The coding of target alignment and scaling in pitch accent transcription

Barbara Gili Fivela

This paper provides data and discussion on the need for teasing apart information that is to be considered part of a phonological transcription of pitch accents and properties that may be coded at a phonetic level. The discussion stems from the study of both production and perception data on the variety of Italian spoken in Pisa, with particular attention to the interplay of alignment and scaling information in the classification of pitch accents.

Acoustic measurements and perception tests show that in some cases, in particular for two accents showing the same starred tone (i.e., H*) in Pisa Italian, the scaling appears to be important for distinguishing pitch accents and for identifying them. This information, both phonetically grounded and significant in distinguishing the patterns, could then be considered important for underscoring the different properties of the pitch accents. Conversely, the presence in both patterns of a low tonal target preceding the peak does not appear to interfere with the subjects' judgments in identifying the two patterns. In the paper it is argued that trying to code all the systematic and significant phonetic properties at the phonological level would lead to a highly complex and redundant system, in which detailed information on scaling could also be included. It will be underscored that in assigning phonological labels, even if they are phonetically transparent, one goal may be to avoid redundancy as far as possible, and to encode only information that distinguishes elements within the system; this would mean that phonetic transcriptions would code all the phonetically relevant characteristics, i.e. the presence of all targets and the significant scaling differences. Within the transcription, then, phonological and phonetic information may be differentiated.[*]

1. Introduction

Differences in the synchronization of tonal events relative to the segmental string and to prosodic units are relevant for the identification of intonational categories, regardless of one's specific framework. This holds true, in fact, for works carried out within the autosegmental-metrical tradition (see, for instance, Bruce 1977; Pierrehumbert 1980) as well as for analyses performed according to the British school of intonation (e.g., Cruttenden 1986) or for studies adopting the tonal movement as the primitive for the description of intonation (see the work by 't Hart *et al.* 1990). In the autosegmental-metrical framework adopted here, the relation between tonal events

Rivista di Linguistica 18.1 (2006), pp. 189-221 *(ricevuto nel settembre 2006)*

and prosodic units is important for determining whether a given event is, for instance, a pitch accent or an edge tone, and which metrical unit it is associated with. In early autosegmental-metrical work (Pierrehumbert 1980), association was assumed to entail phonetic alignment between a tonal event and a unit in the prosodic hierarchy. However, in many subsequent works, especially those representing extensions to ToBI (the labelling system developed for some varieties of English on the basis of Pierrehumbert's model: Beckman & Ayers 1997), many cases in which this assumption is not valid are discussed. For example, there are cases in which either no tonal target or more tonal targets than expected are phonetically aligned with a prosodic unit that is assumed to be associated with a tonal event (such as a syllable associated with a pitch accent). This has led some authors to suggest that tonal alignment cannot be taken as 'the defining characteristic of starred tones, i.e. phonological association' (Arvaniti *et al*. 2000:130), where the starred tone is the tone associated with the metrical unit (e.g., H* for a high tone associated with the metrical unit). Nevertheless, tonal alignment is still considered as a fundamental feature by which different pitch accents can be distinguished, as shown since Bruce's original work on Swedish (Bruce 1977).

In the autosegmental-metrical framework, the scaling of tonal events is relevant to the identification of two main tone levels, high and low, and to the characterization of the interpolation phases between targets found on the two levels. In detailed studies on tone targets, F0 height variations are taken into account for differentiating pitch accents characterized by different F0 levels, such as L* and H* (Gussenhoven & Rietveld 2000), or for identifying bitonal pitch accents, composed of a sequence of targets on two different levels, i.e. high-low or low-high. On the other hand, few studies consider variation in F0 height important for distinguishing pitch accents that are similarly characterized by their main tonal event, regardless of their structure (e.g., for distinguishing L+H* and H*, or H* and H*+L). In the literature on English intonation, reference to a pitch excursion difference is found in relation to the H* and L+H* accents. Bartels and Kingston (1994) report that peak height and raise slope are "by far the most categorical property" as turns out from English listeners' interpretation of non contrastive and contrastive focus; however, other studies more strongly suggest the phonological identity of the two patterns, i.e. the range difference is not considered a distinctive feature (for discussion, see Ladd & Schepman 2003). Thus, in the literature on pitch accent categorization, alignment properties are usually given

more attention, even though some stability in the scaling characteristics is also reported in some works. For instance, in works investigating the segmental anchoring hypothesis, according to which tonal targets are anchored to points at the segmental level, small variations in scaling are expected across speech rates (Ladd *et al.* 1999).[1]

Although it is nowadays widely accepted that phonetic correlates of pitch accents should not be taken as direct indicators of phonological association, it is still the case that alignment characteristics are taken into account for pitch accent categorization. Autosegmental-metrical intonologists still consider alignment characteristics when deciding which label should be used for referring to a specific pitch accent. Moreover, it is still the case that analyses offered within the autosegmental-metrical framework, at least ToBI-like analyses, are supposed to be phonological, i.e. the labels stand for phonological units of the system. However, problems arises because these labels are not unrelated to the phonetic properties of the phonological unit, as opposed to what happens with segmental labels.[2] Rather, they encode information on the presence and main characteristics of pitch accents (e.g., whether one – e.g., H* – or two targets – e.g. L+H* – are systematically produced in the context of a certain prosodic unit, or whether, for instance, the first is low and the second is high – e.g., L+H* vs. H+L*). Thus, on the one hand, intonologists look for distinct labels standing for phonological units in the system, while, on the other, they directly and necessarily encode in those labels some information which is phonetic.

Moreover, apart from the tendency to avoid tritonal pitch accents – in line with Pierrehumbert's (1980) idea of limiting the number of possible tone combinations – and except for the debate on the phonological identity of H* and L+H* in some languages (see Ladd & Schepman 2003 for English), the targets characterizing a pitch accent are usually represented by a corresponding symbol in the label. For instance, no work, to the author's knowledge, suggest labelling as a monotonal accent L* a pattern clearly showing two targets, for instance, a low target systematically followed by a high target that belongs to the accent itself. Thus, in many cases phonological units within intonational systems are represented by distinct and, simultaneously, quite phonetically-detailed labels. The greater importance of the 'distinctive'[3] compared to the 'phonetically-detailed' property of the labels emerges when two patterns show strong phonetic similarities while being different phonological entities. This is what happens in Neapolitan Italian, in which both the pitch accent found in

declarative narrow focus and that used as the focal accent in interrogatives show two targets, L and H, slightly differing in alignment. Clearly giving more importance to having 'distinct' labels, D'Imperio (1999) decided to use L*+H for the nuclear accent of interrogatives after obtaining phonetic evidence for the suitability of L+H* in describing the pitch accent found in declarative narrow focus.

In this work, it will be confirmed that is important to have both 'distinct' and 'phonetically-detailed' phonological labels. However, it will be argued that, although the distinctive properties coded in the labels stem from the phonetic characteristics of the pattern and may be represented transparently, the phonetic characteristics encoded at the phonological level should be as little redundant as possible. On the basis of data from the variety of Italian spoken in Pisa, this work will show that in some cases, above all for two accents showing the same starred tone (i.e., H*), the scaling appears to be important for distinguishing pitch accents and for identifying them. This information, both phonetically grounded and significant in distinguishing the patterns, could then be considered important for underscoring the different properties of the two pitch accents. Conversely, the presence in both patterns of a low tonal target preceding the peak, showing stable characteristics in alignment – stable, for instance, with respect to the prosodic context – does not appear to interfere with the subjects' judgments in identifying the two patterns. Moreover, the low target may actually show a high degree of variability when a different type of factor is considered – in the data shown here, the position in the sentence. Thus, to a certain extent, in assigning 'distinct' and 'phonetically-detailed' labels to the two accents, one might be tempted to give priority to the significant difference in scaling rather than to the presence of a low leading tone target.

The main goal of this paper is to argue that the coding of all the systematic and significant phonetic properties at the phonological level would lead to a highly complex and redundant system, in which detailed information on scaling could also be included. It is underscored that, in assigning phonological labels, even if they are phonetically transparent, one goal may be to avoid redundancy as far as possible, and to encode only information that distinguishes elements within the system; the phonetic transcriptions would then code all the phonetically relevant characteristics, i.e. the presence of all targets and the significant scaling differences. Moreover this work will draw attention to the fact that subjects perceive a large number of cues and may also differ in the cues they pay attention to. On the one hand, this points again to the usefulness of more abstract phonological

representations that do not necessarily code all the information related to one – albeit important – cue, systematically excluding others; on the other hand, it shows that the variability observed in relation to one cue, say alignment, should also be related to the complex interaction with other pitch accent cues, e.g. tone scaling.

2. Phonetic properties and phonological coding

As mentioned in the preceding section, it is nowadays a widespread opinion that phonetic correlates cannot be taken as direct indicators of pitch accent phonological association. Nevertheless, alignment characteristics are obviously still taken into account for pitch accent categorization.

