

Discourse structure and phrase level phenomena: The phonetics of continuation contours in topic-internal position

Barbara Gili Fivela

University of Salento & CRIL – Lecce, Italy
barbara.gili@ateneo.unile.it

Abstract

This paper investigates the structuring of information in task-oriented dialogues, focussing on sequences of phrases produced with final-rising contours conveying continuation. Fragments extracted from dialogues are described with the aim of both showing some unexpected phonetic correlates of such contours and discussing the characteristics of the contexts where these correlates are observed.

A closer look at the phonetic implementation of continuation contours led to the observation that speakers may implement phrases with progressively expanded pitch ranges and with increasing fundamental frequency values for their boundary tones. The trend appears to be produced in well-planned sections of the dialogues, where it seems to be a quite robust phenomenon: the pattern was observed on not adjacent phrases, across quite complex stretches of dialogues. However, it is interesting to notice that the number of phrases showing this quite precise phonetic control was always three, four at the most - this find recalls other observations in the literature on phonetic long distance relations. Nevertheless, even when a sequence of phrases is characterized by rising contours, the topic-initial and topic-final phrases - usually characterized by falling edge tones - are usually realized with pitch range variations in line with those described in the literature. Therefore, the topic-internal increasing of frequency values at the boundaries, and of phrase pitch ranges, does not conflict because of the different phonological choices made by the speakers.

1. Introduction

A number of works in the literature investigate the relationship of suprasegmental information and various aspects of discourse, looking at both speech production and perception. Basically, the suprasegmental characteristics are found to be connected in a number of ways to information related to units larger than sentences: the main findings concern pauses, fundamental frequency (F0), voice quality, speech rate, duration and amplitude [2]. Among the acoustic correlates, here the investigation will be restricted to the analysis of F0 characteristics. F0 variations have been found to be connected with various information, such as those connected to the interactive structure in dialogues, e.g., turn taking, in different languages [3, 4]. Nevertheless, many studies show that F0 correlates to the position of the sentence in paragraphs, i.e. to topic structure.

Many works investigating the correlates of topic structure deal with read speech and monologues, and, at least for acoustic measurements, mainly focus on falling contours, usually signalling finality (although they may also imply continuity; [3], for a review, [7]). As for F0 characteristics, decreasing values have been reported to be connected with the developing of the topic, i.e. the pitch range is wider - and

possibly raised - at the beginning of a topic, and narrower - and possibly lowered - at the end [9,6]. This suggests that F0 values decrease through the topic units and can indicate the expected acoustic correlates of topic internal structuring.

When considering spontaneous and semispontaneous speech, rising contours in connection with topic structuring are more frequent, though they are not given particular attention as far as the phonetic implementation is concerned. Their role in discourse topic structuring has been clearly described as signaling that the unit is not finished, that it implies a following part, that it has to be interpreted on the basis of what follows [3, 11], and so forth, but few works investigated those contexts acoustically. In [10] the authors analyze one task dialogue and found that the F0 values of the boundaries within a 'speech-act-continuation' - the unit they consider - appear to gradually decrease, though they are higher than the values found at the boundaries of other topic classes.

The present study will show that in Pisa Italian phrases realized with continuation contours, and belonging to the same topic, may be characterized by F0 values that, on the contrary, increase through the unit. In these phrases, this is taken to signal the tight connection and the sequential type of the information. In the next sections, some fragments where the pattern was observed will be discussed in order to attest the existence of this alternative phonetic strategy that speakers may follow in topic internal position; moreover, the examples will offer a first suggestion on which factors may play a role in the choice of the strategy.

2. Corpus and methods

Eight Map-Task dialogues were considered for this investigation - for a total of about 2 hours and 22 minutes of speech (see also [7]). Map-Task dialogues [1] are elicited by giving speakers two maps: only one map has a path through the landmarks drawn on, and the task is to interact verbally for reconstructing the path on the other map.

