or slow tempo in all conditions. Indeed, when the other measure of peak alignment (HtoEndV0) is examined, the effects of syllable structure and speech rate are parallel to the ones obtained with HtoEndSyll.

As in Experiment 1, an attempt was made to identify the combined effects of the various factors and to observe the interactions between the independent variables and the dependent variables HtoEndSyll (model 1) and HtoEndV0 (model 2) by means of an univariate linear model taking SYLLABLE TYPE, RATE, and SPEAKER as independent factors. The results of the first model reveal significant main effects of the factors SYLLABLE TYPE (at $p<0.05$) and SPEECH RATE ($F(2,4) = 8.789$, $p<0.05$) on the independent variable HtoEndSyll. The linear model revealed that the effect of SPEAKER was marginally significant (at $F(1,2) = 64.506$, $p<0.06$), revealing that there are no large individual differences in overall alignment. There was a two-way interaction between SYLLABLE TYPE and RATE. An inspection of the means (see Fig. 9) revealed small variations in effect size for the RATE factor across different syllable types as a source of the interaction. Importantly, there were no interactions between SYLLABLE TYPE and SPEAKER, and RATE and SPEAKER, meaning that the effects of SYLLABLE TYPE and RATE on alignment patterns are relatively stable across subjects. Even though speakers showed variation in their mean individual alignment points, the overall effects of both independent variables were the same. Importantly, exactly the same results were obtained using HtoEndV0 as an independent variable.

In sum, the results of Experiment 2 makes clear that both syllable structure and speaking rate have consistent effects on H peak aligment in Castilian Spanish. The main effect of these two variables was statistically robust for the two dependent variables analyzed, namely, HtoEndSyll and HtoEndV0, that is, this result emerges regardless of whether the alignment of H is measured relative to the end of the accented syllable or the end of the accented vowel. Thus the results presented here do not support the prediction of the SAH that alignment of H would not be affected by syllable structure or speech rate.

3.2.5. Syllable duration and rise time data

Several analyses were performed in order to test the hypotheses that the alignment differences reported are exclusively due to syllable duration, and that rise time is invariant. With respect to the first question, the three graphs in Fig. 10 show that the degree of correlation between the two variables is moderately high (R^2 ranges from 0.302 to 0.89, all significant below a probability level of $p<0.004$). It is clear from the graphs that peaks are more retracted in syllables with a coda, indicating that the peak is actually not getting to the end of the syllable in closed syllables. The scatterplots show that, for two of the speakers, the scores do not fit into a straight line, and that they form two separate clouds according to the two syllable type groups.

In order to evaluate the invariant rise time hypothesis, the variable RISE TIME was calculated as the distance from the f0 valley at the beginning of the accented syllable to the f0 peak. Taking the data from Experiment 1, RISE TIME measures were compared across SYLLABLE TYPE conditions, since this factor had exerted a clear effect on HtoEndV0 without significantly modifying segmental durations in the accented syllable. A series of one-way ANOVAs revealed significant RISE TIME differences for all three speakers [$F(1,142) = 46.90$, $p<0.001$ (ES); $F(1,142) = 5.36$, $p<0.05$ (RG); $F(1,142) = 8.88$, $p<0.01$ (SB)]. For speakers RG and SB, mean RISE TIME values were higher for closed than for open syllables, with mean differences of 12 ms for both speakers. Speaker ES, however, consistently displayed a contrary tendency. On average, RISE TIME for this speaker was 30 ms longer in open than in closed syllables. After that, using data from Experiment 2, a series one-way ANOVAs were performed to analyze the effects of syllable structure and speech rate on rise time. The results presented in the following table show that speakers tend to display a significant difference in rise time (at $p<0.05$) for the three speech rate conditions (fast, normal, and slow). Post hoc Bonferroni tests confirmed that

