



**HAL**  
open science

## Anchored down in Anchorage: Syllable structure and segmental anchoring in French

Pauline Welby, H el ene Loevenbruck

► **To cite this version:**

Pauline Welby, H el ene Loevenbruck. Anchored down in Anchorage: Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics / Rivista di linguistica*, 2006, 18, pp.74-124. hal-00371892

**HAL Id: hal-00371892**

**<https://hal.science/hal-00371892>**

Submitted on 30 Mar 2009

**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 destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

Anchored down in Anchorage:  
Syllable structure, rate, and segmental anchoring  
in French

Pauline Welby and H el ene L oevenbruck

Institut de la Communication Parl ee, CNRS UMR 5009

Institut National Polytechnique de Grenoble, Universit e Stendhal

Grenoble, France

October 9, 2006

## *Abstract*

We examined tonal alignment and scaling patterns of the start and end points of the French late intonational rise in two read speech corpora. Our goals were twofold. First, we examined several competing hypotheses for characterizing the late rise: 1. an autosegmental-metrical account in which the rise is a bitonal pitch accent composed of a H\* pitch accent preceded by a L leading tone, 2. a segmental anchoring account where both the L and the H are temporally anchored with respect to points in segmental structure and maintain an invariant  $F_0$  excursion size, and 3. a holistic contour account in which the speaker seeks to achieve a specific slope. A rate manipulation paradigm was used, following Ladd et al. (1999).

We argue that the late rise is a LH\* pitch accent, although the L does not behave like a traditional leading tone to the following associated tone: there was no invariant rise time for any of our speakers. We also found no convincing evidence that speakers sought to achieve a constant slope, although we did notice speaker-specific tendencies for generally steeper or shallower rises. We were unable to identify a plausible segmental anchor for the low starting point of the late rise. In addition, we observed rate effects on  $F_0$  excursion sizes for three speakers across the two corpora. This finding, combined with earlier findings for French, calls into question the segmental anchoring assumption of an invariable  $F_0$  excursion size. These results thus challenge the generalizability of segmental anchoring to all spoken languages.

Our second goal was to more precisely define the alignment of the peak of the late rise, which earlier work had shown to be realized near the end of the last syllable of the accentual phrase (AP). Work on some languages had shown variability in

tonal alignment according to syllable structure, while work on at least one language had shown stability of alignment across syllables structures. For French CV and CVCobstruent syllables, the peak was consistently found at the end of the vowel. For CVCsonorant syllables, however, the position of the peak varied from the end of the vowel to the end of the sonorant coda. Thus, there was a fair amount of variability in the position of this peak, within and across speakers, in contrast to the very stable “segmental anchors” found for other languages. To account for this variable, yet rule-governed behavior, we propose the notion of an “anchorage,” that is, a region within which a tone can anchor. For the peak of the French late rise, the left boundary of this anchorage is near the end of the vowel of the last full syllable of the AP and the right boundary is the end of the AP.

## 1 Introduction

### 1.1 Segmental anchoring

In recent years, the “segmental anchoring” hypothesis has become an influential hypothesis on the nature of fundamental frequency ( $F_0$ ) rises and falls cross-linguistically. It makes a number of strong predictions. It claims that the temporal alignment of BOTH the start and the end of an  $F_0$  movement will be defined with respect to landmarks in the segmental string (Ladd et al. 2000: 2693, 2694):

<petit> [W]hen we look more closely at the differences of detail that have been documented between one syllable structure and another or between one language and another, we find that both the beginning and the end of the  $F_0$  movement can and must be precisely specified [with respect to segmental landmarks].

Studies report a remarkable precision of the start and end of  $F_0$  rises and falls with respect to these “anchor” points, sometimes on the order of a few milliseconds (for example, the low starting point of the Greek prenuclear rise is found approximately 5 ms before the onset of the accented syllable (Arvaniti & Ladd 1995; Arvaniti et al. 1998).

Proponents of segmental anchoring have also assumed that the stability of  $F_0$  movements will also extend to aspects of tonal scaling, a claim that seems to have been originally motivated by the findings of Arvaniti et al. (1998) for Greek. This is explicitly stated in Ladd et al. (1999: 1547):

<petit>

Strictly speaking, the segmental anchoring hypothesis makes no predictions about the size of the  $F_0$  excursion, yet the idea of independent L and H targets leads to the expectation that rate will have little effect on excursion size.

In this study, Ladd et al. (1999) did find some evidence for larger  $F_0$  excursions in slower rates (the effect was significant for one of the speakers in the study), but argued that the “[e]ffects of rate on  $F_0$  excursion size were . . . small and inconsistent” (p. 1543) and concluded that the “primary determinants of the shape of a pitch accent are the alignment and  $F_0$  level of the specific targets that make it up” (p. 1553).

This tonal scaling assumption is crucial to a rate manipulation paradigm used in segmental anchoring studies (e.g., Igarashi 2004). Ladd et al. (1999), who introduced the term “segmental anchoring,” reasoned that given stable temporal alignment and stable  $F_0$  excursion size,  $F_0$  rises in fast speech should have shorter rise times and steeper slopes. Figure 1 schematizes this prediction.<sup>1</sup> Slope comparisons in such studies allow researchers to evaluate autosegmental-metrical/ tonal targets approaches in which tones are primitives and pitch “movements” fall out from interpolation *vs.* holistic approaches where properties like slope or rise time may be taken as primitives. Figure 2 illustrates the predictions of a constant slope hypothesis: given shorter rise times (as in fast rate productions), we should observe lower peaks.

In the current study, we examine these predictions with respect to French, building on our recent work on tonal alignment in the language (Welby 2002, 2003, 2006), which has challenged some of the assumptions of segmental anchoring and of certain aspects of autosegmental-metrical (AM) phonology.

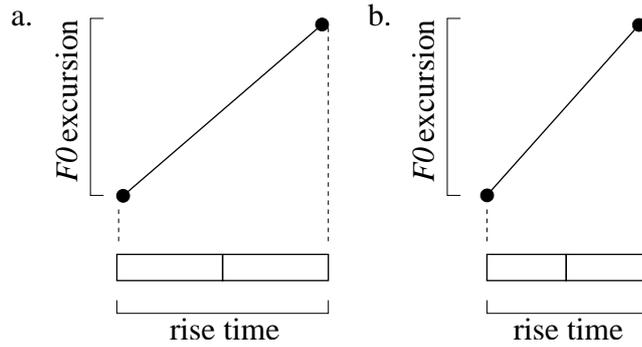


Figure 1: Schematic representation of the predictions of segmental anchoring of the effect of rate on slope. (a) normal rate: longer rise time, shallower slope (b) fast rate: shorter rise time, steeper slope.

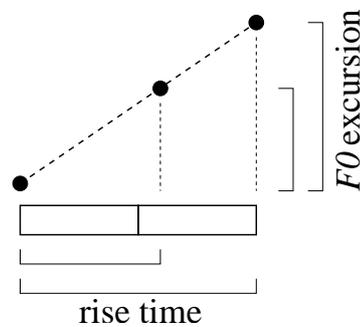


Figure 2: Schematic representation of the predictions of the constant slope hypothesis on the effect of rate on  $F_0$  excursion size. Normal rate: longer rise time, higher peak. Fast rate: shorter rise time, lower peak.

## 1.2 Basics of French intonational structure

All accounts of French intonation agree that the utterance can be divided into smaller units, called Rhythmic Unit (Di Cristo 1998), Accentual Phrase (Jun & Fougeron 2000), Intonation Group (Mertens 1987), Intonème intonatif mineur (Rossi 1985), Phonological Phrase (Post 2000, 2002), Prosodic Word (Martin 1979; Vaissière 1992), etc. by various researchers.<sup>2</sup>

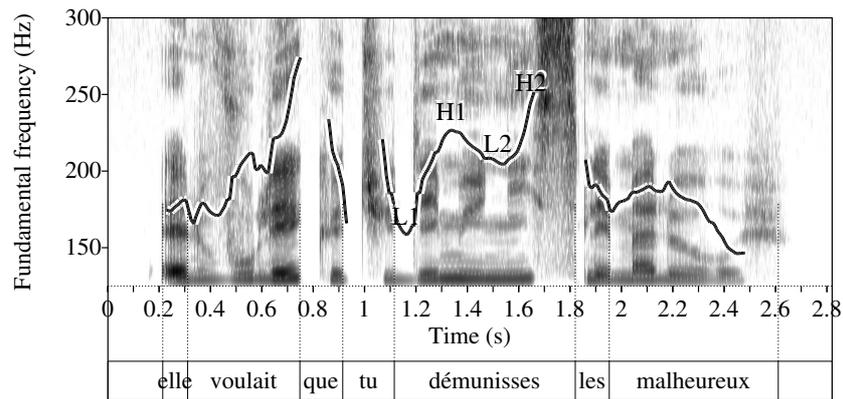
This accentual phrase (AP) is typically characterized by an  $F_0$  rise on the last

syllable of a phrase that is not utterance-final and an optional early rise near the beginning of the phrase. The late (final) rise is a marker of the “primary accent” and is accompanied by syllabic lengthening (Di Cristo 1976, 1985; Padeloup 1990; Jun & Fougeron 2000 *inter alia*). The early (initial) rise is a marker of the “secondary accent” and may be accompanied by strengthening of onset consonants (Padeloup 1990; Mertens et al. 2001 *inter alia*).

