

Modeling intonation in interaction.

A new approach to the intonational analysis of questions in (semi-)spontaneous speech

The aim of this contribution is to present a new approach to the analysis of intonation in (semi-)spontaneous conversational settings, as a complement to the currently widespread analysis of read speech. We explore the idea that studying intonation in a rich pragmatic and interactional context yields precious insights, and we conclude that the benefits of meaningful communicative contexts offset the drawbacks imposed by challenging phonetic analyses.

The data that we discuss were gathered from a *Map Task* administered to 16 native speakers of Genoese (average age= 63; 6 females), a Northern Italo-Romance variety (cf. Forner 1988). Even if a full-fledged intonational description of this dialect is not currently available, a preliminary analysis of Genoese declarative and interrogative sentences is offered by De Iacovo (2017: 54-55) (for recent intonational analyses of the regional Italian variety spoken in Genoa, see Crocco 2011 and Savino 2012). Our *Map Task* was inspired by the *Montclair Map Task*, proposed by Pardo et al. (2019). Following this model, we have cancelled the ‘classic’ instruction Giver’s and Follower’s roles, in order to provide a more natural interactional setting. The speakers were provided with two maps, in which they were supposed to follow an already drawn path and which presented both shared and unshared landmarks (the target words were minimal pairs originally used for gathering data on the phonetic realization of short and long vowels. Since this is not the topic at issue here, we will not discuss it any further). The speakers were instructed to compare the items on their maps in order to ‘spot the differences’ (this is a feature of the *Montclair Map Task* that makes it similar to the ‘spot the difference’ task already used in the CLIPS corpus, cf. Albano Leoni & Giordano 2005, and in the *Diapix*, cf. van Engen et al. 2011; Baker & Hazan 2011).

Recordings were automatically annotated into InterPausal Units using temporal criteria (i.e. pauses of at least 200ms). IPU were then manually verified for boundary position. Out of the 1527 interpausal units composing the corpus, 72 were coded as questions. We excluded a large number of items for phonetic (e.g. inaudible overlaps), structural (e.g. false starts) and functional (e.g. stylized mockery) reasons. The 41 remaining items were judged reliable enough to undergo further analysis. Questions were then annotated with respect to their logical semantic structure (polar, alternative, *wh*-questions) and discourse-pragmatic functions (information-seeking, check, rhetorical), cf. Enfield et al. 2010 and Stivers & Enfield 2010, as well as their possible epistemic biases (biased, unbiased).

Our qualitative analysis has allowed us to recognize some recurring patterns in Genoese questions, such as a steep fall on the last stressed syllable, followed by a final rise, and often preceded by a high pitch region. Savino (2012: 34-35) also documents fall-rises in Genoese Italian, especially for post-nuclear material in polar questions. Note that this pattern was not always recognized as such by the Praat in-built pitch detector, due to the intrinsically “noisier” nature of conversational speech (Figure 1). This problem was solved by manually verifying pitch points as requested by the creation of periograms (Figure 2), where traditional f0 contours are modulated by periodic energy (Albert et al. 2018; Cangemi et al. 2019). In these displays, the curve “is wide and solid at the most periodic portions and it becomes gradually narrower and more transparent as periodic energy drops” (Cangemi et al. 2019: 807), with red portions indicating interpolated material. On the one hand, the absence of the final rise in the off-the-shelf Praat pitch extractor confirms the necessity for high-quality and reliable phonetic representations in the study of conversation. On the other hand, the possibility of extracting reliable and readable f0 representation from spontaneous data confirms that the ease of treatment of carefully elicited monological speech should not be prized beyond reason.

Crucially, in our corpus the steep fall with low rise was present in ca. 39% of the data for polar (Figures 2 and 4), alternative (Figure 3) and *wh*-questions, as well as for information-seeking (Figures 2 and 3) and check (Figure 4) functions. Therefore, it might seem that this pattern is used in Genoese independently of speakers and semantic types of questions as well as, from a discourse-pragmatic perspective, in both information-seeking and check contexts. This pattern does not seem to be attested in rhetorical questions (Figure 5), or in exchanges in which one speaker seems to perform not only a request for information, but also to express other types of meaning (such as specific attitudes and emotions related to the conversational dynamics). This is the case of the item in Figure 6, in which the speaker is asking for the fifth time about the possible presence of a [fry:tu] ‘fruit’ landmark on the interlocutor’s map, after having failed to receive a satisfactory answer to his previous questions.

These cases, anecdotal as they might seem, suggest that the analysis of conversational speech holds great promise for the understanding of intonation. They raise relevant issues for the study of prosody, such as (i) whether question intonation is as frequent in conversation as its high rate of representation in read speech research might lead us to believe; (ii) whether new visualization and representation techniques can make possible the widespread use of conversational speech in intonation research; and (iii) whether question intonation in interaction might be more sensitive to conversational functions than to well-studied semantic types and pragmatic properties, thus suggesting different directions for the future of intonation studies.

Speaker RM

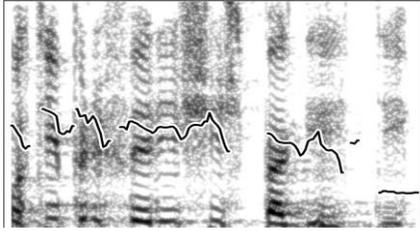


Figure 1. Praat - Polar Question, Information-seeking. *O ti preferisci i spaghetti?* “Or do you prefer spaghetti?”

