Prosodic structure revisited: the need to disentangle rhythm from intonation

Elisabeth Delais-Roussarie

▶ **To cite this version:**

Elisabeth Delais-Roussarie. Prosodic structure revisited: the need to disentangle rhythm from intonation. Romanistisches Jahrbuch, 2022, 73 (1), pp.31-69. 10.1515/roja-2022-0002 . hal-04370523

HAL Id: hal-04370523
https://hal.science/hal-04370523

Submitted on 3 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Published in *Romanistisches Jahrbuch 2022 (ROJA 22)*, edited by Andreas Dufter, Folke Gernert, Daniel Jacob, David Nelting, Christian Schmitt, Maria Selig, Susanne Zepp.
ISSN : 1613-0413.

Prosodic structure revisited: the need to disentangle rhythm from intonation

Elisabeth DELAIS-ROUSSARIE[1]

Université de Nantes & UMR 6310 - LLING
Elisabeth.delais-roussarie@univ-nantes.fr

Abstract
It is well established that speech flow is segmented into prosodic chunks which are crucial for accurately interpreting utterances. Since the eighties, several proposals have been formulated, mostly within *Prosodic Structure Theory* and the *Autosegmental Metrical model*, to account for how prosodic chunks are derived and organized. These theoretical frameworks usually consider prosodic units as phonological primitives and domains within which segmental and suprasegmental phenomena are analysed. Due to the variability observed in the chunking obtained, which partly results from the various criteria used to derive these prosodic units, their exact nature and status can be questioned. In this contribution, it will be argued that prosodic structure and the chunks observed are best analysed as resulting from two types of structure-creating mechanisms, intonation and its relation to meaning on the one hand, and rhythm and its manifestations in languages on the other. This proposal is based on an analysis of French, a language with a strong syncretism between accentuation, metrical patterns and intonation.

Résumé
On admet généralement que le flux sonore est segmenté en groupes prosodiques permettant une bonne interprétation des énoncés. Depuis les années 80, plusieurs formalisations ont été proposées, notamment dans le cadre de la *Théorie Prosodique* et du *modèle Métrique Autosegmental*, pour rendre compte de la formation et de l'organisation de ces groupes. Généralement, ces modélisations considèrent les unités prosodiques comme des primitives phonologiques et des domaines au sein desquels les phénomènes segmentaux et suprasegmentaux sont analysés. En raison de la variabilité observée dans les segmentations obtenues, laquelle s'explique en partie par la multitude des critères utilisés pour dériver les unités prosodiques, leur nature et leur statut exacts peuvent être interrogés. Dans cette contribution, nous soutiendrons que la structure prosodique et les découpages observés gagnent à être analysés comme résultant de deux mécanismes créateurs de structure : l'intonation et sa relation avec le sens d'une part, et le rythme et ses manifestations dans la langue, d'autre part. Cette proposition est basée sur une analyse du français, une langue où il existe un fort syncrétisme entre accentuation, rythme et intonation.

1. Introduction
The speech flow is not continuous, but divided into chunks, which are usually described as structured and are represented in terms of prosodic structure. Prosodic units are mostly considered as phonological primitives within which segmental and suprasegmental phonological phenomena are described (*cf.*, among others, Selkirk 1978/1981, 1986; Nespor/Vogel 1986; Jun 1993 for analyses within *Prosodic Structure Theory* or the *Autosegmental Metrical model*). In so doing, we do not take into account the mechanisms at

[1] I wish to thank the audience at the conferences and seminars in which I intervened in Paris, Toulouse, Zürich, Nancy, Constance and Nantes this last ten years: their comments and suggestions on previous and uncomplete versions of the work presented here were stimulating. Thank you also to Corine Astésano, Bettina Braun, Nicole Dehe, Ingo Feldhausen, Cyrille Granget, Brechtje Post, Patrick Tomatis and Hiyon Yoo for fruitful discussions on issues related to this paper.

work in their formation which can be of a very different, even contradictory, nature (*cf.*, Delais-Roussarie 1996, 2000; Prieto 2005, 2006, among others). Recognising that these units may emerge on the surface from independently motivated constraints or mechanisms would allow us to take variability into consideration.

The examples in (1) illustrate the types of mechanisms that are involved. When one speaks, the sound stream is segmented by pauses and melodic movements in order to better convey the meaning of the utterance. The resulting units are closely related to other levels of linguistic description such as semantics, syntax and information structure (see, among others, Selkirk 1995a, Zubizarreta 1998, Steedman 2000). In (1), the NP adjunct *la semaine dernière* 'last week' may either be syntactically dependent on the first clause *les enfants sont partis en vacances* 'the children went on holiday' or on the second one *j'ai gardé leur chat* 'I took care of their cat'. Prosodic phrasing usually allows one to distinguish between the two cases by encoding the attachment of the NP adjunct. In (1a), a prosodic break (#), which is realised by a continuation rise (H*H%) on the syllable /njɛʁ/ followed by a pause, occurs at the end of the adjunct *la semaine dernière* yielding the interpretation: 'it was last week that the children went on holiday'. By contrast, in (1b), the pause and the continuation rise (H*H%) are associated with the syllable /kɑ̃s/, and the sentence means that 'the children went on holiday, and I took care of their cat last week'.

(1) Les enfants sont partis en vacances - la semaine dernière - j'ai gardé leur chat.

a. Les enfants sont partis en vacances la semaine dernière # j'ai gardé leur chat
H* H* H*H% H* !H* L%

b. Les enfants sont partis en vacances # la semaine dernière, j'ai gardé leur chat.
H* H*H% H* H- H* !H* L%

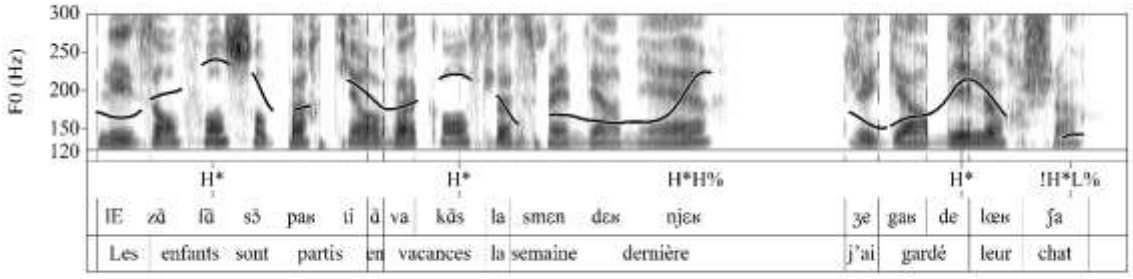


Figure 1. Spectrogram and F0 trace for example (1a).

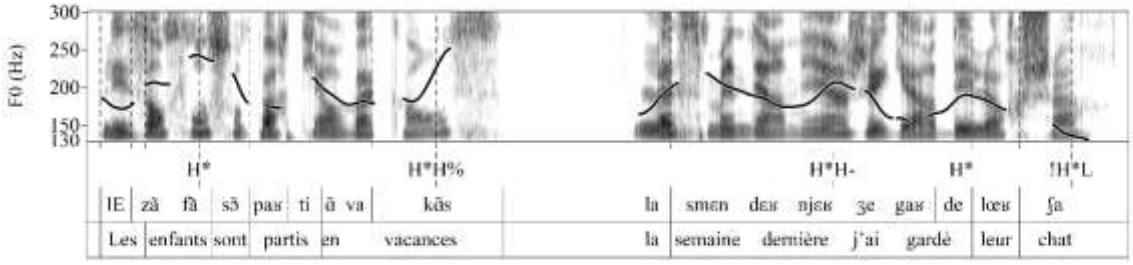


Figure 2. Spectrogram and F0 trace for example (1b).

However, prosodic events, such as the presence of an accent or a pause, may also be implemented so as to better structure speech flow and to anchor it in time; these events are then related to rhythm, which is based on two fundamental characteristics, periodicity and form (*cf.* Fraisse 1974). Different syllables are thus perceived as forming a group, and at the same time giving rise to a form or a pattern. The principles at work in rhythmic groupings respond to motor and cognitive constraints which go beyond strictly linguistic ones, even though these principles may take different forms depending on languages and their metrical structures (i.e. accentuation and rhythmic patterns). Note also that these structure-building principles operate in many fields such as music, poetry, and even dance (*cf.*, among others, Lerdahl/Jackendoff 1983; Cureton 1992). In addition to the structure accounting for the meaning conveyed, there is thus another structure divisible into two substructures, depending on the mechanisms at play: metrical principles governing alternation and grouping in languages, and rhythmic principles. While metrical principles remain abstract and language dependent, the rhythmical principles that derive from them result in an effective manifestation and a clear embodiment.

Considering the distinction between these structure-building mechanisms, I here defend the idea that the surface prosodic structure (i.e. the observed chunks) results from two types of grouping mechanisms which have to be clearly separated, and therefore described and represented according to their own principles and rules. The surface prosodic structure would then be the result of a merging or an alignment between both structures motivated by intonation and its relation to meaning for one, and by rhythm and its manifestations in language for the other. This proposal is reminiscent in some respects of that proposed by Selkirk (1984) and, for the French language, by Dell (1984).

The paper is organized as follows. In section 2, the dominant approaches concerning prosodic structure are briefly presented (i.e. *Prosodic Structure Theory* and the *Autosegmental-Metrical model*), in order to insist on the complex nature of the various units they postulate. The way these models have been adapted to French will also be discussed. Taking concrete examples in French from various corpora, section 3 will describe the mechanisms and principles that account for intonational patterns and meaning on the one hand, and metrical patterns on the other. Section 4 will explain how intonational structure and metrical structure can then be paired, giving rise to the observed prosodic groupings. Such a pairing process supposes a qualitative change given that the resulting structure is anchored in time and clearly embodied.

2. Theoretical background: prosodic structure and rhythm

In order to understand the need to distinguish the two different dimensions, it is necessary to present the main characteristics of prosodic structure propounded in theoretical writing, and more specifically in *Prosodic Structure Theory* (Selkirk 1978/81, 1986; Hayes 1989; Nespor/ Vogel 1986) and in the *Autosegmental Metrical Model* (Pierrehumbert/Beckman 1988; Ladd 1996/2008). It will also be useful to describe the various prosodic units and structuring levels and to show the differences that exist between those under and those above the level of the word, the former being essentially phonological in nature, and the latter deriving generally from morpho-syntactic and semantic information. In a second sub-section, it is shown that metrical principles, as represented in *metrical phonology* (*cf.* Liberman/Prince 1977; Hayes 1995 among others), should be more deeply integrated in the formation of various prosodic units. This will lead us to defend an approach similar to that proposed by Selkirk (1984). The section will conclude with a rapid review of the prosodic "models" proposed for French prosody, special

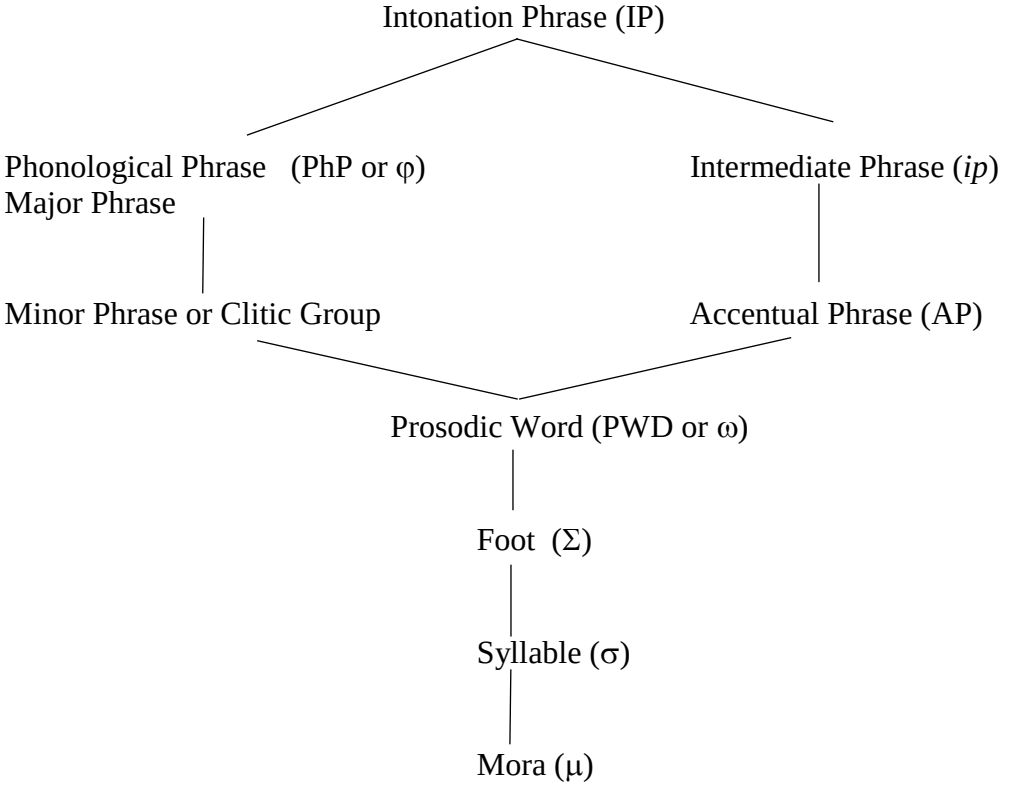
attention being given to the place attributed to metrical and intonational features in the definition of the various units.

