

# A topologic view of Topic and Focus marking in Italian

Gloria Gagliardi<sup>1</sup>, Edoardo Lombardi Vallauri<sup>2</sup>, Fabio Tamburini<sup>3</sup>

<sup>1</sup> Department of Italian Studies, University of Firenze, Italy  
gloria.gagliardi@unifi.it

<sup>2</sup> Department of Linguistics, University of Roma Tre, Italy  
lombardi@uniroma3.it

<sup>3</sup> Department of Linguistics and Oriental Studies, University of Bologna, Italy  
fabio.tamburini@unibo.it

## Abstract

Regularities in position and level of prosodic prominences associated to patterns of Information Structure are identified for some Italian varieties. The experiments' results suggest a possibly new structural hypothesis on the role and function of the main prominence in marking information patterns. (1) An abstract and merely structural, "topologic" concept of Prominence location can be conceived of, as endowed with the function of demarcation between units, before their culmination and "description". This may suffice to explain much of the process by which speakers interpret the IS of utterances in discourse. Further features, such as the specific intonational contours of the different IS units, may thus represent a certain amount of redundancy. (2) Real utterances do not always signal the distribution of Topic and Focus clearly. Acoustically, many remain underspecified in this respect. This is especially true for the distinction between Topic-Focus and Broad Focus, which indeed often has no serious effects on the progression of communicative dynamism in the subsequent discourse. (3) The consistency of such results with the law of least effort, and the very high percent of matching between perceptual evaluations and automatic measurement, seem to validate the used algorithm.

**Keywords:** Information Structure, Prominence, Italian speech.

## 1. Introduction

One of the main functions of acoustic (intonational and accentual) patterns in linguistic utterances is the expression of Information Structure (IS). We have argued elsewhere (Lombardi Vallauri, 2001; 2009) that the level of IS most related to acoustic features is the one mainly referred to in the literature as "Theme-Rheme" or "Topic-Focus", for which we adopt the definitions proposed by Cresti (1992; 2000) and Lombardi Vallauri (2001; 2009), based on which part(s) of the utterance may be regarded as conveying its illocutionary force. We assume that the Focus is *the part of an utterance which carries illocutionary force and realizes the informational purpose of the utterance itself*. The Topic, on the contrary, is *the part of an utterance that has no illocutionary force, whose function is to allow the comprehension of the Focus with respect to the discourse*.

These definitions essentially match those (though not always explicitly expressed) underlying the concepts of Topic and Focus (Theme-Rheme, Topic-Comment) usually dealt with in much literature concerned with the acoustic correlates of IS (e.g. (Halliday, 1989; Ladd, 1978; 1996; Pierrehumbert, 1987; Selkirk, 1984), and, more relevant in relation to our analysis, (Avesani, 2000; Avesani, Vayra, 2004; Avesani, *et al.* 2007; Breen, *et al.* 2010; D'Imperio, 2002b; Féry, Krifka, 2008; Frascarelli, 2000; 2004; Frascarelli, Hinterözl, 2007), etc.). For the purposes of the present study, chunks of linguistic material in utterances from two corpora of spoken Italian have been labeled as Topic or Focus following

essentially two criteria:

- First, the subjective impression (mainly based on the perception of acoustic patterns, but also on negation tests) that a certain part of the utterance conveys illocutionary force, thus being also responsible for the linguistic act carried out by the utterance itself, i.e. for its being an assertion, a question, a request, a command or any other pragmatically relevant act (see (Cresti, 2000), for a list of about 80 illocutionary acts).

- Second, the evaluation of the preceding context, aimed at establishing which information may be considered as active (Chafe, 1987; 1992) at the utterance time, i.e. Given, and consequently less likely to be in Focus, and which information may be considered inactive, i.e. New, and consequently more likely to be in Focus.

Only three typologies of IS were examined, namely Broad Focus (extending to the whole utterance), Topic-Focus and Focus-Appendix (i.e. constructions with a Narrow Focus located to the left of the utterance).

Some studies on the matter directly investigate the relations between IS and phonetic phenomena, while others analyse them through an intermediate, phonological level. (e.g. (Ladd, 1996; Pierrehumbert, 1987) and all studies adopting the ToBI labelling scheme (Beckman, *et al.* 2005)). In this second perspective phonological categories are derived from acoustic parameters, mainly considering intonation, i.e. F0 profiles.

Most studies on Italian belong to the Autosegmental Metrical (AM) paradigm, quite often based on read rather than spontaneous speech. Table 1 outlines the (typical) tonal profiles, mainly pitch accents, of assertive

utterances described by various scholars regarding the Italian varieties examined in this study.

	Broad Focus	Narrow Focus	Contrastive Focus
Rome (Frascarelli, 2004)	H+L*	H* H*+L	H* H*+L
Florence (Avesani, Vayra, 2004)	H+L*	H+L*	L+H* (L+H)* H+H*
Naples (D’Imperio, 2002b)	H+L*	L+H*	L+H*

Table 1: typical tonal profiles of assertive utterances in AM studies.

