Italian intonation: an overview and some questions.

Mariapaola D’Imperio

Laboratoire Parole et Langage, CNRS – Université de Provence, FRANCE

Abstract

This paper presents a selective state of the art for the intonation of Standard and regional varieties of Italian, drawing especially from Neapolitan Italian data. Production and perception experimental data for this variety are employed to show some interesting interactions between focus, accent placement and accent type. The issues are presented within the autosegmental-metrical approach to intonational phonology. Points for future research are suggested.

1 Introduction

Quantitative as well as qualitative data about the intonation of Standard and regional varieties of Italian are still scarce, and a consensus labeling system has only recently been proposed (Grice, D’Imperio, Savino, and Avesani in press) mainly based on Central and Southern variety data. Nevertheless, some of the phonological and phonetic work produced within the last ten years, framed within the Autosegmental Metrical (AM) approach to intonational phonology (see Ladd (1996) for a review), appears to be very promising. Hence, the facts that I shall discuss here are mostly based on solid empiric inspection intended to test specific phonological models of intonation (Avesani 1990; D’Imperio 1995, 1997a, 1997b, 1999, 2000b; Grice 1995a; Grice, Benzmüller, Savino and Andreeva 1995).

Similar to Spanish (see Face, this issue, and Beckman et al., this issue), Italian is a stress accent language, in that metrical prominence is phonetically expressed through a combination of fundamental frequency ($F_0$), intensity, duration and possibly other spectral cues. The use of duration as a stress cue, both in production (Farinetani and Kori 1990; Marotta 1985) and perception (Bertinetto 1980; D’Imperio 2000a) can be compared to the role of the lax/tense vowel opposition which characterizes the first level of the stress hierarchy proposed for English (Beckman and Edwards 1994; Beckman 1996). Stressed syllables are generally penultimate (Lepschy and Lepschy 1977; D’Imperio and Rosenthal 1999), though antepenultimate, final and preantepenultimate stress can also be found.

In Italian, lexically stressed syllables can receive a pitch accent. The positional definition of nuclear accent offered within the AM approach has been the object of controversy in the description of Italian varieties (D’Imperio 2001). According to the standard AM definition, this is the last and most prominent accent of the intermediate phrase. Also, this is the accent immediately preceding the phrase accent (Pierrehumbert 1980). In Italian, instead, the nuclear accent can be defined as the “rightmost fully-fledged pitch accent in the focussed constituent” (Grice et al. in press). This alternative definition allows one to identify later accents within the same intermediate phrase, i.e. post-nuclear accents, which otherwise would not be allowed from a theoretical standpoint.

Different pitch accents cue different pragmatic functions. In broad focus statements the nuclear pitch accent is a fall, generally analyzed as a sequence of a high target followed by a low one on the stressed syllable ($H^*+L^*$) in all of the Italian varieties explored so far (cf. §2.1 for experimental evidence from Neapolitan supporting this analysis). The description of the question tune is instead highly dependent on the regional variety under scrutiny. While most Northern and Central varieties seem to be characterized by a terminal rise (i.e., a low nuclear accent followed by a rising phrase accent), Southern varieties exhibit a local rise on the nuclear accented syllable followed by a later fall. For instance, in the Neapolitan variety this tune has been analyzed as a combination of a $L^*+H$ followed by a HL phrase accent (D’Imperio 1997b, 2000b) (cf. §1.2.1).

The location of pitch accents is not fixed in an Italian prosodic phrase and the focus structure can be entirely signaled by nuclear accent placement. However, late accent placement can lead to ambiguity of focus
interpretation for question tunes. This issue is explored in §2.1. An alternative way of signaling focus is to alter the basic word order in the sentence, since the rheme (new information) of an unmarked declarative generally corresponds with the predicate position (Sobrero 1993). Therefore, in an utterance such as Maria ama Giovanni “Mary loves John”, the speaker can apply narrow focus on Giovanni (the Object) in two different ways. First, one can place a nuclear accent on the Object, as in Maria ama GIOVANNI (capital letters will henceforth indicate nuclear accent placement). Secondly, one can displace the same constituent to the left of the sentence (left topicalization), as in Giovanni, Maria ama. In this case, though, a narrow focus accent will have to be placed on Giovanni, which is somewhat redundant to word order.

An interesting interplay between focus and intonation structure has thus been revealed in some of the varieties of Italian that have been studied so far. First, broad focus nuclear accents of statements are structurally different from prenuclear accents belonging to the same utterance modality, in that they are bitonal (cf. §2.1). Second, it also appears that, as far as statements are concerned, narrow focus nuclear accents are different from broad focus ones, in that the first are rises and the seconds are falls. What is more, this difference appears to help native listeners of the Neapolitan variety to infer the intended focus structure of the utterance (cf. §2.2). In the following sections I shall first briefly review some of the previous approaches to the description of Standard Italian intonation (§1.1). I shall then review work on some regional varieties (§1.2) and, in particular, work conducted on Neapolitan Italian, by which I mean the Neapolitan pronunciation of Standard Italian¹ In the second part of the paper I shall concentrate on the issue of (nuclear) accent structure and its relationship with focus scope by reporting on experimental evidence from both the production and the perception domain (§2). I shall also raise some general theoretical and descriptive issues regarding the intonation system of Neapolitan and of other regional varieties and suggest points for future research.

1.1 Standard Italian intonation

The concept of a Standard Italian accent is a controversial issue (Lepschy and Lepschy 1977). Traditionally, the Tuscan pronunciation (including its intonation) has been referred to as the Standard. Here I shall attempt to draw a line between intonation descriptions that are intended to cover such a variety as opposed to those targeted to a specific regional variety. Before turning to more recent work incorporating the insights of intonational phonology, I shall briefly mention some work on Italian intonation inspired by the British School (O’Connor and Arnold 1961) and the Hallidayan holistic approach to tunes (Halliday 1976). These approaches are very similar in that they do not allow for a decomposition of the tune of an utterance into smaller linguistic elements.

Among the first attempts to describe Standard Italian intonation according to the British system, Chapallaz (1964) provides a classification of the “basic tunes” of questions. In Italian, questions are a good testing ground for the central role of intonation in shaping the pragmatic meaning of an utterance. This is because Italian lacks any morphological or syntactic means for marking yes/no questions, so that intonation by itself can carry the crucial function of distinguishing modality. For example, a sentence such as Maria ama Giovanni “Mary loves John” could be uttered as a question or a statement by exclusively manipulating its tune.

Among question tunes, Chapallaz distinguishes between a “basic pattern I” and a “basic pattern II”. These roughly correspond, respectively, to the “falling” and the “falling-rising” contours of the British tradition. From the notation that the author employs, one can deduce that the typical falling tune could be transcribed as a (H*) H+L* (L-)L% intonation phrase in autosegmental terms. This pattern is described as being typical of wh-questions.² Note that a H+L* accent is employed today in the transcription of the broad focus declarative pattern of most Italian varieties, and will be discussed later. On the other hand, basic pattern II, i.e. the falling-rising pattern, is claimed to be the basic tune of yes/no and alternative questions. In particular, the author claims that “the last stressed syllable may be on a low level pitch with a rise of pitch in the following unstressed syllables” (p. 307), which could be autosegmentally transcribed as L* H-H%.