The presence of a task usually gives the dialogue a pretty straightforward topic structure. Therefore, the task exploited for the recordings is taken here as a reference point for defining what is meant as topic, along the lines of the strategy adopted in other works in the literature. For instance, in instruction monologues on the reconstruction of the front view of a house, a group of sentences dealing with the reconstruction of a particular part of the front view - a building block - is considered as part of the same topic [13]. Similarly, the working hypothesis for the Map-Task dialogues analysed here is that the global path is seen as the sequential reconstruction of smaller path segments, whose starting and ending points may be identified with specific landmarks. In the coding system proposed in [5], and adopted to transcribe the Pisa Italian dialogues, the discourse's stretches where major steps in the drawing of the path are achieved are

identified with transactions. Therefore, a topic will be considered here as a transaction. In procedural terms, this implies that transactions will usually begin and end whenever a landmark is reached (and it is not just exploited as a reference point); it also implies that the analysis will be oriented mainly to sequences of moves containing indications for reconstructing the path (mainly instruct moves). As already mentioned, the attention will focus on those uttered with continuation contours. These intonational contours, according to the analysis given for Pisa Italian within the autosegmental-metrical framework [Pierrehumbert, 1980], are usually implemented with high edge tones.

2.1. Discourse structure and prosodic analysis

The eight dialogues were orthographically transcribed; four out of eight dialogues were then transcribed in terms of transactions [5], on the basis of only textual information. Then, the eight dialogues were auditorily analyzed in order to individuate the contexts relevant for the study, i.e. sequences of instructions where the speakers used a significant number of continuation contours (≥ 2). These sequences were digitized and the F0 traces were inspected.

The intonation analysis was performed according to the autosegmental-metrical framework, and in particular according to the analysis proposed for Pisa Italian [7]. The contexts under investigation are analyzed as characterized by nuclear H+L* or [L+]H* pitch accents and H-H% edge tones. Although continuation contours may be characterized by different nuclear pitch accents, they are here homogeneous as far as edge tones are concerned (in Pisa Italian, not surprisingly, high edge tones may also be found in elliptic questions; in other question types - e.g., in checks - only a high tone is found, and it plays the role of a phrase accent, followed by a low boundary tone).

2.2. Acoustic analysis

The following measurements were performed for each intonation/intermediate phrase in the instruction sequence:

- High edge tones. The last stable F0 point in H-H% contexts was measured, correcting the values due to octave errors.
- Minimum and maximum F0 value. The former was preferably taken in the stable part of vowels, and no spurious F0 point was measured. The maximum was the highest F0 value in phrases, neglecting spurious points.
- Pitch range. The pitch range over the intermediate, or intonation, phrase was measured as the difference between its maximum and its minimum value.

3. Results

3.1. Discourse structure and prosodic analysis: Observations

The dialogues appear to be differently characterized in terms of topic structuring. In fact, in some of them the topic organization is more straightforward than in others. In two dialogues, the instruction-giver divides the path in feasible segments and gives very structured information to the instruction-follower who interacts actively but, basically, accepts the proposed information organization. In the other two dialogues, for instance, slightly different characteristics are found. The main difference appears to be connected with the fact that instruction-followers seem to compete to a

greater extent with the instruction-giver in organizing the flow of information.

Sequences of instructions were analyzed in the four dialogues which were coded as for the transactions level, while the other four dialogues were used to validate the analysis. As expected, high boundaries were reasonably frequent in the dialogues, and in sequences of instructions the speakers were not bound to the use of rising contours to express continuation. This is in line with the findings reported in the literature for other languages [3].

3.2. Acoustic analysis

The measurements of high boundaries' F0 values reflect the existence of two possible phonetic characterizations of the contexts under investigation. On one side, high boundaries may be implemented at variable, often decreasing, F0 values; on the other side, examples of consistently increasing F0 values are also found. The first type of pattern is the most similar to those described in the literature, while the second one represents an instance of the control speakers may perform on phonetic parameters.

Focussing on this second pattern, two examples of sequences of instructions within transactions will be described: phrases uttered with continuation rise contours will be shown as realized with progressively increasing F0 values.

Stretch 1

- High edge tones

In Figure 1, the translation of part of the orthographic transcription of a transaction is given. The slashes marked by numbers correspond to intermediate and intonation phrases in the giver's production (questions are left out), and the underlined phrases are those produced with continuation contours. The figure reports part of the transaction produced to reconstruct the path between two landmarks: *leone* 'leon' and *ristorante Anima Mia* 'the restaurant Anima Mia'.

```
G093: no, you don't have to break it up/17 <pb> go around it/18
F094: #<G093> no, <ah> <pl># yes
G095: then from the bottom/19 then from the right of the leon/20
F096: yes
G097: and you go up up/21
F098: yes
G099: do you have the restaurant 'Anima Mia' ?
```

Figure 1: Translation of part of the transaction (dialogue C03, turns G093-G099).