As is shown, contrastiveness is marked intonationally in Florentine, while in Roman and Neapolitan different pitch accents depend on Focus breadth. It is still unclear whether such differences are due to diatopic variation or to idiosyncrasies of the ToBI transcription scheme. On the one hand ToBI notation seems unable to account for melodic differences clearly perceived by the speakers: Broad Focus of assertive utterances is represented through the same pitch accent although hearers are able to identify the geographic origin of other speakers on the sole basis of intonation (Marotta, 2008). On the other hand, scholars agree on the identification of edge tones and pitch accents, but not about the classification of pitch accents different in nature (Pitrelli, et al. 1994; Syrdal, McGorg, 2000). Disagreement concerns tonal alignment (D’Imperio, 2002a; Gili Fivela, 2002) and tonal target identification, in particular inside plateaux (where a single maximum or minimum cannot be easily discerned) (D’Imperio, 2002a). Information about scaling (i.e. the frequency range within pitch accents) and slope is underestimated, although potentially distinctive (Gili Fivela, 2002).

As suggested in some classical studies (such as Ladd, 1996) and substantiated in more recent investigations (Breen, et al. 2010; Lee, Yu, 2010), a focused item might involve a complex combination of different acoustic cues, namely duration, pitch and intensity, and cannot be analysed only through its intonational profile.

For these reasons, we will try to investigate the correlation between focused items and phonetic features by considering the concept of prosodic prominence as a complex and rich set of acoustic features combined in a sophisticated way. The automatic identification of prominence levels is definitely a complex task.

## 2. Prominence Definition and Automatic Detection

Following e.g. (Couper-Kuhlen, 1986; Jensen, 2004; Kohler, 2006; Mertens, 1991; Terken, 1991), we can define prosodic prominence as a *perceptual phenomenon, continuous in its nature, emphasizing segmental units with respect to their surrounding context, and supported by a complex interaction of prosodic and*

*phonetic/acoustic parameters.*

Due to its methodological rigour, we will primarily refer to (Kohler, 2005) for a description of the interactions between the different prosodic features that determine the perception of prominence. In his view, there are two main ‘actors’ playing a relevant role in supporting sentence prominence (or sentence accent). The first, *pitch accent* (Bolinger, 1958) concerns specific movements in F0 profile. The second, *force accent*, is independent from intonation and is connected with intensity, segmental durations and possibly other parameters. Both phenomena seem to play relevant roles in supporting prominence perception at utterance level (see also Ladd, 1996), reinforcing each other without establishing specific antagonistic or hierarchical roles.

One of the major challenges in predicting syllable prominence is the disentangling of various sources of influence such as fundamental frequency excursions, duration, intensity related parameters and the listeners’ linguistic expectancies. At the acoustic level, various studies (e.g. Bagshaw, 1994; Heldner, 2003; Sluijter, van Heuven, 1996; Streefkerk, 1996) suggest, also cross-linguistically, the dependence of force accents from unit duration and spectral emphasis (spectral tilt or spectral balance), while pitch accents would be supported by specific F0 configurations and by the global intensity inside a particular segmental unit. One of the authors has carried out experiments confirming such relations for some languages (Tamburini, 2005; 2006; 2009).

Assuming this view, we can introduce a prominence function which should be able to assign a continuous prominence level to each syllabic nucleus using only acoustic information:

$$Prom^i = W_{FA} \cdot [SpEmph_{SPLH-SPL}^i \cdot dur^i] + W_{PA} \cdot [en_{ov}^i \cdot (A_{event}^i(at_M, at_m) \cdot D_{event}^i(at_M, at_m))]$$

where  $SpEmph_{SPLH-SPL}$  is the spectral emphasis,  $dur$  is the nucleus duration,  $en_{ov}$  is the overall energy in the nucleus and  $A_{event}$  and  $D_{event}$  are the parameters derived from the TILT model (Taylor, 2000) as a function of the maxima alignment type –  $at_M$  – and the minima alignment type –  $at_m$ . All parameters are referred to the generic syllable nucleus  $i$ . See Table 2 for some details on parameter computation.

The body of the function  $Prom$  contains nine parameters. Five of them can be considered as supporting the prominence phenomenon from a cross-linguistic point of view ( $SpEmph_{SPLH-SPL}$ ,  $dur$ ,  $en_{ov}$ ,  $A_{event}$  and  $D_{event}$ ), while the other four, represented in the vector  $\mathbf{W} = (W_{FA}, W_{PA}, at_M, at_m)$ , can be seen as language specific. In our model,  $W_{FA}$  and  $W_{PA}$  weigh the contribution of the two different accent types, while  $at_M$  and  $at_m$  model the different pitch accent alignments specific for each language (see Fig. 1).

Parameter	Description
Nucleus Duration ( $dur$ )	Time duration of the syllable nucleus normalised by considering the mean and variance duration of the syllable nuclei in the utterance (z-score), computed using the manual segmentation available in the considered corpora.
Spectral emphasis ( $SpEmph_{SPLH-SPL}$ )	Normalised SPLH-SPL parameter (Fant, <i>et al.</i> 2000) (z-score).
Pitch movements	TILT model (Taylor, 2000) representation of pitch movements derived from a pitch contour computed using the <i>ESPS get_f0</i> program (Talkin, 1995).
Overall intensity ( $en_{ov}$ )	RMS energy computed in the frequency band 50-5000 Hz normalised to the mean and variance of intensity inside the utterance (z-score).

