

# The Intonational Marking of Topical Salience in Spontaneous Speech Evidence from Spoken French

Anne Lacheret-Dujour

CRISCO, University of Caen, France

anne.lacheret@crisco.unicaen.fr

## Abstract

Our analysis of prosodic patterns in spontaneous speech in French is based on the hypothesis that intonation, far from being the reflection of syntactical and rhythmical organisation only, also depends on informative structure. This hypothesis is illustrated here by the role of prosodic markedness in the expression of topic. We propose a bottom-up modelisation of prosody, from signal processing (segmentation and labelling of prosodic units according to perceptual constraints) to phonological representation in terms of autonomous principles which reflect the intonational structure as it is co-constructed by speaker and hearer. A functional reading of our prosodic model is then proposed, regarding constraints derived from informative structure for topic construction.

## 1. Introduction

Prosodic modelling of languages in general, and of French in particular, is far from meeting with any general consensus, as is evidenced by the variety of approaches developed for the construction of models. Nevertheless, one point has the agreement of most researchers: the principle of autonomy. According to this principle, defended in particular by Standard Phonologies, a prosodic representation is constructed independently of content. However, in most models proposed today, based on top-down processing, prosodic constituents derive from an underlying syntactic structure considered to be primary in the tradition of Prosodic Phonology [4], [6], [12], [15]. This obvious contradiction between the initial intentions and the methods actually used inevitably leads to circular reasoning such as the postulate that prosodic and syntactic structures may be linked by certain relations even if they are not congruent. A situation like this can be clarified if one bears in mind the fact that prosodic forms participate fully in the segmentation and ordering of the message [14].

We therefore raise this question: can modularity be in the method when it is not in the object<sup>1</sup>? We shall answer in the affirmative and present an autonomous intonational model developed to map the prosodic functioning of spontaneous speech in interview situation. In practice, our model hinges around two modules constructed independently of any previous knowledge of the syntactic organisation of the utterances analysed<sup>2</sup>:

- A phonetic module, based on the segmentation of the sound continuum into intonational groups on the basis of perceptual cues: a global approach via perceptively controlled prosodic patterns.
- A derived phonological module which leads to the formulation of **internal principles** enabling us to discern an ordered intonational structure.

Our purpose is to present the semi-automatic method used for developing these two modules. A functional analysis of the prosodic objects which emerge from the processing is then proposed. We shall represent the frequent non-alignment between prosodic and syntactic structures as the result of instructions generated by the informative structure (SI)<sup>3</sup>. Our attention will focus on the prosodic marking of « topicalisation »<sup>4</sup>. We shall show:

- I. How prosody can become involved in topicalizing an element in contexts in which the syntactic constructions are neutral, thus supporting the hypothesis of [9]: « prosody encodes what syntax fails to mark ».
- II. How contexts of topical overmarking can be identified *a posteriori* by joint use of syntactic and prosodic resources.

A topicalization scale (or topical continuum) will finally be put forward, enabling us to map the various degrees of topic salience which can be evidenced via the typical prosodic and/or morphosyntactic constructions.

## 2. The prosodic module

For our purpose extracts of interviews carried out in French on Radio France International between 1998 and 1999 were selected. We analysed two hours' recording featuring 10 speakers (approximately 10 minutes' speaking-time per subject). The prosodic processing of the data includes a phonetic analysis and a derived phonological representation

---

<sup>3</sup> In a given language exchange, a sufficient background of common knowledge is necessary for speech to function, the SI is responsible for implementing the progression of shared knowledge. This function is actualised by different pragmatic operations (topicalisation, focalisation, comment) which make it possible to attribute various degrees of informative salience. (foreground, background and middle plan).

<sup>4</sup> By « topicalisation » we mean the marked instantiation of the object of the assertion, the **theme**, which can be defined on two levels, communicative and structural (referring to the psychological and positional starting-point of the utterance). By a **marking operation**, we mean the recourse to the various means made available by language (prosodic and syntactic-semantic) to adjust the degree of salience of an element in speech.

---

<sup>1</sup> See [13] for a discussion.

<sup>2</sup> The immediate constituents are not the primitive data of our prosodic model.

which links to the emerging acoustic and perceptive categories a set of organizing principles revealing an ordered and autonomous prosodic structure.