Coherently with Bruce's (1977) and Pierrehmumbert's (1980) proposals, alignment information often plays an important role in deciding which tone in a bitonal pitch accent should be considered the starred one, the tone associated with the bearing unit – see, for instance, the discussion on the analysis of H+L* and H*+L in Grice (1995: 178). However, in a number of studies, alignment details are reported as not offering evidence for deciding which is the starred tone. Examples from different languages show that no target may be aligned with the accented syllable (Arvaniti & Ladd 1995, Arvaniti *et al.* 1998, 2000) or that two targets may be aligned with it (among others, see Grice 1995, D'imperio 1999, Besana 1999, Marotta 2000, 2002 for works on varieties of Italian). D'Imperio (2000) suggested that an acoustic misalignment may be due to an improper view of what a target is, as the acoustic target may be different from the perceived one. However, some authors decided to choose specific labels in order to encode peculiar alignment characteristics, to the extent of modifying the original interpretation of the star notation. This is the case of proposals such as those by Arvaniti *et al.* (2000), Ladd & Schepman (2003), Marotta (2000). Arvaniti *et al.* (2000), for instance, after discarding the choice of the L+H* label for the prenuclear Greek accent, suggest that it be considered L*H* – among several possible notations – on the basis of the observation that "both the L and the H tones are in fact aligned relative to the stressed syllable albeit outside its boundaries" (Arvaniti *et al.* 2000:130). The choice implies that the star is basically meaningless (Ladd, p.c.). Ladd & Schepman (2003) proposed using the label (L+H)* for an English "pitch accent raising across the accented syllable and involving clearly distinct L and H targets". The authors specify that they avoid the L+H* label for

expository reasons only, and make clear that (L+H)* may represent both the L+H* and the H* accents in English, probably involving only a pitch excursion difference. Marotta (2000), on the other hand, suggested labelling Italian rising pitch accents (L+H)* and, conversely, (L+H) in those cases where, respectively, both and no targets are aligned with the syllable. According to Marotta's (2000) analysis these patterns alternate with more transparent ones, where the star is actually transcribed after the tone that is aligned to the syllable. In her proposal, the star would then correspond to alignment characteristics, and the two pitch accents would represent allophonic variants of the same pattern, phonologically encoded as /L+H/.

Observations of the phonetic properties of pitch accents, meaning both alignment and scaling, led some authors to suggest that accents could be differently characterised in terms of their internal structure. Grice (1995) and Frota (2000) focused on the asymmetric characteristics of leading and trailing tones – those, respectively, preceding and following the star tone – and proposed that the pitch accent has an internal structure. In the case of Grice's proposal for the analysis of Palermo Italian, the structure consists in the leading tone being associated with an unstarred syllable and the trailing tone belonging to the same root node as the starred tone; in the case of Frota's proposal for European Portuguese, the said internal structure consists in the trailing tones belonging to the same strong node as the starred tone, while the leading tones are more peripheral in the structure.

More recently the distinctive alignment characteristics of pitch accents have been accounted for by means of secondary association. On the basis of data from Catalan, Neapolitan and Pisa Italian, Prieto *et al.* (2005) show that alignment contrasts in accents involving a high tone associated with the syllable, and raising accents in general, may be accounted for by postulating that the high tone is primarily associated with the syllable, and that, in specific cases – those showing differences in alignment which are contrastive – the tone may get a secondary association to other prosodic units, such as moras and prosodic words. This corresponds to a more transparent phonological coding, which accounts for the presence of contrasts in alignment within the same language or variety of language. Thus, many works point to the need for more transparent transcription and a way to relate phonetic data with phonological coding.

Generally speaking, many experimental studies provide evidence for the presence of targets, such as acoustic targets, articulatory targets or perceptual ones, and – as one of the reviewers observed – an

event either is a target or is not a target. Nevertheless, once an event is recognized as a tonal target, two approaches are possible. Coding the presence of targets may be seen as a way to represent intonational events phonologically. In this case, a one-to-one correspondence between phonetic and phonological targets is assumed, which might be seen as one of the problems of the ToBI labelling system, i.e. the difficulty of separating the phonetic and the phonological information coded by the labels. Alternatively, all targets could be seen as purely phonetic targets, whose presence is related to a phonological, more abstract, specification. In the above case, the phonological representation may be phonetically transparent – directly indicating the presence and characteristics of some targets – but may nevertheless be different from a symbolic description of all the targets. In the paper, this second point of view will be argued for, and it will be suggested that phonological transcription code only information that underscore the difference between elements of the system, and avoid redundancy as far as possible.

This paper will therefore focus on the boundary between, on the one hand, information that is to be considered part of a transparent phonological transcription of pitch accents and, on the other, properties that may be coded as phonetic. The central idea is that phonological analysis stems from phonetic analysis, but should lead to coding in transcription only that information which is necessary to differentiate the contrasting pitch patterns within the phonological system. This means avoiding redundancy as far as possible while preserving phonetic transparency. It will be argued that an attempt to enrich the phonological transcription of all the information related to systematic phonetic properties would lead to a highly complex and redundant phonological coding. Alternatively, the use of phonological labels that are as transparent as possible but encode only 'distinctive' information, makes it possible to code in a strictly phonetic transcription those properties related to the actual realization of the phonological categories. Scaling information, and its interplay with alignment, will play a crucial role in deciding what may be considered part of phonetic transcriptions and what may be seen as important at a phonological transcription level; moreover, observation of the relevance of both alignment and scaling information will suggest a possible explanation for the high variability often observed in alignment (see also Welby and Loevenbruck's proposal (this volume) of an anchorage region within which the target may be aligned).

Both production and perception data will be described in the following sections, with particular attention to both alignment and

scaling details. Section 3 is an overview of the (main) phonetic and functional characteristics of the two pitch accents considered in the paper. Production and perception data on the pitch accents are described with details in 4 and 5, respectively. The acoustic measurements presented in 4 relate to target alignment and scaling considering changes in the nuclear vs. prenuclear position of the pitch accent, and changes of other factors, such as prosodic context. The results discussed in 4.1 show that two accents in a nuclear position, both involving a peak associated with the accented syllable, differ for the actual alignment and scaling of all their targets, yet they are both characterized by a raising to the peak that starts around the beginning of the syllable. The transcription suggested by their formal characteristics would be L+H* and L+H*+L, where the L+ do not contribute to distinguishing the two patterns within the system. However, the scaling of the targets is different, and this information could then be important in the coding of distinctive aspects. Section 4.2 focuses on the first of the two accents, whose alignment and scaling characteristics are observed when it is prenuclear, considering various positions in the utterance. The results of measurements show that, when the accent is realized in the initial position of the utterance, the starting point of the raise may actually be replaced by a straight interpolation from the onset of the utterance. However, when the low target is aligned later than the beginning of the utterance, the raise starting point shows a high degree of variation, depending on the position of the accented syllable. These data suggest that the presence of a low elbow close to the peak does not appear to be a mandatory component of the pitch accent, and moreover its position is highly variable, and definitely more variable than the peak position. Therefore the presence of a high starred tone, implying a raise to a peak in Pisa Italian, may be enough to characterize the pitch accent appropriately, at least from a phonological point of view.

In section 5, the relevance of alignment and scaling information is tested by means of perceptual experiments. The experiment described in section 5.1 relates to the perceptual difference between the two accents considered so far. Results show that the alignment of the peak is significant in relation to the subjects' identification of the pitch accents, and allows subjects to identify the two pitch accents, by differentiating the utterance function and characteristics, regardless of the alignment of the raise onset (that was not manipulated). More interesting, the data also show that scaling information interacts with alignment. It favours pitch accent identification and modifies the reference values for the two pitch patterns, i.e. it affects 'when' the

accents begin to be identified. Moreover, the experiment shows that there is inter-subject variability regarding the cues used to identify the accents, i.e. alignment vs. scaling information.

On the basis of these data it will be proposed that both the presence of the low leading tone target and the information on peak scaling may be left unspecified in a phonological transcription. They may be deduced from the other elements that characterize the phonological unit – such as, respectively, the presence of H* and the monotonal vs. bitonal status of the accents – and may be explicitly coded only within a phonetic transcription. Finally, it will be suggested to explicitly differentiate phonological and phonetic information within the transcription.

3. The pitch accents under investigation

In this paper data will be discussed in relation to two accents that have already been reported as being important for the description of Pisa Italian. The analysis of both read and semispontaneous speech, i.e. Map-Task dialogues,[4] showed that the accents are formally distinct and functionally different, and may be phonologically analysed as H* and H*+L (Gili Fivela 2004).

As far as formal characteristics are concerned, H* and H*+L have in common that they are both characterized by a raise to the peak. In both, the raise starts around the beginning of the accented syllable. In H*, it ends in a peak which is aligned the right end of the syllable, while in H*+L the peak is realized earlier in the syllable and is followed by a low target[5] – see figure 1. Detailed measurements – see section 4.1 – show that the starting point of the raise has all the phonetic properties found in targets usually considered phonologically significant. Nevertheless, the target behaving as a leading tone does not distinguish the two categories, which, however, are unambiguously differentiated by the choice of the starred tone together with the absence/presence of the trailing tone. Moreover, the specific – and contrastive – alignment properties of the two accents may be more precisely accounted for by postulating a secondary association of the high tone (see the work by Prieto *et al.* (2005), mentioned above).

