BASIC MELODIC CONTOURS IN MAJORCAN CATALAN
SHAPE AND ALIGNMENT OF PITCH MOVEMENTS IN DECLARATIVES

MIQUEL SIMONET[1]
University of Illinois at Urbana-Champaign

1. *Introduction*

Research on Catalan intonation has almost always been focused on Central Catalan, spoken in Barcelona and Central Catalonia (Prieto, 2002; Estebas 2003, among many others). This variety, as it occurs with many metropolitan varieties, is the result of leveling processes motivated by intensive dialect contact in the city, due to growth and migration from different dialectal areas. If our long-term goal were to be able to reconstruct previous stages of Catalan (and Romance) intonation, it is clear that we would need data from other dialects as well. This is not unique to research on Catalan, as intonational studies on many Romance languages have been focused on the prestige dialects of these languages. The present paper is a contribution to this task: We take a rather descriptive approach and analyze in detail the shape of melodic contours in declaratives (in read speech) in Majorcan Catalan. Alignment patterns are also analyzed in detail, with respect to two time references.

Majorcan Catalan (MC) is a dialect of Catalan spoken by about 500,000 speakers in the Western Mediterranean island of Majorca. Recasens (1998) points out that the phonetics and phonology of MC present autochthonous characteristics, conservative features (of Old Catalan), and also features of other Romance languages with which the variety was in early contact – Occitan and Italian. The presence in MC of autochthonous and conservative characteristics is unsurprising if we take into consideration the peripheral and isolated status of MC with respect to the dialects spoken in the mainland: the relative isolation of this variety is a likely reason why conservative features have been retained.

The intonation of Balearic Catalan (BC is the umbrella term for Majorcan, Minorcan and Eivissan Catalan, the dialects spoken in the Balearic islands, of

[1] I am grateful to the following people for comments and/or assistance: José I. Hualde, Pilar Prieto, Erin O'Rourke, Laura Colantoni, Erik Willis, one anonymous reviewer, and Eunice Díaz. Thanks are due to the editors of this volume and to the audience of LSRL, 2006.

which Majorca is the largest one) has received more attention in recent times (Mascaró, 1985, 1986; Payà and Vanrell, 2005; Vanrell, 2006). However, there are no instrumental studies of declarative speech. This is the first one.

1.1 *What we know about Romance intonation: a basic typology*

Hualde (2004) offers a brief typology of Romance pitch accents, described within the tenets of the Autosegmental-Metrical approach to intonational phonology (Pierrehumbert, 1980; Ladd, 1996). Abstracting away from the labeling of these pitch accents in AM annotation, the generalization is that most of the Romance varieties that have been studied so far present a basic structure of three melodic contours in declaratives, which seem to be associated with stressed syllables. First, prenuclear or nonfinal pitch accents tend to show rising pitch movements whose peaks usually fall on the post-tonic syllable (the reason/s for this are yet to be identified), and whose valleys tend to coincide with the beginning of the stressed syllable. Second, nuclear or utterance-final pitch accents tend to show falling melodic contours. Again, the beginning and end of the falling gesture seem to be associated with the beginning and end of the host stressed syllable. Third, pitch accents in narrow focus are complex rising-falling gestures in which the peak of the rise occurs within the bounds of the stressed syllable. More options are possible, but these are the most common across the standard dialects of the Ibero-Romance languages that have been studied so far.

How Romance languages make use of these melodic configurations varies from language to language. In Central Catalan, nonfinal melodic gestures are rising movements with a delayed peak (as in many dialects of Spanish, for instance), while final pitch accents show falling contours (as in some Portuguese and Italian dialects, for instance). Estebas (2003) found that, in Central Catalan, delayed peaks are timed with respect to word edges. On the other hand, Prieto (2006), using different materials, demonstrated that there is no *strict* alignment of peaks with respect to word edges, even though word-edge effects on timing are strong. This topic, the role of word edges on pitch peak timing, will also be addressed in the present paper.

2. *Describing Majorcan Catalan declarative intonation: an experiment*

2.1 *Research questions and hypotheses*

In order to describe the basic intonational patterns of MC declaratives, we prepared a list of sentences that four participants (two males and two females) were asked to read as naturally as possible at a self-selected speech rate. Our intention was to analyze the shape of nonfinal pitch gestures, as well as their alignment patterns with respect to the segmental string.

Several research questions and hypotheses inform this paper. Broadly, the research questions are as follows: (i) What is the basic melodic implementation of nonfinal local pitch targets in MC declaratives? Are they similar to those in other Romance varieties? (ii) What are the alignment tendencies in MC? Assuming what has been described for other Catalan varieties, we hypothesize that (i) MC nonfinal pitch accents consist of a rising gesture with a delayed peak, and (ii), regarding alignment preferences, delayed peaks may optionally align with respect to right-hand word edges (Estebas 2003) or syllables (Prieto, 2006). *A priori* both are possible options, since both options have been proposed for other Iberian Romance varieties.