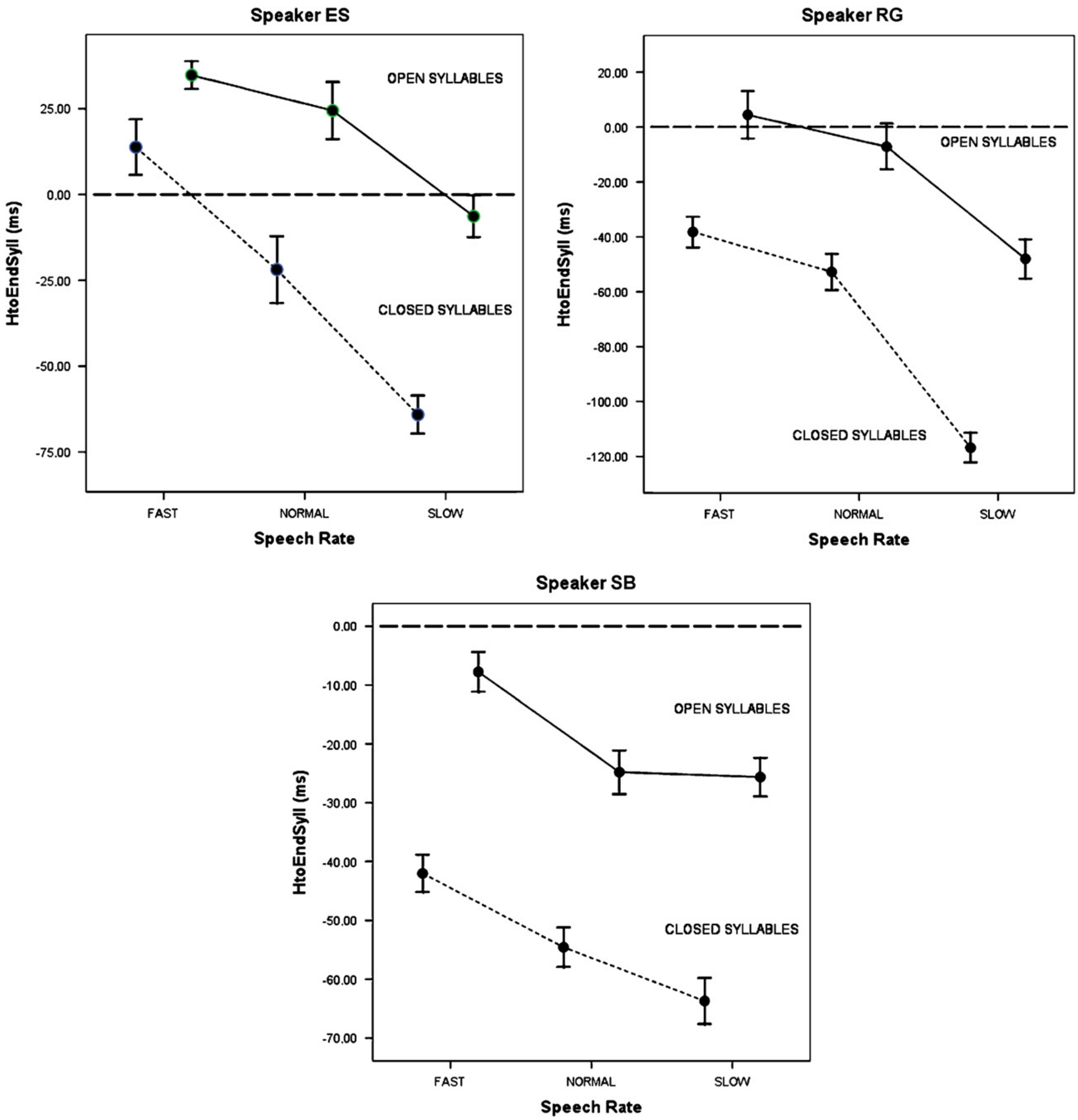

Fig. 9. Mean distance in ms from the H peak to the end of the syllable across syllable structure and speech rate conditions, for the three speakers (ES = upper left graph; RG = bottom graph; SB = upper right). Solid lines link values in open syllables and dashed lines values in closed syllables. The height of the bars represents standard errors and the horizontal dashed line represents the temporal position of the end of the syllable.

the differences were also significant across speech rate groups. As for the effects of syllable structure, the analysis reveals that the effect is significant for only one of the speakers (SB) (Table 5).

For more evidence against the claims of the fixed rise time hypothesis in our data, see the data in Fig. 11 in the following section (Section 4.1). While the alignment of the f0 valley is shown to be relatively stable with respect to the beginning of the syllable, the alignment of the f0 peak is relatively unstable with respect to the end of the syllable.

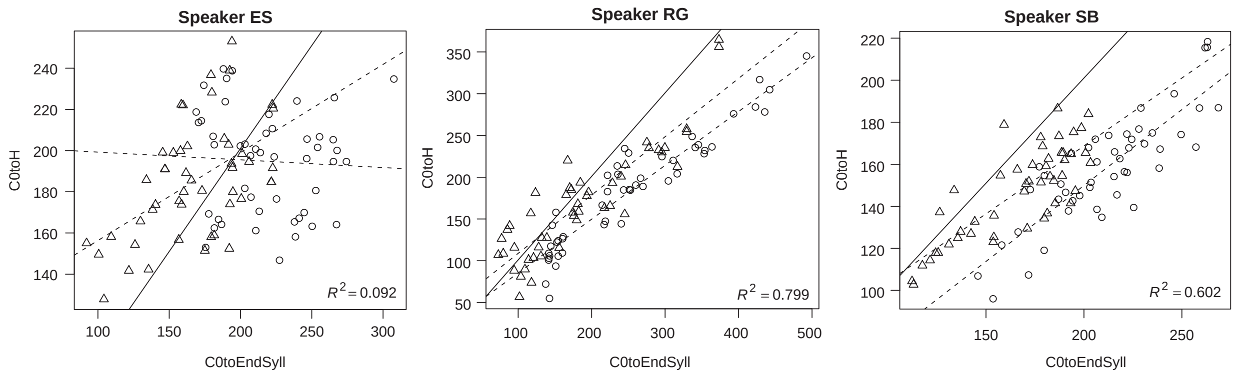

Fig. 10. Peak delay or C0toH (in ms) as a function of syllable duration in two syllable type conditions: open syllables (represented by a triangle) and closed syllables (represented by a circle), for the three speakers. Solid lines represent the temporal position of syllable boundaries ($x = y$), while dashed lines are regression lines for each syllable type.

Table 5
One-way ANOVA summaries of the effects of syllable type (ST) and speech rate (SR) on rise time

Speaker	Rise time/ST	Rise time/SR
ES	$F(1,94) = 3,104$, $p = 0.081$	$F(2,93) = 3.986$, $p<0.022$
RG	$F(1,94) = 1.368$, $p = 0.245$	$F(2,93) = 68.763$, $p<0.001$
SB	$F(1,94) = 5.434$, $p<0.022$	$F(2,93) = 23.960$, $p<0.001$

