COORDINATION PATTERNS BETWEEN PITCH MOVEMENTS AND ORAL GESTURES IN CATALAN

*Pilar Prieto**, *Doris Mücke*^, *Johannes Becker*^ & *Martine Grice*^

*ICREA-Universitat Autònoma de Barcelona, ^I*f*L Phonetik, University of Cologne

pilar.prieto@uab.es, {doris.muecke; becker.johannes; martine.grice}@uni-koeln.de

ABSTRACT

In this paper, we investigate the coordination relations between F0 turning points in bitonal pitch accents and landmarks of dynamically defined articulatory gestures in Catalan, using kinematic and acoustic data on three pitch accent types. Electromagnetic articulography data (EMMA) reveals that the end of pitch movements for two rising and one falling accent are tightly synchronized with the *peak velocity* of the oral closing gestures (such as tongue tip raising) during the production of vowel-consonant sequences.

Keywords: articulatory anchors, tonal alignment, tune-text association.

1. INTRODUCTION

Recent acoustic investigations of tonal alignment have shown that the start and end of a bitonal pitch accent can be synchronized with the segmental structure in consistent ways. For example, there is robust evidence in favor of the stability of L in rising accents, which are found to be consistently 'anchored' with the onset of the accented syllable in a variety of languages (see Arvaniti et al. [1] for Greek; and Estebas-Vilaplana [4] for Catalan, among others). Yet the behavior of H peaks in rising accents is typically more variable. Even when H peaks are free of tonal pressure, they have been shown to be consistently affected by syllable structure in many languages (for Neapolitan Italian, D'Imperio [2, 3]; for Spanish, Prieto & Torreira [10], and references therein).

Recent work by D'Imperio et al. [3] and Mücke et al. [8, 9] on the coordination patterns between F0 targets and supraglottal gestures (Goldstein et al. [5]) in Neapolitan Italian and German has suggested that alignment with articulatory gestures is more stable than alignment with acoustic landmarks. They found that the end of pitch movements in bitonal accents are tightly synchronized with the closing gesture of C2 in CV.C and CVC sequences. However, in those papers only rising accents were investigated. It is our goal here to explore the stability of peak alignment in two rising and one falling pitch accent type in Catalan and ultimately to test whether the tonal targets in all three have anchor points in the articulatory domain. We investigate the behavior of tonal targets in both open and closed syllables, as syllable type has been shown to have an effect on target location in the acoustic domain ([2], [3], [10]).

2. METHOD

2.1. Materials

2.1.1. Target utterances

The following 14 target words containing both open and closed syllables were used for the experiment. Target words include the following segmental string: C = nasal [n, m], V1 and V3 = high vowel [i], target V2 = low vowel [a]).

Block 1: LABIALS

CLOSED		OPEN
[mi.**mam**.zi]	vs.	[mi.**ma**.mi]
[mi.**mam**.li]		
[mi.**mam**.zi.la]	vs.	[mi.**ma**.mi.la]
[mi.**mam**.li.la]		

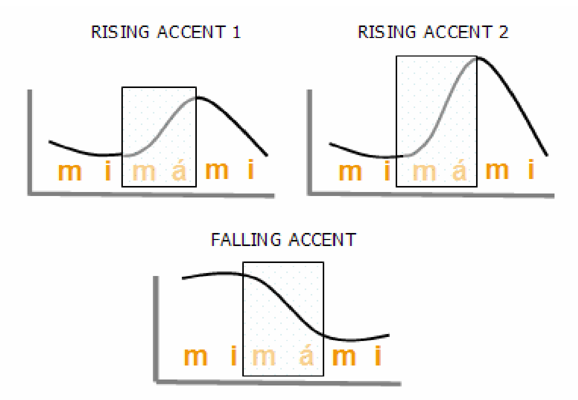
Block 2: ALVEOLARS

CLOSED		OPEN
[ni.**nan**.mi]	vs.	[ni.**na**.ni]
[ni.**nwan**.mi]	vs.	[ni.**nwa**.ni]
[ni.**nan**.mi.la]	vs.	[ni.**na**.ni.la]
[ni.**nwan**.mi.la]	vs.	[ni.**nwa**.ni.la]

2.1.2. Pitch accent types

The three pitch accents under investigation are, on the one hand, two types of rising accents with aligned peaks, namely the *nuclear broad focus* (or rising accent 1) and the *nuclear narrow contrastive focus* (or rising accent 2, with a higher peak), and, on the other, the *nuclear accent in yes-no questions* (a falling accent). Crucially, previous acoustic data on these accent types have shown that both the L and the H turning points are synchronous with either the onset or the offset of the accented syllable (Estebas-Vilaplana [4]). Figure 1 shows a schematic diagram of these accent types:

Figure 1: Target pitch accent types in Catalan. The schematic alignment shows that both L and H turning points are aligned with the syllable edges.