2.1 Prosodic structure and prosodic units

For more than three decades, most studies accounting for the phrasing and intonational patterns associated with an utterance referred to a hierarchically organized prosodic structure as advocated in *Prosodic Structure Theory* (*cf.*, among others, Selkirk 1978/1981, 1986; Nespor/Vogel 1986) or in the *Autosegmental Metrical model* (*cf.*, among others Pierrehumbert/Beckman 1988; Ladd 1996/2008). The internal organization of the prosodic structure is represented in (2), where a distinction between both frameworks is shown.

(2) Prosodic hierarchy and prosodic units

a. *Prosodic Structure Theory* b. *Autosegmental-Metrical Model*



As one can see from (2), both hierarchies present similarities. For example, one may mention the fact that units are considered in both approaches as domains for the application of phonological phenomena, be they segmental, accentual or intonational. Moreover, some units are present in both frameworks: the intonational phrase, the prosodic word, and the three units under word-level, i.e. the foot, the syllable and the mora. As for differences, they mostly concern the name and number of potential levels of constituency between the prosodic word and the intonational phrase (one or two levels of phonological phrases)[2] as well as the way units are described. In some studies (*cf.*, among other, Selkirk 1978/1981, 1986; Nespor/Vogel 1986,

[2] In Selkirk (1978/1981, 1986), no distinction is made between minor and major phonological phrases, even if both phrases could be derived from syntactic structure by different parameters (X_{Head} and X_{Max}). Note also that reference to the *Clitic Group* is only made in Nespor/Vogel (1986) and Hayes (1989).

and Hayes 1989), priority is given to syntax-phonology mapping principles to characterize prosodic units above word-level, in particular phonological phrases, whereas in other studies such as Pierrehumbert/Beckman (1988), Jun (1993), Jun/Fougeron (2000) or Jun (2005), among others, intonation plays a crucial role in the definition of the various units. Note, however, that several experimental studies which compare the phrasing obtained in both approaches show that the predictions made by both of them are quite similar (*cf.*, among others, Hayes/Lahiri 1991; Frota 2000). This has thus led Frota (2012, 257) to say that "despite the clear differences in the underlying principles and definitions of the prosodic constituents, a closer inspection of the hierarchies […] reveals some striking similarities."

Concerning the well-formedness of the prosodic hierarchy, it has been argued that it should be strictly layered and thus conformed to the *Strict Layer Hypothesis* (Selkirk 1978/1981, 1986; Nespor/Vogel 1986). In order to account for the fact that this hypothesis embodied four distinct principles, it has been factored out by Selkirk (1995b) in four constraints in optimality theoretic terms (3):

(3) Constraints on Prosodic Structure (Selkirk 1995b, 443)

LAYEREDNESS:	No constituent C_i dominates a constituent C_j, j>i. Example: No syllable (σ) dominates a foot (Σ).
HEADEDNESS:	Any constituent C_i must dominate a constituent of level C_{i-1}, except if C_i is a syllable (σ). Example: A phonological word (ω) must dominate a foot (Σ).
EXHAUSTIVITY:	No constituent C_i immediately dominates a constituent C_j, j<i-1. Example: No phonological word (ω) immediately dominates a foot (Σ).
NON-RECURSIVITY:	No constituent C_i dominates a C_j, j=i. Example: No foot (Σ) dominates a foot (Σ).

Among these constraints, LAYEREDNESS and HEADEDNESS are universally respected in phonological representations. By contrast, the universality of EXHAUSTIVITY and NON-RECURSIVITY have often been challenged[3], which explains why the *strict layer hypothesis* was factored out as four different constraints as in (3). In intonation-based studies, the internal organization of the prosodic structure is usually less discussed, but some authors clearly argue in favour of a strictly layered structure (Jun p.c.), whereas others refer to a recursive structure to account for the realisation of some IPs, or at least to a more complex one integrating constituent conjunction (Frota 2000 for European Portuguese).

As to the units, they are of a very heterogeneous nature. The mora, the syllable and the foot can be considered as phonological constituents *per se*, since they are defined solely in phonological terms. On the contrary, the prosodic word and the phonological phrase are usually considered to be derived from the morpho-syntactic structure of the utterance, even if they are the domain of application of segmental and suprasegmental phonological phenomena (stress retraction, schwa insertion and phrasal accentuation in French, etc.). Concerning intonational phrase,

[3] Embedded prosodic phrases leading to recursion have been argued for at higher levels of structure in studies on French for instance (*cf.* Di Cristo/Hirst 1996 and Féry/Hörnig/Pahaut 2011).

which is referred to in both approaches, there has generally been little consensus on how to derive it, despite its universality, as pointed out by Ladd (1996: 235):

> "It is universally assumed that one of the functions of prosody is to divide up the stream of speech into chunks or phrases of one sort or another – for the moment, informally, let us call these chunks intonational phrases or IPs. Despite the apparent universality of the chunking function, however, IPs and IP boundaries are remarkably difficult to define and to identify consistently. IP boundaries seem to take on a bewildering variety of manifestations, from a clear pause accompanied by a local F0 fall or rise, to a subtle local slowing or pitch change that defies unambiguous definition. As a result, there is often disagreement about whether a particular IP boundary is or is not present, and definitions of IP boundaries in the literature are frequently circular or vague."

To sum up, even though some issues are not completely settled, the prosodic structure displays several features: (i) prosodic constituents are domains for the application of suprasegmental and segmental processes, be they above or under the word-level; (ii) the structure is usually flatter than the syntactic structure. As for the units themselves, they are of various natures: the syllable and the foot are phonological, whereas the prosodic word and the phonological (or accentual) phrase are related to the morpho-syntactic structure. Note nevertheless that each unit has to be strictly headed, in the sense that one of its constitutive elements can be analysed as a head (e.g. the nucleus in a syllable, the stress syllable in a foot, etc.). Concerning the units above the word-level, they are not isomorphic to the morpho-syntactic and information structures, but derived from them.

2.2 Prosodic units, metrical principles and rhythm

Among the prosodic units, some are constructed according to metrical principles which call for alternation and culminativity at each level. These principles may constrain the composition or size of feet, prosodic words, and even phonological phrases. Thus, in some languages, a foot may contain no more than two syllables or two morae, while in others there can be three; in addition, some languages limit the size of prosodic words and phonological phrases to respectively two feet or prosodic words. The purpose of these various limitations is to guarantee the alternation between strong and weak beats or groups, which can be understood as moraes, syllables, feet, or even higher-level units. In addition, the principle of culminativity is crucial to letting groups or chunks emerge from beats, strong beats being anchor-points that determine unit edges. We shall not here go into detail about the mechanisms used to build up metrical patterns from these two fundamental principles. Note however that the emergence of units such as feet, prosodic words and even phonological phrases can be derived from them. In metrical phonology (*cf.*, among other, Liberman/ Prince 1977; Hayes 1995), these principles are often translated into a limited number of parameters (e.g. right-headed *vs* left-headed foot, for instance; bounded *vs* unbounded foot) and rules (e.g. nuclear stress rule).

In metrical phonology, alternation principles have been formulated in a tree-like representation, often derived from the syntactic structure (see Liberman/Prince 1977 among others). It has also been proposed to apply metrical constraints on alternation within prosodic units (see, among others, Nespor/Vogel 1989) or to derive phonological phrases by taking into account metrical constraints. Thus, Delais-Roussarie (1996) and Prieto (2005, 2006) proposed, within *Optimality Theory* (Prince/Smolensky 1993), to analyse phonological phrasing, in French, Catalan and Spanish respectively, as resulting from the interaction of alignment and metrical constraints. Generally, although these various propositions allowed metrical factors to be more explicitly

considered in the formation of prosodic units, they did not radically question their status: these units remained phonological primitives!

Instead of using metrical trees, some authors (Prince 1983; Selkirk 1984; Dell 1984) proposed to account for metrical phenomena by means of a grid (grid-only model). This change opened the way to a different understanding of these phenomena, since structure and groupings are not primary in a grid-only framework, but emerge from the two metrical principles of alternation and culminativity. The consequences of this paradigmatic change have been clearly formulated by Selkirk (1984, 27):

> "It is necessary now, we believe, to reassess the claims for the existence of those supra-syllabic prosodic constituents, for it is clear that some of the phonological phenomena that were thought to provide motivation for these higher units of structure are better explained in terms of metrical grid alignment of the sentence. Some categories will disappear entirely from the prosodic repertoire; others will be given a much reduced role in phonological description, once the role of rhythmic structure in phonology is fully understood."

And a few pages later she adds (Selkirk 1984, 31):

> "The particular claim we are making about the prosodic constituent structure of phonological representations in English, then, is that the phonological phrase, the prosodic word, and the foot are not units in the hierarchy, but that the syllables and the intonational phrase are."

In addition to having consequences for the way in which certain prosodic units are considered, the grid-only approach makes it easier to draw a parallel between the metrical principles of alternation at work in languages and those observed in music (Lerdahl/Jackendoff 1983), poetry (Cureton 1992), and also in many other cognitive and biological activities (Goldbeter 2010). In fact, the principle of alternation, by leading to differentiation, generates structure, whatever the nature of the basic units (syllable or demi-beat in Selkirk 1984). This mechanism was clearly exposed by Fraisse (1974) who distinguished two types of grouping:

- Subjective groupings: rhythmic grouping is of subjective origin if a regular succession of identical sounds is perceived as forming a group (e.g. the ticking of a clock). In these cases, the grouping results from our perceptual activity;

- Objective groupings: the grouping is of objective origin if it is based on the periodic return of certain dissociable elements (e.g. occurrence of stressed syllables, strong beats in music, etc.).