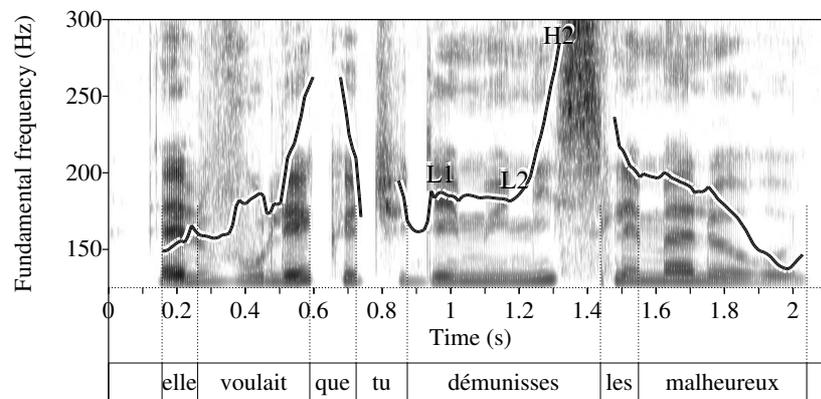
It is widely agreed upon that there is also a level of intonational phrasing above that of the AP, sometimes called the intonation phrase or IP (See Hirst & Di Cristo 1996; Di Cristo 1999, 2000; Jun & Fougeron 2000, 2002; Post 2000, 2002; Mertens 1987, 2002, for example). An utterance may be divided into one or more IPs, which may end in falls (for example, at the end of a declarative utterance) or rises (at the end of an interrogative utterance or to signal continuation). Evidence for this higher level of phrasing comes in part from differences in syllable lengthening: IP-final syllables show a greater degree of lengthening than do syllables that are simply AP-final, *ceteris paribus* (see Di Cristo 1976, 1985; Jun & Fougeron 2002, for example).

Figure 3, drawn from the current study, shows a minimal pair of utterances, differing in the presence or absence of an early rise in the target AP (*que tu démunisses* ‘that you impoverish’). The two pronunciations do not differ in meaning. Each utterance contains a single intonation phrase (IP), divided into three APs. This prosodic analysis is supported by an examination of the  $F_0$  curve as well as by the native speaker judgment of author HL. Our earlier research has shown that naive native speakers segment utterances into units corresponding to the predicted APs (Rolland & Loevenbruck 2002).

In the example, word boundaries are indicated by vertical dotted lines, and the



a.



b.

Figure 3: Spectrogram with a superimposed fundamental frequency curve for a minimal pair of utterances. Target AP *que tu démunisses* ‘that you impoverish’ (a) with an early rise and a late rise (LHLH pattern, Speaker 1, normal rate) and (b) with only the late rise (LLH pattern, Speaker 1, fast rate). The gloss is ‘She wanted you to impoverish the unfortunate ones.’

target AP is tonally transcribed (L(ow)1, H(igh)1 for the early rise; L2, H2 for the late rise).

In addition to the LHLH and LLH patterns illustrated, other patterns are possible. For example, speakers sometimes produce the pattern LHH (without L2), a pattern

characterized by a clear early rise followed by a high plateau extending to the end of the AP. In short APs, speakers often produce the pattern LH, a rise from L1 to H2 (without H1 or L2).<sup>3</sup> APs ending in a fall are also possible (see Martin 1987, 2004; Jun & Fougeron 2000, for example). For a discussion of these patterns, see Jun & Fougeron (2000) and Welby (2003), *inter alia*. In this study, we were concerned only with patterns containing the late rise (specifically LHLH and LLH, illustrated in the example).

### 1.3 Tonal alignment in French

Welby (2002, 2003, 2006) studied the tonal alignment of the starting and end points of the early and late rises in read speech corpora of standard Hexagonal French.<sup>4</sup> Following other researchers who have used evidence from tonal alignment as evidence for or against models of intonational phonology, Welby used her results to argue for a model of French intonation in which the early and late rises are structurally different. The LH of the late rise (L2H2) was argued to be a bitonal pitch accent, whose H is associated to the last full syllable of the AP. The LH of the early rise (L1H1), on the other hand, was claimed to be a bitonal phrase accent whose L is associated to the left edge of the first content word syllable of the AP, with an optional association to the left edge of an earlier syllable (for discussion of the alignment of L1 from an articulatory perspective, see D’Imperio et al. in press). The account is similar to that of Jun & Fougeron (2002), although there are important differences.

For each of these rises, only a single tone showed evidence of segmental anchoring: the H of the late rise (H2) was consistently realized near the end of the last syllable of the AP, and the L of the early rise (L1)<sup>5</sup> was consistently realized at the boundary

between a clitic function word and a content word (or at the beginning of the content word, when there was no preceding function word). The position of the low starting point of the late rise and the position of the high peak of the early rise (H1), however, varied considerably. The late L elbow (L2) was sometimes realized in the last syllable of the AP, sometimes in the penultimate syllable. The H of the early rise was sometimes realized in the first content word syllable of the AP, sometimes in the second, a variability that could not be explained by tonal crowding. Thus neither of these tones showed any evidence of being associated to a segmental landmark.

In addition, H1 and L2 did not behave like leading or trailing tones in the traditional autosegmental-metrical sense (there was no invariant temporal interval between the L and H tones). Welby (2006) nevertheless argued that the early rise and the late rise were both bitonal units, since the unassociated tone never appeared very far from the associated tone.

Based on the observation that for each of the rise types, only a single tone was associated to a segmental landmark, Welby argued against the strong segmental anchoring hypothesis for French, according to which both the start and end of an  $F_0$  rise or fall is anchored to segmental landmarks (Ladd et al. 1999). She questioned the universality of the strong segmental anchoring hypothesis for spoken languages.

Other researchers have suggested that some of the variability in the alignment of L2 (the L of the late rise) may convey pragmatic differences, with later elbows conveying affirmativeness or indignation. Post (2000) for example, discusses this possibility, but this aspect of alignment was not explored in the current study (see p. 126).

#### 1.4 Goals of the current study

The first main goal of the current study was to examine competing hypotheses for characterizing the late rise: 1. an AM account in which the rise is a bitonal pitch accent composed of a H\* pitch accent preceded by a L leading tone, 2. a segmental anchoring account where both the L and the H are temporally anchored with respect to points in segmental structure and maintain an invariant  $F_0$  excursion size, and 3. a holistic contour account in which the speaker seeks to achieve a specific slope.

The second goal was to determine the precise landmark to which the H of the late rise (H2) is aligned, a question left open in our earlier studies. Welby (2003) concluded (rather vaguely) that the H of the late rise was realized “close to the end” of the last syllable of the AP (p. 113). But the interval between H2 and the end of the syllable ranged from just beyond the end of the syllable (in 3% of cases) to over 100 ms before the end of the syllable. We suspected that this variability might be due to differences in syllable structure of the target syllables, a factor that was not manipulated in earlier studies, but is known to affect alignment in other languages. For example, studies of English, Dutch, and Neapolitan Italian have all shown that the presence of sonorant or voiced coda consonants pulls accent peaks to the right (van Santen & Hirschberg 1994, Rietveld & Gussenhoven 1995, and D’Imperio 2000, respectively). For Mandarin, however, Xu (1998) found no influence of syllable structure (CV vs. CVNasal) on tonal alignment, a stability that Xu interpreted as evidence that “the syllable is the proper domain for tone implementation” (p. 179).

Our hypotheses with respect to the alignment of H2 and L2 are detailed below.

#### 1.4.1 Syllable structure and the alignment of H2

We examined whether the alignment of H2 differs according to whether the associated syllable (the last full syllable) is open (CV) as in *démuniât* [de.my.ni], ‘impoverish (3rd pers. sing. imperfect subjunctive)’, closed by a sonorant (CVCson), as in *démuniâmes* [de.my.nim] ‘impoverish (1st pers. pl. simple past)’ or closed by an obstruent (e.g., a fricative, CVCfr), as in *démunisses* [de.my.nis] ‘impoverish (2nd pers. sing. simple past)’. This allowed us to examine three potential “segmental anchors” for H2, given in the hypotheses in (1).

- (1) a. *Hypothesis H2-A*: H2 is anchored with respect to the end of the vowel of the accentual-phrase final syllable.
- b. *Hypothesis H2-B*: H2 is anchored with respect to the end of the sonorant portion of the rhyme (“sonorant rhyme”) of the accentual-phrase final syllable.<sup>6</sup>
- c. *Hypothesis H2-C*: H2 is anchored with respect to the end of the accentual-phrase final syllable.

These hypotheses make very different predictions about the alignment of H2 (the H of the late rise).

If Hypothesis H2-A is correct, H2 should be stably realized with respect to the end of the vowel, regardless of the syllable structure of the final syllable (open or closed), as shown in Figure 4.<sup>7</sup>

If Hypothesis H2-B is correct, H2 should be realized at the end of the sonorant portion of the last syllable, as shown in Figure 5.

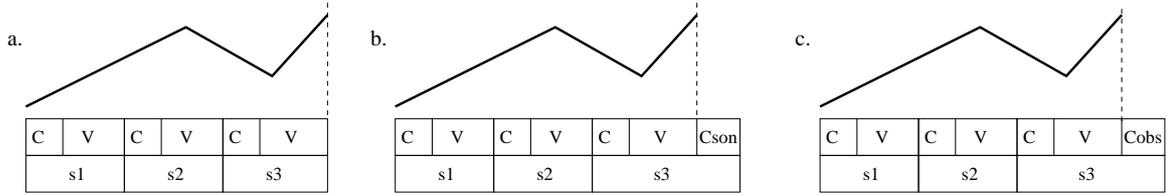


Figure 4: Stylized  $F_0$  curve illustrating the alignment predictions of Hypothesis H2-A (anchor point: end of vowel). The pattern is LHLH.

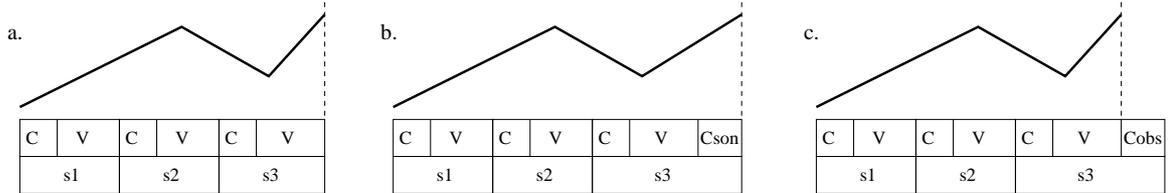


Figure 5: Alignment predictions of Hypothesis H2-B (anchor point: end of sonorant rhyme).

