Tonal structure and constituency in Neapolitan Italian: Evidence for the Accentual Phrase in statements and questions

Caterina Petrone and Mariapaola D'Imperio

Aix-Marseille I & Laboratoire Parole et Langage, CNRS, France
caterina.petrone@lpl-aix.fr, dimperio@lpl-aix.fr

Abstract

In this paper, we report on some acoustic data suggesting the existence of the Accentual Phrase (AP) in Neapolitan Italian. In late narrow focus utterances in which intonation modality was varied, there was a difference in the shape and slope of the interaccentual region linking the H target of the prenuclear rise with the L target of the following nuclear rise. We hypothesized that such difference is due to the presence of a tone marking the end of an Accentual Phrase (AP), whose domain roughly correspond to the phonological phrase (φ). It is tentatively proposed that the phrasal tone is L for statements and HL for questions.

1. Introduction

Within the autosegmental-metrical (AM) theory of intonation [1], tunes are represented by a sequence of one or more pitch accents followed by a phrase accent and a boundary tone. Phrase accents and boundary tones occur at the end of the phrasal constituents, such as the intonation and the intermediate phrase. Nuclear accents are positionally defined as the last and more prominent accents in the intermediate phrase, immediately preceding a phrase accent.

In Standard Italian, the positional definition of the nuclear accent has been revised [2, 3], since this accent can be followed by postnuclear ones within the same intermediate phrase. The nuclear accent has been therefore defined as "the rightmost fully-fledged pitch accent in the focussed constituent" [2]. Moreover, the number and the definition of the phrasing levels in this language are still quite controversial [4]. In fact, though the intonation phrase is well attested in many Italian varieties, there is less evidence for the existence of the intermediate phrase (but see [3] for evidence in Florentine and [5] for Neapolitan Italian). Within a different and less recent approach [6] prosodic constituents smaller than the intermediate phrase have been claimed to exist in Italian, such as the phonological phrase (φ), which includes a lexical head and all its complements on the non-recursive side. More recently, [7] suggests that the end of non final φs in long declaratives is always marked by a L (B_{φ})[1] tonal event, different from the phrase accent of the intermediate phrase. Moreover, evidence for a tonal constituent smaller than the intermediate phrase, the Accentual Phrase (AP), whose domain roughly correspond to that of φ, has been found in languages such as Korean [8] and French [9].

In this paper, we report on data from Neapolitan Italian suggesting the existence of the AP. This constituent would be marked by an edge tone aligning with its right boundary and its break level would be weaker than the for the intermediate phrase. Specifically, we will focus on the acoustic properties of the F_0 region between the prenuclear and nuclear rising accent for both narrow focus questions and statements. We know that in Neapolitan Italian the nuclear accent is later in yes/no questions (L*+H) than in narrow focus statements (L+H*) [2]. Moreover, in long focus constituents, while statements show a constituent medial fall, analyzed as the leading tone of a L+H*, questions show the presence of a HL-phrase accent immediately following the focal L*+H [2, 10].

From informal observations, it appears that the F_0 contour within the region spanning from the prenuclear to the nuclear accent is differently realized in the two modalities. According to AM theory, if no intermediate tonal target is present, we should expect a linear interpolation between the peak of the prenuclear accent (L)H* [1,11] and the L tone of the nuclear rise. However, a different picture emerges, as shown in Fig. 1. In statements (upper), the F_0 rapidly falls from the prenuclear peak (Hp) to the region immediately after the end of the first φ (*La mamma*), with a low turning point followed by a low plateau which continues until the beginning of the nuclear rise. This pattern is similar to that found for broad focus statements [2], thus suggesting that the F_0 minimum following the prenuclear rise could be the surface realization of an underlying tonal target. On the other hand, in questions (lower), the F_0 fall after Hp is shallower, so that the F_0 contour in the immediate postaccentual region assumes a concave shape. After this region, the slope becomes steeper in order to attend the low values for the L* of the L*+H nuclear accent.