2.2 *Data collection, speakers, materials and methods*

Four participants, two males (1M, 2M) and two females (3F, 4F), read a total of 1344 sentences (112 tokens * 4 speakers * 3 repetitions). The four speakers are in their late twenties and they all have college degrees. Speakers 1M and 3F live in the rural centre of the island, where MC is the dominant language, while 2M lives in a touristic town on the island's Eastern coast, where Spanish and MC coexist, and speaker 4F lives in the capital city, Palma, where Spanish is more commonly used, probably due to the Spanish-monolingual immigrant population, who established there during the '60s (e.g. Pieras, 1999). While the four speakers have, to our native intuitions, typical native MC accents, their linguistic background is different: speakers 1M and 3F self-reportedly use MC 100% of the time and used to speak MC with both parents at home as children, speaker 2M used to speak Spanish at home as a child with one of his parents, and speaker 4F used to speak MC at home with both parents, but nowadays self-reportedly uses MC 50% of the time or less, the other language being Spanish, used with friends and at work in Palma. Subjects 2M and 4F claim to feel as comfortable with Spanish as with MC, while speakers 1M and 3F claim to feel relatively uncomfortable when speaking Spanish.

The test sentences contained balanced instances of utterances where the number of following unstressed syllables, i.e. following the target word, was manipulated from 0 to 4. In addition, in all these tonal-context situations, the position of the word boundary was moved as much as constraints on the lexicon would allow us. Thus, target words included oxytones, paroxytones and proparoxytones. A three-way comparison between oxytones, paroxytones and proparoxytones was only possible in tonal-context situations where 2, 3 or 4 unstressed syllables intervened between the target stressed syllable and the following one. These two factors were included to maximize alignment variability, since they both have been shown to affect timing (Prieto, *et al.* 1995; Estebas, 2003), so alignment tendencies can be investigated under high variability. That is, by including materials that are known to increase timing variation, potential findings of alignment stability under these conditions will be highly reliable, since they will be inferred from highly variable data. It should be noted, however, that this variation is not random and that it could be quantified. Nonetheless, the effects of these factors will not be investigated here. All target items were DPs, with a Noun + Adjective combination, in which the N was the target word, i.e. whose pitch gesture will be analyzed, and the A served as a way to control for the N's right-hand prosodic context and phrasing.

All the target words appeared within the sentence predicate, and, since our aim was to study nonfinal, prenuclear accent-lending pitch contours, they never appeared in utterance-final position. Prieto (2005) found that, in Central Catalan read speech, many noun phrases in subject position, as well as many verbs, are followed by a continuation rise or sustained high pitch –analyzed as H- boundary tone, using AM's notation. The gesture of this continuation rise may be blended with the rising gesture of the stressed syllable and thus result in a highly displaced peak. Analyzing this delayed peak as an instance of a regular nonfinal gesture could confound the facts, rather than help clarify them. Thus, in order to minimize this risk, our target pitch accents were placed only on Ns appearing within the predicate.

It was decided to add one more factor to the study. In the beginning, this factor was added as a way to add variability (in order to gain external validity); however, as it will be seen, this resulted in the finding of a significant conditioning factor. All of the sentences had a coordinated DP or a PP added to the right of the N+A combination, that is, at the end of the utterance. These 'larger' sentences were added to the original list of test utterances. This resulted in a group of sentences in which the N+A combination was at the end of the sentence and in a similar group of sentences in which the N+A combination was followed by another phrase. Recall that the target pitch gesture is that of the N, and thus it was never in phrase-final position, but in either penultimate or antepenultimate position. Two examples can be observed in (1).

(1) a. En Joan duia sa [***monéda*** *romána*] (Target: ***monéda***)
'John was carrying the Roman coin'
(Paroxytone followed by 2 unstressed syllables, penultimate position)
b. En Joan duia sa [***monéda*** *romána*] [*en es coll*] (Target: ***monéda***)
'John was wearing the Roman coin on his neck'
(Paroxytone followed by 2 unstressed syllables, antepenultimate position)

The subjects were asked not to 'stop' in the middle of phrases while reading. This was done in order to avoid artificial phrasing. The occurrence of phrase breaks was unavoidable, however: All subjects placed H- boundary tones after subjects in most cases. No other obvious breaks were inserted anywhere else, certainly, not between the target N+A combinations. We are aware, however, of the fact that this is a categorical, arbitrary claim based on simple impressionistic observations. Instrumental research on phrasing issues is needed. Besides this, the participants did not receive direct instruction regarding the how the contour should sound. They were left with the decision of repeating any sentences *they* felt they misread. Recordings of the materials were made through a Shure SM10A head-worn dynamic microphone into a digital solid-state recorder, Marantz PMD 660. The speech signal was digitized at 44.1 kHz and 16-bit quantization. The stimuli were presented to the participants through a slide presentation on a computer screen.

At this juncture, it may be important to note that the data and findings from this study on MC also may be compared to those reported in Simonet (2006) for Spanish, since identically-designed materials were used. Having equivalent data in Spanish will allow us to make straightforward comparison between Majorcan Catalan and Spanish, a language whose prosody has received much more attention, and thus for which more is known.