Table 2: Acoustic parameters used by the prominence identification algorithm.

All the parameters involved in the *Prom*-function computation are normalised inside the utterance, thus the contributions of different speakers and numeric ranges should be factored out. In all the experiments we used  $\mathbf{W} = (1.0, 1.0, 2, 2)$ .

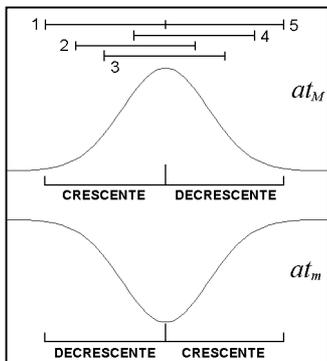


Figure 1: Alignment type parameters between pitch accents and syllable nuclei.

### 3. Experiments

The two experiments presented here were aimed at searching invariances in position and level of the Main Prominence, identified through the automatic algorithm presented in the previous section, compared to the IS assigned to the utterances by an expert annotator.

The first experiment is a pilot study on a limited corpus of spoken Roman Italian. The second experiment was aimed to verify the results for the same kind of Italian on a different corpus, and to extend the analysis to two further diatopic varieties, namely Florentine and Neapolitan Italian. The annotator identified the mandatory unit of Focus and possible units of Topic and Appendix, if present. He also determined Focus breadth and possible contrastiveness. We will consider here

utterances of 3 classes on the basis of IS: (a) TOPIC | FOCUS; (b) BROAD FOCUS; (c) FOCUS | APPENDIX, NARROW FOCUS, CONTRASTIVE FOCUS. The utterances containing retracting, hesitations and speech disfluencies have been discarded.

(a) TOPIC   FOCUS							
Var.- Corp.	Main Prominence on the...						No Main Prom
	LsT	LsF	LsA	IsT	IsF	IsA	
R-B	18	1	-	0	1	-	3
R-C	12	3	-	1	0	-	3
F-C	24	1	-	0	1	-	7
N-C	8	0	-	2	1	-	2
(b) BROAD FOCUS							
Var.- Corp.	Main Prominence on the...						No Main Prom
	LsT	LsF	LsA	IsT	IsF	IsA	
R-B	-	4	-	-	0	-	4
R-C	-	4	-	-	6	-	8
F-C	-	3	-	-	3	-	2
N-C	-	4	-	-	7	-	6
(c) FOCUS   APPENDIX, Narrow F, Contrastive F							
Var.- Corp.	Main Prominence on the...						No Main Prom
	LsT	LsF	LsA	IsT	IsF	IsA	
R-B	-	14	0	-	2	0	0
R-C	-	22	1	-	2	0	2
F-C	-	14	1	-	1	0	2
N-C	-	25	0	-	6	0	0

Table 3: Number of utterances divided by Variety-Corpus pairs (R=Rome, F=Florence, N=Naples; B=Bonvino, C=CLIPS) and configurations (e.g. LsT=Last syl. of Topic, IsF=Internal syl. of Focus). Some combination pairs are not possible; in those cases we have inserted a ‘-’ in the corresponding cells.

#### 3.1 Experiment 1

The data have been extracted from the “Bonvino” corpus, a section of *Ar.Co.Dip.* (Bonvino, 2005). It consists of 12 conversations by speakers from Rome, homogeneous in social level, age, level of education and geographical origin. 47 utterances have been selected from three conversations; the corresponding waveforms have then been extracted, and a reference transcription has been manually added to mark the syllabic nuclei needed for prominence identification.

#### 3.2 Experiment 2

The data have been selected from the spoken dialogue sub-corpus of *CLIPS* (in particular, from the map-task sections), stratified through diatopic and diaphasic dimensions (Albano Leoni, 2003). The choice fell on the labeled texts from Rome, to replicate the first experiment using a different data set, Florence and Naples, so far particularly studied in the autosegmental-metric phonology approach. 184 utterances have been selected: 64 for Rome, 59 for Florence and 61 for Naples.

The results of both experiments, depicted in Table 3 above, show relevant regularities considering the

position of the Main Prominence in relation to the kind of IS. First of all, we can note that, considering each specific IS, there are no relevant differences between the Italian varieties: the distribution of the Main Prominences seems to follow similar patterns in the different Variety-Corpus pairs. Moreover, the position of the Main Prominence tend to be placed at the border between the two IS components for the TOPIC | FOCUS and the FOCUS | APPENDIX IS, while, in case of BROAD FOCUS utterances, the overall picture seems to be less clear, even if a slight tendency of the Main Prominence to be at the end of the utterance can be found. Figure 2 outlines these regularities for three example utterances: Aurelia\_02 (TOPIC | FOCUS), Colosseo\_04 (BROAD FOCUS) and Chiacchiere\_42 (FOCUS | APPENDIX) all from the Bonvino corpus.

