



**HAL**  
open science

## Earlier or Higher? Comparing French rising-falling contour with rising contour in a corpus of conversation

Cristel Portes, Leonardo Lancia

► **To cite this version:**

Cristel Portes, Leonardo Lancia. Earlier or Higher? Comparing French rising-falling contour with rising contour in a corpus of conversation. *Journal of Phonetics*, 2017, 63, pp.35 - 52. 10.1016/j.wocn.2017.04.002 . hal-01569748

**HAL Id: hal-01569748**

**<https://hal.science/hal-01569748>**

Submitted on 27 Jul 2017

**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.

# Earlier or Higher? Comparing French rising-falling contour with rising contour in a corpus of conversation

**Authors:** Cristel Portes<sup>a</sup>, Leonardo Lancia<sup>b</sup>

**Affiliation:** <sup>a</sup> Aix-Marseille Université, CNRS, LPL, UMR 7309, 5 Avenue Pasteur, BP 80975 - 13604 Aix-en-Provence, France

<sup>b</sup> Laboratoire de Phonétique et Phonologie (UMR7018, CNRS-Sorbonne Nouvelle), 19 rue des Bernardins, 75005, Paris, France

**E-mail addresses:** cristel.portes@lpl-aix.fr, leonardo.lancia@univ-paris3.fr

## **Corresponding author:**

Cristel Portes

Laboratoire Parole et Langage

5 Avenue Pasteur

BP 80975 - 13604 Aix-en-Provence, France

Telephone number: +00 33 6 13 04 80 50

Fax number: +00 33 4 13 55 37 44

Email : cristel.portes@lpl-aix.fr

**Keywords:** intonation; intonational meaning; alignment; scaling; corpus analysis; French; Wavelet based Functional Mixed Models

## **Abstract:**

In French intonation, a rising-falling contour (RF) has been described by many authors, but the characteristics of its phonetic realization as well as its phonological status remain controversial. Is its f<sub>0</sub> movement temporally aligned earlier compared to the f<sub>0</sub> movement of the simple rise (R)? Or is it scaled higher in the speaker's pitch range? Does it convey conviction and obviousness while the simple rise rather announces that the speaker has more to say? Firstly, the present study compared the phonetic implementation of RF and R in a corpus of naturally occurring conversation. Through the application of a Wavelet based Functional Mixed Model, we could detect significant differences between the shapes of the f<sub>0</sub> curves corresponding to RF and R contours. Results show that RF and R mainly differ with

respect to the timing of the rise and the amplitude of the falling part. They thus support the claim that these characteristics are more important than the scaling of the pitch peak for the implementation of the contrast between RF and R. Secondly, the results of a forced choice identification task performed by naïve listeners show that they consistently associate the RF contour with the expression of conviction and obviousness and the R contour with the indication that the speaker has more to say.

## 1. Introduction

The study of intonational contours (or *tunes*) is at the core of intonational phonology. A large body of studies is hence devoted to, or at least includes, the inventory of phonological contours in a given language or in a family of languages (see Hirst & Di Cristo, 1998 or Jun, 2005 for typological surveys). Most of these studies are based on the impressionistic observation of selected examples from corpora of spontaneous speech and/or on quantitative analyses of productions elicited in laboratory conditions, but very few rely on a quantitative investigation of naturally occurring corpus data. With the arrival of new tools, the trend in linguistics has turned to exploring larger corpora (Schafer, Speer & Warren, 2005; Bresnan 2007; Coleman, Liberman, Kochanski, Burnard, & Yuan, 2011; Cole & Hasegawa-Johnson, 2012; Jun & Fletcher, 2014). Such an approach provides us with access to linguistic phenomena in their ecological context, and allows a better understanding of the full range of variation in speaker behavior.

In French, one of the intonational contours that can benefit greatly from a corpus quantitative investigation is a rising-falling contour, first described by Delattre (1966) who called it an “intonation d’implication” (implicative intonation). As a matter of fact, its phonetic description and phonological characterization remain controversial in French intonational literature. Specifically, the fact that it contrasts with the rising contour has been said to rely mainly on an earlier temporal alignment of the rising part of the contour (Post, 2000), or mainly on a higher implementation of the pitch peak in the fundamental frequency range (Di Cristo & Hirst, 1996; Rossi, 1999). The aim of the present study was therefore to examine the precise phonetic implementation of this French rising-falling contour (henceforth RF) in a corpus of natural conversation, while comparing it with the rising contour (henceforth R). We pursued our aim in three steps. First, we asked two French phoneticians to identify the instances of RF and R contours in a radio conversation involving five male speakers. Second, we modelled the difference between the two contours via a Wavelet-based Functional Mixed Model (Morris and Carrol, 2006). This technique allowed us to estimate the difference between shapes of  $f_0$  trajectories in a mixed model framework, and therefore to implement a complex random effect structure, as required by the unbalanced nature of corpus data. Third, we validated the classification of the two expert listeners with that of 34 naïve French listeners involved in a forced choice identification experiment.

In the remainder of this introduction, we will review the description of the rising-falling contour in the literature on French intonation (section 1.1.), give some background about alignment and scaling of tonal targets (section 1.2.), and formulate the aims of the study (section 1.3.). In section 2, we describe our corpus study. And in section 3, we report the methods and results of the follow-up perception experiment. Section 4 consists in a general discussion, while section 5 presents concluding remarks.

### 1.1. The rising-falling contour in French intonation

The rising-falling contour under study here is part of the inventory of the nuclear contours (or intonational phrase final contours) of the French intonation system. The matrix of RF is Delattre’s “contour d’implication” (Delattre, 1966) schematized in Figure 1 below, borrowed from the author. Delattre conceived the contour as a rising-falling one, rising through the penultimate syllable, with the f0 peak occurring on the final (primary accented) syllable and then falling slightly on that same syllable.

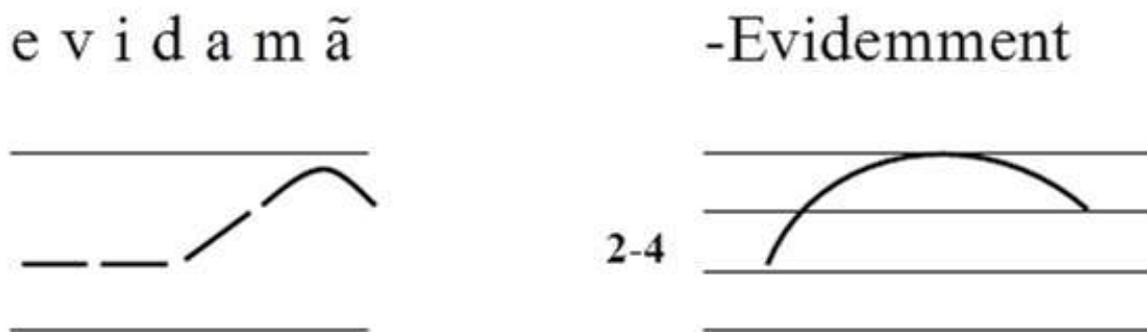


Figure 1. Delattre’s representation of the “contour d’implication”. The graph on the left represents the f0 curve realized on the utterance “Evidemment.” (*obviously or of course*). The right graph represents the stylized form of the contrastive part of the curve “on the last vowel of the meaningful group” (Delattre, 1966: p. 11): it rises between levels 2 and 4 and then falls from level 4 to 3 of the speaker’s range.

According to Delattre, the meaning of the contour links the meaning of the actual utterance to some implicit content which is quite underspecified and can be inferred by the context and may convey different pragmatic attitudes such as obviousness, exasperation or on the contrary politeness.

Although most alternative accounts of French intonation have mentioned a rising-falling contour, its status and description is as yet far from consensual. The main points of discussion, discordance and ambiguity addressed in the present study are the role of f0 temporal alignment, the role of f0 height and the role of the f0 level reached by the fall in the speaker’s range as cues to distinguish RF from the simple rise R. An additional ambiguity concerns the contrast of RF with another rising-falling contour in which the f0 peak is aligned with the penultimate syllable. Post (2000) encoded the phonological difference between the two contours using a

pitch accent contrast, namely H\* in the LH\*L% tune for the RF under study here and H+H\* in the LH+H\*L% tune for the penultimate peak contour. Martin (1999) also proposed to distinguish two rising-falling contours (“contours à courbe en cloche” *bell curve contours*) which share the feature “+ circonflexe” (+ *circumflex*) but differ concerning the feature +/- rising (Martin, 1999). However, other authors did not mention this contrast, for instance, Vaissière (1980) only mentioned one rising-falling contour called a “contour à pic” (*peak contour*). As for Mertens (2008), he actually proposed a contrast between two rising-falling contours HB and HB- (“H” for High and “B” for Bottom), but they differ concerning the level at the bottom of the speaker’s range which is higher in HB and lower in HB-. Unfortunately, we were not able to analyze the contrast between RF and the penultimate peak contour in the present study due to the lack of relevant data (see section 3 below).

There are a number of controversial elements to be taken into account, beginning with Post’s (1999) study of the contrast between RF and R. In order to explore a three-way phonological contrast between LH\*L% (RF), LH\*H% (R) and LH\*0% (a variant of R rising to mid versus to high), Post (1999) conducted a categorical perception experiment (discrimination and identification tasks) on a continuum between RF and R (Figure 2).

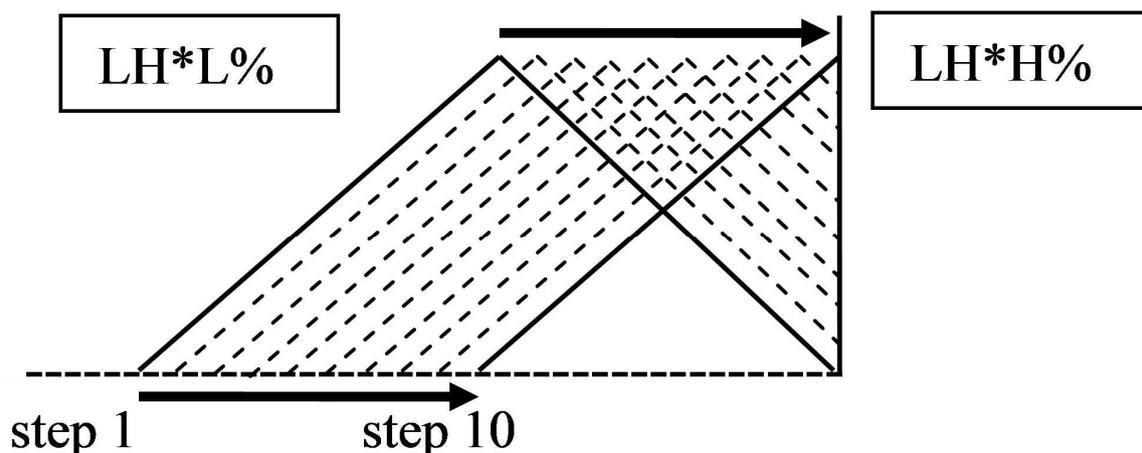


Figure 2. Design of the 10-step continuum between the rising-falling contour LH\*L% (RF) and the simple rise contour LH\*H% (R) used in the categorical perception experiment reported in Post (1999).

As shown in Figure 2, in Post’s proposal, the alignment of both the leading L and the H\* targets are delayed from RF to R; on the other hand, the fall of the RF contour becomes shorter and reaches higher final f0 values when moving from step 1 to step 10 of the continuum, to disappear at step 10 corresponding to the R contour; finally, the height of the f0 peak remains constant in all steps of the continuum. In Post’s 1999 categorical perception study, this continuum was used to test the validity of the experimental procedure, as the contrast between

RF and R was presented as uncontroversial. However, the results of the identification task did not support a categorical perception of the RF and R contours. The results could “be interpreted to indicate that the differences were perceived continuously” (Post 1999, p. 968). This result may be due to the fact that RF is a variant of R. However, this explanation is not in line with a general claim in the literature that the two contours have a different meaning: RF is known to convey an epistemic meaning (Portes & Reyle, 2014; Delais-Roussarie et al., 2015; Sichel-Bazin, 2015) while R conveys continuation or a polar question. One of the aims of the present study is to verify the validity of this form-meaning association and to test whether it is only a broad preference between two possible pragmatic functions or a stronger phonological contour/meaning relation. Another explanation may be that the stylized representation of the contours used in the task, does not control for the relevant cues or at least not for all of them.

