# Intonational cues to discourse structure in Bari and Pisa Italian: perceptual evidence

*Michelina Savino[a], Martine Grice[b], Barbara Gili Fivela[c] and Giovanna Marotta[d]*

[a]Univ. of Bari (Italy), [b]Univ. of Cologne (Germany), [c]Univ. of Lecce (Italy), [d]Univ. of Pisa (Italy)
esavino@poliba.it, martine.grice@uni-koeln.de, b.gilifivela@sns.it,gmarotta@ling.unipi.it

## Abstract

Perception experiments for Bari and Pisa Italian showed that listeners can reliably distinguish between final and non-final utterances in discourse by means of intonation. Bari listeners were also able to distinguish a third category, signalling that the end of the discourse unit is approaching (penultimate position). This was not the case for Pisa listeners, although there was possibly an effect of speaking style.

## 1. Introduction

It is widely agreed that intonation plays a very important role in signalling discourse structure. A clear example is represented by the continuation rise, which "conveys a subordination relationship among phrases within a discourse segment" [1]. Moreover, closer examination of this aspect of prosody has shown that in some languages intonation can also be used to signal pre-finality, i.e. that the end of a discourse unit is *approaching.* This is realised by marking the penultimate item of a sequence with an intonation contour that is different from that in other non-final ones. Such a "pre-finality marking" has been attested in (peninsular) Spanish, Romanian, American English [2], and in Dutch [3].

For Italian, preparatory work on Bari and Pisa varieties based on both task-oriented and read speech materials (see refs in sections 2 and 3) revealed that Bari Italian appears to use specific tonal sequences for distinguishing between non-final and penultimate position, whereas Pisa Italian does not. This paper presents a summary of those aspects of Bari and Pisa intonation involved in discourse structuring, and describes a perception experiment carried out for each variety which largely confirmed our initial observations.

## 2. Non-finality and pre-finality in Bari Italian

Intonation contours conveying information about the hierarchical structuring of discourse have been described for Bari Italian, based on Map Task dialogues and monologues (INSTRUCT and EXPLAIN moves, [4]), and list readings [5], [6]. The most common non-final contour encountered in Map Task dialogues is the continuation rise, signalling that "more is to come" and consisting of a gradual rising movement from the nuclear syllable to the end of the phrase, phonologically analysed as L*L-H%. In addition, a high rise melody (analysed as H*H-H%) was found to convey a different, more restricted meaning, namely that "one more is to come" (pre-finality).

Further observations in Map Task monologues (where the Instruction Givers could convey the instruction sequences without interruption), and in lists of landmark names read aloud, confirmed the use of different contours for non-finality and pre-finality, although the inventory of intonation patterns extended beyond the L*L-H% low rise and H*H-H% high rise. For the non-final contours, two other patterns were attested, both with a peak on the accented syllable, one followed by a fall, (H*L-L%), and one followed by a plateau (H*H-L%). In pre-final position, the contours produced in the monologues have an early peak followed by a low level pitch which stays low or rises towards the end of the phrase (H+L*L-L% and H+L*L-H%, respectively). Pre-final contours are immediately followed by a phrase with a final contour, which, in declaratives, is typically !H+L*L-L%. Table 1 summarises the inventory of Bari Italian intonational sequences (adapted from [6]).

| | NON-LOW BOUNDARY | LOW BOUNDARY |
|---|---|---|
| NON-FINAL | L* L-H%<br>H* H-L% | H* L-L% |
| PRE-FINAL | H* H-H%<br>H+L* L-H% | H+L*L-L% |
| FINAL | | !H+L*L-L% |

Table 1: *Non-final, pre-final and final tonal sequences in Bari Italian*

## 3. Non-finality and pre-finality in Pisa Italian

Background work on Pisa Italian intonation strategies for signalling non-finality and pre-finality stems from analysis of both spontaneous (Map Task dialogues) and read speech, as part of a more comprehensive description of the intonation system of this variety [7]. Unlike what has been described for Bari Italian, it appears that in Pisa Italian no relation can be found between the pattern in continuation contours and its pre-final or non-final position in sequences of instructions. INSTRUCT moves are intonationally realised in Pisa Italian by means of various pitch patterns, such as a falling nuclear pitch accent followed either by low (H+L*L-L%) or high pitch (H+L*H-H%) or a rising pitch accent followed by either high or, in some cases, low pitch (H*H-H% or H*L-L%).

