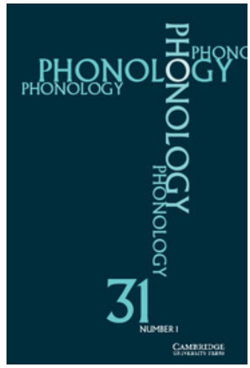

Pitch scaling and the internal structuring of the Intonation Phrase in French

Mariapaola D'Imperio and Amandine Michelas

Phonology / Volume 31 / Issue 01 / May 2014, pp 95 - 122
DOI: 10.1017/S0952675714000049, Published online: 20 May 2014

Phonology **31** (2014) 95–122.
doi:10.1017/S0952675714000049

*Pitch scaling and the internal structuring of the Intonation Phrase in French**

Mariapaola D'Imperio
Aix Marseille University, CNRS, 13604 Aix-en-Provence and Institut Universitaire de France (IUF)

Amandine Michelas
Aix Marseille University and CNRS, 13604 Aix-en-Provence

Within the autosegmental-metrical approach to French intonation, the existence of an intermediate phrase or ip is controversial. Our study provides strong evidence for its existence, by uncovering systematic pitch-scaling effects within this constituent. We first show that the presence of an ip break is responsible for blocking recursive downstep of subsequent AP-final LH* rises in declarative utterances, causing the return of the final H target to the pitch level set by the first accentual peak of the phrase (i.e. complete reset). Additional evidence for the internal structuring of the Intonation Phrase is also provided by partial reset on the postboundary H target as well as by preboundary lengthening on the last syllable of the ip. The pitch-scaling effects are claimed to result from control over the reference pitch level for the entire ip, which can alternatively be modelled through secondary association of the last pitch accent of the domain.

1 Introduction

1.1 Downstep domains and reset phenomena

Intonational modulations in speech can be characterised as sequences of tonal events, modelled as H (high) and L (low) fundamental frequency (F0) targets. This view is espoused by the autosegmental-metrical

* E-mail: MARIAPAOLA.DIMPERIO@LPL-AIX.FR, MICHELAS@LPL-AIX.FR
This work was made possible by a doctoral fellowship from the French Ministry of Education and Research to Amandine Michelas and by support from the Institut Universitaire de France to Mariapaola D'Imperio. This work was also carried out within the Labex Brain and Language Research Institute (ANR-11-LABX-0036) and has benefited from support from the French National Agency of Research (ANR), under the project title 'Investments of the Future' A*MIDEX (ANR-11-IDEX-0001-02). We would also like to thank three anonymous reviewers and the editors of this journal for providing insightful suggestions and comments, which helped to greatly improve the quality of the paper.

model of intonation (Pierrehumbert 1980; see Ladd 2008 for a review). Within this framework, the F0 level (or SCALING) of H and L tonal targets varies as a function of a number of variables, such as local emphasis, iterative downstep/upstep and final lowering, which can be active either in the phonological component or in the phonetic implementation process of the language under investigation. Studies on detailed scaling effects have been reported for a variety of languages (e.g. Liberman & Pierrehumbert 1984, Ladd 1988 for English; Pierrehumbert & Beckman 1988 for Tokyo Japanese; van den Berg *et al.* 1992 for Dutch; Prieto *et al.* 1996 for Mexican Spanish). However, work on the intonational phonology of French is still scarce, in particular regarding scaling (e.g. downstep, upstep and reset phenomena). In this paper we specifically address the issue of the downstep domain in French both within and across phrasal boundaries, as it relates to the presence of an intermediate phrase break.

Within the autosegmental-metrical model, downstep has been modelled either as a strictly local phenomenon, targeting immediately adjacent pitch accents, or as the result of a global lowering of the pitch range level for a specific phrase. On the first view, downstep is an iterative, exponential phenomenon (Liberman & Pierrehumbert 1984), lowering the level of subsequent, non-initial, H targets within rising pitch accents. As a result of pitch reset, initial H peaks return to utterance-initial values for a specific speaker after a phrase break. Downstep can thus be blocked by the presence of a phrasal break, which can either be an INTONATION PHRASE (IP) edge or, as in the case of English, an INTERMEDIATE PHRASE (ip) edge (Beckman & Pierrehumbert 1986). Not all stress-accent languages seem to work in the same way in regards to downstep and reset constraints. For instance, we know that in German (Truckenbrodt 2002) complete reset of a rising pitch accent can occur before, but not after, an IP break.[1] French shows similar behaviour, in that there appears to be cases in which prefinal accentual peaks are scaled so as to reach a specific reference value for the prosodic phrase. For instance, in Fig. 1, the second LH* peak (H2) in the declarative utterance *Le schéma du trois-mâts de Thomas devenait vraiment brouillon* appears to be downstepped relative to H1 (which sets the pitch level of the phrase, i.e. the reference level for the phrase, indicated by the unbroken horizontal line in Fig. 1), though this is not true for H3. Rather, H3 is produced at the same level as H1, once declination has been taken into account. In other words, H3 appears to be completely reset relative to the reference level set by H1. This is unexpected, given that in French H peaks are predicted to be iteratively downstepped up to the end of an IP (Post 2000).

Moreover, work on English, Dutch and German has shown that downstep can be 'nested', in the sense that it does not exclusively target

[1] Note that in Truckenbrodt (2002), complete reset of preboundary L*+H of German is defined as 'upstep'. We will employ the term 'complete reset', for reasons detailed in §2.1.

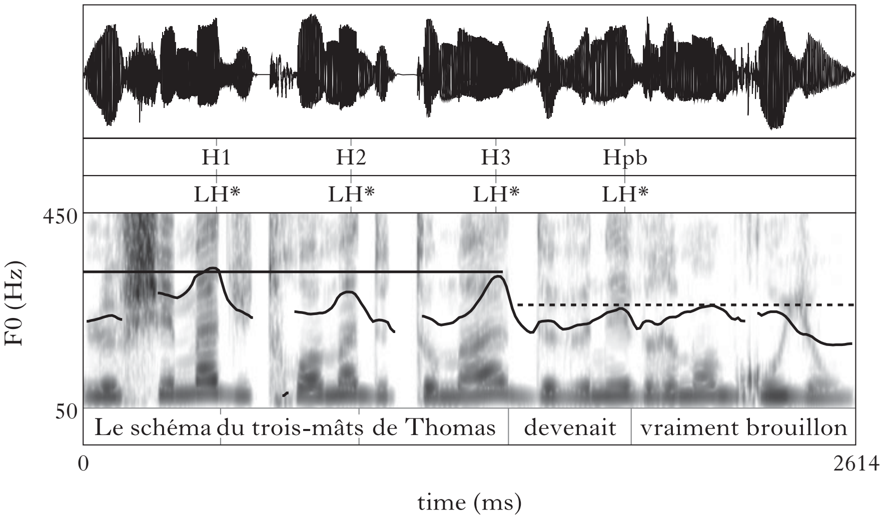

Figure 1

F0 track and spectrogram for the utterance *Le schéma du trois-mâts de Thomas devenait vraiment brouillon*. 'Thomas's sketch of a square-rigger became a real scribble.' The unbroken horizontal line represents the reference pitch level for the first phrase; the dashed horizontal line represents the reference pitch level for the second phrase.

phrase-internal tones, but also the pitch level of entire phrases. Evidence that a non-initial phrase level can be downstepped relative to the preceding phrase comes from scaling of initial postboundary peaks. For instance, note in Fig. 1 that the first H peak of the second phrase (the post-boundary H, or Hpb) appears to be scaled lower than H1, and hence does not undergo complete reset. This phenomenon has been labelled PARTIAL RESET (Ladd 1988, van den Berg *et al.* 1992, Truckenbrodt 2002). Partial reset can be modelled as the consequence of downstep among prosodic domains, targeting the pitch level for a specific prosodic unit. From the early work of O'Shaughnessy & Allen (1983) onwards, it has been noticed that in non-initial IPs the F0 level might not be reset at exactly the same level as the first peak in the utterance, yet still be higher than a previous downstepped accentual peak. This view allows register shifts to be modelled in terms of a return to a reference line (the dashed line in Fig. 1), which is itself downstepped relative to the level of the preceding phrase (the unbroken line). A reference line (or REGISTER LEVEL) can be conceptualised as the abstract reference pitch level for a prosodic phrase (Ladd 2008). This 'tonal space' would constrain the melodic implementation of actual L and H tonal targets within the same prosodic constituent. Interestingly, recent theoretical proposals concerning German (Féry & Truckenbrodt 2005) suggest that phrase-internal downstep would be greater than downstep across subsequent prosodic domains.

In this study, we focused on scaling phenomena concerning both complete and partial reset in French. Specifically, our main goal was to investigate scaling data (i.e. the domain of downstep both within and across prosodic phrases) showing the existence of an intermediate level of phrasing. This constituent was predicted to show different phonological and phonetic properties from constituents at the lower and higher levels, the ACCENTUAL PHRASE (AP) and the IP. Our main hypothesis was that an intermediate phrase break would block recursive downstep of subsequent AP-final LH* rises in declarative utterances, causing the return of a preboundary H target to the pitch level set by the first accentual peak of the phrase (i.e. complete reset). Since duration cues (i.e. preboundary lengthening) can also be employed to mark the location of prosodic boundaries in French (e.g. Jun & Fougeron 2000, Astésano 2001, Michelas & D'Imperio 2012) as well as in other languages (e.g. Lehiste 1973, Klatt 1975, Beckman & Edwards 1990, Wightman *et al.* 1992), preboundary lengthening was also investigated, in order to test the presence of an ip boundary at the target-syllable location. Hence, both tonal-scaling and duration analyses were at the heart of our study, with the aim of testing whether an additional level of phrasing needs to be postulated in French. Secondly, it is widely acknowledged that prosodic phrasing is highly constrained by speech rate (e.g. for French Barthkova 1991, Keller & Zellner 1995, Fougeron & Jun 1998). Thus, as a secondary goal of our study, we tested how speech rate interacts with both tonal scaling and lengthening cues relative to the existence of an ip in French. But before discussing the instrumental study, we will review the basics of French intonational phonology.