Functionally the H* pitch accent is found in both statements and questions. In the former, it may appear both as prenuclear and as nuclear, although mainly in early positions within the utterance, as occurs in the case of left dislocated constituents – see example in figure 1 and 4.1, where it occurs in nuclear position (for further discussion,

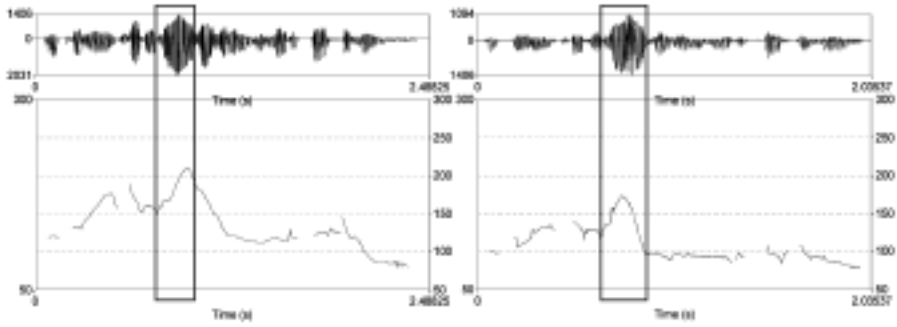

Fig. 1. Two productions of the sentence *La pronuncia di lavaglielo non (la) ricordo mai* "The pronunciation of 'wash it for him', I never remember (it)": the same speaker realizing a broad (H* on the target word) and a contrastive interpretation (H*+L on the target word) of the sentence – respectively, left and right panel (the relevant syllable is highlighted).

see fn.8); the pitch accent is often the nuclear one in patterns expressing continuation, when followed by either high or low edge tones. In statements, a H* may then be used to express narrow focus, but when a contrastive interpretation is at issue, it appears to be related to a weaker illocutionary force variation[6] (the accent sounds less peremptory and conclusive than the H*+L one) and/or to the presence of a syntagmatic contrast rather than a paradigmatic one, e.g. in sentences such as *Pietro e Giovanni vanno in montagna? No, Giovanni va al mare; Pietro va in montagna* 'Do John and Peter go to the montains? No, John goes to the seaside; Peter goes to the mountains' (see also Gili Fivela 1999). In questions, H* pitch accents are found in a nuclear position in elliptic questions, in some confirmation-seeking questions realized with a low illocutionary force, and in questions suggesting a number of alternatives.

The pitch accent phonologically analysed as H*+L expresses narrow focus with a contrastive interpretation, and as such may also be used by speakers to ask for confirmation. In Pisa Italian Map-Task dialogues, in fact, the pitch pattern H*+L followed by low edge tones is interpreted as a confirmation-seeking question when uttered by instruction followers, probably because they explicitly lack information in comparison to instruction givers, and this leads to the utterance being interpreted as a question. However, instruction givers may ask for confirmation by realizing the same pitch accent followed by a different edge tone, i.e. by a high boundary tone. The latter phonological analysis is also used for patterns conveying

incredulity/surprise,[7] although in these cases a high left boundary tone or a later alignment within the nuclear syllable and/or a greater pitch excursion may be found.

Thus, the two accents are functionally distinct, although in some specific contexts they may play partially overlapping roles. As mentioned above, for instance, both H* and H*+L may express narrow focus and some kind of contrast. These two pitch accents will be considered in the next sections, where the results of acoustic and perceptual studies are reported in order to determine both the role of alignment and scaling characteristics in pitch accent production and perception, and finally to indicate what information is important for phonological coding of the tonal events.

4. Production data on alignment and scaling

The issues discussed in this section stem from work carried out on read speech. On this speech style, it was possible to perform measures of the acoustic signal corresponding to pitch accents involving H*, i.e. H* and H*+L, and to treat the results statistically.

In the first section, the patterns are investigated in a nuclear position – the work is described in detail in Gili Fivela (2002, 2004); in the second section, only the H* accent is at issue, when realized in a prenuclear position – the discussion stems from part of a wider work on both Pisa and Neapolitan Italian, presented during the Tone and Intonation in Europe Conference, held in September 2004 (Gili Fivela & D'Imperio manuscript).

4.1. Pitch accents in a nuclear position

Previous studies focused on the alignment and scaling characteristics of tonal targets in two Pisa Italian pitch accents, both showing a high tone associated with the nuclear syllable (Gili Fivela 2002; Gili Fivela 2004). The two pitch accents were analysed when produced in nuclear position and followed by a low edge tone, on the final word of a left dislocated constituent. In that context, a raising accent was obtained by forcing a broad interpretation of the sentence, while a falling accent was realized when a contrastive interpretation of it was suggested by a previous question – see figure 1.

Detailed latency measurements on three Pisa speakers' productions showed that both accents are characterized by the presence of a peak which is, however, aligned earlier in the

contrastive interpretation than in the broad one (at about 50% vs. 100% of the (open) syllable duration; the difference is about 80 ms on average, considering the latency from the syllable onset). Both patterns show a low target leading to the peak, differently aligned with respect to the syllable onset, but stably aligned with respect to the peak itself. In fact, the low target alignment with respect to the peak does not change in relation to the number of preceding syllables, although its latency from the syllable onset differs in the contrastive and the broad interpretation (respectively, 1 ms vs. 25 ms from syllable onset, on average). Due to the experimental design, both pitch accents are followed by a low phrase accent – indeed, they are both produced by the end of a left dislocated constituent, realized in a separate intermediate phrase (Gili Fivela 1999).[8] Although they both show a fall after the peak, only the pitch accent expressing contrast shows a low target that is realized close to the high target. In the contrastive accent, in fact, the latency between the peak and following low does not change in relation to the number of syllables realized between the accent and the constituent boundary (on average, however, the target is aligned 10 ms before the end of the postaccentual syllable onset in the contrastive interpretation, and 133 ms after it in the broad one). The characteristics of the contrastive accent were confirmed by a later study on the varieties of Italian spoken in Pisa and Bari (Gili Fivela & Savino 2003 – see also fn.5).

The two pitch accents also differ with regard to target scaling, in that the pitch accent exploited for contrast shows lower values than the pitch accent realized in the broad interpretation – see, for instance, figure 1. The average F0 height of the three targets in the speaker productions is shown in table 1.

Table 1. Average F0 height of the targets in the productions of the speakers.

	Speaker 1		Speaker 2		Speaker 3	
	Broad	Contrastive	Broad	Contrastive	Broad	Contrastive
Left elbow	207Hz	166Hz	206Hz	144Hz	136Hz	109Hz
Peak	346Hz	234Hz	297Hz	174Hz	170Hz	135Hz
Right elbow	185Hz	155Hz	163Hz	141Hz	108Hz	97Hz

Summing up, the two accents differ in the actual alignment and scaling of all their targets: the accent exploited for contrast shows an

earlier alignment of both the low leading tone and the peak, an earlier low target after the peak (which is actually a low trailing tone) and lower F0 values for all its targets, especially for the start and end point of the raise – see figure 2 for a schematic representation of the main differences. Notice also that the syllable in the contrastive interpretation is longer than in the broad focus one.

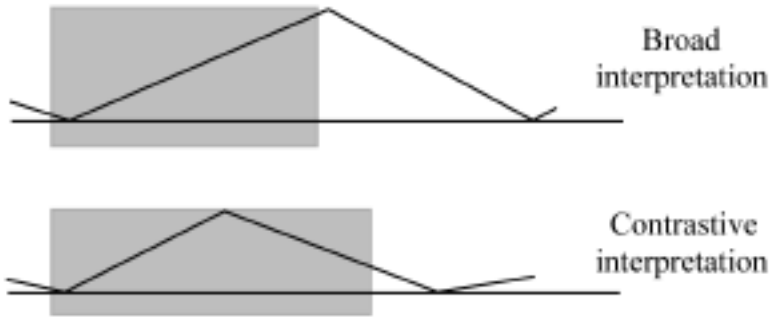

Fig. 2. Schema of the two pitch accents alignment and scaling characteristics (the grey area represents the nuclear syllable) - adapted from Gili Fivela (2002).[9]

Although the alignment details might suggest an analysis of the broad pattern as L+H* L- and of the contrastive pattern as L+H*+L L-, Gili Fivela (2002) proposed that, at a phonological level, the two accents could be distinguished solely on the basis of the absence vs. presence of the trailing tone, and labelled as H* and H*+L. Therefore, the low target, behaving as a leading tone, was considered a structural property of H* in Pisa Italian, and interpreted as redundant in a phonological transcription. On the other hand, when a phonetic transcription of tonal events was at issue, the low leading tone target could be correctly transcribed – in fact, it was phonetically realized. In other words, the low target representing the raise onset was taken as an intrinsic property of H* in Pisa Italian, a pitch accent that is realized thanks to a raise, i.e. not at the end of a plateau. In order to distinguish this redundant information, square brackets were used to highlight the low tone transcription: the two accents were then phonetically labelled as [L+]H* and [L+]H*+L.[10]

These considerations all stemmed from the observation of alignment differences. It is worth noticing, though, that the scaling of F0 targets is systematically and significantly different, and this fact, in principle, could be encoded in a transcription (for instance by means

of arrows – e.g. [L+]H*↑ and [L+]H*↓+L). This topic will be addressed in section 5.1. The next section focuses on the realization of [L+]H* in a prenuclear position. The results of acoustic measurements suggest that the presence of a low elbow close to the peak does not appear to be a mandatory component of the pitch accent, and that, in any case, its position is more variable than the peak position. This supports the idea that a raise to a peak may be enough for the appropriate characterization of the pitch accent. Therefore the peak – implying a raise to it – may be the phonological information to be coded within a transcription aiming at distinguishing patterns while offering phonetically transparent, although not redundant, information.