Taken alone, the temporal alignment shown in Figure 5 would not provide us with enough information to distinguish between Hypotheses H2-B and H2-C. For example, the  $F_0$  peak will not be realized at the very end of the syllable [nis] in *démunisses* [de.my.nis]: the syllable ends in a voiceless obstruent, so there is no vocal fold vibration and  $F_0$  is therefore not defined.<sup>8</sup> The speaker’s intended target, however, may be the end of the syllable or some point beyond the vowel. Comparing tonal scaling in cases like *démunîmes* vs. *démunisses* might allow us to distinguish between these two hypotheses. If Hypothesis H2-C is correct and the speaker’s target is indeed the end of the syllable or some point beyond the vowel and the target is undershot due to physical constraints imposed by the structure of the rhyme, we might find that peaks in syllables closed by an obstruent consonant (e.g., [nis] of *démunisses*) are systematically lower than those that end in sonorant segments (e.g., [ni] and [nim]).<sup>9</sup> This situation is schematized in Figure 6.

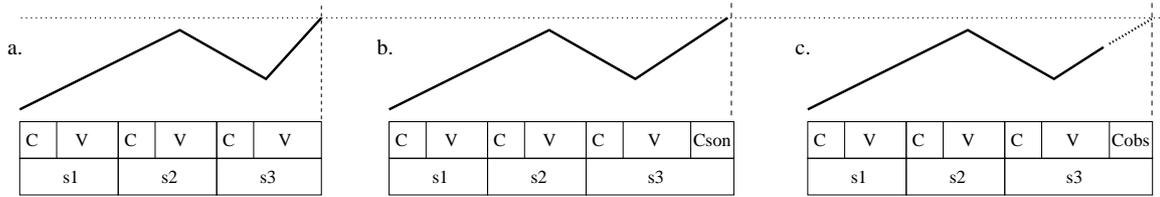


Figure 6: Alignment and scaling predictions of Hypothesis H2-C (anchor point: end of syllable). The dotted line represents the fact that the target is not achieved.

If one of the three hypotheses is supported, we will also be interested in the degree of precision of the alignment with respect to the segmental anchor.

#### 1.4.2 Alignment of L2

A second set of hypotheses concerns the alignment of the low starting point of the late rise (L2). As discussed earlier (§ 1.3), our previous work showed that the alignment of L2 varied considerably. It was therefore not possible to suggest plausible specific segmental anchors *a priori*.

We therefore test slightly modified versions of the hypotheses discussed in Welby (2006). These hypotheses are given in (2). Note that the syllable structure manipulations discussed above with respect to the alignment of H2 are not predicted to influence the alignment of L2.

- (2) a. *Hypothesis L2-A*: L2 is associated to either the ultimate or the penultimate syllable of the accentual phrase.
- b. *Hypothesis L2-B*: L2 is a leading tone of the late rise, in the traditional autosegmental-metrical understanding of “leading tone.”
- c. *Hypothesis L2-C*: L2 and H2 form a bitonal unit, despite the absence of a fairly constant interval between the two.

Hypothesis L2-A refers to the AM concept of association. This concept is not identical to that of anchoring, although both involve alignment relationships between tonal targets and the segmental level. Anchoring predicts precise alignment with respect to segmental landmarks like consonants and vowels. In the AM framework, alignment is used as evidence for phonological association either to a syllable or to the edge of a prosodic unit. Relative alignment is often used to distinguish one tone from another. For example, the English L+H\* and L\*+H pitch accents differ in relative alignment: in the first case, the H is realized within or just beyond the associated syllable; in the second case the L is realized within this syllable (Pierrehumbert 1980, Beckman & Hirschberg 1994). Associated syllables are typically metrically strong, and often perceived as accented. While the strong segmental anchoring hypothesis holds that all turning points (H and L tones) will be stably anchored with respect to segmental landmarks, in the AM framework, not all tones are associated, and so some turning points will not be aligned with respect to a syllable or a prosodic edge. One aspect that the two concepts have in common is that a tone can appear beyond the borders of its associated syllable (AM framework) or landmark (segmental anchoring). Segmental anchoring differs from AM association in that anchoring predicts a high degree of stability of alignment, while association makes no specific claims about alignment stability.

If Hypothesis L2-A is correct, we expect to find that L2 is aligned with respect to a given syllable. If we find evidence for this hypothesis, it may be possible to identify potential segmental anchors.

If Hypothesis L2-B is correct, we expect that L2 will be realized at an “fairly invariant” interval before H2 (Pierrehumbert & Beckman 1988: 123; see also Pierre-

humbert 1980). Rise times should thus vary very little within a given speaker.

Even if no invariant interval between L2 and H2 is found, if both tones are consistently realized near the end of the AP, this will be evidence for Hypothesis L2-C. If the hypothesis is correct, we expect the L2 tone to be consistently realized near the end of the AP (and not, for example, near the middle).

The results for the alignment of L2 will have a direct bearing on our conclusions about the strong segmental anchoring hypothesis, which holds that both the start and the end of an  $F_0$  rise should be anchored to segmental structure — if we do not find evidence for Hypothesis L2-A, this will be evidence against the strong segmental anchoring hypothesis for French and against the idea of the hypothesis as a spoken language universal.

#### *1.4.3 Hypotheses concerning relationship between L2 and H2*

In order to test the claims of segmental anchoring discussed above, we will also examine a number of hypotheses concerning the relationship between L2 and H2, notably whether a constant rise time, constant  $F_0$  excursion size, and or constant slope are found within a given rate and under changes in rate.

## *2 Methods*

In the experiment, native speakers first read a list of sentences at two speaking rates, then a list of paragraph-long stories also at two speaking rates. We measured tonal alignment and scaling of the start (L2) and end (H2) of the late rise to test the hypotheses discussed above.

## 2.1 Materials

### 2.1.1 Sentence Corpus

We designed a corpus of 24 sentences, containing six sets of minimal quadruplets composed of different forms of a given verb. An example set is given in (3), with the phonemic transcription of the target verb given in parentheses.

- (3) a. Elles voulaient que l'on démunît les malheureux. /de.my.ni/  
b. Elles croyaient que nous démunîmes les malheureux. /de.my.nim/  
c. Elles croyaient que vous démunîtes les malheureux. /de.my.nit/  
d. Elles voulaient que tu démunisses les malheureux. /de.my.nis/  
a./d. 'She wanted me/you to impoverish the unfortunate ones.'  
b./c. 'She believed that we/you impoverished the unfortunate ones.'

Each target verb appeared sentence-medially, in a dependent clause introduced by a two-syllable matrix verb that provided the required context for each verb form. The dependent clause was always of the (linear) structure: complementizer *que*, one-syllable pronominal subject, three-syllable target verb, definite article *le*, *la*, or *les*, three-syllable noun. The matrix verbs and pronouns were varied in order to discourage speakers from reading the sentences with an intonation consistent with contrastive narrow focus on the target. The target verb and preceding function words were expected to be produced as a single AP.

Within each quadruplet, the target differed only in the structure of the last syllable: it was either open (*démunît*), closed by the sonorant /m/ (*démunîmes*), or closed by a voiceless obstruent, either the stop /t/ (*démunîtes*), or the fricative /s/ (*démunisses*).<sup>10</sup> The target verbs were either in the simple past (*démunîmes*,

*démunîtes*) or imperfect subjunctive (*démunît, démunisses*).<sup>11</sup> Target verbs with three syllables (i.e., relatively long targets) were chosen to maximize the possibility that a distinct late rise would be produced (i.e., LLH or LHLH, as shown in Figure 3). Target verbs had the structure  $C_1V_1.C_2V_2.C_3V_3(C_4)$ , where  $C_2$  and  $C_3$  were always sonorants and  $C_1$  was either a sonorant or a voiced stop (/d/). Three quadruplets contained target verbs in *-ir* (e.g., *démunir* ‘to impoverish’); three contained target verbs in *-er* (*mouliner* ‘to mill, grind’). Half of the quadruplets thus contained the vowel /i/ in the target syllable (e.g., *démunît*, /de.my.ni/) and half the vowel /a/ (e.g., *moulinât*, /mu.li.na/).<sup>12</sup>

### 2.1.2 Paragraph Corpus

We designed a second corpus with six short paragraphs containing triplets of target words. An example paragraph is given in (4), with the target words underlined.

- (4) La cuisine de Marie-Noëlle regorge de victuailles. Un salami qui a été rapporté de Parme suinte d’une graisse très prometteuse. Sur la desserte, un immense compotier déborde de fruits de toutes sortes. Les vitamines qu’ils recèlent naturellement sont un prétexte pour les gourmands. Une pièce-montée en choux et nougatine trône sur le buffet. La pyramide qui dégouline de caramel fera sûrement des amateurs.

‘Marie-Noëlle’s kitchen is overflowing with provisions. A salami that was brought back from Parma is dripping with very promising grease. On the side table, a giant bowl is brimming with all kinds of fruit. The vitamins that they naturally contain are just an excuse for people who like to eat. A layer cake with puff pastries and nougatine

is prominently displayed on the cabinet. The pyramid, dripping with caramel, will surely be appreciated.’

As with the verb quadruplets of the sentence corpus, the triplets of the paragraph corpus differed in the syllabic structure of the last syllable: it was either open (CV, as in *salami* /sa.la.mi/), closed with a sonorant (CVCson, as in *vitamine* /vi.ta.min/), or closed with a voiced or voiceless obstruent (CVCobs, as in *pyramid* /pi.ra.mid/). All target words were three-syllable nouns or substantive adjectives with the syllabic structure  $C_1V_1.C_2V_2.C_3V_3(C_4)$ , where  $C_3$  was always a sonorant, and to the extent possible,  $C_1$  and  $C_2$  were sonorant or at least voiced. Within a triplet, the sequence  $V_2.C_3V_3$  at the end of the target word (/a.mi/ in the triplet in (4)) was identical. In addition, we sought to include only fairly common words (as judged from word frequency scores in the LEXIQUE database).