As a matter of fact, many authors mentioned emphasis as a property of the rising-falling contour. In the second edition of his book *Intonational Phonology*, Ladd described Delattre’s “intonation d’implication” as “matrix sentences ending with an emphatic or exclamatory high peak” (Ladd, 2008, p. 121). Rossi (1981; 1999) conceived the rising-falling contour as the result of a clustering between one of his *intonational morphemes* (“intonèmes”) called *major conclusion* (“conclusive majeure”) and a special feature called “expressème”, whose role was to add expressivity to the meaning of a falling declarative or interrogative contour by raising it up in the speaker range. The same idea was developed by Di Cristo and Hirst (1996) who distinguished between two rising-falling contours on the basis of the height reached by the f<sub>0</sub> peak: *the contrastive emphasis* (“l’emphase contrastive”) with a very high peak and *the marked assertive variant* (“la variante assertive marquée”) with a normal high peak. In the latter approaches it was hence made explicit that the height of the contour was responsible for its emphatic/expressive function.

As we have already mentioned above, a further distinction was proposed by Mertens (2008) who contrasted two rising-falling contours, HB and HB<sup>-</sup>, on the basis of the lower f<sub>0</sub> values reached at the end of HB<sup>-</sup>. Both contours were said to be used to mark focus, or at least attention centers, and to convey speaker commitment. The functional difference concerned the strength of the discourse frontier, which is maximal for HB<sup>-</sup> but not for HB. Conversely, Ladd suggested that the fall from the peak “seldom, if ever, reaches the bottom of the range” (Ladd 2008, p. 122), negating a phonological role to that dimension of variation.

On the other hand, Mertens (2008) and Ladd (2008) agreed that RF can be followed by a deaccented sequence. So did Jun and Fougeron (2000) who evoked a rising-falling contour as the combination of the final rise LH\* of the Accentual Phrase with the low phrasal tone L-spreading over a deaccented segment until the end of the utterance marked by a low boundary tone L%. In this latter approach, the occurrence of the rising-falling contour is hence dependent

---

<sup>1</sup> This is in line with Post’s (1999) hypothesis of a distinction between LH\*0% (corresponding to HB) and LH\*L% (corresponding to HB<sup>-</sup>).

on the occurrence of deaccentuation on the right-hand part of the utterance. Deaccentuation after the rising-falling contour is mentioned by several authors and is common in corpora. However, since we found no occurrence of this case in our corpus, we will not address this point in the present study.

This short overview shows that the different accounts of the French intonation system differ in how many rising-falling contours they distinguish (one or two) and along which dimensions (alignment of the peak, scaling of the peak, or scaling of the fall). However, they all systematically build a contrast between the rising-falling contour and the rising contour concerning the variation of the two crucial phonetic dimensions for intonation that the autosegmental-metrical theory (AM) enabled to posit explicitly: the pitch scaling of the f<sub>0</sub> targets and their temporal alignment with the segmental string. These are therefore the two dimensions studied in the corpus analysis presented in this paper.

## 1.2. Alignment and scaling of tonal targets

The AM theory assumes that the primitives of the phonological analysis of intonation are not the local f<sub>0</sub> movements (rises and falls) but rather the turning points of the f<sub>0</sub> curve conceived as tonal targets or tones (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988). These tones are pitch accents (T\*) or edge tones (T-, T%) depending on their phonological association with accented syllables or constituents of the prosodic structure. Constraints on the scaling and the alignment of tones in different languages have been extensively studied in the AM literature (see Ladd, 2008, for a review and an extensive discussion).

Concerning scaling, the higher f<sub>0</sub> values than those of prenuclear H\* that are observed at the edges of major prosodic phrases like ip boundaries and IP boundaries are modeled as deriving from an upstep rule applying to the edge tones associated to these boundaries, i.e. H- for the ip and H% for the IP (Ladd, 1988; Van den Berg, Gussenhoven & Rietveld, 1992; Truckenbrodt, 2002, 2007; Post, 2000 and D'Imperio & Michelas, 2015). Since R is modeled as LH\*H% in AM models of French intonation, upstep should apply to R but not to RF which has a L% boundary tone instead of a H%. This should lead to a higher scaling of the pitch maximum for R (corresponding to H%) compared to the pitch maximum for RF (corresponding to H\*), contrary to what is predicted by Di Cristo and Hirst (1996) and Rossi (1999).

As for alignment, the segmental anchoring hypothesis relates the phonological association of tones to the stability of their realization in the vicinity of their segmental anchor points (Arvaniti et al., 1998; Ladd et al., 1999; Xu, 1998; Atterer & Ladd, 2004; Schepman et al., 2006; Welby, 2006). If RF and R align differently with respect to the segmental string as proposed by Post (1999), their L and H\* tones should be anchored to different segmental landmarks.

Although our corpus data cannot be controlled enough to directly address the issues raised by these authors, they will nevertheless enable us to discuss some of their predictions.

## 1.3. Aims of the study

In this study, we conducted two different investigations in order to better capture the phonetic and perceptual differences between RF and R. In a production study, we looked at the phonetic realization of RF and R in a corpus of naturally occurring conversation when they were accessible for ecological observation. Based on an auditory identification of the contours by two experts, we defined which of the contours correspond to RF and R in our corpus. We then analyzed the surface realizations of the identified RF and R thanks to a qualitative analysis. We then compared RF to R in order to decide between the different distinctive features proposed in the literature. Specifically, we verified whether the phonetic implementation of the contours confirms the predictions suggested by Post's (1999) study that the early alignment of the rise (the initial L and H\* tonal targets) is the major feature of RF or conversely, if it gives support to the alternative views by Rossi (1999) and Hirst and Di Cristo (1996) who attributed the major role to the height of the peak. We also examined if the final height of the falling part of RF plays any significant role as suggested by Mertens (2008) and Post (2000). In a perception experiment, we examined if the RF and R categories provided by experts' identification was confirmed by naïve listeners during a categorization task. By using the difference in meaning between RF (conviction/obviousness) and R (continuation) proposed in the literature as the discriminatory factor between the two categories, we also tested the validity and the strength of the contour/meaning relation.

## 2. Corpus based production study

### 2.1. Which corpus and why?

Pierre Delattre (1966) claimed that his “intonation d'implication” is very common in conversation. Our exploration of different types of naturally occurring data revealed that it is even more frequent in one particular type of conversation, namely in argumentative interactions. For instance, many occurrences of RF can be heard in political discourse (Bigi et al. 2012) or in different kinds of debate. This high frequency may be related to the meaning of the contour which is frequently formulated in terms of “evidence”, “certainty”, “rising exclamation” in the French prosodic literature (Delattre, 1966; Rossi, 1999; Mertens, 2008). Moreover, RF has been shown to have argumentative and polemical functions (Portes, 2004; Portes & Reyle, 2014).

The chosen corpus is therefore an item of a radio debate broadcast every Saturday on the French public radio station France Culture and entitled “La suite dans les idées” (French idiomatic expression which could be translated as *Train of thought*). The episode under study was recorded on 25<sup>th</sup> February 2000 using a TASCAM 122MKIII cassette deck, and subsequently digitalized with an ADC & DAC 3000 converter at Laboratoire Parole et Langage in Aix-en-Provence, France. The sampling frequency of the recording is 16000 Hz with a 16 bits dynamic. The recording lasts around 45 minutes. The title of the episode was “L'Europe et son élargissement, question taboue?” (*Is the issue of Europe and its expansion a taboo issue?*). Four guest speakers of different professions and opinions were moderated by two journalists debating the question. All six speakers were recorded on the same channel. Only one of them, a

journalist, was a female. A further advantage was that the corpus includes instances of the contours of interest produced by the six different speakers.

## 2.2. Specificity of corpus data analysis and chosen solutions

Working with corpus data has some disadvantages. In particular, it does not allow for the same detailed control of the data as in a laboratory unless the number of relevant items is huge. Additionally, naturally occurring data shows high levels of variability that must be constrained in different ways.

### 2.2.1. Contour identification

An appropriate method of identifying the instances of both RF and F contours in the corpus in a replicable fashion was required. In order to avoid circularity and make the identification reliable, the first author of the present study was not the only person involved in the identification, another expert of French intonation, who was not otherwise involved in the study, was asked to perform the same task in the same conditions. Delattre's inventory of French intonation contours (Delattre, 1966) was used for the identification task because both experts knew it well. The complete results of the identification task are available in Portes (2004).

Among the intonation contours of the whole corpus, 114 RF and 460 R were identified. Actually, 62 RF contours were identified as such by both experts, but also 38 RF contours were identified by only the first expert, and 14 others identified by only the second. Among these 52 ambiguous RF, the ambiguity concerned a fall in 19 % of the cases (10 items) and a major continuation rise (i.e. R) in 81 % of the cases (42 items). The 10 items identified as a fall by one of the experts all had a peak aligned with the penultimate or even the antepenultimate syllable like the examples displayed in Figure 6 and 8 in section 2.5.1. below. The fact that the ambiguity mainly concerned RF versus R gives additional relevance to the comparison between these two contours, and confirms a potential ambiguity zone between them. We therefore decided to separate the ambiguous contours from the consensual ones in the comparison. Hence, the consensual RF contours, those that have been identified as such by both experts, were henceforth labeled RF, while the ambiguous contours between RF and R, those that have been identified as RF by only one of the experts and as R by the other, were labeled Amb. Indeed, the differences observed in the ambiguous Amb allowed us to collect information about the possible variability in the realization of RF in the space involving the two final syllables of the utterance, which is the location where the contrast between the tunes takes place in French.

In our corpus, the number of R identified by the experts (460 items) were four times the number of RF (114 items). This is not surprising since, in this data, R is mainly used as a clause chaining device within long monological turns alternating between the six speakers. In order to restrict the number of R for the sake of comparison, we selected 177 R among those that were consensually recognized by both experts. The choice was driven by two criteria. Firstly, because RF was always final of a speech act in the corpus, we discarded the instances of R which were not final of a speech act (or discourse move). This selection was made to guarantee that the

chosen R contours were intonational phrase final contours instead of contours associated with lower levels of constituency of French prosodic phrasing (such as the accentual phrase AP or the intermediate phrase ip, see Jun & Fougeron, 2000; Michélas & D’Imperio, 2012; i.a.). Secondly, we retained each instance of R whose final syllable duration and final syllabic nucleus duration remained in the ranges of values observed for RF contours. This was done by retaining all the R items such that the standard deviations of these two parameters obtained by considering all the selected instances of both RF and R contours did not exceed the standard deviations obtained by considering only RF contours.

### 2.2.2. Contour variability

Among the identified contours, a certain amount of variability remains due to different factors such as the effect of prosodic context on the surface realization of the contour (e.g. plateau realization due to the occurrence of a preceding initial rise) or ambiguity with closely related contours (e.g. the similarity of RF with another rise-fall whose f<sub>0</sub> peak is aligned with the penultimate syllable, not with the final one). To account for this kind of variability, we conducted a qualitative analysis of the 114 identified RF items. Specifically, thanks to the visual inspection of each f<sub>0</sub> curve, we classified the items following their global shape. The results are displayed in section 3.1 below.

### 2.2.3. Segmental variability

It is difficult to control for the segmental content of the target syllables in naturally occurring data. This limitation introduced variability in f<sub>0</sub> curves due to the heterogeneity of the segmental content and to the presence of unvoiced segments, which is a problem when working on intonation. To reduce the effects of these limitations, only the voiced parts of the syllables were submitted to statistical modeling.

## 2.3. Data processing

The 45-minute recording was phonetically aligned using an analysis by synthesis method (Malfrère, Deroo, Dutoit & Ris, 2003) and the SYNTAIX text-to-speech system (Di Cristo & Di Cristo, 2001). After the automatic procedure, the alignment was manually corrected. Then, a label identifying each item of the selected contours was manually localized on the sound track at the very end of each utterance carrying one of these contours.

## 2.4. Method for the quantitative comparison of RF, Amb and R

### 2.4.1. Data selection and classification

For the quantitative analysis, we made a selection of the data. First, we discarded the non prototypical items of RF and Amb identified through the qualitative analysis (see section 3.1 below). We also decided to remove the 4 items produced by the unique female speaker of the corpus, and one additional item because of detection problems. A total of 100 tokens of the rising-falling contour were thus kept for the analysis with 58 items classified as RF and 42

classified as Amb. From the selected 177 consensual exemplars of R, we had to remove 9 items that were produced by the female speaker and 4 items because of f0 detection problems. We therefore kept 164 tokens of R for the quantitative analysis.

#### 2.4.2. f0 processing

To compare the f0 curves from consensual RF, ambiguous contours Amb and consensual R, we measured f0 in the portions of signal corresponding to the nuclei of the penultimate and of the last accented syllable (with a time-step of 1 ms). We then linearly normalized the length of each f0 contour to a constant value corresponding to 100 time-points and we concatenated the f0 curves of the two consecutive nuclei. In this way, each obtained curve of length equal to 200 time-points represented the evolution of the f0 on the penultimate and last accented nuclei of one observed accentual domain.