¹I shall be concerned hence with the mere Neapolitan accent and not the Neapolitan dialetto, which differs from Standard Italian also morphologically and syntactically.

²Early acoustic investigations based on Northern varieties of Italian later suggested that wh-questions can also be characterized by a terminal rise, which would make them quite similar to yes/no questions (Magno Caldognetto, Ferrero, Lavagnoli, and Vagges 1978; Canepari 1986; Avesani 1990).
Lepschy (1978) does not limit his study to questions, and adopts Halliday’s framework of description by identifying 5 tunes for Standard Italian: a falling tune (typical of statements), a rising tune (typical of yes/no questions), a level tune (denoting incompleteness or uncertainty), a falling-rising tune (expressing doubt or surprise) and a rising-falling tune (employed as a “contradiction contour” or to correct a claim previously mentioned in the discourse).

Descriptive work on the intonation of different varieties of Italian carried out by Canepari (1986) is based on a system that is reminiscent of the British school. Tunes are divided in two sections, the protonia and the tonia (roughly corresponding to the head and nucleus of the British tradition). The tonia includes the melodic movement from the last accented syllable of the phrase up to and including any following movement on the postaccentual syllables. According to this system, the three essentials tunes of Italian are: a falling tune (expressing completeness, namely that the speaker is done with his/her turn), a rising tune (used for questions) and a suspensive tune (almost level on the nuclear syllable, used to express non-completeness). Using this system, not only does Canepari describe the basic tunes of Standard Italian, but he also manages to describe those of most regional varieties. However, his description of the yes/no question tune of some Southern varieties (such as the one spoken in Naples, Palermo and Bari) lacks a crucial insight. While he describes this tune as finally rising, recent analyses have shown that (unlike Standard Italian) it is in fact a terminally falling tune (Grice 1991; Caputo 1994; D’Imperio 1995).

In the late seventies, acoustic studies on Italian intonation began to appear. Though very interesting, such studies are often difficult to evaluate, the main problem being (as in many studies conducted on English intonation at the time) the attempt to directly relate acoustic measurements (generally \( F_0 \) maxima and minima) to syntactic position (such as Subject, Verb, Object), or pragmatic interpretation (such as contrastive focus, etc.) without the mediation of a well-defined prosodic/phonological level of analysis. Some of the first acoustic studies found that the \( F_0 \) contour of a Subject-Verb-Object statement is characterized by a peak on the Subject, a gradual lowering on the remainder of the utterance up to the Object position, and a substantial lowering on the last accented syllable, i.e. the accented syllable of the Object (Magno Caldognetto et al. 1978; Magno Caldognetto, Ferrero, Vagges and Cazzanello 1983). Yes/no questions are reported to show a final rise, starting from the last accented syllable and continuing on the following unstressed syllables (Magno Caldognetto et al. 1978; Kori and Farnetani 1983). This is, as mentioned above, the typical yes/no question contour described for Standard Italian, which Chapallaz (1964) had already identified. Among some of the first attempts to provide an acoustic characterization of focus, Magno Caldognetto and Fava (1972) and, later, Kori and Farnetani (1983), found that, in Standard Italian, focus is expressed by an \( F_0 \) peak. This finding has been confirmed by later research, such as Avesani (1990) and Avesani (1995), in which the pitch accent expressing narrow focus is analyzed as a peak accent (H*).

Pierrehumbert’s dissertation (Pierrehumbert 1980), and several other influential works on intonational phonology (Beckman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988), produced a shift in perspective in traditional intonation studies. However, it has taken roughly ten years since Pierrehumbert’s dissertation for the research on Italian intonation to be seriously influenced by the new current of ideas. The first attempt to employ Pierrehumbert’s system for describing Italian intonation belongs to Avesani (1990). In this study, the author proposes an inventory of pitch accents and boundary tones for Standard Italian, based on her experience in implementing a text-to-speech system. Avesani has later proposed a partial system based on ToBI, in which a number of interesting issues are raised (Avesani 1995). Specifically, she proposes a preliminary intonation system for Italian based on data derived essentially from one variety (Tuscan) and one speech style (that of professional speakers). The following set of pitch accents are claimed to form the inventory for Standard Italian: H*, L*, H+L*, L*+H and L+H*. According to Avesani, H* is the accent employed to signal narrow focus, since the lexical item carrying it is marked by the speaker as “new information”. This peak accent is also employed in order to transcribe the “default” prenuclear accent in a number of varieties (Grice et al. in press), including Neapolitan. On the other hand, Avesani proposes the use of a H+L* pitch accent to label the nuclear pitch accent of broad focus statements. This is also common usage in other varieties, such as Palermo Italian (Grice 1995a), Neapolitan Italian (Caputo 1994; D’Imperio 1995), Bari Italian (Grice and

---

3 A similar distinction between tonia and protonia has been employed for the description of Spanish intonation by Navarro Tomás (cf. Beckman et al., this issue).

4 Other Central and Northern varieties have also been reported as having terminally falling question tunes (Grice et al. in press).

5 Though most of these studies were conducted on Northern varieties, I include them within the Standard Italian work because they were not intended to be explorations of a specific regional variety as opposed to the Standard.
Savino 1995, 1997), as well as accents are postulated for broad and narrow focus, at least for statements. The nature of the H+L* falling accent of Neapolitan Italian statements is explored in §2.2.

1.2 Intonation in regional varieties of Italian

The description of regional varieties of Italian has recently benefited from close inspection, especially in relation to aspects of intonational phonology which put the descriptive power of different models to the test. Grice (1995a), for instance, has tackled the problem of how to describe the rise-fall accentual pattern of Palermo Italian yes/no questions by directly comparing the traditional British intonational framework against Pierre-humbert’s framework. Experimental data from the Neapolitan variety bearing on the alignment and scaling of tones and segments in the utterances have also been exploited as a means of describing tonal structure.

1.2.1 Question tune

As mentioned above, Standard Italian yes/no question tunes are terminally rising. Specifically, they appear to be cued by a combination of a L* nuclear accent followed by a rising H- phrase accent (plus a H% boundary tone) or, alternatively by a H+L* followed by a L-H% combination (Avesani 1995). In both cases, the rise, held to be responsible for the question meaning, is a property of the terminal part of the contour, as in English yes/no questions. In contrast to this, Southern varieties (at least those that have received attention in the literature so far) encapsulate the question meaning within a specific pitch accent type, which is generally also the nuclear accent of the phrase. This is true of some Sicilian varieties, such as Palermo and Catania Italian (Grice 1991, 1995a), Bari Italian (Grice and Savino 1995) and Neapolitan Italian (Caputo 1994; D’Imperio 1997b). This accent is a rising LH tone, where either the L or the H is starred depending on the variety under consideration.