Each target word appeared sentence-initially, preceded by a form of the definite article (*le, la, les*). The resulting phrase was followed by relative clause beginning with a relative pronoun (*qui, qu’il, qu’elle, qu’ils*), which was followed by the verb phrase of the main clause.<sup>13</sup>

The material following the target was on average 15.1 syllables long. Most post-target regions contained between 13 and 15 syllables. One contained 19 syllables. Each paragraph contained between three and five non-critical sentences, necessary to make semantically coherent paragraphs. Note that unlike the minimal verb quadruplets in the sentence corpus, the triplets in the paragraph corpus were never exact minimal sets. Potential triplets were identified in LEXIQUE with a Perl script written for the purpose, and no exact minimal triplets were found.<sup>14</sup>

### 2.1.3 Points of comparison between the two corpora

While we did not intend to directly compare the two, using two corpora provided us with a number of points of comparison. The targets in the paragraph corpus contained a variety of vowels (/a ε i o/), sonorant coda consonants (/m n l/) and voiced and voiceless obstruent coda consonants (/d z t s/). This variety is relevant to the generalizability of the patterns we find. We can be confident, for example, that the patterns are not specific to intrinsic differences in vowel length (see Benguerel 1971; O’Shaughnessy 1981 on the duration of French vowels). In addition, the inclusion of both voiced and voiceless obstruent codas was expected to allow us to examine whether there are voicing-dependent differences in alignment within obstruents. The two corpora also differed in that the targets in the sentence corpus were verbs in medial position, while the targets in the paragraph corpus were nouns and adjectives in initial position. In our earlier alignment work, we used only noun phrases.

The two corpora also differed in the phonetic context following the target. In the sentence corpus, the target was always followed by a form of the definite article (*le, la, les*) and so the sonorant /l/. In the paragraph corpus, the target was always followed by a relative pronoun (*qui, qu’il, qu’elle, or qu’ils*) and so the voiceless obstruent /k/. Apparent alignment patterns observed in recent pilot recordings for other projects had suggested that the nature of the following segment might play a role in the alignment of H2, at least for some speakers. In particular, we noticed for two speakers reading two separate corpora (one a reiterant speech corpus with only [ma] syllables, one a corpus with sentences designed to maximize sonorant consonants in target sequences), that H2 peaks appeared quite late, not in the last syllable of the phrase, but at the

end of the onset consonant of the first syllable of the following AP. Note that “peak delay,” in which a peak is aligned after the syllable with which it is associated (as is in English (Pierrehumbert & Steele 1989; Ladd et al. 1999), Greek (Arvaniti et al. 1998), German (Atterer & Ladd 2004), and many other languages) is not typically found in French.

Finally, one could argue that reading a list of unconnected sentences is a different (perhaps less natural) task than reading semantically coherent paragraphs.

For these reasons, we were interested in whether the patterns we found for one corpus would hold for the other.

## 2.2 *Participants*

Six native speakers of standard Hexagonal French participated in the experiment. They were all women from Paris or the surrounding region. All participants in the experiment were students or researchers at universities in Grenoble (Université Stendhal, Institut National Polytechnique de Grenoble or École Nationale Supérieure d’Électricité et de Radioélectricité de Grenoble (ENSERG)). The speakers ranged in age from 22 to 36, with an average of 26.8. Three of the speakers had doctoral degrees; the other two had four or five years of university education beyond the *baccalauréat*. Five of the speakers were naive to the hypotheses being tested. Speaker 6 is author HL. Two of the speakers are fluent speakers of a second language: Speaker 3 learned German starting at age 11 and attended a French-German *lycée*. Speaker 6 has near-native fluency in English. Neither of the two speakers uses her second language on a daily basis.

Five of the six speakers reported no speech or hearing difficulties. Speaker 3

reported an occasional problem with tinnitus, although she experienced no difficulty during the recording.

Speakers (except author HL) were paid €10 for their participation. The recording lasted about a half hour for each speaker.

### 2.3 Procedures

The 24 sentences of the sentence corpus were randomized. Each participant first read the sentence corpus aloud two times, first in a self-selected normal rate and then in a fast rate. She then read the paragraph corpus, first in the normal rate and then in the fast rate. The experiment instructions were designed to minimize the possibility that participants would change reading styles from one corpus to the next. In particular, we were concerned that speakers might give a dramatic reading of the paragraphs. We therefore never called the paragraphs “stories” (*histoires*) and specifically instructed participants to “read these paragraphs just like you read the sentences” (*lisez ces paragraphes comme vous avez lu les phrases*).

Speakers were recorded at 44.1 kHz onto digital audio tape (DAT), using a Shure SM10A headworn microphone in a sound attenuated chamber at the Institut de la Communication Parlée. The data were transferred to a computer, then downsampled to 22.05 kHz.

The soundfiles were segmented and each utterance saved as a separate file.  $F_0$  curves and spectrograms were created using Praat speech analysis software (Boersma & Weenink 2005).

Word, syllable, and segment boundaries were tagged for each target word and the syllables immediately preceding and following the target, using waveforms and

spectrograms to guide the segmentation. The beginning and end of each utterance were also tagged. Segment boundaries were generally easy to identify, since the critical regions contained many nasals or liquids followed by vowels, and there were thus abrupt changes in intensity. There were a few cases in which segment boundaries were less clear. For example, the [tl] or [sl] sequences in groups like *démunîtes les* (/de.my.nit.le/) or *démunisses les* (/de.my.nis.le/) often gave rise to a lateral fricative, obscuring the boundary between the end of the /t/ or /s/ and the beginning of the /l/. In this case, we set the criterion for the beginning of the /l/ as the beginning of voicing, visible in the zoomed-in waveform.

Praat scripts were written to semi-automate the segmentation and labelling process. These prompted the user to click on the location of a desired tag, automatically inserted boundaries and tags, and saved the results to a Praat TextGrid file.

The intonational features tagged included H1 (peak of early rise), H2 (peak of late rise), and L2 (start of the late rise) (see Figure 3).

Since each of the H tones was typically preceded and followed by a L tone, the location of a H tone was defined in the Praat labelling scripts as the time of local  $F_0$  maximum<sup>15</sup> and automatically detected. The position of the tags was hand-corrected if necessary. For example, in some cases, H2 was realized as a short, fairly level plateau (defined as a series of  $F_0$  values differing by 3 Hz or less). In these cases, H2 was moved by hand to the beginning of the plateau, if necessary. This decision was motivated by two considerations. First, from a production point of view, the beginning of the plateau is the turning point. Second, according to one of our hypotheses, speakers seek to maintain a constant slope; taking the beginning of the plateau allowed us to obtain an accurate slope measure (slopes would have been shallower if H2 had

been defined at the midpoint or endpoint of a plateau). We are aware that different researchers make different decisions with respect to this measurement point (see, for example, D’Imperio 2000, who takes the end of the plateau, and Atterer & Ladd 2004, who take the beginning of a “slightly rising” plateau).

As many researchers have noted, reliably hand-labelling the starting points of rises (L tones) is often problematic, since they do not always correspond to  $F_0$  minima (Pierrehumbert & Beckman 1988; Xu 1998; D’Imperio 2000; Frota 2002; Welby 2002, 2003, 2006). Of his Mandarin data, Xu (1998: 196, 197) writes:

<petit> As a first approximation, the onset of the rise may be defined as the  $F_0$  minimum right before the rise. However, in some cases... the portion of the contour before the apparent rise is virtually flat and sometimes even rises slightly. In such cases, the  $F_0$  minima seem to be too far away from where the real rise is...

We find the same situation for some of our French data with target APs with a LLH pattern. In these cases, objectively determining the position of L2 by hand would have been nearly impossible. Therefore, as in our earlier work, we used a line-fitting procedure to automatically calculate the position of L2.<sup>16</sup> Due to space considerations, we do not detail the procedure here, but refer the reader to Welby (2003, 2006) for a detailed explanation of the procedure (see also D’Imperio 2000).

Time values for all tags, durations, and  $F_0$  values were automatically extracted from the label files by Praat scripts.

### 3 Results

#### 3.1 Sentence corpus

##### 3.1.1 Rate manipulation

Average speaking rate in syllables/second was calculated for each speaker in each of the two speaking rates, according to the following procedure. Each sentence had a base syllable count of 12 syllables. We listened to each utterance and adjusted the base count based on the actual pronunciation, which sometimes included pronunciation of word-final schwas. This syllable count was divided by the utterance duration (including pauses) to obtain a rate measurement for each utterance. Rate was similarly calculated for each target word, using a base count of three syllables. This count was adjusted for the rare cases in which the speaker produced a schwa at the end of the target word (e.g., *laminâmes* [la.mi.na.mə]).

Results of ANOVA confirmed that all speakers successfully raised their speaking rates from the normal to the fast speaking rate condition, as measured in target words as well across the entire sentence. There was a large effect of rate (for sentences:  $F(1,135) = 889.57, p < 0.001$ , for target words:  $F(1,135) = 288.93, p < 0.001$ ). Figure 7 shows rates for target words for each speaker.<sup>17</sup>

As seen in the figure, speaking rate and the magnitude of the change from one rate to the other varied across speakers. For example, Speaker 2 had the slowest rate in the normal condition, 4.63 syllables/s, but she had the largest percentage change, raising her rate 46% to 6.45 syllable/s in the fast rate. Speaker 3, had the fastest rate in the normal condition, 6.61 syllables/s, but raised her rate only 12%, to 7.36 syllables/s in the fast condition. The effect of Speaker was significant (for sentences:

$F(5,135) = 114.89, p < 0.001$ ; for target words:  $F(5,135) = 16.19, p < 0.001$ ), as was the Rate  $\times$  Speaker interaction (for sentences:  $F(5,135) = 12.28, p < 0.001$ ; for target words:  $F(5,135) = 10.79, p < 0.001$ ).