#### 2.4.3. Design of the statistical model

To evaluate the differences between the contours, we submitted the obtained f0 curves to a Wavelet-based Functional Mixed Model (Morris & Carroll, 2006). The model was designed to estimate the effects of the following predictors on the shape of the f0 curves: the nature of the contour (RF, Amb and R, with Amb as reference level), the presence vs. absence of a coda in the last syllable (coda vs. no-coda, with no-coda as reference level) and their interaction. We chose to include the presence vs. absence of a coda in the last syllable because the presence of a coda modifies the syllable boundary location, which is a particularly crucial factor for tonal alignment (Silverman & Pierrehumbert, 1990; Prieto, van Santen & Hirschberg, 1995; Arvaniti, Ladd & Mennen, 1998, Prieto & Torreira, 2007, i.a.). Although both common wisdom and available literature suggest that the presence and the type of coda consonant(s) in the penultimate syllable, the presence and type of attack consonant(s) in the last accented syllable, the type of the syllabic nucleus and of the coda in that same syllable may all affect to some extent the alignment of the f0 peak, we could not test the effects of these factors. In order to conduct such an extensive analysis, we would need instances of each contour type for each combination of levels of the potentially relevant factors. This was not the case with the corpus under study. Nevertheless, in order to discard the hypothesis that our results depend on the uneven distribution of these features across the three contours under study, we included these factors in the random effects structure of the model. The random intercepts defined were: the identity of the speaker, the identity of the vowel in the last accented syllable, the kind of coda of the penultimate syllable, the kind of onset of the last syllable and the kind of coda of the last syllable. The types of onset and coda consonants were coded according to the presence or absence of a consonant, its sonorant/obstruent nature and the presence of voicing. Due to the adopted coding, each level of these random intercepts represented a particular combination of these three features. An additional level for these random intercepts was introduced to account for consonant clusters (which in our data contain always a voiced-plosive and a liquid). The random effect structure also included one speaker-specific random term for each of the predictors.

#### 2.4.4. Wavelet based functional mixed models

In order to model the effects of controlled factors on time-series sampled on the same equally spaced grid of length  $t$ , we can design a mixed model that simulates values arranged in a vector of length  $t$  as a function of the fixed and random factors and of a noise term. In such a model, each level of the factors provides  $t$  coefficients (each coefficient contributing to the estimation of one point in time). With respect to running one mixed model for each time point considered, a unique model permits capturing the correlation between the values of the same time-series, so that the estimation at a given point in time is informed by nearby points. However, this comes with the price that highly dimensioned covariance matrices need to be estimated, limiting in practice the possibility of fitting complex models.

Morris and Carrol (2006) address this issue by representing the observed time-series via the Discrete Wavelet Transform (DWT) coefficients and by fitting the model on the transformed data. Wavelets are time-varying patterns having the same shape (called the mother wavelet) but differing in duration (i.e. they are scaled versions of the mother wavelet) and located at different time points (i.e. they are translations of the scaled mother wavelet). Wavelet coefficients can be obtained by computing the inner product between the observed time series and the scaled and time-shifted wavelets. In practice, by representing an  $f_0$  curve via wavelet coefficients we are assuming that the time series can be obtained by combining several repetitions of the same pattern with different durations and located at different points in time. A parsimonious representation of an observed curve can be obtained through the Discrete Wavelet Transform (DWT) algorithm introduced by Mallat (1989), which represents a time series of length  $t$  through  $t$  coefficients. Similarly, to Fourier coefficients, the absolute value of each DWT coefficient corresponds to the contribution of a different wavelet (with given scale and location) to the shape of the  $f_0$  curve. Although the DWT coefficients representing a given  $f_0$  curve, taken separately do not refer to any precise feature of the curve, due to the whitening property of the DTW, coefficients are uncorrelated across scales and locations. When running a mixed model whose response is a vector of  $t$  uncorrelated coefficients, we can make the assumption that the covariance matrices are diagonal. This assumption reduces enormously the number of parameters to be estimated, making the estimation of complex models computationally feasible.

Given a model describing the mathematical relation between some factors affecting the data and the values observed, we are usually interested in estimating the model parameters, which quantify the effects of the considered factors on the observed data. In Bayesian estimation, the model parameters are treated as random variables, having their own probability distributions. More precisely, each parameter is associated with two probability distributions: the prior distribution indicates how probable we think that each possible value of the parameter is before looking at the data. The posterior distribution combines the information contained in the data with that contained in the prior distribution and indicates how probable is each value of the model parameter given the observed data and the prior distribution. In order to obtain an estimate of the parameter value we can compute the mode of the posterior distribution (which is the value with the highest posterior probability) or alternatively we can compute its mean.

The 95% of the posterior distribution is used to derive the high posterior density interval or most simply the credibility interval, which determines the range of probable outcomes for the parameter of interest. Deciding whether an estimated parameter value is significantly different from a null value (i.e. if an effect is significant), is as simple as checking if the null value is contained or not in the credibility interval for that parameter.

Since in Bayesian inference the posterior distribution is obtained by combining the information contained in the data with the prior distributions of the model parameters, we need to define a prior distribution for each parameter in the model. In any mixed model we need to estimate two kinds of parameters: the parameters quantifying the fixed effects and the variance parameters. Usually two different kinds of prior distributions are hypothesized. In the approach adopted by Morris and Carroll (2006), the prior distributions for the fixed effects are mixtures of a normal and a Bernoulli distribution with mass at 0 (aka spike-and-slab distributions with medium or heavy tails). The parameters governing these distribution are estimated from the data via an Empirical Bayes approach (see Morris and Carrol, 2006 for details). The priors for the variance components of the model are vague proper priors, meaning that no particular a priori assumption is made on their distribution, and that the variance parameters are estimated essentially from the observed data. Wavelet coefficients estimated with these assumptions on the prior distributions produce smooth curves. Importantly, the smoothing is driven by the importance that the features of the observed time series have in distinguishing the levels of the factors. This is an key feature in modelling time series affected by observational noise, because this is smoothed out in a way that is informed by the structure of the statistical model adopted and ultimately by the underlying theoretical considerations.

The posterior distributions of the fixed effects are transformed via the inverse discrete wavelet transform back to  $f_0$  values so that they can be interpreted in the terms of the hypotheses tested by the model. For each level of the fixed factors and for each point in time we have a posterior distribution describing how probable is each possible effect size given the data and the prior distribution. From this time-varying distribution we can extract the time-varying estimate of the effect size and the credibility interval, allowing to determine if the effect is significantly different from 0. The time-varying estimates and the relative 95% credibility intervals obtained for the predictors considered in our model are displayed in Figure 10 in section 3.2 below.

#### 2.4.5. Predictions

The quantitative analysis aims at verifying which of the competing distinctive features of RF proposed in the literature are relevant. We compared the three categories RF, Amb and R, assuming that Amb would be intermediate between the other two. If this was the case, then the predictions according to the literature would be as follows: As shown in Figure 3a below, according to Post's (1999) model, the  $f_0$  peak H remains at the same height for all three categories RF, Amb and R. The difference concerns the alignment of the rise which should be earlier for RF than for Amb, than for R. If this was the case, the beginning of the rise (henceforth L1) and H should be aligned earlier for RF than for Amb and earlier for Amb than for R. Moreover, a secondary consequence of the model is that the final point of the contour L2 should

be lower for RF than for Amb, and lower for Amb than for R. The predictions derived from Rossi's (1999) and Di Cristo and Hirst's (1996) views are quite different (see Figure 3b). In this view, the contours differ with respect to the height of the f0 peak H, which is supposed to be higher for RF and therefore lower for Amb and for R. Nothing is mentioned concerning alignment. If we assume that the beginning of the rise is the same for all the contours and that the end of the fall is the same for RF and Amb, the two following additional differences are expected: i) the alignment of H occurs earlier in RF than in Amb and it occurs earlier in Amb than in R; ii) L2 is lower in RF and Amb than in R.

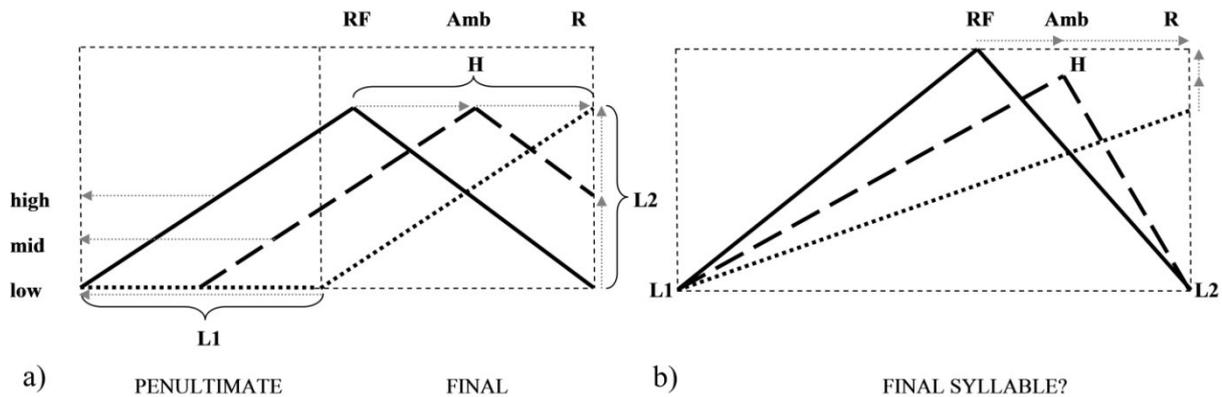


Figure 3. Hypotheses concerning the scaling and alignment of the beginning of the rise L1, the f0 peak H and the end of the f0 curve L2 for the three categories under study: RF, Amb and R. Panel a) shows the hypotheses derived from Post's (1999) model, while panel b) shows the hypotheses derived from the views by Rossi (1999) and Di Cristo and Hirst (1996).

In summary, we can test the following four hypotheses:

H1: the alignment of the f0 peak H is earlier for RF than for Amb and it is earlier for Amb than for R. This is predicted by all approaches (Post, 1999; Rossi, 1999 and Di Cristo & Hirst, 1996).

H2: the alignment of the beginning of the rise L1 is earlier for RF than for Amb and it is earlier for Amb than for R. This is predicted only by Post (1999).

H3: the final value of the f0 curve L2 is implemented lower for RF than for Amb than for R. This is also predicted only by Post (1999).

H4: the scaling of the f0 peak H is higher for RF than for Amb than for R. This is predicted by Rossi (1999) and Di Cristo and Hirst (1996) but not by Post (1999).

Note that only H2, H3 and H4 but not H1 would discriminate between Post's model on the one hand and Rossi's and Hirst and Di Cristo's view on the other hand.

Two important remarks must be added concerning the standard predictions of the AM theory for scaling and alignment. Concerning scaling, if an H% boundary tone models an upstep of the final H target, the prediction should be that the f0 curve of LH\*L% (RF) peaks lower than the

f0 curve of LH\*H% (R) since the f0 maximum corresponds to H\* in the former contour and to H% in the latter contour, an additional rising portion being added to H\* by the implementation of H%<sup>2</sup>. However, Post's (1999) model predicts that the peaks of both contours are equally high. Rossi's (1999) and Di Cristo and Hirst's (1996) predictions even contradict the upstep hypothesis. Concerning alignment, panel a) in figure 3 above shows that a number of predictions about the segmental anchoring landmarks of RF and R can be derived from our interpretation of Post's (1999) model. These predictions are summarized in table 1 below:

	RF (LH*L%)	R (LH*H%)
L	Beginning of the penultimate	Beginning of the accented syllable
H*	Beginning of the accented vowel	End of the accented vowel
L% or H%	End of the accented syllable	End of the accented syllable

Table 1. Predictions concerning the segmental anchoring landmarks of RF and R derived from our interpretation of Post (1999).

Although our data are not controlled enough to test these predictions in full details, because of limitations related to a corpus of spontaneous speech which is only 45 minutes long, we will discuss them in our general discussion in section 4 below.

## 2.5. Results of the corpus study

### 2.5.1. Results of the qualitative analysis

A visual inspection of the f0 curves of the 114 identified RF contours (both consensual and ambiguous) revealed some differences in their phonetic implementation. This gives a good vantage point on the amount and directions of the phonetic variability of RF in the corpus. Moreover, it supports the exclusion of some items for the quantitative analyses.