The F0 values for high boundary tones in giver's productions are plotted in Figure 2. The numbers corresponding to the sequential phrases uttered by the giver are shown on the x axis (they are indicated on the transcription in figure 1 as well), while on the y axis F0 values (Hz) are given. Each circle in the plot represents the F0 value of a phrase produced with a continuation contour, i.e. with high edge tones. This means that not all sequential phrases in the transaction have a corresponding value on this plot, as not all of them end with high edge tones.

Phrases from 17 to 21 belong to the transaction, where the segment between two landmarks is at issue. Notice that the path around *leone* 'leon' is easily described by the giver (*non lo spacca' in due, ci passi intorno* 'don't break it down, just go around it') and the next icon, which should be reached by means of the instructions in phrases 19, 20 and 21, is really close. These phrases present increasing F0 values for their

high boundary tones - see the trend from low to high in figure 2. The path drawn on the follower's map shows that phrases with high edge tones collect all the necessary information to reconstruct the relevant part of the path. Therefore, the speaker makes the same phonological choices along a relevant part of the transactions and he consistently implements the height of the boundary tones.

- Pitch range values

For each phrase in the transaction, the pitch range was calculated. The results are plotted in figure 3, where the phrases under discussion - number 19, 20, 21 - are plotted together with those belonging to the preceding transaction.

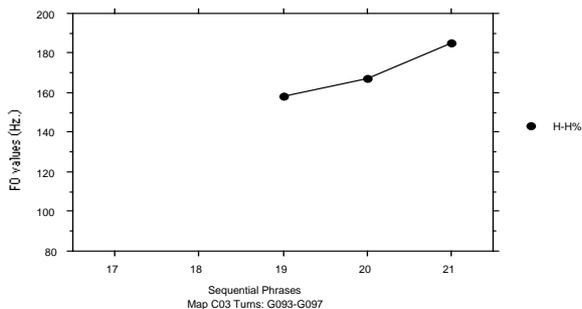


Figure 2: Plot of F0 values at phrases' boundaries (Turns G093-G097): Number of giver's sequential phrases - x axis - and F0 values (Hz) - y axis

The number of giver's sequential phrases - see also figure 1 - are given on the x axis, while the F0 values (Hz) are given on the y axis. Notice that all the phrases in the transactions have been plotted (apart from questions) and those produced with continuation contours are plotted as filled circles. As the graph shows, the pitch range values for the phrases produced with high edge tones show progressively increasing values, i.e., the pitch range appears to be progressively expanded in the instructions within the transaction (in phrase 14 the speaker hesitates, and, in fact, interrupts himself; this may explain why the measures show different values).

The plot also shows the other phrases' pitch range values (empty circles). According to the literature, the beginning of topics is characterized by higher peak values, corresponding to expanded pitch ranges. In figure 3, the first phrase in the transaction under discussion is phrase 17 (no high edge tones). Notice that the pitch range of this phrase is wider than both the previous (phrase 16 vs 17) and the following one (17 vs 18). Moreover, notice that also in the preceding transaction a rising trend may be observed in high edge tone phrases - number 2, 5, 9, and 13 in the plot- and that, on the other hand, phrase 1 (no high edge tones) shows a greater value than both the previous phrase, number 0, and the following, number 2. Therefore, the first phrases in the two transactions, which do not belong to the set of phrases produced with a continuation contour, show pitch range values higher than the neighbouring phrases' values. This is in line with the reported tendency for topics to start with high pitch range values¹.

Thus, it appears that, on the one hand, topic initial phrases with no high edge tones are characterized by particularly high pitch range values - thus, they could potentially start a decreasing series of pitch range values for the following units; on the other hand, the presence of phrases produced with high

¹ Moreover phrase 12 is the first phrase after a nested game and shows a higher pitch range value than adjacent phrases (see [7]).

edge tones, i.e. uttered with continuation contours, appears to influence the pitch range variations in the opposite direction, i.e. the pitch range in those phrases appears to progressively expand within the transaction.

- Minima values

The values do not offer a clear pattern and do not allow one to understand whether the pitch range is only expanded or it is also shifted: The only possible observation relates to a slight rising of the speaker's floor in the last instruction realized with high edge tones in both transactions.