1.2 French intonational grammar

Although there is an ongoing debate on French prosody, there is general agreement on a number of aspects.[2] Descriptions of French intonation generally agree on the existence of a high level of constituency, the Intonation Phrase (Jun & Fougeron 2000, 2002, Post 2000). Moreover, most intonation models of French include a lower unit as well, although different terms and definitions are employed to describe it. In Jun & Fougeron's account of French intonation, this tonally defined unit is labelled the ACCENTUAL PHRASE (Jun & Fougeron 2000, 2002), which in Post's (2000) model corresponds to the PHONOLOGICAL PHRASE,

[2] In this study, we adopt an autosegmental-metrical approach, in which the notion of a universal hierarchy of prosodic types is generally accepted. In this framework, the prosodic hierarchy results in a limited number of abstract prosodic levels that differ across languages. The phonetic and phonological properties of these levels are also language-dependent.

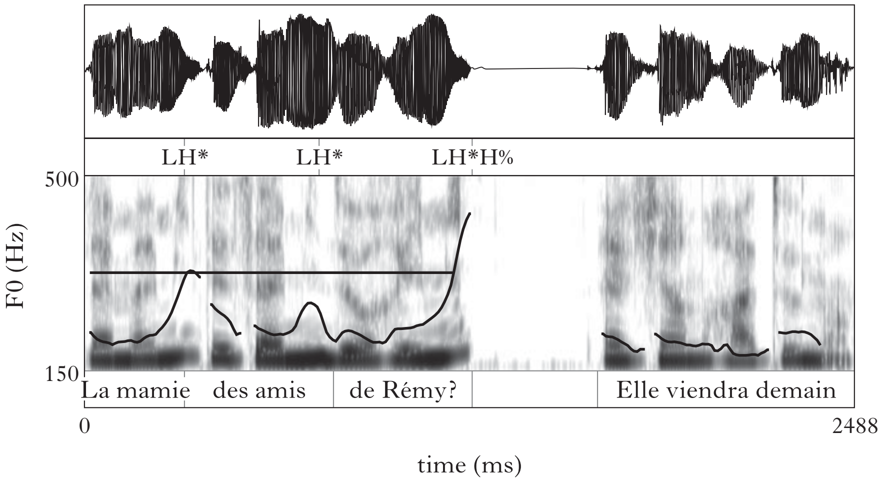

Figure 2

F0 track and spectrogram for the utterance *La mamie des amis de Rémy? Elle viendra demain*. 'The grandmother of Remy's friends? She is coming tomorrow.' The unbroken horizontal line represents the reference pitch level for the first phrase. Adapted from Michelas (2011).

a rhythmic constituent built on the basis of prosodic phonological rules referring to syntax. Thus, in Post's account, Phonological Phrases do not play a role in the tonal description, and only determine which metrically strong syllables receive a pitch accent. Here, in line with recent studies accounting for a tone specification of a unit smaller than the IP (Welby 2006, Michelas & D'Imperio 2012), we adopt Jun & Fougeron's definition of the AP. According to this view, non-final APs are typically marked by a final LH* rise associated to the last metrically strong syllable, i.e. the last full syllable of the phrase (Dell 1984).[3] The AP domain might also be marked by an optional initial, non-accentual, LHi rise (see Welby 2006 for its phonological and phonetic alignment properties). In this paper, we will generally discuss only accentual (final) LH* rises, marking primary stresses at the right edge of an AP.

It is interesting that, regardless of the definition of a unit smaller than the IP, the level of subsequent H targets within the same IP is claimed to be lowered due to iterative downstep (Post 2000), while upstep is responsible for raising the level of an H% boundary tone relative to preceding accentual rises in both question and continuation contours. As shown in Fig. 2 for the utterance *La mamie des amis de Rémy? Elle viendra*

[3] Note that the tonal grammar proposed by Post (2000) treats this tonal event as a simple H*. Here we adopt Jun & Fougeron's (2000) analysis, in which the underlying structure of the rising pitch accent is bitonal, and hence the LH* notation will be employed from now on in what follows.

demain, the H% at the right edge of a confirmation question is generally upstepped relative to the previous accentual H targets (see also Michelas 2011).

As shown in Fig. 1, downstep of phrase-internal subsequent LH* peaks is blocked before a prosodic boundary which appears to be weaker than an Intonation Phrase break, and is labelled as an intermediate phrase boundary in autosegmental-metrical models (Beckman & Pierrehumbert 1986). Jun & Fougeron (2000) first proposed the existence of an ip for French, but in their model this constituent is linked to marked syntactic constructions (e.g. left-dislocated sentences) in which either a low or a mid pitch plateau can be found after a LH* rise. More recent work has shown that an ip break can be found in unmarked syntactic constructions, i.e. at the boundary between major syntactic constituents such as a subject Noun Phrase and a Verb Phrase (the NP/VP break), when prosodic length constraints are satisfied (D'Imperio & Michelas 2010, Michelas & D'Imperio 2012).

Here we investigate the tonal scaling of H targets in unmarked SVO declaratives, to test whether the existence of an ip break is necessary to account for register level shifts within the French IP. The paper is organised as follows: in the following section we present our research questions, the details of the experimental study and the results, in §3 we offer some elements of discussion and §4 provides a brief conclusion.

2 Instrumental study

2.1 Background and experimental questions

The aim of our study was twofold: first, to determine whether systematic variation in pitch-scaling and final lengthening cues supports the hypothesis of the existence of an ip boundary in French, and second, to test how speech rate affects the presence of its phonological and acoustic-prosodic cues.

To test our hypotheses, we employed a set of highly controlled sentences, in which we manipulated the length of the subject NP in order to obtain either a sequence of two APs (as in sentence (1a); see Fig. 3a) or three APs (as in sentence (1b); see Fig. 3b). We predicted that the target syllable (in boldface in (1)), corresponding to the H2 position in Fig. 3, would be followed by a right ip boundary only in (1a), since both a syntactic alignment constraint and a prosodic length constraint would conspire to ensure its placement. Conversely, the placement of an ip boundary is not predicted in (1b), where the target syllable is contained within the subject NP. Note that this contrastive analysis concerns only the scaling of the second accent within the ip (i.e. H2).

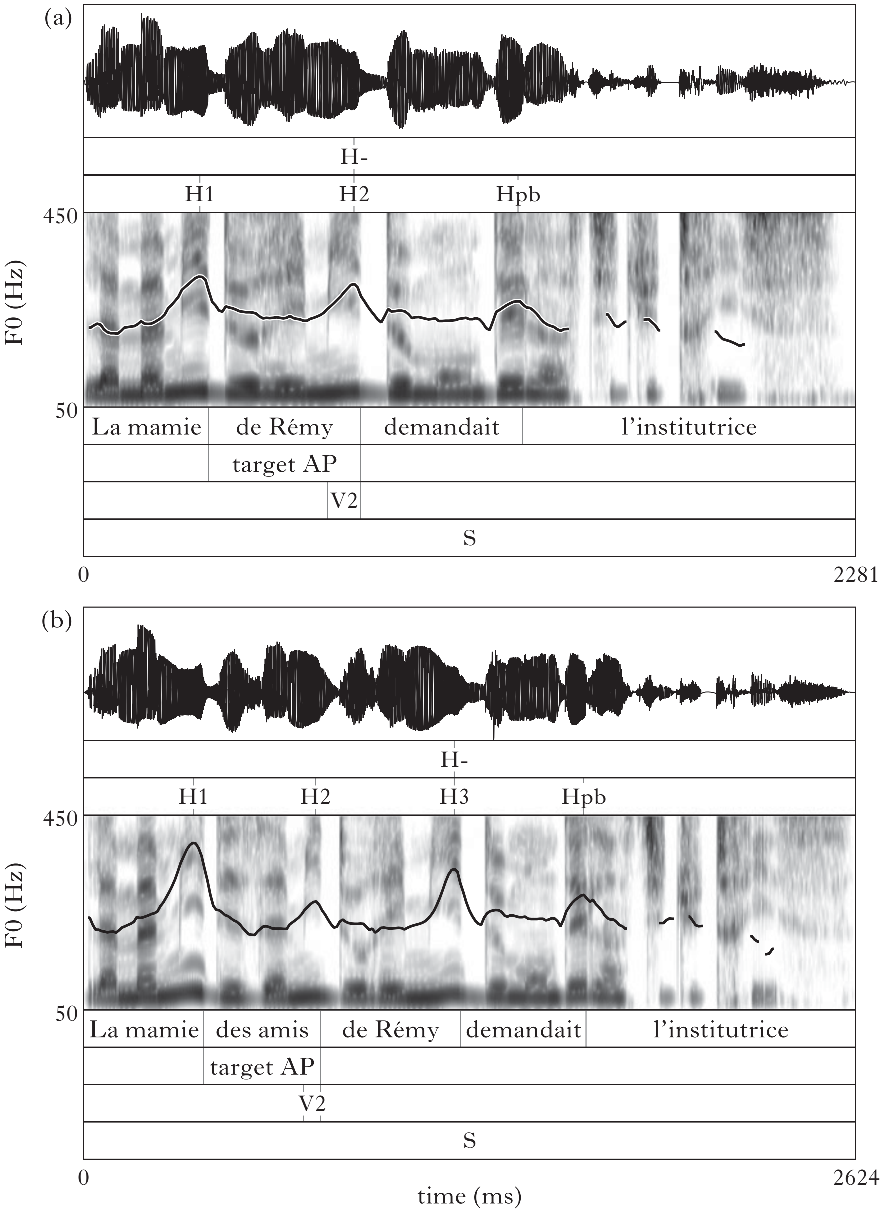

Figure 3

F0 tracks and spectrograms for (a) *La mamie de Rémy demandait l'institutrice* and (b) *La mamie des amis de Rémy demandait l'institutrice*, produced by two female speakers and exemplifying (a) the 2-AP condition and (b) the 3-AP condition. Note that H2 precedes an ip boundary only in the 2-AP condition. V2 is the duration of the last full vowel of the second AP of the utterance, S the duration of the whole utterance and target AP the duration of the second AP of the first prosodic phrase.