4.2. Raising accent in prenuclear position

A study of the raising accent in prenuclear position was presented by Gili Fivela and D'Imperio during the Conference "Tone and Intonation in Europe", held in Santorini, Greece, in September 2004. The study was conducted on the varieties of Italian spoken in Naples and Pisa, but in this paper the discussion will only refer to data on the latter variety. Results of the analysis of Pisa Italian recordings confirm the data described in the literature, i.e. that pitch accent position in the utterance and prosodic context play a major role in changing the phonetic characteristics of the accent itself. In particular, data on the Pisa Italian pitch accent realized at the beginning of the utterance show that the starting point of the raise may be variably aligned, given that it can be realized even at the very beginning of the utterance rather than around the accented syllable boundary. This is taken to be an example of phonetic effects, rather than evidence of a different phonological category. In fact, the H* always has to be preceded by a raise to the peak.

The corpus collected for studying the accent variations in prenuclear position consisted of phonotactically acceptable target words, inserted in a carrier sentence favouring their interpretation as proper names or brand names. The target words varied with regard to various factors. Those relevant for the present study are the following:

– number of preaccentual syllables: the accented syllable could be preceded by from 0 to 2 unstressed syllables (e.g., *Lana* 'wool' or or proper/brand name, *Melana* as a proper/brand name, *La Melana* as a proper/brand name);
– position of the target word in the sentence, which determined the following conditions:

'1s': the word was the first – and only – one in a simple noun phrase (e.g., *'Lana' è il libro di Giovanni* '"Lana" is John's book');
'1c': the target word was the first one in a complex noun phrase (e.g., *'Lana della Melanide' è il libro di Giovanni* '"Lana of Melanide" is John's book');
'2c': the word was the second one in a complex noun phrase (e.g., *'Il melone Lana' è il libro di Giovanni* '"Lana melon" is John's book');

– realization: speakers were asked to read the corpus both naturally and pretending to be far from the listener, i.e. raising the pitch range.[11]

Two Pisa (and two Neapolitan) speakers read aloud five repetitions of the eighteen target sentences. Both manual and automatic labelling was carried out on the digitised utterances. Segment boundaries, F0 peak position, and two reference points for the automatic insertion of the low elbow were manually labelled, while, as just mentioned, the low elbow was automatically marked. The algorithm inserted the elbow according to a two slope best fit regression analysis, basically when no linear interpolation between the beginning of the utterance and the peak was found. The labels allowed the authors to measure segment duration, latencies – between tonal events, and between tonal events and segment boundaries – and F0 values of tonal events.

Results on Pisa Italian show that in a number of cases a linear interpolation between the utterance onset and the peak was found, and no elbow was realized right before the peak. Considering stimuli with 1 and 2 preaccentual syllables, in fact – see table 2 – the percentage of cases of linear interpolation appears to be higher when the target word is the first one in the utterance, especially when it is first in a complex phrase; moreover, the percentage is much greater when the utterances are realized with a raised pitch range.

Table 2. Number of utterances showing a linear interpolation between the utterance F0 onset and the peak – i.e. no elbow was detected before the peak.

	First in complex NP	First in simple NP	Second in complex NP
Normal pitch range	4/40 (10%)	2/40 (5%)	0/40 (0%)
Raised pitch range	11/40 (27.5%)	10/40 (25%)	1/40 (2.5%)

Thus, the fact that the elbow may actually be replaced by a straight interpolation from the beginning of the utterance suggests that the presence of a stable low elbow – a typical low leading tone – does not appear to be a mandatory component of the pitch accent, and that a raise to the peak may be enough for the pitch accent to be (acoustically) realized.

Nevertheless in many cases the low elbow was actually produced – see figures 3 and 4 – and detected by the algorithm used for the automatic labelling. Both in the case of natural pitch range productions and in the case of raised pitch range ones, a low elbow preceding the peak may be observed, although its alignment characteristics change. The elbow, when clearly visible, is realized earlier than expected on the basis of the measures discussed in the previous section.

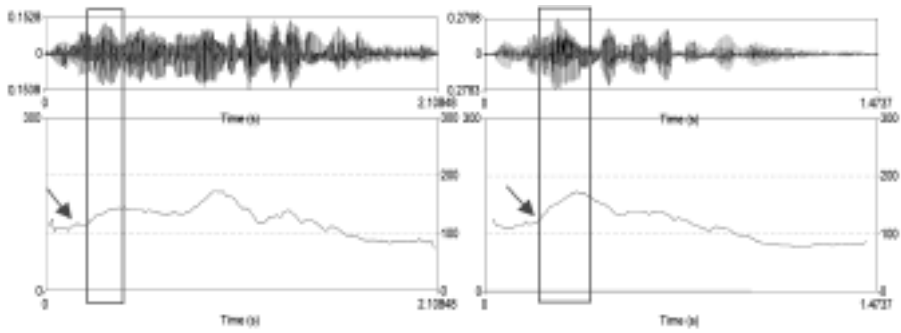

Fig. 3. Natural pitch range: waveform and F0 track of utterances showing target words in first position in a complex noun phrase – left panel, *La Melanda della Melandide è il libro di Giovanni* "'La Melanda of Melandide" is John's book' – and in first (and only) position in a simple noun phrase – right panel, *La Melanda è il libro di Giovanni* "'La Melanda" is John's book' – the relevant syllable and the elbow position are highlighted.

ANOVAs carried out on the measurement results, in fact, showed that the target word position always affects tonal target characteristics, both in normal and raised pitch range productions. The main results for normal pitch range production are summarized in what follows.

The accent realized at the very beginning of the utterance, and preceding a second pitch prominence produced within the same phrase, is characterized by a gradual raise. In this case the peak represents a sort of subordinated accent to the second and final one

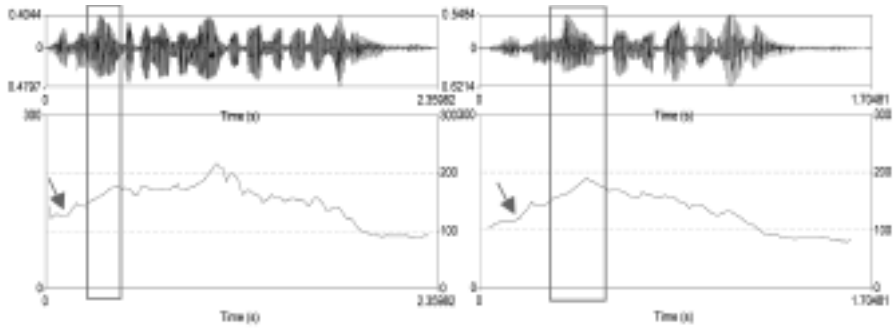

Fig. 4. Raised pitch range: waveform and F0 track of utterances showing target words in first position in a complex noun phrase – left panel, *La Melanda della Melandide è il libro di Giovanni* '"La Melanda of Melandide" is John's book' – and in first (and only) position in a simple noun phrase – right panel, *La Melanda è il libro di Giovanni* '"La Melanda" is John's book' – the relevant syllable and the elbow position are highlighted.

in the phrase, it is realized by the very end of the syllable, and its position is not affected by the number of preceding syllables. The elbow is not always realized, i.e. the target may correspond to the beginning of the utterance in 10% of the investigated contexts (an even greater percentage – 27.5% – is found in the case of raised pitch range); when it is realized, it is aligned before the accented syllable, and its distance from the syllable onset and the peak increases when more preaccentual syllables are produced.

Similar characteristics with respect to the prosodic context are found when the pitch accent is at the beginning of the utterance and is the only one within the phrase, despite the difference in alignment and scaling. In this position, the accent is characterized by a greater pitch excursion, the peak is aligned somewhat earlier with respect to the syllable boundary, but its position is not affected by the number of preceding syllables. The elbow is not realized in a small number of cases, i.e. the target may correspond to the beginning of the utterance in 5% of the contexts investigated (although an even greater percentage, 25%, is found in the case of a raised pitch range); when the elbow is realized, it is aligned at the beginning of the accented syllable, but, again, its distance from the syllable onset and the peak increases when more preaccentual syllables are produced.

On the other hand, the pitch accent realized in second, and final, position within the phrase shows a different sensitivity to the prosodic context, specifically meaning the number of preaccentual

syllables (actually the prosodic context is different when considering the tonal context, as, in this case, another peak prominence precedes the pitch accent). The pitch accent is characterized by a smaller F0 excursion, although the peak reaches values similar to those found in '1s' condition:[12] in this case part of the raise from the F0 onset has already been achieved during the first pitch accent production. As in the accent found in '1s' condition, the peak is aligned before the syllable right boundary, and, similarly to both the contexts just discussed, its position is not affected by the number of preceding syllables. Nevertheless, unlike the previously mentioned patterns, in this position the elbow is always realized, i.e. the target never corresponds to the beginning of the utterance (although, in the case of raised pitch range, a small percentage of cases of straight interpolation from the utterance F0 onset – 2.5% – is found); the elbow is aligned slightly later within the syllable compared with the previous contexts, but its position with respect to the syllable onset and the peak is not affected by the number of preceding syllables.[13] The formal characteristics of this realization of the accent are similar to those found in nuclear position – see 4.1.