## 2.1. Phonetic processing

In spite of the heterogeneousness and apparent instability of the material under analysis [3], the phonetician must devise a segmenting and labelling strategy which transforms a continuum into segmented data. According to the hypothesis which underlies our work, a prosodic structure is basically actualised by a succession of intonational pitch movements which occur continuously in time and in frequency. In other words, the two parameters analysed are the fundamental frequency

modulations and the variations in duration. On these bases, the signal was segmented and transcribed into phonological words (lexical head-driven units whose final full syllable may carry tonic stress), the syllables carrying final stress prominence were segmented (marking of primary stress) and finally the different types of pauses (air intake, silence, or the hesitating “um”) were noted (see [7] for the presentation of the software programmes analysing the data). Upon conclusion of the processing, we consider that there are as many intonational groups (IG) as there are occurrences of final stress prominence.

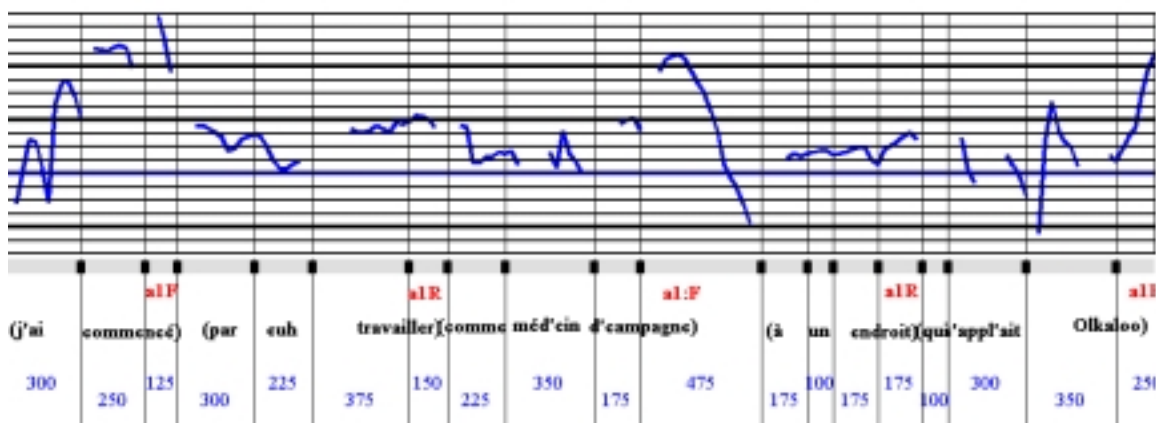


Figure 1: Sample viewing of a signal processed by ANALOR, with  $F0(st)$  on the Y-axis and on the abscissa the segments and their respective durations; *a1* markers mapping the final stress of the IG (rising: ‘*a1R*’, falling: ‘*a1F*’, the colon symbol marking the lengthening of a syllable).

## 2.2. Phonological processing

The purpose of phonological processing is to put forward an autonomous<sup>5</sup> formal representation of the prosodic structure which may then be used as a keystone for its linguistic interpretation. First, the pitch movements carried by the final syllables of the IG are described in terms of perceptively apprehended features. In order to achieve this, we drew our inspiration from the feature-driven outline coding method as expounded by [10] and from IPO [17] methodology. According to the theory underlying the latter method, prosodic categories are derived from categorial perception processes<sup>6</sup>; it is therefore possible to reduce the acoustic variation to a limited number of intonational segments<sup>7</sup>. These features are used to calculate degrees of stress for the final prominencies, which serve as anchorage points for two types of structuring processes: the segmenting of utterances into intonational periods (IP)<sup>8</sup> and the intonational organization of these. The first type of processing is carried out automatically according to the algorithm herewith:

Segmentation into periods occurs if and only if the following four conditions are fulfilled :

- Occurrence of a pause of at least 300 ms;
- Detection of an  $F0$  pitch movement attaining a certain amplitude, defined as the difference in height between the last  $F0$  extremum and the mean  $F0$  over the entire portion of the signal preceding the pause. The amplitude exceeds a threshold of 4 st;
- Detection of a “jump” defined as the difference in height between the last  $F0$  extremum preceding the pause and the first  $F0$  value following the pause (in excess of a threshold of approx. 3 st);
- Absence of « um » in the immediate vicinity of the pause.

It must be emphasized that the decision to recognize a cut does not depend on the exact values of the thresholds but on the order of their size. In other words, when one parameter is very slightly the threshold, segmentation can occur only if the other parameters have values distinctly above the threshold.

Next, to calculate the internal organization of the period and, notwithstanding the apparent linearity of the data, bring to light the integration and ordering processes of the IG, a generic dominance principle (Head-driven DOMI) was manually applied<sup>9</sup>. According to this principle, any IG dominates the IG immediately adjacent to it on the left if it is

<sup>5</sup> Independent of any constraint affecting segmentation into syntactic constituents.

<sup>6</sup> See also [1] for a clear presentation of perceptual approaches.

