TONAL REALIZATION OF SYLLABIC AFFILIATION IN SPANISH

Francisco Torreira

University of Illinois at Urbana-Champaign
ftorrei2@uiuc.edu

ABSTRACT

This study explicitly addresses the hypothesis that pitch accents are aligned to syllables by examining the F0 shape of rising accents over similar segmental material differing in syllabic affiliation. A consistent correlate of syllabic affiliation is found in the timing of the rising F0 movement with respect to an independent segmental landmark. Our results provide phonetic evidence of resyllabification and of the role played by the syllable as an abstract unit guiding speech production.

Keywords: prosody, intonation, syllable, tonal alignment, Spanish.

1. INTRODUCTION

Research in tonal alignment during the last decade has repeatedly addressed the hypothesis that tonal events are coordinated with defined points in the speech chain. For the most part, these studies have tried to underpin the alignment of local F0 movements with reference to acoustic landmarks such as segmental boundaries. One recurrent finding in these studies is that the start of rising accents is consistently located at the beginning of their host syllable, even when the accents are produced under adverse conditions like fast speech rates or tonal pressure [11, 1]. On the other hand, considerable variability has been found in the alignment of F0 peaks, especially in the case of prenuclear accents [11, 7, 12]. This lack of invariance at the acoustic level has motivated a series of articulatory studies aimed at capturing possible coordination patterns between tonal events and oral gestures [3, 4]. In [9], for instance, it is claimed that the timing of F0 turning points in German high accents is better captured with reference to minima and maxima in the kinematic signals of a consonantal constriction than to acoustically defined events. Yet at the present time it is still unclear to what extent tones and pitch accents are aligned with acoustic or gestural landmarks.

The results of the studies above also lead to think that tonal events might be coordinated with more abstract units of prosodic organization such as the syllable. However, this possibility still remains largely unexplored. To our knowledge, only [8] could explicitly address the hypothesis that pitch accents are aligned with the syllable *as opposed to* other units or events in speech (e.g. oral constrictions or the end the vowel *per se*). This study effectively showed that the location of pitch valleys was sensitive to the syllabic affiliation of the medial part of English pairs such as *Norma Nelson vs. Norman Elson*. Even if such findings clearly support the syllabic F0 alignment hypothesis for English, more evidence is needed before its conclusions can be extended to other prosodic contexts and languages. For example, it is still not known if other syllabic phenomena such as resyllabification or syllabification contrasts between hiatus and diphthongs have correlates in F0 timing. The present study explicitly addresses this issue by examining the production of rising pitch accents in Spanish in similar segmental sequences differing in syllabic affiliation.

2. THE STUDY

The goal of this study was to examine the effect of syllabic affiliation on the timing of rising pitch accents. More particularly, our hypothesis was that, all other things being equal, changes in the location of the syllable boundary would alter the location of the start of the F0 rise.

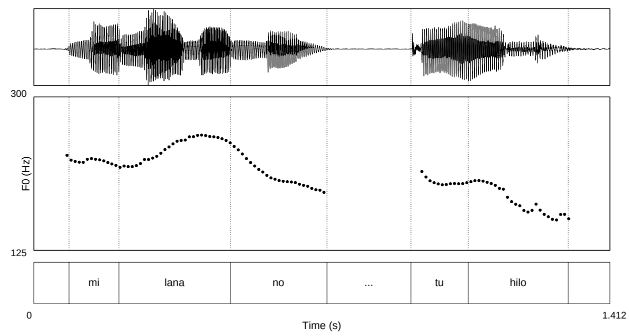
2.1. Materials

In order to test our hypthesis, we selected five utterance types fairly similar in segmental composition but differing in syllabic structure. Given the scarcity of instrumental studies on correlates of the syllable in Spanish (but see [10]), we decided to follow syllabification schemes found in conventional phonological analyses [5]:

1. CANONICAL syllable: *mi lana* [mi.**la**.na]
2. RESYLLABIFIED consonant: *el ama* [e.**la**.ma]
3. GEMINATE consonant: *el lama* [el.**la**.ma]
4. HIATUS: *mi li.a.na* [mi.li.**a**.na]
5. DIPHTHONG: *Emiliana* [e.mi.**lja**.na]

20 test utterances were chosen following these models. All of them contained sequences composed of the following elements: 1) a singleton or geminate lateral coronal consonant /l/ ∼ /l.l/ 2) one of