The most frequent description of RF in French mentions that at least part of the rise and the fall both occur on the last syllable of the utterance, or of the intonational phrase (Delattre, 1966; Vaissière, 1980; Di Cristo & Hirst 1996; Post, 1999; 2000)<sup>3</sup>. This is the actual location of 91% of the 114 contours identified as RF contours in our corpus (i.e. 104 items). Figure 4 below shows a prototypical example. Note that, in this example, the first part of the rise occurs on the

---

<sup>2</sup> It is the case in Post's model (Post, 1999; 2000) when H% replaces 0% to account for the difference in scaling between the rise to high LH\*H% and the rise to mid LH\*0%.

<sup>3</sup> We do not discuss here the case in which RF is followed by a deaccented part in the utterance because we have no cases of this in the corpus. In such cases, the syllable hosting the rising and the falling movement of RF could be intermediate phrase final rather than intonational phrase final (see Jun & Fougeron, 2002).

penultimate syllable as predicted by Post’s model. We will test this property in the quantitative analysis.

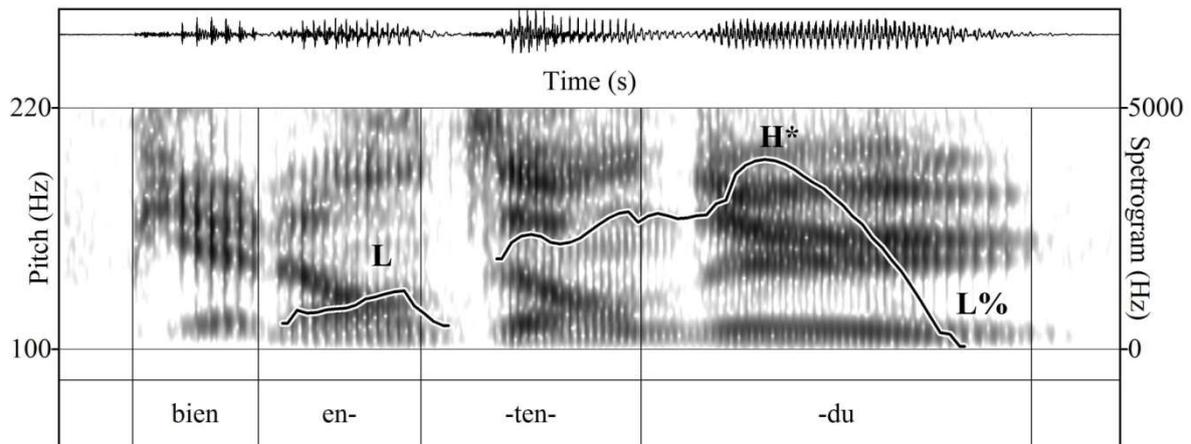


Figure 4. A prototypical example of RF on the expression “bien entendu” (*of course*). The frontiers of the syllables are marked by vertical lines, the tones L, H\* and L% are approximately located at the relevant turning points of the f0 curve.

The remaining 9% of the identified RF (i.e. 10 items) show a different pattern of the f0 curve, so we will regard them as atypical. Among these 9%, one has the curve of a simple rise without any falling part. Its phonetic realization is illustrated in figure 5 below. The identification of this item as a consensual RF by both experts despite its non falling curve, could be due to the early beginning of the rise: although the early L target is not easily detectable due to the unvoiced onset of the penultimate syllable, the fact that the penultimate is realized clearly higher than its preceding syllable is a strong cue in favor of this hypotheses. Like the other atypical items, this one will be removed from the quantitative analysis.

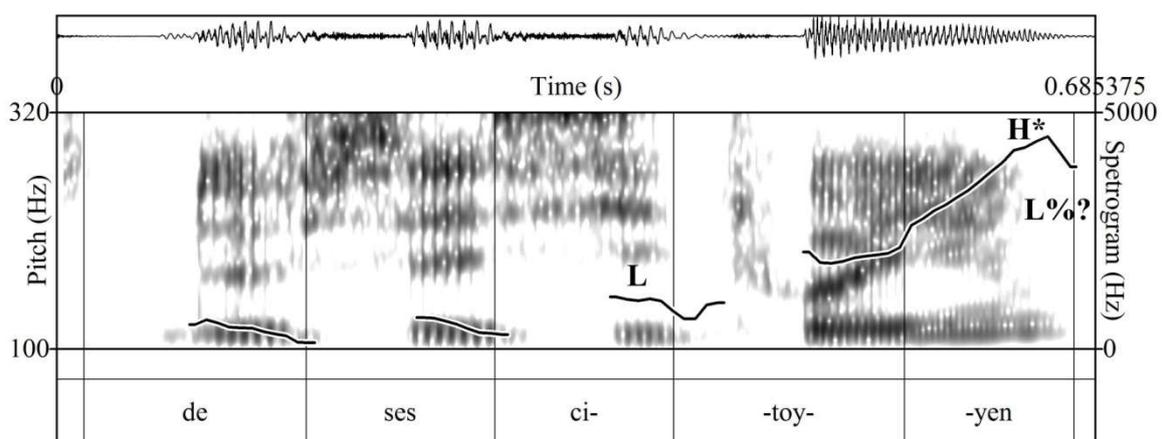


Figure 5. A consensual example of RF with almost no falling part of the f0 curve. The upstepped part of the curve on the syllable “-toy-” may explain the consensual identification.

Another atypical item, an ambiguous one this time, shows the curve of a clear fall beginning from a high point located on the antepenultimate syllable. As it appears clearly in figure 6, the f0 curve then decreases progressively down to the last syllable, so this could hardly be identified as a case of deaccenting after RF. A possible explanation of its identification as RF by one of the experts would rely on the emphatic effect of this contour which is aurally very strong and could be explained by the very high realization of the high target on the /ki/ syllable (around 324 hz). We mentioned (in section 1.1.) the emphatic/expressive function attributed to RF which could be at stake here.

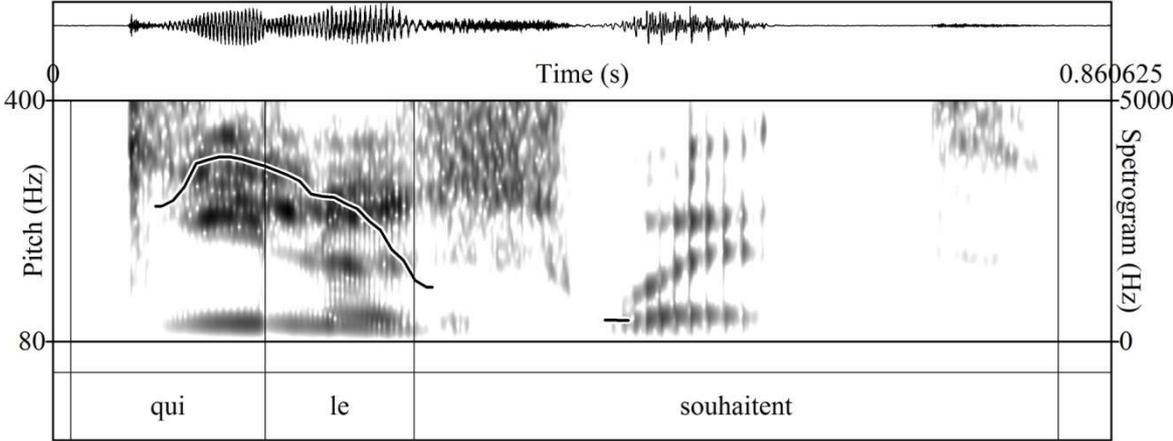


Figure 6. Ambiguous atypical item consisting of a fall from the antepenultimate syllable “qui” *who* in the relative clause “qui le souhaite” *who wish to do so*. The word “souhaitent” is realized as one syllable with the following phonological composition /swEt/. This is why the f0 peak is on the antepenultimate.

Another item, also consensual, shows a high plateau followed by a fall on the last syllable as shown in figure 7. In this case, its identification as RF probably relies on a perceived fall through the accented syllable’s rime.

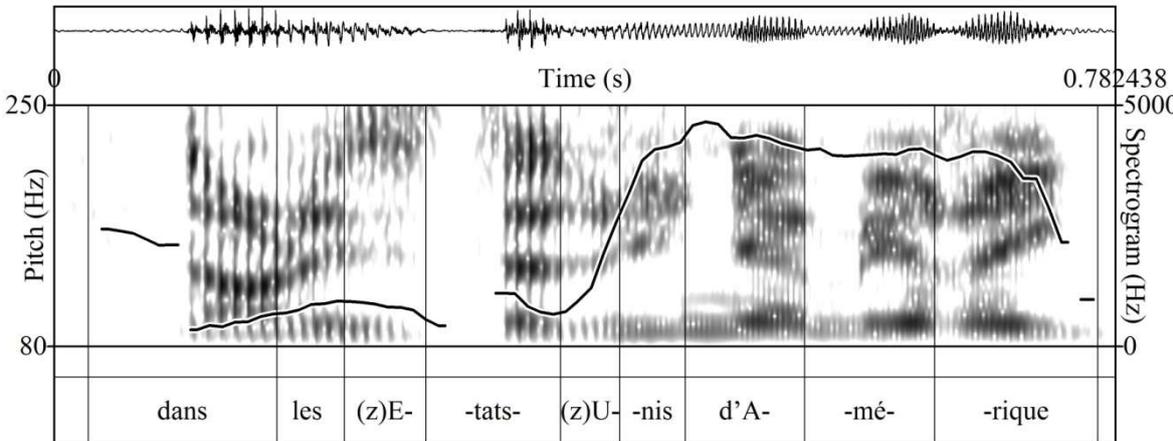


Figure 7. Consensual atypical item consisting of a high plateau beginning on syllable “-nis” of “Unis” up to the last syllable of the prepositional phrase “dans les Etats-Unis d’Amérique” *in the United States of America*. The fall on the rime of the last syllable “-rique” could explain the consensual identification of RF.

Finally, the most interesting non prototypical items identified as RF are the seven rising-falling contours whose high peaks align with the penultimate syllable, not the final one. Figure 8 shows one of these.

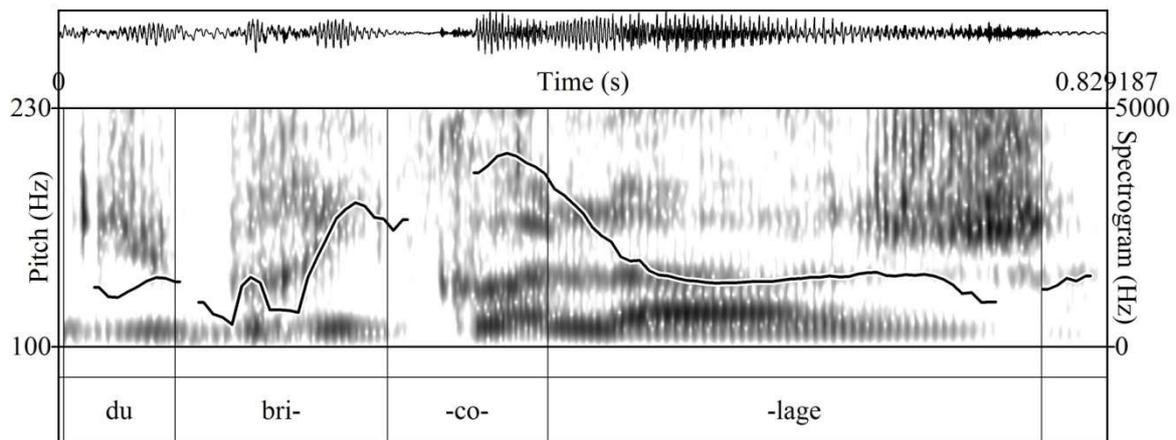


Figure 8. Example of an atypical contour identified as RF but having the f0 peak on the penultimate syllable “-co-”, not on the last syllable “-lage” in the prepositional phrase “du bricolage” *of odd jobs*.

In section 1.1., we mentioned that, in the literature on French intonation, these penultimate rise-falls are not always phonologically distinguished from the “implication contour”. Unfortunately, we did not find enough tokens of this variety to establish a clear quantitative comparison in the present corpus. Instead, we propose a brief qualitative account. Globally, we found 15 items of this penultimate peak contour in the corpus. They were not specifically identified by our identification task since it was based in Delattre’s inventory which does not include penultimate peak contours. As mentioned above, seven of them were identified as RF by expert A, 4 as F (falling contour) by expert C, and 2 as R by both experts. The remaining two items were not identified as members of the inventory. We will come back to this contour in the discussion of the corpus study (section 2.6).