In order to elicit the three pitch accent types, the target string was placed in meaningful utterances designed to trigger the intended target pitch accent.

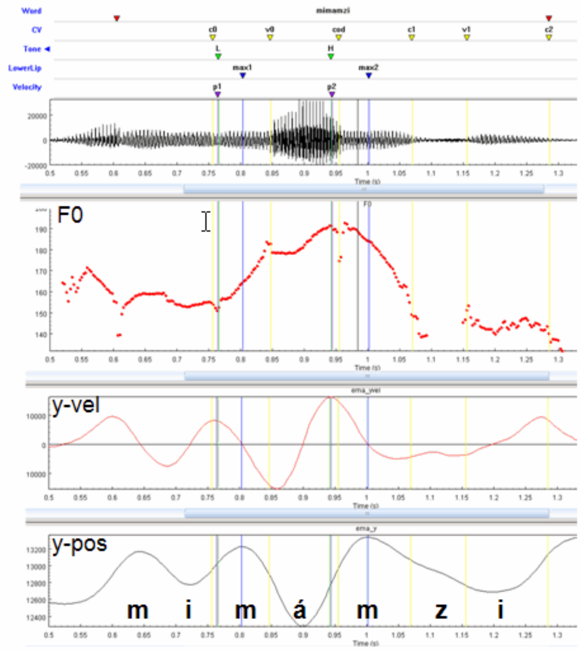
2.2. Recordings

One female speaker of Central Catalan (PP) participated in the experiment. Recordings were made with the Carstens articulograph AG100 and a time-synchronized DAT recorder at the University of Cologne. Sensors were placed on the nose, lower lip, tongue tip and tongue body (4cm behind the tip). The kinematic data were recorded at 400Hz, downsampled to 200Hz and smoothed with a 40Hz low-pass filter. The acoustic data were digitized at 44.1kHz. The subject read the target utterances from a monitor at a normal rate of speech, for a total of 210 tokens (14 target utterances x 3 accent types x 5 repetitions).

2.3. Labeling

The acoustic and articulatory data were hand labeled with Praat and the speech database system, EMU. For each data file, the following acoustic measures were taken: (1) *c0*, or start of the initial consonant in the accented syllable; (2) *v0*, or start of the vowel in the accented syllable; (3) *k0*, or start of the coda in the accented syllable; (4) *c1*, or start of the initial consonant in the postaccentual syllable. As for the F0 data, the two basic L and H pitch turning points were labeled. In the kinematic data, the following timing points were located for the consonantal gestures: (1) *max1*, or maximum degree of constriction (lower lip and tongue tip) in C1; (2) *max2*, or maximum degree of constriction in C2; (3) *p1*, or peak velocity of constriction movement (lower lip or tongue tip) during the production of C1; (4) *p2*, peak velocity of constriction movement during the production of C2. Articulatory maxima were located at zero crossings in the respective velocity trace and peak velocities at zero crossings in the respective acceleration trace. Figure 2 illustrates the labeling scheme used for closed syllables in narrow focus accents:

Figure 2: Labeling scheme, acoustic waveform, f0 contour, velocity and position (displacement) for the vertical movement of lower lip (high displacement indicates that lips are closed). Test word *mimamzi*.



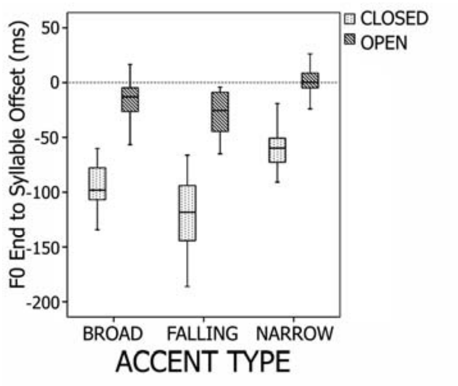
3. RESULTS

All 210 tokens were included in the analysis.