Two additional points should be mentioned concerning the formation of metrical and rhythmic structures. First, structuring does not result from a single hierarchy, but emerges from several autonomous hierarchical principles. In their work on Western tonal music, Lerdahl/Jackendoff (1983) defined four rhythmic components that correspond to different structuring principles: metre, grouping, reduction and prolongation. In his study of rhythm in poetry, Cureton (1992) has taken up three of these components (metre, grouping and prolongation) and emphasized that they are the basis of all rhythmic perception and construction. In any activity perceived as rhythmic, at least one of these components must intervene, but in more elaborate rhythmic

constructions (music, poetry, or even literary prose), the three components intervene simultaneously. Rhythm is thus born of a double movement: the movement of each of the components, and the movement of the relations between the different components. As a result, Cureton describes rhythm as a "multidimensional" composition.[4] In languages, as in music, recurrent melodic patterns (e.g. repetition of a tonal movement, for instance) contribute to the emergence of rhythmical forms or patterns, as does the repetition of stressed syllables or duration patterns.

Second, the units that derive from the rhythmic alternation principle must be analysed first as metrical, in the sense that they correspond to subdivisions that can be represented by abstract templates (three- or four-beat meters in music, iambic or trochaic feet in languages and poetry, but also musical phrases, accentual groups, verses, etc.). In all cases, these forms or groups are in some way external to the speaker or listener. In order for these subdivisions to become rhythmic patterns or forms, it is necessary to transpose metrical indications into pulsations, as Gagnepain (2020, 26-27) makes clear in an essay on rhythm in musical interpretation:

> "Vouloir établir une distinction entre pulsation et subdivision peut sembler relever d'un débat purement théorique. En réalité, la différence est essentielle, et pourrait être exprimée par la formule suivante, pour réductrice qu'elle soit : la pulsation se vit, la subdivision se pense. (…)
> Donnée véritablement corporelle, la pulsation a pour vocation d'être transmise à l'auditeur ; ce qui implique qu'elle soit physiquement ressentie par l'interprète. "

In my view, the distinction between pulsation and subdivision also makes sense if one wants to account for observed prosodic chunking. At a relatively abstract level, metrical patterns must be generated from metrical and linguistic principles that specify how alternation and culminativity are instanciated within the lexical and morpho-syntactic units of any given language. But, the groupings actually produced are rooted in temporality and clearly embodied. Although it would be presumptuous to model this, I will make some propositions in this direction in sections 3 and 4, making a distinction between metrical patterns and rhythm. Rhythmic forms or patterns that emerge at the surface will be sketched in section 4, while the modalities of formation of abstract metrical patterns will be specified in section 3.2.

2.3. Prosodic and metrical units in French

Among the various studies on French prosody, few offer a comprehensive description of the different units of the prosodic hierarchy given in (2). To our knowledge, reference to a constituent of a level comparable to the foot is only made in the work from Hirst and Di Cristo (*cf.*, among others, Hirst/Di Cristo 1984; Di Cristo/Hirst 1996; Di Cristo 1998). Constituents at a level equivalent to or higher than the prosodic word are often described, but the terminology is often confusing. Indeed, for some authors, the *prosodic word* is named *rhythmic word* and, for others, the terms *prosodic word* and *rhythmic word* refer to a constituent close to the *accentual phrase* (or minor phonological phrase). Similarly, the terms *phonological phrase*, *accentual phrase* (or group) and *rhythmic group* can refer to the same level of structuring. These

[4] "(...) I assume that rhythmic structure is *componential*. Rhythmic perception is not monolithically hierarchical; it is composed of a small number of different hierarchies. Given this organization, the complexity of a rhythmic response can come from three different sources: (1) interactions among rhythmic forms (both horizontally and vertically) within the same rhythmic hierarchy; (2) interactions among rhythmic forms (both horizontally and vertically) across different rhythmic hierarchies; and (3) interactions among rhythmic forms and the presentation of phenomena by the perceptual medium." (Cureton, 1992, p.123)

differences in terminology may also be accompanied by differences in the way prosodic units are understood. In fact, in many works, prosodic units are characterised according to the intonational patterns associated with them, even if they are the domain in which metrical phenomena (such as stress assignment or stress retraction) apply.

In much of Hirst and Di Cristo's work (Hirst/ Di Cristo 1984; Di Cristo/ Hirst 1996 among others), the units used to analyse French prosody are frequently described in terms of melodic or intonational patterns. The foot-equivalent unit (i.e. the tonal unit *UT*), for example, is described according to its tonal form LH. As for the intonational unit, corresponding to the IP, it is given an LL or LH pattern depending on whether it is in an utterance terminal position or not (Di Cristo/ Hirst 1996). As for the accentual phrase (AP) (or phonological phrase), it is frequently described in terms of the tonal pattern associated with it, i.e. LHLH (*cf.* Jun/ Fougeron 2000; Delais-Roussarie et al. 2015 among others). On the whole, many units are described on the basis of its tonal structure, even those that Selkirk (1984) questions, judging that they would benefit from being treated on metrical bases (*cf.* § 2.2). Many authors justify this position on the basis of the existence of a syncretism between accentuation and intonation in French (cf., Di Cristo 2016 for a review on this issue). However, this idea of syncretism is disputable, and even confusing. For example, there is often confusion between different categories of 'intermediate phrases' which have different metrical and intonational characteristics (*cf.* on this point § 3.2 as well as Delais-Roussarie/Post/Yoo 2020).

In parallel to the work on intonation, some studies on French prosody focus on metrical phenomena and define prosodic units. In Pasdeloup (1992) and Delais-Roussarie (1995, 1996), constituents are formalised relatively to accentual phenomena, in particular the distribution of stressed syllables. Dell (1984) and Delais-Roussarie (2000) are pursuing the same goal, but within a grid-only approach. It is nevertheless important to note that most of these studies do not refer to intonation. Therefore, higher order units, which are named rhythmic sequences (*sequence rythmique*) in Pasdeloup (1992) and Delais-Roussarie (1995) often refer to syntactically defined major phonological phrases.

We have seen so far that most units used in the prosodic models of French are characterised either by tonal information or by accentual and metrical phenomena, but few works have integrated the two aspects of structure creating mechanism (*cf.*, among others, Post 2000, 2011; Mertens 2008). In the following section, I propose a model which integrates both metrical information and intonation in the construction of prosodic units. In addition, this model pays a special attention in clearly differentiating metrical mechanisms from other mechanisms at work in the formation of the observed surface prosodic chunks.

3. Toward a new approach

As previously mentioned, it is important to propose a prosodic model which allows, at least at an abstract level, to distinguish prosodic phrases resulting from metrical principles – even though they may also be constrained by the morpho-syntactic structure – from those associated with intonational contours and derived from information structure and discourse. In this section, I propose such a model for French. It should allow accounting for intonational and metrical structures separately, both being built in parallel. They then will be in some ways unified at the surface level (see § 4). This proposal is inspired by the work of Selkirk (1984) and Dell (1984), in which it is argued that among the various units of the prosodic structure only the syllable and the intonational phrase are necessary, metrical patterns being derived within a metrical grid (see also § 3.2). The characteristics of intonational phrases and the principles and rules for their

derivation are presented in section 3.1. As for the phrases that emerge from metrical constraints, they are described in section 3.2, the mechanisms at play in their emergence being formalised in metrical grids.

3.1 Intonational units in French

In French prosody theory, the terms *intermediate phrase, intonational phrase, groupe intonatif, unité intonative* and *syntagme intonatif* refer to a wide variety of constituents. Some of them would not even be considered as intonational phrases on the basis of the prosodic events that apply in them or of the relations they entertain with the morpho-syntactic and information structure. In fact, long subject NPs are sometimes considered as intonational phrases (*cf.*, among others Mertens 1987, 1992 or Verluyten 1982), and sometimes even single syllables uttered in isolation and followed by pauses as in (4), see Portes/Bertrand (2011).

(4) *poliment*
uttered as [[po]$_{IP}$ [li]$_{IP}$ [ment]$_{IP}$]$_{IP}$

As to *intermediate phrases*, they refer to very different phrases in the prosodic and linguistic structures: they can correspond to long subject NPs (*cf.*, among others, Michelas/ D'Imperio 2010; Michelas 2011) as well as to syntactic constituents corresponding to topic or focus phrases (*cf.* Jun & Fougeron 2000, Delais-Roussarie et al. 2015, for instance). Even though these different types of intermediate phrases share phonetic features (resetting, lengthening of the last syllable, form of the final melodic movement), they should not be treated as similar. It is indeed source of confusion because they behave differently. It is therefore preferable to define intonational and intermediate phrases on a clear basis and to avoid circularity as much as possible. Consequently, the definitions retained here are in line with the propositions that have been made to overcome circularity and vagueness in the derivation of intonational phrases. Imprecise criteria related to size or length as well as phonetically-driven features are abandoned in favour of mapping rules and principles (*cf.*, among others, Selkirk 2005; and for French, Mertens 2008; Delais-Roussarie/Post 2008).

Following previous work (Delais-Roussarie 2021; Delais-Roussarie/Post/Yoo 2020), the proposition I defend here is based on a distinction between two types of intonational phrases: major intonational phrase (MaIP) and minor intonational phrase (MiIP).

Thus, any independent proposition or clause (5), any elliptical sequence (6) or any syntactic constituent containing the informational focus (7) is wrapped within a major intonational phrase (MaIP).

(5) a. *Le président serbe est arrivé dans la matinée.*
[(le president serbe)$_{AP}$ (est arrive)$_{AP}$ (dans la matinée)$_{AP}$]$_{MaIP}$

b. *Où est-ce que tu vas, comment tu vas y aller et à quelle heure tu vas rentrer*?
(from Delais-Roussarie et al. 2015: 68)
[(où est-ce que tu vas)$_{AP}$]$_{MaIP}$ [(comment)$_{AP}$ (tu vas y aller)$_{AP}$]$_{MaIP}$ [(et à quelle heure)$_{AP}$ (tu vas rentrer)$_{AP}$]$_{MaIP}$

(6) A: *Qui est venu hier*?
B: *Pierre*. [Pierre]$_{MaIP}$

(7) a. A: *Qu'est-ce que vous voulez?*
B: *je voudrais des oranges, s'il vous plaît madame*.
[[(je voudrais)$_{AP}$ (des oranges)$_{AP}$]$_{MaIP}$ (s'il vous plait)$_{AP}$ (madame)$_{AP}$]$_{MaIP}$

b. A: *Vous voulez des citrons?*
B: Non, ce sont des oranges que je veux.
[non]$_{MaIP}$ [[(ce sont des oranges)$_{AP}$]$_{MaIP}$ (que je veux)$_{AP}$]$_{MaIP}$

Analysing these different sequences as MaIPs is justified based on the intonational contours that occur at their right edge. In fact, only terminal contours are observed at the edge of such phrases. Here terminal contours refer to those presented in Table 1 (*cf.* also Post 2000).[5]

	Rising	Rising-falling	Falling	Falling after penultimate peak
Terminal contours	LH*H% 	LH*L% 	L*L %, !H*L% 	H+!H*L% 
Non-terminal contours	LH*H- 			

TABLE 1 : Terminal and non-terminal tonal contours in French

This analysis may seem obvious for examples (5a,b) and (6), but probably less obvious for those in (7). But, as Martin (1987), Delais-Roussarie/Post (2008) and Delais-Roussarie/Rialland (2005) have shown, the contour realised at the right edge of any IP containing the informational focus is a terminal contour, which is then generally copied at the end of the utterance (e.g. after *veux*). Note that the sequence '*que je veux'* corresponds to an appendix (or a tail) and receives the intonation observed in post-focus sequences (*cf.*, among others, Di Cristo/ Jankowski 1999; Jun/Fougeron 2000 and Beyssade et al. 2004,). Moreover, such a sequence may be analysed as forming an incomplete IP, named *segment d'IU*, by Di Cristo/Hirst (1996).

