

A CONTRIBUTION TO THE ESTIMATION OF NATURALNESS IN THE INTONATION OF ITALIAN SPONTANEOUS SPEECH

Mario Refice¹, Michelina Savino^{1,2}, Martine Grice²

¹DEE - Politecnico di Bari, Italy

²Institut fuer Phonetik, University of the Saarland, Germany

E-mails: refice@poliba.it, esavino@poliba.it, mgrice@coli.uni-sb.de

ABSTRACT

This paper focusses on the intonation of yes-no questions in a local non-standard variety of Italian: that spoken in Bari. It has been claimed in any early study [7, 8] that Bari Italian (BI) has a final rise on yes-no questions. However, subsequent accounts of BI [5, 6] have found a predominance of final falls on such questions. Since the former study was based on a corpus read aloud and the latter on spontaneous dialogues, it was decided to compare read and spontaneous questions produced by the same speaker. Spontaneous questions by six speakers were extracted from recordings of task-oriented dialogues and presented for reading, both in list form and in specially constructed contexts. It was found that the final tonal contour was predominantly falling in the spontaneous questions and predominantly rising in the corresponding read questions. These results throw light on the discrepancy in the literature as to the typical yes-no question intonation for this variety of Italian. It is argued that the falling final contour is more natural than the rising one since it is typical of spontaneous speech.

1. INTRODUCTION

In Italian intonation plays a particularly important role, it being the only means for signalling the difference between statement and question, whereas in other languages, such as English, syntactic and/or lexical markers are also used.

In so-called "standard" Italian, information seeking yes-no questions have a final rising contour [1], [2], [3], whereas other regional varieties (such as the ones spoken in Palermo and in Bari) have a rising-falling contour starting from the nuclear syllable ([4] for Palermo Italian, [5] and [6] for Bari Italian). In particular, spontaneous yes-no questions in Bari Italian (henceforth BI) are characterised by the tonal sequence L+H* L-L% (as opposed to statements which have H+L* L-L%, i.e. a falling movement starting from the nuclear syllable). If the accented syllable is final in the phrase, the final fall is substantially curtailed, as illustrated in [5] and [6]. These studies on BI also reported some cases where the terminal fall is replaced by a rise, these cases being described as: L+H* L-H%. This last case seems to be in better agreement with

previous studies on this variety, showing a falling-rising contour in yes-no questions read aloud [7], [8].

In this paper, we present the results obtained in an experiment carried out in order to determine the role of speaking style (spontaneous vs read) in the melodic realisations of terminal intonation contours in BI yes-no questions. Our hypothesis is that, since in Italian read speech is normally associated with a more "standard-like" pronunciation whereas spontaneous speech is associated with a more regional accent [9] [10], this difference is also reflected in yes-no question intonation in BI. Influence of speaking styles on melodic productions (especially for terminals) in questions has also been reported for other languages, such as American English [11].

Moreover, this difference could suggest an interpretation for final rises in spontaneous yes-no questions found in [5], [6], which could be classified as "less natural" as opposed to the "more natural" ones characterised by a final fall.

2. EXPERIMENTAL PROCEDURE

2.1 Data collection

Six speakers of Bari Italian aged between 24 and 37 (4 male and 2 female), all with a university education, were recorded in two sessions. Five of the six speakers were unaware as to the purpose of the experiment. In the first session, pairs of speakers were involved in a task-oriented dialogue (Map Task [12]). In a second session, each speaker was asked to read aloud the questions which s/he had produced during the dialogue in session one. Each speaker thus read aloud a different set of questions. In order to make sure the focal structure of the read questions was the same as that in the original spontaneous questions, the word containing the nuclear accent in the spontaneous question was underlined. All six speakers read their questions in two randomised lists (each randomisation was read once). Four of the six also read two repetitions of each their questions embedded in a specially produced text. Attention was paid in the construction of the text to providing information through the context on the focal structure of the target question. A number of questions were incorporated into the same text. A sample of a spontaneous context and the corresponding constructed context for two questions is given in Appendix 1.

2.2 Type of yes-no questions

Yes-no questions selected for intonation analysis are the ones referring to the move categories QUERY-YN (information-seeking yes-no questions), CHECK and ALIGN (confirmation-seeking yes-no questions), according to the Map Task coding scheme [13], [14]. As reported in [15], in BI pitch accent type reflects the degree to which the speaker believes whether the current question contains shared material. Therefore, in cases of confirmation-seeking questions, when the speaker is lacking in confidence as to the status of information requested, the same intonation contour, as in information-seeking questions, is used. i.e. a rising-falling nuclear pattern (L+H* L-L%). Our intonation analysis concerns both information-seeking and confirmation-seeking yes-no questions, the latter of the type described above.