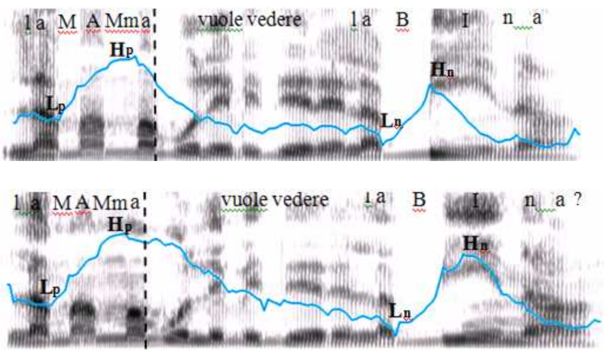

Figure 1: *F_0 contour of a narrow focus statement (upper) and a yes/no question (lower) for the utterance"La mamma vuole vedere la Bina" (The mother wants to see the Bina). Lp/n=prenuclear/nuclear L; Hp/n=prenuclear/nuclear H (speaker OM)*

Two hypotheses might account for the slope and shape differences within the interaccentual region. First, in statements a L tone would associate with an AP "La mamma", and realized through an L target at the right edge of the AP,

[1] More precisely, Nespor proposes each non final φ consists of a H* accent (A*=*alto* "high") and a L (B_{φ}= *basso* "low") tone.

while no such a tone would be present at the same structural position for questions. This hypothesis is disfavored since if no tonal specification would be present in this region for questions, the interpolation between Hp and L* (Ln) would be linear. According to an alternative hypothesis, a different tone would mark the right edge of the AP in both intonation modalities, i.e. a falling L_A for statements and a high H_A for questions. This difference would be in line with the regularities observed for long focus constituents [2, 10]. The latest hypothesis appears to be supported by our data.

2. Corpus and methods

A corpus of read speech was employed, originally recorded for other purposes. Eleven real words were embedded in the same carrier sentence, with intended narrow focus on the last NP:

$[\text{La mamma}]_{NPAP}$ $[\text{vuole vedere}]_{VP/\,AP}$ $[\text{la X}]_{NP/\,AP}$
"(the) mother wants to see the X"

All the target words are paroxytones proper names, in which stressed syllable onset type as well as vowel height were orthogonally varied (ex. *Dina* vs. *Nina*; *Dina* vs *Dana*). The sentences were read as either a yes/no question or a narrow focus statements, with a nuclear accent on X and a prenuclear (L)H* on *La mamma*. Each sentence was repeated 7 times by 4 Neapolitan speakers. Here results for two speakers (OM and AS) will be presented.

From a corpus of 308 sentences (2 intonation modalities X 11 words X 7 repetitions X 2 speakers), a total of 244 sentences were analyzed. The acoustic analysis was performed through Praat [12]. We manually marked the boundaries of the APs and those of their syllables and segments. The onset (Lp) and the peak (Hp) of the prenuclear rise as well as the onset (Ln) and the peak (Hn) of the nuclear rise were also marked (Fig.1). F_0 peaks were automatically detected as maxima within the host AP, whereas an automatic procedure described in [13] was used to detect the valleys (elbow) locations. We then measured: (1) the alignment and scaling of such targets relative to the different acoustic boundaries (the onset and offset of the prosodic word, and of its accented syllable and vowel); (2) the slope of the nuclear and prenuclear F_0 rises and of the region between Hp and Ln. We also modeled the shape of the interaccentual region through different regression techniques described below.

3. Results

Statistical analyses (including ANOVA and regression modeling) were performed on the fall of prenuclear accents, separately for OM and AS. Modality (questions vs. statements, **Q/S**) was the only predictor variable. The cutoff point for significance was 0.05.

3.1. Tonal alignment and F_0 scaling.

Before examining the interaccentual slope, we carefully analyzed the acoustic characteristics of the rise in prenuclear and nuclear accents. In fact, if the differences in interaccentual slope between questions and statements are genuine, we expect that changes in the slope gradient would be independent of timing or melodic adjustments of the two tones delimiting the interaccentual region, ie. the H target of the prenuclear (L)H* accent (Hp) and the L target of the LH nuclear rise (Ln).

As for tonal alignment, Hp was aligned around the end of the geminate [m:] of *mamma* in questions as well as in statements for speakers (Fig. 1). When alignment was measured relative to stressed syllable onset, an ANOVA showed **Q/S** effect for OM. On the other hand, for AS Hp was significantly later in questions (172 ms) than in statements (128 ms) [$F(1,135) = 200.19$, $p < 0.05$]. Ln was stably anchored to the accented syllable onset independent of intonation modalities for OM, though nuclear peaks were later in questions than in statements. For AS, the whole nuclear rise was significantly later in questions than in statements for AS (Ln: [$F(1,125) = 28.67$, $p < 0.05$]; Hn [$F(1,125) = 537.32$, $p < 0.05$]).

These results suggest that tonal alignment does not play any role in affecting the interaccentual slope for OM, since the alignment of Hp and Ln is stable across intonation modality. On the contrary, if an edge tone is present at the end of *La mamma*, its realization could be undershot in AS questions because of the late alignment of prenuclear peak (Hp), ie. the parabolic shape of the F_0 postaccentual region will be flatter for AS than for OM.