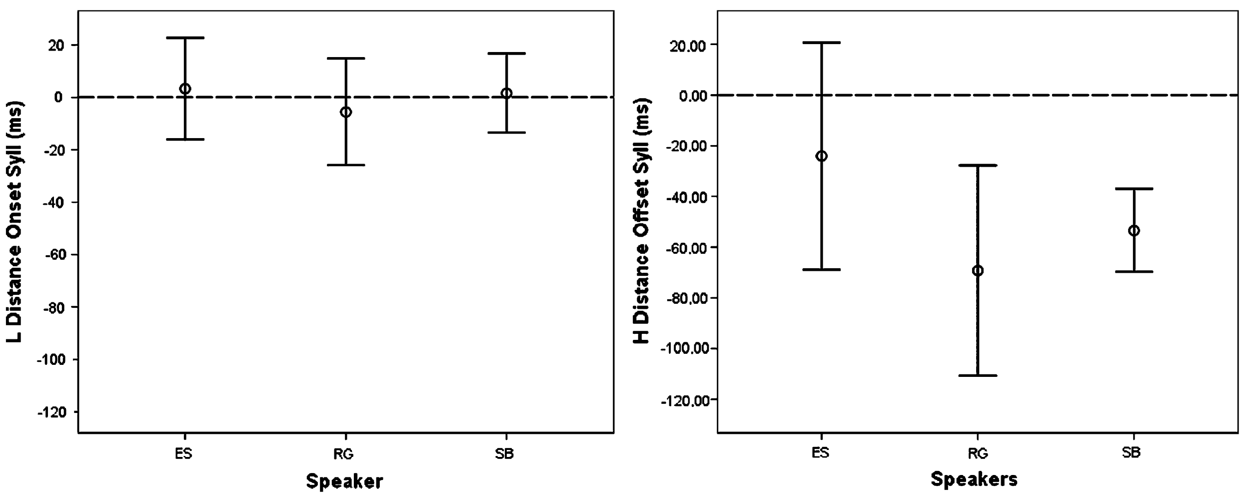

Fig. 11. Mean distance in ms from the L valley to the onset of the accented syllable (left graph), and mean distance from the H peak to the offset of the accented syllable (right graph) for the three speakers. The height of the bars represents standard deviations and the horizontal dashed line represents the temporal position of the beginning and end of the accented syllable (left and right graphs, respectively).

4. Discussion

4.1. The segmental anchoring hypothesis revisited. In-phase and anti-phase synchronizations with syllable structure

Even though recent tonal alignment work has suggested that the temporal location of tonal targets relative to the segmental string might be governed by principles of synchrony and stability, the results of the present experiments taken together reveal that the alignment patterns of H prenuclear peaks in Spanish are sensitive to the syllable structure of the target accented syllable and also to speech rate. These effects were statistically robust for the two dependent variables under study, namely, HtoEndSyll and HtoEndV0. These results have clear implications in relation to the predictions of the SAH, as they demonstrate that tonal H turning points are not anchored at acoustic segmental landmarks such as the vocalic or the syllabic offsets.

The evidence presented by our two experiments shows that the end of the f0 rise of a Spanish utterance-initial pitch accent is not aligned with segmental 'anchors' but is clearly affected by factors such as syllable structure and speech rate. To a certain extent, one had good reason to expect that such effects would be at work in our data. As is well known, previous work on tonal alignment has reported a contrast between L and H targets in rising accents, namely, consistent alignment of L targets with the syllable onset, while H placement is found to be variable to a greater or lesser extent (and strongly influenced by segmental duration and right-hand prosodic environment). For example, the findings of Caspers and van Heuven (1993) for Dutch, Prieto et al. (1995) for Spanish, Arvaniti et al. (1998) for Greek, D'Imperio (2000, 2002) for Neapolitan Italian, Ladd et al. (1999), Ladd et al. (2000), Atterer and Ladd (2004) for English, and Xu (1998, 1999, 2002) for Chinese all suggest that f0 movement toward a tonal target starts at the onset of the syllable. In our data, this contrast is also clear. The two graphs in Fig. 11 plot the mean distance from L to the onset of the syllable (left graph) for all syllable types and the mean distance from H to the end of the syllable in closed syllables. Only closed CVC syllables from Experiment 2 have been taken into account for the analysis. By using the same scale in the two graphs, one can visually appreciate the quantitative difference in data dispersion between the two anchoring points. While the majority of L tones were realized within 20 ms of the onset of the syllable, H tones were more variable and were realized within 120 ms from the end of the syllable.

Even though the findings reported in this paper, together with a range of crosslinguistic empirical findings, make a strict anchoring hypothesis less attractive than it once seemed, we believe that "there is a genuine phenomenon of 'segmental anchoring', whereby the duration of pitch movements in speech is finely adjusted to the duration of the accompanying segmental material." Yet, how can the anchoring view account for the contrast in behavior between a stable synchronization of L valleys with the onset of the syllable and a more variable synchronization between H peaks and the end of both open and closed syllables?

The Articulatory Phonology (AP) model originally proposed by Browman and Goldstein (1986) and developed over the last two decades (Browman & Goldstein, 1988, 1992; Goldstein & Fowler, 2003; Byrd, 1996, among many others) may provide the tools to adequately account for some of the alignment patterns reported by recent work on tonal alignment. We take the view expressed by Ladd (2006) that pitch accents can also be considered from an articulatory point of view as speech gestures, and that we can test their behavior and compare it to the behavior of supraglottal gestures. This line of research has recently started with the work by Prieto (2006b) and Mücke, Grice, Becker, Hermes, and Baumann (2006) on the coordination of tonal and supraglottal gestures in Neapolitan Italian and German, and in Gao's (2006) work applying Articulatory Phonology to tonal specification. Ladd (2006) points out that "in one way or another, it appears that we must acknowledge that pitch movements are 'gestures' in the sense of Articulatory Phonology, and seek to explain segmental anchoring in terms of a more general theory of gestural coordination." In this model, the phonological representation of an utterance is viewed as an organized pattern of dynamically defined articulatory gestures. These gestural constellations allow the specification of temporal relations among articulatory gestures and thus have the advantage of accounting for the temporal properties of speech. Gestural units are modeled as dynamic systems, using the Task Dynamics Model (Browman and Goldstein, 1995b; Saltzman, 1991; Saltzman and Kelso, 1987; Saltzman and Munhall, 1989) and gestural duration is specified largely by factors intrinsic to individual gestures (such as "stiffness") and to factors that affect all gestures (such as speech rate).