The H+L*L-L% contour is usually found in final position within a discourse segment (note that it corresponds to the Bari Italian !H+L*L-L%), but may also be found earlier in a sequence. Contours characterised by a final rise, analysed as H+L*H-H%, or by rising nuclear accents, analysed as H*H-H% or H*L-L%, are also found in INSTRUCT move utterances. These latter patterns are not specifically used for signalling pre-finality. They can be found in any non-final utterance.

A further inspection of Map Task monologues (and reading of lists of landmarks, as described for Bari Italian speakers above) has revealed that Pisa Italian speakers produced all the contours described above except H*L-L% (which was, however, not frequently found in the previously analysed corpora). A difference was also observed in the distribution of contours in relation to utterance position. While the H+L*L-L% pattern is produced in all positions, in pre-final position H*H-H% appears to be preferred over H+L*H-H%. The latter pattern is otherwise relatively rare. Therefore, in monologues, patterns still appear to be

interchangeable, apart for the preference of one pattern in prefinal position (H*H-H%). For a summary of Pisa Italian intonational categories, see Table 2.

| | NON-LOW BOUNDARY | LOW BOUNDARY |
|---|---|---|
| NON-FINAL AND PRE-FINAL | H+L*H-H%<br>H*H-H% | H+L*L-L%<br>H*L-L% |
| FINAL | | H+L*L-L% |

Table 2: *Non-final, pre-final and final tonal sequences in Pisa Italian*

## 4. Perceptual experiment

Background analyses on Bari and Pisa speech materials revealed that their intonational systems use different strategies to cue non-finality and pre-finality. This interpretation is based on both contextual analysis and native speaker intuitions. In order to test such interpretations, a perceptual experiment was carried out, requiring native speakers of each variety to judge the position of a number of utterances as final, pre-final or non-final.

### 4.1. Materials

For both varieties, utterances were extracted from Map Task monologues (with a silent partner) and from a list of landmarks read aloud after the monologue task.

#### *4.1.1. Bari Italian*

The stimuli consisted of five INSTRUCT moves containing non-final, prefinal and final utterances, with a total of 28 stimuli, 23 of which are giving instructions (for example "vai a destra...", *"go to the right..."*), the remaining 5 ones are the final mention of the landmarks drawn on the map, and therefore consist of one NP (i.e. the landmark name, for example "dimora per animali...", *"animals' house..."*). Classification of the stimuli as belonging to each of 3 pragmatic classes under examination was guided by considering both its position within the sequence and an analysis of the intonation contour.

It must be stressed that the main aim of the experiment was to verify perceptually whether Bari Italian has the means to differentiate intonationally between the three categories, finality, pre-finality and non-finality. Thus, inclusion of all possible contour types for pre-finality and non-finality was not considered a priority. Further, since it is well known that information on discourse structure can be conveyed not only prosodically but also by means of specific lexical markers [8], all lexical markers and discourse connectors like "poi" (*then*), "dopodichè" (*afterwards*), "e" (*and*) etc. were eliminated from the stimuli.

#### *4.1.2. Pisa Italian*

A similar procedure was adopted for creating the Pisa Italian test materials. A set of 28 stimuli was selected, consisting of non-final, pre-final and final utterances (in a sequence of instructions) from two speakers. Twenty stimuli were taken from Map Task monologues, 8 from lists of landmark names.

The main criterion for selecting stimuli was that the intonation contour was representative of the three major phonological sequences observed in the production data, namely the typical (final) pattern found in declaratives, i.e. H+L*L-L%, and the two realisations of continuation contours, i.e. H+L*H-H% and H*H-H%.

### 4.2. Experimental procedure

For Bari Italian, 23 informants (9 males, 14 females) took part in the experiment, aged between 19 and 38, all born and living in Bari. They were students and University staff, none had a background in linguistics or phonetics. Informants were presented 3 different randomised sequences of the 28 stimuli (plus an initial set of 12 stimuli as training) for a total of 84 stimuli. Stimuli were presented in groups of 12, preceded by two warning tones and followed by 10 seconds of silence. Each stimulus was preceded by one warning tone and followed by 4 seconds silence during which subjects made their judgement. Judgements involved crossing a box on a sheet indicating whether the utterance was the "last one uttered", the "penultimate one uttered" or the "intermediate one uttered". Before starting, informants were presented written instructions containing preliminary information on the situation in which stimuli were produced (the Map Task), about the classification requested (giving a definition of the three categories "final", "pre-final" and "non-final" in non-technical terms), and a description of the stimuli presentation. They were then familiarised with the type of material to be presented in the experiment, by hearing three excerpts from the Map Task monologues (two consisting of instructions sequences, and one sequence with objects mentioned in the listing task). Each session lasted approximately 20 minutes.