2.3 *Acoustic analysis*

The sound files were analyzed using the Praat signal-processing package (Boersma and Weenink, 2006). Segmental landmarks were manually labeled based on time-synchronized displays of sound waves and wide-band spectrograms. The labeled landmarks were: (i) Onset of the stressed syllable of the target word, (ii) offset of the stressed syllable, (iii) right-hand word edge.

Instrumental studies of intonation are usually carried out through the examination of pitch pivot points, i.e. comparable points in the curve, such as peaks and valleys (Nooteboom, 1997). Pitch pivot points are coordinates on a *pitch*time* two-dimensional plot. In our case, discrete points in the pitch contours were extracted using an automatic technique. Their *pitch*time* coordinates were stored together with the other three prosodic landmarks. First, a smoothing algorithm was applied to *F0* contours. Then, Praat scripts were written that would extract the pitch contour from 50 ms before the onset of the tonic to 50 ms after the word edge, in steps of 10 ms, and then calculate the first derivative of the curve, which is a representation of the velocity of the pitch curve (Xu and Liu, in press). Derivatives were used to automatically detect relevant discrete points in the pitch curve, such as absolute valleys, absolute peaks, and points of maximum rising and lowering velocity. The study of the timing of pitch turning points such as peaks and valleys is supported by findings such as those in Arvaniti *et al.* (1998), Ladd *et al.* (1999, 2000) and Xu (1998), according to which pitch turning points tend to be aligned with respect to syllable or segment boundaries. According to Xu (2005), Xu and Xu (2005) and Xu and Wang (2001), other discrete points in the contour may be as informative as pitch turning points, such as points of maximum rising or lowering velocity and maximum acceleration. The assertion is that, after a point of, say, maximum rising velocity, deceleration begins and, even though inertia may still provoke the continuation of the rise, asymptotic approximation to the next low target may already have been activated (Xu and Liu, *in press*). It was decided to include in our analyses the study of points of maximum velocity, which were studied merely in order to describe pitch contours in more detail: velocity points are approximately in the middle of pitch rises or falls.

A valley was defined as a zero-crossing point in a *rising* curve in the derivative, in which velocity is zero, while a peak was defined as a zero-crossing point in a *falling* curve, also in the derivative, in which velocity is also zero. A point of maximum rising velocity and one of maximum lowering velocity correspond to maxima and minima, respectively, in the derivative between two zero-crossing points. All sentences were re-checked and hand-corrections were applied when needed. A maximum of five points, defined as follows, were automatically located for a given gesture: *valley*^{1}, a valley before a rise; *risevel*, a point of maximum rising velocity; *peak*, a peak; *fallvel*, a point of maximum lowering velocity; and, *valley*^{2}, a valley after a fall. These labels are used in order to avoid reification, which the use of H and L as they are used in the AM model would promote, since we do not *a priori* know whether these pitch points correspond to actual tones or not. Reification is a common problem in recent intonation studies (Xu, 2006). For instance, it is not always obvious what H or L stand for: actual *F0* peaks and valleys or phonological targets.

Five pitch landmarks were located on clear rising-falling contours, while in rises (followed by a high plateau or a very smooth pitch lowering) and in falls (preceded by a sustained plateau), only three pitch landmarks were found: *valley*^{1}, *risevel* and *peak* for rises, and *peak*, *fallvel* and *valley*^{2} for falls.

3. *Results*

3.1 *Overall shape of pitch gestures: Rising and falling nonfinal accents*

While analyzing the sound files, we were surprised by the finding of falling contours on many of these nonfinal target words. Falling contours are typical of phrase-final, nuclear pitch accents in languages such as Italian, Portuguese, and also Central Catalan, but they do not typically appear in prenuclear position in Romance languages. Careful observation suggested that falling and rising gestures were in complementary distribution (at least clearly so for two speakers) as a function of utterance position. That is, our impression was that penultimate pitch gestures (***monéda*** *romána*) showed always rising contours while antepenultimate ones (***monéda*** *romána en es coll*) were more variable and many presented falling movements, with a steep descending trajectory during the tonic syllable. An example is displayed in Figure 2, which shows a contour with a fall during the stressed syllable of the word *monéda* ‘coin’.

Figure 3 displays average contours for each of the four participants as a function of utterance position: solid lines represent penultimate pitch gestures, while dashed lines stand for antepenultimate accents. The contours are generated by plotting average pitch (ERB), relative to the minimum pitch value in each sentence, with time values of pitch points, except for antepenultimate gestures of speakers 1M and 3F, for which only points *peak* to *valley*[2] are plotted (this is due to the fact that most falling contours for these two subjects were preceded by high plateaux, and thus points *valley*[1] and *risevel* were not extracted). In Figure 3, each column plots the same pitch points timed with reference to a different landmark: onset of tonic, offset and word edge.