Our results show that the patterns of tune-text coordination found are similar to other types of articulatory gesture coordination found in the AP literature. Browman and Goldstein (1995b) proposed that there are distinct modes of gestural coordination for gestures in an onset position as opposed to those in a coda position, and that these modes are the manifestations of syllable structure. There is extensive evidence that in syllable onsets a synchronous mode of coordination dominates—consonant gestures tend to be synchronous with one another and with the following vowel—while in coda position, a sequential mode dominates. For example, it has been shown that nasals and laterals both show differential timing betweeen their component gestures syllable-initially and syllable-finally. For syllable-initial nasals (e.g., *see* ***m****ore*) the *end* of the velic lowering gesture roughly coincides with the *end* of the lip closing movement, whereas in syllable-final nasals (e.g., *seem more*) the end of the suprasegmental gesture coincides with the *beginning* of the lip closing movement (Krakow, 1999). Within AP, the coupling modes hypothesis argues that there are two possible ways of combining a consonant gesture and a vowel gesture in an intrinsically stable mode: (1) *in-phase* coupling, hypothesized for C–V onset relations; and (2) *anti-phase* coupling, hypothesized for V–C coda relations (for similar views about distinct C–V and V–C modes, see Tuller & Kelso, 1991).[2]

Even though there are no direct observations of gestural dynamics in our data, we hypothesize that the beginning of the rising f0 movement (L target) is phased more synchronously with commands of the supralaryngeal articulator, as the syllable onset is the point at which the greatest synchrony is achieved between gestures (the initial consonant, the first vowel and the start of the accent gesture). In contrast, the end of the rising f0 movement (H target), which occurs towards the end of the syllable, is not so tightly coordinated with supraglottal gestures. This lack of synchronicity makes tonal gestures in the coda weak and unstable, and tonal undershoot is expected to occur more easily under greater time pressure, that is, it is expected that a pitch gesture will fall short of reaching its ideal target.

When conceived in this way, one would also expect that in *falling* accents H turning points should be more closely coordinated with the syllable onset than L ending points. Interestingly, recent results on tonal alignment in falling accents of Central Catalan (Prieto, 2006a) and of Majorcan Catalan (Vanrell, 2006) provide evidence for this point. In these two dialects of Catalan, the beginning of the f0 falling movement (H target) in falling nuclear accents of yes–no questions is closely coordinated with the onset of the accented syllable, while the end of the falling gesture (L target) is more variable. Again, a clear effect of syllable structure is found which is the mirror image of the results obtained in this article: while in open syllables the end of the fall is aligned roughly with the end of the accented vowel, in closed syllables it is aligned somewhere within the coda consonant (see Prieto, 2006a).

We believe that segmental anchoring can be interpreted in a less constrained way as a consequence of the general coordination between f0 gestures and units of speech production. Following Xu (1998) Xu and Liu (2006), and D'Imperio (2002), some pitch accent gestures are associated and synchronized with the syllable unit. For these models, if these accents are produced with no tonal pressure, the prediction is that both the beginning and the end of the rise (or the fall) will seek to align with the syllable edges.[3]

In its turn, the intricate behavior of tonal alignment can be explained by general properties of intergestural timing, which can provide a unifying explanation for (1) 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 (2) the right-hand tonal pressure effects and 'undershoot' patterns displayed by peaks at the ends of syllables and other prosodic domains. This view calls for the further collection and analysis of articulatory data to explore issues of intergestural coordination and phasing relations involving tonal gestures. Preliminary results on the alignment of f0 turning points with articulatory gestures reveal that H targets of nuclear rises in Neapolitan statements and questions are more closely phased with the articulatory dimension of between-lip distance than with the most commonly employed acoustic landmarks for tonal alignment (i.e., onset and offset of stressed vowel (D'Imperio, Espesser, et al., 2007; Mücke et al., 2006).

[2]This hypothesis predicts a variety of observations about coupling between competing C gestures, namely, that this coupling is more stable in onset clusters than in coda clusters (Byrd, 1996). Thus the hierarchical structure of syllables is interpreted as a consequence of combining gestures using stable coupling modes. There is now a fairly extensive literature relevant to this topic (for a review, see Browman & Goldstein, 1995a, 1995b; Krakow, 1999; Byrd, 1996, and references therein).

[3]In the time structure model of the syllable proposed by Xu and Liu (2006), which is an extension of the target approximation model (Xu & Wang, 2001) does not hypothesize that f0 peak or valley is anchored to either syllable onset or offset. Rather, the hypothesis is that an underlying tonal target, which is usually just a linear level or slope, is approached in synchrony with the entire syllable. But a certain "anchoring" behavior can be predicted by this hypothesis, given a particular tonal target and its surrounding tones. For example, given a Rising tone with a simple linear rising slope, when articulated in a High ___ Low context, an f0 peak is predicted to align quite consistently with the offset of the syllable.

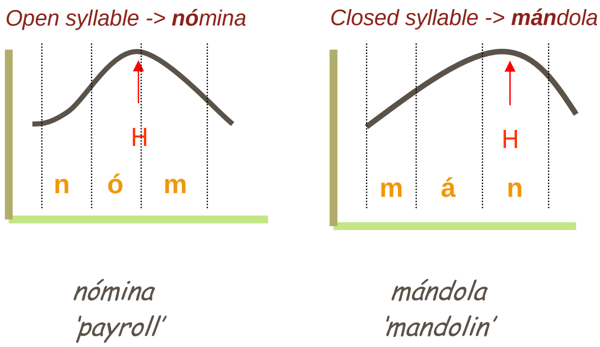

Fig. 12. Schematic representation of the alignment of peaks (H) relative to the accented syllable of the Spanish test words (based on results of Experiments 1 and 2). Segment durations are idealized.

4.2. Interpreting syllable structure effects: prosodic domains of tonal timing coordination

Previous work on peak alignment patterns in Spanish has mainly focused on the effects of right-hand prosodic environment and prosodic word boundaries on peak alignment, but has not dealt with syllable structure effects (see Prieto et al., 1995; Prieto, Estebas-Vilaplana, & Vanrell, 2006, among others). In our experiments, H alignment patterns were qualitatively different in open and closed syllables, regardless of syllable duration. The difference between the two types of alignment is illustrated in Fig. 12: while the H tonal target for ***nó**mina* is realized in the vicinity of the vocalic offset, the H target for ***mán**dola* is realized within the coda consonant. If the predictions of the SAH were true, one would expect that the H target would be aligned either to the vowel or the syllable offset in both cases.