(1) a. La mamie de Ré**my** demandait l'institutrice.

$]_{\text{AP}}$ $]_{\text{AP/ip}}$ *prosodic structure*

$]_{\text{NP/VP break}}$ *syntactic structure*

'Rémy's grandmother was asking for the teacher.'

b. La mamie des a**mis** de Rémy demandait l'institutrice.

$]_{\text{AP}}$ $]_{\text{AP}}$ $]_{\text{AP/ip}}$ *prosodic structure*

$]_{\text{NP/VP break}}$ *syntactic structure*

'The grandmother of Rémy's friends was asking for the teacher.'

The main goal of our study was to test whether an intermediate level of prosodic phrasing exists in French. Note that the presence of an ip boundary is predicted at the right edge of an XP (X-phrase or maximal projection; D'Imperio & Michelas 2010), which in this case corresponds to the NP/VP boundary, when the syntactic phrase is produced as a sequence of at least two APs. We formulated two main hypotheses, the Complete Reset Hypothesis in (2) and the Partial Reset Hypothesis in (4).

(2) *Complete Reset Hypothesis*

The pitch level of a LH* rising accent immediately preceding an ip boundary is scaled at the level of the first LH* peak of the ip in which it is contained.

The Complete Reset Hypothesis concerns the tonal scaling of the last LH* rise before an NP/VP break (i.e. H2 in Figs 3a and 4a, and H3 in Figs 3b and 4b). Specifically, we predicted that when H2 and H3 immediately preceded an ip boundary they would have the same F0 value as H1. The first peak within the ip was therefore taken to set the reference level for the phrase. On the other hand, if the H peak did not precede an NP/VP break, and hence there was no ip boundary, it was predicted to be lowered relative to the preceding H. Given that only H2 could appear both in phrase-final and in phrase-internal position, we predicted this outcome only for phrase-internal H2 and not for H3, as it was always phrase-final in our material. In other words, H2 was predicted to be scaled lower than H1 in the 3-AP condition but not in the 2-AP condition.

Note that two alternative hypotheses can be formulated, as in (3).

(3) a. *Downstep Hypothesis*

The pitch level of a LH* rising accent immediately preceding an ip boundary is downstepped relative to the preceding peak within the same ip.

b. *Downstep Blocking Hypothesis*

The pitch level of a LH* rising accent immediately preceding an ip boundary is scaled at the same level as the preceding peak within the same ip.

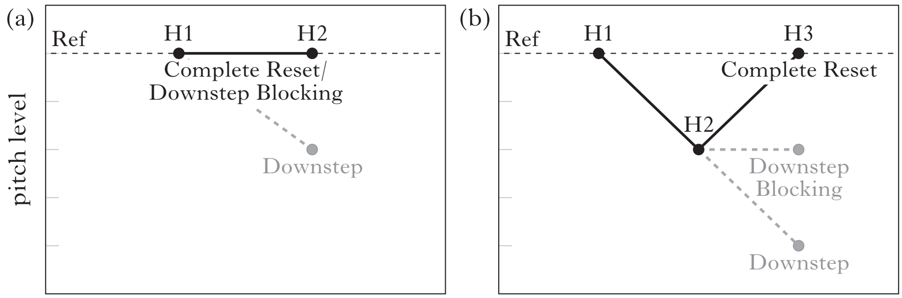

Figure 4

Schematisation of the Complete Reset Hypothesis (in black) and of the alternative hypotheses (Downstep Blocking and Downstep; in grey) for utterances in (a) the 2-AP condition and (b) the 3-AP condition.

Our main hypothesis, the Complete Reset Hypothesis in (2), and the alternative hypotheses in (3) are illustrated in Fig. 4 for both the 2-AP and 3-AP conditions.[4] While the Downstep Hypothesis would predict iterative downstep on all H peaks following H1 up to the end of the Intonation Phrase, Downstep Blocking would predict H2 and H3 to be set at the same level as the preceding H target. Note that the tonal behaviour of H3 (in 3-AP utterances) is crucial for testing our Complete Reset Hypothesis. According to this hypothesis, H3 would be completely reset, and hence be scaled at the same level as H1 (not that of H2) in the 3-AP condition. Alternatively, H3 could either be lowered relative to H2 (in line with the Downstep Hypothesis) or be realised at the same height as H2 (in line with the Downstep Blocking Hypothesis).

We first compared the F0 level of the second LH* rise (H2) in both the 2-AP and the 3-AP conditions. Since we postulated that an ip boundary would be placed after H2 only in the 2-AP condition but not in the 3-AP condition, we predicted H2 to be lower in the 3-AP condition than in the 2-AP condition, due to downstep triggered by H1. This is because the Complete Reset Hypothesis predicts pitch reset to occur at the target syllable only when the latter is aligned with the right edge of an ip. The preboundary H was predicted to be completely reset, independent of the number of APs in the NP, so that we predicted complete reset to apply also to H3, which was always in potential ip-final position in our material.

Apart from predicting complete reset targeting the preboundary H, we also made specific predictions concerning the pitch level of the post-boundary peak (Hpb). Specifically, we hypothesised that phrasal down-step is realised through partial reset, which is formulated as (4).

[4] An additional hypothesis could be formulated, that of H-Raising (cf. Laniran & Clements 2003), by which the preboundary H3 would be raised (relative to H2) by a variable amount. This hypothesis is not explicitly tested here, though it is briefly discussed in §3.

(4) *Partial Reset Hypothesis*

The level of the first LH* peak after an ip boundary (Hpb) is scaled lower than the IP-initial LH* accent (H1), but higher than H2.

Following Féry & Truckenbrodt (2005), we thus predicted that the downstep of two subsequent LH*s within an NP (i.e. between H1 and H2) would be greater than downstep across an ip boundary. As a consequence, the first LH* pitch accent of the second ip (Hpb) was predicted to be scaled lower than H1 in 2-AP phrases, but higher than H2. We therefore also measured the pitch level for Hpb, in order to test whether it would be affected by partial reset. We will also discuss whether this effect can be modelled as a return to a downstepped reference line for the second ip, as proposed by Féry & Truckenbrodt.

In order to provide additional evidence for the presence of an ip break, we also measured the duration of the stressed vowel of the target syllable, corresponding to the H2 position, in both the 2-AP and the 3-AP conditions. This is because the target syllable was predicted to be longer in the 2-AP condition (being ip-final) than in the 3-AP condition (being AP-final). Our third hypothesis is formulated as in (5).

(5) *Preboundary Lengthening Hypothesis*

Target vowels (V2) contained in syllables immediately preceding an ip boundary (i.e. at the H2 location in the 2-AP condition) are longer than ip-internal/AP-final syllables.

We expected to find greater preboundary lengthening for the target syllable associated to H2 in the 2-AP condition than in the 3-AP condition.

The second aim of our study was to test how speech rate interacts with both phonological and acoustic-prosodic cues to ip boundary realisation. Speech rate is generally viewed as an important influence on prosodic phrasing (see for French Barthkova 1991, Keller & Zellner 1995 and Fougeron & Jun 1998). For instance, fast rate has been claimed to produce restructuring of prosodic phrases (i.e. prosodic boundary erasure and/or boundary demotion; Fougeron & Jun 1998). Another possible outcome was that phonetic cues to phrasing would be reduced at fast speech rate. Hence, it was important to test whether the three hypotheses in (2), (4) and (5) are affected by an increase in speech rate. Thus speech rate was manipulated as a factor in our study by having our speakers produce the material at two speech rates (normal and fast). Following Michelas & D'Imperio (2012), who observed only minimal prosodic restructuring at fast rate, we formulated the hypothesis in (6).

(6) *Fast Speech Rate Hypothesis*

Fast speech rate does not erase ip-boundary insertion.

Given this hypothesis, we predicted that both complete reset and preboundary lengthening would be realised on our target syllables (H2 in the 2-AP condition and H3 in the 3-AP condition), independent of speech rate. Additionally, we predicted that rate would not interact with scaling effects for Hpb, i.e. we expected partial reset to occur after the ip boundary, independent of speech rate.

2.2 Methodology

2.2.1 *Corpus.* Given that the goal of our experiment was to precisely evaluate both the F0 level of an H peak within an LH* and preboundary lengthening, it was crucial to employ test materials in which factors influencing prosodic phrasing – length of phrase, syntactic, discourse and speech-rate factors – were highly controlled. Thus our materials included a set of 16 pairs of SVO utterances, in which the structure of the subject NP was varied in order to obtain either a sequence of two or three APs, as exemplified in (1) above.

In order to avoid potential effects of AP length on phrasing (Astésano 2001, Welby 2003), each AP contained only three syllables. For the example in (1b), each of the three APs composing the subject NP, *la mamie*, *des amis* and *de Rémy*, contained three syllables. Syntactically, the target syllable for H2 always preceded a major syntactic break corresponding to an NP/VP boundary (i.e. the right edge of an XP) in the 2-AP condition, while it only preceded a minor syntactic break (i.e. a boundary between a head and its complement) within the NP in the 3-AP condition. Like preboundary H2 in 2-AP phrases, H3 was always associated to a major syntactic break in 3-AP phrases.

