



HAL
open science

Prosodic constituents in French: a data-driven approach

Jacqueline Vaissière, Alexis Michaud

► **To cite this version:**

Jacqueline Vaissière, Alexis Michaud. Prosodic constituents in French: a data-driven approach. I. Fónagy, Y. Kawaguchi and T. Moriguchi. Prosody and syntax, John Benjamins, pp.47-64, 2006, coll. "Usage-based linguistic informatics". hal-00130794

HAL Id: hal-00130794

<https://hal.science/hal-00130794v1>

Submitted on 13 Feb 2007

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.

Reference Type: Book Section

Authors: Vaissière, Jacqueline; Michaud, Alexis

Year: 2006

Title: Prosodic constituents in French: a data-driven approach

Editors: Fónagy, Ivan; Kawaguchi, Yuji; Moriguchi, Tsunekazu

Book Title: Prosody and syntax

City: Amsterdam

Publisher: John Benjamins

Pages: 47-64

Series Title: Usage-based linguistic informatics

Prosodic constituents in French: a data-driven approach

Jacqueline Vaissière*, Alexis Michaud*

*Laboratoire de Phonétique et Phonologie, UMR 7018 CNRS/Sorbonne Nouvelle

Fortcoming in: *Prosody and syntax*, ed. by I. Fónagy, Y. Kawaguchi and T. Moriguchi, “Usage-based linguistic informatics” series, Amsterdam:John Benjamins.

Abstract

This paper aims (i) to summarise essential facts about the *syntactic prosody* of French as seen within the broader picture of French prosody, (ii) to provide a cross-linguistic perspective, by bringing out characteristics which sharply distinguish French from English, and drawing their implications for the thorny issue of cross-linguistic prosodic description, which arguably holds the key to substantial progress in our understanding of prosody. The essentials of a *superpositional* model of intonation for French are briefly set out.

1. Introduction

The aim of this paper is to provide a cross-linguistic perspective on prosody and syntax, starting out from French data. The present account of the relations between prosody and syntax in French builds on research that deals specifically with this language; it is however written with a view to contributing to cross-linguistic investigation: prosodic typology arguably holds the key to major progress in this field. Despite common underlying principles (Bolinger 1978, Vaissière 1995), there are major cross-language differences. The discussion brings out characteristics that sharply distinguish that sharply distinguish French from languages that possess lexically distinctive stress, addressing the thorny issue of cross-linguistic prosodic description.

1.1. Use of terms

How terms are defined is especially crucial in prosodic studies. *Prosody* as defined here consists of *accentuation*, *intonation* and several *performance factors* (including *rhythm*). *Accentuation* includes all nonphonemic lexically distinctive properties, i.e. (depending on the language) stress, as in English, tone, as in Mandarin, pitch accent, as in Japanese and Swedish, voice quality register, as in Southeast Asian languages such as Mon. *Intonation*, which is often (and perhaps somewhat abusively) identified with the parameters whereby it manifests itself—and especially with fundamental frequency—, is a complex, abstract structure, that can usefully be divided into (i) two sub-systems of structuration: *syntactic intonation*, which essentially reflects syntax in the broader sense (the relationship between syntax and syntactic prosody will be elaborated on below), and *pragmatic intonation*, which

reflects information structure; (ii) *attitudinal* and *emotional* dimensions, that convey speaker attitudes and emotions¹.

These definitions elaborate on proposals by Coustenoble and Armstrong 1937, Delattre 1965, 1966, and closely resemble those put forward by Rossi 1967, 1999 (see also the terminological discussion in Di Cristo 1998, and the introduction to the volume [Hirst and Di Cristo 1998]).

The phrase “syntactic intonation” may appear as somewhat of a misnomer, insofar as syntax and intonational phrasing do not stand in a strict, one-to-one relationship with syntactic units, as was already noted in the early classics of phonetics (Grammont 1933; see also, more recently, Selkirk 1972, 2000:231, Martin 1981). The phrase “syntactic intonation” is nonetheless retained in view of the fact that knowledge of a sentence’s syntax offers a sufficient basis for the synthesis of an acceptable fundamental frequency contour (Vaissière 1971).

The acoustic correlates of prosody are many. They include the variations in *fundamental frequency*, *duration* and *intensity*, *voice quality* (mode of vibration of the vocal folds), and also the allophonic variations in the realisation of the segments (intrinsic and cointrinsic characteristics, which are uncontrolled, should be factored out of prosodic analysis). Said differently, prosody has correlates at the respiratory level, at the glottis, and at the supraglottic level. All parameters take part in prosody simultaneously, to a greater or lesser extent.