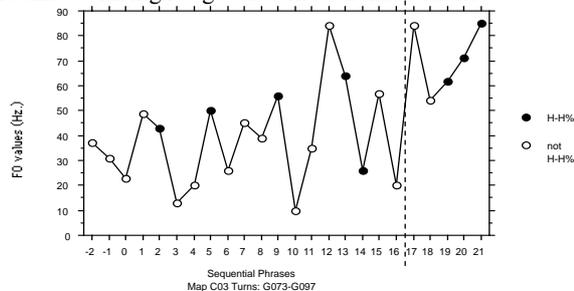


Figure 3: Pitch range values in two transactions (Turns G073-F092 and G093-G097): Number of giver's sequential phrases - x axis - and F0 values (Hz) - y axis

Stretch 2

- High edge tones

An extra example relates to the transaction reconstructing the path from the landmark *giardino delle visite* to *miniera* - see figure 4. Similarly to what observed above, the phonetic implementation of the boundaries appears to clearly go from lower to higher F0 values - see figure 5 - and notice that the number of phrases characterized by this pattern is four (phrases 3, 5, 9 and 12).

G033: *okay then you go/2 <pb> <eeh> while going above the 'giardino delle visite'/3 go towards/4 <pb> the right side/5*
 F034: *but from 'bar da Liola' then / then I go+ 'giardino delle visite' and then I overcross it*
 G035: *yes <pb> but yougo there going/6 <pb> doing this long path/7*
 F036: *#<G035> yes <pl># #<G037> <mhmh>#*
 G037: *then, go above the garden/8 and go towards the right side/9 <pb> untill/10 <pb> boh <pb> three quarters of the page/11*
 F038: *#<G037> yes <pb> # yes <ii>*
 G039: *then you go up/12*
 F040: *<mhmh>*
 G041: *go up untill the 'miniera'/13, do you have 'miniera'?*

Figure 4: Translation of part of transaction (dialogue C03, turns G033-41)

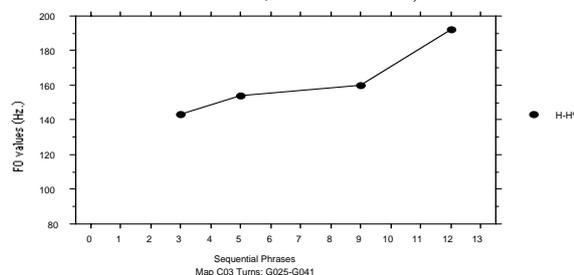


Figure 5: Plot of F0 values at phrases' boundaries in transaction (Turns G025-G041): Number of giver's sequential phrases - x axis - and F0 values (Hz) - y axis

- Pitch range values

The pitch range values for the transaction are shown in the next plot – see figure 6. In the transaction, phrase 1 (*da quel punto li* ‘from that position’) shows a wide pitch range in comparison to both the following and the previous phrase, that is actually the last phrase in the preceding transaction. Therefore, the beginning of this transaction is expressed by means of a wider pitch range, as is usually reported in the literature. After that, the sequence of continuation rises, starting with phrase 3 - whose range is quite narrow - and continuing with phrases 5, 9 and 12, shows increasing pitch range values. Therefore, both the beginning of a new topic and the internal structuring of the topic are signalled.

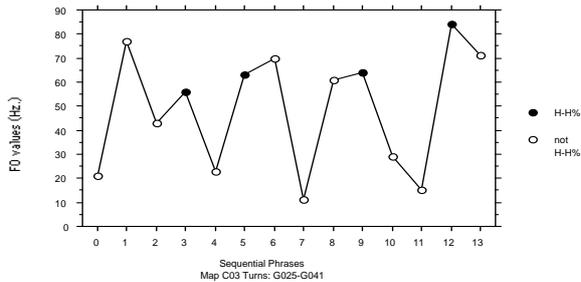


Figure 6: Pitch range values in two transactions (Turns G025-G041): Number of giver's sequential phrases - x axis - and F0 values (Hz) - y axis

- Minima values

The minima values of phrases produced with high edge tones do not appear to follow any particular pattern throughout the sequence, apart from the fact that the last instruction is characterized by a higher minimum.

Notice that sequences of phrases with at least two elements show an inter-phrase difference in semitones equal, or greater, than 1.5 – see table 1, for the examples above. This is one of the threshold considered as meaningful for a perceptually relevant difference [12]. Although these results have to be taken as purely indicative, and systematic perception experiments should verify the relevance of the trend observed in the data, it appears that the rising values found at phrases boundaries could be perceptually relevant, at least when the whole sequence is concerned.