In sum, in our corpus, 91% of the 114 contours identified as RF by one of the experts (Amb) or both (RF), i.e. 104 items, show a prototypical rising-falling movement on the last full syllable of the utterance. The remaining 9% (10 items) correspond to various non prototypical configurations that were removed from the set of relevant data for the quantitative analysis.

For the sake of comparison, we finally display an example of a typical R contour in figure 9:

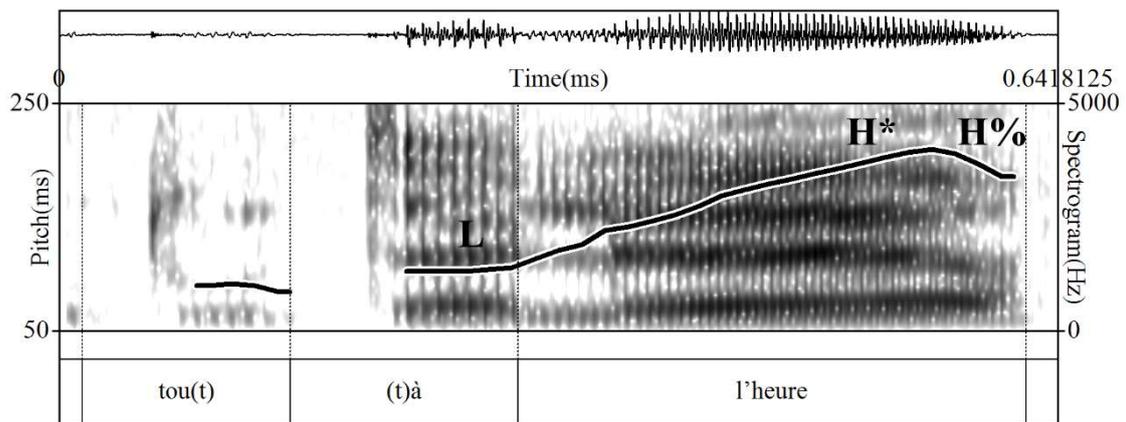


Figure 9. A prototypical example of R on the expression “tout à l’heure” (*a moment ago*). The frontiers of the syllables are marked by vertical lines, the tones L, H\* and H% are approximately located at the relevant turning points of the f0 curve.

This example illustrates that, for R, the beginning of the rise corresponds to the beginning of the accented syllable and its end corresponds to the end of that same syllable.

### 2.5.2. Results of the quantitative analysis

The estimated time-varying average values of the model effects and their HPD intervals are displayed in Figure 10 below. In each panel, we show the time-varying mean and extension of the HPD interval for the effect of a given predictor. Figure 10 Amb displays the intercept of the model which, due to the adopted coding of the factors, represents the f0 curve estimated for the ambiguous contour in the no-coda condition. The continuous red line indicates the time-varying mean of the posterior distribution of f0 values and the pink band indicates the lower and upper bounds of the corresponding HPD intervals. The curves in the other panels indicate the point-estimates of the effects. Deviations from 0 correspond to changes in f0 induced by the corresponding factor levels. An effect is considered significant at time  $t$  if the corresponding HPD interval does not include 0 (i. e. if at time  $t$  the pink band does not include 0).

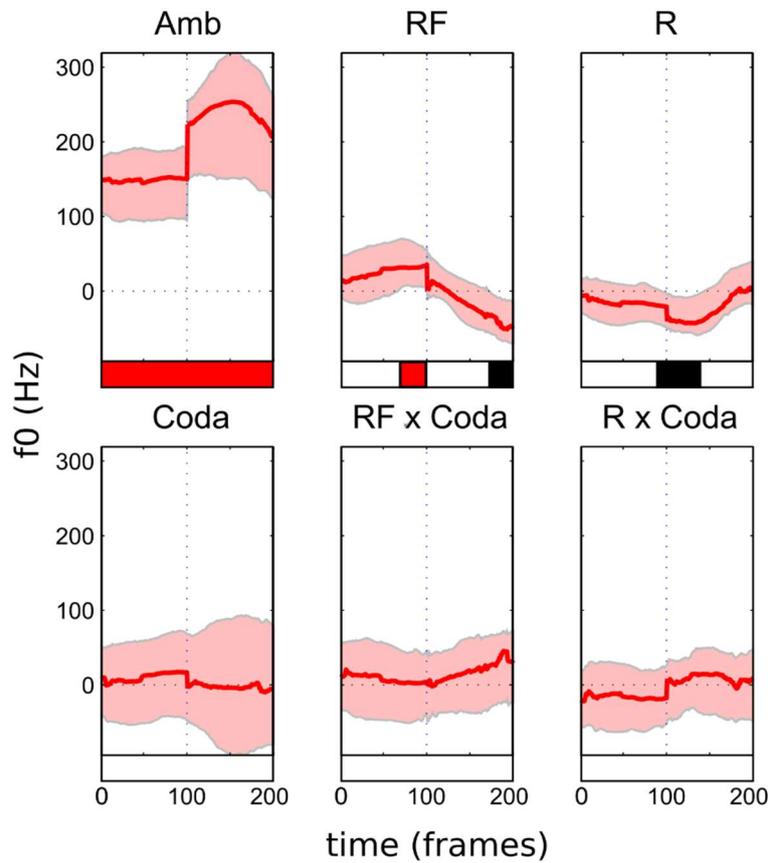


Figure 10. Evolution over time of the effects of the model predictors. The titles above each panel indicate the corresponding levels of the model predictors. The abscissa indicates relative time. The nucleus of the penultimate syllable ends at time point 100, where the nucleus of the last syllable begins. The continuous red line indicates the time-varying mean of the posterior distribution for each predictor. The pink bands indicate the time-varying lower and upper limits of the corresponding HPD intervals. The red and black rectangles below each panel indicate time intervals where the effects are significantly larger (red) or smaller (black) than zero.

By combining the model effects, we obtain the estimated average curves for the different classes of f0 contours considered (cf. Figure 11).

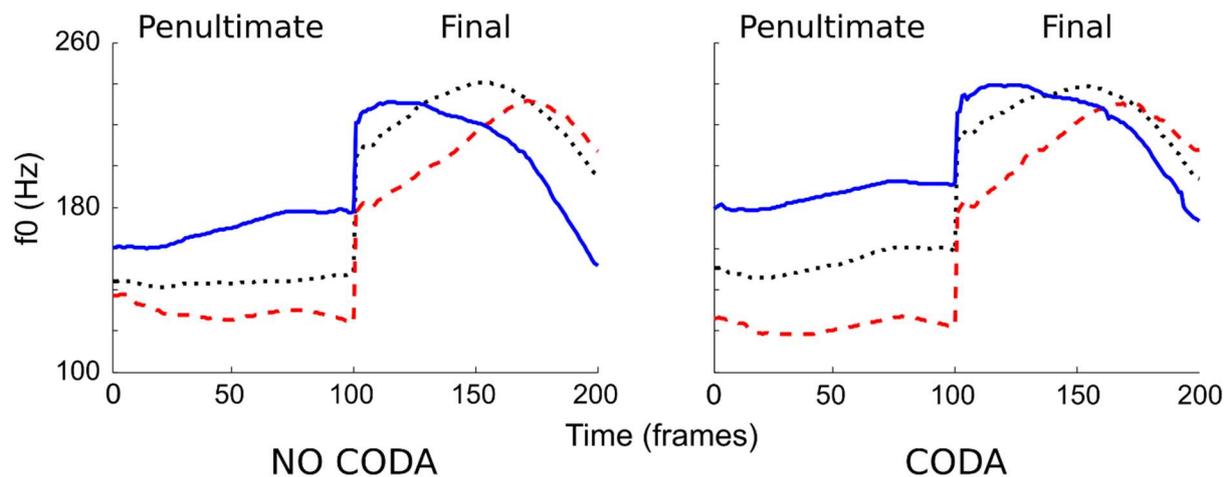


Figure 11. Estimated trajectories of the  $f_0$  curve on the nuclei of the penultimate and the final syllables of RF (solid blue line), Amb (dotted black line) and R (dashed red line), in the no-coda condition on the left panel and in the coda condition on the right panel. The dramatic change between the penultimate and the final part of the curve is due to the fact that eventual coda consonants of the penultimate syllable and onset consonants of the final syllable are not depicted.

The curve for the ambiguous contour Amb in the no-coda condition (dotted line in Figure 11-No Coda) corresponds to the model intercept evolving over time (displayed in Figure 10-Amb together with its HPD). The curve for the RF contour in the same no-coda condition (continuous line in Figure 11-No Coda) is obtained by adding to the model intercept the time-varying effect of the RF contour (Figure 10-RF). This effect is significantly positive at the end of the penultimate vowel (red rectangle), where the RF contour is higher than the ambiguous contour. This is consistent with alignment of the beginning of the rise L1 being earlier in RF than in Amb (hypothesis H2). Moreover, this effect is significantly negative at the end of the final accented vowel (black rectangle), where the RF contour is lower than the ambiguous contour. This is consistent with a lower scaling of the final part of the fall in RF than in Amb (hypothesis H3) which, then, is already visible on the syllable nucleus. The curve for the R contour in the no-coda condition (dashed line in Figure 11-No Coda) is obtained by adding the time-varying effect of the R contour (Figure 10-R) to the model intercept. The effect is significantly negative at the beginning of the final accented vowel, where the R contour is significantly lower than the Amb contour. The differences observed in the last accented vowel between the  $f_0$  curves for RF and R and the curve for Amb are both consistent with hypothesis H1 (earlier alignment of the  $f_0$  peak H in RF than in Amb and earlier alignment of that same  $f_0$  peak in Amb than in R). The  $f_0$  curve for the ambiguous contour in the coda condition (dotted line in Figure 11-Coda) is obtained by adding the effect of that condition (Figure 10-Coda) to the model intercept. No significant change is observed in the ambiguous contour between the coda and no-coda condition. The  $f_0$  curve for the RF contour in the coda condition (continuous line in Figure 11-Coda) is obtained by adding to the model intercept the effect of the coda condition (Figure 10-Coda) and the effect of the interaction between the coda condition and the RF contour (Figure

10-RF × Coda). Although the RF contour observed in the no-coda seems to be generally higher than that observed in the coda condition the interaction between the coda condition and the RF contour is not significant. The f<sub>0</sub> curve of the R contour in the coda condition (dashed line in Figure 11-Coda) is obtained by adding to the model intercept the effect of the coda condition (Figure 10-Coda) and the effect of the interaction between the R contour and the coda condition (Figure 10-Coda × R).

## 2.6. Discussion of the corpus study

The corpus study has shown that the final syllable rising-falling contour of French intonation called “implication contour” by Delattre (1966) RF is consistently identifiable in naturally occurring data, and is distinguishable from the final continuation rise R, despite the absence of categorical perception found between these two contours by Post (1999). Moreover, 91% of the identified RF display the canonical realization of the contour with a rising movement within the penultimate syllable ending in a pitch peak at the beginning of the final nucleus followed by a fall towards the rime of the final accented syllable. The quantitative analysis confirmed that the formal difference between RF and R mainly relies on the earlier alignment of the rise and a lower falling movement for RF. Interestingly, the RF items that were identified as ambiguous between RF and R emerge as a third phonetic form from the quantitative analysis.

Specifically, the fact that the f<sub>0</sub> curves corresponding to the Amb category in Figure 11 (dotted black lines) and in the final nucleus is not a flat line, indicates that the differences between this contour and the other modeled contours are not the product of the variability of the curves falling in that category. Conversely, their phonetic implementations group around an intermediate form. However, there is no reason to hypothesize that this intermediate form corresponds to a phonological category: no proposal in the literature makes such a claim. The informal observations that we can make about the ambiguous items of our corpus indicate that this intermediate form could be used as an attenuated version of RF which is known to convey contrast and disagreement with the addressee (Portes & Reyle 2014; Sichel-Bazin, 2015). A comparison between two speakers of the corpus gives some clue about what could be at stake here: speaker BE uses more Amb than RF (16 items versus 13) while speaker AL uses more RF than Amb (18 versus 10). The important fact is that AL, who is a well-known politician and used to public arguments, sounds much more self-confident and more combative than BE, an academic historian, who seems to argue with much more caution. This difference may rely on the difference between RF and Amb at least partially. Of course, this interpretation is highly speculative and should be investigated in future work.