In recent work, Grice and her colleagues have concentrated their investigation on two varieties, namely Palermo and Bari Italian. In doing so, they have extensively relied upon the HCRC Map Task procedure for data gathering. In Bari Italian, the nuclear pitch-accent of yes/no questions is described as a L+H* “which involves a low pitch target just before a high accented syllable” (Grice and Savino 1997, p. 30). The star notation here must be just taken as a convenient way to mark that the H is aligned with the stressed syllable, since no contrast has been found with another LH rise accent in this variety.

\[ \text{Discussion about the necessity for postulating a phrase accent in Italian is offered in Grice et al. (in press).} \]

\[ \text{Map tasks consist of task-oriented dialogues obtained from speakers involved in a “conversational game”.} \]
Similar to Palermo Italian, Neapolitan appears to employ a L*+H accent for yes/no question tunes. The shape of this L*+H accent is quite similar to that of narrow focus statements. A rise-fall pattern, with a very salient and discernible peak, characterizes the section of the contour from the accented syllable up to the end, as can be seen in Figure 1. Though both nuclear accents have been analyzed as LH rises, the temporal alignment of the L and H targets (and even of the final L) is different and perceptually relevant to the purpose of signaling question vs. statement modality. Specifically, the entire rise-fall appears to be timed later (relative to the stressed vowel) in questions than in statements (D’Imperio 2000b). The Neapolitan L*+H appears to be similar in shape to the “scooped” L*+H accent of American English, despite obvious timing differences for the L and the H targets. While in Neapolitan the L tends to occur at or just before the onset of the stressed syllable, with the rise to the peak being entirely contained within the boundaries of the stressed syllable, the H target of the English L*+H is generally reached much later, usually in the postaccentual syllable.

Note that the analysis of the statement narrow focus pitch accent as a rising L+H*, proposed in D’Imperio (1999), was instrumental in deciding the starredness status of the L in the L*+H of questions. In fact, from alignment facts alone, it was not clear whether the L, the H or both tones within the rise are associated to the stressed syllable, since both the L and the H target are realized within the syllable boundaries. However, since evidence for another LH rising accent for narrow focus statements has been found, the use of a contrastive notation is justified. Additional details on peak alignment difference in Southern varieties are offered in Grice et al. (in press). Future experimental studies of tonal alignment in the different varieties of Italian are very desirable as a tool for discovering contrasting tonal categories.

1.2.2 Truncation, compression and tonal repulsion

Recent progress in the description of the intonation system of Romance languages has benefited from the target-interpolation assumption of the AM framework. Within this approach, tonal targets are specified both in the time (alignment) and in the fundamental frequency (scaling) domains, while the intervening contour is the result of (mostly) linear interpolation. These targets represent, more or less directly, the tones postulated by the phonological analysis. Regarding the temporal domain, recent work (Silverman and Pierrehumbert 1990; Arvaniti and Ladd 1995; Prieto, van Santen, and Hirschberg 1995) has shown that it is possible to quantitatively model the alignment of certain F0 peaks and valleys with the stressed syllables. These studies concentrated on the interaction of a number of diverse factors (such as duration of the stressed syllable, distance to the next stressed syllable, distance to the end of the word, etc.) in determining systematic alignment differences.

Among tonal alignment works, Silverman and Pierrehumbert (1990) propose possible explanations for the variable location of tonal targets within the same pitch accent category of English. This problem is also currently faced by researchers who are engaged in the description of Italian varieties, especially of those lacking an explicit phonological analysis. According to Silverman and Pierrehumbert, underlying tonal targets (those postulated by the phonological analysis) can undergo some readjustments in the process of their phonetic implementation. As a consequence, tonal targets can be aligned earlier/later than predicted, and/or they can be scaled higher/lower in the F0 domain.

In order to explain such phenomena, the authors tested a number of hypotheses. Among them we find those of tonal undershoot and tonal repulsion. Briefly, both hypotheses stem from the idea, already very common in the segmental literature (Lindblom 1963), that two sequential articulatory gestures can interfere with each other when they are very close in time (gestural overlap). According to the tonal undershoot hypothesis, if a tonal target is produced as a result of two competing tonal gestures (such as a rise and an immediately contiguous fall), such a target can be undershot, i.e., its resulting F0 value can be lower than otherwise expected. Another possible outcome of gestural overlap is that the target peak (or valley) might be displaced earlier in time. This hypothesis is known as tonal repulsion.

Tonal repulsion and undershoot can be invoked to account for the realization of final low targets in Southern Italian rise-fall question tunes. Some examples are here drawn from Neapolitan. Compare the final rise-fall of the question Vedrai il NANO dopo? (where nano is initially stressed) “Will you see the DWARF afterwards?” with that of Vedrai MA? “Will you see MOM?”, respectively in the upper and lower panel of Figure 2. In the first case, three syllables follow the stressed NA- before reaching the end of the intonation phrase, hence the rise-fall can be fully realized. On the other hand, when less or no segmental material follows the stressed syllable, two outcomes are possible. If the rise-fall is placed on a monosyllable (such as the word fal “bonfire”), the fall can be curtailed.
This is shown in the lower panel of Figure 2. Such “truncation” is also found in Palermo and Bari Italian (Grice et al. in press). Hence, when it comes to choosing a strategy for temporal reorganization due to tonal crowding, some Southern varieties of Italian appear to be “truncating” rather than “compressing” (according to the typological classification proposed by Grønnum (1991)). If, on the other hand the accented syllable is followed by segmental material, the outcome of tonal crowding would be a “compression” of the tonal gesture, and hence a temporal repulsion of its target(s). For instance, it is possible to compress the entire rise-fall movement, and temporally anticipate the location of the peak even on a word such as nano “dwarf” when it occurs phrase finally. This is shown in the lower panel of Figure 3. Here the utterance Vedrai il NANO? “Will you see the DWARF?” is compared to the utterance Vedrai il NANO dopo? (which is repeated for ease of comparison). Note that when nano is in absolute phrase final position (lower panel), the accent peak is realized earlier relative to the offset of the stressed vowel. The compression of the rise-fall sequence can be seen as an outcome of tonal repulsion.

The degree of tonal repulsion does not seem to directly depend on the number of postaccentual syllables, though this result is only suggestive (cf. D’Imperio (2001)). Tonal repulsion has also been employed as a means to test the structure of broad focus accents in Neapolitan. This is the topic of the experiment presented in §2.1. Therefore, the situation appears quite complex in Neapolitan Italian, since both truncation and compression can be found, with truncation mainly found in intonation phrase final monosyllables or finally accented words, while compression (hence tonal repulsion) is mainly found in intonation phrase final non-mono-syllabic words.