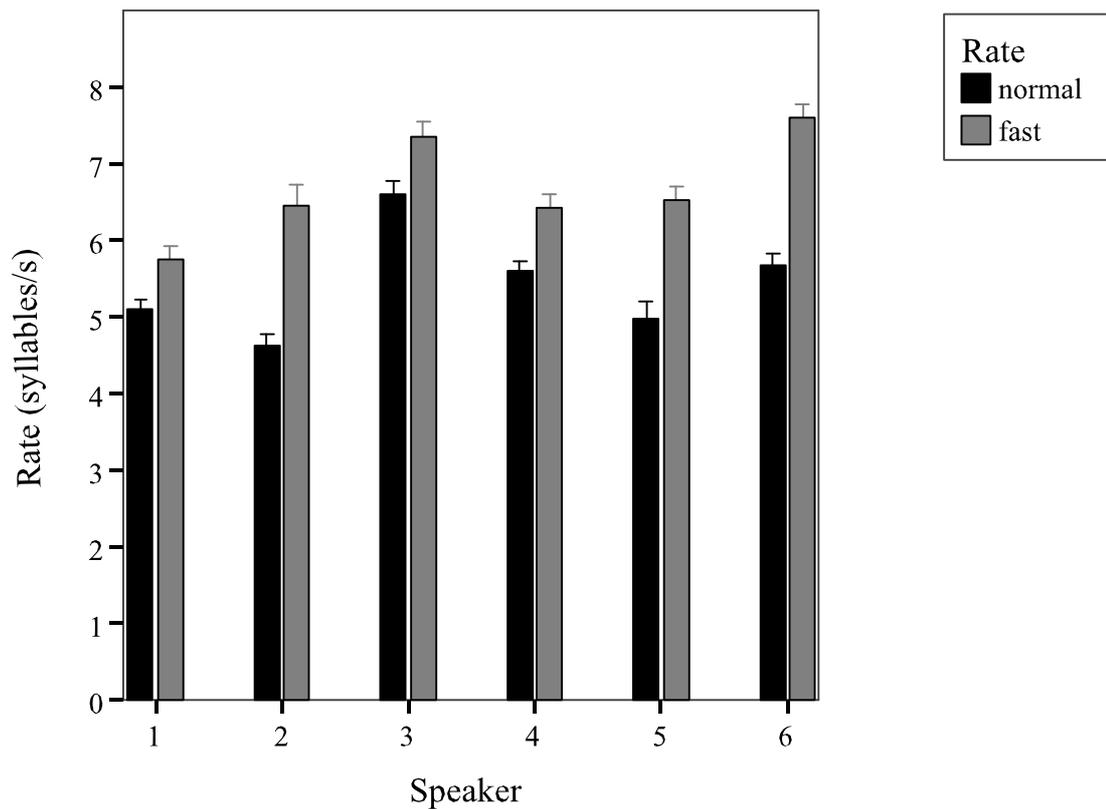


Figure 7: Speaking rates for the target words in the sentence corpus. Bars show standard error of the mean.

### 3.1.2 Items available for further analyses

Not all of the 288 utterances of the sentence corpus (24 sentences  $\times$  2 rates  $\times$  6 speakers) were available for the analyses. One item was excluded because it contained a hesitation, and another was unusable due to a data transfer problem. In addition,

items without a clear late rise in the target AP were excluded.<sup>18</sup>

A total of 110 items (or 38% of the total available) without late rises were excluded. We were somewhat surprised by the relatively low rate of LHLH patterns produced in the target APs, given reports in the literature and what we had observed in our own previous studies (Rolland & Løevenbrück 2002; Welby 2002, 2003). Given a corpus with three-syllable long content words in target position, we expected speakers to produce the vast majority of target APs with the two-rise LHLH pattern.<sup>19</sup> Another unexpected pattern in the current data was that in most of the non-LHLH cases, the target AP was produced with a LHH pattern, which was fairly uncommon in our other studies. Of the remaining 176 items with a clear late rise in the target AP, only 15 contained the pattern LLH. We therefore decided to focus our investigation on items with the pattern LHLH in the target AP. We chose not to examine the two late rise patterns (LLH and LHLH) together since there is some (limited) evidence that the alignment of the turning points of the late rise may be affected by tonal crowding (see Welby 2006). Of the remaining 161 items with the LHLH in the target AP, very few – only seven – were produced by Speaker 4. We therefore excluded Speaker 4’s data from these analyses. Finally, of the remaining 154 items, an additional seven items that were produced with a clear final schwa in the target AP were also excluded, since the presence of a schwa changes the syllable structure of the final syllable of the target AP.<sup>20</sup>

### *3.1.3 Tonal scaling, rise time, slope analyses*

For each of the remaining 147 items, we measured RISE TIME (temporal interval between L2 and H2),  $F_0$  EXCURSION (difference in Hz between L2 and H2), and slope

Speaker	rise time		$F_0$ exc.		slope		N (pairs)
	(ms)		(Hz)		(Hz/ms)		
	normal	fast	normal	fast	normal	fast	
1	119	114	52	71*	0.47	0.66**	20
	(8)	(5)	(3)	(4)	(0.04)	(0.06)	
2	94	87	28	24	0.30	0.36	8
	(14)	(19)	(7)	(5)	(0.06)	(0.08)	
3	80	39**	25	12**	0.31	0.34	5
	(8)	(9)	(4)	(3)	(0.09)	(0.04)	
5	97	74	25	18	0.25	0.22	5
	(12)	(2)	(11)	(2)	(0.02)	(0.07)	
6	98	88	30	29	0.31	0.35	13
	(4)	(5)	(2)	(3)	(0.02)	(0.04)	

Table 1: Mean  $F_0$  excursion size, rise time, and slope by speaker, separated by rate, sentence corpus (paired comparisons). Standard error is shown in parentheses.

\* $p < 0.001$ , \*\* $p < 0.01$ .

( $F_0$  excursion/ rise time). These analyses are relevant to the segmental anchoring hypothesis, the constant slope hypothesis and the hypothesis of L2 as a leading tone (Hypothesis L2-B).

Means were calculated for each of the five speakers (recall that the data from Speaker 4, who produced only seven LHLH patterns for this corpus, are not included) at each rate, and paired t-tests were performed to examine whether there were systematic differences across rates. Means are given in Table 1.

The statistical analyses were somewhat complicated by the unexpectedly low rate of LHLH patterns in target APs. Comparison of the normal *vs.* fast rate productions was not possible for all pairs. For example, Speaker 1's normal and fast productions of the target word *démunisses* (Figure 3) could not be used in the paired comparisons since the target AP was not produced with the LHLH pattern in the fast rate. For some participants, this left very few pairs for the paired t-tests. The number of pairs available for these analyses is indicated in the last column of Table 1.

#### *Rise time*

The segmental anchoring hypothesis predicts a clear effect of rate: the shorter segmental durations in the fast rate should lead to shorter rise times. Although rise times were numerically longer in the normal rate than in the fast rate for all speakers, results of paired t-tests failed to show consistently significant differences. For only one speaker was there a significant rate effect on rise time. Speaker 3's late rises were significantly longer in the normal rate than in the fast rate ( $t = 5.30, df = 4, p < 0.01$ ). We acknowledge that the low Ns for many of the speakers may increase the chance of a type II error (i.e., failing to get a significant result when the hypothesis is in fact true) but we note that no effect of rise time was found for Speakers 1 and 6,

for whom the most pairs were available for comparison. In addition, we note that in addition to the paired t-tests, we also performed t-tests on all the available data (the 147 items), treating the data for each rate within a given speaker as an independent sample, and the same pattern emerged: rise time was significantly different between the two rates only for Speaker 3 (rise time, normal rate: 83 ms, fast rate: 43 ms;  $t = 3.26, df = 17, p < 0.01$ ).

We also note that although most speakers did not show a significant difference in rise time between rates, rise times varied considerably within a single speaker (as shown by the standard errors in both the paired samples t-tests (Table 1) and the independent samples t-tests performed on all 147 items). This variability is shown in the scatterplot of all 147 items in Figure 8. It is clear from the plot that there was no “fairly invariant” interval between L2 and H2 for any speaker; for example, Speaker 1, for whom we have 44 tokens, had rise times ranging from a minimum of 80 ms to a maximum of 207 ms. This lack of invariance is evidence against Hypothesis L2-B, according to which L2 is a leading tone of H2.

#### *F<sub>0</sub> excursion*

We performed another series of paired t-tests to test the effect of rate on  $F_0$  excursion size. These analyses and all subsequent analyses involving  $F_0$  were performed first using hertz as the unit of measure, then ERBs (equivalent-rectangular-bandwidth, Hermes & van Gestel 1991), then semitones (’t Hart et al. 1990).<sup>21</sup> These three units of measure were used to allow comparison with results in the literature, since different researchers have used different units. In the tables, values are given in hertz, in part to make it possible to compare the slopes found in the current study with those found in Fougeron & Jun (1998) and Welby (2006).

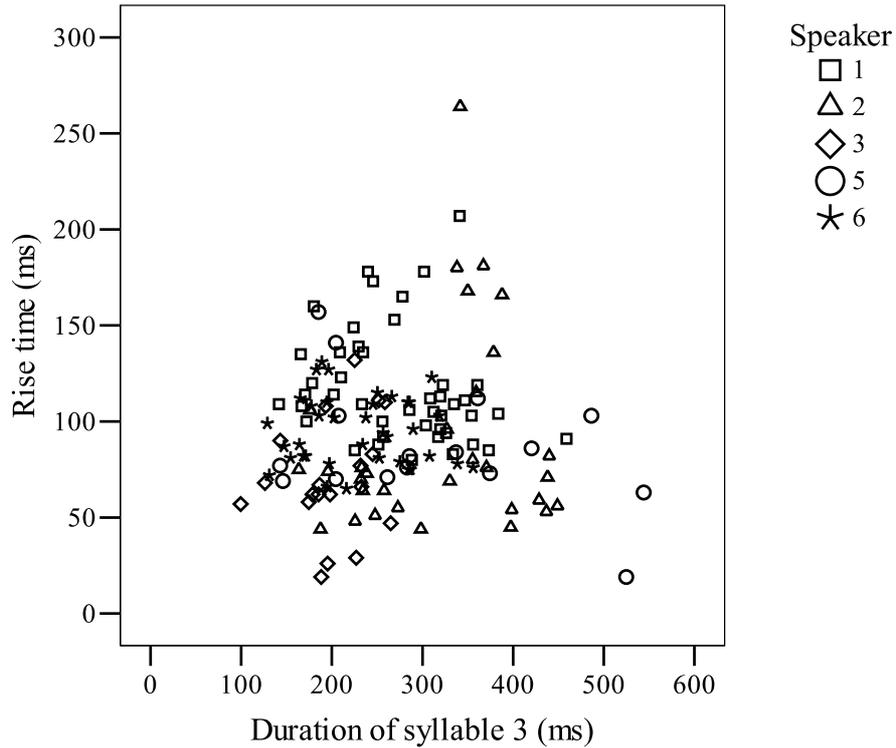


Figure 8: Rise time of the late rise as a function of the duration of the last syllable in the AP (syllable 3), sentence corpus.

