

A dynamic view of the prosodic structure: The example of French

Philippe Martin

EA 3967 CLILLAC-ARP, UFR Linguistique, Université Paris Diderot, France

<https://doi.org/10.36505/ExLing-2010/03/0028/000148>

Abstract

The sentence prosodic structure is traditionally viewed globally in intonation phonology, taking into account all the prosodic events at once from the beginning to the end, without explicitly integrating the fact that these events occur one after the other in a time sequence. Indeed, whereas the speaker can achieve some planning ahead while reading a written text, it is barely the case for the listener, who has to process the sequence of syllabic units perceived one by one along the time scale. It is suggested here that prosodic events are used as acoustic signals triggering partial processing and storage of the already perceived syllables, followed by a concatenation of linguistic units (normally stress groups) already stored. This paper looks in some details into the mechanism of this decoding process, focusing on the role of prosodic events in the specific case of French.

Key words: Prosodic structure, French, F₀ contour, cognitive process

Introduction

The formal representation of the prosodic structure in intonation phonology normally extends from the beginning to the end of the sentence. Whether hypothetical-deductive or inductive, prosodic theories using the concept of prosodic structure consider prosodic events (PE) statically as part of the sentence as a whole. For instance, for French, the autosegmental-metrical approach assigns specific tones to intonation phrases normally aligned on syntactic boundaries (Post, 2000; Jun and Fougeron, 2002), whereas other theories (Martin, 1975; Mertens, 1993) assign rising or falling melodic contours to stressed syllables according to some mechanism such as the contrast of melodic slope depending on the declarative or interrogative terminal contour.

Applied to the encoding and decoding processes performed by the speaker and the listener, most of these formal accounts of sentence intonation assume that the entirety of the prosodic structure of each sentence is known in advance. Whereas this might be true to some extent for read speech, where punctuation marks and visual word detection may give some indications on the prosodic structure to be associated with the text, it is less the case for spontaneous speech, except perhaps for short sentences containing only a few number of stress groups.

A dynamic cognitive view

From the point of view of the listener, prosodic events are perceived one after the other in function of time. Therefore their categorization (i.e. the identification of the class they belong, assuming of course that all PE are not identical) depends only on past and not on future events.

The prosodic structure reconsidered dynamically results from a process by which strings of syllables are hierarchically assembled thanks to the identification of each PE as belonging to classes known by the listener. This process involves 2 mechanisms: a) processing of the string of syllables into a linguistic unit such as a stress group followed by the storage of their processed unit and 2) concatenation of this unit with all units belonging to the same level (i.e. whose storage was triggered by same class of PEs).

Among all PEs occurring in French, only those located on the last syllable of stress groups take part in this mechanism, at the exclusion of events located on the first syllable of lexical words, which are treated as secondary stress (a stress group contains either a Noun, a Verb, an Adjective or an Adverb). This means that the identification of PEs implied in the dynamic prosodic structure elaborated by the listener involves the identification of a syllable as stressed in order to qualify the associated PE as a triggering signal to process to storage-concatenation mechanism.

The Storage-Concatenation process

The dynamic view leads to consider that each PE, instantiated by various contrasts of variation and melody height, duration, of intensity, or vocalic quality, appears as a signal triggering the processing and storage of syllables perceived since the appearance of the last prosodic signal belonging to the same class. This syllabic storage is accompanied by concatenation of the elements already present at the same prosodic “address”.

We can look into this mechanism in some details with the simple example schematized on Fig. 1 which gives a diagram of the storage-concatenation mechanism relative to a syllabic sequence $\sigma_1 \sigma_2 \dots \sigma_n$. In this sequence, some syllables are stressed and are therefore associated with a PE in the prosodic structure. When one of these stressed syllables is perceived, the PE attached to that syllable is identified by a grid of perception acquired by the listener.

In our examples, this grid has 4 classes and pertains to the knowledge of the linguistic system of French, but can easily be adapted to other systems as well.