We also had good reason to expect that our experiments would reveal substantial effects of syllable structure on H location, given that parallel findings had been reported for other languages (see Ladd et al., 2000 for Dutch; D'Imperio, 2000 and D'Imperio, Petrone, et al., in press for Neapolitan Italian; Gili-Fivela & Savino, 2003 for Bari and Pisa Italian; Welby & Lœvenbruck, 2005, 2006 for French; Hellmuth, 2005, 2006 for Egyptian Arabic). Similar to the Italian case, Ladd et al. (2000), found that for Dutch prenuclear LH accents the peak was aligned at the end of the following consonant when the accented vowel was short. On the other hand, when the vowel was long, the peak was anchored at the end of this vowel. Given that these two vowels did not differ in duration phonetically, Ladd et al. suggested that this difference in anchoring could be due to differences in syllabic phonological structure. Another study that describes the effects of syllable structure is Welby and Lœvenbruck (2005, 2006). They found that the peak of the French late rise varied across syllable types. They studied three different structures: (a) open syllable (e.g., *sala**mi***), (b) closed by a sonorant consonant (e.g., *vita**min**(e)*) and (c) closed by an obstruent (e.g., *pyra**mid**(e)*). Whereas in CVCobs and CV syllables H was realized within 20 ms of the end of the vowel, in CVCson syllables it was generally realized before the end of the coda consonant.

Work on other non-European languages such as Japanese and Arabic offers further evidence in favor of the active role of syllable structure on H alignment. Ishihara's (2003) work on the accentual phrase-initial rise in Tokyo Japanese shows that the alignment of the f0 peak depends on syllable structure. As Ishihara (2003) argues, "while for CV.CV utterances with an accent in the first syllable the peak was aligned with the start of the vowel in the second syllable, in CVN.CV utterances with the same stress pattern the peak was located between the first vowel and the nasal coda. The proposed explanation of these data was that in both cases the H tone was aligned with the onset of the second mora after the beginning of the accented syllable."[4] Finally, Hellmuth (2005, 2006) finds that H alignment in Egyptian Arabic varies across syllable types (light

[4]Crucially, in his interpretation of the data, Ishihara makes reference to the moraic structure in order to provide an 'anchoring' generalization in reference to prosodic structure, not segmental structure: "Given that onset consonants belong directly to the syllable, rather than to the mora, it can be claimed that the f0 peak is anchored to the beginning of the second mora" (Ishihara, 2003).

vs heavy): ''in open stressed syllables the H peak falls well outside the CV syllable, (...), whilst in both CVC and CVV syllables the H peak falls just inside the stressed syllable (within the coda in CVC and just before the end of the stressed vowel in CVV).''

In sum, crosslinguistic evidence seems to suggest that, all other things being equal, H peaks tend to occur earlier within the syllable in closed than in open syllables when measured relative to the syllable end. This finding is consistent with the expectation that phonological domains are relevant for the coordination of f0 gestures. Let us assume that L and H turning points of a rising accent are seeking to align with the edges of the accented syllable, in a 'lax' interpretation of the anchoring hypothesis. As illustrated above, one would expect the coordination between tonal gestures and supraglottal gestures to be tighter (perhaps in-phase) in syllable-initial position and more variable in syllable-final position. In coda position, gesture coordination is less stable and undershoot is expected to occur more easily under the effects of time pressure (i.e., the typical right-hand prosodic effects reported in the literature). Yet why should there a consistent difference in alignment in closed syllables than in open syllables?

Studies of gestural overlap within the AP model have suggested that just as the syllable frontiers influence intergestural timing, other types of prosodic boundaries have been shown to influence it too (for a review, see Byrd, 1996 and Krakow, 1999). 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 subjects appeared to use a diminished gestural stiffness as the main source of lengthening before prosodic boundaries. Extrapolating their results, pushing H alignment back within the syllable or the prosodic word could be the result of a decrease in the stiffness of all articulators (including the glottal movements) in syllable-final position.

Importantly, all this evidence suggests that prosodic structure domains should be viewed as domains of articulatory organization. As Ladd (2006) asks, ''[do] right-context effects operate at the level of the foot, the (prosodic) word, or some larger prosodic unit like the intonation phrase?'' We would like to suggest that tonal domains are not concerned exclusively with syllables, and that there is evidence that higher levels of prosodic structure influence tonal intergestural coordination. Support for this view is the fact that acoustic studies of a variety of languages have shown that *within-word position* has a very robust effect on the position of the peak of rising prenuclear accents, with the details of the alignment depending 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 et al., 1995; de la Mota, 2005; Estebas-Vilaplana & Prieto, 2007; Simonet & Torreira, 2005; Simonet, 2006 for Spanish; Prieto, 2006b for Catalan; Arvaniti et al., 1998 for Greek; Ishihara, 2003, 2006 for Japanese). In all these languages, peaks tend to shift backwards 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. In order to correct for the potentially confounding effects of stress clash (or distance to the next accented syllable), Prieto et al. (1995) analyzed a subset of the data which contained test syllables in different positions in the word (*número*, *numero*, *numeró*) and which maintained a distance of two unstressed syllables from the next accented syllable (*número rápido*, *numero nervioso*, *numeró regular*). A significant effect of word position within the sentence on peak delay was found in all of the comparisons (though it was stronger for one of the three speakers). Similarly, in Silverman and 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.

Crucially, the effect of upcoming prosodic word boundaries on H alignment is parallel to that exerted by upcoming syllable boundaries. This 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. Even though there is well-documented evidence of this in supraglottal gestures, very little research has been done in the tonal domain.

Finally, within this view it follows that timing relations are not universal and languages can differ from one another in the timing of articulatory gestures. As argued by Browman and Goldstein (1986, p. 222), timing of articulatory gestures is relevant in terms of how languages differ from one another. Accordingly, in the tonal alignment domain, one would expect to find consistent small differences between languages and between varieties of the same language. Indeed, recent research has revealed that not all rises align in the same way with the associated syllable (Atterer & Ladd, 2004; Arvaniti & Gårding, 2007; see also D'Imperio, Espesser, et al., 2007 comparing French and Neapolitan Italian).[5] Also, phasing relations are intrinsically variable, and thus intra-speaker variability is expected to show up in the alignment data. In this vein, the concept of *anchorage* proposed by Welby and Lœvenbruck (2005, 2006) successfully captures the inter- and intra-speaker variation found in the current Spanish data.