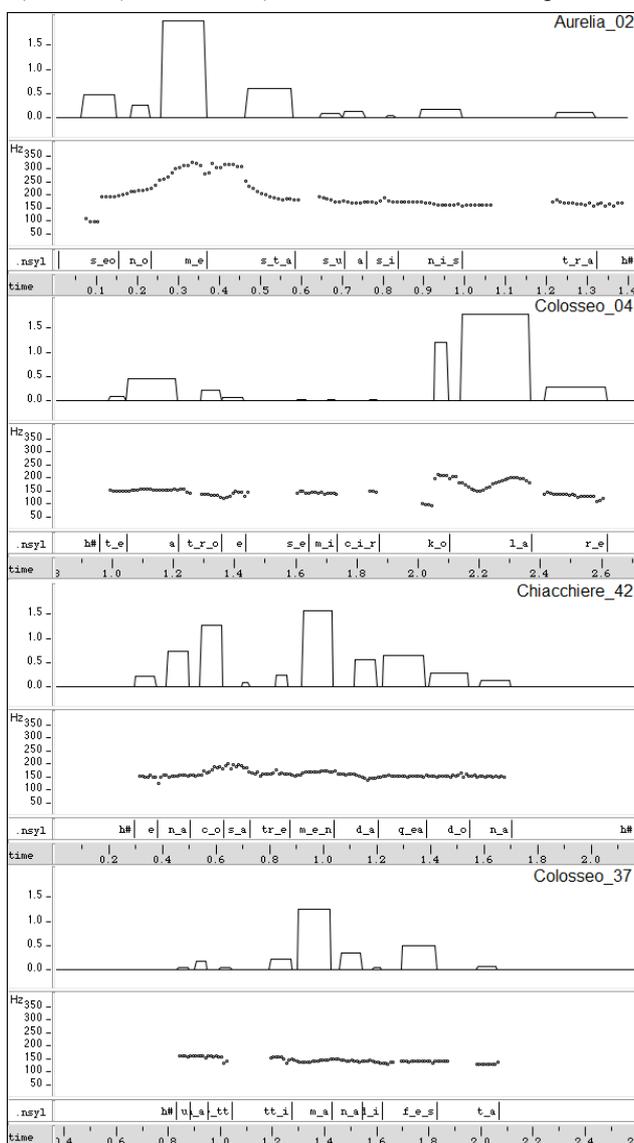


Figure 2: The prominence function profiles – *Prom* – and pitch profiles for some utterances considered in this study. Aurelia\_02: “*Secondo me<sub>T</sub> | stava sulla sinistra<sub>F</sub>*”. Colosseo\_04: “*Il teatro è semicircolare<sub>F</sub>*”. Chiacchiere\_42: “*E’ una cosa tremenda<sub>F</sub> | quella donna<sub>A</sub>*”. Colosseo\_37: “*Una settimana<sub>F</sub> | di festa<sub>A</sub>*”.

It is worth to note that a relevant number of the Main Prominences considered here (e.g. 14 samples out of the 47 extracted for this study from the “Bonvino” corpus) are supported mainly, or uniquely, by force-accents, as shown by the utterance Colosseo\_37 in Fig. 2, meaning that no intonational phenomena contributed to support them.

These regularities showed to be highly relevant also when testing them by the Fisher exact test.

## 4. Discussion

The results we obtained are by no means absolute. The matching between perception and measurement reveals strong tendencies, but it is never complete. In our opinion, when working on real corpora of spoken language, neat results where the prosodic patterns associated to Topic and Focus are perfectly consistent can only arise from *ex post* procedures, i.e. when measurement is made first, and then labeling is made on its basis. That is to say, when all utterances whose measurement gives the same pattern are given the same label (say, Topic-Focus; or Broad Focus; etc.). If labeling is made first on perceptual bases, some surprises are bound to come up when measurements are made.

However, from the results just exposed some provisional consequences can be drawn.

### 4.1 A functional interpretation: demarcation rather than culmination

As it can be seen in Table 3, the comparison between perceptual evidence about the utterances in the corpus and their automatic measurement made by means of our algorithm lead to the following results:

#### Topic-Focus

- the majority of utterances have the Main Prominence at the Right end of the Topic;
- a minority seems not to distinguish between the two units, with comparable Prominences.

#### Narrow Focus (at the Left)

- it is always marked by the Main Prominence at the Right of the Focus.

#### Broad Focus

- about half of the utterances have the Main Prominence at the Right;
- the other half have no Main Prominence, but several minor/equivalent Prominences.

In sum, only constituents located at the left of the utterance (Topic or Narrow Focus), and more precisely the right end of such constituents, seem to be steadily associated to the Main Prominence.

A possible explanation is the following: the primary function of the Main Prominence may be *demarcation*, rather than *culmination*. In other words, its first, immediate effect may be that of drawing a boundary between two information units, rather than “describing” one of them.

This doesn't mean that different intonation patterns

cannot express different kinds of Focuses and Topics, effecting different types of illocutions and pragmatic functions. But the *bare presence and position* of the Main Prominence (as it results from our measurements) may suffice to signal if the utterance contains a boundary between Information Units, and where. Then, once the Main Prominence has signaled a boundary between two units, for the recognition of which kind of units they are it is sufficient that the contour of the one located to the right signals if it is a Focus or an Appendix.

The minimal cues that can suffice to make the boundaries between information units recognizable to the addressee are shown in Table 4.