A schematic representation of the characteristics found in the three different utterance positions for normal pitch range productions is shown in figure 5. The schema shows the differences in alignment of both the peak and the previous low elbow, and the differences in both target scaling. The schema refers to cases where an elbow between the utterance F0 onset and the peak was found. Nevertheless, it is worth recalling that in the normal pitch range condition, 10% of productions do not show any elbow when the accent is realized in complex phrase first position (1c), 5% of productions do not show the elbow when the accent is the only one in the phrase (1s), while elbows are always found when the accent is the second and final one in the phrase (2c) – see table 2, above.

Thus the pitch accent alignment and scaling characteristics change, not surprisingly, depending on its position in the utterance. Nevertheless, the three patterns found in different utterance positions are considered allophonic variants of the same phonological pitch accent. In particular, the similar behaviour with respect to prosodic context observed when the patterns are realized in utterance initial position, despite alignment and scaling differences (conditions '1c' and '1s'), and, on the other hand, the observation of similarities in alignment and scaling when the accents are the main ones in the phrase (conditions '1s' and '2c') suggest that the patterns be considered allophonic variants of the same phonological unit.

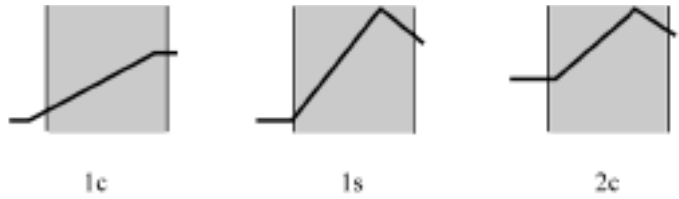

Fig. 5. Schematic representation of the alignment and scaling characteristics found in normal pitch range productions for the three different utterance positions: first in complex NP ('1c'), first in simple NP ('1s'), and second in complex NP ('2c'). The grey area represents the accented syllable.

While these observations do not support other hypotheses, such as the one that considers the initial raise as a different phonological entity, e.g., a phrasal tone (Frota 2003). Thus the accent characteristics depend on accent position in the utterance (and in the phrase) and on the metrical relation with the other – if any – elements in it. This is in line with one of the major assumptions of autosegmental intonational phonology (Pierrehumbert 1980), i.e. the existence of an inventory of pitch accents that may appear in various positions in the utterance.

Of relevance to the point being made here, the production of the accent in utterance initial position appears to affect the low target alignment. Again this is taken to be a phonetic effect. The beginning of the raise may actually be aligned earlier, even at the beginning of the utterance, while the peak, which also appears to be more variably aligned in utterance initial position, is realized within the syllable.

5. Perception effects of target alignment and scaling variation

The interplay of alignment and scaling variations and the perceptive salience of the low target were tested by means of a perception test described in section 5.1. All manipulations were performed using PRAAT – developed at the Institute of Phonetic Sciences of the University of Amsterdam by Paul Boersma and David Weenink – and through PSOLA resynthesis. The experiment, reported with details in Gili Fivela (2005), exploits the contrast between accents phonetically characterized by similar targets ([L+]H* vs. [L+]H*+L), and focuses on the importance of alignment and scaling characteristics.

5.1. Accents in utterance initial position

The interplay of target alignment and scaling is explicitly investigated in Gili Fivela (2005). The study focuses on the [L+]H* and [L+]H*+L pitch accents, that showed significantly different acoustic alignment and scaling of all their targets. A perception experiment carried out by exploiting manipulated stimuli showed that the F0 height of tonal targets affects subject judgements, and plays a role in the identification of the two pitch accents.

In this experiment, subjects were asked to listen to stimuli that had been manipulated to obtain a gradual shift from an [L+]H*+L accent to an [L+]H* L- pattern, i.e. an [L+]H* accent followed by a low phrase accent. The manipulation was carried out on the basis of mean latency values obtained for tonal targets in a previous study (Gili Fivela 2002).[14] The relevant interval for realizing the intermediate characteristics between the two pitch accents was found by considering, for each pitch pattern, the mean values related to the first low target – starting point of the raise – and to the peak – the end of the raise.[15] Starting from an utterance corresponding to a contrastive interpretation of the sentence *Ho detto velava* 'I said veiled' (i.e. the pitch accent realized on the word *velava* was [L+]H*+L), the pattern corresponding to a [L+]H* L- was obtained by varying the alignment of both the peak and the falling phase of the [L+]H*+L pitch accent in 8 steps, and by modifying the scaling of all the targets in 2 steps[16] – see figure 6, left and right drawing respectively. Thus the 9 stimuli corresponding to the eight-step manipulation plus the natural stimulus – which was, however, resynthesized – were either maintained unchanged with regard to the target scaling (coherently with a [L+]H*+L accent pitch range) or manipulated in order to obtain both the natural height of targets for a [L+]H* accent and an intermediate value.

The 27 stimuli were then presented five times (a total of 135 stimuli) to subjects who were instructed to listen to each stimulus and judge whether it represented a peremptory and conclusive correction or a neutral statement. It was made clear that the neutral statement option could also be considered as a sort of correction in case it was followed by other material – as a sort of motivation for the correction – closely linked to the preceding one from a prosodic point of view. As discussed in 3, in the [L+]H* L- pattern, the pitch accent often appears utterance initially, is prominent and, among other functions, may easily convey narrow focus, making the utterance potentially sound like a correction, although not a peremptory and

conclusive one. Thus, on the one hand, considerable ambiguity was expected – and confirmed by a small pilot perception test – if subjects were asked only to judge the suitability of the pattern to a context of correction; on the other hand, if subjects were asked to focus only on the conclusive feature – recalling the finality/non finality opposition – it was not possible to ensure that the pitch accent properties rather than the phrase accent phonetic characteristics were involved in the subject judgements (moreover, data on the perception of discourse structure in Pisa Italian showed that subjects do not appear to rely on pitch accent type to judge the utterance position within the discourse (Savino *et al.* 2006). Finally, the hypothesis that, in Pisa Italian, pitch accent type may be associated to degrees of strength in the speaker purpose, and may be considered as peremptory to different extents (Gili Fivela 2004), was clearly related to pragmatic factors. Thus the question subjects were asked was necessarily a composite one – involving a reference to both a function and a difference in strength of speaker's purpose.

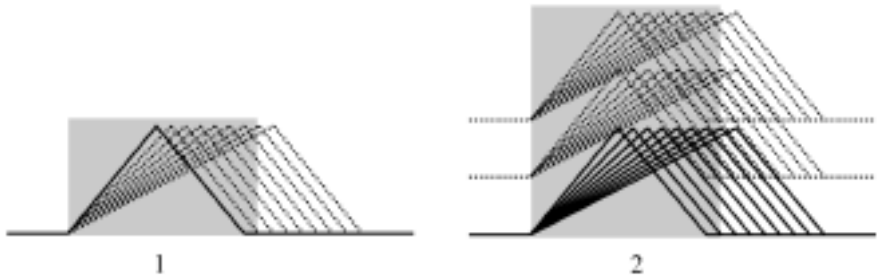

Fig. 6. Schematic representation of the acoustic manipulations (the grey area represents the nuclear syllable).

The mean of the answers given by the 10 Pisa Italian subjects who took part in the experiment are shown in figure 7, where the mean of all the subjects' answers in favour of the contrastive interpretation are plotted, calculated on the mean of all the five answers given by subjects to each stimulus (expressed by a value between 0 and 1, equal to the number of positive answers divided by five). The data are shown in relation to the steps of alignment variation and divided by the scaling differences ('PR0' corresponds to the scaling characteristics of the contrastive interpretation, 'PR2' to those found in the arrival [L+]H* L- pattern, and 'PR1' to an intermediate scaling). The results

show that alignment is significant in relation to the subjects' answers, and allows subjects to identify the two pitch accent, by differentiating the utterance functions and characteristics. Moreover, the data also show that scaling information interacts with alignment, modifying the reference values for the two pitch accent patterns, i.e. affects 'when' the two accents begin to be identified. As may be seen in the plot, subjects identify two accents although the graphs are not the typical s-shaped ones, usually shown in case of categorical perception. This is assumed to be due to the fact that the answers plotted in the graphs relate to stimuli in which only one acoustic parameter was manipulated at a time, say alignment, while the other parameter, say scaling, was kept constant. An s-shaped graph really showing a categorical shift would probably be more easily obtained for answers to stimuli representing a gradual shift from all the characteristics of an accent to all those of the other accent. In contrast, here only one parameter is manipulated at a time. This procedure not only allows one to create a continuum of variation, but also generates a continuum of variation plus a systematically ambiguous correlate. Figure 7 shows that the best s-shaped line is obtained for stimuli where the scaling information was intermediate between the two patterns – condition 'PR1'. The 'neutral' value for this correlate may have favoured a greater attention to alignment differences. However, the graph also shows that subjects were clearly identifying the stimuli having coherent correlates, interpreting those stimuli as either contrastive or not according to the (coherent) alignment and scaling characteristics. Indeed, the mean of answers in favour of the contrastive interpretation for stimuli having both alignment and tone scaling coherent with the contrastive pitch accent is about 0.85 (see values for step of align manipulation equal '0' and '1', in condition 'PR0'), while the mean of answers for stimuli having both alignment and scaling coherent with the other pitch accent is about 0.15 (see values for step of align manipulation equal '7' and '8', in condition 'PR2').