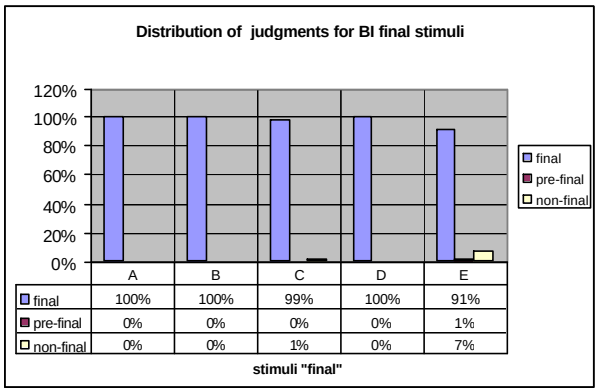
For Pisa Italian, 26 Pisan subjects (12 females and 14 males) took part in the perceptual test. They were all born and living in Pisa, and were students and University staff, aged from 21 to 37. None had any special competence in prosody or linguistics. The procedure for carrying out the experiment was the same as described for Bari Italian with one exception: since Pisa Italian materials were selected from two different Map Task monologues instead of one, stimuli were presented to the Pisan subjects in two consecutive sessions. In the first session informants were presented the 12 stimuli produced by one of the two speakers, and in the second session the remaining 12 stimuli uttered by the other. The excerpts in the training phase were different in the two sessions, so that the same speaker produced the following test stimuli.

### 4.3. Results

#### *4.3.1. Bari Italian*

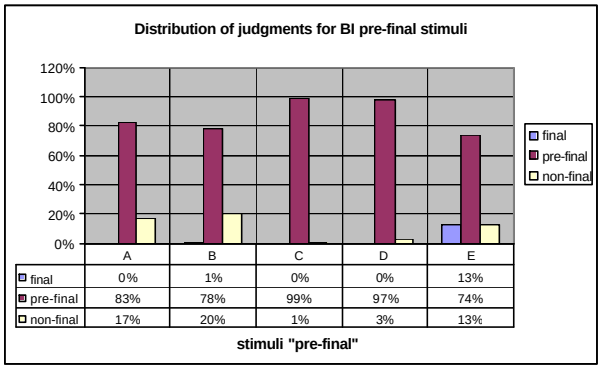
Figures 4a, 4b, 4c, and 4d show the distributions of judgments by all Bari Italian informants for stimuli related to the pragmatic classes "final", "pre-final", and "non-final" respectively. In all figures, letters referring to each stimulus also identify the sequence from which the stimulus was extracted. In particular, all stimuli identified by the letters A, B, C, D for each of the 3 pragmatic categories refer to one of the 4 sequences of giving-instructions utterances, whereas stimuli with letter E refer to the read lists. The larger amount of stimuli for the category "non-final" in each sequence is because each sequence can have just one final utterance and one pre-final utterance, whereas the number of non-final ones can vary, according to the speaker's discourse planning strategies. Overall test results confirm our hypothesis that Bari Italian speakers are able to recognise the position of a discourse segment by means of its intonation contour. As expected, the most significantly recognised is the !H+L*L-L% final contour with 100% agreement in almost all cases (Figure 4a). Second is the pragmatic category of pre-finality, with agreement ranging from 99% to 74% (Figure 4b). Also, there was no obvious preference for either a high or a low ending contour (H+L*L-L% and H+L*L-H%). Perceptual

evidence for the non-finality category can be clearly derived from the data, which show agreement ranging between 72% and 99% (Figures 4c and 4d). The lowest rate of agreement for non-finality has been assigned to stimulus B3, which is confused with pre-final utterances in the 41% of the cases. This may be due to the fact that it entails some disfluency.