We also know that prosodic boundary strength is affected by discourse position, with more substantial cues – such as greater preboundary lengthening, longer pauses and higher F0 peaks – at the beginning of a discourse segment than at its end (Hirschberg & Pierrehumbert 1986, Crystal & House 1990, Swerts 1997, den Ouden *et al.* 2009). To avoid any effect of discourse structure on prosodic phrasing, we therefore employed a set of isolated utterances, for which pragmatic structure was held constant (i.e. ALL-FOCUS utterances).

In order to minimise possible microprosodic effects due to obstruent perturbations, voiceless gaps or intrinsic F0 perturbations, the segmental composition of the target syllables was kept constant for each utterance pair, so that the consonants preceding and following the target vowel would always be voiced. The set of 32 experimental sentences is given in the Appendix. Note that, in order to increase statistical power, each sentence was presented four times to each of the nine speakers. There were thus 2304 tokens in total: 32 sentences (with 2-AP and 3-AP subject NPs) × 4 repetitions × 2 speech rates × 9 speakers. The experimental sentences were interspersed with 90 filler items, in the form of declarative sentences with a variety of syntactic structures. Two lists were created, such that each member of a pair of experimental sentences appeared in a different list. Each participant was presented with both lists, with half of the participants reading List 1 first and the other half reading List 2 first. Finally, in each list, the order of appearance was randomised by participant.

2.2.2 *Experimental procedure.* Speakers were recorded individually in a soundproof room at the *Laboratoire Parole et Langage* at Aix Marseille University, using a Shure SM10A head-mounted microphone. Speech samples were digitally recorded at 32,000 Hz, 16-bit. Participants were seated at a computer screen and were first asked to read aloud the set of sentences at a self-selected normal rate. They were also told to repeat the utterance if they misread part of the text. The experimenter instructed participants to read the sentences as if conveying new information. This allowed us to obtain all-focus utterances in reading style, avoiding possible effects of topic or focus intonation. In the second part of the experiment, speakers read the same set of utterances as in the first part, but at a fast, self-selected, speech rate. The instructions were given as follows: 'You are going to see the same set of utterances as in the preceding part of the experiment. Your task is to read these sentences as fast as you can, as if you were talking with someone who is in a hurry'. Sentences appeared ten at a time on a computer screen; a metronome was not used to control for rate. Before the experiment began, participants read six sentences at both normal and fast speech rates, in order to test the material and become acquainted with the procedure. Post hoc measurements were performed in order to be sure that all participants had increased their speech rate from normal to fast (see §2.3.1 below).

Each individual utterance was then isolated and saved as a separate file. Phonetic segmentation was automatically carried out using EasyAlign (Goldman 2011) and subsequently hand-corrected by the second author after inspection of both waveforms and spectrograms. Nine native speakers of Southern French (three males and six females, aged 24 to 34) participated in the experiment. They were all born in Southern France, where they had spent most of their life. Most of the participants were graduate students at Aix Marseille University.

2.2.3 *Measurements.* The F0 values associated to target syllables were measured first, in accordance with the widely accepted view that F0 is a direct correlate of tonal variation, both for pitch accents and for boundary tones (Pierrehumbert 1980, Ladd 2008). The location for each target H tone (H1, H2, H3 or Hpb) was labelled, and F0 values were automatically extracted, using Praat scripts (Boersma & Weenink 2009). Moreover, in order to evaluate the tonal target values not only in local terms but also in relation to the reference pitch level for the phrase, we also calculated the H2/H1, H3/H1 and Hpb/H1 ratios. This is because the first peak in the IP, i.e. H1, can be considered to set the reference level for the phrase. Note that ratios greater than or equal to 1 indicate that the tonal target is not downstepped relative to H1 (as for downstep blocking and/or complete reset). On the other hand, ratios smaller than 1 indicate that the tonal target is downstepped relative to H1. In order to test the Complete Reset Hypothesis, H3 was also labelled in 3-AP NPs, and its maximum F0 value was subsequently measured.

Finally, in order to test the Partial Reset Hypothesis in non-initial intermediate phrases, we labelled the H peak of the first LH* rise immediately following the ip boundary (Hpb) in both 2-AP and 3-AP conditions. The tonal labelling scheme was as in (7) (see also Fig. 3 above).

(7) H1: F0 maximum in the region of the last full syllable of the first AP of the utterance

H2: F0 maximum in the region of the last full syllable of the second AP of the utterance

H3: F0 maximum in the region of the last full syllable of the third AP of the utterance (only for the 3-AP condition)

Hpb: F0 maximum in the region of the last full syllable of the first AP following the ip boundary

Recall that the first aim of our study was to provide additional evidence for the existence of an intermediate level in French by uncovering the role of scaling relationships among subsequent pitch accents. Note though that tonal scaling is not the only cue to phrasing, since duration is an additional robust cue to prosodic boundary marking in a variety of languages, including French (Vaissière 1983, Jun & Fougeron 2000, Michelas & D'Imperio 2012, among others). The duration of the vowel in the H2 location was also measured (it is labelled 'V2' in Fig. 3 and elsewhere), so that the 2-AP and 3-AP conditions could be compared in terms of potential preboundary lengthening. Since we know that lengthening preferentially affects the syllable rhyme in French (Astésano 2001), and given that we included only open syllables in our study, final vowel-duration measurements were taken to adequately reflect degree of preboundary lengthening. A Praat script was employed in order to automatically extract V2 duration in both conditions.

In order to evaluate whether speakers effectively modified speech rate from normal to fast, the duration of the entire utterance (labelled 'S' in Fig. 3) was also measured, as well as that of the target APs. Speech rate was then calculated as utterance duration divided by the total number of syllables produced.

2.2.4 *Statistical analyses.* Statistical analyses were carried out using linear mixed-effects models, which provide a powerful tool for the analysis of grouped data (Pinheiro & Bates 2000, Quené & van den Bergh 2004), although the question of degrees of freedom for these models continues to be a subject of controversy in the literature. Hence a Monte Carlo Markov Chain Technique (Baayen 2008, Bates & Maechler 2010) was used to obtain p-values (MCMC p-values). All data analyses were carried out using the R statistical package (R Development Core Team 2008).

	AM		CD		CDe		CP		NM	
	S	AP	S	AP	S	AP	S	AP	S	AP
normal rate	6·0	6·4	5·8	6·1	5·9	6·8	5·5	5·9	7·1	7·2
fast rate	7·6	8·4	8·1	8·6	8·1	9·8	6·6	7·2	8·5	9·1
rate increase (%)	26·7	31·2	39·7	41·0	37·3	44·1	20·0	22·0	19·7	26·4

	PP		RF		SM		YL	
	S	AP	S	AP	S	AP	S	AP
normal rate	6·1	6·0	7·1	7·3	7·0	7·1	6·1	6·1
fast rate	7·8	7·7	8·5	9·5	8·0	7·9	7·7	7·4
rate increase (%)	27·9	28·3	19·7	30·1	14·3	11·3	26·2	21·3

Table I

Averages for normal and fast speaking rates (in syllables per second) for both entire utterances (S) and target Accentual Phrases (AP) for each speaker (AM, CD, CDe, CP, NM, PP, RF, SM, YL). The increase from normal to fast speaking rate is given as a percentage.

2.3 Results

2.3.1 *Speech rate.* In order to test if speech rate would influence the prosodic cues under investigation, we instructed our speakers to read the sentences at two self-selected speech rates (normal and fast). To assess whether speakers successfully increased their speech rate from normal to fast, we calculated average rate in syllables per second for the entire utterance. Average rate was also calculated for target APs, to check whether whole utterances and APs were similarly affected by rate change. We first calculated the total duration of each utterance and each target AP. Subsequently, each utterance and each AP was assigned a phone count, which was then adjusted to reflect actual pronunciation, including the presence of word-final schwas or elisions.

As shown in Table I, speaking rate and the magnitude of the change from one rate to the other varied across speakers, depending on the size of the unit under investigation (entire utterances *vs.* target APs). For example, Speaker CP had the slowest rate for both entire utterances (5·5 syll/sec) and target APs (5·9 syll/sec). Speakers NM and RF had the fastest rate for entire utterances (8·5 syll/sec), while Speaker CDe showed the fastest rate for target APs (9·8 syll/sec). A series of t-tests confirmed that all speakers successfully increased their speech rate from normal to fast in both entire utterances and target APs (p values were always less than 0·05). These values are in line with previous work on speech rate for French

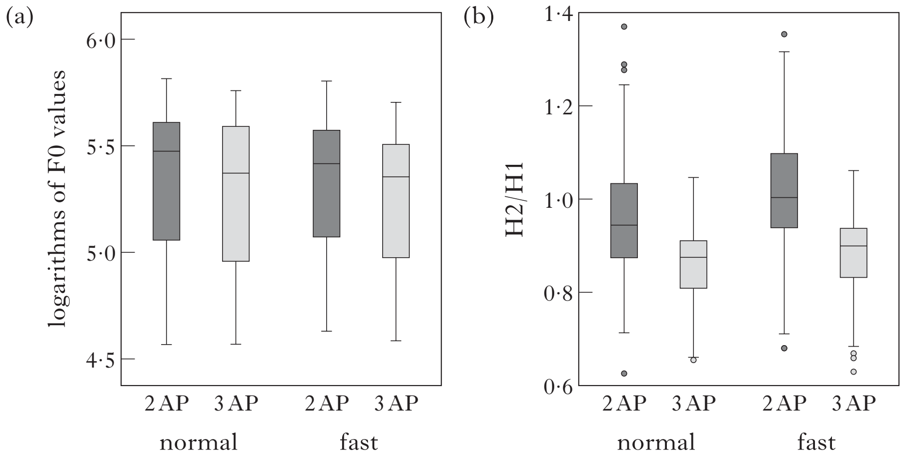

Figure 5
Boxplots of (a) log-transformed F0 values for H2 and (b) H2/H1 ratios according to AP Number (2 *vs*. 3) and Rate (normal *vs*. fast).