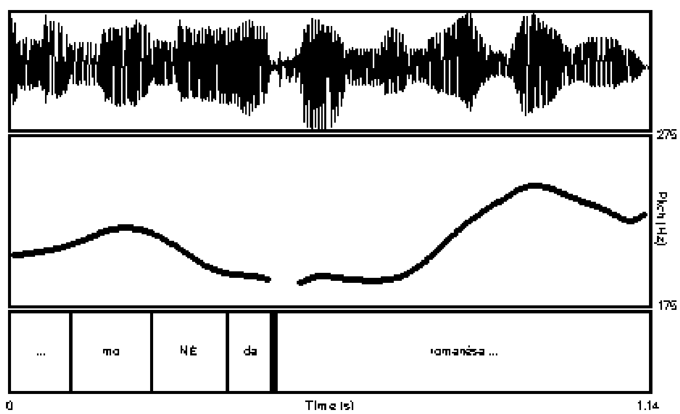

Figure 2. Sequence ***monéda*** *romanésa* 'Romanian coin' uttered by speaker 3F, in which *monéda* is in antepenultimate position. A falling gesture can be observed on *monéda*.

For subject 2M, penultimate and antepenultimate pitch gestures are very similar in shape: rising-falling. However, for the other three speakers, utterance position proves to be a relevant factor in that it triggers major differences in overall shape (direction) of the contour, as impressionistically observed. For 1M and 3F, pitch gestures in penultimate position were rising gestures, while those in antepenultimate position have clearly falling contours. Speaker 4F's melodic gestures in penultimate position were rising, while those in antepenultimate position were either falling (as for 1M and 3F) or rising (as for 2M). This variability in the shape of pitch contours of 4F is responsible for the average values displayed in the figure, which are neither clear falls nor clear rises: they are, in fact, averages of rising plus falling contours, and thus do not represent real contours. In other words, two patterns emerge: speakers 1M and 3F, who show a complementary distribution of falling and rising melodic gestures as a function of utterance position, and speaker 2M, who shows rising pitch movements in all positions. Speaker 4F is less consistent and stands in the middle, since she seems to freely select between falling and rising gestures in antepenultimate position, while always using rises in penultimate. This was attested by plotting all the actual *F0* contours on a graph –not shown here.

Regarding rising penultimate melodic gestures, a careful observation of the panels in Figure 3 suggests the following impressionistic alignment patterns: for all speakers, *valley*[1] is aligned with respect to the onset of tonic syllables; for the two male speakers, 1M and 2M, *peak* is aligned with the syllable's offset, while for the two female speakers, 3F and 4F, *peak* seems to be aligned towards the right-hand word edge. The maximum rising velocity point, *risevel*, tends to coincide with the middle of the tonic syllable, which agrees with the view that the maximum velocity may be reached at the point of maximum sonority in the syllable, where pitch changes are more clearly perceptible.

Regarding falling antepenultimate pitch contours (taking into consideration only those of speakers 1M and 3F, who present clear falls in this position), *peak* is aligned with respect to the onset of stressed syllables. For speaker 1M, *valley*[2] coincides with the offset of tonic syllables, while for speaker 3F it is aligned with the word edge. Points of maximum lowering velocity, *fallvel*, seem to occur at the middle of stressed syllables. In a sense, for speakers 1M and 3F, falling and rising gestures present the same alignment tendencies: the first pitch point (*valley*[1] for rising and *peak* for falling) occurs at the onset of tonic syllables; regarding the third pitch point (*peak* for rising and *valley*[2] for falling), speaker 1M aligns the two points with respect to the offset of stressed syllables while speaker 3F places them by the right-hand word edge. Timing patterns, however, need to be analyzed quantitatively, rather than by cursory examination of plots of average values.

3.2 *Alignment of pitch gestures: testing two reference points*

Linear and multiple regression models were applied to the dataset in order to investigate potential patterned correlations between pitch points and segmental/syllabic landmarks. In a number of studies, the location of pitch peaks relative to syllable onsets were plotted against syllable duration or *word-part* duration, which is the time lapse between the onset of the tonic syllable and the right-hand word edge (e.g. Estebas, 2003; Silverman and Pierrehumbert, 1990; Prieto *et al.*, 1995; Xu and Wang, 2001). As Xu and Wang (2001) note, in studies such as these, since both, say, peak location and syllable duration are measured relative to the same reference point (syllable onset), plotting the location of peaks as a function of syllable duration may reveal how peaks, or any other reference pitch points, align with respect to onsets and/or offsets of syllables. Models such as these, in addition to the *y*-intercept, provide a measure of the slope of the least-squares regression line. Slope coefficients can then be interpreted as follows: a slope of 1 means that peaks move fully in synchrony with offsets, while a slope of 0 means that they are synchronized with onsets, and a slope of 0.5 suggests that points are aligned with the syllable’s midpoint.