As far as Minor Intonational Phrases (MiIP) are concerned, they are aligned with the right edge of syntactic phrases with a specific informational status and entering in certain constructions (Dislocated XPs, appositions, parentheticals, etc.) (*cf.*, among others, Mertens 2008; Delais-Roussarie/Yoo/Post 2011; Delais-Roussarie 2021 for a similar position). Despite the great variability in the way IP boundaries located at the right edge of syntactic phrases within specific constructions (dislocations, parentheticals) are phonetically realized, there are two distinct arguments in favour of such a phonological analysis. First, these boundaries cannot be easily erased in order to allow restructuring within larger phrases for eurhythmicity, in contradistinction to boundaries occurring after NP subjects (cf. § 3.2 and also Delais-Roussarie 1996, 2000, 2021). On this issue, the analysis of dislocated XPs or adjuncts in subordinate

[5] A parallel can be drawn with Delattre 1966. Among the contours he referred to, *echo* and *parenthèse* are associated with parentheticals (or *segments d'UI*) and do not refer to any of the contours in table 1. In addition, the *minor continuation* has to be analysed as non-terminal. So, would refer to terminal contours all the other contours, which would be grouped according to their form, i.e. rising, falling, etc.

clauses is very interesting. In (8) and (9), the strongest break within the utterance occurs at the MiIP boundary, i.e. after the PP adjunct *dans ce role-là* in (8) and the dislocated NP *un enseignant* in (9), despite their syntactic embedded position.

(8) Mais je crois que dans ce rôle-là // il y a des choses à dire (ACSYNT, COAP)

[[(mais je crois)$_{AP}$ (que dans ce role-là)$_{AP}$]$_{MiIP}$ (il y a des choses à dire)$_{AP}$]$_{MaIP}$

(9) J'ai le sentiment qu'un enseignant il est quand même dans cette idée (Avanzi 2012 :162)

[[(j'ai le sentiment)$_{AP}$ (qu'un enseignant)$_{AP}$]$_{MiIP}$ // (il est quand même)$_{AP}$ (dans cette idée)$_{AP}$]$_{MaIP}$

The second argument relies on the fact that the form of the contour occurring at the edge of MiIPs is constrained by the contours occurring at the end of the MaIPs in which they are embedded. There is thus a dependency relation between MiIPs and MaIPs, which does not occur in cases of rhythmically-driven phrases (*cf.* also Delais-Roussarie/Post/Yoo 2020 on this issue).[6] Consider for instance (10). The tonal contour occurring at the end of the frame topic *quand je reste trop longtemps à Carcassonne*, which is wrapped within a MiIP, is falling, as a rising contour is realised at the end of the MaIP aligned with the right edge of the clause from which the MiIP depends.

(10) *Quand je reste trop longtemps à Carcassonne, j'en ai ras le bol, mais*..(From Delais-Roussarie et al. 2015: 73)

[[(quand je reste)$_{AP}$ (trop longtemps)$_{AP}$ (à Carcassonne)$_{IP}$]$_{MiIP}$ (j'en ai ras le bol)$_{AP}$]$_{MaIP}$ [mais…]

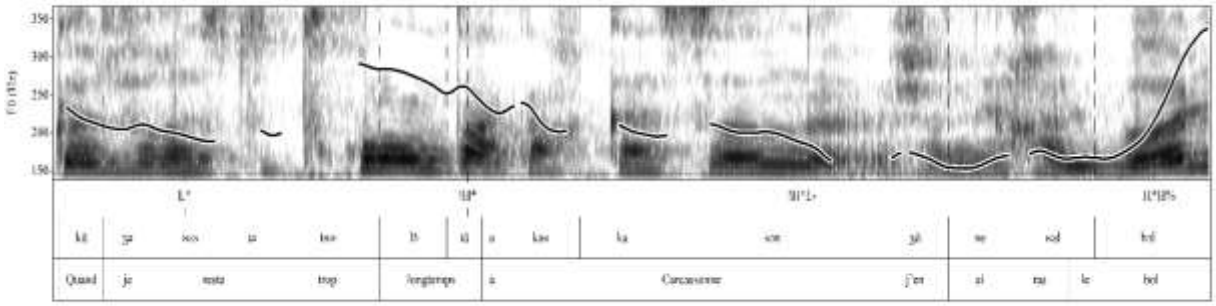


Figure 3. Spectrogram and F0 trace for example (10).

In (11), a sentence very similar to (10) but uttered in isolation and not in a chain of clauses, the contour realised at the end of the utterance (i.e. MaIP boundary) is falling, so that the contour at the end of the MiIP *quand je reste trop longtemps à Carcassonne* is rising.

[6] The dependency relation that applies between tonal contours corresponds to what is argued for by Martin (1987), among others. Note also that the falling realisation of the tonal contour at the end of *Carcassonne* in (10) can be considered as an instance of a minor continuation in Delattre's terminology. It was said in Delattre (1966) that such a rising contour may be realised falling.

(11) *Quand je reste trop longtemps à Carcassonne, j'en ai ras le bol.*

[[(quand je reste)$_{AP}$ (trop longtemps)$_{AP}$ (à Carcassonne)$_{IP}$]$_{MiIP}$ (j'en ai ras le bol)$_{AP}$]$_{MaIP}$

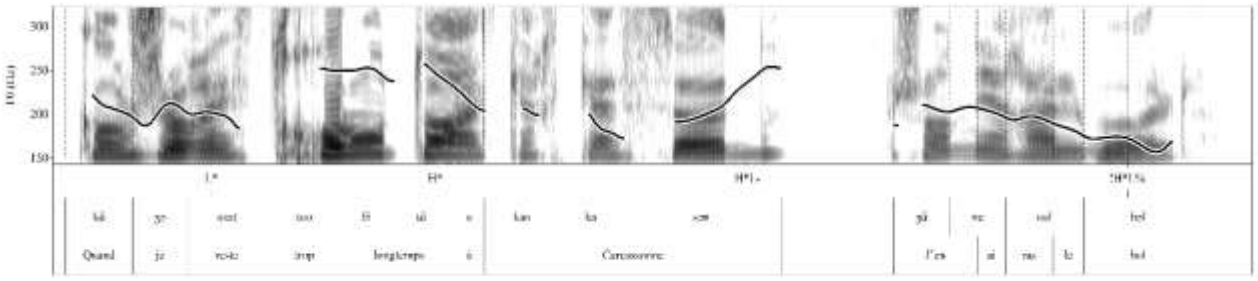


Figure 4. Spectrogram and F0 trace for example (11).

Thus, the form of the tonal contour occurring at the right edge of minor IPs is partly determined by the form of the terminal contour at the end of the major IPs within which the minor IPs are embedded, even if a rising contour may sometimes be realised as a default contour. I therefore propose to analyse MiIPs, whose right boundary aligns with the edge of particular constructions (frame-topics, dislocated NPs, etc.) as a recursive subcategory of major IPs. A parallel could then be drawn with that proposed to distinguish minor from major phonological phrases by Ito/Mester (2013), *cf.* Delais-Roussarie/Post/Yoo (2020) for arguments in favour of such an analysis.

To sum up, I propose to restrict the intonational structure to two levels of embedded phrases, the minor intonational phrase (MiIP) and the major one (MaIP). They are both derived from informational and discourse-based features. MaIP boundaries coincide either with the right edge of the syntactic phrase containing the informational focus in case of narrow focus or with the right edge of independent clauses, be they uttered in isolation or in a chain. As for MiIP boundaries, they are aligned with the right edge of syntactic phrases entering into specific discourse constructions (dislocated XPs, Topic phrases, parentheticals, etc.). Regarding the form of intonational contours, a terminal contour, whose form is often related to the meaning of the utterance, is realised at the end of any MaIP, whether it be located at the end of a clause or a syntactic phrase containing the informational focus, whereas a non-final contour, whose form is often derived from the form of the terminal contour, is realised at the end of any MiIP. These two related phrasal categories, which can be considered as recursive, are sufficient for the interpretation of any utterance.

3.2. Metrical principles and rhythmic patterns

In studies on French prosody, metrical patterns are often described within prosodic phrases, more specifically the accentual or phonological phrase (*cf.*, among others, Pasdeloup 1992; Delais-Roussarie 1996; Post 2000). Moreover, accentuation or phrasal stress are often considered as playing a crucial role in intonation, because of the strong syncretism between intonation, accentuation and phrasal stress (*cf.*, among others, Di Cristo/ Hirst 1996; Di Cristo 1998, 2016; Post 2011). Indeed, stressed syllables are often realised by means of a change in melodic height. Such tonal features can be seen as a way to phonetically distinguish stressed syllables from unstressed, just as much as lengthening. Phonetically, and even perceptually, this melodic change differs from a tonal contour in the sense that it is more static (*cf.* on this issue

Di Cristo 1999, 2016). However, in spontaneous everyday speech stressed syllables are often marked by syllabic lengthening alone.

For all the reasons mentioned in § 2.2 regarding rhythmic and metrical patterns, a grid only framework will be used to account for metrical phenomena. Such an approach has two advantages: (i) metrical phrases – which are named rhythmic words, rhythmic groups, accentual phrases, phonological phrases in the literature on French prosody – are considered as emerging at the surface level from metrical and alignment principles; and (ii) variability that occurs in the rhythmical parsing of an utterance is more easily treated (*cf.* also Dell, 1984; Delais-Roussarie, 2000). Before presenting the metrical and alignment principles used to build up the grid, it is worth noting that such a representation remains very abstract and does not really capture the way rhythmic structures emerge concretely (see § 2.2). Indeed, studies in cognitive sciences and in psychology insist on the fact that rhythm is dynamic, anchored in temporality, and, for the speaker and the listener, clearly embodied, whereas grid representation remains relatively static.

Analyses of different corpora (*cf.*, among others, Delais-Roussarie 1996, 2000) showed that in 40% of the cases the phrasing obtained did not correspond to that predicted by deriving accentual phrases (AP) by morpho-syntactic parameters only, i.e. the $]_{XHead}$ parameter in French (i.e. by aligning AP right edge with the right edge of heads of syntactic constituents, *cf.* Selkirk 1986). Differences occurred in the following cases:

- Derived accentual phrases with less than three syllables: in the whole corpus, 240 of the derived accentual phrases had less than three syllables (49 monosyllabic and 191 bisyllabic), but in the speaker's productions, only 69 APs had less than three syllables (4 monosyllabic and 65 bisyllabic). See table 2 (from Delais-Roussarie 2000).

Texts	**Derived phrasing (AP $]_{XHead}$)**		**Phrasing observed in APs**		**AP $]_{XHead}$ realized in the data (in percent)**
	Mono-syllabic φ	Bi-syllabic φ	Mono-syllabic φ	Bi-syllabic φ	
Le Petit Prince	21	121	2	32	25 %
Newpapers	20	40	2	22	40 %
France Info	8	30	0	11	29 %
Σ	**49**	**191**	**4**	**65**	**28,75 %**

TABLE 2 : short AP distribution, with a distinction between derived and observed one

- Derived accentual phrases with more than six syllables: 12 derived APs contained more than six syllables (7 were composed of 7 syllables, 5 of 8 syllables and more). In the observed data, only 3 accentual phrases had more than six syllables (7 syllables). See table 3 (from Delais-Roussarie 2000).

Texts	**Derived phrasing (AP $]_{XHead}$)**		**Phrasing observed in APs**		**AP $]_{XHead}$ realized in the data (in percent)**
	7 syllables	8 syllables and more	7 syllables	8 syllables and more	
Le Petit Prince	4	4	2	0	25 %
NewsPapers	2	1	1	0	33 %
France Info	1	0	0	0	0 %
Σ	**7**	**5**	**3**	**0**	**25 %**

TABLE 3 : long AP distribution, with a distinction between derived and observed one

A closer look at the differences showed that three distinct strategies were used by speakers in cases of differences between derived and observed phrasing:

- in some cases, it appears that the $]_{XHead}$ parameter derives more accentual phrases than actually observed. Restructuring thus occurs as shown in (12a, b, c). The analysis of the data showed that this was frequently observed when derived accentual phrases had less than three syllables. Restructuring can thus be seen as a mechanism to avoid stress clash.