In concluding this study it is important to note a key caveat. The findings presented here are based on the performance of only three subjects. Yet the fact that results from other studies on closely related languages point in the same direction as the patterns exhibited by these Spanish speakers represents evidence in favor of the stability and robustness of the results (see Prieto et al., 2006a for similar results in both Catalan and Spanish).

5. Conclusion

The experimental evidence presented in this paper demonstrates that, even when tonal pressure effects are controlled for, factors like syllabic structure and speech rate continue to exert consistent effects on H alignment in prenuclear peaks in Spanish. Both factors have shown statistically robust effects for the 3 speakers for the two dependent variables under study, namely, the distance from the peak relative to the end of the accented vowel and the distance to the end of the accented syllable (HtoEndV0 and HtoEndSyll, respectively), and in two separate experiments. In CV syllables the peak was located around the end of the accented vowel, while in CVC syllables around the beginning-mid part of the sonorant coda, though still far from the syllable boundary. With respect to the effects of rate, peaks were located earlier in the syllable as speech rate decreased. In sum, the evidence provided by the two experiments demonstrates that the end point of the f0 rise is not anchored at acoustic segmental landmarks such as the vocalic or the syllabic offsets, and thus do not provide support for a strict view of the SAH. A strict interpretation of the 'anchoring' hypothesis cannot account for the robust effects of syllable structure and speech rate on alignment exhibited by our data.

In light of the results reported in this study (together with empirical results from a variety of languages) we conclude that the resulting f0 alignment patterns can be better understood by closely analyzing the patterns of gestural coordination of f0 gestures with the supraglottal gestures. First, the syllabic domain is taken as the appropriate domain of gestural coordination and a crucial part of our understanding of the coordination of pitch gestures with the segmentals. It is clear that crosslinguistically the durations of the segments making up the target-accented syllable seem to determine peak delay. In general, the longer the syllable, the longer the rise time. Second, following the Task Dynamics model, different coordination patterns are predicted to occur at syllable onsets as compared with syllable offsets. We have suggested that the asymmetry between alignment patterns in syllable-initial vs. syllable-final (and also in higher-level prosodic domains such as the prosodic word) might be attributable to general properties of intergestural coordination.

Finally, we have suggested that future research within the gestural model may be the key to fully understanding the temporal positioning of tonal targets and its relationship with syllable structure and higher prosodic domains. This would allow us to respond to two fundamental issues in tonal alignment research, namely, the discovery of the domains of tonal gesture coordination and its source. This is indeed a potentially fruitful area for future research that would allow progress toward a general model of tonal timing.

Acknowledgments

Parts of this study were presented at the ESF International Conference on Tone and Intonation (Santorini, September 2004), the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology

[5]Atterer and Ladd (2004) found that German rising prenuclear accents were aligned consistently later than those in English and Dutch, and within German, such accents were aligned consistently later in Southern varieties than in Northern varieties. The effects are small but significant, and they crucially affect both the beginning and the end of the rise. Similar differences have been reported for Southern California and Minnesota varieties of American English by Arvaniti and Gårding (2007).

(Bloomington, Indiana, September 2004), and The Tenth Annual Midcontinental Workshop on Phonology (Evanston, Illinois, October 2004) and at talks at the Laboratoire de Parole et Langage (Aix-en-Provence, April 2005) and Institut de la Communication Parlée (Grenoble, November 2005). We are grateful to the audience in these conferences, and especially to M. D'Imperio, G. Elordieta, B. Gili-Fivela, M. Grice, C. Gussenhoven, S. Hellmuth, D. Hirst, J. I. Hualde, K. Iskarous, J. Kingston, D. R. Ladd, M-H. Lœvenbruck, C. Petrone, and P. Welby, and Y. Xu for very useful feedback. We are also indebted to Eva Estebas for help in contacting and arranging the recording sessions with the 3 Castilian subjects and also for conducting some extra recordings. Finally, we also thank the SEA (Servei d'Estadística de la Universitat Autònoma de Barcelona) for their help with the statistical analysis of the data and to the editor, G. Docherty and the three anonymous reviewers for their thorough revision of the paper. This research was funded by Grants 2002XT-00032, 2001SGR 00150, and 2001SGR 00425 from the Generalitat de Catalunya and HUM2006-01758/FILO from the Ministry of Science and Technology to the first author and a research grant from the Department of Linguistics at Université Libre de Bruxelles to the second author, where the first experiment described in this paper was submitted as an undergraduate thesis in 2004.

Appendix A. Speech materials

A.1. Experiment 1 test sentences (test syllable in boldface and nonsense words marked with an asterisk)