1.2. Empirical basis for the present proposal: A data-driven approach based on read speech

The adequacy of data-driven prosodic models is assessed on the basis of whether they attain their immediate goal and account for the body of data chosen as corpus. Our own research into intonation originally hinged on syntax, for the sake of synthesis (Vaissière 1971), language comparison (Vaissière 1983), language recognition (Vaissière 1988), and more recently for teaching French prosody to students of French, and using prosody in speech therapy. To a certain extent, all data-driven models depend on the corpus they are based on. Studies of *read speech* quite predictably focus on *syntactic intonation*. The reader has access to the structure of the sentence as a whole, and can evaluate the length of its parts and their semantic relations, and organise his production in consequence. The preplanning which this allows is reflected in the structuring of the sentences; *prosodic structures* emerge (the plural *structures* will be discussed below). In the 1970s, text-to-speech synthesis (together with the advent of generative grammar) fuelled interest in the relationship between syntactic and intonational constituents in read speech; the two are similar but not equivalent. (The development of automatic recognition of continuous speech also directed the researchers’ attention to these same issues.) The use of read speech was a useful abstraction: the

¹ The interesting topic of how speaker-specific habits come to pattern into a personal *style*, and beyond, into dialect-specific characteristics, will not be addressed here.

congruence between syntactic and prosodic units is stronger in read than in spontaneous speech, in sentences outside context than in continuous texts.

2. An overview of syntactic intonation in French

A representation of syntactic prosody should be detailed enough to specify all the contrasts that lead to a difference in the syntactic interpretation of the sentence (and hence provides a basis for a synthesis-by-rule program which aims to maximally enhance the intelligibility of the output signal).

2.1. *The salience of intonational phenomena in French, due to the absence of lexically distinctive stress*

French lends itself especially well to a study of postlexical prosodic structuration, because the constraints imposed by the lexical level (*i.e.* accentuation; see our definition above) are minimal.

The sense-group is the basic unit. Since the earliest accounts of French prosody, researchers have noted the predominance of *groupe de sens*, *i.e.* sense-group boundaries, over word boundaries in French (Grammont 1933, Coustenoble *et al.* 1937, Delattre 1966). The words may lose—partly or completely—their acoustical identity to a higher-level constituent (*i.e.* word boundaries may go unmarked).

French is generally considered as a ‘rising’ language with final lengthening. An insightful account of French prosodic phrasing was proposed that by Delattre 1966 (usefully complemented by Fónagy 1980). The French ear is trained to perceive *continuation* at the end of a prosodic phrase. In this sense, French is a ‘rising’ language: each prosodic phrase inside a sentence tend to end with a sharp rise (Delattre’s *continuation majeure*), or a smaller rise, or a high F_0 value (Delattre’s *continuation mineure*). In the babbling of French infants, rising F_0 contours and final-syllable lengthening are most frequent (whereas falling F_0 contours and an absence of final lengthening are most frequent for Japanese children: see Hallé, de Boysson-Bardies *et al.* 1991). To an adult French listener, the quasi-regular recurrence of strongly stressed syllables in English is striking, because the closest phonetic equivalent in French is emphatic stress, hence an impression of unceasing emphasis. To the untrained French ear, Japanese rhythm is somewhat puzzling, because the duration of a vowel primarily depends on its phonemic identity, not on the presence or absence of a boundary, and vowel lengthening may therefore occur in any position, whereas in French it is a cue to an intonational boundary.

2.2. *The units defined by F_0 fluctuations*

The units that are here proposed for the syntactic prosody of French and other languages are the *prosodic paragraph*, the *sentence*, the *breath group*, the *melodic phrase*, the *prosodic phrase* or phonological syntagm (variously called *minor group*, *accent group*, *sense group*, *syntagma* in Russian, *buntetsu* in Japanese, or simply *phonological phrase*), the *prosodic*

word, the *foot* (for a language like English), the *syllable* and the *rhyme*. We propose below a definition for each of them, based on observations concerning carefully read long sentences embedded in paragraphs, a speaking style which brings out different degrees of boundaries and prominences especially well. There are major discrepancies across researchers in the use of these terms; however appreciative we are of the research of others, we hold fast to our own definitions in this paper, for the sake of coherence and simplicity (hence the high number of self-references). We do not systematically attempt to provide cross-references to the terms used by other authors.

1) The paragraph is the largest unit. The highest F_0 value in each sentence tends to decline from the first to the last sentence in a *paragraph*, in French and in other languages (Lehiste 1975). The end of the paragraph typically ends on an extra-low F_0 (often leading to a change in voice quality) and intensity.

2) The sentence level is the next unit. The neutral, affirmative statement is taken as the basic, archetypal pattern (according to sentence mode, significant departures from this basic pattern are observed: Thorsen 1980). The F_0 curve for the sentence rises to a peak typically located within the first lexical word, i.e. on one of the sentence's first syllables. Falls and rises in F_0 then alternate, within a gradually narrowed range. A final lowering marks the end of the sentence. (This corresponds to Tune 1 as described for English by Armstrong and Ward 1926.) The realisation of the final fall constrains the two final content words: the rise starts at the end of the penultimate word, and the contour over the last word is falling.