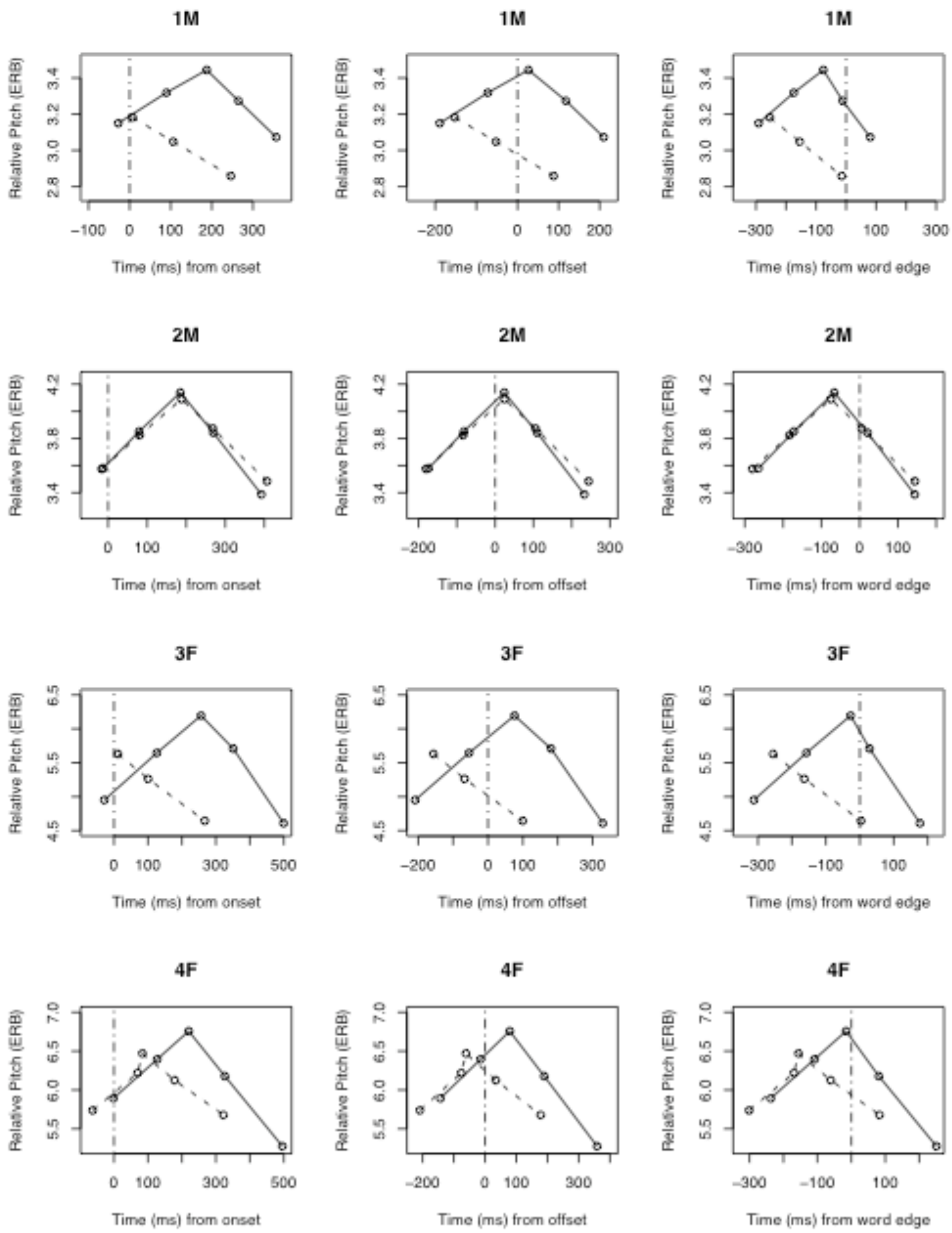

Figure 3. Typical patterns on *pitch*time* plots, conflated over all sentences, broken down by speaker and phrase positions: penultimate (solid line) and antepenultimate (dashed line).

Due to space limitations, we will report on the analyses of the timing of only two pitch points from each pitch gesture. For rising gestures, *risevel* and *peak* were examined, while for falls, the point of maximum *fallvel* and the *valley*^{2} were. Linear regression models were calculated in which the dependent variables, i.e. the time lapse between each of the two points and the onset of tonic syllables, were predicted with the use of two competing independent variables: (i) duration of the syllable or *syllable duration* [calculated as: offset – onset], and (ii) time lapse between the onset of the stressed syllable and the word boundary, or *word-part duration* [calculated as: word edge – onset].

For the study of **falling** pitch accents, a subset was extracted in which only sentences whose target word was in antepenultimate position were included. Only data from 1M and 3F were taken into consideration, since these two speakers showed a very consistent behavior in antepenultimate pitch movements. For 1M, syllable duration was not a significant predictor of the location of *fallvel* ($F(1,121) = 0.168$, *ns*; $R^2 = .001$), while word-part duration ($F(1,121) = 9.956$, $p<.01$; $R^2 = .07$) was. However, the amount of variance accounted for by word-part duration is still small. For 1M, word-part duration ($F(1,121) = 107.9$, $p<.001$; $R^2 = .47$) clearly outperformed syllable duration ($F(1,121) = 1.796$, *ns*; $R^2 = .01$) in significantly predicting the location of *valley*^{2}. Results for 3F were very similar, syllable duration was not a significant predictor of the placement of *fallvel* ($F(1,113) = 0.025$, *ns*; $R^2 = .0002$), while word-part duration ($F(1,113) = 11.16$, $p=.001$; $R^2 = .08$) was. Again, the amount of variance accounted for by word-part duration as the predictive variable is small. Regarding the timing of valleys (*valley*^{2}), word-part duration ($F(1,113) = 91.98$, $p<.001$; $R^2 = .44$) outperformed syllable duration ($F(1,113) = 0.769$, *ns*; $R^2 = .006$) in significantly predicting their placement. For both speakers, the model in which word-part duration was used as the predictive variable can explain 44% to 47% of the variance in the relative timing of *valley*^{2} with respect to syllable onsets; *fallvel*, on the other hand, remained largely unpredicted by both competing models. Slope coefficients of the successful models for both speakers were close to .70 (.71 for 1M, and .70 for 3F), which may suggest that valleys occur at 70% of the time latency between the onset and the word edge. Using Xu and Wang's (2001) paradigm, we could say that these valleys move more in synchrony with word edges than with syllable offsets or onsets, even though they are not strictly aligned with any of the three prosodic landmarks.