Another interesting result of the qualitative analysis is the identification as RF of seven items that actually could be classified as LH+H\*L% contours in Post’s (2000) grammar since their f<sub>0</sub> peak is aligned with the penultimate syllable (see figure 8 in section 3.1). As a matter of fact, RF and this specific penultimate pattern are far from being systematically distinguished in the literature. For instance, among the authors cited in the present study, Post (2000), Martin (1999) and Ladd (2008) make the difference between a rise-fall aligned on the penultimate and a rise-

fall aligned on the final syllable, but Rossi (1999), Di Cristo & Hirst (1996), and Mertens (2008) do not mention it. Like Post, Martin and Ladd, we believe that the rising-falling pattern with the peak on the penultimate may constitute a distinct phonological category in French. Unfortunately, we found only 15 examples of this pattern in the whole corpus (among 1400 identified contours, including all types of major and minor contours), so we will keep its study and its comparison with RF for future work.

The quantitative analysis shows that the alignment of the  $f_0$  peak on the nucleus of the final syllable is a clear distinctive feature of the difference between the RF and R contours. This is not surprising since both Post's model and the views of Rossi and Hirst and Di Cristo predict this alignment difference, even if the latter authors don't mention it explicitly. Moreover, this is likely to be the main cue of the contrast since it happens on the unique part of the syllable that is always voiced and then stable for  $f_0$  perception. Moreover,  $f_0$  on the penultimate nucleus is higher for RF than for Amb in both no-coda and coda conditions, and there is a tendency for Amb to have a higher  $f_0$  on the penultimate nucleus than R, even if this difference is not statistically significant. Consequently, it is likely that  $f_0$  level is higher for RF than for R on the penultimate nucleus, regardless of the presence of a coda consonant. A higher  $f_0$  on the penultimate is a cue that the rise begins earlier for RF than for R. This was predicted by Post's model which is right in assuming that the alignment of the whole rise is an important cue of the contrast. Thirdly, the hypothesis that the final value of the  $f_0$  curve L2 is implemented lower for RF than for Amb and for R is also confirmed both in no-coda and coda conditions which is predicted by all authors since R is not supposed to have a falling part. However, as shown in Fig. 10, the phonetic implementations of this contour have a falling portion and this feature is already present in the nucleus, even in syllables containing a voiced coda. So the presence of a falling part of the curve cannot reliably distinguish RF from R. However, the two contours can be distinguished by the height of the  $f_0$  value at the end of the nucleus (or of the syllable if a voiced coda is present). Finally, the quantitative analysis does not demonstrate that the  $f_0$  peak H is higher for RF than for R. Hence, Rossi's (1999) and Di Cristo and Hirst's (1996) claim that the peak height is an important distinctive feature of the difference between RF and R is not supported by our results.

Let us now come back to the ambiguous items (Amb). Figure 11-NoCoda shows that, in the *no-coda* condition, Amb is similar to R in the penultimate syllable and in the final part of the curve in the nucleus, however its  $f_0$  peak (H) is aligned earlier than for R. As a consequence, the falling part of the contour is longer and could be more perceptually salient. This makes Amb a more expressive example of R that corresponds to Rossi's (1999) description of the "expressème" when it is derived from a final rise. On the other hand, in the *coda* condition (Figure 11-Coda), Amb is more similar to RF than to R on the penultimate and in the beginning of the final syllable. This profile may correspond to a non emphatic version of the RF category, that could explain the difference between speaker BE and speaker AL reported in the discussion above.

Although the results discussed so far permit to determine the differences between the phonetic implementations of the f0 contours under study, they leave two important questions unanswered. The first question is related to the association of different meanings to the RF and the R contours. The statistical analyses were meant to identify the f0 features relevant to the distinction made by the two experts under the assumption that their judgments were driven by the meaning of the utterances. However, we cannot exclude that theoretical expectations concerning the shape of the f0 curves biased the selection of the RF and R contours. Therefore, further analyses of the association between meaning and form by naïve listeners is required. The second question is related to the status of the ambiguous contour. Although its shape results distinct from that of the two other contours, it is unclear whether this f0 pattern is ambiguously perceived by each expert, or if it is just differently perceived by the two experts. In order to verify if the results of the corpus analysis were relevant regarding an untrained and unbiased population, we designed a perception experiment using the corpus material. This experiment is described in the following section.

### 3. Perception experiment

If RF and R are different phonological categories among the inventory of French intonational tunes, they should convey different pragmatic meanings. In section 1.1, we reported claims found in previous studies (Portes & Reyle, 2014; Delais-Roussarie et al., 2015; Sichel-Bazin, 2015), that RF conveys an epistemic meaning while R rather conveys continuation or a polar question. More precisely in our corpus data, RF can be said to convey a strong commitment of the speaker (henceforth conviction) and/or a statement of the obvious (henceforth obviousness), and R exclusively expresses continuation<sup>4</sup>. To test this hypothesis, we designed a forced choice perceptual identification experiment in which listeners were asked to identify the meaning conveyed by 30 stimuli extracted from our corpus on the basis of their intonation. Each stimulus consisted in a portion of speech signal corresponding to an accentual phrase (AP) containing one of the contours analyzed in the previous sections of this paper. After hearing each stimulus, the listener had to decide if it bore a melody conveying conviction/obviousness or a melody conveying continuation. Listeners were separated in two groups according to the instructions received. The example of the conviction/obviousness meaning (labelled “C/O” in the following figure) presented to listeners in Group I was an instance of the RF contour while that presented to listeners in Group II was an instance of the R contour. Likewise, the example of the continuation meaning (labelled “Cont”) presented to listeners of the Group I was an instance of the R contour while that presented to the listeners of Group II was an instance of the RF contour. With this kind of experimental design, we can formulate the following hypotheses on the association between meaning and f0 contours:

#### Hypotheses concerning RF and R

---

<sup>4</sup> The few occurrences of rising yes-no questions in the corpus have already been discarded of our data set in the corpus study.

H0: There is no form-meaning association between RF and the C/O meaning and between R and the Cont meaning. This would be the case if participants associate forms with meanings by chance.

H1: The training part of the experiment plays a major role in the form-meaning mapping. This would be the case if participants of Group I relate RF to C/O and R to Cont, while participants of Group II relate RF to Cont and R to C/O.

H2: RF is associated with the C/O meaning and R is associated with the Cont meaning. This would be the case if participants of each group relate RF to C/O and R to Cont.

### Hypotheses concerning Amb

H0: Amb is not perceived as ambiguous and is associated with one of the two meanings. This would be the case if participants preferentially relate Amb contours to the C/O meaning or, alternatively, to the Cont meaning.

H1: Amb is perceived more ambiguously than R or RF by each speaker, considered separately. This would be the case if participants relate Amb to one or the other meaning less consistently than RF or R.

H2: Amb is consistently associated with one of the two meanings by each participant considered individually, but consistency is lost at the inter-individual level.

## 3.1. Material and method

### 3.1.1. Stimuli

Among the items identified in the corpus study, we selected those whose  $f_0$  curves were most similar to the curves obtained through statistical modelling and displayed in Figure 11. To this aim we implemented a two-steps procedure similar to that described in section 2.2.1. In the initialization step, we computed a distance score by comparing each observed  $f_0$  curve to the modelled shape of  $f_0$  (see section 2.5.2) obtained for the same combination of contour type and coda condition characterizing the observed curve. The distance score was obtained by deforming the time scale of the observed curve as described in section 2.4.2 and by computing the mean point-wise distance between the modelled curve and the time-scaled version of the observed  $f_0$  curve. In the iterative step of the procedure, at each iteration, we selected one item per contour and coda condition by choosing the one with the smallest distance score. By reiterating the procedure five times, we identified 30 realizations of the three contours. We then extracted a portion of the speech signal corresponding to the accentual phrase containing each selected item. To build our final stimuli we scaled the intensity of the extracted portions of signal so that each stimulus had the same RMS value.

### 3.1.2. Participants

Thirty four native speakers of French (26 females and 8 males) participated in the experiment. They were between 18 and 40 years old (mean: 21,91; SD: 5,83). At the time of recording they were undergraduate students at Aix-Marseille University. None of them had any known speech

or hearing problems. None of them had received any training in prosody. Seventeen listeners were assigned to Group I and the other seventeen to Group II.

### 3.1.3. Procedure

Participants were told that they were going to listen to parts of sentences produced with two different melodies conveying two different meanings. In the instruction phase of the experiment, each participant could listen to four repetitions of one RF item and of one R item. Both the items were produced by the same speaker and had the same segmental content corresponding to the word *European* (“européenne”). Each time an item was presented, the associated meaning was displayed. Each item was presented as conveying a particular meaning depending on the group of participants. In the instructions to participants belonging to Group I, the item with the RF contour was associated with the C/O meaning and the item with the R contour was associated with the Cont meaning, consistently with what proposed in the literature. In the instructions to participants belonging to Group II, the item produced with the RF contour was associated to the Cont meaning and the item produced with the R contour was associated to the C/O meaning. The order of presentation of the items in the instruction phase was random, with the constraint that each presentation of one contour was followed by a presentation of the other contour. In the test part of the experiment, listeners were asked to evaluate 10 times each of the 30 stimuli selected for the experiment. After hearing each stimulus, they had to decide whether it had a conviction/obviousness conveying melody (C/O) or a continuation conveying melody (Cont). Each category was associated with one of the right or left edge buttons of a five buttons box, which participants had to press as soon as possible after identifying the tune<sup>5</sup>. The association between meanings and buttons of the box varied randomly across participants. Stimuli were arranged in 10 blocks, each block containing one repetition per stimulus. The order of stimuli was randomized in each block.

### 3.1.4. Statistical analysis

In order to factor out rapidly vanishing effects of the instruction phase of the experiment, we did not analyze the responses given to the first block of stimuli. A logistic mixed effects model was designed to predict the probability that listeners perceive the stimuli as conveying the Cont meaning. The model’s predictors were the listeners’ group (Group I vs. Group II, reference level: Group I) and the type of contour (RF, Amb, R, reference level: Amb). The interaction between the predictors was also considered. Due to the high number of data points (N = 9180) we initially designed a model with a maximal random effects structure (Barr, Levy, Scheepers & Tily, 2013). We then applied the procedure introduced by Matuschek, Kliegl, Vasishth,

---

<sup>5</sup> Reaction times were also recorded. However, we did not expect to find significant differences between the contours because we could not control several factors affecting this kind of behavioural measure. For example, both the number of words and syllables varied across the stimuli, as well as the location of the uniqueness point within each word. This prediction was confirmed through visual inspection of the reaction times distributions.

Baayen and Bates (2017) to obtain a more parsimonious random effect structure, which could capture the covariance relations in the data (and thus is expected to maintain an error rate similar to that of the full model) while guaranteeing a higher statistical power. The final model included a random intercept for the identity of the listener, a random intercept per stimulus and a listener-specific random slope for the effect of the contour. In order to determine if Amb stimuli were perceived more ambiguously than RF or R stimuli by each participant considered separately, we submitted the responses given to each kind of stimuli by each participant to a separate logistic mixed effects model with a random intercept per stimulus. No predictor was included in the models as we were interested in their intercept being different from 0. The sign of the intercept indicates if the stimuli were perceived more often as C/O or as Cont. The corresponding Z statistic (obtained by dividing the intercept value by its standard error) deviates from zero as the probability that the intercept is significantly different from zero increases. This means that when the magnitude of Z ( $|Z|$ ) is big enough, the value of the intercept is considered to be significantly different from 0 and the participant associated forms with meaning in a consistent fashion. However, when  $|Z|$  is low, the participant perceived the stimuli ambiguously (and associated forms with meanings by chance). In other words,  $|Z|$  can be used as an index of how consistently each speaker judged each type of contour. By running the models separately on the responses given by each participant to stimuli characterized by the same type of contour, we obtained 108 values of Z ( $34 \times 3$ ). The magnitudes of the Z statistics were submitted to a linear mixed effects model in which the predictors were the type of contour (reference level: Amb) and the participant's group (reference level: Group I). The participant's identity was considered as a random intercept<sup>6</sup>.

### 3.2. Results

The bar plots in Figure 12 represent the number of C/O and Cont responses obtained for each type of contour by each group of listeners. Regardless of the instructions received by the listeners, the stimuli with the RF contour elicited more often a C/O response than a Cont response, while the stimuli with the R contour elicited more often a Cont response than a C/O response. The latter was also true for stimuli with an Amb contour, although for these stimuli the difference between the two responses is smaller than for stimuli with an R contour. The magnitude of the differences between the responses to the same contour is smaller in Group II with respect to Group I.

---

<sup>6</sup> We initially fitted a model including also the interaction between stimulus' type and participant's group. However, we moved to a simpler model because the interaction did not result significant and it did not contribute significantly to the model's fit (as determined by a Chi-square test comparing the errors of the models with and without the interaction).