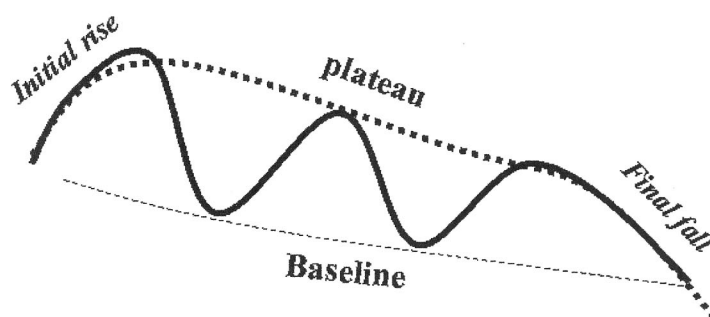


Figure 1. General outline of the F_0 curve of an affirmative statement (after Vaissière 1983)

The sentence is further divided into breath groups and melodic phrases by returns to the baseline. A long sentence may be divided into two or more breath groups by inspiratory pauses; the structure of a breath group recalls that of the sentence as a whole. A breath group, whether sentence-final or not, is acoustically characterised at its beginning by a resetting of the baseline, an initial rise, generally ending at the beginning or end of the first content word, and by the return to the baseline (further characterised below). A melodic phrase is similar to a breath group except that *inspiration* does not actually take place as at the end of the breath group; inspiration is *simulated*—much as a large excursion in F_0 can function as a signal of effort, even though it only simulates vocal effort: physiological necessity and linguistic structure interact closely (see Gussenhoven 2002 and references; Gendrot 2005). The *melodic*

phrase is essentially equivalent to the group ending either in Delattre's *majeure* (realised as an F_0 rise during the final syllable before a major boundary), or in Delattre's *finalité* (a falling F_0 movement spread over the word-final syllables).



Figure 2. Division of the sentence into two melodic phrases. General outline (left), and example with a randomly chosen sequence of syllables 'rides' over the general shape (right). Triangles indicate word-initial syllables, empty circles indicate word-medial syllables, full circles word-final syllables, stars indicate grammatical words.

What is common to both types of melodic phrases (at least under our description) is a return to the *baseline*. F_0 reaches and even goes under the baseline at the very end of the *sentence-final melodic phrase*. *Nonfinal melodic phrases* end in a *continuation* (see 2.1 above; this corresponds to Tune 2 of Armstrong *et al.* 1926). The return to the baseline before the final major continuation rise takes place (i) over the last syllable, or (ii) at the end of the penultimate (thereby suppressing the initial rise of the last disyllabic word), or again (iii) at the end of the preceding function word or of the penultimate word (the falling contour over the penultimate content word thus contrasting with the rising contour on the final word, a fact highlighted by Martin 1981).

No hard-and-fast phonological evidence can be adduced to support this division into levels. There is no single, well-defined domain of application for rules such as *liaison* (i.e. whether a word-final consonant is pronounced or not when the following word begins with a vowel): sometimes they apply across prosodic phrases, sometimes they do not. "Les soldats anglais" generally forms a single prosodic phrase (but two prosodic words) and the *liaison* is optional (sometimes replaced by a glottal stop).

The baseline is a major notion in our description. Our way of modelling the observed data relies strongly on the baseline (also called declination line). F_0 values tend to decline slightly during the course of a sentence, partly due to the decrease in sub-glottal pressure (Lieberman 1967) and to the tracheal pull (Maeda 1976). The baseline is speaker-dependent. Using what is commonly referred to as 'Maeda's method' (Maeda 1976), the baseline is calculated visually by superposing the F_0 contour of a number of isolated, declarative sentences of similar length and determining the upper line (*plateau*) and lower line (baseline). Our description gives a central role to the baseline for two reasons: it is relative stable as compared to the topline (or 'plateau') in the production data (for English, see, again, Maeda 1976); French listeners seems to be very sensitive to the fact that a syllable does or does not actually hit the baseline (Vaissière 1976). The declining baseline seems to serve a perceptual role as a reference line for the listeners (Pierrehumbert 1979).

In a sentence-final melodic phrase, the baseline is reached at the very end of the melodic phrase. In a nonfinal melodic phrase, some variation is observed in the return to the baseline before the final rise: it takes place at the beginning or middle of the last syllable, or on the penultimate syllable (anticipatory lowering often takes place on the penultimate syllable of a content word before a major continuation rise), on the function word preceding the final word, or at the end of the penultimate content word.

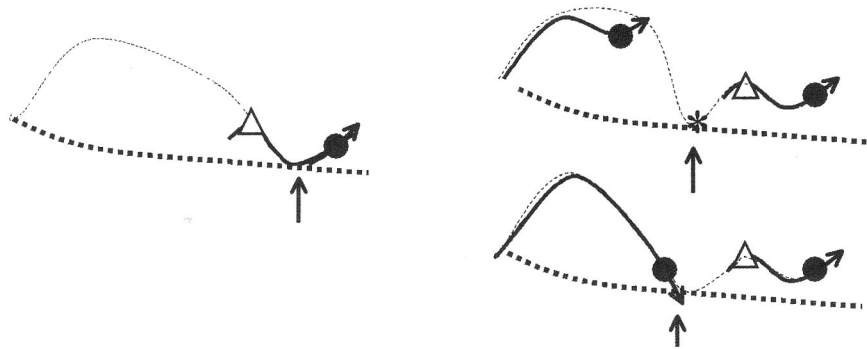


Figure 3. Variation in the timing of the return to the baseline in French. Left: at beginning of last syllable; top right: on function word; bottom right: on end of penultimate content word (after Vaissière 2002)

4) Melodic phrases can be further divided into two or more prosodic phrases. The *prosodic phrase* corresponds to the sense-group (and to Delattre's *minor continuation*). It is composed of a single word, or of two or more semantically related words. The main acoustic difference between a melodic phrase and a prosodic phrase is the absence of a return to the baseline within the latter. Minor continuation is typically realised by a rise or a peak that is not preceded by a return to the baseline. The overall shape of a prosodic phrase strongly recalls that of a breath group, though the final rise is less salient. Final lengthening is generally (though not always) found at the end of a prosodic phrase.