IS unit	beginning marked by:	end marked by:
Topic	beginning of utterance / intonational contour	MP on last stressed syllable of the Topic
Right Focus after T	MP on last stressed syllable of the Topic	end of utterance / intonational contour
Broad Focus	beginning of utterance / intonational contour	end of utterance / intonational contour
Narrow Focus (at the Left)	beginning of utterance / intonational contour	MP on last stressed syllable of the Focus, and beginning of Appendix flat contour
Appendix	MP on last stressed syllable of the Focus, and beginning of Appendix flat contour	end of utterance

Table 4: Minimal perceptual cues for the recognition of IS units.

This would provide us with a quite simple explanation of:

- Why Topics are marked more strongly than both Broad Focuses and Right Focuses after a Topic, though the communicative import of Focuses is greater than that of Topics: this is because Topics, unlike Right Focuses, are followed by another major Information Unit within the same utterance, so that the boundary between the two needs to be signaled.

- Why Narrow Focuses (at the Left) are also strongly marked: this is for the same reason, since also Left Focuses are followed by a boundary between Information Units within the utterance.

The explanation we propose is an *exquisitely structural* one, more precisely a “topologic” one, of how the Main Prominence (at least in some Italian varieties) may allow recognition of Information Units; i.e. an explanation based only on the *presence and position*, not on the quality of Prominence and intonation contours:

#### A Topologic Hypothesis on Main Prominence

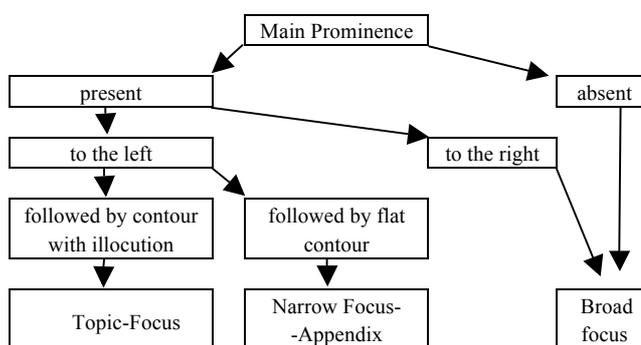
*"What is marked through the Main Prominence is the boundary between Information Units within the utterance."*

Strictly speaking, the only *qualitative* difference needed in order to recognize the Information Structure of an utterance is that between the marking of a Topic and the marking of a Left (Narrow) Focus, because both are

followed by another unit. That difference can be effected either by the different intonation contours of the following units (respectively a Right Focus or an Appendix), or (also, with some redundancy) by the specific intonational contours of the Topic and the Left Focus themselves.

The absence of a Main Prominence, or its being located on the last stressed syllable of the utterance, both signal a Broad Focus (not preceded by a Topic), whose boundaries in principle do not need to be signaled by a Main Prominence, since they match the boundaries of the whole utterance.

The steps by which the addressee can “compute” the Information Structure of an utterance are proposed in Scheme 1.



Scheme 1: Minimal steps for the recognition of IS units

	Utterances corresponding to the description	Utterances not corresponding to the description
Rome – Bonvino	40 (85.10%)	7 (14.90%)
Rome – Clips	46 (71.88%)	18 (28.12%)
Florence – Clips	42 (71.19%)	17 (29.81%)
Naples – Clips	43 (70.49%)	18 (29.50%)
TOTAL	170 (73.59%)	61 (26.41%)

Table 5: Foreseen vs. unforeseen results for IS acoustic realization in the corpus.

In this interpretation, speakers obey to a (non-)surprising extent to the *law of least effort*. The only elements strictly needed are (a) a Main Prominence per utterance, and (b) the difference between an “illocutionary” Focus contour and the contour of an Appendix, devoid of illocution. Now, since the different Focus contours are independently needed to express the different illocutions of utterances (i.e. the different linguistic acts), the specific cost required for expressing Information Structure is very low. Marking each information unit with a culminative Prominence would cost more effort than simply marking the boundaries, because:

- distinguishing Topic from Focus would require two different Prominences (one for each) instead of just one (at the boundary);
- distinguishing Broad Focus from Narrow Focus would require two recognizably different Prominences, because also Broad Focuses would need a “dedicated”

Prominence.

Instead, language prefers to work in a more economic way, namely marking only... the marked element (i.e. Narrow Focus). This situation is well represented in the corpus, as shown in Table 5.

But there is more, which we will expose in the next section.

#### 4.2 A continuum, rather than discrete alternatives

As it can be seen in Table 3 above, a minority of the utterances in the corpus that are perceived as Topic-Focus have no Main Prominence. And a minority of the utterances evaluated as Broad Focuses have an internal Main Prominence, in a position similar to that of Topic-Focus structures.

In other words, utterances acoustically measurable as Broad Focuses can be perceived as Topic-Focus, and vice versa. This can be explained: Topic-Focus and Broad Focus are not separate and reciprocally exclusive structures, rather the extremes of a continuum. The middle of the continuum is occupied by utterances where the boundary between the units is not neatly marked, and the distinction between the two possible Information Structures remains under- or unspecified.