Paroxytonic target words	
Open syllables	
Eu**la**lia Fernández es su hermana mayor.	'Eulalia Fernández is his/her older sister.'
Pau**li**na Martínez es profesora de inglés.	'Paulina Martínez is an English teacher.'
Pa**lo**ma Gutiérrez trabaja en el banco.	'Paloma Gutiérrez works in a bank.'
La **lu**na y el sol se ven en el cielo.	'The moon and the sun can be seen in the sky.'
A**ma**lia Marín es prima mía.	'Amalia Marín is my cousin.'
E**mi**lio Rodríguez terminó ayer.	'Emilio Rodríguez finished yesterday.'
Ma**no**lo Moreno empezó ayer.	'Manolo Moreno started yesterday.'
La **mu**la y la burra se distinguen fácilmente.	'The mule and the donkey can be easily distinguished.'
Ha ha**bla**do y hablado toda la tarde.	'He/she has talked and talked all afternoon.'
Ne**bli**na de Duero es un pueblo de Castilla.	'Neblina de Duero is a village in Castile.'
Ta**blo**nes López se llama su empresa.	'His/her company is called Tablones López.'
Ha**blu**na y Lorca son pueblos de Granada.	'Habluna and Lorca are villages in Granada.'
La **lá**mina blanca está en la mesa.	'The white foil is on the table.'
El **lí**mite máximo será cincuenta.	'The upper limit will be fifty.'
Sa**ló**nica y Rodas están en Grecia.	'Salonica and Rhodes are in Greece.'
El **ú**nico precio es el que ves.	'The only price is the one you see.'
Closed syllables	
Ar**man**do Martínez es mi amigo.	'Armando Martínez is my friend.'
Do**min**go Fernández se llama su novio.	'Domingo Fernández is her boyfriend's name.'
Bel**mon**do Rodríguez son sus apellidos.	'His/her surnames are Belmondo Rodríguez.'
Rai**mun**do Noguer es colega mío.	'Raimundo Noguer is a colleague of mine.'
Rei**nal**do Marín es su marido.	'Reinaldo Marín is her husband.'
Bru**nil**do Segundo fue un rey.	'Brunildo the Second was a king.'
El **mol**de y la llave se descubrieron allí.	'The mould and the key were found there.'
La **mul**ta y la pena le llegarán el lunes.	'He/she will be informed of the fine and the sentence on Monday.'
El **blan**do y el duro se parecen poco.	'The soft one and the hard one are quite different.'
Se **blin**da y protege como bien puede.	'He/she protects him/herself as well as he/she can.'

La **blon**da amarilla me parece bien.	'The yellow cloth is fine.'
La **blun**da y la honda fueron utilizadas.	'The *blunda and the sling were used.'
Proparoxytones	
Open syllables	
El **Má**laga B jugará la final.	'Málaga B will play in the final game.'
La a**ní**mica gata maulló en la calle.	'The lively cat meowed in the street.'
La **nó**mina suya parece mayor.	'His/her salary seems higher.'
El **nú**mero siete fue el ganador.	'The winner was number seven.'
Ha**blá**bales rápido y con mucha prisa.	'He/she talked to them fast and in a rush.'
Ha**blí**grafo y grafo se consideran barbarismos.	'*Habligafo and *grafo are considered barbarisms.'
Ha**bló**grafo y grafo se consideran neologismos.	'*Hablografo and *grafo are conidered neologisms.'
El **blú**mental blando es su favorito.	'Soft blumethal is his/her favorite.'
Closed syllables	
Mi **lán**guido niño no se encuentra bien.	'My languid child does not feel well.'
El **ín**dice Nasdaq se desplomó ayer.	'The Nasdaq index plummeted yesterday.'
La al**bón**diga ésa era enorme.	'That meatball was huge.'
El **hún**garo rojo fue encarcelado.	'The red Hungarian was imprisoned.'
El **nál**garo azul se alimenta de insectos.	'The blue *nalgaro feeds on insects.
El **níl**daro rojo habita en Asturias.	'The red *nildaro lives in Asturias.'
La **nól**dina roja se alimenta de insectos.	'The red *noldina feeds on insects.'
La **múl**dina verde canta al amanecer.	'The green *muldina sings at dawn.'

A.2. Experiment 2 test sentences

La **lá**mina blanca está en la mesa.	'The white foil is on the table.'
El **Má**laga B jugará la final	'Málaga B will play the final game.'
La **Mó**nica hija no vendrá.	'Her daughter Monica will not come.'
La **nó**mina suya parece mayor.	'His/her salary seems higher.'
El **nó**dulo central tiene averías.	'The central node is damaged.'
La **mó**nada única se formó enseguida.	'The only monad formed quickly.'
El **nó**mada griego desapareció hace tiempo.	'The Greek nomad disappeared long ago.'
El **mó**dulo único es una de sus características.	'The single module is one of its characteristics.'
Mi **lán**guido niño no se encuentra bien	'My languid child does not feel well.'
Al Ándalus árabe y Castilla cristiana	'Arab Al Andalus and Christian Castile.'
El ámbito suyo es muy interesante	'His/her environment is very interesting.'
La ***mán**dola mágica suena muy bien	'The magic *mandola sounds very good.'
La **nól**dina roja se alimenta de insectos	'The red *noldina feeds on insects.'
El **ván**dalo nórdico invadió Europa	'The nordic Vandal invaded Europe.'
La al**bón**diga ésa era enorme	'That meatball was huge.'
El **nór**dico blanco le costó mucho	'The nordic white one cost him/her a lot.'

References

Arvaniti, A., & Gårding, X. (2007). Dialectal variation in the rising accents of American English. In J. Cole, & J. Hualde (Eds.), *Papers in laboratory phonology*, Vol. 9 (pp. 547–576). Berlin, New York: Mouton de Gruyter.

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 f0: Evidence from German. *Journal of Phonetics*, *32*, 177–197.

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. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*, 140–155.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*, 155–180.

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

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.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.

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.

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.

D'Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment*. Ph.D. thesis, Department of Linguistics, The Ohio State University.

D'Imperio, M. (2002). Language-specific and universal constraints on tonal alignment: The nature of targets and "anchors". In *Proceedings of the 1st international conference on speech prosody*, Aix-en-Provence, France, 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. Change in phonology*, Vol. 9 (pp. 577–608). Berlin, New York: Mouton de Gruyter.

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

Estebas-Vilaplana, E. (2000). *The use and realisation of accentual focus in Central Catalan*. Doctoral dissertation, University College London.

Estebas-Vilaplana, E., & Prieto, P. (2007). Production and perception of word-edge tones in Catalan and Spanish. Manuel González González, Elisa Fernández Rei and Begoña González Rei, (Eds.), Actas del III Congreso de Fonética Experimental, Santiago de Compostela, October 24–26, 2005 (pp. 279–290). Xunta de Galicia: Universidad de Santiago de Compostela.

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 Sp-ToBI". *Journal of Portuguese Linguistics* (special issue on Prosody of Iberian Languages, G. Elordieta & M. Vigário (Eds.), *5*(2), 1–32.

Gao, M. (2006). Gestural representation and alignment patterns of mandarin tones. Paper presented at *the 10th conference on laboratory phonology*, Paris, June 29–July 1, 2006.