3. RESULTS

3.1 Intonation analysis

As in the above mentioned studies on spontaneous BI speech [5], [6] the intonation analysis here uses a modified version of the ToBI transcription system [16]. Analysis of the spontaneous yes-no questions indicates that the predominant pattern is the typical rising nuclear pitch accent followed by a final fall (L+H* L-L%). The same pitch accent is also found in the corresponding read productions, but most of them have a terminal rise (L+H* L-H%). Examples of a spontaneous production and one read aloud in context of the utterance: “Sopra il museo delle BAMbole?” (“Above the DOLLs museum?”) are shown in Figures 1 and 2.

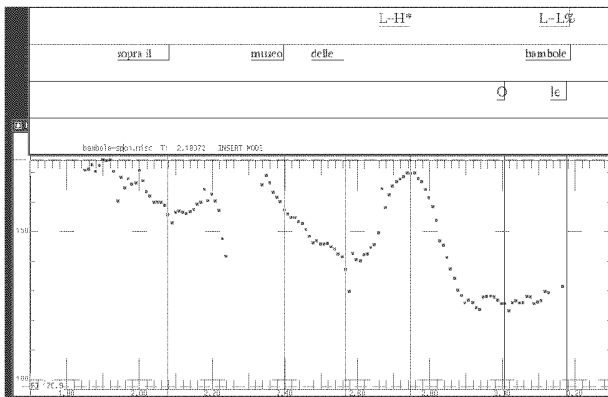


Figure 1 Spontaneous production of the utterance: *Sopra il museo delle BAMbole?* (“Above the DOLLs museum?”)

In all but one case (speaker ED), speakers used L+H* pitch accents. Owing to the variability in ED’s spontaneous data and to the fact that he was not naive as to the purposes of the experiment, his data were not further analysed. Only the data from the remaining five appear in the histograms in Figures 5.a and b below.

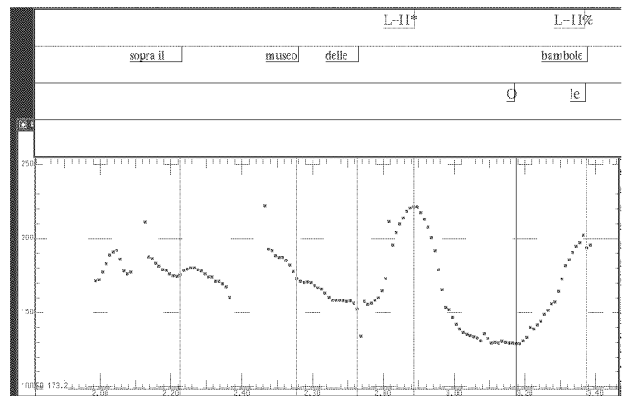


Figure 2 Production read aloud in context of the utterance: *Sopra il museo delle BAMbole?* (“Above the DOLLs museum?”)

The results in this paper refer to the pitch trace found in the final part of each utterance. In cases where the nuclear syllable is followed by one or more unstressed syllables, there is an actual fall or rise in F0 as well as a clearly perceived fall or rise in the questions transcribed with L+H* L-L% and L+H* L-H% respectively.

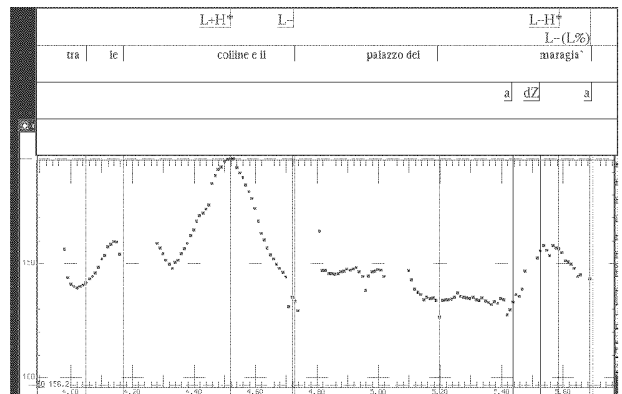


Figure 3 Spontaneous production of the utterance: *Tra le colline e il palazzo del maraGLA?* (“Between the hills and the maharaja’s palace?”), with final nuclear syllable.