<sup>7</sup> See [2] for an application to the prosody of French.

<sup>8</sup> Represents the root of the prosodic structure (*i.e.* the field selected by prosodic categories).

<sup>9</sup> See [11] for the basic principles of our approach.

marked either by a terminal frequency excursion into the high-pitch level, higher by at least one tone, or by a dropping into the extra-low level, or if its terminal outline is of contrasting slope<sup>10</sup>, or, finally, if it responds to a progression principle marked by a lengthening by at least 50% of the final syllable compared to that of the preceding IG<sup>11</sup>. The application of DOMI results in the grouping of IGs into packages (PIG); otherwise, the intonational construction reveals a flat linear structure (LIG). The intonational structure internal to the period is thus obtained by the recursive left-right test of this inclusion principle (comparison of IG1 with IG2, of IG2 with IG3, etc.), the presence of the extra-low feature being the condition upon which the algorithm is stopped.

Table 1 : Example of internal segmentation of the period *J'ai commencé par travailler comme médecin de campagne à un endroit qui s'appelait Olkalo (I started by working as a country doctor at a place called Olkalo). Where SLO: grouping by slope reversal; PRO: grouping by syllable lengthening, DOM: grouping by intonational dominance of IG2 on high-pitch level (see fig.1).*

[[[[[j'ai commencé)] <sub>IG1</sub>	
(par travailler)] <sub>IG2</sub> ] <sub>PIG1</sub>	SLO
(comme médecin de campagne)] <sub>IG3</sub> ] <sub>PIG2</sub>	SLO,PRO
(à un endroit)] <sub>IG4</sub> ] <sub>PIG3</sub>	SLO
(qui s'appelait Olkalo)] <sub>IG5</sub> ] <sub>PIG4</sub>	DOM

### 3. Intonational-Syntactical alignment

If intonational periods may coincide with syntactic sentences<sup>12</sup>, our corpus nevertheless contains many cases of non-congruency which must be accounted for. Indeed, out of the 69 periods detected, 17 coincide with the sentence, be it simple (10 occurrences) or complex (7 occurrences). If, in the remaining cases, the prosodic structures can encompass several syntactic sentences in one and the same period, it also frequently happens that a sentence is broken up into several periods. We have pinpointed this phenomenon, disturbing at first sight, as a basis for our study of the topical marking of the syntactic subject, in some cases accompanied by locative framing.

#### 3.1. The prosodic configurations of the subject

The application of the phonological principles presented above shows 3 types of possible prosodic relations between the subject and the governing verb: (i) the two constituents are included in the same IG or the same PIG: (1) *un homme ne vaudra pas que sa femme aille travailler (a man will not let his wife go out to work)*, (ii) the two syntagmas are not linked by any relation of prosodic dependency but are included in the same period (LIG) : (2) *Les japonaises ne connaissent pas les odeurs sucrées (Japanese women are not familiar with sweet smells)*, (iii) thirdly, the prosody may violate the basic syntactic constraint according to which the subject as a compulsory constituent cannot on its own fulfil the prosodic selection criteria for an autonomous period: (3)

*le commissaire du district, le préfet du lieu ne voulait pas de femme non mariée (the district police superintendent, the local police chief did not want an unmarried woman).*

#### 3.2. The notion of topical continuum

Following the notion of “windowing of attention” proposed in [16], these various configurations, considered independently of syntax, bring to light a scale of degrees of thematic salience: the theme is neutral when it forms a PIG with the verb, the LIG construction indicates a movement towards topical marking (TM), and exoperiodicity (left dislocation of the subject in an autonomous period) constitutes an explicit TM.

Next, the combining of prosodic and syntactic factors enables us to pinpoint finer distinctions on our topical scale : theme saliency is heightened by the combination of prosodic and syntactic marking. Thus, the formation of a LIG combined with the use of the presentative structure in (4) *c'est un mouvement qui descend du nord vers le sud (it is a movement which travels down from north to south)* constitutes more salient thematic marking than (2). A higher level of topical marking is achieved when left exoperiodicity is combined with morphosyntactic dislocation with anaphoric reiteration in (5) *Il y a tout une activité féminine qui n'apparaît pas au grand jour (a large sector of women's labour escapes public attention)* indicating stronger topic saliency than in (3), etc. (table 2).

Table 2: Intonosyntactic configurations and topic marking (topic weight indicated in rising order)

Prosodic configuration	Syntactic marking	Topic weight
(1) PIG	-	0
(2.a) LIG	-	1
(2.b)	+	2
(3.a) EXO	-	3
(3.b)	+	4

This table illustrates the full relevance of linguistic redundancy : the more markers involved (both intonational and syntactic), the higher the topic weights. Nevertheless, if we compare contexts (2.b) and (3.a), we may postulate that this weight is greater in 3.a than in 2.b, although there is no syntactic marking. In other words, the exoperiodic construction seems, on its own, to be the ultimate mark of topic saliency, regardless of the syntactic structure associated with it.