(12) a. *Tout au plus pouvait-on dire à Brejnev ce que les parents faibles disent à leurs enfants*. *(le Monde)*

Predicted phrasing: (ce que les parents)$_{AP}$ (faibles)$_{AP}$ (disent)$_{AP}$ (à leurs enfants)$_{AP}$

Observed phrasing: (ce que les parents faibles)$_{AP}$ (disent à leurs enfants)$_{AP}$

b. *Et la grande personne était bien contente de connaître un homme aussi raisonnable*. *(Le Petit Prince)*

Predicted phrasing: (Et la grande personne)$_{AP}$ (était)$_{AP}$ (bien contente)$_{AP}$ (de connaître)$_{AP}$ (un homme)$_{AP}$ (aussi raisonnable)$_{AP}$

Observed phrasing: (Et la grande personne)$_{AP}$ (était bien contente)$_{AP}$ (de connaître un homme)$_{AP}$ (aussi raisonnable)$_{AP}$

c. *Mon ami sourit gentiment avec indulgence*. *(Le Petit Prince)*

Predicted phrasing: (Mon ami)$_{AP}$ (sourit)$_{AP}$ (gentiment)$_{AP}$ (avec indulgence)$_{AP}$

Observed phrasing: (Mon ami)$_{AP}$ (sourit gentiment)$_{AP}$ (avec indulgence)$_{AP}$

- in some cases, derived accentual phrases were segmented in two distinct APs. Such strategy usually applied when the $]_{XHead}$ parameter did derive phrases with more than 6 syllables. The division of derived φ phrases in two is usually justified for metrical reasons, as it prevents stress lapses. Examples are given in (13a, b).

(13) a. *Les médias décrivaient en termes apocalyptiques les conséquences d'éventuels bombardements*. *(Le Monde)*

Découpage prédit: (Les média)$_{AP}$ (décrivaient)$_{AP}$ (en termes)$_{AP}$ (apocalyptiques)$_{AP}$ (les conséquences)$_{AP}$ (d'éventuels bombardements.)$_{AP}$

Découpage réalisé: (Les média)$_{AP}$ (décrivaient)$_{AP}$ (en termes)$_{AP}$ (apocalyptiques)$_{AP}$ (les conséquences)$_{AP}$ (d'éventuels)$_{AP}$ (bombardements.)$_{AP}$

b. *J'avais été découragé dans ma carrière de peintre par les grandes personnes*. *(Le Petit Prince)*

Découpage prédit : (J'avais été découragé)$_{AP}$ (dans ma carrière)$_{AP}$ (de peintre)$_{AP}$ (par les grandes personnes)$_{AP}$

Découpage réalisé : (J'avais été)$_{AP}$ (découragé)$_{AP}$ (dans ma carrière)$_{AP}$ (de peintre)$_{AP}$ (par les grandes personnes)$_{AP}$

- in some other cases, the phrasing obtained cannot be easily derived from readjustment rules (in contradistinction to examples (12) and (13)). Such an example is given in (14). This strategy is mostly motivated by the guaranteeing of a certain rhythmic regularity or eurhythmicity. In (14), for instance, the first phrase (*et le petit prince*) contains five syllables that respect the following stress-pattern (σ σ **σ** σ **σ**), which is also observed in

almost all the subsequent phrases of the utterance: (y.ɛ̃.**tʁE**.ʒo.**li**) (e.**kla**.də.**ʁiʁ**) (ki.**mi**.ʁi.**ta**) (bo.**ku**)

(14) *Et le Petit Prince eut un très joli éclat de rire qui m'irrita beaucoup. (Le Petit Prince)*
Predicted phrasing : (Et le Petit Prince)$_{AP}$ (eut)$_{AP}$ (un très joli éclat)$_{AP}$ (de rire)$_{AP}$ (qui m'irrita)$_{AP}$ (beaucoup)$_{AP}$
Observed phrasing : (Et le Petit Prince)$_{AP}$ (eut un très joli)$_{AP}$ (éclat de rire)$_{AP}$ (qui m'irrita)$_{AP}$ (beaucoup)$_{AP}$

The analysis of the data showed that metrical and linguistic principles play a continuous role in prominence distribution. Both types of principles intervene in parallel, and the construction of metrical patterns is dynamic and progressive. It is therefore difficult to support the idea that prosodic constituents are constructed prior to stress assignment. In order to model this, I argue for the following procedure: two distinct constraints hierarchies are proposed, and possible parsings are evaluated against both of them. Preferred parsings (or optimal candidates) should best satisfy each hierarchy independently. Before demonstrating how this works through some concrete examples, I will present the two hierarchies of constraints.

As for the metrically-based hierarchy, metrical constraints are formalised in order to guarantee the principles of alternation and culminativity. This is achieved by two constraints referring to a metrical interpretation of the obligatory contour principle **OCP$_{Rhythm}$** (*cf.* also Laks 1997). As alternation is not interpreted in the same way at lower and higher levels in the metrical grid, a distinction is made between **OCP$_{Rhythm-Inf}$** and **OCP$_{Rhythm-Struc}$.** Indeed, at lower levels in the grid (i.e. syllables), rhythm may be binary or ternary; at higher levels, when groupings have already emerged, only binary rhythm is accepted. The **OCP$_{Rhythm-Inf}$** constraint stipulates that alternance may be binary or ternary. A sequence of six syllables may thus be parsed in one of the ways exemplified in (15).

(15) Possible patterns for a sequence of six syllables (using grid-like representations)

1.
```
    *     *     *
(*  *)  (*  *)  (*  *)
```

2.
```
      *           *
(*  *  *)  (*  *  *)
```

As for the constraint **OCP$_{Rhythm-Struc}$,** it ensures that constituents emerging at higher levels in the grid (2, 3, 4 or more) do not contain more than two constituents generated at the immediately lower level. Patterns in (15) could then lead to the parsing in (16).

(16) Possible patterns according to **OCP$_{Rhythm-Struc}$**

1.
```
    *           *
[   *]  [   *   *]
(*  *)  (*  *)  (*  *)
```

2.
```
          *     *
 [  *     *]  [  *]
(*  *)  (*  *)  (*  *)
```

3.
```
                  *
[       *         *]
(*  *   *)  (*  *   *)
```

As for the constraints based on the alignment with the morpho-syntactic structure, they are formulated in such a way as to show that accentuation in French has a demarcative function. Moreover, the emergence of the surface metrical phrases when aligning with the syntactic structure is done in a retrograde manner, i.e. from right to left. Indeed, the listener must wait for the occurrence of a lexical head in order to process the items that precede it. Psycholinguistic studies on ambiguous sentences and PP attachment generally rely on this retrograde nature of processing (*cf.*, among other, Frazier/Clifton 1996). In example (17), the lexeme *enfants* cannot be accurately parsed into the syntactic structure before the appearance of *sont* in (17a) and *ils* in (17b).

(17) a. *Sans attendre les enfants **sont** partis.*
b. *Sans attendre les enfants **ils** sont partis.*

The alignment constraints formulated are of the family *Align* and *Edgemost* in Optimality Theory (*cf.*, among others, Delais-Roussarie 1996; Selkirk 2005). The alignment constraints indicate which edge of any morpho-syntactic domain has to be aligned with a prominent position in the grid. They are given in (18), (19) and (20).

(18) **ALIGN$_{\textbf{Lex}}$**: Align (Lex, R, ś)
It stipulates that the last metrical syllable of a lexical word must align with a strong position in the grid.
"*Les mé**dias** décri**vaient** en **termes** apocalyp**tiques** les consé**quences** d'éventu**els** bombarde**ments**.*"

(19) **ALIGN$_{\textbf{XHead}}$**: Align (X_{HEAD}, R, ś)
It stipulates that the last metrical syllable of any lexical head of maximal projection must align with a strong position in the grid.
*"On s'inter**roge** grave**ment** sur ses **mines** de psy**chiatre** et ses pro**pos** défini**tifs**"*

(20) **ALIGN$_{\textbf{XMax}}$**: Align (X_{Max}, R, ś)
It stipulates that the last metrical syllable of any maximal projection must align with a strong position in the grid.
"*Les **cours** reprendront le huit sep**tembre**.*" *(France Info)*

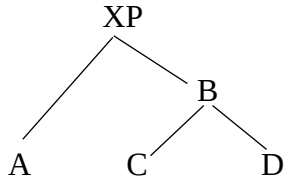
The Edgemost constraints indicate which syllable is the most prominent in a syntagmatic domain (Xhead, Maximal projection, etc.). They are formulated in (21) and (22).

(21) **EDGEMOST$_{\textbf{XHead}}$**: Edgemsot (X_{Head}, Right)
The last metrical syllable of any lexical head of maximal projection is the most prominent in the domain it generates. (e.g. for *les enfants de mon voisin*, the generated domains are *(les enfants)* and *(de mon voisin)*).
"*J'avais été découra**gé** dans ma car**rière** de **peintre** par les grandes per**sonnes**.*" *(Le Petit Prince)*

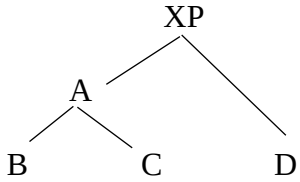
(22) **EDGEMOST$_{\textbf{XMAX}}$**: Edgemsot (X_{Max}, Right)
The last metrical syllable of any maximal projection is the most prominent in the domain corresponding to the maximal projection (in the NP *le president serbe*, the syllable /sEʁb/ must be the most prominent).
"*Le président **serbe** fait le **mort**.*" *(Courrier)*

EDGEMOST constraints serve to account for an asymmetry that has been pointed by Dell (1984). Given the structures in (23), it is possible to parse constituent A with the element C in (23a), whereas D cannot be phrased together with C, leaving B isolated in (23b).

(23) a.



b.



[A C] [D] but * [B] [C D]

This asymmetry has been formulated by the *principe du bord droit* in Martin (1987). According to it, the NP subject *ses voisins* may be parsed with the verb *ont regardé* as in (24a), whereas the PP *de Pierre* cannot be parsed without the rest of the branching subject NP as in (24b).[7]

(24) a. [Ses voisins ont regardé] [un film américain]

b. * [Les amis] [de Pierre viennent]

The EDGEMOST constraints account for such asymmetries. In (24a), EDGEMOST is satisfied as no syllable is more prominent than the syllable /zɛ̃/ within the subject *ses voisins*. By contrast, in (24b), the syllable /pjɛʁ/, which is at the edge of the subject NP, is less prominent than the syllable /mi/.

As already said, possible parsings are evaluated against both constraint hierarchies. The metrical hierarchy is given in (25). The constraint **OCP$_{\text{Rhythm-Struc}}$**, which has almost always been respected in the data, dominates the constraint **OCP$_{\text{Rhythm-Inf}}$**. The latter may be violated in a sequence of four syllables like *à la maison*, in which three consecutive syllables may remain unstressed.