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.

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.

Hellmuth, S. (2005). Pitch accent alignment in Egyptian Arabic: Exploring the boundaries of cross-linguistic alignment variation. Paper presented at *II phonetics and phonology in Iberia*, PaPI 2005, Bellaterra, June 20–21.

Hellmuth, S. (2006). *Intonational pitch accent distribution in Egyptian Arabic*. Ph.D. School of Oriental & African Studies, University of London. ⟨http://www.sfb632.uni-potsdam.de/homes/samhellmuth/⟩.

Igarashi, Y. (2004). 'Segmental anchoring' of f0 under changes in speech rate: Evidence from Russian. In B. Bel, & I. Marlien (Eds.), *Proceedings of the speech prosody 2004 conference*, Nara, Japan, March 23–26, 2004, pp. 25–28. ISCA Archive ⟨http://www.isca-speech.org/archive/sp2004⟩

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*, Vol. 1 (pp. 615–618). Barcelona: Causal Productions.

Ishihara, T. (2006). *Tonal alignment in Tokyo Japanese*. Ph.D. thesis, University of Edinburgh.

Knight, R.-A., (2003). *Peaks and plateaux: The production and perception of intonational high targets in English*, Ph.D. dissertation, University of Cambridge.

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

Ladd, D. R. (2003). Phonological conditioning of f0 target alignment. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the XVth international congress of phonetic sciences*, Vol. 1 (pp. 249–252). Barcelona: Causal Productions.

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., 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.

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.

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. (2006a). The relevance of prosodic structure in tonal articulation. Edge effects at the syllable and prosodic word level in Catalan and Spanish. In: *International conference on the typology of tone and intonation (TTI)*. ZAS and University of Potsdam, September 7–9, 2006. Berlin, Germany.

Prieto, P. (2006b). 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., Estebas-Vilaplana, E., & Vanrell, M. M. (2006). The role of tonal alignment and slope of the rise on word-boundary identification in Catalan and Spanish. Paper presented at *the 10th conference on laboratory phonology*, Paris, June 29–July 1, 2006.

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

Rietveld, T., & Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: Effects of syllable structure. *Journal of Phonetics*, *23*, 375–385.

Saltzman, E. (1991). The task dynamic model in speech production. In H. Peters, W. Hulstijn, & W. Starkweather (Eds.), *Speech motor control and stuttering* (pp. 37–52). Amsterdam: Excerpta Medica.

Saltzman, E., & Kelso, J. A. S. (1987). Skilled actions: a task dynamic approach. *Psychological Review*, *94*, 84–106.

Saltzman, E. L., & Munhall, K. G. (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, *1*(4), 333–382.

Schepman, A., Lickley, R., & Ladd, D. 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.

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

Steele, S. A. (1986). Nuclear accent f0 peak location: Effects of rate, vowel, and number of following syllables. *Journal of the Acoustical Society*, *80*(Suppl. 1), s51.

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

Tuller, B., & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research*, *34*, 501–508.

van Santen, J. P. H., & Hirschberg, J. (1994). Segmental effects on timing and height of pitch contours. In *Proceedings of the 1994 international conference on spoken language processing*, pp. 719–722.

Vanrell, M. M. (2006). *The phonological role of tonal scaling in Majorcan Catalan interrogatives*, MA thesis, Department of Catalan Philology, UAB. 〈http://seneca.uab.es/ggt/Tesis/MVanrell.pdf〉

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

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

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

Xu, Y. (1999). F0 peak delay: When, where and why it occurs. In *Proceedings of the 14th international congress of phonetic sciences*, (pp. 1881–1884). San Francisco.

Xu, Y. (2002). Articulatory constraints and tonal alignment. In B. Bel, & I. Marlien (Eds.), *Proceedings of the 1st international conference on speech prosody* (pp. 91–100). France: Aix-en-Provence, Laboratoire Parole et Langage.

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.

Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, *33*, 319–337.

Further reading

Andreeva, B., & Oliver, D. (2005). Information structure in Polish and Bulgarian: Accent types and peak alignment in broad and narrow focus. In S. Franks, F. Y. Gladney, & M. Tasseva-Kurktchieva (Eds.), *Formal approaches to Slavic Linguistics 13: The South Carolina meeting* (pp. 1–13). Ann Arbor, MI: Michigan Slavic Publications.

Arvaniti, A., Ladd, D. R., & Mennen, I. (2000). What is a starred tone? Evidence from Greek. In M. Broe, & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Acquisition and the Lexicon* (pp. 119–131). Cambridge: Cambridge University Press.

Boersma, P., & Weenink, D. (1992–2001). *Praat: A system for doing phonetics by computer*, available at ⟨http://www.praat.org⟩.

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.

Fougeron, C., & Jun, S.-A. (1998). Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics*, *26*, 45–69.

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

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.), *Action to language via the mirror neuron system*. CUP, pp. 215–249.

House, D. (1990). *Tonal perception in speech*. Lund: Lund University Press.

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

Ladd, D. R., & Schepman, A. (2003). Sagging transitions between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, *31*, 81–112.

Petrone, C. (2006). The role of tonal alignment in the realization of Radoppiamento Sintattico in Italian. Paper presented at *the Xth conference on laboratory phonology*. Laboratoire de Phonétique et Phonologie, CNRS/U. Paris 3. June 29–July 1, 2006.

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., Grice, M., Mücke, D., & Becker, J. (in prep). *Articulatory alignment patterns for rising and falling pitch accents in Catalan*.

Prieto, P., & Torreira, F. (2004). Anchoring effects in Spanish. In *ESF international conference on tone and intonation*. Santorini, Greece, September 9–11, 2004.

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

Welby, P. (2003). *The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. Doctoral dissertation, The Ohio State University.

Welby, P. (2004). The structure of French intonational rises: A study of text-to-tune alignment. In B. Bel, & I. Marlien (Eds.), *Proceedings of the 1994 speech prosody conference* (pp. 127–130). Japan: Nara.