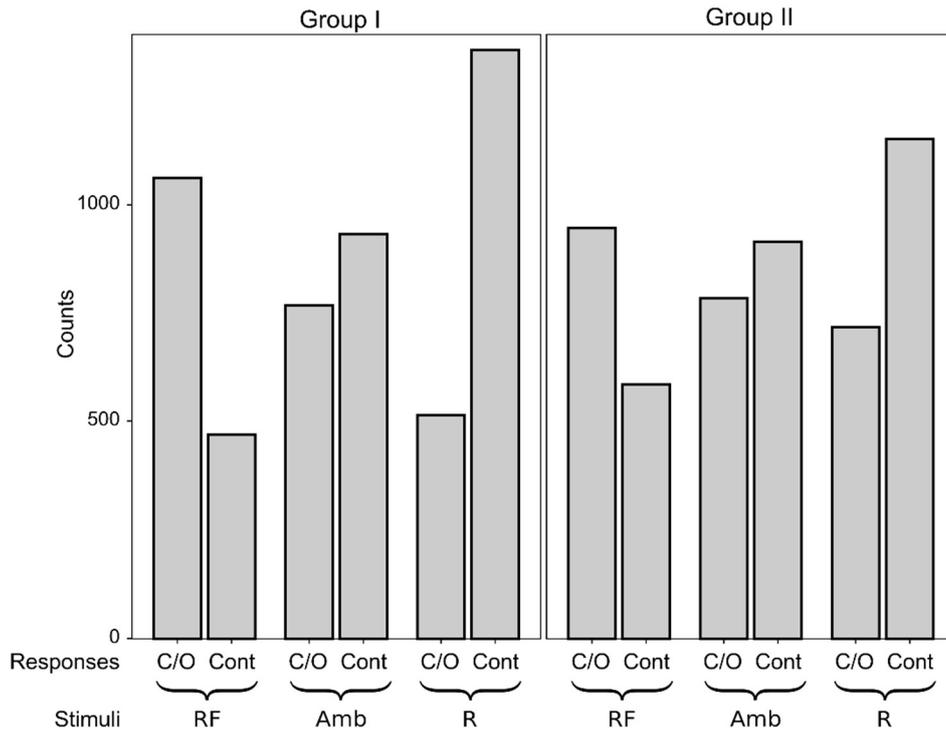


Figure 12. Counts of responses C/O and Cont over contour type. Left panel: responses obtained from listeners of Group I. Right panel: responses obtained from listeners of Group II.

Statistical analyses confirmed that the effect of the contour was significant both when comparing the responses to Amb stimuli with the responses to RF stimuli (odds ratio: -1.06, S.E. = 0.14,  $|Z| = -6.46$ ,  $p < 0.01$ ) and when comparing the responses to R stimuli with the responses to Amb stimuli (odds ratio = 0.589, S.E. = 0.1,  $|Z| = 8.85$ ,  $p < 0.01$ ). This means that listeners of Group I answered more often C/O in response to RF stimuli than in response to Amb stimuli and more often Cont in response to R stimuli than in response to Amb stimuli. The effect of the group was not significant (odds ratio = -0.05, S.E. = 0.2,  $|Z| = -0.26$ ,  $p = 0.79$ ), meaning that probability of observing a Cont response after the presentation of a Amb stimulus did not decrease significantly from Group I to Group II. The interaction between the listeners' group and the contour was not significant when the difference between the Amb contour and the RF contour was considered (odds ratio = 0.34, S.E. = 0.23,  $|Z| = 1.47$ ,  $p = 0.14$ ) but it was significant when the difference between the R contour and the Amb contour was considered (odds ratio = -0.56, S.E. = 0.14,  $|Z| = -3.92$ ,  $p < 0.01$ ). This means that the difference between the responses to Amb and those to R stimuli is lower for listeners of Group II compared to listeners of Group I.

The boxplots in Figure 13 summarize the distributions of the Z statistic values obtained by modeling separately responses given by different listeners to different stimuli types. The lowest values are observed for the responses to Amb stimuli, indicating that these stimuli are perceived more ambiguously by listeners considered separately. Differences seem to be smaller for listeners belonging to Group II.

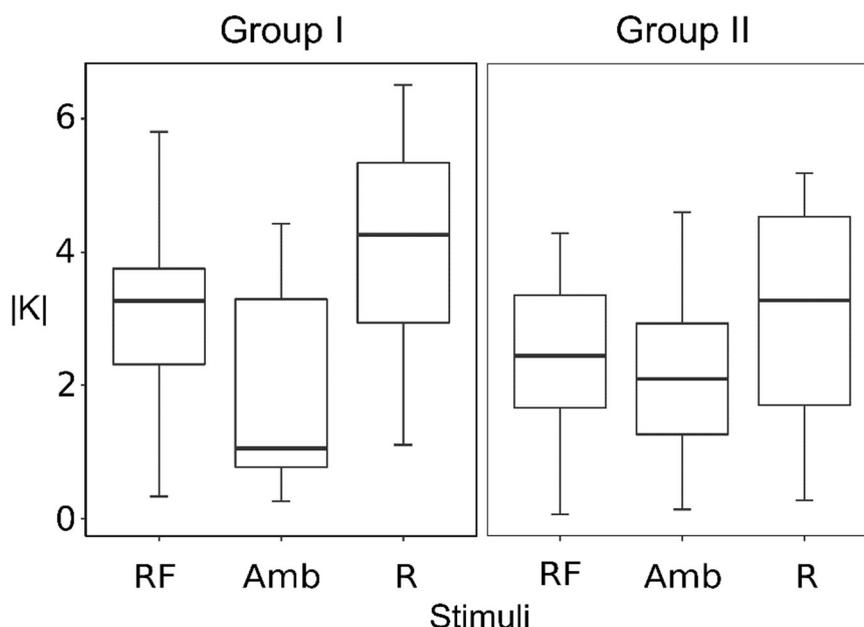


Figure 13. Box plots of the consistency index  $|K|$  over contour type. Left and right panels group responses according to the listeners' group.

Statistical analyses confirm that  $|K|$  values obtained when modeling the responses to RF stimuli or to R stimuli are higher than those observed when modeling the responses to Amb stimuli (RF vs. Amb: estimate = 0.72, SE = 0.34,  $t=2.11$ ,  $p < 0.05$ ; R vs. Amb: estimate = 1.58, SE = 0.34,  $t= 4.65$ ,  $p < 0.01$ ). However, the effect of the speaker's group was not significant.

### 3.3. Discussion of the perception study

By showing that the *C/O* meaning was perceived more often after hearing stimuli with the RF contour than after hearing stimuli with the Amb or the R contours, our results provide the first experimental evidence for the claim that RF is reliably interpreted as conveying conviction and/or obviousness (Martin, 1999; Mertens, 2008; Portes & Reyle, 2014; Delais-Roussarie et al., 2015; Sichel-Bazin, 2015). Importantly, even listeners instructed to associate the R contour with the *C/O* meaning and the RF contour with the *Cont* meaning preferred the expected association between RF and *C/O* on the one hand and between R and *Cont* on the other hand. This result shows that the association between the meanings and the two contours is anchored strongly enough in the participants' behavior to counteract the effect of a reversed priming.

The proportion of *Cont* responses to the Amb contour was significantly higher than the proportion of *Cont* responses to the RF contour and significantly smaller than the proportion of *Cont* responses to the R contour. This means that the difference between the probabilities to observe the two responses is smaller for the Amb contour than for the two other contours. The interaction between the listeners' group and the contour for the difference between R and Amb shows that the choice of the *Cont* meaning for R stimuli was indeed affected by the instruction

phase, but not enough to change the preferred form/meaning association. Moreover, the analyses conducted separately per speaker and stimulus type support the hypothesis that the Amb contour is perceived as truly ambiguous by each speaker contra the hypothesis that, although the perception of this contour is inconsistent across listeners, each listener consistently associates the contour to one of the two meanings.

#### 4. General discussion

Post's (1999) categorical perception study on French tunes cast doubt on the phonologically contrastive nature of RF versus R in this language. The results of our corpus study and of our perception study provide new evidence in favor of a phonological contrast between these two contours. The former shows that the alignment of the rising part of the contour and the scaling of the end of the  $f_0$  curve consistently distinguish  $f_0$  patterns that have been consensually classified as RF or R. The latter shows that naïve listeners are able to identify RF as a contour conveying conviction/obviousness and R as a contour conveying continuation, even when they are primed to perform the reverse contour-meaning association. In this section, we would like to discuss some implications of these new results on the phonology of French intonation in confronting them to some predictions derived from the auto-segmental view of intonational phonology (see section 2.4.4.), and to French ToBI as proposed in Delais-Roussarie et al. (2015). We will address the three following topics: the scaling of H% in R contours, the segmental anchoring of RF versus R, and the tonal coding of both tunes. Finally, we will briefly discuss the status of the ambiguous pattern Amb.

As we mentioned in section 2.4.4., H% is often assumed to require an upstep rule after another H tone to model that the level of the H% tune is higher than that of the immediately preceding H accent instead of involving a sustained high level between the two H (Ladd 2008, p. 103). In Post (1999; 2000) R is modelled as an LH\*H% and so H% should logically involve an upstep in her model. However, this is not the case since the continuum she used for her categorical perception study of the RF versus R contrast (Post, 1999) does not assume a higher scaling of H% in LH\*H% (R) than for H\* in LH\*L% (RF; see figure 2 above). A possible explanation for this would be that H\* in RF corresponds to what Roettger and Grice (2015) called a “high pitch”, i.e. a highly scaled H tone involved in a nuclear rise fall and expressing interrogativity or contrastivity in a variety of languages. This would be congruent with the fact that RF has sometimes been claimed to convey contrastivity (Di Cristo & Hirst, 1996; Portes & Reyle, 2014).

An alternative proposal by Jun and Fougeron (2002) is that in IP final position, the final H\* “is pre-empted by the boundary tone of the IP” (Jun & Fougeron, 2002, p. 153). This would work for R which would be labelled LH% instead of LH\*H%, but not for RF which would lose its H tone becoming LL%. The LH% would eliminate the difficulty related to an upstepped H%. Moreover, if we assume an LH\*L% tonal composition for RF and an LH% composition for R, this would also account for the fact that the rise is earlier in RF than in R. In this case, we could explain this earlier alignment by the tonal pressure exerted by the presence of the boundary tone

L% on the pitch accent LH\* as it is assumed by the segmental anchoring hypothesis (Arvaniti et al., 1998, Ladd et al., 1999; Xu, 1998; Atterer & Ladd, 2004; Schepman et al., 2006; Welby, 2006).

Unfortunately, our corpus study did not allow us to systematically measure the scaling of H% for R because we could measure only what happened on the voiced nucleus. However, what we found is that, when there is a coda (our coda-condition), the f<sub>0</sub> reaches the same maximum level for RF as for R. One can speculate that, in case of a voiced coda, the frequency will have room to continue rising after the nucleus. This would mean that H% would be upstepped and would give support to an LH\*H% coding of R. Further experimental production studies on more controlled and fully voiced material are therefore needed to verify whether or not H\* in RF and H% in R are scaled at the same level, and which of these explanations is relevant for French.

Another phonological device that has been proposed by the segmental anchoring literature is the secondary association of tones not only with the accented syllable or the constituent boundary, but also with other phonological landmarks (Pierrehumbert & Beckman, 1988; Grice et al., 2000). In table 1 in section 2.4.4., we spelled out the predictions that can be derived from Post (1999) about possible segmental landmarks for the secondary association of the tones composing RF and R contours. The results of our corpus study did not enable us to examine the precise location neither of the beginning of the rise L nor of the boundary tone T% (L% for RF and H% for R) since the f<sub>0</sub> curves used for the phonetic analysis only involved the nucleus of the target syllables that are always voiced. However, what happens on the nuclei allowed us to argue: 1) that the beginning of the rise L should align with the beginning of the penultimate syllable in RF and with the beginning of the final syllable in R, and 2) that the f<sub>0</sub> peak H\* is aligned with the beginning of the accented vowel in RF and with the end of the accented vowel in R. If we assume that the rising part LH\* of both contours primarily associates with the final accented syllable of the accentual phrase, our results can be interpreted as providing some preliminary evidence in favor of a potential secondary association of the initial L with the beginning of the penultimate syllable for LH\*L% and with the beginning of the final accented syllable for LH\*H%. As for H\*, they plead in favor of a secondary association with the beginning of the final nucleus for LH\*L% and with the end of that same nucleus for LH\*H%. On the other hand, our data give us no useful information on the anchoring of the boundary tones. Overall, the predictions derived from Post's (1999) model and displayed in Table 1 are confirmed by our results. However, here again, further research on carefully controlled voiced material is required to make clear proposals about secondary association of the tones composing RF and R (for careful experimental investigations of the French LH\* rise but in *prenuclear* contours, see Welby & Loevenbruck, 2005 and Welby, 2006).