5) In turn, prosodic phrases are divided into prosodic words. The prosodic word corresponds, roughly speaking, to a content word. The alternation of lexical words with grammatical words (the latter realised less strongly, with lower F_0) plays a role in French prosody that is to some extent comparable with the alternation of stressed and unstressed syllables in English. The division of a prosodic phrase into two prosodic words is realised phonetically either by *final lengthening* at the end of the first word, a *strengthening* of the beginning of the following word (glottal onset in the case of an initial vowel), or again by an F_0 fluctuation aligned with the edge of one of the words. The feature “+ Strong” was borrowed from Straka with a view to covering the following manifestations: an F_0 jump, a longer syllable onset, glottalisation, lesser nasalisation, less voicing, a stronger contact of the articulators (Vaissière 1986; for more recent results on French, Fougeron 2001).



Figure 4. Typical subdivision of a prosodic phrase into two prosodic words by F_0 fluctuations. Left: Noun + Adjective sequence, right: Adjective + Noun (with indication of possible variants; see figure 6 for details).

Word-final lengthening can thus be the only marker of the division into prosodic phrases, without an F_0 excursion: experiments in synthetic speech show that longer duration of the first syllable is a sufficient cue to the distinction between *bordures* (“rims”), [bɔ:ɛdy::r], and *bords durs* (“hard edges”) [bɔ:ɛdy::r], between *Jean-Pierre et Jacques* (two names) and *Jean, Pierre et Jacques* (three names). (More precisely: if the first syllable is short, the phrase is ambiguous; beyond a certain length threshold, two prosodic phrases are heard; see Bacri and Banel 1993.) There are English equivalents, such as *coffee-cake and honey* vs. *coffee, cake and honey* (Lehiste, *op. cit.*).

2.3. The role of duration

Like F_0 fluctuations, the different degrees of final lengthening reflect prosodic structure; the two are not strictly equivalent, however. The use of prosodic parameters in automatic speech recognition has shown that informations on F_0 and on duration are both essential in French to distinguish between a left- or right-boundary. The consonant-to-vowel length ratio also has to be taken into account. The three ‘classical’ intonational parameters, F_0 , duration and intensity, must all be adduced (Vaissière 1988, Nasri 1992, Langlais 1995); indeed, they can be usefully supplemented by other information, such as positional allophonic variation. (Even so, in the end, automatic separation cannot be achieved in all cases.)

An important structural dimension of the sentence is encapsulated in *pauses* alone—what Monnin and Grosjean 1993 call ‘performance structure’. The degree of word-final lengthening distinguishes among at least three degrees of boundaries: sentence-internal prepausal lengthening (the longest :::), phrase-final lengthening (::), and word-final lengthening (:). Six levels of syllable rhyme length can usefully be distinguished: (1) lengthening at the end of a non-sentence-final breath group; (2) sentence-final lengthening; (3) phrase-final lengthening; (4) word-final lengthening; (5) default length, on initial syllables of lexical words; (6) shortened duration for grammatical words (and word-internal syllables). As for syllable onset, it is longer in the initial syllable of a lexical word (Duez and Nishinuma 1985).

Figure 5 presents a caricatural but revealing example, that of three sentences that are quasi-homophonous at the phonemic level: the sequence [sɛtɔmɛtenɔrmɛmɔbɛt] can correspond, depending on the intonation, to *Cet homme est énormément bête* (“This man is immensely stupid”), *Cet homme est énorme et m’embête* (“This man is huge and annoys me”), or to a highly implausible *Cet homme et Ténor m’aiment en bête* (“This man and Tenor love me as a

beast”). The figure presents a dot for each syllable: an asterisk “*” for grammatical words (which show a strong tendency to be shorter, with lower F_0 and intensity), and, for lexical words, a triangle for the initial syllable and empty circles for the next syllables, up to the final syllable (filled circle). The F_0 curve (measured from one sentence) is stylised by retaining only one point per syllable, except for the parts of the curve which correspond to a right boundary, if the syllable concerned is lengthened and therefore has the potential to carry an audible pitch movement within its rhyme): these are highlighted by a thicker line. Such examples allows for a direct comparison of observations on fundamental frequency and duration: it brings out clearly the fact that the syllable (more precisely the syllable *rhyme*) with the strongest rising slope corresponds to a major boundary (Delattre), and is accompanied by lengthening. This should be taken as an illustration of the *potential* of prosody as a cue to syntactic structure, rather than an indication of its actual role in communication, where there is hardly ever a threat of confusion between such pairs of sentences. The lengthened syllables are indicated.