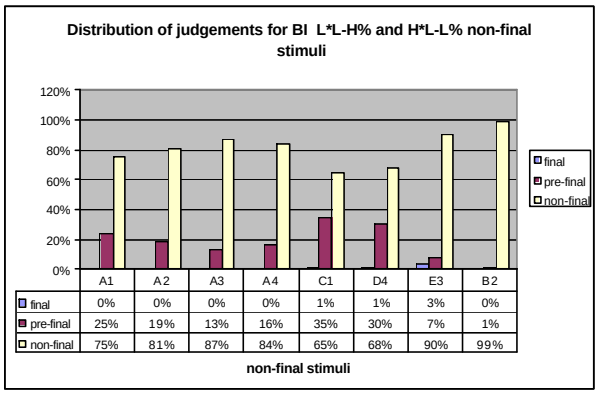

| | A | B | C | D | E |
|---|---|---|---|---|---|
| final | 100% | 100% | 99% | 100% | 91% |
| pre-final | 0% | 0% | 0% | 0% | 1% |
| non-final | 0% | 0% | 1% | 0% | 7% |

Figure 4a: *Distribution of judgements (in %) for Bari Italian !H+L*L-L% final stimuli*




| | A | B | C | D | E |
|---|---|---|---|---|---|
| final | 0% | 1% | 0% | 0% | 13% |
| pre-final | 83% | 78% | 99% | 97% | 74% |
| non-final | 17% | 20% | 1% | 3% | 13% |

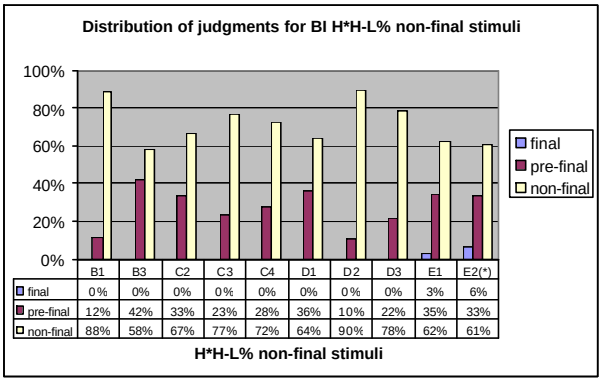
Figure 4b: *Distribution of judgements (in %) for Bari Italian pre-final stimuli. A, B, D are H+L*L-H% stimuli, and C, E are H+L*L-L% stimuli*




| | A1 | A2 | A3 | A4 | C1 | D4 | E3 | B2 |
|---|---|---|---|---|---|---|---|---|
| final | 0% | 0% | 0% | 0% | 1% | 1% | 3% | 0% |
| pre-final | 25% | 19% | 13% | 16% | 35% | 30% | 7% | 1% |
| non-final | 75% | 81% | 87% | 84% | 65% | 68% | 90% | 99% |

Figure 4c: *Distribution of judgements (in %) for Bari Italian non-final stimuli. A1, A2, A3, A4, C1, D4 are L*L-H% stimuli, B2 is the only H*L-L% stimulus.*

It appears that no one particular tonal sequence is preferred for either pre-finality or non-finality, as high response agreement was found for all three intonation patterns involved in the perception test. Note that the non-final stimulus E2 is ambiguous between H*H-L% and H*L-L%, since the nuclear syllable is phrase final, leading to possible truncation of the L% (see [9] for a discussion on truncation and compression).




| | B1 | B3 | C2 | C3 | C4 | D1 | D2 | D3 | E1 | E2(*) |
|---|---|---|---|---|---|---|---|---|---|---|
| final | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 3% | 6% |
| pre-final | 12% | 42% | 33% | 23% | 28% | 36% | 10% | 22% | 35% | 33% |
| non-final | 88% | 58% | 67% | 77% | 72% | 64% | 90% | 78% | 62% | 61% |

Figure 4d: *Distribution of judgements (in %) for all Bari Italian H*H-L% non-final stimuli. Asterisk on E2 signals uncertainty in intonation analysis (see section 4.3.1 for details)*

Table 2 provides confusion matrices over the whole Bari stimulus set.

| | *final* | *prefinal* | *non-final* |
|---|---|---|---|
| final stimuli | 98% | <1% | 2% |
| pre-final stim. | 3% | 86% | 11% |
| non-final stim. | <1% | 23% | 76% |

Table 2: *Confusion matrices for Bari Italian (numbers rounded to the closest integer value)*