Speaker RM

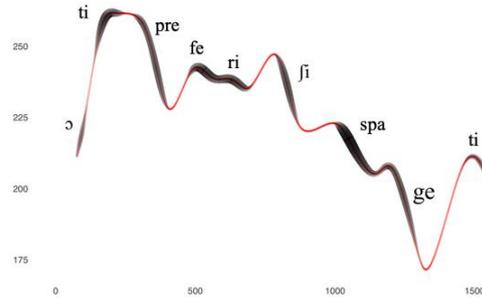


Figure 2. Periogram - Polar Question, Information-seeking. *O ti preferisci i spaghetti?* “Or do you prefer spaghetti?”

Speaker CG

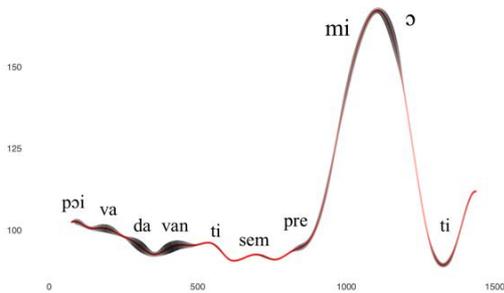


Figure 3. Alternative Question, Information-seeking. *[Poi vado avanti] sempre mi o ti?* “Then do I go on, always me or you?”

Speaker BG

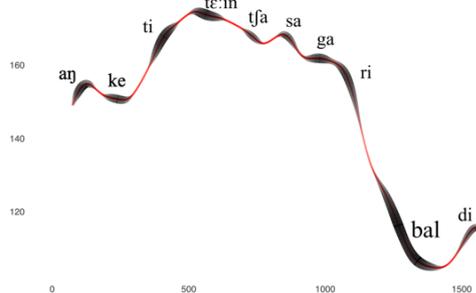


Figure 4. Polar Question, Check. *Anche ti t'è in Ciassa Garibaldi?* “Are you in Garibaldi Square too?”

Speaker GM

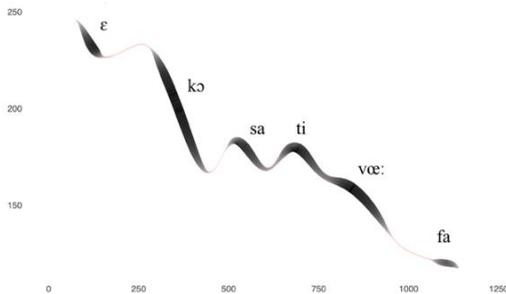


Figure 5. Wh-question, Rhetorical. *Eh, cosa ti voe fa?* “Eh, what do you want to do?”

Speaker BG

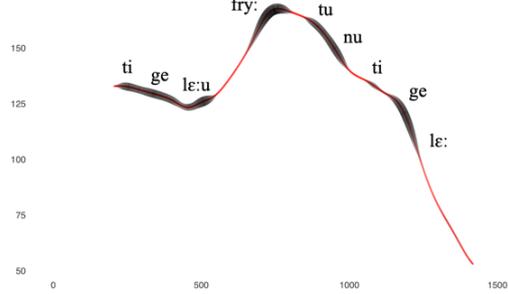


Figure 6. Alternative Question, Check. *Ti ghe l'è u frytu o nu ti ghe l'è?* “Do you have the fruit, or don't you have it?”

References

- Albano Leoni, F. & Giordano, R. (eds.) (2005). *Italiano parlato. Analisi di un dialogo*. Napoli: Liguori Editore.
- Albert, A., Cangemi, F. & Grice, M. (2018). Using periodic energy to enrich acoustic representations of pitch in speech: A demonstration. In *Proc. 9th International Conference on Speech Prosody*, 804-808.
- Baker, R. & Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior research methods*, 43(3), 761-770.
- Cangemi, F., Albert, A. & Grice, M. (2019). Modelling intonation: Beyond segments and tonal targets. In *Proc. 19th International Congress of Phonetic Sciences*, 572-576.
- Crocco, C. (2011). Profili melodici della varietà genovese. In *Atti del VII convegno dell'Associazione Italiana di Scienze della Voce*. Bulzoni Roma, 188-199.
- De Iacovo, V. (2017). *Intonation analysis on some samples of Italian dialects: an instrumental approach*. Doctoral thesis, Università degli Studi di Genova e Torino.
- Enfield, N. et al. (2010). Question–response sequences in conversation across ten languages: An introduction. *Journal of Pragmatics*, 42 (2010), 2615-2619.
- Forner, W. (1988). Areallinguistik I: Ligurien. In Holtus, G., Metzeltin, M. & Schmitt, C. (eds.), *Romanistischen Linguistik (LRL)*, vol. IV, Tübingen: Niemeyer, 453-469.
- Pardo, J. S et al. (2019). The Montclair map task: Balance, efficacy, and efficiency in conversational interaction. *Language and speech*, 62(2), 378-398.
- Savino, M. (2012). The intonation of polar questions in Italian: Where is the rise?. *Journal of the International Phonetic Association*, 42(1), 23-48.
- Stivers, T., Enfield, N. (2010). A coding scheme for question–response sequences in conversation. *Journal of Pragmatics* 42(10), 2620-2626.
- Van Engen, K. J. et al. (2010). The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and speech*, 53(4), 510-540.