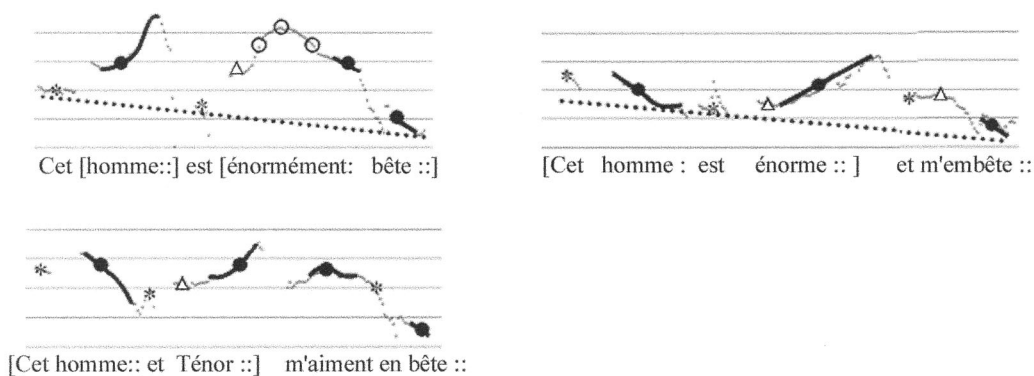


Figure 5. F_0 contour for sentences “Cet homme est énormément bête”, “Cet homme est énorme et m’embête” and “Cet homme et Ténor m’aiment en bête”, stylised using a star for grammatical words, a triangle for word-initial syllables (which can optionally host intonational intensification), an empty circle for lexical syllables that are neither initial nor final, and filled circles for word-final syllables. Each sentence is divided into two prosodic phrases (there is a full return to the baseline).

Within a sentence, the longer the final syllable of a word, the more rising its F_0 curve, and the stronger the perceived boundary.² Conversely, a falling contour indicates continuity, connectedness with what follows. Figure 6 illustrates how, in conjunction with the division into phrases, intonational variants convey semantic nuances, the syntactic dimension of intonation interacting with its pragmatic dimension.

² There exists, however, a variant of the same intonational morpheme, written as X↓ by Rossi 1999:73-75.

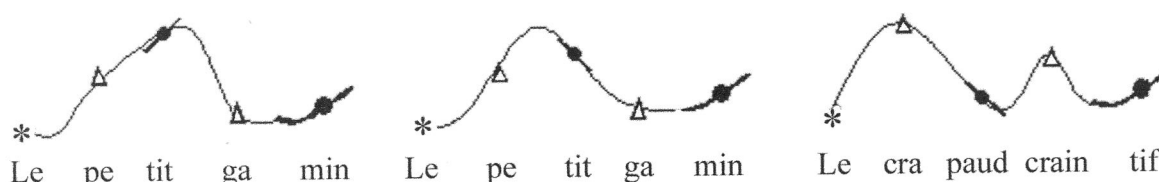


Figure 6. Left: a high-rising tune ('late peak') at the end of the first contour (le petit gamin) indicates a degree of semantic independence between the two words. Middle: a high-falling tune ('early peak') at the end of the first word indicates a degree of dependency. Right: a falling pattern indicates semantic dependency of the first word relative to the second. An initial jump may or may not be realized on the second word. If it is realized (as is here the case on *craintif*), it increases the perceptual distance between the two words; an adjective that comes after the noun it determines tends to be realised with an initial jump. (After Vaissière 2002.)

It seems, however, that the syntactic and pragmatic components of intonation as actualised by F_0 and durational variations do not always combine into a single, well-groomed 'phonological' structure.

2.4. The interplay of syntax, rhythm, and speaking style

The eurhythmic tendencies may in some cases prevail over syntax. The statement made in the previous section, that division into intonational components can be predicted from the syntax, is somewhat of an oversimplification. In early synthesis experiments at IBM France, two types of information were actually used: the syntactic bracketing of the sentence was supplemented by an indication of the number of syllables corresponding to each node in the structure. The need for the latter information clearly indicates the lack of a one-to-one correspondence between prosody and syntax: it originates in part in the rhythmic tendency to build prosodic units of roughly equivalent length (Gee and Grosjean 1983), and with a repetition of the same F_0 contour, stretched over words of different lengths. For instance, the major prosodic boundary (to use Delattre's term) tends to occur in-between the subject Noun Phrase and the Verb Phrase (at least in isolated sentences), but may be deferred until after the verb if the subject NP is short.

Some choices differ across speakers. Besides, given one syntactic phrasing and eurhythmic tendencies, intonational phrasing is to some extent left to the speaker's appreciation: a sentence may be realised at a go, without a clear division into MELODIC PHRASES by a *continuation rise*³, only subdividing the sequence of words into a series of equivalent prosodic phrases (the so-called parallel structure, Vaissière 1975). The more familiar a speaker is with a phrase, the lower the probability that (s)he will place an intonational boundary mid-way through the phrase: recordings of the phrase "l'Institut de Technologie du Massachusetts" reveals that readers who have some familiarity with the notion at issue hardly place any boundary within the group, whereas others divide it into up to three prosodic words. A speaker may freely choose among several rhythmic strategies: smoothing entire breath groups (which often goes hand in hand with a high speaking rate), separating individual prosodic

³ The continuation rise is an *intonational morpheme* (to use Rossi's phrase: see Rossi 1999) that marks a boundary.

phrases, or even bringing out sharply the division into individual prosodic words (in hyper-articulated, slow-rate elocution).