Advocates of the segmental anchoring hypothesis have argued that rate should have little effect on  $F_0$  excursion size (see Figure 1). By contrast, a constant slope hypothesis predicts that  $F_0$  excursions should be larger in normal speaking rates than in fast rates, since the speaker has more time to reach a peak.

$F_0$  excursion for Speaker 1 was significantly larger in the fast rate than in the normal rate ( $t = 5.56, df = 19, p < 0.001$ ). This was also true for the ERB and semitone analyses. The direction of this difference was the reverse of that expected by a constant slope hypothesis. Such a difference, in which fast rate rises have larger excursion sizes than normal rate rises was not observed by Fougeron & Jun (1998) or Welby (2006) for any of the speakers in those studies. For Speaker 3,  $F_0$  excursion

was significantly greater in the normal rate than in the fast rate ( $t = 5.582, df = 4, p < 0.01$ ). The rate-dependent difference in  $F_0$  excursion was also significant in the independent t-test analyses for Speaker 1 (normal rate: 71 Hz, fast rate: 50 Hz,  $t = 4.348, df = 42, p < 0.001$ ). For the other speakers, no significant differences were found across rates.

### *Slope*

We examined whether speakers seek to maintain a constant slope in two ways. We first performed a series of paired t-tests to test whether slope differed significantly across rates for each speaker. Slope was calculated by dividing  $F_0$  excursion by rise time for each item examined.

For one speaker, Speaker 1, there was an effect of rate on slope: fast rate late rises had steeper slopes than normal rate rises ( $t = 3.78, df = 19, p < 0.01$ ). The same result was found for the ERB and semitone analysis, as well as the independent t-test analyses (normal rate: 0.45 Hz/ms, fast rate: 0.64 Hz/ms;  $t = 2.756, df = 42, p < 0.01$ ). The difference in rate for this speaker is in line with the direction predicted by the segmental anchoring hypothesis. For the remaining speakers, the slope did not vary significantly, contrary to the hypothesis.

To further examine whether speakers sought to achieve a constant slope, we performed correlation analyses to examine whether rise time and  $F_0$  excursion size were positively correlated, as predicted by a constant slope hypothesis (as described in Ladd et al. 1999).

The results, presented in Table 2, show that it was not generally the case that rise time was correlated with  $F_0$  excursion size. The correlation was significant for two speakers, Speakers 3 and 5. This pattern held for the semitone and ERB analyses.

Speaker	r	p	N
1	-0.091	0.279 ns	44
2	0.258	0.08 ns	31
3	0.434	0.03*	19
5	0.706	0.001*	16
6	0.016	0.464 ns	37

Table 2: Correlations between rise time and  $F_0$  excursion (measured in Hz) for the sentence corpus, both rates combined.

We note that while there was considerable variability in the slope of the late rise (as shown by the size of the standard error), these mean values did seem to be fairly stable across rates for some speakers. We have observed the same pattern in earlier work (Welby 2006).

*H2 scaling in CVCobs vs. CV and CVCson syllables*

According to Hypothesis H2-C, H2 is aligned with respect to the end of the last syllable of the AP. We reasoned that in this case, H2 would be undershot in syllables with obstruent rhymes (CVCobs, i.e., CVCst and CVCfr). To test the hypothesis, we compared the scaling of H2 in syllables with sonorant rhymes (CV and CVCson) and those with obstruent rhymes (CVCst and CVCfr), performing independent samples t-tests for each of the five speakers (excluding Speaker 4). Two speakers showed a significant effect of rhyme type. For Speaker 1, H2 was significantly higher in syllables with obstruent rhymes than in syllables with sonorant rhymes (266 Hz *vs.* 250 Hz;  $t = 3.14, df = 42, p < 0.01$ ). This difference is in the opposite direction of the hypothesis; H2 was clearly not undershot in CVCobs syllables for this speaker. For

Speaker 2, H2 was significantly lower in syllables with obstruent rhymes (204 Hz *vs.* 225 Hz;  $t = 3.68$ ,  $df = 29$ ,  $p < 0.01$ ), a difference in line with the hypothesis.

#### 3.1.4 L2 alignment

We next examined the alignment of L2, the low starting point of the late rise, first plotting the results by speaker and rate. Surprisingly, the plots (not shown here, due to space constraints) indicated that L2 was always realized in the last syllable of the AP, never in the penultimate syllable. This was different from the pattern we had observed in our earlier studies, where L2 was often realized in the last syllable but sometimes in the penultimate syllable. In addition, the plots showed a potential effect of rate for some speakers; we conducted a series of paired t-tests to examine that possibility. We chose to examine the alignment of L2 with respect to two segmental landmarks: the beginning of the last (third) syllable of the AP (begS3toL2) and the beginning of the vowel of this syllable (begV3toL2). Our choice of hypothesized landmarks was motivated by the fact that L2 was often (though not always) near the beginning of syllable 3 (s3), sometimes in the onset consonant. The results are shown in Table 3, with all syllable structures combined (since we do not hypothesize L2 alignment to be affected by our syllable structure manipulations). Negative mean values for begV3toL2 indicate that L2 tended to be realized in the onset consonant of the last syllable (s3).

Hypothesis L2-A and segmental anchoring predict that L2 will be consistently aligned with a segmental landmark. This alignment should remain stable across rates. For two of the speakers, however, the alignment of L2 was affected by rate. For both Speaker 1 and Speaker 6, L2 was aligned later in the the normal rate than

in the fast rate (Speaker 1 begS3toL2:  $t = 5.36, df = 19, p < 0.001$ , begV3toL2:  $t = 3.21, df = 19, p < 0.01$ ; Speaker 6 begS3toL2:  $t = 4.41, df = 12, p < 0.01$ , begV3toL2:  $t = 2.16, df = 12, p = 0.054$ ). Other speakers also showed trends in the direction of later L2s in the normal condition, although these differences were not significant in the paired t-tests. However, in independent samples t-tests with all 147 items, this pattern was significant not only for Speakers 1 and 6, but also Speaker 2 (begS3toL2 normal rate: 132.79 ms, fast rate: 94.25 ms,  $t = 1.92, df = 29, p = 0.064$ , begV3toL2: normal rate: 43.42 ms, fast rate: 6.50 ms  $t = 2.85, df = 29, p < 0.01$ ).

### 3.1.5 H2 alignment

We next examined the alignment of H2, with respect to the competing hypotheses discussed in the introduction. We began by making plots of the data for each speaker by rate and by syllable structure. The plots revealed an apparent alignment difference based on syllable structure, but no apparent effect of rate. We therefore decided to examine the syllable structure differences with both rates combined; none of our hypotheses predicts an alignment difference based on rate.

We examined the alignment of H2 with respect to two potential segmental anchors, the end of V3 (endV3toH2) and the end of the sonorant rhyme (endSonRhymetoH2). Alignment with respect to these landmarks is relevant to distinguishing between Hypotheses H2-A and H2-B (end of syllable vs. end of sonorant rhyme). Means are shown in Table 4. For syllable structure CVCson, the sonorant rhyme was comprised of V3 and the following coda consonant; for all other syllable structures, the end of V3 was identical to the end of the sonorant rhyme. A separate mean for endSonRhymetoH2 is therefore given only for CVCson.

Speaker	Norm.	Fast	Norm.	Fast	N (pairs)
	begS3	begS3	begV3	begV3	
	toL2	toL2	toL2	toL2	
1	85	53*	5	-15**	20
	(7)	(3)	(7)	(3)	
2	123	111	34	11	8
	(22)	(18)	(15)	(9)	
3	38	34	-18	-13	5
	(6)	(6)	(6)	(2)	
5	59	47	-15	-15	5
	(4)	(6)	(4)	(5)	
6	72	31**	-5	-23***	13
	(8)	(4)	(7)	(3)	

Table 3: Alignment of L2 (in ms) relative to the beginning of syllable 3 and relative to the beginning of the syllable 3 vowel, sentence corpus (paired comparisons). Standard error is shown in parentheses. \* $p < 0.001$ , \*\* $p < 0.01$ , \*\*\* $p = 0.054$ .

For these analyses, we performed a series of t-tests for each speaker, treating latency values for each syllable as independent samples. Relying on paired comparisons would have drastically reduced the data available for the analyses, since we would

Speaker	S3 structure	endV3	endSonRhyme	N
		toH2	toH2	
1	CVCfr	0 (1)		11
	CVCst	1 (1)		12
	CVCson	41 (5)	-36 (7)	10
	CV	0 (7)		11
2	CVCfr	-6 (3)		7
	CVCst	-14 (4)		9
	CVCson	66 (21)	-37 (8)	7
	CV	-15 (4)		8
3	CVCfr	-13 (4)		4
	CVCst	-18 (9)		6
	CVCson	-11 (16)	-74 (15)	6
	CV	-22 (3)		3
5	CVCfr	-16 (6)		4
	CVCst	-14 (6)		5
	CVCson	5 (19)	-102 (48)	4
	CV	-16 (10)		3
6	CVCfr	-23 (2)		8
	CVCst	-13 (1)		8
	CVCson	-10 (3)	37-70 (6)	12
	CV	-16 (1)		9

Table 4: Alignment of H2 (in ms) relative to the end of the syllable 3 vowel and

be examining only cases in which a LHLH pattern was produced in the target AP of all items of a quadruplet. We reasoned that if the effects of syllable structure were strong enough, they would emerge even in independent samples analyses (which, unlike paired t-tests, do not control for cross-item differences, such as intrinsic vowel length differences between an [a] and an [i]).