Evidence for tonal readjustments comes also from the analysis of the focus constituent final fall of Neapolitan yes/no questions. When the constituent in focus is made of a single word, the fall of the rise-fall pattern occurs immediately after the pitch accent rise and seems to mark the end of the focus constituent. When the focus constituent is longer, the rise and fall appear to separate, with the rise staying anchored to the focal initial stressed syllable while the fall occurs later, reaching its target in the vicinity of the constituent right boundary. Figure 4 shows such a contrast by using different words. The utterances depicted are Vedrai [mamma] domani? “Will you see [mom] tomorrow?” with narrow focus on the word mamma (left) and the utterance Vedrai [la bella mano di Mammola] domani? “Will you see [Mammola’s beautiful hand] tomorrow?” (right), with narrow focus on the longer NP constituent la bella mano di Mammola. Note how, in the right panel, the pitch stays high after the word mano. On the other hand, as shown in the left panel, the pitch falls abruptly after mamma.

Figure 2: Word labels, $F_0$ curve and waveform for the utterances Vedrai il NANO dopo? (upper) and Vedrai MA? (lower). Stressed vowel offsets are lined up at the dashed line.
A production study (D’Imperio 1997b, 2001) concentrated on the properties of the final constituent fall (as well as the initial rise) in early focus questions with different focus constituent size. It was found that the final fall (which I analyze as a HL phrase accent) is anchored to the last stressed syllable of multi-word focus constituents when this is available, thus resembling a falling pitch accent8. However, when there is only one stressed syllable in the focus constituent, the existence of two separate H targets is obscured, since there are not two separate docking sites for them. In this case, only a low-high-low (LHL) sequence is found (such as in the left panel of Figure 4), due to the merging of the H targets of the L*+H and the HL- phrasal fall. Specifically, the L target for the phrasal fall is reached later in single-word focus constituents, as if it were displaced by the preceding L*+H rise. It was also found that the initial rise reaches its H target later in multi-word than in single-word focus constituents. This has been interpreted as evidence for tonal repulsion, which would cause the temporal anticipation of the H target when this is immediately followed by the HL- fall (i.e., in single-word constituents).

---

8In this case, the notation employed is H(*)L-.
Figure 5: $F_0$ curve and waveform for the broad focus statement *Mamma andava a ballare da Lalla* “Mom used to go dancing at Lalla’s”.

2 Focus and intonation: some experimental evidence from Neapolitan Italian

2.1 Broad focus statements: evidence for an accent fall

A first proposal for adapting the ToBI framework to describe Neapolitan Italian intonation can be found in Caputo and D’Imperio (1995). In this work, it is proposed, among other things, that a H+L* should be used to describe the nuclear accent of statements, where the L tone is scaled at the speakers’ baseline (the bottom of the speaker’s range). A H+L* has also been employed in the description of broad focus statements in other Romance languages, such as European Portuguese (Frota 1997, 2000, this issue), as well as of some English varieties (see the discussion in Grice (1995b)). The legitimacy of a HL fall analysis for the nuclear accent of Neapolitan statements is questioned in this section.

Several earlier investigations (D’Imperio 1995, 1997a, 1997b, 2000b, 2001) revealed an interesting interaction between focus scope and modality for utterances with late nuclear accent. In contrast to the narrow focus nuclear accent, the nuclear H+L* accent of broad focus statements appears to be acoustically “downstepped” (for a discussion of downstep, see §3.1) and less prominent than the prenuclear H* accent preceding it. Moreover, the H target of the H+L* is very difficult to discern given the lack of a clear peak. This is the case shown in Figure 5, presenting the $F_0$ contour for a broad focus statement (*Mamma andava a ballare da Lalla* “Mom used to go dancing at Lalla’s”) uttered by a Neapolitan speaker. Italian speakers hear a clear falling accent on the stressed syllable *Lall-* of *Lalla* at the location marked by H+L* and the arrow. Note that the $F_0$ forms a mid-high plateau, whose starting point (which is quite difficult to pin down) appears to be around *da*. This plateau ends at the location where the $F_0$ curve starts to rapidly fall.

An analysis that treats the prenuclear and nuclear accents of the same utterance type as being structurally different (i.e., a monotonal H* for the prenuclear accent vs. a bitonal H+L* for the nuclear one) in the absence of a meaning difference is problematic for models assigning a given semantic interpretation to each pitch accent (and phrase tone) (Hirschberg and Pierrehumbert 1986). In other words, the prenuclear H* and the nuclear H+L* could not belong to the same “natural class” according to these models. An alternative analysis of those differences models the shape of both accents as H*. Crucially, this analysis accounts for the earlier timing of the H target in the nuclear H+L* as a consequence of tonal repulsion due to an upcoming L tone associated with the remainder of the utterance (a L- phrase accent). A similar account has been proposed for the nuclear H* of English. The data presented by Silverman and Pierrehumbert (1990) seem in fact to support a parallel phonological and phonetic treatment of nuclear and prenuclear accents for English, shedding light on the contextual factors that can affect both.

Regarding the further observation that nuclear accent peaks are still earlier than prenuclear, they offer two possible explanations: 1) nuclear syllables always show a greater lengthening, therefore an earlier peak; 2) in nuclear position, H* tends to be somehow repelled by the upcoming L- phrase accent. This analysis
predicts that precise timing within the lexically stressed syllable in the nuclear accent will depend on such things as the number of following unstressed syllables before the end of the intonational phrase and also that the same alignment constraints will be exhibited by the different pitch accent categories used in statements and questions. Hence, I tested the hypothesis that the broad focus statement contour of Neapolitan is the result of two separate tonal events, i.e., a $H^*$ nuclear accent followed by a $L^-$ phrase accent (D’Imperio 1995; D’Imperio 1996). The hypothesis was tested by measuring the temporal latency of the target peak, relative to the onset of the stressed syllable, within the nuclear $H^*L^*$ accent of broad focus statements. Measurements were carried out for various prosodic and intonational contexts. It was hypothesized that, if the $H$ peak is retracted when the stressed syllable is closer to the phrase boundary (while the $L$ target remains temporally stable), this would be supporting evidence for a $H^*L^-$ analysis. On the other hand, if no such tonal repulsion is found, an alternative analysis can be sought, such that the fall is the result of a unitary accent fall. Moreover, if we find a correlation between the alignment of the $H$ and the $L$ target for the fall (i.e., if they are displaced in the same direction for a given context) this would be compelling evidence for a $H^*L^*$ fall analysis.

2.1.1 Materials and experimental procedure

The corpus consisted of a set of ten words which differed in the number of postaccentual syllables and stress position within the word (either initial or non-initial stress). Basic templates and target words are shown in Table 1.