Style plays an important role. In public addresses (typically by journalists, politicians and teachers), *initial accentuation* is extremely frequent (...*le PRÉsident de la REpublique*...): it appears to be no other than the generalisation of *emphatic stress*, with the effect of conveying speaker involvement. Though originally a *stylistic* and not a *syntactic* intonational phenomenon, this emphatic stress, as it becomes almost systematic at the beginning of an intonational group, takes on some of the functional load of *intonational phrasing*, a striking instance of the interplay between syntactic and pragmatic intonation (see Lucci 1979). This also results in exceptions to the general principle according to which grammatical words are prosodically weak: the above example can just as well be realised as *LE président DE la république*, which is, if anything, still more emphatic. Emphatic stress may in fact be aligned with either the first or last syllable of disyllabic words (“Je vais a PaRIS, pas à Londres”, just as well as “Je vais à PaRIS, pas à Londres”).

Lastly, speaking rate also influences phrasing: the faster the rate, the more likely it is that the speaker will overlook the fine detail and bunch up several units together, resulting in groupings of 7 or 8 syllables, whereas in careful, deliberate speech, prosodic word boundaries are found every 3rd or 4th syllable (Vaissière 1971). As the rate of speech increases, a unit at one level may be progressively merged with a neighbouring (following) similar unit, the two being united by the melody into one larger unit.

Figure 7 shows three realisations of the phrase “Le président directeur général” (“the CEO”).

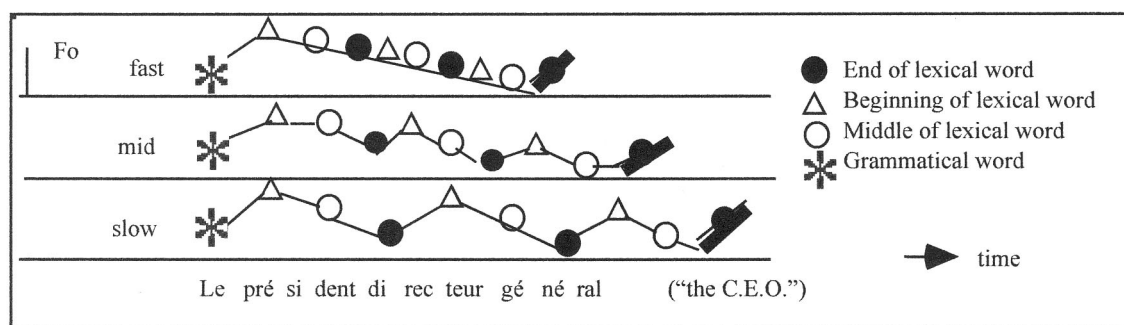


Figure 7. A schematic illustration of the influence of speaking rate on the division of the breath group into prosodic words (after Vaissière 1997)

Note that the term ‘major’ (as well as ‘minor’) is best adapted for the use of prosody in automatic speech recognition, when the rate of speech varies: the *major* rise, although sometimes reduced to a single peak on a word-final syllable, can still be detected as *the major boundary*, by comparison with what happens on the final syllables of the other words. Note that in very rapid and excited speech, the physiological constraints take over: the speaker only breathes in when actually out of breath.

3. The debate over the transcription of intonation

3.1. *French and English: a different perspective on prosodic parameters*

Approximating a sentence's prosody by means of its fundamental frequency alone yields reasonably acceptable results in English because, in this particular language, duration and intensity tend to be strongly correlated with F_0 : all three (F_0 excursions, lengthening and increase in intensity) tend to cluster on one and the same syllable—the most prominent syllable within an accentual phrase (Palmer 1922). By contrast, in French, there are (at least) two positions within a polysyllabic word that have a potential for hosting an *intonational morpheme*: the beginning of the word may receive emphatic stress (a morpheme of *intonational intensification*), which typically manifests itself phonetically by an articulatory strengthening of the consonant (resulting in an increase in the consonant-to-vowel length ratio) and an increase in subglottal pressure, among other correlates (Carton, Hirst *et al.* 1976, Fónagy 2001); the last syllable is where *an intonational morpheme that marks continuation* may be realised; it typically manifests itself by lengthening, an F_0 excursion, and a *decrease* in intensity, or at least no increase (Delattre 1938). To sum up, the complexity of the phenomena at issue is a formidable challenge.

3.2. *A model of intonation as superposition*

Experiments in speech engineering tend to support *models of intonation as superposition*. These models (which go back at least to Öhman 1967 and Fujisaki and Nagashima 1969) are referred to as “Contour Interaction models” by Ladd 1992, a somewhat restrictive term, since the emphasis of superpositional models is in fact not so much on the primitives of intonational description (contours *vs.* level tones) as on the recognition of the interplay of several levels of structure and the use of global and semi-global components, superimposed onto local ones. In using prosodic information for speech recognition, the degree of juncture between successive syllables is computed relative to the sum of junctures observed over the whole sentence; there is no fixed number of levels from the point of view of production—a standpoint which makes sense from a perceptual point of view, since it is known that listeners can go by fine details in their perception of boundaries (see Lehiste 1979).

The effects of prosodic groupings are not simply local. Sentence-finality affects (at least) the last two phonological words. The closeness of the link between successive words is to be estimated relatively to the realisation of broader constituents and to the speaker's habits.