Figure 1: Carrier sentence for the utterance *mi lana* 'my wool'.



the vocoid sequences /i.a/, /ja/ or /a/. 3) a nasal consonant /n/ ∼ /m/. These sequences were all preceded by a front vowel /e/ or /i/ and followed by either /a/ or /o/. The test utterances were embedded in the carrier sentence *X no, tu Y*, 'not X, your Y', where X was the test utterance and Y a random noun (e.g. *mi lana no, tu hilo* 'not my wool, but your thread'). In Standard Peninsular Spanish, this kind of sentence was expected to display a distinct intonation contour: a steep rise from the first stressed syllable of our target utterances up to the postaccentual syllable, and then a downstepped high accent on the word *no* followed by another prosodic phrase irrelevant for our purposes. The carrier sentence is illustrated in Fig 1.

The recordings were conducted in a silent room using a Shure SM110 head-mounted microphone and an M-Audio 410 FireWire external soundcard connected to a computer. 5 speakers of Peninsular Spanish (3 female and 2 male) read 5 times a list of sentences displayed on a computer screen. No particular instructions were given as how to read the sentences. Tokens exhibiting critical disfluencies were rerecorded at the end of the session. The total database consisted of 500 tokens (5 utterance types x 4 testword variants x 5 repetitions x 5 speakers).

2.2. Measurements

F0 contours were extracted using the autocorrelation method available in Praat set to default values except for a time step of 5 ms instead of 10 ms. No smoothing procedure was applied to the F0 contours. Given that our hypothesis explicitly involved the timing of the start of the rising accent, its location relative to a landmark common across all utterance types appeared as a crucial measurement for the experiment. Two main obstacles encouraged us to discard manual marking of this point. First, F0 perturbations in the critical region of measurement, that is around the consonantal syllable onset, made segmentation decisions too arbitrary in many occasions. Second, F0 contours often rose in a gradual way, with little increase in F0 velocity occurring during a considerable part of the F0 rise. To this we must add possible auto-suggestion effects, since the segmentation was to be carried out by the author himself. In view of all these difficulties, it was decided that the beginning of [l] would be our only manual marking for all of our utterance types. This point was defined as the decrease in formant and overall amplitude which characterizes the vowel-to-[l] transition. Since its acoustic characteristics were sufficiently stable, it would provide a robust reference point for our F0 timing measurements. The TARGET CONTOUR (TC) was then defined as the stretch of F0 going *from the start of [l] to the upcoming F0 peak for all utterance types*. The reason for choosing this interval is that we expected syllable boundaries to vary with utterance type precisely within its bounds (e.g. [emi.lja.na] vs. [la li.a.na]), and that this variation would allow us to identify syllabic affiliation differences by examining the shape of our TC. Fig. 3 shows raw TARGET CONTOURS before normalization for four utterance types as produced by speaker MC.

Once all the TC were extracted, we decided to focus on a basic property of their shape that would allow us to differentiate early from late rises with respect to our segmental [l] landmark: TC were normalized in time and fitted with quadratic polynomials of the form $ax^2 + bx + c$, where a can be readily interpreted as representative of curvature, b as a measure of initial slope and c as the intercept or initial F0. All other things being equal, early rises should display low or even negative curvature, while late rises should be better fitted by models with high curvature coefficients. With regard to our hypothesis, we expected that the later the syllable boundary would lie with respect to the [l] landmark, the greater the F0 curvature would be. This rationale is illustrated in (Fig. 2).

Figure 2: Ideal TARGET CONTOURS over different syllabic materials as predicted by our hypothesis.

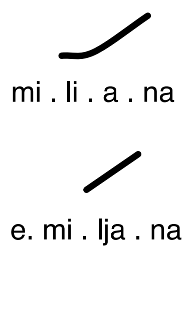


Table 1: Mean curvature coefficients with standard deviations for each speaker across every utterance type.