(25) Rhythmic constraint hierarchy

OCP$_{\text{Rhythm-Struc}}$ << OCP$_{\text{Rhythm-Inf}}$ (for OCP$_{\text{Rhythm-Struc}}$ dominates OCP$_{\text{Rhythm-Inf}}$)

As for the syntactically-driven constraints, they have been ordered in such a way as to account for the fact that EDGMOST constraints are almost always respected, whereas alignment constraints may be violated. The resulting hierarchy is given in (26):

[7] It is worth noting that this principle is not effective when the remaining constituent is in sentence final position. In such a case, it may be assumed that the principle is nevertheless respected since the sentence final position may call for a promotion of the beat position in the grid as in the example below :

								*
	*			*				*
*	*	*	*	*	*	*	*	*
ʒE	vy	lE	zɑ̃	fɑ̃	də	mɔ̃	vwa	zɛ̃

(26) Hierarchy of the syntactically-driven constraints

EDGEMOST$_{XMAX}$ << EDGEMOST$_{XHead}$ << ALIGN$_{XMax}$ << ALIGN$_{Xhead}$ << ALIGN$_{Lex}$

As a first approximation, I propose to implement the procedure used to choose optimal parsings as follows. For any utterance, several well-formed metrical grids are generated. These grids are in conformity with the metrical principles, i.e. OCP$_{RHYTHM}$. They are built up by assuming that every metrical syllable is associated with a grid position. In parallel, a syntactically-driven grid, which is in conformity with the morpho-syntactically driven constraints, is generated for every utterance. Then metrical grids and the syntactically-driven one are merged in order to generate metrical patterns that best satisfy both constraint hierarchies. Two concrete examples which inform our understanding of how it works are given in (27) and (28).

(27) *Les cours reprendront le huit septembre.*

Assuming that each syllable corresponds to a demi-beat (or a grid position), several well-formed metrical grids may be generated for (27). Among them, consider (27a) which already respects Align$_{Lex}$.

(27a) Well-formed metrical grid for (27)

								(*)	
				*				*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
lE	kuʁ	ʁə	pʁɑ	dʁɔ̃	lə	ɥi	sɛp	tɑ̃	bʁə

This grid violates some of the morpho-syntactically driven constraints. According to Align$_{XMAX}$, the syllable /kuʁ/ should be more prominent than the syllable /dʁɔ̃/. Three possibilities may be considered:

- as it is not possible to satisfy both hierarchies, precedence may be given to the rhythmic one, and the grid (27a) could thus be chosen as optimal.

- the second possibility (27b) will consist in promoting the position or syllable / kuʁ/, while the position /dʁɔ̃/ is demoted.

  (27b) Possibility 2

								*	
	*							*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
lE	kuʁ	ʁə	pʁɑ	dʁɔ̃	lə	ɥi	sɛp	tɑ̃	bʁə

  This possibility is not completely satisfactory metrically as it generates a lapse at level two in the grid: the positions /dʁɔ̃/ and / ɥi/ are at the same level.

- The third possibility (27c) consists in promoting the position /kuʁ/ in such a way as not to have to demote position /dʁɔ̃/.

(27c) Possibility 3

								*	
	*							*	
	*			*				*	
	*			*		*		*	
*	*	*	*	*	*	*	*	*	<*>
lE	kuʁ	ʁə	pʁɑ	dʁɔ̃	lə	ɥi	sɛp	tɑ̃	bʁə

Possibility (27c) represents the optimal candidate as soon as precedence is given to the syntactically-driven constraint hierarchy. In comparison to (27b), it best satisfies both hierarchies. The phrasing that emerges at the surface corresponds to the one observed in the data: (le cours)$_{AP}$ (reprendront)$_{AP}$ (le huit septembre)$_{AP}$.[8]

Let us now consider example (28)

(28) *J'ai sauté sur mes pieds comme si j'avais été frappé par la foudre.*

The syntactically driven grid generated relatively to the morpho-syntactic constraints is given in (28a).

(28a) Syntactically driven grid

																*
					*											*
		*			*								*			*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ʒE	so	te	syʁ	mE	pje	kOm	si	ʒa	vE	ze	te	fʁa	pe	paʁ	la	fudʁ

A possible metrical grid is proposed in (28b).

(28b) Possible metrical grid

											*					
					*						*					*
		*			*			*			*		*			*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ʒE	so	te	syʁ	mE	pje	kOm	si	ʒa	vE	ze	te	fʁa	pe	paʁ	la	fudʁ

The syntactically driven grid (28a) is not metrically well-formed: a sequence of seven unstressed syllables, which corresponds to grammatical words (preposition, conjunction, auxiliaries, etc.) occurs in the grid (see grey cells). The problem may be solved in one of the following ways:

- the syllables /ʒa/ and /te/ (in grey) may be accented, as it is the case in the grid (28c). This is problematic since the most prominent syllable (/te/) should not be accented according to

[8] A secondary stress is realised on the syllable /ɥi/.

the morpho-syntactically driven constraints. The constraints $Align_{XMAX}$ and $Align_{XHEAD}$ are indeed violated.

(28c) Possibility 1

											*					
					*						*					*
		*			*			*			*		*			*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ʒE	so	te	syʁ	mE	pje	kOm	si	ʒa	vE	ze	te	fʁa	pe	paʁ	la	fudʁ

- the syllables /ʒa / and /te/ may be accented in order to avoid a violation of OCP_{RHYTHM} as in (27d); but the promotion of these positions should be done in a such way as not to violate the constraints $Align_{XMAX}$ and $Align_{XHEAD}$. Even if some violations may occur, they have to be in conformity with the hierarchy in (21).

(28d) Possibility 2

																(*)
					*											*
		(*)			*						*					*
		*			*			*			*		*			*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
ʒE	so	te	syʁ	mE	pje	kOm	si	ʒa	vE	ze	te	fʁa	pe	paʁ	la	fudʁ

Among the solutions, possibility 2 (28d) is the optimal one. The metrical constraints are respected and the morpho-syntactic structure appears relatively clearly in the prosodic structure that emerges at the surface level: (J'ai sau**té** / sur mes **pieds**) (comme si **j'a**vais é**té**) (fra**ppé** / par la **foudre**).

From the examples analysed above, it appears that the evaluation of the candidates has to be formalised in such way as to allow a certain correspondence in terms of prominence: prominent syllables in the syntactically driven grid should preferably be prominent in the 'surface grid', but prominence relation may change (i.e. the syllable /pe/ is prominent in (28d), but not as much as syllable /te/).

To sum up, the proposal argued for here is that two distinct structures are generated for any given utterance. One of them results from the structure-building mechanism described in section 3.1. It is derived from information structure and it plays a crucial role in conveying the meaning associated with the utterance. The other is constrained by the metrical and alignment principles described in section 3.2. Note however that both types of structure are not always respected on surface. Indeed, the chunking that is realised will result from the merging of both structures. In the following section, I will attempt to explain how the two structures (informationally-based and metrically-based) can merge during the production process.

4. How the two structures can be merged

The present model provides two distinct representations (or structures) for any given utterance, one accounting for the way it is segmented into intonational phrases and one derived from

metrical patterns. In this section, which is still very prospective, I will explain how both structures may merge. This will be done by differentiating two broad strategies that can be observed in real communicative situations: a merging process in which the speaker gives precedence to the meaning conveyed, and another one where priority is given to rhythm over meaning. The first strategy will be presented in § 4.1, and the second in § 4.2.

Before explaining the merging mechanism, it is important to note that the distinction between both strategies is only relevant in cases where utterances or discourses are divided into several intonational phrases. In very short utterances, which are segmented into a single intonational phrase, the surface form will often correspond to the structure derived from the optimal grid (see, for instance, examples (27) and (28), § 3.2).

A second issue, which is worth noting, is related to the nature of the merging process, which corresponds in some way to a shift from metrical to rhythmical: the result is clearly anchored in temporality. Moreover, it is embodied, in the sense that it depends on the communicative intentions of the speaker, and to a lesser extent, the hearer. Indeed, this merging process is reminiscent of the distinction made by Gagnepain (2020) (*cf.* § 2.2.) between subdivision and pulsation. Whereas the structures generated from metrical and linguistic principles and constraints (*cf.* § 3.1 and 3.2) are essentially abstract grammatical representations (which can be related to *language* or *competence* in Saussurean and Chomskyan terms respectively), the result of the merging process is the fruit of a concrete speech act, carried out in actual circumstances and by particular speakers.

4.1 Meaning over rhythm

In everyday communicative situations, speakers usually aim to ensure that their interlocutors grasp what they want to say as correctly as possible, and it is with this aim in mind, i.e. to better highlight essential information, that elements are detached (as in dislocations, topicalizations), set back (by the use of incises or parentheticals), etc. As mentioned in § 3.1, all these constructions generate intonational phrases, which are like islands isolated from each other, and indicated by the realisation of specific tonal events. Consequently, the rhythmic pattern (tempos, effective frequency of pulses, etc.) adopted at the beginning, and more precisely while uttering the first intonational phrase of the utterance, may be difficult to maintain throughout the speaking process. Various phenomena, already observed and mentioned in conversational or prosodic analyses, such as presence of pause, change in speech rate, etc., may be used to facilitate the conservation of rhythm. A few examples are presented below, but more detailed studies on this issue, be they qualitative or quantitative, are clearly necessary and should provide interesting insights.

Any utterance or speech segmented into several IPs may be rhythmically unbalanced, the number of syllables in each IP being very different. Consider examples (10) and (11) repeated in (29): the first IP contains 9 or 11 syllables, but the second one 4 or 5. [9] Several strategies are available to speakers to try to maintain eurhythmy. Accents may be realised on the final syllables of *reste*, *longtemps*, *Carcassonne* and *bol*, so that the prosodic groups that emerge are fairly similar in terms of number of syllables (3 or 4 syllables): (quand je reste) (trop longtemps) (à Carcassonne) (j'en ai ras le bol).

[9] The number of syllables depends on the realisation of schwa in *je* (/ʒə/) and *reste* (/ʁɛst(ə)/) in the first IP, and in *le* (/lə/) in the second one.

For utterances where such a solution is difficult to implement, speech rate may vary so that short IPs are pronounced more slowly and long ones more rapidly. In addition, pauses may be inserted to ensure isochrony, i.e. a sort of eurhythmicity. In (29), a change in speech rate occurs, but remains limited. As shown in figure 5 (repetition of figure 3), the temporal distance between the accents (encoded L* or H*) is quite close, regardless of their number in each IP (three in the first one, and one in the second one).

(29) *Quand je reste trop longtemps à Carcassonne, j'en ai ras le bol….*
[[(quand je **reste**)$_{AP}$ (<u>trop</u> long**temp**s)$_{AP}$ (à Carca**ssonn**e)$_{IP}$]$_{MiIP}$ (j'en ai ras le **bol**)$_{AP}$]$_{MaIP}$

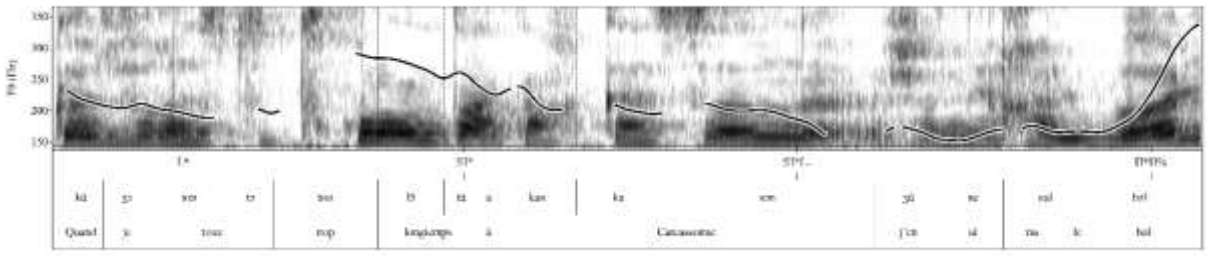


Figure 5. Spectrogram and F0 trace for example (29).

It is likely that pauses and speech rate variations allow each IP to have approximately the same duration; or at least to allow the duration of short IPs to correspond to that of APs in long IPs. Though further systematic investigation needs to be done, the idea which I defend here was mentioned by Wioland (1991, 38), in the example given in (30). He explains that the first IP is usually performed with a slower speech rate than the second one, and with a pause after the syllable /vjẽ/, which is lengthened, all these insuring a sort of isochrony.