3.1. Acoustic results

In this section, we examine the alignment behavior of the end of the pitch movement with respect to the end of the syllable and the potential effects of syllable structure. Figure 3 shows the median and quartiles of the distance in ms from the end of the F0 movement to the end of the accented syllable for open and closed syllables in the three accent types. The results confirm that the end of the F0 movement is strongly retracted in closed syllables. It preceded the end of the syllable by, on average, 89ms for falling accents, 78ms for broad focus rises, and 60ms for narrow focus rises. By contrast, in open syllables, the end of the F0 movement is roughly synchronized with the end of the syllable. A factorial ANOVA was run with F0-End to Syllable Offset as the dependent variable and Syllable Type as the main factor and Accent Type as an additional factor. Results reveal a significant main effect of Syllable Type [F(1, 199) = 594.045, p<0.001] and Accent Type [F(2, 199) = 64.413, p<0.001], with a significant interaction between the two factors.

Figure 3: Median distance in ms from the end of the F0 movement to the end of the accented syllable for closed and open syllables and the three accent types.

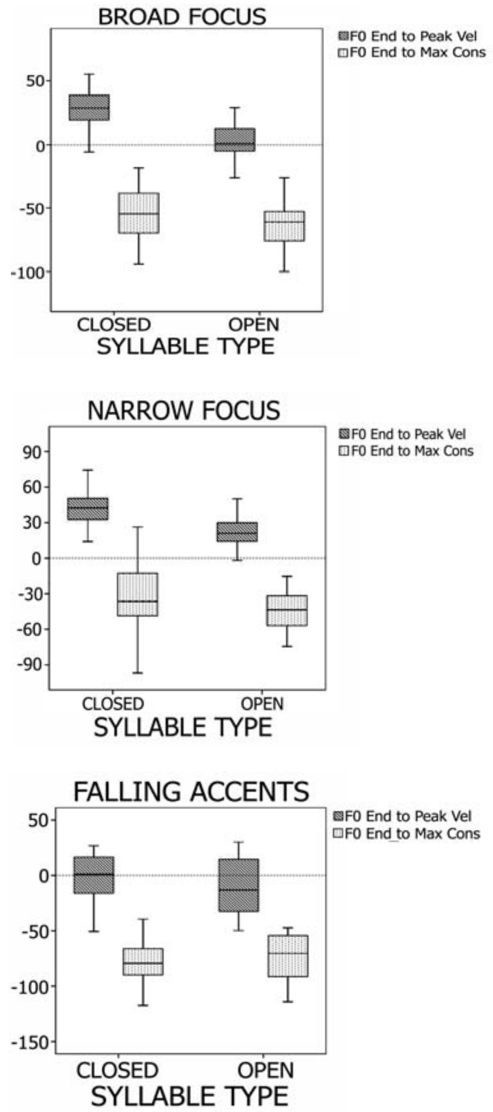


3.2. Articulatory results

In this section we investigate the coordination patterns between the end of the three pitch accents with respect to two potential articulatory anchors, namely, the maximum constriction and the peak velocity of labial [m] and alveolar [n] consonant gestures for C2. The three box plots in Figure 4 show for each accent the median distance (in ms) from the end of the F0 movement to the maximum constriction of C2 and to the peak velocity of the C2 constriction movement for both open and closed syllables.

The results show that for the three pitch accent types, the end of the F0 movement, either H or L, is better synchronized with the peak velocity of the C2 closing gesture (lower lip and tongue tip movement) than with the target (a value of 0 represents cooccurrence). However, there are differences across the three accent types. In narrow focus accents, we found an effect of syllable structure on the alignment of H relative to the C2 closing gesture: H was later in relation to peak velocity in closed syllables than in open ones. This might be due to the fact that in closed syllables there are two consonants to be realised (e.g. *mimamzi* vs. *mimami*). This could affect the stiffness of the closing movement, causing the peak velocity to be reached earlier.

Figure 4: Median distance in ms from the end of F0 movement to the maximum constriction for C2 (striped bars) and to the peak velocity for C2 (dotted bars) for both open and closed syllables.



The H target for the narrow focus rising pitch accent is somewhat later aligned than the H for the broad rising one. However, since narrow focus accents have a higher peak, it is not surprising that the peak is later (see [6], where delay can substitute for height). The difference between these two accents could be expressed through the height of the peak, as later alignment might be a result of the peak height difference. Finally, we found the tightest synchronization in falling accents. L coincides with the peak velocity of the C2 closing gesture.