Speaker	Canon.	Resyll.	Hiatus	Diphthong	Geminate
AM	6.8 (16.3)	10.2 (12)	42.3 (17.3)	12.1 (11.5)	52.2 (28.1)
CP	43.5 (27.5)	51.2 (44.1)	104.3 (22.3)	36.1 (25.3)	100.7 (31.1)
MC	10.1 (15.6)	9.5 (20)	49.1 (20.5)	8.9 (17)	16.2 (16.2)
MM	-1.6 (25.1)	0.5 (26.3)	130.9 (34.3)	-30.2 (34.3)	32.6 (31.5)
SM	15.9 (17.7)	20.9 (28.5)	67.7 (23.3)	30.1 (16.6)	52.4 (18.4)

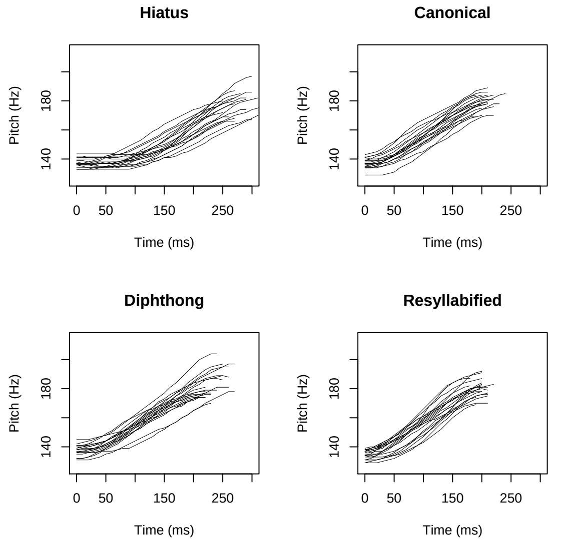
3. RESULTS

Our hypothesis predicted that differences in syllable affiliation would imply differences in the shape of our TARGET CONTOURS, since these were all extracted from a common landmark independent of the syllable boundary location. For example, in the utterance *mi liana* /mi li.**a**.na/, with stress on the third syllable, we expect our TC to exhibit a level pitch well into the sequence [li] and then start rising when approaching stressed vowel [a]. On the other hand, in stressed diphthongs as in *Emiliana* [e.mi.**lja**.na], the TC should display an earlier or even immediate rise (Fig. 2). Another prediction implied by our hypothesis is that canonical syllables and syllables with resyllabified [l] onsets would exhibit similar F0 shapes. Finally, TC including clusters of two [l] sounds (e.g. *el lama*) should display a late rise when compared to the TC of canonical syllables, since only the second [l] in geminate clusters properly belongs to the stressed syllable according to our working assumptions. It was expected that all of these patterns would translate into F0 curvature values.

Table 1 shows mean curvature values and standard deviations for each speaker across the different utterance types. Separate analyses of variance were carried out for each speaker with CURVATURE as the dependent variable and UTTERANCE TYPE as the only independent factor. Statistically significant effects were found in each case (speaker AM: $F(4,94) = 26.82, p < .001$; speaker CP: $F(4,94) = 22.16, p < .001$; speaker MC: $F(4,94) = 18.32, p < .001$; speaker MM: $F(4,94) = 80.92, p < .001$; speaker SM: $F(4,94) = 21.16, p < .001$). Tukey HSD post-hoc comparisons revealed the following patterns relevant for our hypothesis:

- CANONICAL, RESYLLABIFIED and DIPHTHONG utterance types displayed similar curvature coefficients for four of our five speakers. DIPHTHONG values in speaker MM were different from all other groups.
- HIATUS and DIPHTHONGS displayed different curvature coefficients for all speakers.

Figure 3: TARGET CONTOURS for four utterance types (Speaker MC). The contours are shown before time normalization.



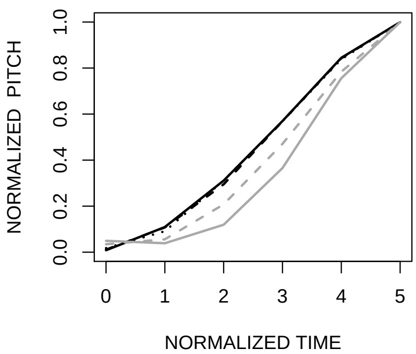
- GEMINATE was different from CANONICAL, RESYLLABIFIED and DIPHTHONG for all speakers except for speaker MC.

Fig. 4 summarizes the data. In order to draw the lines in Fig. 4, six equidistant F0 points were extracted within each TARGET CONTOUR. These F0 values were then normalized with respect to the F0 range of the rising accent and averaged across all repetitions and speakers. It can be seen from this figure that the three black lines representing CANONICAL, RESYLLABIFIED and DIPHTHONG utterances largely overlap, while the two grey lines representing HIATUS (solid) and GEMINATE (dashed) types display a later rise.