(30) *Viens, on est en retard.*

[Viens,]$_{MaIP}$ [on est en retard]$_{MaIP}$
1 syllable 4 /5 syllables

When each IP contains approximately the same number of syllables, the same rhythmic pattern can be maintained throughout the utterance. In (8) and (9), repeated in (31) and (32) respectively, eurhythmicity occurs, particularly in terms of syllable number, so that similar rhythmic patterns are observed in the different IPs (MiIP or MaIP).

(31) *Mais je crois que dans ce rôle-là // il y a des choses à dire* (ACSYNT, COAP)

[[(mais je **crois**)$_{AP}$ (que dans ce role-**là**)$_{AP}$]$_{MiIP}$ (il y a des <u>choses</u> à **dire**)$_{AP}$]$_{MaIP}$

(32) *J'ai le sentiment qu'un enseignant // il est quand même dans cette idée* (Avanzi 2012:162)

[[(j'ai le sentiment)$_{AP}$ (qu'un enseignant)$_{AP}$]$_{MiIP}$ // (il est quand même)$_{AP}$ (dans cette idée)$_{AP}$]$_{MaIP}$

It is likely that the rhythm adopted and observed in speech, which is partly encoded by the temporal distance between accents, is also encoded in the gestures produced by speakers. Rohrer et al (2022) found in a study of the alignment between beat gestures and accents that the

frequency of gestures and accents is relatively similar. This similarity confirms the multimodal and highly embodied nature of rhythm in speech.

In the cases cited so far, the IPs analysed were minor IPs. It is however worth investigating how rhythm is maintained after a major IP boundary in a non-final position, i.e. in the case of narrow focus. Although more systematic work on corpora should be carried out in a near future, previous studies showed that accents and phrase boundaries are encoded in post-focus sequences by means of syllable lengthening rather than by the presence of a pitch movement (*cf.* Jun/Fougeron 2000, Delais-Roussarie et al. 2002). This can nevertheless be analysed as a way of preserving rhythmic balance.

To sum up, in actual communication situations, speakers often give priority to meaning, and therefore to the information-based predicted chunking at the IP level (*cf.* § 3.1), even if it leads to an utterance which is metrically unbalanced. This being the case, it is not uncommon for strategies (occurrence of pauses, change in articulation rate) to be used to maintain eurhythmicity or isochrony as much as possible.

4.2 Rhythm over meaning

In most communicative situations, speakers' primary intention is to convey the meaning of the utterance, although this may require the use of strategies to maintain a sense of rhythm. However, it happens that the speaker's primary intention is to preserve a certain tempo or rhythm, related in most cases to the meaning being conveyed and this rhythm (or momentum) is then also transmitted to the interlocutors/listeners. This is similar to what Gagnepain (2020, 27-30) says about the rhythmical beats that the music performer wishes to convey to the audience during concerts. In such situations, the merging of both structuring principles (intonational and metrical) should be chosen so as to ensure the preservation of tempo and rhythm, even if direct access to the literal meaning becomes difficult.

Among the language productions in which rhythm is of primary importance, one may mention eloquent speeches (eulogies, declamations, etc.) and poetic creations (versified or non-versified poetry, slam, rap). Without going into detail on this issue, it is worth pointing out that figures such as *enjambment* (i.e. when a syntactic phrase overflows from one metrical unit, the verse, to another) always allow the preservation of rhythm.

In some cases, while creating the effect of a certain distance between syntax and metrical structure, *enjambment* does not go as far as transgressing $\text{EDGEMOST}_{\text{XMAX}}$ (see (22), § 3.2). Thus, in (33), to align the end of the first verse with the lexical head (*contorsions*) of the NP "*les contorsions de tous ces grands faiseurs de protestations* " creates a gap between metrical and linguistic mechanisms. Nevertheless, the right boundary of the NP (at the end of *protestations*) coincides with a metrical position of equal level, since it falls at the end of a verse.

(33) *Et je ne hais rien tant que les contorsions*
De tous ces grands faiseurs de protestations.
Molière, Le Misanthrope (vers 43-44)

In some other cases, as in (34), the $\text{EDGEMOST}_{\text{XMAX}}$ constraint and alignment with major IP boundaries are not observed, creating a clear mismatch, and even a feeling of disharmony. Indeed, the end of the subject NP coincides with the end of a verse, whereas the right edges of

the two clauses "*L'empereur se tourna vers Dieu*" and "*L'homme de gloire trembla*" are not aligned with the end of a verse, which can be analysed as a MaIP boundary.

(34) *L'empereur se tourna vers Dieu ; l'homme de gloire*
Trembla ; Napoléon comprit qu'il expiait.
Victor Hugo, *Les Châtiments*, L'expiation I (62-63)

In speech used to guide meditation and yoga practice, rhythm is also given a significant place, (rhythm being used here in a very general sense; breathing, for example, being also included). In these situations, the speaker's intention is to lead the audience to interiority (or silence); and this is often reflected prosodically by the recurrent use of a very slow tempo and rate of speech, by the presence of many relatively long pauses, and by a rather compressed register in which melodic movements are of limited slope (cf. Delais-Roussarie/Yoo, in prep.).

Silences or pauses can allow everyone (speaker and audience) to harmonise with his/her own respiratory rhythm. Example (35) gives an indication of the duration of pauses in comparison to effective speaking time for an utterance extracted from a yoga practice.[10] It shows that in a very short extract with relatively few long pauses, the duration of the pauses (i.e. silent time) represents more than 50% of the total duration. In addition, when listening to the extract, the meaning of the whole utterance does not come over clearly, in particular because of the long pause after *inspiré.*

(35) *Observez ainsi où s- où se diffuse le souffle inspiré et le souffle expiré en vivant ces temps de suspension.* [InteriorityCorpus - MIT-yo2]

observez ainsi où s'(1,217s) #[0,14s] où(0,221s) #[0,221s] se diffuse le souffle inspiré (2,038s) ##[1,195s] et le souffle expiré en vivant (1,651s) #[0,515s] ces temps (0,583s) #[0,304s] de suspension (0,919s) ## [4,825s]

Other features often present in such types of discourse clearly indicate the priority given to rhythm over syntactic and informational structures and the meaning they convey in the chunking process. First of all, acceleration in articulation rate is not as important in such discourse type, which is in contradistinction to that mentioned in § 4.1. In ordinary speech, these changes allow one to maintain the tempo and rhythm adopted, regardless of the size of the derived IPs. In meditations and yoga practises, change in speech rate only happens when the utterances convey necessary and relatively long technical or practical information which are in a way outside the guidance to silence. In addition, it should be noted also that unlike the examples cited in the previous section, lengthening and pauses can occur here in positions that are generally unaccentuated in other speech styles. Thus, in (36), the monosyllabic adverb *juste* is lengthened and preceded and followed by long pauses, but according to linguistic and metrical constraints this word would rather remain unaccented in such a sentence.

(36) *Il n'y a rien à faire juste à se laisser porter*.
Il n'y a rien à faire # juste # à se laisser porter.

On this issue, it is important to note that nominative sentences and anacoluths are also very frequent in this speech style. As for the relation to meaning, in some cases, it is not literal;

[10] In (35), pause duration is indicated in seconds in the square brackets after the symbol #, whereas speech duration is given in brackets after the orthographic transcription of each chunk.

instead nouns and adjectives uttered in isolated chunks have an evocative power, sometimes enhanced by the rhythm and tempo chosen. Further research in this domain is currently in progress.

As just seen, in some speech situations and styles, precedence is clearly given to rhythm over meaning. In such cases, a direct access to the syntactic and informational structures may not be given priority but preference is given to the images and meanings evoked by the words and rhythm patterns themselves.

5. Conclusion and perspectives

In this contribution, it is argued that surface prosodic chunking should be analysed as rhythmically structured, in the sense that it is anchored in time and clearly embodied. In order to account for the observable chunking, I proposed a model based on a distinction between a surface level, which concerns speech or performance, and a more abstract underlying level. At the underlying level, the prosodic organisation of any utterance is constructed from two distinct structures, an intonational structure derived from the informational structure and a metrico-syntactic structure represented within a metrical grid in which groupings emerge from the interaction between metrical principles of alternation and alignment constraints with the morpho-syntactic structure. During the realisation of an utterance, the two abstract structures have to merge. During the merging process, it is possible either to enhance access to the meaning conveyed by the utterance or to focus on metrical and rhythmic regularities. In the first case, the boundaries of the constituents of the intonational structure are clearly marked by means of tonal events, even if eurhythmicity, which is usually encoded at the grammatical level in metrical divisions, is not respected. Occurrences of pauses or accelerations in speech rate, i.e. phonetically implemented events, can then favour a certain isochrony between the different intonational constituents and markers (pitch accents, boundary tones, etc.). In the second case, rhythm and metrical principles are privileged even though meaning does not come over clearly. Another, probably more symbolic, meaning may then come across, transmitted by words uttered in isolation and tempos, whereas intonational events related to information structure are less salient.

In both cases, however, rhythm, in its temporal dimension, tends to be respected. Indeed, the temporal distance between melodic events tends to be balanced in order to ensure isochrony. It is therefore interesting in future research to study more precisely and in depth the relations between rhythmic units of a metrical nature, often of a low level (the mora, the syllable and the foot), on the one hand, and on the other units of a tonal nature, in particular the accentual phrase which is often defined as having both a tonal form and a rhythmic dimension in French, and also in other languages. Moreover, as the rhythmic organisation of utterances seems to respond to constraints that go beyond the strictly linguistic and grammatical dimension, it is interesting to encourage work integrating other elements of human communication such as gestures, whether or not they are referential. Finally, in order to better understand what rhythm refers to and how it is instantiated in language and speech, exploring parallels with music or other rhythmically structured human activities could also be very promising.

6. Bibliography

Avanzi, Mathieu (2012): *L'interface prosodie/syntaxe en français. Dislocations, incises et asyndètes*. Bruxelles, Peter Lang (Collection Gramm-R).

Beyssade, Claire/ Delais-Roussarie, Elisabeth/ Doetjes, Jenny/ Marandin, Jean-Marie et al. (2004): “Prosody and Information in French”, in: Francis Corblin/ Henriett de Swart (eds), *Handbook of French semantics.* Stanford, CSLI, 477–500.

Cureton, Richard D. (1992): *Rhythmic phrasing in English Verse*, English Language Series 18, London/New York, Longman.

Delais-Roussarie, Elisabeth (1995): *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française*, Thèse de Doctorat (PhD Dissertation), Université de Toulouse-Le Mirail.

Delais-Roussarie, Elisabeth. (1996): “Phonological Phrasing and Accentuation in French”, in: Marina Nespor/ Norval Smith (eds), *Dam Phonology: HIL Phonology Paper II*, Den Haag, Holland Academic Graphics, 1–38.

Delais-Roussarie, Elisabeth (2000): “Vers une nouvelle approche de la structure prosodique”, in: *Langue Française* 126, 92–112.

Delais-Roussarie, Elisabeth (2021) : “Intonational phrasing in French in the light of oral corpora”, in : Alexander Teixeira-Kalkhoff/ Maria Selig/ Christine Mooshammer (eds), *Prosody and Conceptional variation*, Frankfurt am Main, Peter Lang Verlag, 61–73.

Delais-Roussarie, Elisabeth/ Rialland, Annie / Doetjes, Jenny /Marandin, Jean-Marie (2002): “The Prosody of post-focus sequences in French”, in: Bernard Bel/ Isabelle Marlien (eds), *Proceedings Speech Prosody 2002,* Aix en Provence, April 2002.