Finally, we calculated multiple regression models in which, in addition to the continuous variable predictors (syllable duration or word-part duration), we added two categorical factors: (i) stress configuration, and (ii) number of following unstressed syllables. Thus, two multiple regression models were compared: (a) syllable duration + stress configuration + number of following unstressed syllables; and (b) word-part duration + stress configuration + number of following unstressed syllables. These multiple regression models accounted for about 70% of the variance of *fallvel* (73% for 1M and 68% for

3F), and 93.5% (93% for 1M and 94% for 3F) of the variance of *peak*. Both models performed equally well. However, less weight was put on the categorical factors in model (b). This suggests that word-part duration is a better predictor overall, also observable in multiple regression models.

For **rising** penultimate pitch gestures, we extracted a subset in which only sentences whose target word was in penultimate position were included. Data from all four speakers were used. For rising accents, we analyzed only points *risevel* and *peak*. For 1M, both syllable duration ($F(1,103) = 18.32$, $p<.001$; $R^2 = .15$) and word-part duration ($F(1,103) = 50.23$, $p<.001$; $R^2 = .32$) are significant predictors of *risevel* location. The amount of variance accounted for by word-part duration, however, is larger. Still for 1M, both syllable duration ($F(1,105) = 13.79$, $p<.001$; $R^2 = .11$) and word-part duration ($F(1,105) = 33.87$, $p<.001$; $R^2 = .24$) give significant results in predicting *peak* placement. Again, word-part duration accounts for more of the dependent variable's variance. For 2M, both variable predictors were significant in accounting for the placement of *risevel*: syllable duration ($F(1,134) = 7.7$, $p<.01$; $R^2 = .05$) and word-part duration ($F(1,134) = 35.84$; $p<.001$; $R^2 = .21$). Once again, word-part duration outperforms syllable duration in the amount of variance accounted for. Regarding *peak*, however, only word-part duration was a significant predictor of its timing ($F(1,134) = 138.3$, $p<.001$; $R^2 = .50$), while syllable duration ($F(1,134) = 2.8$; *ns*; $R^2 = .02$) was not. Word-part duration accounts for a considerable amount of the variance of the dependent variable. For 3F, both syllable duration ($F(1,118) = 42.8$, $p<.001$; $R^2 = .26$) and word-part duration ($F(1,118) = 28.17$, $p<.001$; $R^2 = .19$) seem to be significant predictors of the timing of *risevel*. This time, however, syllable duration provides a higher R^2 value. Regarding the alignment of *peak*, results for 3F were identical to those for 2M: syllable duration was not a significant predictor ($F(1,119) = 2.62$, *ns*; $R^2 = .02$), while word-part duration was ($F(1,119) = 129.1$, $p<.001$; $R^2 = .54$). Again, the amount of variance accounted for by this predictor is considerable. Finally, for 4F, regarding the timing of *risevel*, word-part duration ($F(1,133) = 74.03$, $p<.001$; $R^2 = .35$) outperforms syllable duration ($F(1,133) = 23.41$, $p<.001$; $R^2 = .14$) as a predictor in that it accounts for a greater amount of the variance. With respect to the location of *peak*, again word-part duration ($F(1,135) = 113.2$, $p<.001$; $R^2 = .53$) is a much better predictor than syllable duration ($F(1,135) = 6.89$, $p<.01$; $R^2 = .04$). The R^2 value of the linear model in which word-part duration is the main predictor is considerable. In summary, in most cases word-part duration was a better predictor than syllable duration of the placement of both *risevel* and *peak* for the four speakers. There is only one exception: the timing of *risevel*, for speaker 3F, which seems to better correlate with syllable duration. In any case, what seems to be clear is that peaks move in synchrony with the duration of word-parts. The amount of variance accounted for by these simple linear models is considerable in most cases: 24%

in the case of 1M, and 50%, 54% and 53% in the case of speakers 2M, 3F and 4F, respectively.