In other words, *the speaker is not bound to decide* between Topic-Focus and Broad Focus. At least not prosodically, possible disambiguation remaining entrusted to pragmatic and contextual factors.

This is even more true if we consider that the speaker and the addressee can evaluate prosodic cues differently, and the speaker is always aware of this. As a consequence, (s)he knows in advance that the perception of IS may be subject to a certain amount of fuzziness.

More radically, there is no reason to think that a content must necessarily be either 100% or 0% focused. Instead, any content can be focused at an unlimited variety of degrees (Daneš, 1967, 1974; Firbas, 1966, 1987, 1989; Sgall 1975; Sgall *et al.* 1973), or even at a degree that simply remains underspecified.

Thus, no surprise if the Main Prominence is not always clearly recognizable. One should always expect for some utterances to have intermediate status between Topic-Focus and Broad Focus. And the status of a certain amount of information, typically “in the middle”, will remain uncertain.

In sum, Topic vs. Focus seems not to be a black & white story, rather one in a grey scale.

This is the case for the utterances in Figure 3.

The absence of a clear-cut distinction between Topic-Focus and Broad Focus corresponds to their being structures often possible in the same contexts, and to their often not influencing subsequent discourse in a decisively different way. Moreover, a general remark may be made: the fact that the categories of IS remain underspecified in actual communicative exchanges is not problematic at all, since the same obviously happens for

other aspects of the semantic/pragmatic interpretation of utterances.

For instance, if I say "the car was stopped by Tom", my addressee can perform any kind of free enrichment in interpreting my utterance, leading to different representations, such as Tom being the driver of the car, a policeman commanding to stop, an elephant crossing the road, etc. Even information less pragmatic in nature may remain unspecified. For instance, in many languages verbal tense can remain not overtly expressed, leading to different possible interpretations (often not totally disambiguated by the context) of the temporal coordinates of the event expressed by each utterance.

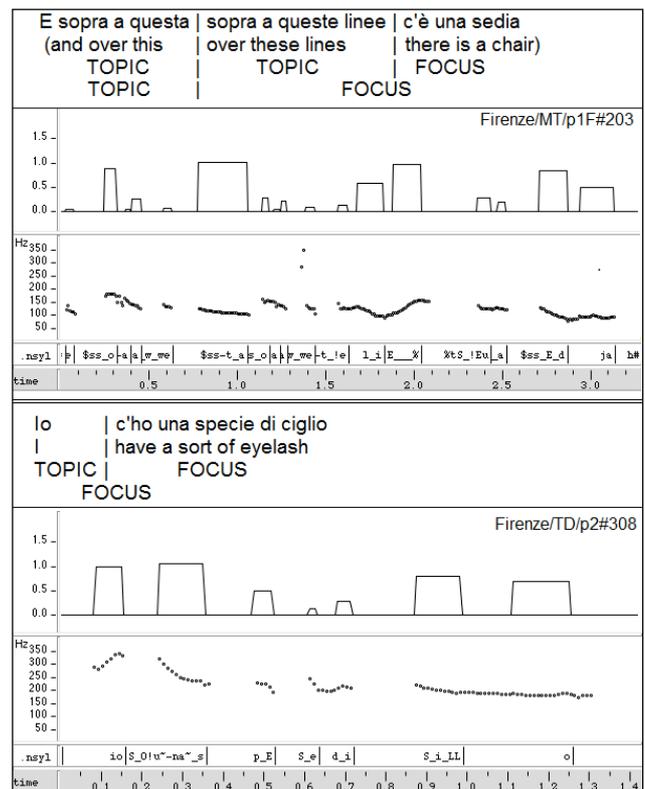


Figure 3: utterances underspecified between Topic-Focus and Broad Focus.

	Utterances corresponding to the description	Utterances not corresponding to the description
Rome – Bonvino	43 (91.49%)	4 (8.51%)
Rome – Clips	55 (85.94%)	9 (14.06%)
Florence – Clips	53 (89.83%)	6 (10.17%)
Naples – Clips	53 (86.89%)	8 (13.11%)
TOTAL	170 (87.88%)	28 (12.12%)

Table 6: Foreseen vs. unforeseen results for IS acoustic realization in the corpus (including the continuum between Topic-Focus and Broad Focus)

Even more obviously, the identity of the participants to an event may remain unspecified in languages where overt Subjects are not the rule and the Verb has no morphological marking for the Person. The following Japanese example contains both ambiguities:

*Tokyo-e ikimasu*

Tokyo-to go

"I/you/(s)he/we/they go/will go to Tokyo

Now, if we consider all cases in our corpus where Information Structure remains underspecified between Topic-Focus and Broad Focus as consistent with the model, we obtain the new figures depicted in Table 6.

This means that almost 90% of the utterances present one of the following matchings between their perceptive evaluation and the results of measurement:

- structures evaluated as Topic-Focus, with Main Prominence at the right end of the Topic;
- structures evaluated as Focus-Appendix, with Main Prominence at the right end of the Focus;
- structures evaluated as Broad Focus, either with no Main Prominence or with Main Prominence at the right end;
- structures evaluated either as Topic-Focus or as Broad Focus, with no evident Main Prominence.