Delais-Roussarie, Elisabeth/ Rialland, Annie (2005): “Metrical organization, tonal association and focus in French”, in: Stefan Blaauw / Frank Drijkoningen (eds.), *Romance languages and linguistic theory 2005*. John Benjamins Publishing, 73–98.

Delais-Roussarie, Elisabeth/ Post, Brechtje (2008): “Unités prosodiques et grammaire de l’intonation : vers une nouvelle approche”, in: *Actes des XXVIIème Journées d’Etudes sur la Parole (JEP-TALN 2008)*, Avignon, June 2008. (http://www.afcp-parole.org/doc/Archives_JEP/2008_XXVIIe_JEP_Avignon/PDF/avignon2008_pdf/JEP/077_jep_1565.pdf)

Delais-Roussarie, Elisabeth/ Yoo, Hiyon/ Post, Brechtje (2011): “Quand frontières prosodiques et frontières syntaxiques se rencontrent”, in: *Langue Française* 170/2, 29 – 44.

Delais-Roussarie, Elisabeth/Post, Brechtje/ Avanzi, Mathieu/ Buthke, Caroline et al. (2015): “Chapter 3 - Intonational phonology of French: Developing a ToBI system for French”, in: Sonia Frota/ Pilar Prieto (eds), *Intonation in Romance*, Oxford, Oxford University Press, 63–100.

Delais-Roussarie, Elisabeth/ Post, Brechtje/ Yoo, Hiyon (2020): “Unités prosodiques et grammaire intonative du français : vers une nouvelle approche”, in: *Actes des 33ème Journées d’études sur la Parole (JEP-TALN-Recital)*, Nancy, June 2020, 145–153

Delais-Roussarie, Elisabeth/ Yoo, Hiyon (in prep.): “Quelle prosodie pour guider vers l’intériorité ?”, unpublished ms.

Delattre, Pierre (1966): “Les dix intonations de base du français”, in : *The French Review* 40 (1), 1-14.

Dell, François (1984): “L'accentuation dans les phrases en français”, in: François Dell/ Daniel Hirst/ Jean-Roger Vergnaud (eds), *Forme sonore du langage: structure des représentations en phonologie*, Paris, Hermann, 65–122.

Di Cristo, Albert (1998): “Intonation in French”, in: Daniel Hirst/ Albert Di Cristo (eds), *Intonation Systems: A Survey of Twenty Languages*, Cambridge, Cambridge University Press, 195-218.

Di Cristo, Albert (1999): “Le cadre accentuel du français contemporain”, in: *Langues* 3 (2), 184-205 and *Langues* 4 (2), 258-267.

Di Cristo, Albert (2016): *Les Musiques du Français Parlé : Essais sur l'accentuation, la métrique, le rythme, le phrasé prosodique et l'intonation du français*, Berlin/Boston, De Gruyter.

Di Cristo, Albert/ Hirst, Daniel (1996): “Vers une typologie des unités intonatives du français”, in: *Actes des XXièmes Journées d’Études sur la Parole – JEP 96*, Avignon (France), 219-222.

Di Cristo, Albert/ Jankowski, Ludovic (1999): “Prosodic organization and phrasing after focus in French”, in: *Proceedings of the* XIVth *International Congress of Phonetic Sciences (ICPhS)*, San Francisco (USA), 1565-1568.

Féry, Caroline/Hörnig, Robin/Pahaut, Serge (2011): “Correlates of phrasing in French and German from an experiment with semi-spontaneous speech”, in: Christoph Gabriel/Conxita Lleó (eds.), *Intonational phrasing in Romance and Germanic. Cross-linguistic and bilingual studies*, Amsterdam/Philadelphia, Benjamins, 11–41.

Fraisse, Paul (1974): *Psychologie du rythme*, Paris, PUF: Presses Universitaires de France.

Frazier, Lyn/ Clifton, Charles (1996): *Construal*, Cambridge (MA), MIT Press.

Frota, Sónia (2000): *Prosody and focus in European Portuguese: phonological phrasing and intonation*, New York, Garland.

Frota, Sónia (2012): “Prosodic structure, constituents and their representations”, in: Abigail Cohn / Cécile Fougeron/ Mary Huffman (eds.), *The Oxford Handbook of Laboratory Phonology*, Oxford, Oxford University Press, 255–265

Gagnepain, Xavier (2020): *Du musician en général… au violoncelliste en particulier*, Paris, Philharmonie de Paris Editions.

Goldbeter, Albert (2010): *La vie oscillatoire : au coeur des rythmes du vivant*, Paris, Odile Jacob.

Hayes, Bruce (1989): “The Prosodic Hierarchy in Meter”, in: Paul Kiparsky/ Gilbert Youmans (eds), *Rhythm and Meter*, Orlando, Academic Press, 209–260.

Hayes, Bruce (1995): *Metrical stress theory*, Chigago, The University of Chicago Press.

Hayes, Bruce/Lahiri, Aditi (1991): “Bengali Intonational Phonology”, in: *Natural Language and Linguistic Theory* 9/1, 47–96.

Hirst, Daniel/ Di Cristo, Albert (1984): “French Intonation: A parametric approach”, *Die Neuren Sprachen* 83 (5), 554-569.

Ito, Junko/Mester Armin (2013): “Prosodic Subcategories in Japanese”, in: *Lingua* 124, 20–40.

Jun, Sun-Ah (1993): *The Phonetics and Phonology of Korean Prosody*, PhD. Dissertation, Ohio State University.

Jun, Sun-Ah (2005): “Prosodic Typology”, in: Jun, Sun-Ah (ed.), Prosodic Typology: The Phonology of Intonation and Phrasing, Oxford, Oxford University Press, 430–458.

Jun, Sun-Ah/ Fougeron, Cécile (2000): “A phonological model of French intonation”, in A. Botinis (ed.), *Intonation: Analysis, Modelling and Technology*, Dordrecht: Kluwer Academic Publishers, 209-242.

Ladd, Robert (1996/2008): *Intonational Phonology*, Cambridge: Cambridge University Press.

Laks, Bernard (1997): *Phonologie Accentuelle : Métrique, Autosegmentalité et Constituance.* Collection Sciences du Langage, Paris, Editions du CNRS.

Lerdhal, Fred/ Jackendoff, Ray (1983): *A generative Theory of Tonal Music*, Cambridge (MA), MIT Press.

Liberman, Mark/ Prince, Alan (1977): “On stress and linguistic rhythm”, in: *Linguistics Inquiry* 8, 249–336.

Martin, Philippe (1987): “Prosodic and rhythmic structures in French”, in: *Linguistics* 25, 925–949.

Mertens, Piet (1987): *L'intonation du français. De la description linguistique à la reconnaissance automatique*, Unpublished PhD Dissertation, Leuven, University of Leuven.

Mertens, Piet (1993): “Accentuation, Intonation et Morphosyntaxe”, in: *Travaux de Linguistique* 26/93, 21-69.

Mertens, Piet (2008): “Syntaxe, prosodie et structure informationnelle: une approche prédictive pour l'analyse de l'intonation dans le discours”, in: *Travaux de linguistique* 2008/1(56), 97–124.

Michelas, Amandine (2011): *Caractérisation phonétique et phonologique du syntagme intermédiaire en français: de la production à la perception*, Thèse de doctorat d'Aix-Marseille Université.

Michelas, Amandine/D'Imperio, Mariapaola (2010): “Indices phonétiques et contraintes phonologiques : caractérisation du syntagme intermédiaire en français”, in: *Actes des XXVIIIème Journées d’Etude sur la Parole*, Mons, 25-28 May 2010, 77-80.

Nespor, Marina/Vogel, Irene (1986): *Prosodic phonology*, Dordrecht, Foris.

Nespor, Marina/Vogel, Irene (1989): “On clashes and lapses”, in: *Phonology* 6, 69–116.

Pasdeloup, Valérie (1992): “A prosodic model for French text-to-speech synthesis: a psycholinguistic approach”, in: Gérard Bailly/ Christian Benoit/ T.R Sawallis (eds), *Talking Machines: Theories, Models, And Designs*, Elsevier Sciences Publishers, 335-348.

Pierrehumbert, Janet/Beckman, Mary (1988): *Japanese Tone Structure*, Cambridge (MA), MIT Press.

Portes, Christel/ Bertrand, Roxane (2011): “Permanence et variation des unités prosodiques dans le discours et l'interaction”, in: *Journal of French Language Studies*, 21(1), 97-110.

Post, Brechtje (1999): “Restructured Phonological Phrases in French. Evidence from Clash Resolution” in: *Linguistics* 37 (1), 41-63.

Post, Brechtje (2000): *Tonal and phrasal structures in French intonation*. Den Haag: Holland Academic Graphics.

Post, Brechtje (2011): “The multi-faceted relation between phrasing and intonation in French”, in: Christoph Gabriel/Conxita Lleó (eds), *Intonational Phrasing at the Interfaces: Cross-Linguistic and Bilingual Studies in Romance and Germanic*, Amsterdam, Benjamins, 44–74.

Prieto, Pilar (2005): “Syntactic and eurhythmic constraints on phrasing decisions”, in: *Studia Linguistica* 59, 194–222.

Prieto, Pilar (2006): “Phonological phrasing in Spanish”, in: Sonia Colina/Fernando Martínez-Gil (eds.), *Optimality-theoretic advances in Spanish phonology*, Amsterdam, Benjamins, 39–60.

Prince, Alan (1983): “Relating to the grid”, in: *Linguistic Inquiry* 14, 19–100.

Prince, Alan/ Paul Smolensky. (1993/2004): *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell Publishers (2004), Technical Report, Rutgers University Center for Cognitive Science and Computer Science Department, University of Colorado at Boulder (1993).

Rohrer, Patrick/Prieto, Pilar/ Delais-Roussarie, Elisabeth (2022): “Le rythme prosodique guide le rythme gestuel”, in: *Actes des Journées d’études sur la Parole - JEP 2022*, Ile de Noirmoutier, June 2022.

Selkirk, Elisabeth (1978/1981): “On prosodic structure and its relation to syntactic structure”, in: Thorstein Fretheim (ed.), *Nordic Prosody II: Papers from a symposium*, Trondheim, TAPIR, 111–140.

Selkirk, Elisabeth (1984): *Phonology and syntax: the relation between sound and structure*, Cambridge (MA), MIT Press.

Selkirk, Elisabeth (1986): “On derived domains in sentence phonology”, in: *Phonology Yearbook* 3, 371–405.

Selkirk, Elisabeth (1995a): “Sentence prosody: intonation, stress, and phrasing”, in: John Goldsmith (ed), *The Handbook of phonological theory*, Cambridge, Basil Blackwell, 550-569.

Selkirk, Elisabeth (1995b): “The prosodic structure of function words”, in: Jill Beckman/ Laura W. Dickey/ Susan Urbanczyk (eds.), *Papers in Optimality Theory*, Amherst (MA), GLSA Publications, 439–470

Selkirk, Elisabeth (2005): “Comments on intonational phrasing”, in: Sónia Frota/Marina Vigário/Maria João Freitas (eds.), *Prosodies: with special reference to Iberian languages*, Berlin, De Gruyter, 11–58.

Steedman, Mark (2000): “Information Structure and the Syntax-Phonology Interface”, in: *Linguistic Inquiry* 31/4, 649–689.

Verluyten, Sylain-Paul (1982*)*: *Investigation on French Prosodics and Metrics*, Phd Dissertation, Antwerpen, University of Antwerpen.

Wioland, François (1991): *Prononcer les mots du français : des sons et des rythmes*, Paris, Hachette.

Zubizarreta, Maria-Luisa (1998): *Prosody, Focus and Word Order*. Cambridge (MA), MIT Press.