4. DISCUSSION

In general terms, the results of the experiment supported our hypothesis. First of all, resyllabification was found to apply at the tonal level for all of our five speakers. Similarly low levels of curvature in the CANONICAL and RESYLLABIFIED utterance types suggest that the start of the F0 rise occurred briefly after the beginning of [l] in both of these groups. The comparison between HIATUS and DIPHTHONGS followed the predictions made by the syllabic alignment hypothesis. HIATUS utterances exhibited the highest curvature coefficients of all groups in four speakers, while DIPHTHONG patterned with CANONICAL and RESYLLABIFIED in four of our five speakers. Finally, GEMINATE tokens displayed higher curvature than CANONICAL, RESYLLABIFIED and

Figure 4: Normalized contours averaged across all speakers. Black lines represent CANONICAL, RESYLLABIFIED and DIPHTHONG utterance types. Grey lines represent HIATUS (solid) and GEMINATE (dashed) utterance types.



DIPHTHONG utterances for four of our five speakers. All in all, it can be said that traditional syllabification schemes agreed well with the F0 shapes (and therefore with the F0 timing patterns) found.

We see two main implications in these results. First, F0 appears as a robust correlate of resyllabification. In our Spanish sentences, F0 rising contours over canonical and resyllabified onsets had a similar shape. This is an important finding, especially if we consider that numerous studies have failed to find consistent correlates of resyllabification in languages where it is widely assumed to occur (e.g. French). Second, our results undermine the possibility that tonal movements are coordinated with raw supraglottal gestures (e.g. consonantal constrictions) regardless of the syllabic organization of the utterance. Two minimal pairs in our database, [mi.li.a.na] vs. [e.mi.lja.na] and [e.la.ma] vs. [el.la.ma] involved rather similar articulators and articulator trajectories, but clearly differed in terms of F0 (and most probably interarticulator) timing. It follows from this that F0 timing is most sensitive to the location of an abstract unit of gestural, or perhaps perceptual, organization such as the syllable [2, 6, 13]. Only once this association has been identified will it be possible to understand local coordination patterns between F0 movements and particular oral gestures.

5. CONCLUSION

This study has provided evidence for alignment between the start of rising pitch accents and the syllable in Spanish. In a production experiment, the syllabic organization of a segmental sequence was clearly identified as capital factor governing the start of rising accents. Finally, we have suggested that raw gestures or segmental boundaries *per se* are not likely to act as F0 alignment anchors without consideration of the syllabic organization of the utterance.

6. REFERENCES

[1] Arvaniti, A., Ladd, R. D., Mennen, I. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26, 3–25.
[2] Browman, C., Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* 201–251.
[3] D'Imperio, M., Espesser, R., Lœvenbruck, H., Menezes, C., Nguyen, N., Welby, P. to appear. Are tones aligned with articulatory events? evidence from Italian and French. In: Hualde, J. I., Cole, J., (eds), *Papers in Laboratory Phonology 9*.
[4] Gao, M. 2006. Gestural representation and alignment patterns of mandarin tones. Presented at the 10th Conference on Laboratory Phonology, Paris, 2006.
[5] Hualde, J. I. 2005. *The Sounds of Spanish*. Cambridge University Press.
[6] Krakow, R. 1999. Physiological organization of syllables: a review. *Journal of Phonetics* 27, 23–54.
[7] Ladd, R. D., 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.
[8] Ladd, R. D., Schepman, A. 2003. "sagging transitions" between high pitch accents in english: experimental evidence. *Journal of Phonetics* 31, 81–112.
[9] Mücke, D., Grice, M., Becker, J., Hermes, A., Baumann, S. 2006. Articulatory and acoustic correlates of prenuclear and nuclear accents. *Proceedings of Speech Prosody 2006, Dresden*.
[10] Prieto, M. 2002. *Phonetic Correlates of the Syllable: Evidence from Spanish*. PhD thesis University of Illinois at Urbana-Champaign.
[11] Prieto, P., van Santen, J., Hirschberg, J. 1995. Tonal alignment patterns in spanish. *Journal of Phonetics* 23, 429–451.
[12] Prieto, P., Torreira, F. to appear. The segmental anchoring hypothesis revisited: syllable structure and speech rate effects on peak timing in spanish. *Journal of Phonetics*.
[13] Xu, Y. in press. Tonal alignment, syllable structure and coarticulation: Toward an integrated model. *Italian Journal of Linguistics*.