Only in 10% of the cases, automatic measurement gave results where the Main Prominence had different positions. These can probably be considered as remaining "noise" in the procedure: the existence of a minority of cases with different patterns is expected, because (i) there reasonably must have been human errors in the first phase (assessing the distribution of Information Units in utterances through subjective sound perception and context evaluation), (ii) a certain amount of data are bound to be subject to the typical "flaws" of speech, such as imperfect production, changes of intention, etc., and (iii) the efficiency of the automatic algorithm in assigning prominence levels to the syllables cannot be 100%.

## 5. Conclusions

The following conclusions, based on the examined Italian varieties, can be drawn from the described experiments and their possible interpretation given above:

1. An abstract and merely structural, "topologic" level of Prominence can be conceived of, where its *mere location* is endowed with the function of demarcation between units, before (instead of?) that of their culmination and "description".  
This aspect of Prominence may suffice to explain much of the process by which speakers interpret the Information Structure of utterances in discourse. Further features, such as the specific intonational contours of the different Information Units, may thus represent a certain amount of redundancy.
2. Real utterances do not always signal the distribution of Topic and Focus clearly. Acoustically, many remain underspecified in this respect. This is especially true for the distinction between Topic-Focus and Broad Focus, which indeed often has no serious effects on the progression of communicative dynamism in the

subsequent discourse.

3. The consistency of such results with the law of least effort, and the very high percent of matching between perceptual evaluations and automatic measurement, seem to validate the used algorithm.

## 6. References

- F. Albano Leoni. 2003. Tre progetti per l'italiano parlato. In *Atti del XXXIV Congresso SLI*, Firenze, pages 675-683.
- C. Avesani. 2000. Costruzioni marcate e non marcate in italiano. Il ruolo dell'intonazione. In D. Locchi, A. Giannini, M. Pettorino, (eds.), *Atti delle X giornate di studio del GFS, Il parlante e la sua lingua*, pages 1-14.
- C. Avesani and M. Vayra. 2004. Focus ristretto e focus contrastivo in italiano. In F. Albano Leoni, F. Cutugno, M. Pettorino, R. Savy (eds.), *Il Parlato Italiano. Atti del Convegno Nazionale*, Napoli, pages 1-20.
- C. Avesani, M. Vayra, C. Zmarich, R. Paggiaro and D. Sperandio. 2007. Le basi articolatorie della prominenza accentuale in italiano. In V. Giordani, V. Bruseghini, P. Cosi (eds.), *Atti del III convegno AISV*, Trento, pages 1-22.
- P. Bagshaw. 1994. *Automatic prosodic analysis for computer-aided pronunciation teaching*. PhD thesis, University of Edinburgh.
- M.E. Beckman, J. Hirshberg and S. Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In S. Jun (ed.), *Prosodic models and transcription: Towards prosodic typology*, Oxford University Press, pages 9-54.
- D. Bolinger. 1958. A theory of pitch-accent in English, *Word*, 14:109-149.
- E. Bonvino. 2005. *Le sujet postverbal. Une étude sur l'italien parlé*, Paris, Ophrys.
- M. Breen, E. Fedorenko, M. Wagner and E. Gibson. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes*, 25 (7/8/9): 1044-1098.
- W. Chafe. 1987. Cognitive Constraints on Information Flow, in R.S. Tomlin (ed.), *Coherence and Grounding in Discourse*, Benjamins, pages 21-51.
- W. Chafe. 1992. Information Flow in Speaking and Writing. In P. Downing, S.D. Lima, M. Noonan (eds.), *The Linguistics of Literacy*, Benjamins, pages 17-29.
- E. Couper-Kuhlen. 1986. English prosody. Arnold, 1986.
- E. Cresti. 1992. Le unità d'informazione e la teoria degli atti linguistici. In G. Gobber, (ed.), *Atti del XXIV Congresso SLI*, Bulzoni, pages 501-529.
- E. Cresti. 2000. *Corpus di italiano parlato*, Firenze, Accademia della Crusca.
- F. Daneš. (1967), Order of Elements and Sentence Intonation. In *Studies to Honor Roman Jakobson*, The Hague-Paris, Mouton, pages 499-512.
- F. Daneš. 1974. Functional Sentence Perspective and the Organization of the Text. In F. Daneš (ed.), *Papers on Functional Sentence Perspective*, Prague: Academia /The Hague: Mouton, pages 106-128.
- M. D'Imperio. 2002a. Language-specific and universal