The materials were recorded by three native speakers of the Neapolitan variety of Italian, one male (M1) and two females (F1 and F2), in a quiet room at the CIRASS laboratories of the University “Federico II” of Naples (Italy). The subjects were all brought up in Naples and were naïve as to the purposes of the experiment. Each speaker produced, in randomized order, five repetitions of a carrier sentence containing the target words, which were all phrase final (nuclear accent position). The same utterances were produced either as broad focus...
Figure 7: Left: stressed vowel duration and peak latency values for all speakers (target words with different syllable structure and stress pattern are plotted separately); Right: mean peak latency and low latency values plotted separately for initially and non-initially stressed words (standard error bars are shown; the vertical line indicates the onset of the stressed syllable).

A clear location for the H target was difficult to locate consistently by hand, given the fact that the statement contours generally present a high plateau (as mentioned above) before the fall. Also the exact location of the following L target was often hard to determine. In order to overcome this problem, an automatic procedure was employed. In order to determine the location of the L target, two straight lines were fitted to the segment going from the beginning of the accent fall (begin) to the location marked as “end” (see Figure 6). The x-intersection of these lines was made to correspond to the L target (elbow).

Once the “elbow” measure was determined, the location for the H target was calculated in a similar way. Namely, two straight lines were fitted to the segment starting from the plateau (beginH) up to the end of the fall (elbowH), whose x-intersection was made to correspond to the H target (elbowH). An example of straight line fitting for the last case is given in Figure 6. Note that the H target is measured at the elbowH location, as determined by the automatic procedure. Hence, in this work the H target will correspond with the location at which the contour starts to fall (elbowH), while the L target location will correspond with the “elbow” formed by the contour at the end of the fall.

2.1.2 Results and discussion

Figure 7 shows H peak latency (y axis) relative to stressed vowel duration (x axis). Target words are here grouped in terms of number of syllables to the right of the stress and intra-word stress location. Each group is

---

9 A program originally written by Mary Beckman and slightly modified by the author was employed in order to render the measurement more objective. A similar procedure was employed already in Pierrehumbert and Beckman (1988). See D’Imperio (2000b) for more detail.

10 The parameters of the two linear models were estimated by means of conventional linear least-squares methods. To estimate the elbow position, i.e., the intersection of the two fitted lines, two linear regressions were computed for each possible elbow location (from 1 to n, where n is the number of samples within the F0 segment). The location eventually selected as the “elbow” was the one leading to the smallest total modeling error. The elbow was then automatically inserted in the “label file” of a specific utterance, at the location on the F0 curve corresponding to the x-intersection of the fitted lines.
Figure 8: Scatter plot of low latency versus peak latency measurements. The straight line was fitted to the data by means of a conventional least-squares method.

shown by a different symbol. While English data show a positive correlation between the two measures (Steele 1986; Silverman and Pierrehumbert 1990), the results presented here do not show the same contextually-governed variation. In fact, there was neither a significant effect of number of postaccentual syllables nor of vowel duration on peak latency. In other words, there was no relationship between vicinity to the intonational phrase boundary and this measure. If we analyze the fall onto the stressed syllable as a sequence of a H* accent and a L- phrase accent, we would expect a tonal repulsion effect affecting the H target. This effect would be proportional to the vicinity of the H target to the end of the utterance. However, such an effect was not found. Note in the left panel of Figure 7 a single large cloud of overlapping values in the latency dimension (y axis), despite the fairly big range of durational vowels for the stressed vowel (x axis) and the different word sizes. For instance, note the small detached cloud of values for the target word showing the longest stressed vowel, i.e., maLAría. In contrast to our expectations, this word pattern did not show the greatest peak latency values. This can be taken as further evidence for the absence of an effect of vowel duration on peak latency.

On the other hand, intra-word stress location seemed to influence peak placement in Neapolitan Italian. In contrast to English, H (and L) latency appeared to be determined by the number of unstressed syllables preceding the nuclear stress, rather than the ones following it. The effect is shown in the right panel of Figure 7. Here, mean H peak (i.e., elbowH) latency (x axis) is plotted against mean L (i.e., elbow) latency (y axis) separately for words with initial (filled square) and non-initial (open circle) stress. Note that in words with initial stress the H peak is realized soon after the onset of the stressed vowel (marked by the dotted vertical line). In words with non-initial stress (such as maNDIamoglielo) the peak is realized much earlier, before the onset of the stressed vowel (generally within the preceding syllable). Interestingly, L latency appears to be strictly correlated with H peak latency, since L targets were earlier in words with non-initial stress. This pattern, which is already visible in the mean results shown in the left panel of Figure 7, was later confirmed by a linear regression model fitted to the data (r = 0.7). The results of the correlation are shown in Figure 8. Here, the correlation value was calculated for a subset of the data which did not include outliers. The results further support an analysis of the fall as a unitary accent gesture, which can be transcribed as a H+L*. To summarize then, the vicinity of the right edge of the intonational phrase as a whole does not seem to influence the alignment of accent peaks in Neapolitan broad focus statements. Therefore, the early placement of the accent peak within the nuclear stressed vowel cannot be explained in terms of a tonal repulsion effect

11Note that the only word group showing a seeming correlation between stressed vowel duration and peak latency is a subset of the non-initially stressed words indicated by the diamond symbol. This pattern was mostly caused by a segmental effect (the initial high vowel [i] in firmiamolo caused the F0 curve to start falling earlier from the higher values in the [i]).
12The correlation value was quite high even when it was calculated on the raw data (r = 0.6).
13A similar regularity was found for question tunes (D’Imperio 1995; D’Imperio 1996), whose pitch accent is a LH rise starting later in initially stressed words than in non-initial ones. What is more, both the accent fall of the statements and the question LH rise exhibited a constant length in the different prosodic contexts.
caused by the presence of a final L phrase accent. In fact, the alignment of the fall in words with either zero, one, two or three postnuclear syllables was equally affected by the accent being final to the phrase. Also, the total duration of the stressed vowel, to which the pitch accent is associated, did not significantly influence peak latency. The results seem to support the hypothesis of an invariant location of the High (H) and Low (L) turning points in a specific prosodic context, and therefore support a bitonal HL structural analysis of the broad focus pitch accent.

A surprising effect was exhibited by the Neapolitan data relative to the lexical stress location in the target word (regardless of the number of unstressed postnuclear syllables). A systematic variation in peak latency appeared to be triggered by the vicinity to the left edge of the word. What is suggested here is that there is a timing difference concerning the realization of the HL fall between Neapolitan Italian words with initial stress and those with non-initial stress, independent of the semantic/pragmatic context. Hence, I propose that the left-hand prosodic context (i.e. vicinity of word beginning) plays an important role in Neapolitan Italian intonation, in contrast to languages like English where the right-hand prosodic context (i.e. vicinity of intonational phrase end) has shown to be relevant in alignment phenomena. Moreover, different pitch accent categories (statement and yes/no question) show similar alignment effects.