(Fougeron & Jun 1998, Welby 2006, Welby & Lœvenbruck 2006, Michelas 2011).

2.3.2 *F0 results*. In order to test the Complete Reset Hypothesis for tonal values, a linear mixed-effects model was fitted to the logarithms of the F0 values for H2, including AP Number (2 *vs*. 3) and Rate (normal *vs*. fast) as fixed effects.[5] Random intercepts were also included for the effects of Speaker and Syllable. 2304 observations were included by the model. In line with both the Complete Reset Hypothesis and the Downstep Blocking Hypothesis, the results show a main effect of AP Number, independent of Rate (see Fig. 5a). At both normal and fast speech rates, F0 height was significantly lower for the 3-AP condition than for the 2-AP condition (normal speech rate: $\beta = -0{\cdot}066$, $se = 0{\cdot}006$, $t = -11{\cdot}51$, pMCMC $< 0{\cdot}001$; fast speech rate: $\beta = -0{\cdot}088$, $se = 0{\cdot}006$, $t = -15{\cdot}36$, pMCMC $< 0{\cdot}001$). Hence, in line with the Fast Speech Rate Hypothesis, rate did not influence phrasing effects in our corpus. It also did not influence F0 height ($\beta = -0{\cdot}0001$, $se = 0{\cdot}006$, $t = -0{\cdot}02$, pMCMC $= 0{\cdot}9012$), so that the effect of AP Number was observed at both rates.

5 F0 values were log-transformed, in order to normalise the range of variability found both within and across speakers. Note that there is no generally accepted approach to normalising F0 data across speakers, especially in a group including both males and females. Here, following recent studies analysing F0 values in a corpus including speakers of different genders (e.g. Reubold *et al*. 2010), we used a common logarithm transformation (also known as decadic logarithm transformation). Note that this scale is mathematically equivalent to a semitone conversion, since both scales are logarithmic transformations. The formula employed for the semitone conversion uses a base-$2^{1/12}$ (or equivalently 12 times the base-2 logarithm) while the formula employed here used a base-10.

Recall that in order to evaluate pitch scaling within subject NPs, we also measured H2 as a ratio of the first target peak (H1). A linear mixed-effects model was fitted to the log of the H2/H1 ratios.[6] We retained AP Number and Rate as fixed factors and Speaker and Syllable as random factors. The model included 2304 observations. Results show a main effect of AP Number (see Fig. 5b), in that H2/H1 ratios were smaller in the 3-AP condition than in the 2-AP condition at both normal ($\beta = -0{\cdot}094$, $se = 0{\cdot}005$, $t = -17{\cdot}98$, $p\text{MCMC} < 0{\cdot}001$) and fast speech rates ($\beta = -0{\cdot}129$, $se = 0{\cdot}006$, $t = -22{\cdot}036$, $p\text{MCMC} < 0{\cdot}001$). Values estimated by the model are shown in Table II.[7]

	normal	fast
2-AP condition	0·94	1·00
3-AP condition	0·85	0·88

Table II

H2/H1 ratios estimated by the linear mixed-effects model for both 2-AP and 3-AP conditions, depending on Rate.

As shown in Table II, H2/H1 ratios for 2-AP utterances at fast rate, with values equal to 1, clearly support both the Downstep Blocking Hypothesis and the Complete Reset Hypothesis, but not the Downstep Hypothesis. At normal speech rate, H2/H1 ratios were lower than 1 (0·94), but higher than the values estimated for the 3-AP condition at the same rate (0·85). Hence, though the results do not lend support to the Downstep Hypothesis, they nevertheless do not allow us to distinguish between the Downstep Blocking Hypothesis and the Complete Reset Hypothesis at normal rate.

Supplementary analyses of H3 height in the 3-AP condition were then carried out in the 3-AP condition. A linear mixed-effects model was fitted to the log of the F0 values of H1, H2 and H3 for 3-AP utterances. H Target Type (H1, H2, H3) and Rate were included as fixed factors and Speaker and Syllable as random factors. There were 6912 observations. The model first showed that H3 was higher than H2 at both rates (normal: $\beta = 0{\cdot}106$, $se = 0{\cdot}006$, $t = 17{\cdot}88$, $p\text{MCMC} < 0{\cdot}001$; fast: $\beta = 0{\cdot}110$, $se = 0{\cdot}006$, $t = 18{\cdot}63$, $p\text{MCMC} < 0{\cdot}001$). This is in contrast with both the Downstep Hypothesis and the Downstep Blocking Hypothesis. However, although the difference is very small, the linear mixed-effects model fitted on the logarithms of F0 values (Fig. 6a) showed that H3 was lower

[6] Taking the log transformation of ratios ensures a normal distribution of the residuals (see Baayen 2008).

[7] Note that values estimated by the model are not mean values, but medians, since we used logarithms. Indeed, back transformations of logarithms correspond to geometric means, which are equal to medians (according to the properties of a log-normal distribution).

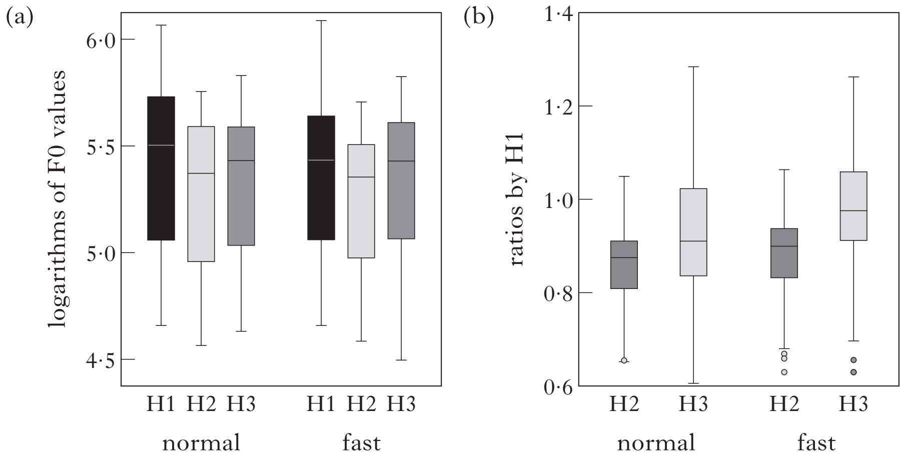

Figure 6

Boxplots of (a) log-transformed F0 values and (b) ratios by H1 according to H Type and Rate (normal *vs*. fast) for the 3-AP condition.

than H1 at both rates (normal: $\beta = -0{\cdot}048$, $se = 0{\cdot}006$, $t = -8{\cdot}03$, $p\text{MCMC} < 0{\cdot}001$; fast: $\beta = -0{\cdot}012$, $se = 0{\cdot}006$, $t = -2{\cdot}19$, $p\text{MCMC} < 0{\cdot}05$).

Since the linear mixed-effects model reports a significant difference between H1 and H3, the results do not immediately appear to validate the Complete Reset Hypothesis (which predicts that F0 values for H3 should be equal to those for H1). However, the estimated difference between H1 and H3 at normal speech rate was only 2 Hz, and the estimated difference between H1 and H3 at fast speech rate was only 1 Hz. Given the magnitude of the difference, the results suggest that the Complete Reset Hypothesis is confirmed if a plausible declination effect is taken into account (see the discussion in §3 below).

Given that we predicted that the H3/H1 ratio would be equal to 1, we also fitted a linear mixed-effects model on the log of the ratios for utterances of the 3-AP condition, with H Target Type (H2 and H3) and Rate as fixed factors and Speaker and Syllable as random factors. The model included 4608 observations. The results showed that H1/H3 ratios were significantly greater (i.e. closer to 1) than H2/H1 ratios at both rates (see Fig. 6b). The estimated value for H3/H1 at normal speech rate was 0·91 (*vs*. 0·85 for H2/H1), whereas the estimated value for fast speech rate was closer to 1 (i.e. 0·93 for H3/H1 *vs*. 0·88 for H2/H1). Hence, the results disconfirm both alternative hypotheses (Downstep and Downstep Blocking), since H3 was always higher than H2 (which was in turn downstepped relative to H1 in the 3-AP condition). Again, as for log-transformed F0 values, it appears that the Complete Reset Hypothesis cannot be completely validated either at fast or at normal speech rate, given that a full return to the reference level set by H1 was not found.

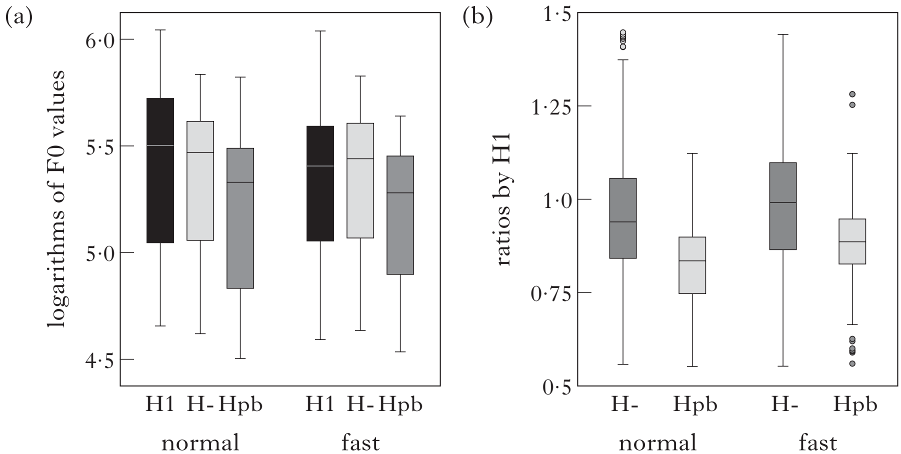

Figure 7

Boxplots of (a) log-transformed F0 values and (b) ratios by H1 by H Type (H- conflates H2 in the 2-AP condition and H3 in the 3-AP condition) and Rate for both 2-AP and 3-AP conditions.