The means are shown in Table 4. For each speaker, we performed the six syllable structure comparisons possible for endV3toH2 and the two possible for endSonRhymetoH2. We therefore set the alpha value to 0.0083 for the endV3toH2 comparisons and to 0.025 for the endSonRhymetoH2 comparisons, applying the Bonferroni correction to a base alpha value of 0.05 ( $0.05/6 = 0.0083$  and  $0.05/2 = 0.025$ ) to correct for multiple comparisons.

The general pattern that emerged from the analyses is that H2 tended to be aligned later in the CVCson syllable structure: the endV3toH2 value for the CVCson structure was often greater than those for the other structures. For the CVCfr, CVCst, and CV structures, H2 tended to be realized within 20 ms of the end of the vowel. Examples of this difference are given in Figures 9 and 10. For the CVCson structure, standard error was generally much larger than for other structures: H2s in this condition seem to have more flexibility in where they can appear. They sometimes appeared at the end of the vowel, sometimes in the early or middle part of the coda consonant, and more rarely at the end of the coda consonant.

Although we did not directly compare alignment across speakers, there did seem to be differences. In particular, Speaker 1's endV3toH2 values are very close to 0, showing that she tended to realize H2 at the very end of the vowel. For the other speakers, these values are negative, indicating that H2 was reached somewhat before

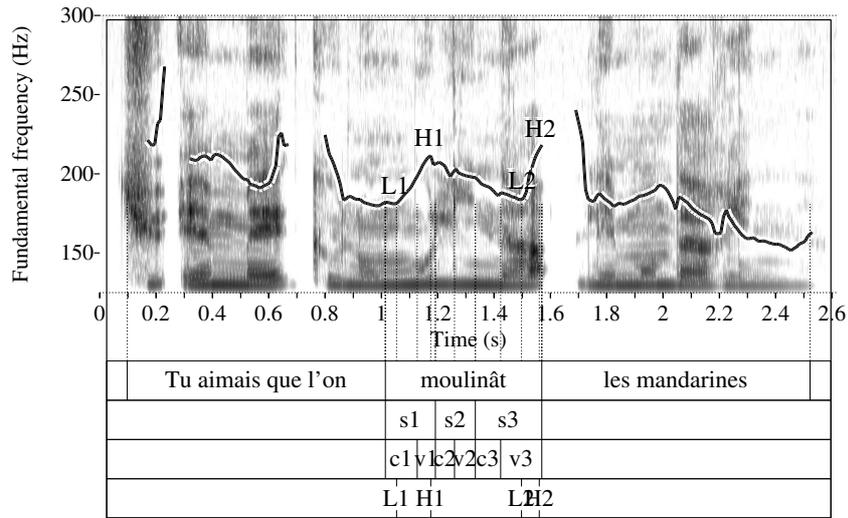


Figure 9: Illustration of the alignment of H2 at the end of the vowel of a CV target syllable ([na]) (Speaker 2, normal rate). The gloss is ‘You liked one to grind the mandarines.’

the end of the vowel. Finally, for the structure CVCson, the positive means for endV3toH2 combined with the negative means for endSonRhymetoH2 values indicate that H2 was often reached within the coda consonant. The relatively large absolute values of endSonRhymetoH2 indicate that H2 was generally realized well before the end of the syllable for CVCson syllables.

Quantifying these differences, for Speaker 1, endV3toH2 in the CVCson condition was significantly different from this latency in all other structures (CVCson *vs.* CVCfr:  $t = 8.85, df = 19, p < 0.001$ ; CVCson *vs.* CV:  $t = 4.56, df = 19, p < 0.001$ ; CVCson *vs.* CVCst:  $t = 8.30, df = 20, p < 0.001$ ). The same was true for endSonRhymetoH2 (CVCson *vs.* CVCfr:  $t = 5.23, df = 19, p < 0.001$ ; CVCson *vs.* CV:  $t = 3.50, df = 19, p < 0.0083$ ; CVCson *vs.* CVCst:  $t = 5.34, df = 20, p < 0.001$ ). No

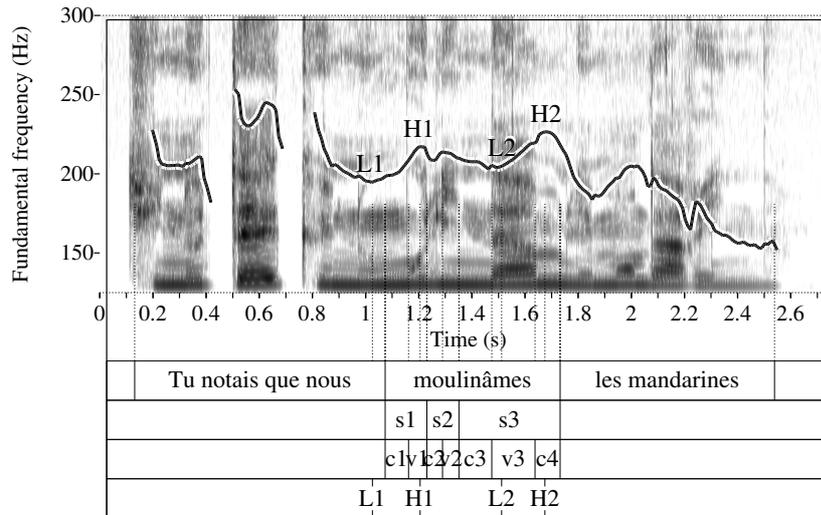


Figure 10: Illustration of the alignment of H2 in the coda consonant of a CVCson target syllable ([nam]) (Speaker 2, normal rate). The gloss is ‘You noted that we had ground the mandarines.’

other comparisons approached significance.

For Speaker 2, endV3toH2 in the CVCson condition was significantly different from that in the CV condition ( $t = 3.818, df = 13, p < 0.0083$ ) and from that in the CVCst condition ( $t = 3.763, df = 14, p < 0.0083$ ). The endV3toH2 comparison between CVCson and CVCfr failed to reach significance at the set alpha value ( $t = 3.41, df = 12, p = 0.014$ ). For endSonRhymetoH2, the CVCson *vs.* CVCfr comparison showed a significant difference ( $t = 3.74, df = 12, p < 0.0083$ ). The CVCson *vs.* CVCst comparison ( $t = 2.83, df = 14, p = 0.013$ ) and the CVCson *vs.* CV comparison ( $t = 2.72, df = 13, p = 0.017$ ), however, did not reach significance. No other comparison approached significance.

For Speaker 3, we considered only the comparison between CVCson and CVCst,

since the other two structures were represented by only three or four tokens. This comparison was not significant for endV3toH2 ( $t = 0.402, df = 10, p = 0.696$ ), but approached significance for endSonRhymetoH2 ( $t = 3.211, df = 10, p < 0.009$ ).

We were unable to make formal comparisons for Speaker 5, since there were so few tokens for each condition. We note, however, that the data for this speaker seemed to follow the general pattern found for other speakers, with later alignment of H2 in the CVCson condition.

Finally, for Speaker 6, the CVCson *vs.* CVCfr comparison was significant for endV3toH2 ( $t = 3.212, df = 18, p < 0.0083$ ): H2 was realized later in CVCson syllables than in CVCfr syllables. Somewhat surprisingly, the CVCfr *vs.* CVCst comparison was also significant for endV3toH2 ( $t = 3.372, df = 14, p < 0.0083$ ); H2 peaks were realized later in CVCst syllables than in CVCfr syllables. An examination of the spectrograms for these cases suggests a possible explanation for this unexpected result. In some cases, there was a weak voicing bar early in the closure of the coda /t/ in the CVCst syllables. This allowed the  $F_0$  peak to be realized slightly after the end of the vowel. This was true for a number of speakers, but for Speaker 6, the phenomenon was particularly apparent. All of the comparisons for endSonRhymetoH2 were significant (CVCson *vs.* CVCfr:  $t = 7.85, df = 18, p < 0.001$ ; CVCson *vs.* CV:  $t = 9.42, df = 19, p < 0.001$ ; CVCson *vs.* CVCst:  $t = 9.84, df = 18, p < 0.001$ ).

To summarize, H2 was generally aligned later in CVCson syllables than in syllables with other rhyme structures. In CVCson syllables, H2 was often, though not always, realized in the coda consonant. It was rarely realized at the very end of this coda consonant, although this alignment was possible.

Certain aspects of H2 alignment are not reflected by the alignment analyses, so we

will briefly discuss them here. For example, the mean endV3toH2 latencies (and low standard error) for Speaker 6 show that this speaker tended to align H2 near the end of the vowel, even in syllables with the CVCson structure. However, it is worth noting that this speaker sometimes had a high plateau or slightly falling plateau following the initial H2 at the end of the vowel. Across speakers, about 9% of H2s were followed by clear plateaux. These plateaux were an average of 50 ms in length, and typically occurred in syllables with the CVCson structure.

Finally, in her CV items, Speaker 1 occasionally aligned H2 with the end of the onset consonant of the post-accentual syllable (i.e., the [l] of the definite article following the target verb). As we noted in the introduction, this alignment pattern is uncommon: it is not observed in most speakers, and we have not seen it mentioned in the literature.

## 3.2 *Paragraph corpus*

### 3.2.1 *Rate manipulation*

Rate was calculated according to the procedure described for the sentence corpus. For this corpus, rate was calculated across entire paragraphs, across target sentences, and across target words. Results of ANOVA showed that all speakers successfully raised their speaking rate from the normal to the fast rate for all three sets of comparisons (for paragraphs:  $F(1,30) = 754.55$ ,  $p < 0.001$ ; for target sentences:  $F(1,102) = 379.87$ ,  $p < 0.001$ ; for target words:  $F(1,102) = 107.54$ ,  $p < 0.001$ ). Figure 11 shows rates for the target word for each speaker.