A question still remains regarding the origin of the “downstepped” level of the H tone postulated for the H+L* accent. This discussion is deferred to §3.1. Another question regards the origin of the fall from the prenuclear H* accent to the plateau level preceding the nuclear H+L* fall. This could be mere interpolation from the H* to the downstepped level of H+L*, but this issue is is beyond the scope of this paper and is left to future research.

2.2 Broad and narrow focus: perceptual evidence for tonal structure differences

Current theories of intonational phonology assume that when more than one accent is present in a phrase, it is the last one (the nuclear accent) that will be associated with the designated “most stressed” syllable (Designated Terminal Element, DTE). Another important defining feature of the nuclear accent in this framework is that it is the pitch accent which immediately precedes the phrase accent in languages such as English. Under such a hierarchical view, the prediction is that nuclear accented words will be more prominent than prenuclear accented words, which in turn will be more prominent than unaccented words. Hence, different pitch accent categories will be characterized by the same degree of prominence (or at least by degrees of prominence that do not reverse the pattern of relative prominence of the prenuclear and nuclear stress within the same phrase) when in the “special” DTE position. In other words, according to this view a H* would render its associated word as prominent as a word to which a !H* (a downstepped accent) would be associated. This would be true only by virtue of purely positional factors. By extension, the same hypothesis would predict that the H+L* of broad focus Neapolitan statements would be as prominent as a narrow focus statement accent (i.e., L+H*), when in nuclear position, despite their obvious acoustic differences.

Moreover, one of the claims of the standard theory of the relationship between focus and accent is that the phonological form of an utterance with late accent placement will be ambiguous between a broad focus interpretation and a late narrow focus interpretation (Jackendoff 1972). For instance, a pattern such as Giovanni ama MARIA “John loves MARIA” (with nuclear accent on MARIA) would be potentially ambiguous between a late narrow focus and a broad focus reading. In other words, this utterance could be the answer to any of the questions in (1), and could equally be the phonological expression of all of the focus structures in (2):

1a. Che succede? “What’s up?”
1b. Cosa fa Giovanni? “What does John do?”
1c. Chi ama Giovanni? “Whom does John love?”
2a. [Giovanni ama MARIA]F
2b. Giovanni [ama MARIA]F
2c. Giovanni ama [MARIA]F
Utterances that are felicitous in relation to question (1a) will have the focus structure in (2a), while utterances that are felicitous with question (1b) will have the focus structure shown in (2b). Finally, utterances that are felicitous with question (1c) will have the focus structure in (2c). All of these utterances will share a late nuclear accent placement, despite the fact that the intended focus is broad (on the entire sentence) in (2a) and (on the entire verb phrase) in (2b), while it is narrow (on the object noun) in (2c).

According to the prominence hierarchy, all the late accents of the utterances in 2) are expected to be equally prominent, thus contributing to the ambiguity in the interpretation of focus structure. Note that this potential ambiguity is caused by the lack of a one-to-one correspondence between focus and accent placement, and represents a problem for a theory that would directly derive one from the other. The question is, then, from a perceptual point of view: “How do listeners determine the breadth of focus, given a specific accent placement?”.

We also know that accent type differences have been observed for statements in a number of Italian varieties including Neapolitan (as well as for a number of other languages, cf. §2.1 above). As we saw in the section above, there is evidence that the Neapolitan broad focus accent is a H+L* fall, while the narrow focus statement accent has been recently analyzed as a L+H* rise\(^{14}\) (D’Imperio 1999), as the one shown in Figure 1. Note that, in addition to the direction of the melodic contour, the broad focus utterance is characterized by a relatively shallow \(F_0\) variation within the nuclear accented syllable (marked with H+L*), as opposed to the greater \(F_0\) excursion within the narrow focus L+H*. Narrow focus statements differ from broad focus ones in that they present a more acoustically salient accent. Hence, the utterance Giovanni ama MARIA would have a H+L* on Maria only if the intended focus is broad, and a L+H* accent if the focus is narrow. Accent type considerations might predict that the acoustically salient L+H* of narrow focus statements will facilitate the task of focus extraction, and that the converse would be true for the less acoustically salient H+L* of broad focus statements. Questions do not seem to show an analogous pitch accent difference between narrow and broad focus (cf. D’Imperio(2001)). If there is indeed no parallelism between the broad/narrow focus accent opposition between Neapolitan Italian questions and statements, we can predict that listeners will be confused between a narrow and a broad focus reading only in questions.

To sum up, the standard accounts of nuclear stress can be understood as a purely “positional definition”. The nuclear stress is the syllable in DTE position. It is also the syllable positioned to hear the last pitch accent in the intonation contour, or the accent positioned just before the phrase accent. The common thread underlying all of these definitions is that sentence stress (as represented, e.g., in the grid) is independent of accent type, even if it is not independent of intonation, as assumed in early Generative Phonology. I shall refer to this hypothesis as the “positional hypothesis”. Moreover, the traditional generative view of the relations between the pragmatics of focus and accent placement claims that an accent placed on the last element of an utterance can ambiguously signal a broad or a late narrow scope of the focus. By the same token, the accent structure of stimuli with either (intended) broad or late narrow focus will be equally perceptually ambiguous. I shall refer to this as the “(late accent) ambiguity hypothesis”. The data presented here tested these hypotheses for Neapolitan Italian. If the hypotheses are not supported by the data, a contrastive phonological analysis of the broad and narrow focus pitch accents would be supported, thus excluding a common H* analysis for the two. Our results provide evidence against the ambiguity hypothesis, though only for statements.

2.2.1 Materials and Procedure

The stimuli consisted of a set of sentences with different word number, focus pattern and modal intonation (question vs. statement). Four groups of sentences, with four sentences in each group, were created. Group I consisted of three-word sentences, uttered as statements; group II consisted of two-word sentences, uttered as statements; groups III and IV were uttered as questions, and included respectively three-word and two-word sentences. All of the sentences had either an SV (Subject-Verb, e.g. Mario esce “Mario goes out”) structure or an SVO (Subject-Verb-Object, e.g. Maria ama Giovanni “Maria loves John”), depending on the number of words in the sentence. The words employed had a variable number of syllables as well as a variable lexical stress pattern (initial vs. non-initial). As shown in Figure 9, each of the sentences for each type was uttered as either a neutral utterance with broad focus (Broad) or as a narrow focused utterance, where scope of focus was limited to a single word, i.e., either Subject (Subject focus), Verb (Verb focus) or Object (Object focus) in the

---

\(^{14}\)Note, though, that this accent had been previously analyzed as H*+L (D’Imperio 1997a). See also §3.2.
case of three-word utterances, not shown in the schema). The set of sentences was produced by two speakers of Neapolitan Italian (the author and a male speaker), each of whom read half of it.