Mean values for F_0 scaling were similar for prenuclear peaks (Q=156 Hz; S=158 Hz) for speaker OM, whereas Hp was lower in questions (207 Hz) than in statements (227 Hz) for AS. In fact, an ANOVA showed a significant effect of **Q/S** on Hp only for AS [$F(1,135) = 264.17$, $p < 0.05$]. Hence, for OM, Hp did not show relevant melodic adjustments which might justify the difference in the steepness of the slope in the region immediately following Hp. However, we expect the F_0 excursion size from Hp to the edge tone will be smaller for spaerker AS.

Finally, we verified whether utterance duration might have affected interaccentual slope. Questions speech rate has been reported to be faster than that of statements in a variety of languages, including Neapolitan Italian [13]. Therefore, it is possible that if a low target is present at the end of the first AP in questions and if questions are uttered more rapidly, this target might be undershot. However, this hypothesis must be discarded since mean utterance duration (measured from the first amplitude peak to the last glottal pulse in the sentence) was longer in questions (OM: 1.507 s.; AS: 1.439 s.) than in statements (OM: 1.443 s.; AS: 1.358 s.). This difference was significant both for OM [$F(1,105) = 29.7$, $p<0.5$] and AS [$F(1, 135) = 51.1$, $p<0.5$].

3.2. The interaccentual slope

In works couched within the AM theory, F_0 slope is often calculated as the ratio between the F_0 excursion and the rise time (Hz/ms) since it is assumed that the interpolation between two tonal targets is linear. However, at a first approximation, linear interpolation rules are unsatisfactory to account for shape of the interaccentual slope in our data, especially for questions (Fig. 1). Hence, the dynamics of this section of the contour was inspected by extracting the F_0 values of 20 equidistant points from Hp to the middle of the interaccentual region (MIR).This choice was based on the observation that the F_0 shape difference between questions and statements was most evident in the region from the end of *La mamma* to the end of the following word, *vuole*. Moreover, by focussing on this specific region, we minimized possible effects of microprosody on F_0 modelling. In fact, if no tone target is present between Hp and Lp, we expect that this region can be modelled in terms of a straight line for both modalities (Fig.2).

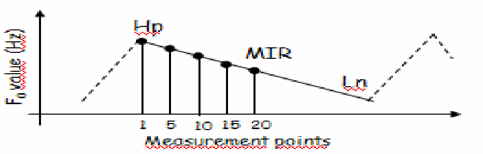

Figure 2: *Schematic representation of F_0 values extracted at 20 points of measurements.*

Time-normalized plots of the mean F_0 values (*y*-axis) at the 20 points (*x*-axis) are shown in Fig. 3 (OM) and Fig. 4 (AS). For each speaker, results are separately plotted for statements (left) and questions (right). Our exploration started by fitting a local weighted polynomial regression (*lowess*) to the Hp-to-MIR for both intonation modalities. The *lowess* regression is a robust statistical technique to smooth curves without prior assumption about their shape.

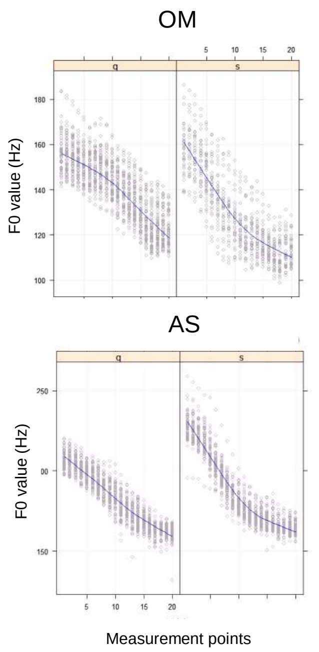

Figure 3 and 4: *F_0 values at 20 measurement points for OM (upper panels) and AS (lower panels). Results are shown for statements (left) and questions (right). The lowess smoother is indicated by the solid line.*

In both speakers the F_0 curve was differently realized in questions vs statements. In statements, the F_0 rapidly falls from the pitch peak to the F_0 region around points 10-15, where the slope abruptly changes and becomes shallower. In questions, the shape of the F_0 postaccentual region is realized in a slight different way between the two speakers. For OM, such a region appears as a concave downward parabola, whose inflection point is located around points 11-12. The parabola is flatter in AS questions, where tonal alignment and melodic variations might have affected the realization of the interaccentual slope. Here, in fact, the F_0 seems to decline in an almost linear way until points 9-11, where a small change in slope occurs. Note, though, that the slope of the F_0 region before the inflection point is still much shallower in questions than in statements.