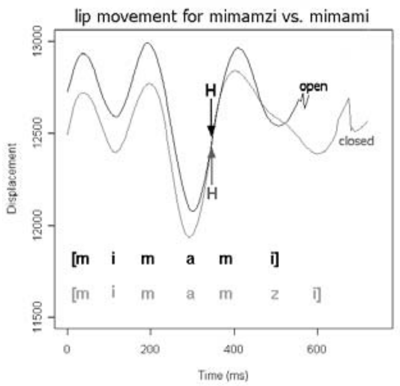
A factorial ANOVA was run with F0-End to Peak Velocity as the dependent variable, Syllable Type as the main factor and Accent Type as an additional factor. Results show a significant main effect of Syllable Type [F(1, 201) = 33.688, p<0.001], with a significant interaction between the two factors. However, for closed syllables, latencies in the articulatory domain are considerably smaller than in the acoustic one (see section 3.1).

4. DISCUSSION AND CONCLUSIONS

Results from articulographic data for one Catalan subject thus reveal that the final target for three types of bitonal pitch accents is relatively well synchronized with the peak velocity of the closing gesture during the production of C2.

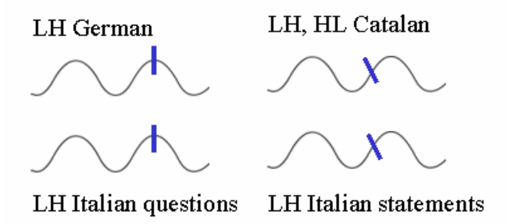
Crucially, peak velocity was the articulatory anchor for both open CV.C and closed CVC syllables. This stability of alignment contrasts with the higher variability found in the acoustic domain. The averaged lower lip trajectories in Figure 5 (broad focus accents) illustrate how the position of the H peak coincides with the peak velocity in both open and closed syllables.

Figure 5: Averaged contours of vertical lower lip trajectories for open and closed syllables for the target words [mimam.zi] vs. [mima.mi]. Arrows mark the location of H relative to the lower lip movement.



Finally, it is interesting to note that landmarks in articulatory movements for consonants are restricted to the closing gesture –something which is true for other languages besides Catalan. However, German and Catalan differ in that the former aligns the end of F0 movements with the maximum constriction, while the latter aligns them with peak velocity of the same closing gesture. That is, there are different anchors within the same articulatory gesture. Such an apparently small alignment difference may be used to make (or contribute towards making) phonological distinctions, as in Neapolitan, where H in L*+H (questions) aligns with the maximum constriction, and H in L+H* (statements) with peak velocity (see D'Imperio et al. [3]). This is shown schematically in Figure 6.

Figure 6: Alignment patterns for end of pitch movement relative to C2 closing gestures in different languages.



However, we show that such alignment differences can be found in cross-language comparisons (German, Catalan), providing further evidence for the small but systematic phonetic alignment differences discussed by Ladd [7].

5. REFERENCES

[1] Arvaniti, A., Ladd, D.R., Mennen, I. 1998. Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics* 26, 3-25.

[2] D'Imperio, M. 2000. The role of perception in defining tonal targets and their alignment. Doctoral dissertation. The Ohio State University.

[3] D'Imperio, M., Espesser, R., Lœvenbruck, H., Menezes, C., Nguyen, N., Welby, P. in press. Are tones aligned with articulatory events? Evidence from Italian and French. In: Cole, J., Hualde, J.I. (eds), *Papers in Laboratory Phonology IX*. The Hague: Mouton.

[4] Estebas-Vilaplana, E. 2000. The use and realisation of accentual focus in Central Catalan. Doctoral dissertation. University College London.

[5] Goldstein, L., Byrd, D., Saltzman, E. 2006. The role of vocal tract gestural action units in understanding the evolution of phonology. In: Arbib, M. (ed), *Action to Language via the Mirror Neuron System*. Cambridge: Cambridge University Press, 215-249.

[6] Gussenhoven, C. 2004. *The phonology of tone and intonation*, Cambridge: CUP

[7] Ladd, D. R. 2007. Segmental anchoring of pitch movements: autosegmental association or gestural coordination? *Italian Journal of Linguistics* 18.1, 19-38.

[8] Mücke, D., Grice, M., Becker, J., Hermes, A., Baumann, S. 2006. Articulatory and acoustic correlates of prenuclear and nuclear accents. In: *Proc. of Speech Prosody 2006* Dresden, 297-300.

[9] Mücke, D., Grice, M., Becker, J., Hermes, A. (submitted). Sources of variation in tonal alignment: evidence from acoustic and kinematic data.

[10] Prieto, P., Torreira, F. 2007. The segmental anchoring hypothesis revisited. Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics.*