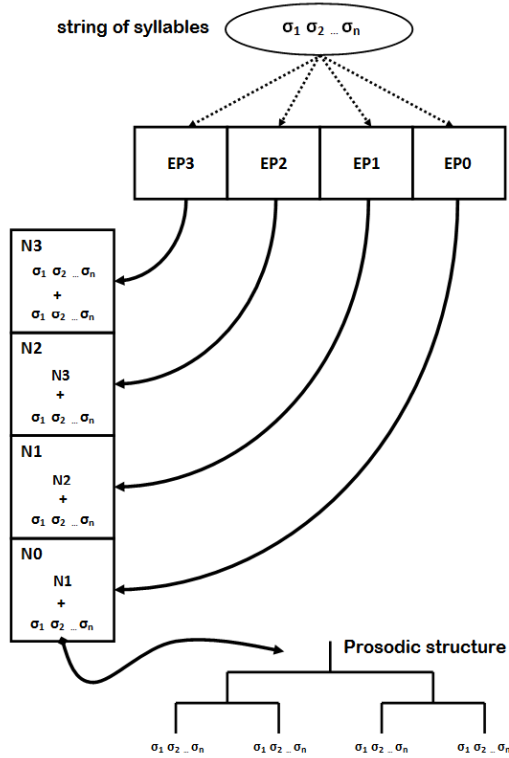


Figure 1. Schema of the storage-concatenation process

Each one of these identifications by the grid starts the following operation:

- 1) Storage of the syllabic sequence appeared since the last PE;
- 2) Conversion of the sequence into a linguistic unit (typically corresponding to a stress group, containing a verb, noun, adverb or adjective and some grammatical words depending on them);
- 3) Concatenation with the converted linguistic units stored at the lower levels, if they exist, and deletion of the existing strings of syllables already stored. This procedure implies a hierarchy between the PEs: $EP0 > EP1 > EP2 > \dots > EPn$.

In the example, EP1 assembles all the sequences stored in the lower memories N2, N3 taken in the order with the sequence of syllables appeared since the last PE. This operation can be generalized on all levels. We will call this general procedure Storage-Concatenation Process (SCP).

It also results from this dynamic cognitive view that prosodic marks enter a network of relative contrasts, and do not need invariant acoustic characteristics, as long as they are identified by the listener as belonging to the same class. Another argument going in this direction comes from the large variety of the styles of production of word. An emotionally depressed speaker will realize little or no melody variations. In whispered voice, contrasts between prosodic marks will have to be realized by other means than laryngeal frequency variations, with segment durations, about 50 to 70% higher than the equivalent achievements of duration in not whispered voice. There cannot thus be invariant characteristics for the prosodic marks, as opposed to what many of researchers in this field seem to admit.

Some cognitive consequences

The storage-concatenation dynamic view leads to a better understanding of some constraints governing the prosodic structure (Martin, 2009). First, the production of a string of syllables is not only constrained by the respiratory capacity of the speaker (Gilbert and al., 2007), but as well by the human immediate memory possibilities, which prevent stress groups to exceed some 7 or 8 syllables (Gilbert and al., 2010). Furthermore, the syntactic collision rule, which does not allow the stress groups whose syntactic units are dominated by distinct nodes in the syntactic structure to be concatenated by the prosodic structure (Martin, 2009) prevents the listener to build a misleading sentence structure at the first step of the concatenation process.

References

- Gilbert, A.C. Boucher, V.J. 2007. What do listeners attend to in hearing prosodic structures? investigating the human speech-parser using short-term recall", In INTERSPEECH-2007, 430-433.
- Gilbert, A.C., Boucher, V.J., Jemel, B., Lalonde, B. 2010. Segmentation de la parole : des groupes rythmiques et des énoncés, pas des mots et des phrases. Présenté au 78e Congrès de l'ACFAS, Montréal, Québec.
- Jun, S.-A., Fougeron, C. 2002. The Realizations of the Accentual Phrase in French Intonation, *Probus* 14, 147-172.
- Martin, Ph. 1975. Analyse phonologique de la phrase française, *Linguistics*, (146) Fév. 1975, 35-68.
- Martin, Ph. 2009. *Intonation du français*, Paris: Armand Colin, 256 p.
- Mertens, P. 1993. Intonational grouping, boundaries, and syntactic structure in French, in House, D., Touati, P. (eds.), *Proc. ESCA Workshop on Prosody, September 27-29, 1993, Lund (S) Working Papers 41* (Lund Univ., Dept. of Linguistics), 156-159.
- Post, B. 2000. *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics, 289 p.