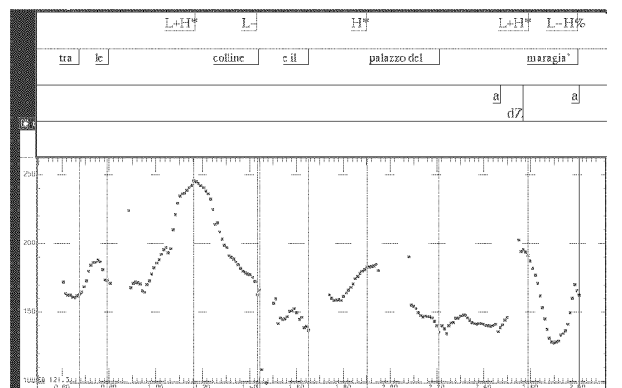


Figure 4 Production read aloud in context of the utterance: *Tra le colline e il palazzo del maraGLA?* (“Between the hills and the maharaja’s palace?”), with final nuclear syllable.

The results presented below are independent of any phonological analysis of the intonation contour, as they simply involve a decision as to whether the F0 rises or not during the final part of the utterance.

By contrast, in cases where the nuclear syllable is final, the relationship between the phonological analysis and the F0 trace is not always straightforward. In the spontaneously produced utterances, there was a rise starting before the accented syllable reaching a peak before the end of the syllable, followed by a fall. Although in some cases the fall was only slight, it typically reached roughly the middle of the speaker's current range. We have referred to cases where falls are not fully realised as "curtailed falls" [5] and analyse them in the same way as full falls where the nuclear syllable is followed by unstressed syllables (L+H* L-L%). Such an example is in Figure 3.

In the read data, there was a tendency for some speakers to produce not only a full fall but also a following rise. An example of an accentual rise followed by a full fall and rise is in Figure 4. Here the rising-falling-rising is realised on the final accented syllable "gia". In such cases, the final syllable is substantially lengthened, presumably to accommodate the more complex intonation contour.

However, not all speakers use this strategy when reading. Some (notably DD) curtailed the contour such that it was difficult to make a decision as to whether the uncurtailed contour would have otherwise been a simple fall or a fall-rise.

3.2 Terminal rises and falls

The tendency to produce terminal falls in spontaneous yes-no questions and terminal rises in read ones is consistently observed in all 5 speakers analysed.

Figure 5.a shows the productions of the 2 speakers who performed the two list-reading tasks, whereas Figure 5.b reports the results for the remaining 3 speakers who performed the additional text-reading task. In the histograms, the bars refer to spontaneous, read-in-context (only in Figure 5.b), the first randomised list, and the second randomised list.

One exception is found in the performance of speaker DD, where the distribution of final rises and no rises is equal in the context reading task. This could be explained by the difficulty of classifying the intonation of terminals for this speaker, because of her particularly fast speaking rate during this task.

This effect was enhanced by the presence, in her set of yes-no questions, of a substantial number of utterances with a nuclear syllable in final position. As mentioned in section 3.1, DD tended to curtail the tonal contour in final accented syllables.

From the data shown it can also be observed that final rises occur independently of whether they are produced in sentences embedded in a context or in simple lists. This means that the final rise is not simply an artefact of "list intonation".

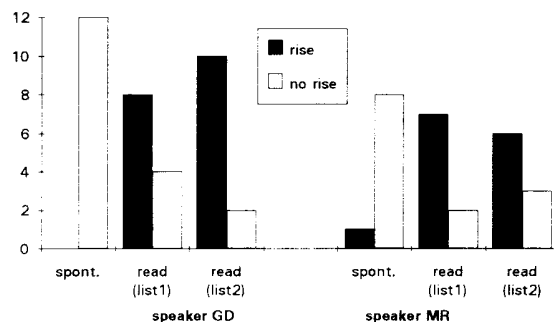


Figure 5.a

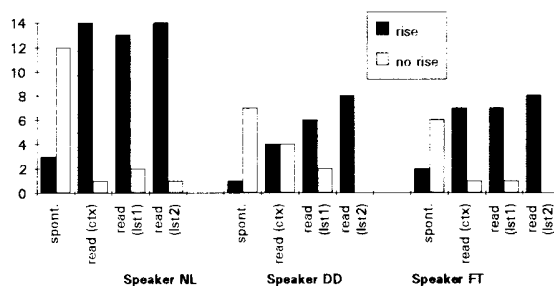


Figure 5.b

4. CONCLUSIONS

From the results obtained, it appears that in Bari Italian yes-no questions are characterised by the following typical contours:

spontaneous ⇒ low boundary tone / final fall

read ⇒ high boundary tone / final rise

One explanation of these results could be that regional Italian speakers tend to associate reading style with a more "standard-like" Italian pronunciation. Since "standard" Italian is characterised by a terminal rise in yes-no questions, Bari Italian speakers tend to produce such a rise on this kind of utterances in a reading task. This interpretation has been proposed also for the intonation of read yes-no questions in another variety of Italian, that spoken in Naples [17].