### *4.3.2. Pisa Italian*

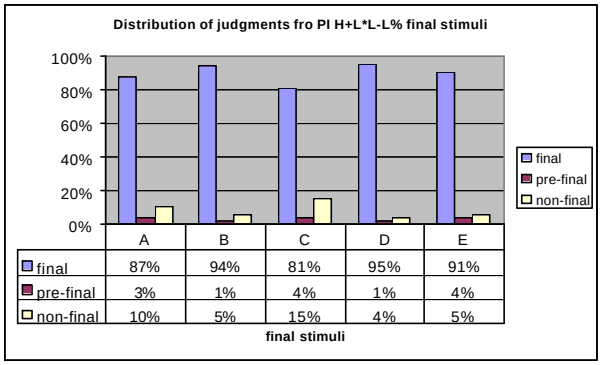
Figures 5a, 5b, and 5c show the distributions of judgments by Pisa Italian subjects for the classes “final”, “pre-final”, and “non-final” respectively. The judgements relate to 26 out of 28 stimuli, as 2 stimuli for the broad focus contour (H+L*L-L%) taken in the ‘non-final’ position set were discarded because their lexical composition might have given information on the possible interpretation. In each table, stimuli are grouped by contour type. Stimuli F,G,I and G,H,L in table 5b and 5c are taken from the read lists.

As Figure 5a shows, stimuli which were in final position are correctly recognised as final, even though agreement is not as high as in Bari Italian. No difference is observed in relation to the speaking style, that is, between stimuli taken from the read lists and those taken from instructions.

For pre-final position, results in Figure 5b show two different tendencies: for some stimuli (A, B, C, H) there is a fairly strong agreement that they should be classified as pre-final; for the remaining ones (D, F, G, I) they are mainly judged as non-final. The distribution of judgments across the two phonological categories makes it difficult to ascribe different interpretations to a specific intonation pattern. In fact, three H+L*H-H% and one H*H-H% are classified as pre-final, and three H*H-H% and one H+L*H-H% as non-final. However, if we look at speaking style, stimuli correctly classified as pre-final (A, B, C, H) are all taken from instruction-giving utterances, whereas those classified as non-final (F, G, I) are taken from read lists. It appears, then, that speaking style (giving instructions vs reading a list) may also play a role.

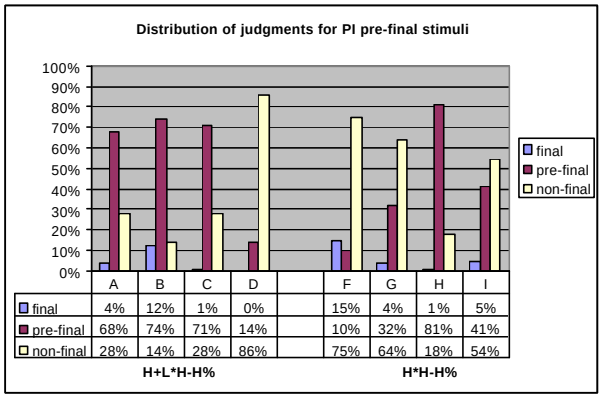
Figure 5c shows results for stimuli taken from non-final position. The highest agreement for the non-final classification is given for stimuli intonationally realised with the final-rising pattern (H+L*H-H%), ranging between 85%

and 62%. A slightly lower performance is observed for H*H-H% stimuli, where the percentage of agreement ranges from 78% to 65%. Further, there are two stimuli which are classified as non-final in 59% (I) and 47% (E) of cases, respectively. Looking at speaking style again, stimuli among those with greater agreement as non-final (G, H, L, in the H*H-H% set) are extracted from lists. However, both H+L*H-H% and H*H-H% stimuli taken from non-final position tend to be mainly recognised as non-final. Finally, non-final stimuli realised with a falling pattern (H+L*L-L%) are always classified as final, with agreement ranging from 85% to 81%. These results clearly show that the falling tonal pattern cannot be reliably associated with non-finality.



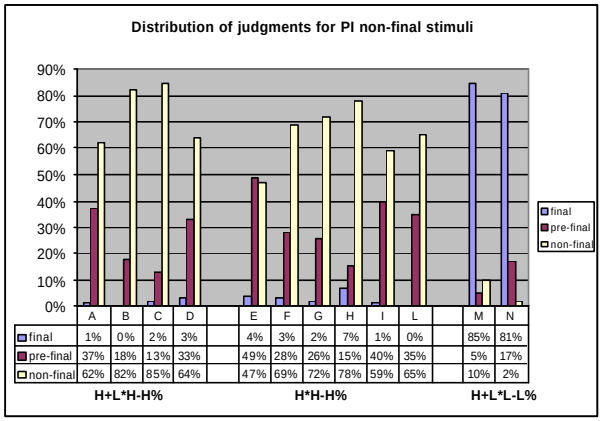

| | A | B | C | D | E |
|---|---|---|---|---|---|
| final | 87% | 94% | 81% | 95% | 91% |
| pre-final | 3% | 1% | 4% | 1% | 4% |
| non-final | 10% | 5% | 15% | 4% | 5% |