The recordings were made at the Linguistics Laboratory of The Ohio State University, where they were also digitized at 16 kHz on a SUN Sparc Station using ESPS Waves+. The stimuli were coded and pseudo-randomized using a Latin square design and were then recorded on a tape, which was later played to 22 Neapolitan subjects. The subjects performed a forced choice task, that is they were asked to mark only the one word that appeared to be the most “important” (even when uncertain) in the utterance heard.15 For additional details on the method, see D’Imperio (1997a).

2.2.2 Results and discussion

Here, only the results relative to two-word utterances will be presented since the three-word results were very similar (for details about three-word results, see D’Imperio (1997a)). The results were transformed into percentages of “assigned importance” to the Subject (S), since in Subject-Verb utterances only one of the two elements can receive prominence (i.e., either the Subject or the Verb).

Figure 10 shows the interaction data for two-word questions and statements. Here, “percent of assigned importance to the Subject” is plotted on the y axis, for either Broad, Subject or Verb focus utterances. Separate scores for questions and statements are shown. As expected from the predictions of the positional hypothesis, the highest score was found in utterances with Subject focus, while the lowest was found in Verb focus utterances. Nevertheless, contrary to the expectations, Verb focus statements received a mean of 22.5% responses of assigned importance to the Subject (against a score of 6.7% for Verb focus questions).

An additional result not predicted by either the positional or the ambiguity hypothesis was the conspicuous difference between scores for broad focus questions and statements. While broad focus statements received a percentage of assigned importance to the Subject that was around chance (55%), the score for broad focus questions was much lower (16.25%), closer to Verb focus question scores (6.7%). The results for three-word utterances replicated those for two-word utterances.16 What is more, the results were even replicated through a different (“question-matching”) task (D’Imperio 1997a).

To sum up, the experiment attempted to assess the percept of prosodic prominence through the notion of “importance”. However, the results are difficult to interpret in this light, and seem to suggest that listeners were responding in terms of focus structure, and not simply in terms of accent structure. Contrary to the predictions of current phonological theory, a significant difference was found in the patterning of the data between utterances with broad focus and late narrow focus (on the verb in two-word utterances and on the object in

---

15 Use of linguistic terms such as “prominence” and “focus” was avoided.
16 The ambiguity between broad and verb focus questions was actually even stronger in three-word utterances with late accent placement.
three-word utterances). Moreover, while a traditional view of the relationship between accent placement and focus structure predicted that late accent utterances in general would show identical responses for both broad and late narrow focus (ambiguity hypothesis), our results supported this only for questions.

This can be explained with the observed differences in accent type between questions and statements, on the one hand, as well as between narrow focus and broad focus on the other. Broad focus statements in Neapolitan Italian possess a nuclear accent that has a downstepped quality and is acoustically less salient than narrow focus accents. This could explain the uncertainty on the part of the speakers in a focus extraction task, hence the chance level results. Downstepped pitch accents are also more difficult to transcribe. In English, for instance, transcribers either disagree on the placement of the downstepped \!H\* or fail to recognize its presence in the utterance altogether (Pitrelli, Beckman, and Hirschberg 1994). This suggests that a simple positional definition of nuclear stress prominence is not enough and that accent type considerations will have to be formalized to integrate the theory. Certainly, the problem of the relation between accent status and prominence on the one hand and accent structure and focus extraction on the other is particularly intricate for Italian. While there is no doubt that Italian possesses some kind of sentence stress, we are still not sure about the constraints operating on it. Among other things, it is still controversial whether or not the nuclear accent, is, as in English, the pitch-accent that immediately precedes a “phrase accent”, or if a different characterization is needed. Future research will try to address these issues.

3 Open issues: downstep

3.1 Downstep and broad focus statements

The suggestion that downstepped accents might mark less prominent syllables is complicated by the fact that an explicit account of downstep for Italian has yet to be proposed. How can we then account for the “downstepped” level of the H+L* nuclear pitch accent in statement tunes? While the L target of the H+L* lies on the speaker’s baseline, its H target is not as “high” as the peak of accents such as H*, L+H* or L*+H. This might suggest that the H target of H+L* has undergone downstep, and one might be tempted to transcribe it as \!H+L*. It is not clear, though, whether a downstep process needs to be postulated for the Italian intonational
system, and what the conditions for its triggering would be.

However, if no downstep is postulated, we are left with a problem of phonetic implementation of what is transcribed as a H tone. Such a downstepped quality is even more evident in long sentences, as we can notice from the utterance shown in Figure 5 above. As we can see, the $F_0$ stretch from the fall (the one preceding da Lalla) following H* to the high target of H+L* is not visibly tilted downwards.

The implementation of the $F_0$ level in the H target of H+L* could be attributed to an overall pitch range compression in the accented region. However, if range compression is the account we choose, we would need a model of compression acting only upon the topline (the higher part of the speaker’s range), similarly to downstepped English accents. Moreover, the origin of such a postulated range compression would be unclear. Another option is to describe H+L* as an “intrinsically” downstepped pitch accent. This choice, though, might lead us to implicitly acknowledge the existence of a “third” tonal level, intermediate between a H and a L.

These are all questions which cannot be answered at this point, but which need to be answered through a more careful phonological as well as phonetic investigation of this accent in different phrasal and segmental contexts. For instance, downstep would have to be excluded as a possible hypothesis if upon observing the same target level in an absolute initial position in the phrase, and comparing it to the H target of a H* or of a L+H* in the same position, it transpired that the H in H+L* was consistently lower than in any other instantiation of a H tone for different pitch accents.

In those cases where H+L* is in absolute initial position, we evidently cannot invoke the effect of a preceding pitch accent as a downstep trigger. That is because downstep, at least in the formulation proposed by Pierrehumbert (1980) is a stepwise lowering of immediately successive melodic peaks in an utterance. In English, the process is triggered by a preceding bitonal pitch accent consisting of opposite tones (Pierrehumbert and Beckman 1988). Some informal observations on Neapolitan Italian point to the fact that the seemingly downstepped level of the H in H+L* is not due to position relative to a preceding pitch accent, since an identical phonetic implementation of the H is found when H+L* is in absolute initial position in a phrase. These results are still preliminary and more research is needed.

3.2 Downstep and postfocal accents

The suggestion that downstepped accents might mark less prominent syllables is also especially intriguing when we compare the role of downstep in Swedish and in Neapolitan Italian. Previous studies have shown that, similarly to Swedish, early focal accents of Neapolitan yes/no questions are not followed by the predicted flat melodic configuration following nuclear accents in languages such as English. In Swedish, in fact, the focal accent is not positionally defined as is the nuclear accent of English (since it does not have to be the last accent in the phrase) but is marked by a separate tonal event in a sequence of accents, i.e. the “sentence accent” (Bruce 1977).