These considerations lead to the further suggestion that, in Bari Italian spontaneous yes-no questions, terminal rises can be interpreted as "less natural", whereas terminal falls are considered to be "more natural".

The same reason can account for the adoption of a "standard-like" Italian pronunciation even in a spontaneous context, where speakers may be influenced by the fact that they are being recorded and therefore that their speech is the focus of attention. In such cases,

they may change their regional (i.e. "natural") intonation into a more "standard-like" ("less natural") Italian intonation.

This hypothesis is supported by the fact that the speaker whose data were most variable was aware that the experiment was centred on his pronunciation.

The concept of "naturalness" may therefore be considered as a further dimension along the ones already described for spontaneous speech, for example in [18].

APPENDIX 1

1) Spontaneous context for question 1 (in bold):

IF: e io poi c'h.. eh legnaiolo...io ce l'ho il legnaiolo
vabbè, poi? andiamo avanti. **Ce l'hai il bambù?**

IG: il bambù non ce l'ho

IF: non ce l'hai. Io ce l'ho.

2) Spontaneous context for question 2 (in bold):

IF: **I mobili di vimini, ce li hai?**

IG: no. Tu il museo del velivolo ce l'hai?

IF: ce l'ho, e anche il mulino.

3) Constructed context for questions 1 and 2:

Successivamente Daniela ha domandato ancora a Marina: "**Ce l'hai il bambù?**" e Marina ha risposto: "No, non ce l'ho". Daniela ha continuato a chiedere: "**I mobili di vimini, ce li hai?**". "No", ha risposto Marina. Ormai era diventato evidente per entrambe che le due mappe non fossero identiche.

IG = Instruction Giver

IF = Instruction Follower

5. REFERENCES

- [1] M. Chapallaz, "The Pronunciation of Italian. A Practical Introduction", Bell & Hyman. London, 1979.
- [2] A. D'Eugenio, "Major problems of English phonology. With special reference to Italian-speaking learners", Atlantica, Foggia, 1982.
- [3] C. Avesani, "A contribution to the synthesis of Italian intonation", Proc. ICSLP 90, vol.1, pp. 833-836, Kobe, 1990.

- [4] M. Grice, "The intonation of interrogation in Palermo Italian: implication for intonation theory", Niemeyer. Tuebingen, 1995.
- [5] M. Grice & M. Savino, "Intonation and communicative function in a regional variety of Italian", *PHONUS*, vol 1, pp. 19-32, 1995.
- [6] M. Grice, R. Benzmueller, M. Savino, B. Andreeva, "The intonation of queries and checks across languages: data from Map Task dialogues", Proc. ICPhS '95, vol. 3, pp. 648-651, Stockholm, 1995.
- [7] L. Canepari, "Manuale di pronuncia italiana", Zanichelli. Bologna, 1992.
- [8] L. Canepari, "Italiano standard e pronunce regionali", CLEUP, Padova, 1980
- [9] L. Lepschy & G. Lepschy, "La lingua italiana", Bompiani. Milano, 1993.
- [10] M. Savino & M. Refice, "L'intonazione dell'italiano di Bari nel parlato letto e in quello spontaneo". to appear in: Atti delle VII Giornate di Studio dell'Associazione Italiana di Acustica: "Fonetica e fonologia degli stili del parlato", Napoli, 1996.
- [11] N.A. Daly, V.W. Zue, "Statistical and linguistic analyses of F0 in read and spontaneous speech", Proc. ICSLP 92, pp.763-766, Banff, 1992
- [12] A. H. Anderson, M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H. Thompson, R. Weinert, "The HCRC Map Task Corpus", *Language and Speech*, 34, (4): 351-366, 1991
- [13] J. Kowtko, S. Isard and J. Doherty-Sneddon, "Conversational games within dialogues", HCRC research paper, Edinburgh, 1992.
- [14] J. Carletta et alii (in prep.), "The coding of dialogue structure in a corpus". Twente Workshop on Language Technology on Corpus-Based Approaches to Dialogue Modelling, Twente, 1995.
- [15] M. Grice & M. Savino, "Can pitch accent type convey information status in yes-no questions?", to appear in Proc. ACL workshop "Concept to Speech Generation systems", Madrid, 1997.
- [16] M. Beckman & G. Ayers, "Guidelines for ToBI labelling", Ohio State Univ., 1994.
- [17] P. Maturi, "L'intonazione delle frasi dichiarative ed interrogative nella varietà napoletana dell'italiano", *Rivista Italiana di Acustica*, vol. XII, n.1, pp. 13-30, 1988.
- [18] M. Eskenazi, "Trends in speaking style research", Proc. Eurospeech 93, pp. 501-509, Berlin, 1993.