However, note that values for both rates were very close to 1 (0·93 for fast and 0·91 for normal rate). We will discuss these results in terms of a plausible declination effect.

In order to test the Partial Reset Hypothesis, we also measured the height of the first rising LH* accent immediately following the ip boundary, i.e. Hpb. Log F0 values for H1, H- (which conflate values for preboundary H2 and H3) and Hpb are illustrated in Fig. 7a, while simple Hpb-/H1 ratios are shown in Fig. 7b. Note in Fig. 7a that Hpb is lower than both H1 and H-, showing that the amount of reset was not complete but only partial. A linear mixed-effects model was fitted to the log of F0 values as well as to the log of H-/H1 and Hpb/H1 ratios. As in the previous statistical analyses, H Target Type, AP Number and Rate were included as fixed factors, while Speaker and Syllable were included as random factors. The two models included 6912 and 4608 observations respectively. Results showed that Hpb values were significantly lower than H- at both rates and for both the 2-AP and the 3-AP conditions (*p*MCMC values were all less than 0·001), ranging from 0·78 for 3-AP utterances at normal rate to 0·87 for 2-AP utterances at fast rate. Both AP Number and Rate were also significant, with lower Hpb values for the 3-AP condition than for the 2-AP condition ($\beta = -0{\cdot}027$, $se = 0{\cdot}007$, $t = -3{\cdot}84$, *p*MCMC $< 0{\cdot}05$) and for fast speech rate relative to normal speech rate ($\beta = -0{\cdot}0241$, $se = 0{\cdot}007$, $t = -3{\cdot}47$, *p*MCMC $< 0{\cdot}05$).

We then measured Hpb relative to H2 in the 3-AP condition. This test was intended to verify whether H2 downstep was greater than Hpb downstep, relative to H1 level. Log F0 values for H2 and Hpb, as well as

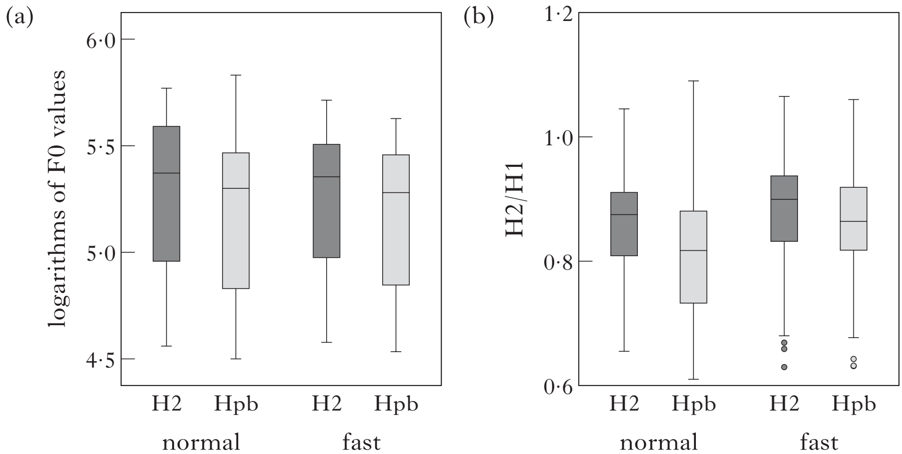

Figure 8

Boxplots of (a) log-transformed F0 values and (b) ratios by H1 according to H Type (H2/Hpb) and Rate for the 3-AP condition.

for H2/H1 and Hpb/H1 ratios, are reported in Fig. 8. Note in Fig. 8 that, unexpectedly, Hpb is lower than H2 for both log F0 values and ratios, which does not confirm that downstep across phrase boundaries is weaker than phrase-internal downstep. A linear mixed-effects model was fitted on the log F0 values (with H Target Type and Rate as fixed factors and Speaker and Syllable as random factors). The model included 2304 observations. The results show that Hpb was significantly lower than H2 at both normal ($\beta = -0{\cdot}0801$, $se = 0{\cdot}006$, $t = -12{\cdot}55$, $p\text{MCMC} < 0{\cdot}001$) and fast speech rates ($\beta = -0{\cdot}227$, $se = 0{\cdot}006$, $t = -3{\cdot}56$, $p\text{MCMC} < 0{\cdot}05$). The same result was found for ratio values for which Hpb/H1 was significantly higher than H2/H1 at both speech rates.

2.3.3 *Duration results*. Finally, we examined the effect of the number of APs on V2 duration, in order to test the Preboundary Lengthening Hypothesis in (5) above. A linear mixed-effects model was fitted to the log of duration values for V2 (at the H2 location). The model included two fixed factors (AP Number and Rate) and two random factors (Speaker and Syllable). 2304 observations were included by the model. The linear mixed-effects model showed a main effect of the number of APs. In line with the Preboundary Lengthening Hypothesis, and as shown by the duration data in Fig. 9, target vowels in the H2 position (V2) were always longer in the 2-AP than in the 3-AP condition, for both normal ($\beta = -0{\cdot}405$, $se = 0{\cdot}020$, $t = -25{\cdot}92$, $p\text{MCMC} < 0{\cdot}001$) and fast speech rates ($\beta = -0{\cdot}22633$, $se = 0{\cdot}01964$, $t = -14{\cdot}47$, $p\text{MCMC} < 0{\cdot}001$). This confirms that the target syllable was immediately followed by an ip boundary in the 2-AP condition, but not in the 3-AP condition. Note that the estimated

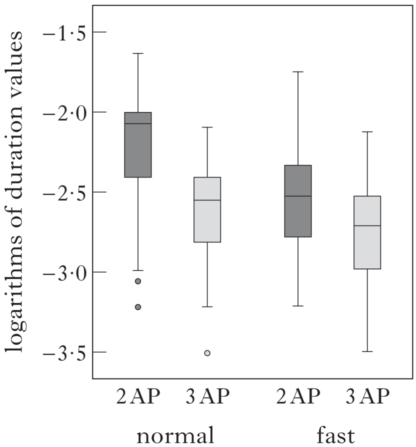

Figure 9
Boxplot of the log of V2 duration by Number of APs (2 *vs*. 3) and Rate (normal *vs*. fast).

values were also greater at normal rate (2-AP condition: 109 ms; 3-AP condition: 74 ms) than at fast rate (2-AP condition: 79 ms; 3-AP condition: 55 ms). The effect of Rate was also significant ($\beta = -0{\cdot}323$, $se = 0{\cdot}020$, $t = -20{\cdot}64$, $p\text{MCMC} < 0{\cdot}001$).

3 Discussion

Taken together, our scaling and duration data suggest that French speakers signal the presence of an intermediate phrase boundary by means of both preboundary lengthening and complete pitch reset. Specifically, the results of the tonal analysis for H2 in the 2-AP fast rate condition show that, in line with the Complete Reset Hypothesis and the Downstep Blocking Hypothesis, H2 targets were not downstepped relative to H1, since in this case H2 immediately preceded an ip boundary. Moreover, the pitch level of the preboundary H2 was completely equivalent to that of H1. On the other hand, H2 was downstepped in the 3-AP condition, given that in these sentences it was followed by an LH* accent (H3) occupying the preboundary position.

Note though that preboundary H2 was not completely reset at normal rate. In fact, the H2/H1 ratio was equal to 0·94 (rather than 1), which appears to disconfirm both the Complete Reset Hypothesis and the Downstep Blocking Hypothesis. However, the 6% ratio difference is in line with what has been argued as being UPSTEPPED values (i.e. COMPLETELY RESET in our terminology) by Truckenbrodt (2002) for German, invoking an F0 declination effect (see also Pierrehumbert & Beckman 1988 for Japanese and Audibert *et al*. 2010 for French) which gradually affects H targets from the beginning to the end of the utterance. The scaling difference between the two rates can hence be accounted for by the fact that at fast rate the 2-AP sequence was globally shorter (for APs the number of syllables per second was on average 2% higher at fast rate for the various speakers), hence the declination effect is expected to be weaker than for normal rate.

Results for H3 pitch level allowed us to discard both alternative hypotheses (Downstep and Downstep Blocking), while providing evidence in favour of the Complete Reset Hypothesis. Similar to preboundary H2 behaviour, the H3 level was not lowered (downstepped) relative to the preceding pitch accent (H2), which goes against the Downstep Hypothesis. Moreover, given that H3 was also always higher than H2, the Downstep Blocking Hypothesis can also be discarded. Similar to the H2 pitch values, a difference was found between the fast and the normal rate conditions for both the H3 log-transformed F0 data and the H3/H1 ratios. Unlike the H2 results, H3/H1 ratios were never equal to 1, either at normal (0·91) or at fast (0·93) speech rate. However, note that the actual F0 difference estimated by the model for both fast and normal rates was never greater than 2 Hz. Recall too that the values for medial peaks were always lower than the 0·90 threshold, showing that complete reset is never found away from the ip boundary.

Hence, the H3 results confirm our main hypothesis of a complete return to the reference level set by H1. This 'return to the reference line' for the prosodic phrase is a phonological process which also occurs in languages other than French, such as German (Truckenbrodt 2002, 2007) and Dutch (van den Berg *et al.* 1992). Here, though, this scaling phenomenon concerns a prosodic domain that is lower in the hierarchy than the IP, i.e. the ip. We propose that the right edge of a non-final ip is marked in French by the presence of an H- phrase accent blocking the iterative downstep on ip-internal, AP-final LH* rises and raising the H target to the level established by the first peak of the IP (H1).