Table 1: Difference in semitones (ST, calculated as in [8]) of high boundary F0 values in 'adjacent' phrases.

Transact.	Ph. n	Ph. n+1	Ph. n F0 at H%	Ph. n+1 F0 at H%	ST
Stretch1	19	20	158	167	0,9590
Stretch1	20	21	167	185	1,7721
Stretch2	3	5	143	154	1,2829
Stretch2	5	9	154	160	0,6616
Stretch2	9	12	160	192	3,1564

4. Discussion and conclusions

In the previous sections, some phonetic characteristics of topic structure in dialogues were investigated. Some fragments of dialogues were described with the aim of both showing a peculiar acoustic strategy speakers appear to follow in topic internal position and offer a first sketch of the factors that may play a role in the choice of the strategy. The examples show that a possible strategy for signalling the inclusion in the same topic, i.e. the same transaction, is to

systematically increase the F0 values for the high boundary tones and the pitch range values for each subsequent phrase. The minima measurements presented too much variation to make it possible to understand if the pitch range was both expanded and up-shifted. However, these topic internal cues do not seem to contradict more 'peripheral' cues to topic structure. In fact, for low edge tones phrases, more compressed pitch range values in a topic final position and more expanded range values in a topic-initial were observed.

It is worth to point some observations out, at least schematically. First of all, the trend described above appears to be produced in well-planned sections of the dialogues, where the giver seems to have clear in mind what is necessary to reconstruct the path segment. In fact, the examples were not equally distributed in the corpus. Differences in the planning of the information due to the participants capabilities, or to the more or less successful communication skills of the paired speakers, seem to have a very high influence on the frequency of occurrence of the pattern. Moreover, the number of phrases produced with high edge tones showing the phonetic control described in the previous sections was always three, four at the most, and the phrases collect all the necessary information to reconstruct the relevant part of the path. Interestingly enough, sequences composed of 3-4 'units' have already been reported as reflecting the organization into sub-topics [3]; this invites the hypothesis of a sort of upper limit in the set of instructions which may belong to such a well-planned sequence.

The data described here only exemplify a phonetic strategy that speakers may follow in topic internal position and further work is needed in order to offer statistical data on both the phonetics and the frequency of the pattern observed.

5. References

- [1] Anderson A., Bader M., Bard E., Boyle e., Doherty G., Garrod S., Isard S., Kowtko J., McAllister J., Miller J., Sotillo C., Thompson H., Weinert R. 1991. The HCRC map task corpus. In *Language and Speech*, 34, 4, 351-366.
- [2] Blaauw E. 1995. *On the perceptual classification of spontaneous and read speech*. PhD thesis, Utrecht: OTS
- [3] Brown G., Currie K., Kenworthy J. 1980. *Questions of Intonation*. Croom Helm, London.
- [4] Caspers J. 1998. Who's next? The melodic marking of question vs. continuation in Dutch. In *Language and Speech*, 41, 375-398.
- [5] Carletta J., Isard A., Isard S., Kowtko J., Doherty-Sneddon G., Anderson A. 1996. HCRC Dialogue structure coding manual. In *Human Communication Research Center, Technical Report 82*, University of Edinburgh, June, 1-27.
- [6] Cruttenden A. 1986. *Intonation*. Cambridge University Press
- [7] Gili Fivela B. 2004. The phonetics and phonology of intonation: The case of Pisa Italian. PhD Diss., Scuola Normale Sup., Pisa.
- [8] 't Hart J., Collier R., Cohen A. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody*. Cambridge University Press.
- [9] Lehiste I. 1975. The phonetic structure of paragraphs. In A.Cohen, S.G.Nooteboom (eds.), *Structure and Process in Speech Perception*, Springer-Verlag, 195-206
- [10] Nakajima S., Allen J. 1993. A study on Prosody and discourse structure in cooperative dialogues. In *Phonetica*, 50, 197-210.
- [11] Pierrehumbert J., Hirschberg J. 1990. The meaning of intonational contours in the interpretation of discourse. In P.R.Cohen, J.Morgan, M.E.Pollack (eds.), *Intentions in communication*, The MIT Press, cap.14, 271-310
- [12] Rietveld A.C.M., Gussenhoven C. 1985. On the relation between pitch excursion size and prominence. In *Journal of Phonetics*, 13, 299-308.
- [13] Swerts M., Gelyuykens R. 1993. The prosody of informational units in spontaneous monologue. In *Phonetica*, 50, 189-196.