Let's now turn to a detailed discussion of the results. Stimuli realized with the target scaling of the contrastive accent (condition 'PR0') tend to be identified as non-contrastive, i.e. as realized as [L+]H* L-, when the peak is aligned later in the syllable (around 95% of the syllable, considering the characteristics of the first stimulus whose score is lower than 0.5, the chance level – see stimulus number 6). On the other hand, stimuli realized with the scaling of the broad interpretation (condition 'PR2') and the alignment characteristics of the contrastive interpretation obtain a score that reaches only chance level. Looking for the step after which speakers beging to identify the

broad pattern, it is possible to find it at the second alignment step, i.e. the stimuli begin to be perceived as corresponding to the broad pattern, [L+]H* L-, when the peak is early in the syllable (around 66% of the syllable – see stimulus number 2). On the other hand, the intermediate condition with regard to the scaling of targets (condition 'PR1') appears to represent a sort of neutralization of the scaling factor, perhaps forcing the subjects to focus more carefully on the alignment characteristics: the subjects begin to identify the [L+]H* accent when the peak is aligned early in the syllable (again, around the 66%-73% of the syllable – see stimuli number 2 and 3), and the responses appear to be more s-shaped in comparison to those given in condition 'PR0', showing a more abrupt increase in answers in favour of the broad pitch accent.[17]

Moreover, the experiment shows that different subjects exploit different cues to identify the accents. The majority of subjects appear to focus more on alignment characteristics, but few others (2 out of 10 subjects) appear to pay greater attention to target scaling variations – see figure 8. Subject 'FA' – left panel – appears to favour the broad focus interpretation according to the alignment characteristics, whatever the target scaling is, while subject 'AR' – right panel – distinguishes the contrastive and broad interpretation on the basis of the target scaling, almost irrespective of target alignment.

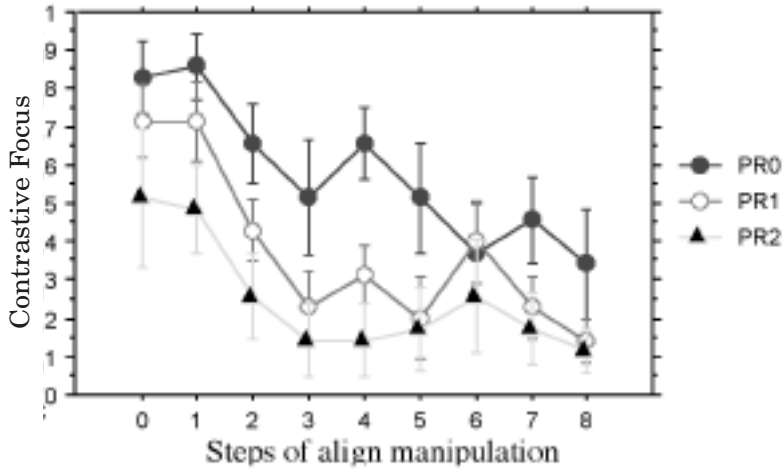

Fig. 7. Mean and standard error of positive answers in favour of the contrastive interpretation in relation to stimuli manipulated as for the peak (and following low) alignment and as for the target scaling.

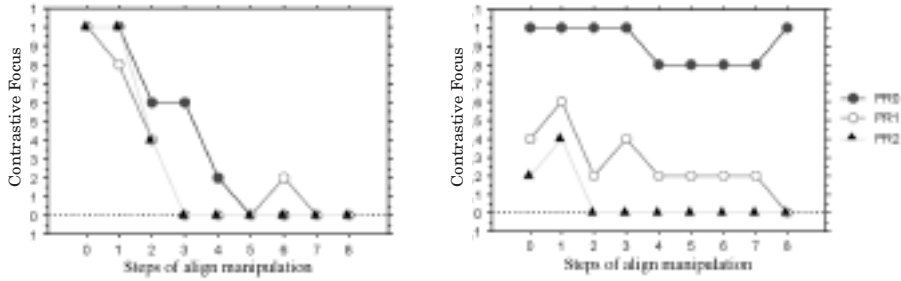

Fig. 8. Mean of answers in relation to stimuli manipulated as for the peak (and following low) alignment, divided by steps of target scaling manipulations: subject FA (left panel) and subject AR (right panel).

The global results show that the scaling of targets highly influences the alignment characteristics that appear to be interpreted as important for identifying two different accents.[18] As shown above, the scaling characteristics of the broad pattern cause troubles in the identification of the contrastive one, even when the alignment characteristics give quite unambiguous information – the score is attested at about the chance level; conversely, with the same scaling characteristics, the broad pattern begins to be identified as such when the alignment has been slightly modified from the base stimulus, i.e. the stimuli starts to be perceived as corresponding to the broad pattern when the peak is early in the syllable, in comparison to its alignment in [L+]H* accent on average. On the other hand, targets scaled as for the contrastive accent induce problems in giving a contrastive interpretation for intermediate alignment steps, and cause a lower score in the perception of the broad focus accent even in stimuli whose alignment characteristics correspond to a broad interpretation. Finally, the lack of manipulation of the low leading tone target did not appear to cause trouble in the interpretation of the patterns.[19]

Therefore pitch height appears to contributing to the identification of these pitch accents. At the moment, it is not possible to exclude the possibility that the role of scaling is (mainly) related to the paralinguistic information used by subjects in performing the task, i.e. it could be important for the perception of peremptory and conclusive patterns. The lower scaling of the high target could be associated with more peremptory and conclusive patterns, in line with predictions deriving from Gussenhoven's (2002) biological codes (the frequency and

production code variation may be interpreted as implying, respectively, assertiveness and finality). However, production data show that the two accents are functionally distinct, although in some specific contexts they may play partially overlapping roles (narrow focus and contrestive one); moreover, in perception, subjects were clearly interpreting the stimuli as either contrastive or not in case they had both alignment and scaling characteristics coherent with one pitch accent. Thus the two patterns are considered phonologically distinct.

Thinking of a phonological transcription of such patterns, and having in mind a phonetically transparent transcription that gives priority to information that distinguish patterns, the different targets scaling rather than the presence of the same low leading tone target should be considered as important, in that makes the specific characteristics of the two accents clear. Looking at properties that distinguish the accents under investigation, in facts, the scaling characteristics rather than the presence of the raise onset appear to be more salient. However, both characteristics could be left unspecified in a phonological transcription, as they may be deduced from the other elements that characterize the phonological unit, i.e. H* vs. H*+L. The low target may be accounted for by the presence of a H*, by stating that any peak in Pisa Italian is realized by means of a raise whose onset may be aligned around the syllable left boundary; the peak height could be related to the monotonal vs. bitonal status of the accents, with a lower value for the bitonal accent involving a low trailing tone target.[20] These characteristics can be unspecified within a phonological transcription, avoiding redundancy, while they may be coded in a phonetic transcription.

Moreover the need of differentiating phonological and phonetic information in the analysis of intonation has already been reported in the literature. In her analysis of Korean intonation, Jun (2005:210) proposes to label tonal events both on a phonological tier and on a phonetic tier. Tonal events reported on the latter "are categorical (i.e., H and L) and their distributions are limited". The proposal made here is highly similar in that only systematic and significant events are suggested to be labelled within a phonetic transcription. Nevertheless, here it is also proposed to explicitly differentiate phonetic and phonological information within the same transcription, with no need for a different transcription tier. For the two accent considered in this study, a phonological analysis may correspond to H* and H*+L, while phonetic additional information may be explicitly encoded by means of parenthesis – e.g. [L+] in case of the onset of the rise – or arrows – e.g. H*↓ +L in case of tone height.

6. Discussion and conclusion

This paper focuses on the need to tease apart information that is to be considered part of a transparent phonological transcription of pitch accents and properties that may be coded as phonetic. The discussion stemmed from both production and perception data on the variety of Italian spoken in Pisa, with particular attention to the interplay of alignment and scaling information in the identification of pitch accents.

Within autosegmental-metrical analyses, especially ToBi-Like ones, the labels that identify patterns stand for phonological units of the system, but, although to varying extents, they also encode the main phonetic characteristics of pitch accents. It is then generally the case that phonological units within intonational systems are represented by distinct and, simultaneously, quite phonetically-detailed labels. In this work it was argued that, although the properties that distinguish contrastive patterns may be transparently coded in the labels and stem from phonetic characteristics of the pattern, the phonetic characteristics encoded at the phonological level should be as little redundant as possible. The priority is then given to labels that distinguish contrastive patterns within the system and, as far as possible, are not redundant.

The core of the argument was the observation of the difference in tone scaling between two patterns found in Pisa Italian, both showing a rise to peak. The scaling characteristics, phonetically grounded and significant in both producing the two pitch accents and allowing subjects to identify them, could then be considered important in underscoring the different properties of two patterns. On the contrary, both patterns are produced with a low tonal target preceding the peak and are identified by subjects even though this target is kept unvaried in alignment. Thus, to a certain extent, in assigning 'distinct' and 'phonetically-detailed' labels to the two accents, one might be tempted to give priority to the significant difference in scaling rather than to the presence of a low leading tone target. Moreover, the low leading tone target, at least in the one accent investigated in prenuclear position, may actually show a high degree of variability when its position in the sentence is changed.