Apart from tonal cues related to the realisation of both preboundary H2 and H3, tonal values for the first postboundary pitch accent (Hpb) offer additional support to the existence of a prosodic break which is stronger than a simple AP right edge, but weaker than an IP break. Our results in fact partially confirm the Partial Reset Hypothesis in (4), in that partial reset affects Hpb, though its tonal value is lower, rather than higher, than downstepped phrase-internal H2. We propose that the right edge of an ip is additionally marked by a phenomenon of partial reset targeting the first accentual H peak within the following ip. These results are partially in line with what has been shown for English, Dutch and German (Ladd 1988, van den Berg *et al.* 1992, Féry & Truckenbrodt 2005), suggesting the existence of different types of downstep: a local downstep, targeting subsequent H peaks, and a phrasal downstep, targeting reference lines within a larger IP. For English, Ladd (1988) first accounted for this phenomenon as downstep among abstract register features assigned to sister nodes. The model of embedded downstep (the 'wheels within wheels' model) has

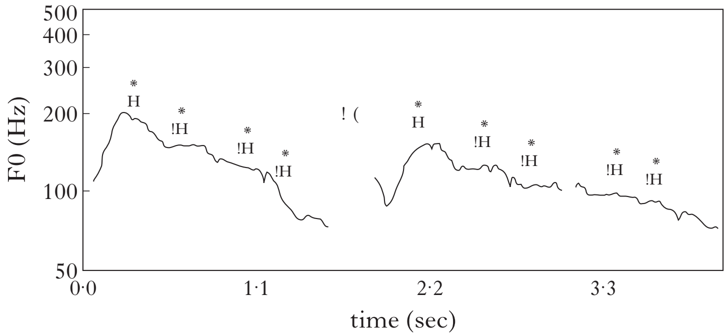

Figure 10

F0 contour for the Dutch utterance *Merel, Nora, Leo, Remy, en Nelie, Mary, Leendert, Mona en Lorna* 'Merel, Nora, Leo, Remy, and Nelie, Mary, Leendert, Mona and Lorna'. From van den Berg *et al.* (1992: 334).

since been empirically strengthened and theoretically modified, in particular by van den Berg *et al.* (1992) for Dutch and by Féry & Truckenbrodt (2005) for German. In Dutch, as shown in Fig. 10, the fifth H target (the one beginning the second phrase, on *Nelie*) is lower than the first H, yet higher than the downstepped second H of the utterance.

Féry & Truckenbrodt, on the basis of the Dutch data and their own German data, propose that downstep must be greater among prosodic sister nodes which are lower in the hierarchical representation ('the deeper, the steeper'). This hypothesis specifically predicts that downstep within an intermediate phrase would be greater than downstep among intermediate phrases, since in the latter case downstep would be internal to a higher-level domain, i.e. the IP.

As opposed to both Dutch and German, in our data the Hpb was scaled lower than downstepped H2, indicating that downstep across phrases is greater than phrase-internal downstep, rather than less. The scaling effects observed are schematised in Fig. 11.

Note, though, that Féry & Truckenbrodt's claim is based on the observation of mean target values pooled over the five speakers of Truckenbrodt's (2002) original study, in which speaker-specific differences might have been present. Indeed, Truckenbrodt (2007) later shows that, while some speakers displayed greater internal downstep, others either showed no difference between the two domains or displayed the opposite effect (e.g. Speaker 3 in that study). Our data appear to be generally in line with the idea of a partial reset in non-initial prosodic phrases as a consequence of phrasal downstep, though the effect appears to be greater (not smaller) among sister prosodic phrases than within the ip itself. It is therefore possible that the amount of downstep for the reference line of a non-initial ip is independently specified, and is a result neither of prosodic sisterhood relations nor of syntactic relations, as first proposed by Ladd

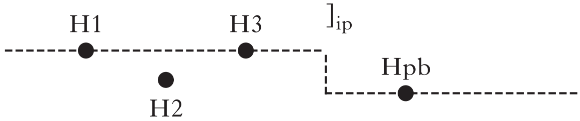

Figure 11

Schematisation of complete reset observed on the last accentual H target before the ip boundary (H3) and partial reset affecting the first accentual H target located after the ip boundary (Hpb) for French in the 3-AP condition.

(1988). In other words, while internal phrase downstep might be more systematic and 'grammaticalised', downstep among phrases might be subject to other variables, such as individual differences or pragmatic enhancement.

Our analysis is also supported by duration data, showing that the rhyme of the preboundary syllable is longer than the rhyme of syllables that are phrase-internal. In fact, our Preboundary Lengthening Hypothesis in (5) predicted that target vowels in primary stressed syllables in H2 position would be longer in the 2-AP condition than in the 3-AP condition, giving additional evidence for the presence of an ip break. Though a global rate effect was found, the duration difference between preboundary and phrase-internal H2 was significant, independent of the rate effect. Since both locations are also AP-final, while not being IP-final, the difference can only be interpreted in terms of the presence of the right edge of a prosodic constituent intermediate between the AP and the IP levels.

These data add to the body of evidence on durational effects and phrasing in French, given that differential preboundary lengthening has been previously shown for unstressed AP-final and IP-final syllables (Jun & Fougeron 2000) but not for ip-final syllables (see also Michelas & D'Imperio 2012). Furthermore, the crucial prosodic boundaries were not obliterated at fast rate, confirming our Fast Speech Rate Hypothesis in (6). This contradicts previous results (Fougeron & Jun 1998), while supporting more recent work (Michelas & D'Imperio 2012). In fact, a greater effect of tonal reset was found in the preboundary location at fast rate, accompanied by smaller preboundary lengthening. This result was unexpected, and could be explained by a compensatory phenomenon involving tonal cues, which might be enhanced to signal boundary location in the absence of adequate duration cues.

Note also that our data do not support the hypothesis that the target syllable might be variably upstepped relative to the previous peak (HIGH RAISING; see Laniran & Clements 2003), since H- values were not variably higher than H1, but rather targeted a specific tonal level. Our data are also compatible with the more recent model of tonal scaling proposed by Truckenbrodt (2007: 371), who claims that 'if a H tone is associated with prosodic level X, it is scaled to Ref(X)', and reference lines are lowered for each new prosodic phrase. However, contrary to what happens in

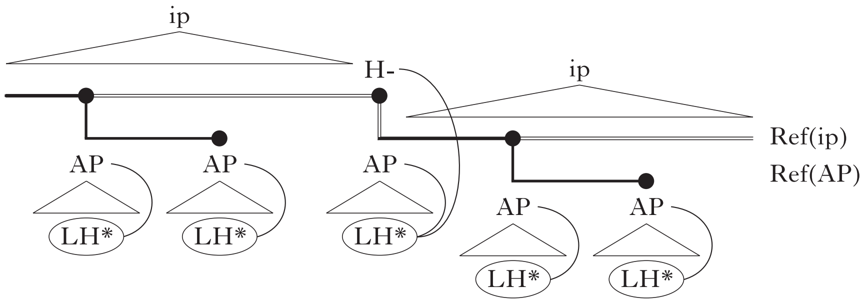

Figure 12

Schematisation of downstep between LH* rises within and across an ip boundary in French. Adapted from Truckenbrodt (2007).

German, the targeted domain in French is the ip, not the IP. More generally, our study is in line with the claim that prosodic constituents affect the assignment of scaling reference lines to target tones. Hence our data are not incompatible with Truckenbrodt's model, in which tone scaling is determined by secondary association of a pitch accent to a phrasal domain, such as an ip. This is shown in Fig. 12, adapted from Truckenbrodt (2007).

Taken together, our tonal scaling and lengthening results are not in line with the predictions of the French intonation model proposed by Post (2000), which includes only two levels of phrasing, the Phonological Phrase and the Intonation Phrase. Moreover, in Post's model the domain of downstep is the IP, hence H targets are predicted to be iteratively lowered up to the end of the phrase, for which our results provide counterevidence. In her model, upstep is in fact only envisaged for H% edge tones (i.e. in yes/no questions and continuation rises). More specifically, the model incorrectly predicts iterative downstep on H2, H3 and Hpb, independent of the presence of an ip boundary, while predicting neither complete reset on H- nor preboundary lengthening at this location.

On the other hand, our results are consistent with a prosodic model of French which includes an intermediate level of phrasing between the IP and the AP, signalled by both local preboundary lengthening and global tonal structuring of pitch targets. In fact, the results of the tonal analysis indicate that F0 values are significantly greater for the target syllable (H2 or H3) when the latter is crucially placed at the end of an AP that is also ip-final, rather than merely AP-final. Our results also confirm that the presence of an ip break in French is not limited to specific syntactic or discourse structures, in contrast to what has been proposed in earlier accounts (see Jun & Fougeron 2000). Recall that in our data the crucial ip boundary co-occurs with a major syntactic break (i.e. the subject/predicate break), which indicates that a syntax/prosody constraint is active in French phrasing, and guides the placement of ip boundaries. In other words, an ip break can be placed in simple SVO utterances once both syntactic and prosodic constraints are satisfied.

4 Conclusion

Our results show an interesting structuring of subsequent accentual rises within complex subject NP phrases in French, which is analysed as the result of the presence of an intermediate phrase break at the NP/VP boundary. We have therefore claimed that an IP-internal boundary, i.e. the ip boundary, is signalled by phonological and phonetic cues which are crucially weaker than those associated with the IP boundary, but stronger than those found for the AP boundary. Specifically, target syllables were produced with lower F0 values when in non-final position within the ip, while they were fully reset when immediately preceding an ip right edge (complete reset). Moreover, as expected, greater lengthening values for the target vowel were found within ip-final syllables than within non-final ones. We hence propose that right-hand ip boundaries are signalled in French by a rising phrase accent (H-), characterised by a return to the pitch reference level set by the first H peak of the phrase (H1), as well as by preboundary lengthening. We also propose that partial reset in the initial H target after the ip-break (Hpb) is additional evidence for an internal structuring of the IP, though our data do not support the hypothesis of greater downstep within, rather than across, phrasal domains suggested for German (Féry & Truckenbrodt 2005). Finally, our data are also compatible with a local scaling model of the H- phrase accent, due to secondary association with a higher-level prosodic domain (i.e. the ip), while suggesting an independent scaling specification for the reference level of each separate intermediate phrase.