Concerning the tonal composition of RF and R, French ToBI as developed by Delais-Rousarie et al. (2015) must also be discussed here. Our results give a picture of the phonetic realization of RF which is closer to the pattern labelled H+H\*L% in Delais-Rousarie et al. (2015) than to their LH\*L% pattern. Their proposal is interesting because the leading tone H of the bitonal pitch accent H+H\* accounts nicely for the higher f<sub>0</sub> on the penultimate syllable for RF versus

for R that we found in our corpus data. Moreover, H+H\*L% is said to convey “statements of the obvious” which is one of the meaning features that the participants of our perception experiment associated with RF. H+H\*L% therefore appears as an appealing alternative coding for RF.

Finally, concerning the items that were classified as Amb by the two experts, the results of the corpus study showed that their phonetic realization was intermediate between the phonetic realizations of RF and R, and the results of the perception experiment showed that naïve listeners hesitated in interpreting them as the expression of a convinced obvious statement or as the expression that the speaker has more to say (continuation). Two competing explanations can be given to that perceptual ambiguity: either Amb is a third phonological category distinct from RF and R and conveys a third meaning; or, alternatively, Amb is a manifestation of the variability of the realization of RF and R in naturally occurring data. The first possible explanation would thus be that, confronted to a forced choice between C/O and Cont, listeners of the Amb contour could not decide between these two meanings because Amb is actually related to a third meaning which is not proposed in our experiment. Following this idea, Amb could for instance correspond to Rossi’s “*expressème*” (Rossi, 1999) and convey a specific emotional meaning like enthusiasm. Conversely, following the idea that Amb would manifest the inherent variability of phonological categories in conversational use, the behavior of the participants to the experiment could be interpreted as a true hesitation between the C/O and the Cont meaning. However, even if we assume this second explanation, nothing prevents speakers to play on the phonetic variability between two clearly distinct forms in order to stay underspecified between two clearly distinct meanings. Hence, using Amb instead of using RF to convey an attenuated conviction (or a partial disagreement) would be comparable to using half a smile instead of an open smile to convey a partial acceptance of the funny quality of a joke. At this point, the difference between the two competing explanations just mentioned could be conceived as a difference between two degrees of grammaticalization of an emergent category Amb through its pragmatic specialization. Clearly, the perception experiment presented here does not allow answering these questions, which might be addressed in future research.

Another limitation of the present study concerns the role of factors depending on the segmental composition of the syllables on which the contours’ tones are anchored (such as the category of onset and coda consonant and the category of vowels of the two last syllables). Due to the limitations of our corpus, a systematic investigation of their potential effects on the alignment and on the scaling of the f<sub>0</sub> targets could not be conducted. This again constitute a question to be addressed in future research.

## 5. Conclusion

In this study, we investigated the contrast of two contours belonging to the French intonational inventory: a rise-fall RF and a rise R, by means of a production study on a corpus of naturally occurring data and of a perception experiment based on the corpus material. We compared the phonetic implementation of these contours by modeling the f<sub>0</sub> curves through a Wavelet-based Functional Mixed Model. Despite the lack of control over the data, and especially on the

segmental composition of the accented carrier syllable, we could show that RF and R mainly differ with respect to the timing of the rise and the amplitude of the falling part. These results support the hypothesis, drawn from Post's (1999) proposal, that these characteristics are more important than the scaling of the pitch peak for the implementation of the contrast between RF and R, contradicting the reverse claim defended by Di Cristo and Hirst (1996) and by Rossi (1999). Moreover, we have shown that some of the observed contours displaying an intermediate implementation and sharing features with both RF and R contours, are perceived ambiguously by French naïve listeners. This latter result led us to speculate that speakers could make use of ambiguous realizations between two close intonational categories resulting in specific pragmatic effects, opening the way for future research. We could also show how the application of methods from functional data analysis allows us to extract the regularity underlying highly variable intonation contours unfolding over time in naturally occurring corpus data. By combining the flexibility of functional data representation with the power of Bayesian estimation, Wavelet-based Functional Mixed Models open the way to a new, more data driven approach to the analysis of intonation patterns in controlled laboratory conditions as well as in naturally occurring corpus data. The results of our perception categorization task using the corpus material provided the first experimental validation that naïve listeners associate the RF contour with the expression of conviction and obviousness and the R contour with the indication that the speaker has more to say.

## Acknowledgements

This work was funded by a Mercator Fellowship attributed to CP by the German Science Foundation (DFG), Sonderforschungsbereich 732 Incremental Specification in Context, Project A6, at the University of Stuttgart. LL's work, carried out within the Labex BLRI (ANR-11-LABX-0036), ASLAN (ANR-10-LABX-0081) and EFL (ANR-10-LABX-0083), has benefited from support from the French government, managed by the French National Agency for Research (ANR), under the program "Investissements d'Avenir". We would like to thank two anonymous reviewers and the editor of this journal for their suggestions, which substantially helped us improve the quality of the paper.

## References

- Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics*, 26(1), 3–25.
- Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: Evidence from German. *Journal of Phonetics*, 32(2), 177–197.

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3), 255-278
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3(01), 255–309.
- Bigi, B., Portes, C., Steuckardt, A., & Tellier, M. (2013). A multimodal study of answers to disruptions. *Journal on Multimodal User Interfaces*, 7(1-2), 55-66.
- Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In S. Featherston, & W. Sternefeld (Eds.), *Roots: Linguistics in Search of Its Evidential Base* (pp. 77-96). Studies in Generative Grammar, Berlin: Mouton de Gruyter.
- Cole, J., & Hasegawa-Johnson, M. (2012). Corpus phonology with speech resources. In A. Cohn, C. Fougerson, & M. Huffman (Eds.), *Handbook of Laboratory Phonology* (pp. 431-440). Oxford, UK: Oxford University Press.
- Coleman, J., Liberman, M., Kochanski, G., Burnard, L., & Yuan, J. (2011). Mining a Year of Speech. Paper presented at VLPS 2011: *New Tools and Methods for Very-Large-Scale Phonetics Research*, University of Pennsylvania, January 29-31.
- Delattre, P. (1966). Les dix intonations de base du français. *French Review* 40(1), 1-14.
- Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., ... & Sichel-Bazin, R. (2015). Intonational Phonology of French: Developing a ToBI system for French. In S. Frota & P. Prieto (Eds.), *Intonation in Romance* (pp. 63-100). Oxford, UK: Oxford University Press.
- Di Cristo, A., & Hirst, D. (1996). Vers une typologie des unités intonatives du français. In *Actes des XXIèmes Journées d'Etude sur la parole*, Avignon, 219–222.
- Di Cristo, A., & Di Cristo, Ph. (2001). SYNTAIX: une approche métrique-autosegmentale de la prosodie. *TAL* 42(1), 69–112.
- D'Imperio, M., & Michelas, A. (2014). Pitch scaling and the internal structuring of the Intonation Phrase in French. *Phonology*, 31(01), 95–122.
- Grice, M., Ladd, D. R., & Arvaniti, A. (2000). On the place of phrase accents in intonational phonology. *Phonology*, 17(02), 143–185.
- Hirst, D., & Di Cristo, A. (eds) (1998). *Intonation Systems: A Survey of Twenty Languages*. Cambridge University Press.
- Jun, S. A. (ed.) (2005). *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford University Press.

- Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, Modeling and Technology* (pp. 209–242). Dordrecht: Kluwer Academic Publishers.
- Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French. *Probus*, 14(1), 147–172.
- Jun, S. A., & Fletcher, J. (2014). Methodology of Studying Intonation: From Data Collection to Data Analysis, in S-A. Jun (Ed.), *Prosodic Typology II: New Developments in the Phonology of Intonation and Phrasing* (pp. 493–519). Oxford: Oxford University Press.
- Ladd, R. D. (1983). Phonological features of intonational peaks. *Language* 59(4), 721–759.
- Ladd, R. D. (2008). *Intonational Phonology*. Second edition. Cambridge: Cambridge University Press.
- Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant “segmental anchoring” of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America*, 106 (3 Pt 1), 1543–1554.
- Malfrère, F., Deroo, O., Dutoit, T., & Ris, C. (2003). Phonetic alignment: speechsynthesis-based versus Viterbi-based. *Speech Communication*, 40(4), 503-517.
- Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7), 674-693.
- Martin, P. (1999). Intonation in spontaneous speech in French. In *Proceedings of the 14th ICPHS San Francisco*, vol. 1, 17-20.
- Mertens, P. (2008). Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de Linguistique* 56(1), 87-124.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H. & Bates, D. (2017) Balancing Type I error and power in linear mixed models, *Journal of Memory and Language*, 94, 305-315.
- Michelas, A. & D'Imperio, M. (2012). When syntax meets prosody: Tonal and duration variability in French Accentual Phrases. *Journal of Phonetics*, 40(6), 816-829.
- Morris, J. S., & Carroll, R. J. (2006). Wavelet-based functional mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68(2), 179-199.
- Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. Doctoral thesis, Massachusetts Institute of Technology. [Distributed 1987 by Indiana University Linguistics Club].

- Pierrehumbert, J. B., & Beckman, M. E. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
- Portes, C. (2004). *Prosodie et économie du discours : Spécificité phonétique, écologie discursive et portée pragmatique de l'intonation d'implication*. PhD thesis, Aix-Marseille Université.
- Portes, C., & Reyle, U. (2014). The meaning of French “implication” contour in conversation. *Proceedings of 7<sup>th</sup> Speech Prosody 2014*, Dublin, 413-417.
- Post, B. (1999). Solving a controversy in the analysis of French rising pitch movements. *Proceedings of ICPhS1999*, San Francisco, 965-968.
- Post, B. (2000). *Tonal and phrasal structures in French intonation* (published PhD dissertation). The Hague: Holland Academic Graphics.
- Prieto, P., van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23(4), 429–451.
- Prieto, P., & Torreira, F. (2007). The segmental anchoring hypothesis revisited: Syllable structure and speech rate effects on peak timing in Spanish. *Journal of Phonetics*, 35(4), 473–500.
- Roettger, T. B., & Grice, M. (2015). The role of high pitch in Tashlhiyt Tamazight (Berber): Evidence from production and perception. *Journal of Phonetics*, 51, 36-49.
- Rossi, M. (1981). Vers une théorie de l'intonation. In M. Rossi, A. Di Cristo, D. Hirst, P. Martin & Y. Nishinuma (Eds.), *L'intonation: de l'acoustique à la sémantique*, (pp. 179–183). Paris: Klincksieck.
- Rossi, M. (1999). *L'intonation, le système du français : description et modélisation*. Paris: Ophrys.
- Schafer, A.J., Speer, S.R., & Warren, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. Tanenhaus, & J. Trueswell (Eds.), *Approaches to Studying World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Tradition* (pp. 209-225). Cambridge: MIT Press.
- Schepman, A., Lickley, R., & Ladd, D. R. (2006). Effects of vowel length and “right context” on the alignment of Dutch nuclear accents. *Journal of Phonetics*, 34(1), 1-28.
- Sichel-Bazin, Rafèu (2015). *Prosodic systems in contact: Occitan and French*. Unpublished PhD thesis. Universität Osnabrück – Universitat Pompeu Fabra.

- Silverman, K. E., & Pierrehumbert, J. B. (1990). The timing of prenuclear high accents in English. In J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 72–106). Cambridge: Cambridge University Press.
- Truckenbrodt, H. (2002). Upstep and embedded register levels. *Phonology*, 19(01), 77–120.
- Truckenbrodt, H. (2007). Upstep of edge tones and of nuclear accents. In C. Gussenhoven & ; T. Riad (Eds.) *Tones and tunes*, 2, (pp. 349–386). Berlin: Walter de Gruyter.
- Vaissière, J. (1980). La structuration acoustique de la phrase française. *Annali della Scuola Normale Superiore di Pisa. Classe di Lettere e Filosofia*, 10(2), 529-560.
- Van den Berg, R., Gussenhoven, C., & Rietveld, T. (1992). Downstep in Dutch: Implications for a model. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, G. J. Docherty, and D. R. Ladd (eds.), 335–367. Cambridge: Cambridge University Press.
- Venables, W. N., & Ripley, B. D. (1999). *Modern applied statistics with S-PLUS*. Springer: New York.
- Xu, Y. (1998). Consistency of tone-syllable alignment across different syllable and speaking rates. *Phonetica*, 55, 179–203.
- Welby, P., & Løevenbruck, H. (2005). Segmental “anchorage” and the French late rise. *Proceedings of Interspeech 2005*, 2369-2372.
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343-371.