As the data suggest, coding all the systematic and significant phonetic properties at the phonological level would lead to a highly complex and redundant system, in which detailed information on scaling would also have to be included. The solution proposed in the paper aims both at avoiding such heavy phonological representation and at keeping track of the phonetically-detailed information. In

order to get distinct phonological labels which are also, simultaneously, phonetically-detailed it may be important to leave unspecified in the phonological transcription all the information that may be deduced from other elements coded in the phonological unit. In the case under investigation, the low leading tone target and the information on peak scaling may be left unspecified in a phonological transcription. They may be deduced from the other elements that characterize the phonological unit, such as, respectively, the presence of H* and the monotonal vs. bitonal status of the accents.

The usefulness of a clear distinction between more abstract phonological representations on the one hand, and phonetic representations, on the other, was also emphasized by the observation that subjects perceive a large number of cues, and that they may differ in the cues they pay attention to – in this case, both alignment and scaling. This provides an additional argument for considering only information that distinguishes patterns as represented within the phonological system, avoiding both the coding of all the correlates and the coding of only those usually considered more important than others. The representation of the 'distinct' and 'not redundant' properties of the patterns is then preferred to the systematic coding of characteristics related to one – although important – correlate of pitch accent, i.e. the stable alignment of the low tone rather than the peculiar scaling differences of the targets. In fact, the actual phonetic shape of a given a phonological unit depends on the effect of various factors, and, despite its variability, may be identified thanks to the redundancy of speech, in this case to the interaction of various correlates.

For the two accents considered in the paper, a phonological analysis may correspond to H* and H*+L, where H* are implemented as peaks, and have to be preceded by a raise. One further step could be taken to make phonological analysis as phonetically transparent as possible. According to Prieto *et al.*'s proposal (submitted), secondary association of the starred tone may account for the contrast in alignment of the peak in H* and H*+L, avoiding the ambiguity that potentially relates to the shared H* tone. In principle, the presence of H* in both patterns could be taken to imply pitch accent identity with regard to the high tone alignment – i.e. the high tone could be expected by the end of the syllable in both cases, and the trailing tone of H*+L could be expected after the syllable boundary. However, the peak is aligned significantly earlier in the H*+L pitch accent, and the trailing tone follows it quite closely. Prieto and colleagues suggest coding the contrastive alignment through

high tone secondary association (for discussion, see Prieto *et al.* (2005)). Therefore, in Pisa Italian, the tone is associated with the syllable right boundary in the H* pattern (H*]$_{\sigma}$, where ']$_{\sigma}$' represents the syllable boundary) and to the first mora in the H*+L pitch accent (H*]$_{\mu}$+L, where ']$_{\mu}$' represents the mora boundary). Thus a transparent phonological coding of the accents considered in this paper corresponds to the following: H*]$_{\sigma}$ vs. H*]$_{\mu}$+L. The transparency may be seen clearly for the accents sharing the H* tone: it may be stated that H*s in Pisa Italian correspond to a high peak realized by the end of a raise, starting around the beginning of the syllable; the different alignment of the peak is accounted for by means of secondary association, and this also makes it possible to obtain the slope of the raise; the shape after the peak is then represented by the absence/presence of the trailing tone, realized at a certain distance from the peak itself. However, all the information may be explicitly coded and differentiated within a narrow – or a broad – phonetic transcription, where everything but 'L', 'H', '+', '*', '-', '%', and possible codes of secondary association, such as ']$_{\mu}$', or ']$_{\omega}$', represents purely phonetic properties. This means that '[L]' or '[H]' may correspond to tonal targets with no phonological role, and '↓' or '↑' may code the systematic characteristics in terms of scaling.

Address of the Author:

Barbara Gili Fivela, Dipartimento di Lingue e Letterature Straniere, Università degli Studi di Lecce, Via Taranto 35, 73100 Lecce – Italy <barbara.gili@ateneo.unile.it>

Notes

* I am really grateful to Bob Ladd and Pauline Welby for their comments, which have helped to greatly improve the clarity and quality of the paper. A special thank goes to Cinzia Avesani for precious observations on an early version of this work. All remaining errors are of course mine.

1 To make another example, some stability in the scaling characteristics is also assumed in Marotta's (2005) discussion on the relevance of scaling variation as a sociolinguistic marker for the varieties of Italian. Marotta argues that the information on pitch range, in fact, should be a structural part of the intonational model.

2 For instance, /e/ denotes a unit produced with certain, say, formant values, but nothing in the symbol codes what the values are.

3 Throughout the paper, reference will be made to the importance usually pay to having distinct labels for phonological elements in intonational systems. Thus, in

some cases, 'distinctive' may also be used in referring to labels denoting patterns having a distinctive function.

[4] Dialogues recorded with the Map-Task correspond to verbal exchanges between two speakers who are given a map. One speaker map has a path among icons drawn on it, and the speaker owning it – the instruction giver – gives the interlocutor – the instruction follower – instructions to reconstruct the path on the other map, which only shows icon names (Anderson *et al.* 1991).

[5] Gili Fivela & Savino (2003) compare the characteristics of the H*+L accent produced by speaker of Pisa and Bari Italian, and show that, apart from alignment variations, the three targets are stably realized by the speakers of both varieties.

[6] In line with Searle & Vanderveken's (1985) definition, with illocutionary force is meant the combination of illocutionary point, i.e. the purpose of the speaker (e.g., to assert something or to attempt to get someone to do something), and the presupposition and attitude that accompany that point, including the strength of the illocutionary point. In what follows, 'illocutionary force' will be used to indicate the 'strength of speaker's purpose' (mainly speaker's attitude), the two expressions will be considered as interchangeable, and the illocutionary force will be considered as possibly related to emphasis.

[7] Within the Map-Task coding, these moves – realized with disbelieving intonation – are identified as object moves (Grice & Savino 1997).

[8] Taking into account a positional criterion for identifying the nuclear accent, in the two utterances under investigation both the accents precede a phrase accent and are nuclear (Beckman & Pierrehumbert 1986). Considering auditory cues – i.e. prominence – or pragmatic criteria over the whole utterance, the two pitch accents may be differently characterized: while the pitch accent exploited for contrast is the most prominent event and is associated to the most important information, the other accent may also be less prominent than the utterance final one, and the information that it signals may not be the most important in the sentence. In the latter case, a detailed investigation of speaker judgements should be taken into account in order to decide which is/are the most prominent/s pitch accent/s in the utterance, and, finally, which accent may be considered nuclear on the basis of auditory cues.

[9] The schema was proposed in Gili Fivela (2002) on the basis of the mean latency measured. Nevertheless, due to either the method followed in order to get the peak measurement when a sort of small plateau rather than a real peak was realized or to inter-speaker variation, the peak in [L]+H* turned out to be aligned, on average, after the syllable boundary. Nevertheless, the peak is often aligned within the syllable boundary, as the F0 track in figure 1, panel left, shows – for discussion see Gili Fivela (2004).

[10] A recent proposal argued for in Prieto *et al.* (2005) is to analyse phonologically the two accents as also having a different starred tone secondary association. The proposal aims at finding more transparent phonological codings that account for the presence of contrasts in alignment within the same language, as in the case under discussion. In order to account for the earlier alignment of the peak in the H*+L accent with respect to the H* one, Prieto *et al.* suggest that the high tone is secondarily associated with the right syllable boundary in the case of H* and with the first mora in case of H*+L (see Prieto *et al.* (2005) for discussion on the bimoraic analysis of the syllable bearing the contrastive accent, and for arguments supporting the hypothesis of the peak secondary association rather than other hypotheses, such as the high tone repulsion in H*+L, due to the presence of the trailing tone).

[11] The raising was actually observed in the Pisa Italian production of the pitch

accent investigated. A more global characterization of the pitch range over the utterance, in terms of raising or expanding, is beyond the goal of this paper.

[12] This result resembles the one found by Ladd & Johnson (1987) for sentence initial accents in English.

[13] As suggested by one reviewer, the slightly later alignment of the low target in case the accent is the second one of the phrase rather than the first and only one could be due to tonal crowding, and tonal crowding effects are rarely observed on low tones. Nevertheless, if tonal crowing was the reason for the later alignment, one could expect a significant effect of the number of preaccentual syllables on the low target alignment. This is not the case. Alternatively, the later alignment could also be related to the fact that the following raise to peak has a smaller F0 excursion.

[14] The measurements relevant for this study are latencies between tonal targets, latencies between tonal targets and syllable boundaries, and F0 height of targets.

[15] The values relate to productions by the same speaker who uttered the utterance manipulated in order to get the stimuli for the perception experiment.

[16] The alignment of both the peak and the following target was manipulated (steps of 15 ms). In fact, a falling phase may be produced also after the [L+]H* accent: realizing the manipulation on a word with antepenultimate stress position, the low target following the peak ends up being aligned by the end of the final syllable. This is coherent with the characteristic found in the [L+]H* L- pattern, produced on words with antepenultimate stress position. As for the scaling characteristics, they were varied differently for each target: steps of 13 Hz for the first low elbow, 17 Hz for the peak, and 6 Hz for the second low target.

[17] The perception test was also carried out with speakers of a different variety of Italian (the one spoken in Turin, northern Italy). Results will not be discussed here, but they are highly comparable to those obtained by Pisa Italian speakers. The only difference relates to the judgements in the case of intermediate target scaling. Turin subjects, in comparison to Pisa Italian ones, identify the [L+]H* accent in the case of a later alignment of the peak within the syllable.