4. Conclusion

The research community has now been made aware of the many functions of intonation (Fónagy, this volume). As Rossi 1999:9 points out, the issues of the relation of intonation and syntax, on the one hand, and intonation and pragmatics, on the other, have often been addressed, whereas interactions between syntax and pragmatics have received little attention. At present, most studies (including the present one) tackle only two or three factors, such as

phrasing and sentence mode, or *phrasing and contrastive accent* and describe a limited number of observed regularities. An increasing number of distinctive prosodic patterns (at the pragmatic level, and the expressive level) come to light, for French as well as for other languages (witness the abundance of papers about prosody in *Phonetica* and *Journal of Phonetics*, at the International Congresses of Phonetic Sciences and the Proceedings of the biannual Speech Prosody conferences); how they cohere together, and to what extent they make up a *system*, is not yet fully known.

Towards a typological overview: In view of the prominent role of intonational-syntactic boundaries in French, it is tempting to place this language in a typological category of *boundary languages*, as opposed to *stress languages* such as English (a suggestion put forward by Vaissière 2002). French is by no means exceptional in this respect: numerous languages (though admittedly less studied, e.g. Newar, a language of Nepal) are likewise non-tone, non-stress languages. However, the name 'boundary language' is misleading, in that all known languages (as far as we know) whatever their accentual system, have demarcative intonation, and are thus entitled to being called "boundary languages"; much as the opposition between "tone languages" and "intonation languages" is flawed, in that all languages, including tone languages, possess intonation (as is known since Chao Yuen-ren 1933 for Chinese, for instance). It therefore seems advisable to provide a negative typological characterisation of French, as a *non-tone, non-stress language*. It has been speculated that French is currently undergoing a major change in its prosodic system (Fónagy 1980); in the present state of the language, due to its flexibility, it lends itself to a host of complex, to some extent speaker-specific strategies (Fónagy 1982). In the 1970s, the project to achieve speaker-independent, rate-independent recognition of prosody in French ran up against the evidence of the variety of individual strategies: some speakers favour syntactic phrasing, others go mainly by pragmatic intonation, others tend to build units that have a roughly equal number of syllables, yet others favour a small set of contours which they tend to reproduce, sometimes as an alternation of rises and falls. This, however, does not detract from the regularity of syntactic intonation when considered in isolation. As synthesis by concatenation reaches its limits in terms of naturalness, specialists are again facing fundamental issues, and wish to feed more phonetic/linguistic knowledge into synthesis systems; this state of affairs may foster a growing interest in superpositional accounts of prosody.

References

- Armstrong, L. and I.C. Ward. 1926. *Handbook of English Intonation*. Cambridge: Heffner.
- Bacri, N. and M.-H. Banel. 1993. "Rhythmic patterns and lexical parsing in French." *Proceedings of ESCA Workshop on Prosody, Prosody-1993*, 120-123.
- Bolinger, D.L.M. 1978. "Intonation across languages." *Universals of Human Language*, vol. 2: *Phonology*, ed. by J.H. Greenberg, 471-524. Stanford: Stanford University Press.

- Carton, F., D. Hirst, A. Marchal and A. Séguinot, eds. 1976. L'accent d'insistance. *Studia Phonetica* 12. Montréal: Didier.
- Chao Yuen-ren. 1933. "Tone and intonation in Chinese." *Bulletin of the Institute of History and Philology* 4:3.121-134.
- Coustenoble, H. and L. Armstrong. 1937. *Studies in French intonation*. Cambridge: Heffer.
- Delattre, P. 1938. "L'accent final en français: accent d'intensité, accent de hauteur, accent de durée." *The French Review* 12:2.141-145.
- Delattre, P. 1965. *Comparing the phonetic features of English, French, German and Spanish: An interim report*. Heidelberg: Julius Groos Verlag.
- Delattre, P. 1966. *Studies in French and Comparative Phonetics*. The Hague/London: Mouton and co.
- Di Cristo, A. 1998. "Intonation in French." *Intonation systems: a survey of twenty languages*, ed. by D. Hirst and A. Di Cristo, 195-218. Cambridge: Cambridge University Press.
- Duez, D. and Y. Nishinuma. 1985. "Le rythme en français: alternance des durées syllabiques." *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, vol. 10, 151-169.
- Fónagy, I. 1980. "L'accent français, accent probabilitaire: dynamique d'un changement prosodique." *L'accent en français contemporain*, ed. by I. Fónagy and P. Léon: *Studia Phonetica* 15.
- Fónagy, I. 1982. "Variation et normes prosodiques." *Folia Linguistica* XVI:1-4.19-39.
- Fónagy, I. 2001. *Languages within language*. Amsterdam-Philadelphia: Benjamins.
- Fougeron, C. 2001. "Articulatory properties of initial segments in several prosodic constituents in French." *Journal of Phonetics* 29:2.109-135.
- Fujisaki, H. and S. Nagashima. 1969. "A model for the synthesis of pitch contours." *Annual Report of the Engineering Research Institute, University of Tokyo* 28.53-60.
- Gee, J.P. and F. Grosjean. 1983. "Performance structures: A psycholinguistic and linguistic appraisal." *Cognitive Psychology* 15.411-458.
- Gendrot, C. 2005. *Aspects perceptifs, physiologiques et acoustiques de différentes catégories prosodiques en français*, Ph.D., Université de la Sorbonne Nouvelle, Paris.
- Grammont, M. 1933. *Traité de phonétique*. Paris: Delagrave.
- Gussenhoven, C. 2002. *Intonation and Interpretation: Phonetics and Phonology*. *Proceedings of Speech Prosody 2002*, ed. by B. Bel and I. Marlien, Aix en Provence, 47-58.
- Hallé, P., D. de Boysson-Bardies and M.M. Vihman. 1991. "Beginnings of prosodic organization: intonation and duration patterns of disyllables produced by Japanese and French infants." *Language and Speech* 34:4.299-318.
- Hirst, D. and A. Di Cristo. 1998. *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press.