### 4. Conclusions

In this article we expound our bottom-up method for the interpretation of the prosodic structure of spontaneous speech events, proceeding from phonetic processing to global utterance interpretation. This method has proved to be sufficiently enlightening to explain the non congruency between syntax and prosody, an explanation which must needs go beyond merely taking rhythmical constraints into account. The enunciative dimension indeed constitutes a vital frame for prosodic constructions: the intonational organization reflects relations which are not expressed by verbal syntax and the processing of information, which is inevitably achieved through the variable degrees of saliency in speech. To illus-

<sup>10</sup> See [10] for the presentation of the inclusion principle.

<sup>11</sup> See [8] for a detailed illustration.

<sup>12</sup> Unit constructed around a verbal nucleus, including its arguments and satellites.

trate this phenomenon, we have focused our attention on the various prosodic constructions of the theme; we have shown that the structural points of convergency between syntax and prosody and/or the non congruency between the two structures serves, in any event, the needs of the IS. We think that an approach such as this, giving rise to a scale of topicality, can be refined by taking into account, over and above the various degrees of prosodic segmentation (PIG, LIG, exo-periodicity), of other types of prosodic marking such as the amplitude and direction of the final pitch movements of themes). According to the theory we finally suggest, this approach can be generalized to the processing of the whole set of pragmatic categories associated with IS, since other scales can be constructed on the same model, distinguishing for example between the variable degrees of focalisation or, on the contrary, of parenthesisating.

## 5. References

- [1] Bailly, G.; Aubergé, V.; Morlec, Y., 1997. Des représentations cognitives aux représentations phonétiques de l'intonation. In *Polyphonie pour I. Fonagy*, J. Perrot, éd. Paris, l'Harmattan, 19-28.
- [2] Beaugendre, F., 1994. *Une étude perceptive de l'intonation du français*. Thèse de Doctorat, Paris XI.
- [3] Blanche-Benveniste, Cl., 1990, ed. *Le français parlé*. Paris : CNRS Editions.
- [4] Dell, F., 1984. L'accentuation dans les phrases françaises. In *Forme sonore du langage*, F. Dell et al., ed. Paris: Hermann.
- [5] Di Cristo, A., 1998. Intonation in French. In [6].
- [6] Hirst, D.; Di Cristo, A., 1998, eds. *Intonation Systems. A Survey of Twenty Languages*. Cambridge University Press.
- [7] Lacheret-Dujour, A.; Victorri, B., 2001. La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques. Forthcoming in *Verbum*, M. Charolles, ed.. Nancy : Presses universitaires de Nancy.
- [8] Lacheret-Dujour, A.; François, J., 2002. Circonstance et prédication en français parlé : contraintes sémantico-pragmatiques et corrélations prosodiques. Forthcoming in *Revue de syntaxe et sémantique*. Caen : Presses Universitaires de Caen.
- [9] Lambrecht, K., 1994. *Information Structure and Sentence Form*. Cambridge University Press.
- [10] Martin, Ph., 1981. Pour une théorie de l'intonation : l'intonation est-elle une structure congruente à la syntaxe ? In *L'intonation : de l'acoustique à la sémantique*, M. Rossi, ed. Paris : Klincksieck.
- [11] Mertens, P., 1987. *L'intonation du français : de la description linguistique à la reconnaissance automatique*. Thèse de Doctorat, Université de Louvain.
- [12] Nespor, M.; Vogel, I., 1986, *Prosodic Phonology*. Dordrecht : Foris.
- [13] Nölke, H., 1999, ed. *Approches modulaires : de la langue au discours*. Paris : Delachaux & Niestlé.
- [14] Rossi, M., 1999. L'intonation, le système du français. Paris : Ophrys.
- [15] Selkirk, E., 1984. *Phonology and Syntax: the Relations between Sounds and Structure*. Cambridge: MIT Press.
- [16] Talmy, L., 2000. *Towards a Cognitive Semantic Structuring System*. Cambridge: MIT Press.

- [17] T'Hart, J.; Collier, R.; Cohen, A., 1991. *A Perceptual Study of Intonation, an Experimental-Phonetic Approach to Speech Melody*. Cambridge University Press.

## Acknowledgments

I wish to thank Bernard Victorri for his invaluable help and proficiency in the development of ANALOR.