Regarding the inspection of slope coefficients, which, assuming Xu and Wang's (2001) paradigm, provide information about the detailed alignment of pitch points, only the slopes of models in which word-part duration was the main predictor were taken into account, since they were, overall, more robust. In this respect, only peaks were examined investigated. Slope coefficients (*S*) were as follows: for 1M, *S*=.41; for 2M, *S*=.55; for 3F, *S*=.59; and for 4F, *S*=.53. Slope coefficients suggest that peaks occur at about 50% of the time latency between the onset of the stressed syllable and the word edge. Clearly, peaks do not go all the way down to word edges. Thus, slope values around .50 suggest that peaks are not aligned with respect to word edges or to onsets of stressed syllables, but with the midpoint of word parts, which we defined as the time lapse between the onset of the tonic and the right-hand word boundary. In other words, while word-part duration is a significant predictor or peak placement in rising pitch accents, there is no strict alignment of turning points with word boundaries or any other boundary.

Finally, we calculated multiple regression models in which two categorical factors (stress configuration and number of following unstressed syllables) were added to the two competing continuous variable predictors. This resulted in the comparison of two models as follows: (a) syllable duration + stress configuration + number of unstressed syllables; and (b) word-part duration + stress configuration + number of unstressed syllables. Results show that both models perform equally well, with R^2 values between .94 and .97. In all models, both stress configuration and number of following unstressed syllables have a significant predictive power, which suggests that there are important timing differences that are a result of these factors. Thus, we were not able to select one model as better than the other. Word-part duration seems to be a better predictor only in simple regression models and not when relevant categorical factors are included in multiple regression models.

4. *Discussion and conclusions*

The present descriptive study of the shape and alignment of nonfinal pitch gestures in declaratives in MC has arrived at the following two conclusions: (i) the position in the utterance of host words seems to determine the shape of pitch gestures in that penultimate contours have always a rising shape while antepenultimate ones tend to have a falling shape (consistently for two speakers and less consistently for one, while the fourth speaker shows no evidence for falling pitch accents), and (ii) pitch turning points such as peaks in rises and valleys in falls are timed with respect to the duration of word parts, which is the time stretch between the onset of the tonic and the word edge.

In Romance, falling accents are typical of final position. The question is thus, why do we find falling contours in nonfinal position in MC? The explanation may lie on phrasing tendencies. A possibility is that these nonfinal pitch gestures are in fact phrase-final, *nuclear* pitch accents, and not typical *prenuclear* gestures. If this were true, there would be no room to claim that there is a phonological difference in the inventory of prenuclear pitch accents between MC and other Ibero-Romance varieties. One could claim that after target words in antepenultimate position in our dataset there could be a L-boundary tone in MC (for speakers 1M and 3F), which might have triggered the preceding pitch accent to be switched into a nuclear falling accent. In order to examine whether these falling accents were in fact nuclear accents, we analyzed the duration of stressed syllables with falling accents, comparing it with that of syllables with rising accents. However, a paired t-test of syllable duration in antepenultimate (falling) *vs*. that in penultimate (rising) accents for speakers 1M and 3F revealed no significant difference in duration between the two. Syllables in nuclear position are expected to be longer than those in prenuclear position, a phenomenon known as final lengthening. We are aware of the fact that our data do not allow us to claim that we have found a new phonological category for prenuclear pitch accents in Ibero-Romance (falling accents), but assuming that falling accents *need* to be in fact nuclear pitch accents just because this is so in other varieties, even if in MC they occur in the middle of sentences and there is no evidence of final lengthening, is also a flawed assumption. More research is needed in order to show the nature of these falls: it may well be that they are in fact nuclear pitch accents, but only future research will tell. This is the reason why we have carefully used the terms ‘final’ and ‘nonfinal’ to refer to these accents, and also ‘penultimate’ and ‘antepenultimate’ to refer to their position.

Let us now focus on timing patterns in both nonfinal falls and rises. It has been shown that pitch turning points (valleys in falls, and peaks in rises) are not strictly aligned with syllable offsets or word edges. However, by taking into account the duration of word parts (as defined in this paper and also used in Silverman and Pierrehumbert, 1990; and Estebas, 2003), we were able to account for about 50% of the variance of the location of turning points. Multiple regression models that included factors such as stress configuration and number of following unstressed syllables were able to explain about 93% of the variance. Valleys in falling accents were found to move in synchrony with word edges and peaks in rising gestures move with word-part midpoints. Before making any claims about the *phonological* nature of the timing of turning points in MC, we should note that recent research has cast some doubt on the assumption that speakers can consciously select the position of pitch turning points and thus that points are actual phonological entities (Xu, 2005). Xu proposes that there might be tonal targets that extend throughout the whole syllable (or foot, mora, etc.) and that the apparent alignment between pitch points and syllabic boundaries may reflect the beginning and end of the *activation* of a target. Surface pitch contours would thus be a result of asymptotic approximation to these targets.

Whatever the theoretical interpretation, what is clear is that, in agreement with findings in Central Catalan (Estebas, 2003; Prieto, in 2006) and Spanish (e.g. Prieto *et al.*, 2005; Simonet, 2006), our data show that turning points move in synchrony with word edges. Since word-edge effects have been found for several Ibero-Romance varieties, including a conservative, isolated dialect such as MC (for both falling and rising accents), we would suggest that they might be a property of most Romance varieties. We would also propose that these effects might have been a property of proto-Romance prosody. This may indicate that, in Iberian Romance, pitch accents may be more appropriately seen as units applied to words rather than to stressed syllables.