Figure 5a: *Distribution of judgements (in %) for Pisa Italian final stimuli. Stimuli A, D are taken from object name list readings*




| | A | B | C | D | F | G | H | I |
|---|---|---|---|---|---|---|---|---|
| final | 4% | 12% | 1% | 0% | 15% | 4% | 1% | 5% |
| pre-final | 68% | 74% | 71% | 14% | 10% | 32% | 81% | 41% |
| non-final | 28% | 14% | 28% | 86% | 75% | 64% | 18% | 54% |

Figure 5b: *Distribution of judgements (in %) for Pisa Italian pre-final stimuli, grouped by contour type. Stimuli F, G, I are taken from object name list readings*




| | A | B | C | D | E | F | G | H | I | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| final | 1% | 0% | 2% | 3% | 4% | 3% | 2% | 7% | 1% | 0% | 85% | 81% |
| pre-final | 37% | 18% | 13% | 33% | 49% | 28% | 26% | 15% | 40% | 35% | 5% | 17% |
| non-final | 62% | 82% | 85% | 64% | 47% | 69% | 72% | 78% | 59% | 65% | 10% | 2% |

Figure 5c: *Distribution of judgements (in %) for Pisa Italian non-final stimuli, grouped by contour type. Stimuli G,H, L are taken from object name list readings*

Thus, apart from the H+L*L-L% contours, speakers do not appear to rely on phonological cues to distinguish between non-final vs pre-final distinction, indicating that it may be phonetic cues which they are attending to. There is an indication of this in the stimuli: within the phrases taken from pre-final position, those showing higher peak values and F0 range values tend to be recognised as pre-final.

Table 3 shows confusion matrices over the whole Pisa stimulus set.

| | *final* | *pre-final* | *non-final* |
|---|---|---|---|
| final stimuli | 90% | 3% | 8% |
| pre-final stim. | 5% | 49% | 46% |
| non-final stim. | 16% | 26% | 58% |

Table 3: *Confusion matrices for Pisa Italian (numbers rounded to the closest integer value)*

## 5. Conclusions

For Bari Italian we have shown that listeners can rely on phonological cues in the intonation for distinguishing between three different positions within an utterance, final, penultimate and non-final. For Pisa listeners this does not appear to be the case. They did, however, reliably recognise finality, although they appeared to misinterpret some contours which were in fact non-final as final. Further analysis is needed to ascertain whether this was a function of planning strategies on the part of the speaker (not all non-final utterances have to be marked as such, since a speaker may change strategy mid-utterance). Further, more controlled experiments are needed to ascertain whether phonetic cues are used.

## 6. References


[1] Hirschberg, J.; Pierrehumbert, J., 1986. The intonational structuring of discourse. *Proc. of the 24th Annual Meeting of the ACL*, 136-144.

[2] Bolinger, D., 1989. *Intonation and Its Uses*. Palo Alto: Stanford University Press.

[3] Swerts, M.; Collier, R.; Terken, J., 1994. Prosodic predictors of discourse finality in spontaneous monologues. *Speech Communication* 15, 79-90.

[4] Carletta, J.; Isard, A.; Isard, S.; Kowtko, J.; Doherty-Sneddon, G.; Anderson, A., 1997. The reliability of a dialogue structure coding scheme. *Computational Linguistics* 23, 13-32.

[5] Savino, M. 2001. Non-finality and Pre-finality in Bari Italian Intonation: a Preliminary Account. *Proc. of the Eurospeech 2001*, vol.2, 939-942.

[6] Savino, M. 2004. Intonational Cues to Discourse Structure in a Variety of Italian. In *Regional Variation in Intonation*, P. Gilles & J. Peters (eds.) Tuebingen: Niemeyer, 145-159.

[7] Gili Fivela, B. 2004. *The phonetics and phonology of intonation: the case of Pisa Italian*, PhD dissertation, Scuola Normale Superiore, Pisa.

[8] Swerts, M. & Ostendorf, M. 1997. Prosodic and lexical indications of discourse structure in human-machine interaction. *Speech Communication* 22, 25-41.

[9] Grice, M.; D'Imperio, M.; Savino, M.; Avesani, C., 2005. Strategies for intonation labelling across varieties of Italian. In *Prosodic Typology: the Phonology of Intonation and Phrasing*, Sun-Ah Jun (ed.), New York: Oxford University Press, 362-389.