Appendix: Experimental utterances used for the corpus

For each pair of utterances, the subject Noun Phrase is composed of 2 APs in the (a) example, and 3 APs in the (b) example. Target accented syllables are given in boldface.

(1) a. La mamie de Ré**my** demandait l'institutrice.
  b. La mamie des a**mis** de Rémy demandait l'institutrice.

(2) a. Le mari de Car**ry** deviendra un grand docteur.
  b. Les carries du ma**ri** de Carry demandaient des soins urgents.

(3) a. Le sauna d'Hélé**na** deviendra le plus connu.
  b. La nana du sau**na** d'Héléna devenait vraiment méchante.

(4) a. Le trois-mâts de Tho**mas** devance tous les autres bateaux.
  b. Le schéma du trois-**mâts** de Thomas devenait vraiment brouillon.

(5) a. La mairie de Pa**ris** demandait des travaux urgents.
  b. Le parvis de la mai**rie** de Paris demandait des travaux urgents.

(6) a. Le tabac de Cu**ba** dessèche énormément la bouche.
b. Le tabac de Cu**ba** d'Alba dessèche énormément la bouche.

(7) a. Les parents de Lau**rent** devaient présenter leur dossier aujourd'hui.
b. Le garant des pa**rents** de Laurent devait présenter leur dossier aujourd'hui.

(8) a. Le roman de Clé**ment** deviendra un best-seller.
b. Le roman de la ma**man** de Clément deviendra un best-seller.

(9) a. Les colis d'Emi**lie** devraient arriver dans la semaine.
b. Les colis du Ma**li** d'Emilie devraient arriver dans la semaine.

(10) a. Les soirées du cu**ré** devenaient de plus en plus connues.
b. Les soirées du cu**ré** de Poiret devenaient de plus en plus connues.

(11) a. La villa de Car**la** demandait des travaux permanents.
b. La villa de Car**la** d'Angola demandait des travaux permanents.

(12) a. Le palais de Ca**lais** devient de plus en plus abîmé.
b. Les volets du pa**lais** de Calais deviennent de plus en plus abîmés.

(13) a. Le valet d'Aza**lée** devrait être un peu plus discret.
b. Le gilet du va**let** d'Azalée devrait être un peu plus discret.

(14) a. Les marmots d'Ada**mo** devenaient de plus en plus agaçants.
b. Le némo des mar**mots** d'Adamo devrait être acheté dans la journée.

(15) a. Le hameau de Ré**mo** demeurait un des plus joli du village.
b. Le hameau des ju**meaux** de Rémo demeurait un des plus joli du village.

(16) a. Le kiné de Re**né** devenait vraiment insupportable.
b. La journée du ki**né** de René devenait vraiment insupportable.

REFERENCES

Astésano, Corine (2001). *Rythme et accentuation en français : invariance et variabilité stylistique*. Paris : L'Harmattan.

Audibert, Nicolas, Véronique Aubergé & Albert Rilliard (2010). Prosodic correlates of acted *vs*. spontaneous discrimination of expressive speech : a pilot study. *Proceedings of Speech Prosody 2010*. Available (January 2014) at http://speechprosody2010.illinois.edu/papers/100097.pdf.

Baayen, R. H. (2008). *Analyzing linguistic data : a practical introduction to statistics using R*. Cambridge : Cambridge University Press.

Barthkova, Katarina (1991). Speaking rate modelization in French application to speech synthesis. In *Proceedings of the 12th International Congress of Phonetic Sciences*. Vol. 3. Aix-en-Provence : Université de Provence. 482–485.

Bates, Douglas M. & Martin Maechler (2010). Package 'lme4' (Version 0.999375-34) : linear mixed-effects models using S4 classes. http://cran.r-project.org/web/packages/lme4/lme4.pdf.

Beckman, Mary E. & Jan Edwards (1990). Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I : between the grammar and physics of speech*. Cambridge : Cambridge University Press. 152–178.

Beckman, Mary E. & Janet B. Pierrehumbert (1986). Intonational structure in Japanese and English. *Phonology Yearbook* **3**. 255–309.

Berg, Rob van den, Carlos Gussenhoven & Toni Rietveld (1992). Downstep in Dutch : implications for a model. In Gerard J. Docherty & D. Robert Ladd (eds.) *Papers in* *laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 335–359.

Boersma, Paul & David Weenink (2009). *Praat: doing phonetics by computer* (version 5.1.34). http://www.praat.org.

Crystal, Thomas H. & Arthur S. House (1990). Articulatory rate and the duration of syllables and stress groups in connected speech. *JASA* **88**. 101–112.

Dell, François (1984). L'accentuation dans les phrases en français. In François Dell, Daniel Hirst & Jean-Roger Vergnaud (eds.) *Forme sonore du langage*. Paris: Hermann. 65–122.

D'Imperio, Mariapaola & Amandine Michelas (2010). Embedded register levels and prosodic phrasing in French. *Proceedings of Speech Prosody 2010*. Available (January 2014) at http://speechprosody2010.illinois.edu/papers/100879.pdf.

Féry, Caroline & Hubert Truckenbrodt (2005). Sisterhood and tonal scaling. *Studia Linguistica* **59**. 223–243.

Fougeron, Cécile & Sun-Ah Jun (1998). Rate effects on French intonation: prosodic organization and phonetic realization. *JPh* **26**. 45–69.

Goldman, Jean-Philippe (2011). EasyAlign: an automatic phonetic alignment tool under Praat. *Proceedings of Interspeech 2011*. 3233–3236.

Hirschberg, Julia & Janet B. Pierrehumbert (1986). The intonational structuring of discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics, Columbia University*. Association for Computational Linguistics. 136–144.

Jun, Sun-Ah & Cécile Fougeron (2000). A phonological model of French intonation. In Antonis Botinis (ed.) *Intonation: analysis, modelling and technology*. Boston: Kluwer. 209–242.

Jun, Sun-Ah & Cécile Fougeron (2002). Realizations of accentual phrase in French intonation. *Probus* **14**. 147–172.

Keller, Eric & Brigitte Zellner (1995). A statistical timing model for French. In Kjell Elenius & Peter Branderud (eds.) *Proceedings of the 13th International Congress of the Phonetic Sciences*. Vol. 4. Stockholm: KTH & Stockholm University. 302–305.

Klatt, Dennis H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *JPh* **3**. 129–140.

Ladd, D. Robert (1988). Declination 'reset' and the hierarchical organization of utterances. *JASA* **84**. 530–544.

Ladd, D. Robert (2008). *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.

Laniran, Yetunde & G. N. Clements (2003). Downstep and high raising: interacting factors in Yoruba tone production. *JPh* **31**. 203–250.

Lehiste, Ilse (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa* **7**. 107–122.

Liberman, Mark & Janet B. Pierrehumbert (1984). Intonational invariance under changes in pitch range and length. In Mark Aronoff & Richard T. Oehrle (eds.) *Language sound structure*. Cambridge, Mass.: MIT Press. 157–233.

Michelas, Amandine (2011). *Caractérisation phonétique et phonologique du syntagme intermédiaire en français: de la production à la perception*. PhD dissertation, Aix Marseille University.

Michelas, Amandine & Mariapaola D'Imperio (2012). When syntax meets prosody: tonal and duration variability in French Accentual Phrases. *JPh* **40**. 816–829.

O'Shaughnessy, Douglas & Jonathan Allen (1983). Linguistic modality effects on fundamental frequency in speech. *JASA* **74**. 1155–1171.

Ouden, Hanny den, Leo Noordman & Jacques Terken (2009). Prosodic realizations of global and local structure and rhetorical relations in read aloud news reports. *Speech Communication* **51**. 116–129.

Pierrehumbert, Janet B. (1980). *The phonetics and phonology of English intonation*. PhD dissertation, MIT.

Pierrehumbert, Janet B. & Mary E. Beckman (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.

Pinheiro, José C. & Douglas M. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

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

Prieto, Pilar, Chilin Shih & Holly Nibert (1996). Pitch downtrend in Spanish. *JPh* **24**. 445–473.

Quené, Hugo & Huub van den Bergh (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication* **43**. 103–121.

R Development Core Team (2008). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available at http://www.r-project.org.

Reubold, Ulrich, Jonathan Harrington & Felicitas Kleber (2010). Vocal aging effects on F_0 and the first formant: a longitudinal analysis in adult speakers. *Speech Communication* **52**. 638–651.

Swerts, Marc (1997). Prosodic features at discourse boundaries of different strength. *JASA* **101**. 514–521.

Truckenbrodt, Hubert (2002). Upstep and embedded register levels. *Phonology* **19**. 77–120.

Truckenbrodt, Hubert (2007). Upstep on edge tones and on nuclear accents. In Carlos Gussenhoven & Tomas Riad (eds.) *Tones and tunes*. Vol. 2: *Experimental studies in word and sentence prosody*. Berlin & New York: Mouton de Gruyter. 349–386.

Vaissière, Jacqueline (1983). Language-independent prosodic features. In Anne Cutler & D. Robert Ladd (ed.) *Prosody : models and measurements*. Berlin: Springer. 53–66.

Welby, Pauline (2003). *The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. PhD dissertation, Ohio State University.

Welby, Pauline (2006). French intonational structure: evidence from tonal alignment. *JPh* **34**. 343–371.

Welby, Pauline & Hélène Lœvenbruck (2006). Anchored down in Anchorage: syllable structure, rate, and segmental anchoring in French. *Italian Journal of Linguistics* **18**. 73–124.

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *JASA* **91**. 1707–1717.