Moreover, the inflection point of the parabola, has a lower F_0 value than that of the prenuclear peak for both OM (138 Hz) and AS (187 Hz) . Because of this large F_0 excursion (27 Hz for OM and 20 Hz for AS), we might therefore exclude the hypothesis that we are dealing with a plateau configuration of the H prenuclear target. Another hypothesis is that in both intonation modalities an AP-final tone is present, though being differently specified. According to this view, the concave shape of questions is due to an intervening high H_A tone, which is responsible for maintaining high F_0 values until point 11. After this point, the F_0 slope linearly decreases until the L* of the nuclear L*+H accent is reached. If this is true, we expect the interaccentual region in questions to be more accurately modeled through non-linear regression, ie. by a logistic regression, while a piecewise regression should yield more satisfactory results for statements.

To test this, we first fitted a logistic curve to the Hp-to-MIR region, separately for questions and statements. In line with our predictions, the logistic regression showed a good fit for questions (Fig. 5; Residuals Standard Error for OM: 8; for AS: 5.6), whereas the model did not converge for statements neither for OM nor AS. Note that, as expected, the inflection point of the curve was at point 13 for OM and at 9 for AS.

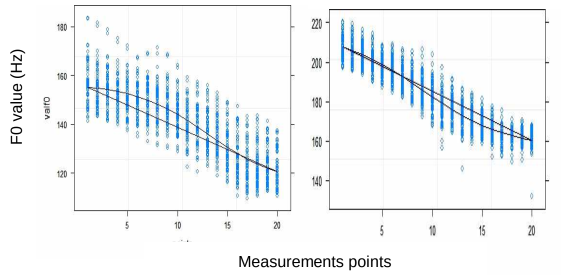

Figure 5: *Observed (dots) and expected (solid line) F_0 values for the logistic regression for OM (left) and AS (right).*

Moreover, in order to directly compare the two intonation modalities, a piecewise regression was fitted to the Hp-to-MIR region. For questions, the breakpoint was placed close to the the inflection point (at point 13 for OM and at 9 for AS). Since we needed to find what the most likely breakpoint is for statements, a series of models (one for each possible breakpoint) was fitted for this intonation modality. Therefore, we selected the model for which the deviance was smallest, i.e. the one in which the breakpoint was located at point 11 for both speakers.

Finally, we fitted another piecewise regression separately for the two speakers, in which **Q/S** was the predictor variable. The breakpoint was placed at 12 for OM and at 10 for AS, i.e. the mean distance between the breakpoints for questions and statements. For both speakers, slope was steeper in statements than in questions in the region before the breakpoint (mean values for OM: S=-3.7, Q=-1.63; for AS: S= -6, Q= -3.1), but the inverse was true after it (OM: S=-1.08, Q=-2.65; AS: S=-2.4, Q=-1.73). An ANOVA of the model confirmed a significant difference in slope due to intonation modality both before [OM: $F(2, 2134) = 2443.5$, $p< 0.5$; AS: $F(2, 2734) = 1206.6$, $p< 0.5$] and after [OM: $F(2,2134) = 135.2$, $p< 0.5$; AS: $F(2, 2734) = 496.7$, $p< 0.5$] the breakpoint. Moreover, for AS, questions were better modeled by a piecewise regression than by a simple linear one, thus supporting the existence of a true F_0 breakpoint even for this speaker [$F(2, 1418) = 43.7$, $p< 0.5$].

4. Discussion

In this study we showed that in late focus statements and questions of Neapolitan there appears to be a difference in the shape and slope of the interaccentual region linking the H target of the prenuclear rise with the L target of the nuclear rise. In particular, we were faced with a puzzling situation when the interaccentual slope in questions was considered because of the F_0 shape assumed by the region immediately following the prenuclear accent. Since interpolation between tonal targets is assumed to be linear within the AM approach [1], we asked whether the difference could be accounted for by some phonological/phonetic difference of the prenuclear accent for each modality or whether it would reflect differences in the tonal marking of a phrasal constituent smaller than the intermediate phrase. The first hypothesis was discarded for both speakers. For OM, we found no difference in the phonetic implementation of the prenuclear accent. However, a strong change in the curvature of the F_0 postaccentual region between questions and statements was still found, thus suggesting that the difference in interaccentual slope is genuine. The parabolic shape is less concave for AS than for OM. This might be due to target undershoot: since for AS prenuclear peaks are later and lower in questions than in statements, the transition between Hp and the following edge tone will also be shorter and flatter. In fact, even for this speaker a difference in F_0 slope is still noticeable before and after the inflection point of F_0 curve.