- Ladd, R. 1992. "An introduction to intonational phonology." Papers in laboratory phonology II: Gesture, segment, prosody, ed. by G. J. Docherty and R. Ladd, Cambridge U.K.: Cambridge University Press.
- Langlais, P. 1995. Traitement de la prosodie en reconnaissance automatique de la parole. Ph. D., Université d'Avignon.
- Lehiste, I. 1975. "The phonetic structure of paragraphs." Structure and Process in Speech Perception, ed. by A. Cohen and S. G. Noteboom, 195-206. Berlin: Springer.
- Lehiste, I. 1979. "Perception of sentence and paragraph boundaries." Frontiers of Speech Communication Research, ed. by B. Lindblom and S. Öhman, 191-201. London: Academic Press.
- Lieberman, P. 1967. Intonation, Perception and Language. Cambridge, Massachusetts: MIT Press.
- Lucci, V. 1979. "L'accent didactique." *Studia Phonetica* 15.107-121.
- Maeda, S. 1976. A Characterization of American English Intonation. Ph. D. dissertation, M.I.T., Cambridge, MA.
- Martin, Ph. 1981. "Pour une théorie de l'intonation." L'intonation, de l'acoustique à la sémantique, ed. by M. Rossi, A. Di Cristo, D. Hirst, P. Martin and Y. Yishinuma, 234-271. Paris: Klincksieck.
- Monnin, P. and F. Grosjean. 1993. "Les structures de performance en français: caractérisation et prédiction." *L'Année Psychologique* 93.9-30.
- Nasri, M.K. 1992. L'architecture du système de reconnaissance automatique de la parole DIRA. Thèse de Docteur ingénieur, Université de Grenoble.
- Öhman, S. 1967. Word and sentence intonation: a quantitative model. Stockholm: Speech Transmission Laboratory Quarterly Progress and Status Report, KTH, 2-3. 20-54.
- Palmer, H.E. 1922. English Intonation, with Systematic Exercises. Cambridge: Heffer.
- Pierrehumbert, J. 1979. "The perception of fundamental frequency declination." *Journal of the Acoustical Society of America* 66.363-369.
- Rossi, M. 1967. "L'accent, le mot et ses limites." Nouvelles perspectives en phonétique, ed. by B. Malmberg, D. B. Fry and R. Lancia, 81-85. Bruxelles: Presses Universitaires de Bruxelles.
- Rossi, M. 1999. L'intonation, le système du français: description et modélisation. Gap/Paris: Ophrys.
- Selkirk, E. 1972. The phrase phonology of English and French. Cambridge, Massachusetts: MIT.
- Selkirk, E. 2000. "The interaction of constraints on prosodic phrasing." Prosody: Theory and Experiment, ed. by M. Horne, 231-261. Dordrecht: Kluwer Academic Publishers.
- Thorsen, N. 1980. "A study of the perception of sentence intonation: Evidence from Danish." *Journal of the Acoustical Society of America* 67.1014-1030.

- Vaissière, J. 1971. Contribution à la synthèse par règles du français. Ph. D., Université de Grenoble.
- Vaissière, J. 1975. "Further note on French prosody." *Research Laboratory of Electronics, MIT, Quarterly Progress Report* 115.251-262.
- Vaissière, J. 1976. Quelques analyses perceptives en français. Proceedings of VIII^e Journées d'Etudes sur la Parole, Aix-en-Provence, 193-208.
- Vaissière, J. 1983. "Language-independent prosodic features." Prosody: Models and Measurements, ed. by A. Cutler and R. Ladd, 53-66. Berlin: Springer Verlag.
- Vaissière, J. 1986. "Variance and Invariance at the Word Level." Invariance and Variability in Speech Process, ed. by J. S. Perkell and D. Klatt, 534-539. Lawrence Erlbaum Associates.
- Vaissière, J. 1988. "The use of prosodic parameters in automatic speech recognition." Recent advances in speech understanding and dialog systems, Berlin: Springer.
- Vaissière, J. 1995. "Phonetic explanations for cross-linguistic similarities." *Phonetica* 52.123-130.
- Vaissière, J. 1997. "Langues, prosodie et syntaxe." *Traitement Automatique des Langues* 38:1.
- Vaissière, J. 2002. "Cross-linguistic prosodic transcription: French vs. English." Problems and methods of experimental phonetics. In honour of the 70th anniversary of Pr. L.V. Bondarko, ed. by N. B. Volskaya, N. D. Svetozarova and P. A. Skrelin, Moscow.