[18] Nevertheless, by changing the peak alignment, the slope of the raising phase was also varied and the slope may play a role in the pattern perception.

[19] In this experiment, the choice was to keep the alignment of the low leading tone target unvaried. However, the check of the perceptual relevance of such a target would be problematic because of the lack, in Pisa Italian, of two patterns mainly differing for the presence/absence of the low leading tone. One possibility, investigated in Gili Fivela (*to appear*), would be to exploit the comparison between the contrastive [L+]H*+L accent and the broad focus H+L* pattern, where no [L+] target is observed. Nevertheless, such contrast may be difficult to control in case of manipulated stimuli because of the simultaneous variation of peak alignment, slope of the fall and pattern scaling.

[20] Moreover, the information on the slope of the raising phase may be deduced by the peak alignment specification - see the secondary association proposal in Prieto *et al.* (2005) – although implementation rules for the low target point should, of course, account for the observed inter-accent differences and for the systematic changes depending on the pitch accent role and position within the utterance.

Bibliografical references

ANDERSON Anne H., Miles BADER, Ellen G. BARD, Elizabeth BOYLE, Gwyneth M. DOHERTY, Simon C. GARROD, Stephen D. ISARD, Jacqueline C. KOWTKO, Jan MCALLISTER, Jim MILLER, Catherine SOTILLO, Henry S. THOMPSON & Regina WEINERT 1991. The HCRC map task corpus. *Language and Speech* 34, 4. 351-366.

ARVANITI Amalia & D. Robert LADD 1995. Tonal alignment and the representation of accentual targets. In *Proceedings of the 13th International Congress of Phonetic Sciences* (ICPhS'95). Stockholm. Sweden. August 13-19. 220-223.

ARVANITI Amalia, D. Robert LADD & Ineke MENNEN 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26. 3-25.

ARVANITI Amalia, D. Robert LADD & Ineke MENNEN 2000. What is a starred tone: Evidence from Greek. In BROE Michael & Janet B. PIERREHUMBERT (eds.). 119-131.

BARTELS Christine & John KINGSTON 1994. Salient pitch cues in the perception of contrastive focus. In BOSCH Peter & Rob VAN DER SANDT (eds.) 1994. *Focus and Natural Language Processing*. IBM Deutschland Informationssysteme GmbH Scientific Centre, Institute for Logic and Linguistics. Working Papers of the Institute for Logic and Linguistics. Vol. 1: Intonation and Syntax. 1-10.

BECKMAN Mary E. & Gayle AYERS 1997. *Guidelines for ToBI Labelling*. The Ohio State University.

BECKMAN Mary E. & Janet B. PIERREHUMBERT 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255-309.

BESANA Sveva 1999. *Towards an Analysis of Turinese Italian Intonation and Theoretical Implications for Intonational Phonology*. PhD thesis. MIT.

BROE Michael B. & Janet B. PIERREHUMBERT 2000. *Papers in Laboratory Phonology V*. Cambridge: Cambridge University Press.

BRUCE Gösta 1977. *Swedish Word Accents in Sentence Perspective*. CWK: Gleerup.

CRUTTENDEN Alan 1986. *Intonation*. Cambridge Textbooks in Linguistics: Cambridge University Press.

D'IMPERIO Mariapaola 1999. Tonal structure and pitch targets in Italian focus constituents. In *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS'99). San Francisco. USA. Vol. 3. 1757-1760.

D'IMPERIO Mariapaola 2000. *The Role of Perception in Defining Tonal Targets and their Alignment*. USA. Ohio State University. PhD thesis.

FROTA Sónia 2000. *Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation*. New York: Garland.

FROTA Sonia 2003. The phonological status of initial peaks in European Portuguese. *Catalan Journal of Linguistic*s, 2 (Special volume on Romance Intonation, ed. by P. Prieto). 133-152.

GILI FIVELA Barbara 1999. The prosody of left-dislocated topicalized constituents in Italian read speech. In *Proceedings of the European*

Conference on Speech Communication and Technology (EuroSpeech'99). Budapest. Hungary. Vol. 1. 531-534.

GILI FIVELA Barbara 2002. Tonal alignment in two Pisa Italian peak accents. In *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence. France. April 11-13. 339-342.

GILI FIVELA Barbara 2004. *The phonetics and phonology of intonation: the case of Pisa Italian*. Tesi di Perfezionamento: Scuola Normale Superiore, Pisa.

GILI FIVELA Barbara 2005. La percezione degli accenti: il ruolo dell'allineamento e dello 'scaling' dei bersagli tonali. In *Proceedings of Convegno Nazionale AISV (Associazione Italiana di Scienze della Voce) "Misura dei parametri"*. Padova. Italy. December 2-4. 313-326.

GILI FIVELA Barbara 2005. 'Scaling' e allineamento dei bersagli tonali: la percezione di due accenti discendenti. In *Proceedings of Convegno Nazionale AISV (Associazione Italiana di Scienze della Voce) "Analisi prosodica: teorie, modelli e sistemi di annotazione"*. Salerno. Italy. November 30, December 1-2.

GILI FIVELA Barbara & Mariapaola D'IMPERIO 2004. Manuscript. Tonal alignment of prenuclear accent in Italian: A study on the varieties spoken in Pisa and Naples. In *EFS Conference on Tone and Intonation in Europe*. Santorini. Greece. September 8-12.

GILI FIVELA Barbara & Michelina SAVINO 2003. Segments, syllables and tonal alignment: A study on two varieties of Italian. In *Proceedings of the 15th International Congress of Phonetic Sciences* (ICPhS'03). Barcelona. Spain. August 4-9. 2933-2936.

GRICE Martine 1995. *The Intonation of Interrogation in Palermo Italian. Implications for Intonation Theory*. Tübingen: Max Niemeyer Verlag.

GRICE Martine & Michelina SAVINO 1997. Can pitch accent type convey information status in yes-no questions? In *Proceedings of Concept to Speech Generation Systems*. 29-37.

GUSSENHOVEN Carlos 2002. Intonation and interpretation: Phonetics and phonology. In *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence. France. April 11-13. 47-57.

GUSSENHOVEN Carlos & Toni RIETVELD 2000. The behaviour of H* and L* under variation in pitch range in Dutch raising contours. *Language and Speech* 43, 2. 183-203.

JUN Sun-Ah 2005. Korean intonational phonology and prosodic transcription. In JUN 2005. 201-229.

JUN Sun-Ah 2005. *Prosodic Typology. The Phonology of Intonation and Phrasing*. New York: Oxford University Press.

LADD D. Robert 1996. *Intonational Phonology*. Cambridge Studies in Linguistics. Cambridge: Cambrige University Press.

LADD D. Robert & Catherine JOHNSON 1987. 'Metrical' factors in the scaling of sentence-initial accent peaks. *Phonetica* 44. 238-245.

LADD D. Robert, Dan FAULKNER, Hanneke FAULKNER & Astrid SCHEPMAN 1999. Constant 'segmental anchoring' of F_0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106, 3. 1543-1554.

LADD D. Robert & Astrid SCHEPMAN 2003. 'Sagging transitions' between high pitch accents in English: Experimental evidence. *Journal of Phonetics* 31. 81-112.

MAROTTA Giovanna 2000. Allineamento e trascrizione dei toni accentuali complessi: una proposta. In *Proceedings of X Giornate di Studio del Gruppo di Fonetica Sperimentale*. Napoli. Italy. December 13-15, 1999. 139-149.

MAROTTA Giovanna 2002. L'intonation des énoncés interrogatifs ouverts dans l'italien toscan. In *Proceedings of the Speech Prosody 2002 Conference*. Aix-en-Provence. France. April 11-13. 475-478.

MAROTTA Giovanna 2005. L'illusione prosodica. In Giovanna MAROTTA (ed.) 2005. *Atti del convegno di studi in memoria di Tristano Bolelli*. Pisa. Italy. Novembre 28-29, 2003. *Studi e saggi linguistici*. XL-XLI. 237-258.

PIERREHUMBERT Janet B. 1980. *The phonology and phonetics of English intonation*. PhD thesis. MIT, published in 1988 by IULC.

PRIETO Pilar, Mariapaola D'IMPERIO, Barbara GILI FIVELA 2005. Pitch accent alignment in Romance: primary and secondary association with metrical structure. *Language and Speech,* Special Issue: Intonation in Language varieties, guest editor Paul Warren, 48. 359-396.

'T HART Johan, René COLLIER & Antoine COHEN 1990. *A perceptual study of intonation. An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

SAVINO Michelina, Martine GRICE, Barbara GILI FIVELA & Giovanna MAROTTA 2006. Intonational cues to discourse structure in Bari and Pisa Italian: Perceptual evidence. In *Speech Prosody 2006*. Dresden. Germany. May 2-5. 114-117.

SEARLE John R. & Daniel VANDERVEKEN 1985. *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

WILTSHIRE Caroline & Joaquim CAMPS 2002. *Romance phonology and variation*. Selected papers from the 30th Linguistic Symposium on Romance Languages. Amsterdam: John Benjamins.

Finito di stampare nel mese di Maggio 2007
presso le Industrie Grafiche della Pacini Editore S.p.A.
Via A. Gherardesca • 56121 Ospedaletto • Pisa
Telefono 050 313011 • Telefax 050 3130300
Internet: http://www.pacinieditore.it