These results led us to examine whether a phrasal tone, without the percept of an intermediate phrase break [5], marks the region immediately following the prenuclear peak. We assumed that if no additional tonal target is present, the intervening F_0 contour could be modeled by a simple linear regression. However, we needed more sophisticated modeling techniques to account for the differences in the interaccentual region. Results from a logistic regression suggest the presence of an AP final H_A tone in questions, which would be responsible for the shallow slope in the region immediately following the prenuclear rise. Moreover, the failure of the logistic model in fitting the interaccentual region of statements, as well as results from a piecewise regression, suggest that an AP-final L_A tone characterizes statements and is responsible for the rapid fall from the H prenuclear peak to the end of the AP. This difference in tonal specification can also account for the change in slope value after the breakpoint: in questions, the slope becomes steeper because the F_0 must fall from the H_A target to the L* of the nuclear accent; in statements, the slope shallowly declines from the AP-final L_A to the leading L of the nuclear L+H*.

Internal and external evidence hence suggests that there is an additional tonal target, beyond the prenuclear and nuclear accents, in late narrow focus questions and statements reflecting the presence of an underlying edge tone marking the end of the AP/φ. First, in both questions and statements the inflection point of the interaccentual slope is located in the region immediately following the end of the first AP, *La mamma*. Moreover, this analysis is congruent with evidence from long focus constituents in which the F_0 does not fall after the prenuclear rise in questions, while it falls rapidly in statements [2]. Also, the existence of a prosodic constituent smaller than the intermediate phrase has been proposed for another Romance language, though typologically different, such as French.

Finally, we might ask why the AP-final tone should be differently specified for statements and questions. A possible hypothesis is that the use of a different tone would enhance the phonological contrast between modalities, for which no morphosyntactic differences are exploited and whose main cue is the different alignment of the nuclear accent.

5. Conclusion

In summary, in Neapolitan Italian there appears to be a difference in the slope and shape of the F_0 region between two consecutive LH (prenuclear and nuclear) accents, within the same intermediate phrase. Results from tonal alignment and scaling confirm that this difference is genuine, ie. not due to phonological/phonetics variation in the neighboring tones. Results from the combined use of non linear and linear modeling techniques suggest that this difference is due to the presence of a differently specified tone, L_A in statements vs. H_A in questions. This tone would mark the right edge of the AP, thus supporting the existence of constituents smaller than the intermediate phrase in Neapolitan Italian.

6. References

[1] Pierrehumbert, J. B. ,1980. *The Phonology and Phonetics of English intonation*. Ph. D. thesis, MIT.

[2] D'Imperio, M. Italian intonation: An overview and some questions,2002. *Probus* 14(1), 37-69

[3] Grice, M.; D'Imperio, M. Savino, M.; Avesani, C., 2005. Towards a strategy for labelling varieties of Italian. In *Prosodic Typology and Transcription: A Unified Approach*, S.-A. Jun, (ed.). Oxford: Oxford University Press, 55-83.

[4] D'Imperio, M; Gili Fivela, B., 2003 How many levels of phrasing ? Evidence from two varieties of Italian. In *Papers in Laboratory Phonology VI*, J. Local; R. Ogden; R. Temple (eds.). Cambridge: CUP, 130-144.

[5] D'Imperio, M.; Elordieta, G.; Frota, S.; Prieto, P.; Vigario, M., 2005. Intonational phrasing and constituent length in Romance. In Frota, S.; Vigàrio, M.; Freitas, M. J., *Prosodies*, The Hague: Mouton de Gruyter., 59-98.

[6] Nespor, M.; Vogel, I., 1986. *Prosodic Phonology*. Foris: Dordrecht.

[7] Nespor, M., 1993. *Fonologia*. Bologna: il Mulino.

[8] Jun, S-A, 1993. *The Phonetics and Phonology of Korean Prosody*. Ph.D. Thesis. The Ohio State University.

[9] Jun, S-A; Fougeron, C., 2002. The Realizations of the Accentual Phrase in French Intonation. *Probus* 14, 147-172.

[10] D'Imperio, M., 2003. Tonal structure and pitch targets in Italian focus constituents, *Catalan Journal of Linguistics*, 2, 55-65.

[11] Gili Fivela, B.; D'Imperio, M. Tonal alignment of prenuclear accents in Italian, 2004. TIE Conference, Santorini, Greece.

[12] Boersma, P., David, W., 2007. Praat (Version 4.6.38) [Computer program]. Retrieved November 2007, http://www.praat.org/.

[13] D'Imperio, M. *The Role of Perception in Defining Tonal Targets and their Alignment.* PhD Thesis, The Ohio State University