- constraints on tonal alignment: the nature of targets and anchors. In *Proc. Speech Prosody 2002*, pages 101-106.
- M. D'Imperio. 2002b. Italian Intonation: An overview and some questions, *Probus*, 14(1):37-69.
- G. Fant, A. Kruckenberg and J. Liljencrants. 2000. Acoustic-phonetic Analysis of Prominence in Swedish. In A. Botinis, (ed.), *Intonation*, Kluwer, pages 55-86.
- C. Féry and M. Krifka. 2008. Information structure. Notional distinctions, ways of expression. In P. van Sterkenburg (ed.), *Unity and diversity of languages*, Benjamins, pages 123-136.
- J. Firbas. 1966. On Defining the Theme in Functional Sentence Analysis. *Travaux Linguistiques de Prague*, 1: 267-280.
- J. Firbas. 1987. On the Delimitation of the Theme in Functional Sentence Perspective. In R. Dirven, V. Fried (eds.), *Functionalism in linguistics*, Amsterdam-Philadelphia, Benjamins, pages 137-156.
- J. Firbas. 1989. Degrees of communicative dynamism and degrees of prosodic prominence (weight). *Brno Studies In English*, 18: 21-66.
- M. Frascarelli. 2000. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Studies in Natural Language and Linguistic Theory 50. Kluwer.
- M. Frascarelli. 2004. L'interpretazione del Focus e la portata degli operatori sintattici. In F. Albano Leoni, F. Cutugno, M. Pettorino, R. Savy (eds.), *Il Parlato Italiano. Atti del Convegno Nazionale*, B06, Napoli.
- M. Frascarelli and R. Hinterhölzl. 2007. Types of Topics in German and Italian. In S. Winkler, K. Schwabe (eds.), *On Information Structure, Meaning and Form*, Benjamins, pages 87-116.
- B. Gili Fivela. 2006. Tonal alignment in two Pisa Italian peak accents. In *Proc. of Speech Prosody 2002*, pages 339-342.
- M.A.K. Halliday. 1989. *Spoken and Written Language*, Oxford, Oxford University Press.
- M. Heldner. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics*, 31: 39-62.
- C. Jensen. 2004. *Stress and Accent*. Phd thesis, University of Copenhagen.
- K.J. Kohler. 2005. Form and Function of Non-Pitch Accents. In *Prosodic Patterns of German Spontaneous Speech*, AIPUK, 35a: 97-123.
- K.J. Kohler. 2006. What is emphasis and how is it coded? In *Proc. Speech Prosody 2006*, Dresden, pages 748-751.
- D.R. Ladd. 1978. *The Structure of Intonational Meaning*, Indiana University Press.
- D.R. Ladd. 1996. *Intonational Phonology*, Cambridge University Press.
- Y. Lee and Y. Xu. 2010. Phonetic Realization of Contrastive Focus in Korean. In *Proc. of Speech Prosody 2010*, Chicago.
- E. Lombardi Vallauri. 2001. La teoria come separatrice di fatti di livello diverso. L'esempio della struttura informativa dell'enunciato. In *Atti del XXXIII Congresso SLI*, Napoli, pages 151-173.
- E. Lombardi Vallauri. 2009. *La struttura informativa. Forma e funzione negli enunciati linguistici*, Carocci.
- G. Marotta. 2008. Phonology or non phonology? That is the question (in intonation), *Estudios de Fonética Experimental*, Universitat Autònoma de Barcelona, XVII:177-206.
- P. Mertens. 1991. Local prominence of acoustic and psychoacoustic functions and perceived stress in French. In *Proc. ICPhS'91*, Aix-en-Provence, pages 218-221.
- J. Pierrehumbert. 1987. *The Phonology and Phonetics of English Intonation* (Ph.D. thesis 1980), Indiana University Linguistics Club.
- J.F. Pitrelli, M.E. Beckman, J. Hirschberg. 1994. Evaluation of Prosodic Transcription Labelling Reliability in the ToBI Framework, In *Proc. ICSLP'94*, Yokohama, pages 123-126.
- E. Selkirk. 1984. *Phonology and Syntax: The Relation between Sound and Structure*, MIT Press.
- P. Sgall. 1975. Conditions of the Use of Sentences and a Semantic Representation of Topic and Focus. In E. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge, Cambridge University Press, pages 297-312.
- P. Sgall, E. Hajicová, E. Benesová. 1973. *Topic, Focus and Generative Semantics*, Kronberg Taunus, Scriptor.
- A. Sluijter and V. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress, *J. Acoustical Society of America*, 100:2471-2485.
- B. Streefkerk. 1996. Prominent accent and pitch movements. *Inst. of Phon. Sciences Proceedings*, University of Amsterdam, 20:111-119.
- A. Syrdal and J. McGorg. 2000. Inter-transcriber reliability of ToBi prosodic labeling, In *Proc. ICSLP2000*, Beijing, pages 235-238.
- F. Tamburini. 2005. Automatic Prominence Identification and Prosodic Typology. In *Proc. InterSpeech 2005*, Lisbon, pages 1813-1816.
- F. Tamburini. 2006. Reliable Prominence Identification in English Spontaneous Speech. In *Proc. Speech Prosody 2006*, Dresden, PS1-9-19.
- F. Tamburini. 2009. Prominenza frasale e tipologia prosodica: un approccio acustico. In *Proc. XL Congresso SLI*, Vercelli, pages 437-455.
- P.A. Taylor. 2000. Analysis and Synthesis of Intonation using the Tilt Model, *J. Acoustical Society of America*, 107:1697-1714.
- D. Talkin. 1995. A robust algorithm for pitch tracking (rapt). In Kleijn, W. and Paliwal, K. (eds), *Speech coding and synthesis*, New York, Elsevier, pages 495-518.
- J. Terken. 1991. Fundamental Frequency and perceived prominence parameters, *J. Acoustical Society of America*, 87:1768-1776.