In fact, note that early focus questions (cf. Figure 2 and Figure 3, upper panel, shown above) present a sharp pitch obtrusion to a pronounced peak (L*+H) on the focused constituent, followed by a smaller peak on the last stressed syllable of the intonational phrase (H*). In an earlier study it was shown that such !H* is postnuclear, since the word to which it is associated is never selected as the most prominent one in the utterance (D’Imperio 2001). Hence, in a manner similar to Swedish, where postfocal accents are downstepped, we might predict a process of downstep in postfocal position for Neapolitan. A production study concentrated on the properties of the focus constituent final fall (as well as its initial rise) in early focus questions with different focus constituent size. It was found that the fall (which I analyze as a HL-) is anchored to the last stressed syllable of multi-word focus constituents, thus behaving as a regular pitch accent. However, when there is only one stressed syllable in the focus constituent, the nuclear L*+H will take over, leaving the HL- sequence to be realized as an appendix of the rise. A tonal event similar to the sentence accent of Swedish has been then proposed to account for the HL- fall (D’Imperio 2001). Namely, it was hypothesized that the fall at the end of the focus constituent is analogous to the sentence accent of Swedish, in that this tone marks the end of the constituent and contributes to the perceived prominence of the focal accent.

A later reanalysis of the narrow focus statement accent as a L+H* rise (D’Imperio 1999) was at the basis of

---

Note though that initial downstepped accents are allowed by other AM theories (cf. Ladd (1996)).
a phrasal reanalysis of its immediately following fall. At first, the narrow focus statement accent was analyzed as a $\text{H}^*\text{+L}$ (D’Imperio and House 1997) since the fall is generally completed within the stressed syllable and the peak is quite early (close to the stressed syllable onset). A subsequent investigation of the medial $F_0$ valley of long focus constituents, shown by the circled section of the $F_0$ contour in the left panel of Figure 11, appears to support a LH analysis of this pitch accent (D’Imperio 1999). This contrasts with the previous HL fall analysis. Hence, the final fall must be attributed to a tonal event that is separate from the pitch accent proper, which is a simple rise. By Occam’s razor, I postulate that this fall is the same tonal event as the HL- of questions.18

The proposal for a downstep process in postfocal position has been recently extended to other Italian varieties (Grice et al. in press) and to postfocal accents of statements. In other words, postfocal accents in Italian statements would not be automatically suppressed through “deaccenting”. This also implies that even if a phrase accent is postulated for a given variety, this accent would not melodically control the region from the end of the focal word up to the end of the intermediate phrase. Rather, the seemingly flat postfocal contour would instead show a compressed accent. Since the postfocal statement accent would be an even more “downstepped” version of the already non-salient H+L*, this would explain why the accent is almost undetectable in the majority of cases. This proposal presents potential problems for descriptions by which the focal accent is followed by an intermediate phrase boundary (marked by a phrase accent), since such an analysis implies that downstep can apply across a prosodic break.

Such a unified account seems to be justified by informal observations. An example is shown in Figure 12. Here the utterance *Vedrai [NONO] e mamma* “You will see [NINTH] and mom” is shown, with focal accent on *nono*. Notice that the postfocal contour is not entirely flat, rather a fall from 163.5 Hz to 142 Hz Hz can be discerned on the word *mamma*. This fall could be transcribed as a !H+L*. As evidence for the postfocal/postnuclear nature of the !H+L*, note the peak alignment of the preceding L+H* accent. As it can be observed, the H peak of the L+H* rise is reached within the stressed vowel. This alignment is typical of statement narrow focus accents (D’Imperio 2000b), and is not the typical alignment of a prenuclear H* (whose H target is generally late, often in the postaccentual vowel). This dismisses an analysis by which the final !H+L* is indeed a regular H+L* nuclear accent. Similar examples appear to exist for other Italian varieties. Avesani (1995), for instance, postulates postfocal accents, which she transcribes as L*. The motivation she invokes for such a labeling choice is that these postfocal accents are usually accompanied by virtually no tonal movement.

The issue then appears to branch into two different questions: “how do we define an accent in Italian?” and “is pitch excursion/level an integral part of the accent definition or is duration (and maybe intensity) a sufficient cue for the presence of an accent?”. The results presented in D’Imperio (2000a) seem to point to the fact that, at least perceptually, prominent lexical items can be signaled by correlates other than $F_0$, though further research is needed to address this issue.

---

18An alternative analysis would be to consider the fall as part of the nuclear pitch accent, which would necessarily be tritonal (LHL). This alternative analysis is rejected here for reasons that go beyond the scope of this paper.

19In English, instead, downstep is blocked across the intermediate phrase (Beckman and Pierrehumbert 1986).
4 Conclusion

In this paper I have outlined most of what we know and what we still do not know in the description of Italian intonation and its varieties. In doing this, I heavily drew upon experimental results within the AM approach. A study on the tonal alignment of the broad focus statement accent in Neapolitan has been taken as further evidence for a H+L* analysis, against an alternative analysis as a fall from a H* accent to a L- phrase accent. This analysis has been widely employed for broad focus statements in other varieties of Italian. Hence, in contrast to languages such as English, and similarly to other Romance languages, narrow and broad focus statements are signaled by different pitch accents, i.e. a rising accent in the first case and a falling accent in the second. The results also showed the influence of lexical stress position on the alignment of the H target within the H+L*. That is, non-initially stressed words appear to show an earlier alignment of the accent fall relative to initially stressed words. This suggests the importance of the left edge prosodic context in languages such as Italian.

In a second study, I employed the results of a perception experiment to support the analysis of the broad focus nuclear accent of statements as being tonally different from the nuclear accent of narrow focus statements. Here, I invoked the notion of acoustic salience. On the one hand, when focus is broad, i.e. when the focused constituent is larger, the task of parsing focus is more difficult. On the other, inherent acoustic prominence due to accent type differences might help the listener infer the focus structure of the utterance. In fact, only late accent questions turned out to be ambiguous in terms of prominence (hence, focus) pattern, while statements did not show such an effect. An explanation must be sought in the fact that narrow focus statement accents are salient rises, while broad focus statement accents are inherently downstepped, though this is an issue for future research.

It was also proposed that downstep might apply “across the board” in the postfocal contours of early focus statement as well of yes/no questions. Among the open questions for research, I underlined the need of a thorough investigation of tonal alignment, downstep (especially concerning the melodic level of the broad focus statement accent) and phrasal falls in the various Italian varieties. Also, the suggested interaction between intra-word lexical stress position and the alignment of the nuclear H+L* fall should be further researched. Finally, perception experiments are needed in order to test postulated contrasts among tonal categories.

Acknowledgments

The experimental results of §2.1 were first presented at the Fall 1995 meeting of the Acoustical Society of America in St. Louis, Missouri (USA). The experiment presented in §2.2 was supported by a Cognitive Science Fellowship of The Ohio State University (OSU). Both works were carried out while the author was at the Department of Linguistics of OSU. The author warmly thanks Mary Beckman for discussion and encouragement and her colleagues at OSU for feedback and support. Thanks also to Tim Face, José Ignacio Hualde, Pilar Prieto, Erik Willis and an anonymous reviewer for useful suggestions.
References