5. *References*

Abulafia, D. 1994. "A Mediterranean Emporium: The Catalan Kingdom of Majorca". Cambridge: Cambridge U. Press.

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

Boersma, P. & D. Weenink. 2006. Praat: a system for doing phonetics by computer. [www.praat.org]

Estebas, E. 2003. "The modeling of prenuclear accents in Central Catalan declaratives". *Catalan Journal of Linguistics* 2: 97-114.

Hualde, J.I. 2004. "Romance intonation from a comparative and diachronic perspective". *Contemporary Approaches to Romance Linguistics*, ed. by Auger, J., J.C. Clements and B. Vance. Amsterdam & Philadelphia: John Benjamins. 217-237.

Ladd, D.R. 1996. "Intonational phonology". Cambridge: Cambridge U. Press.

-------- I. Mennen & A. Schepman. 2000. " Phonological conditioning of peak alignment of rising pitch accents in Dutch". *Journal of the Acoustical Society of America* 107, 2685-2696.

-------- D. Faulkner, H. Faulkner & A. Schepman. 1999. "Constant 'segmental anchoring' of F_0 movements under changes in speech rate". *Journal of the Acoustical Society of America* 106, 1543-1554.

Mascaró i Pons, I. 1985. "Ciutadella-Maó. Greu vs. agut en dos parlars menorquins: plantejament de la qüestió." *Randa* 21: 197-211.

Nooteboom, S. 1997. "The prosody of speech: melody and rythm." The Handbook of Phonetic Sciences, ed. by Laver, J. & W. Hardcastle. Oxford: Blackwell. 640-673.

-------- 1986. "Introducció a l'entonació dialectal catalana." *Randa* 22:5-38.

Payà, M. and M.M. Vanrell. 2005. "Yes/no questions and echo questions intonation in Majorcan and Minorcan Catalan." Poster presented at PaPI-2. Barcelona. June 20-21.

Pieras, F. 1999. “Social dynamics of language contact in Palma de Mallorca.” PhD Dissertation, The Pennsylvania State University.

Pierrehumbert, J. 1980. “The phonology and phonetics of English intonation”. PhD dissertation, Massachusetts Institute of Technology.

Prieto, P. 2006. “Word-edge tones in Catalan”. *Italian Journal of Linguistics*. 18(1): 39-71.

-------- 2005. “Syntactic and eurhythmic constraints on phrasing decisions in Catalan”. *Studia Linguistica* 59(2): 1-29.

-------- 2002. “Entonació”. *Gramàtica del català contemporani*, ed. by Solà, J., M.R. Lloret, J. Mascaró & M. Pérez Saldanya. Barcelona: Edicions 62, vol.I: 393-462.

-------- E. Estebas and M.M. Vanrell. 2005. "The role of tonal alignment and slope of the rise on word-boundary identification in Catalan and Spanish." Xth Conference on Laboratory Phonology. Lab. de Phonétique et Phonologie, CNRS/U. Paris 3. June 29-July 1, 2006.

-------- J. Van Santen & J. Hirschberg. 1995. “Tonal alignment patterns in Spanish.” *Journal of Phonetics* 23: 429-451.

Recasens, D. 1998. “Factors històrics d’algunes característiques fonètiques i fonològiques del mallorquí.” *Estudis de llengua i literatura en honor de Joan Veny*, Barcelona: PAM, vol. II: 541-558.

Silverman, K. & J. Pierrehumbert. 1990. “The timing of pre-nuclear high accents in English”. *Papers in Laboratory Phonology I*. ed. by Kingston, J. & M. Beckman. 1990. Cambridge: Cambridge U. Press: 72-106.

Simonet. M. 2006. “Word-boundary effects on pitch timing in Spanish”. *Selected Proceedings of the 9th Hispanic Linguistics Symposium*, ed. by Sagarra, N. & A.J. Toribio. Cascadilla Proceedings Project: 103-112.

Vanrell, M.M. 2006. “The phonological role of tonal scaling in Majorcan Catalan interrogatives.” MA thesis, Universitat Autònoma de Barcelona.

Xu, Y. 2006. “Principles of tone research.” *Proceedings of International Symposium on Tonal Aspects of Languages.* La Rochelle, France. 3-13.

-------- 2005. “Speech melody as articulatorily implemented communicative functions”. *Speech Communication* 46: 220-251.

-------- 1998. “Consistency of tone-syllable alignment across different syllable structures and speaking rates”. *Phonetica* 55, 179-203.

-------- & F. Liu. *in press*. “Tonal alignment, syllable structure and coarticulation: Toward an integrated model.” *Italian Journal of Linguistics*

-------- & C.X. Xu. 2005. “Phonetic realization of focus in English declarative intonation”. *Journal of Phonetics* 33: 159-197.

-------- & Q.E. Wang. 2001. “Pitch targets and their realization: Evidence from Mandarin Chinese”. *Speech Communication* 33: 319-337.