As in the sentence corpus, speaking rates and the magnitude of changes from one

rate to the other varied from speaker to speaker. There were main effects of Speaker (for paragraphs:  $F(5,30) = 58.55$ ,  $p < 0.001$ ; for target sentences:  $F(5,102) = 27.81$ ,  $p < 0.001$ ; for target words:  $F(5,102) = 14.38$ ,  $p < 0.001$ ), as well as significant Rate  $\times$  Speaker interactions (for paragraphs:  $F(5,30) = 14.09$ ,  $p < 0.001$ ; for target sentences:  $F(5,102) = 7.51$ ,  $p < 0.001$ ; for target words:  $F(5,102) = 3.49$ ,  $p < 0.01$ ).

But as a comparison of Figures 7 and 11 shows, speaking rate was generally slower in the paragraph corpus than in the sentence corpus. A series of t-tests for independent samples showed that this difference was significant for all but two comparisons ( $p < 0.001$  for seven comparisons,  $p < 0.01$  for one comparison,  $p < 0.05$  for two comparisons). The exceptions are the normal rates for Speaker 2 and Speaker 5, which did not differ across corpora ( $p = 0.14$  and  $p = 0.78$ , respectively).

### 3.2.2 Results

#### 3.2.3 Items available for further analyses

Of the 216 items, a total of 55 items were produced without the late rise and were excluded from the analyses. In most of these cases, the target AP was produced with a LHH pattern, although there were also many cases of falls across the target AP. Of the remaining 161 items, an additional 25 items produced with the LLH pattern in the target AP were excluded, leaving 136 items. Finally, three items that were produced with a clear final schwa in the target AP were excluded from these analyses. Thus, a total of 133 items remained for the analyses.

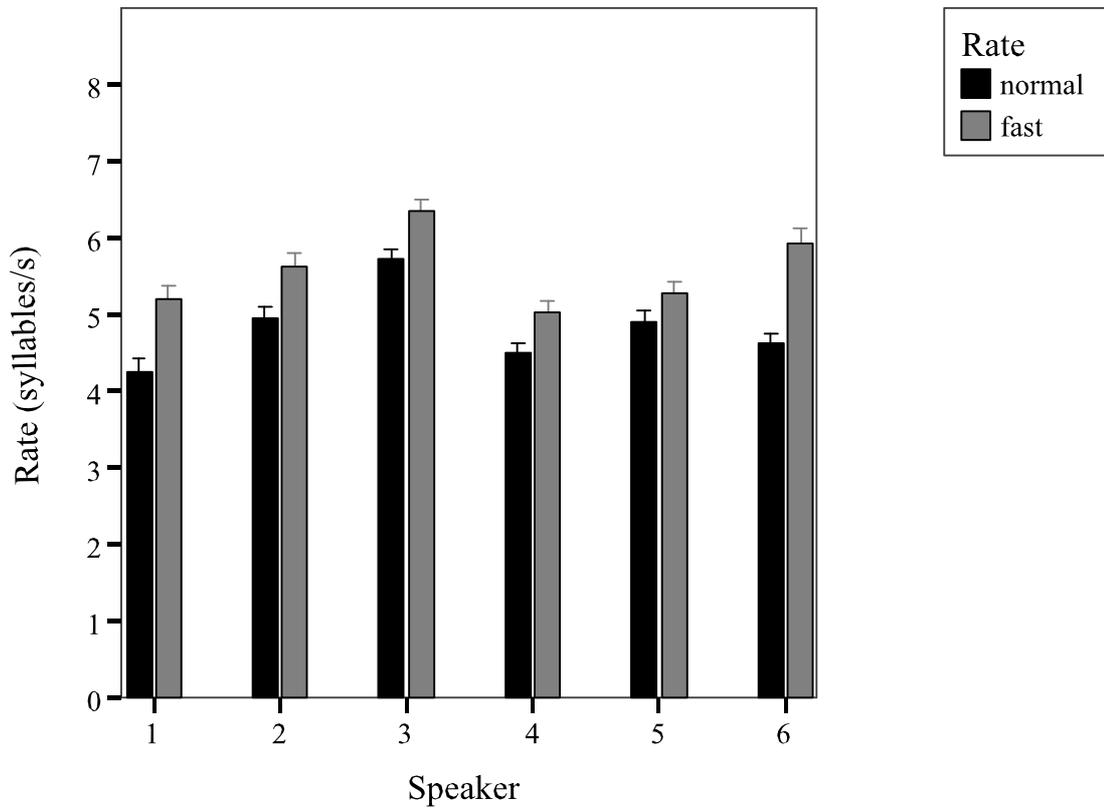


Figure 11: Speaking rates for the target words in the paragraph corpus. Bars show standard error of the mean.

#### 3.2.4 Tonal scaling, rise time, slope analyses

As with the sentence corpus, we performed paired comparisons between the normal and the fast rate means for  $F_0$  excursion, rise time, and slope. The means are given in Table 5. Speaker 6 is not included, since she had only two pairs produced with the LHLH pattern in the target AP (of a total of seven items produced with the LHLH pattern in the normal rate and three in the fast rate).

##### *Rise time*

Speaker	rise time		$F_0$ exc.		slope		N (pairs)
	normal	fast	normal	fast	normal	fast	
1	138	126	119	95	0.93	0.75	5
	(13)	(12)	(19)	(15)	(0.23)	(0.12)	
2	127	95**	69	48*	0.57	0.51	10
	(11)	(7)	(9)	(4)	(0.08)	(0.04)	
3	109	96	28	27	0.26	0.28	7
	(8)	(11)	(3)	(4)	(0.03)	(0.03)	
4	116	100	35	38	0.31	0.39*	9
	(9)	(7)	(3)	(3)	(0.02)	(0.03)	
5	114	91*	30	31	0.28	0.38*	16
	(7)	(7)	(2)	(3)	(0.03)	(0.04)	

Table 5: Mean  $F_0$  excursion (Hz), rise time (ms), and slope (Hz/ms) by speaker, separated by rate, paragraph corpus (paired comparisons). Standard error is shown in parentheses. \* $p < 0.01$ , \*\* $p < 0.05$ .

As in the sentence corpus, rise times were numerically longer in the normal condition than in the fast condition for all speakers, but results of paired t-tests showed that this difference was statistically significant for only two speakers (Speaker 2:  $t = 3.28, df = 9, p < 0.01$ ; Speaker 5:  $t = 3.02, df = 15, p < 0.01$ ). Independent t-tests, using all 133 items with the LHLH pattern in the target, also showed a significant rate difference for these speakers (Speaker 2 normal rate: 134 ms, fast rate: 93 ms,  $t = 3.317, df = 22, p < 0.01$ ; Speaker 5 normal rate: 112 ms, fast rate: 92 ms,  $t = 2.175, df = 32, p < 0.05$ ).

As in the sentence corpus, speakers in the paragraph corpus did not maintain a constant rise time. Due to space constraints, we do not include a scatterplot for the 133 LHLH items in this corpus. But we observe the same pattern of variability. In the 21 items available for Speaker 1 in the paragraph corpus, for example, rise times for the late rise varied from a minimum of 104 ms to a maximum of 167 ms. This is further evidence against Hypothesis L2-B and a leading tone analysis of L2.

#### *F<sub>0</sub> excursion*

Another series of paired t-tests tested the effect of rate on  $F_0$  excursion size. As for the sentence corpus,  $F_0$  analyses were performed in hertz, ERBs and semitones. Values are reported in hertz.

One speaker, Speaker 2, showed a significant effect of rate on  $F_0$  excursion, with larger  $F_0$  excursions in the normal rate ( $t = 2.49, df = 9, p < 0.05$ ). The same results were found using semitones and ERBs. Speaker 2 was also the only speaker to show a significant difference in the independent t-test analyses of all 133 LHLH items (normal rate: 71 Hz, fast rate: 47 Hz;  $t = 2.720, df = 22, p < 0.05$ ).

#### *Slope*

Speaker	r	p	N
1	-0.140	0.272 ns	21
2	0.442	0.015*	24
3	0.527	0.010*	19
4	0.316	0.062 ns	25
5	-0.139	0.216 ns	34
6	0.477	0.082 ns	10

Table 6: Correlations between rise time and  $F_0$  excursion size (measured in Hz) for the paragraph corpus, both rates combined.

For two speakers, there was a significant effect of rate on slope (Speaker 4:  $t = 3.11, df = 15, p < 0.01$ ; Speaker 5:  $t = 5.27, df = 8, p < 0.01$ ). In both cases, slopes were steeper in the fast condition. The rate effect was also found for these speakers (and only these speakers) with the ERB and semitone analyses.

In the independent t-test analyses, a marginally significant rate effect was found for Speaker 5 (normal rate: 0.27 Hz/ms, fast rate: 0.37 Hz/ms;  $t = 3.495, df = 32, p = 0.062$  for the Hz analysis). The same pattern was found for the ERB and semitone analysis. The difference across rates for Speaker 5 reached significance for the semitone analysis (normal rate: 2.35 st, fast rate: 2.67 st;  $t = 5.897, df = 32, p < 0.05$ ).

As in the sentence corpus, we examined whether rise time and  $F_0$  excursion size were positively correlated. The results are shown in Table 6. The correlations were significant for two speakers, Speakers 2 and 3, and the correlation for Speaker 4 approached significance.

### *H2 scaling in CVCobs vs. CV and CVCson syllables*

As in the sentence corpus, we compared H2 scaling in syllables with obstruent rhymes *vs.* syllables with sonorant rhymes. For only one speaker was there a significant difference. For Speaker 4, H2 was significantly lower in CVCobs syllables than in CV and CVCson syllables (213 Hz *vs.* 222 Hz;  $t = 2.64, df = 23, p < 0.05$ ), a difference in the direction predicted by Hypothesis H2-C.

### *3.2.5 L2 alignment*

We next examined the alignment of the L2, first plotting the results by speaker and rate. L2 was almost always realized in the last syllable of the APs, although in this corpus, it was occasionally realized in the penultimate syllable. Paired t-tests were conducted for each speaker to examine the potential effect of rate on alignment of L2. The means for the five speakers examined are given in Table 7.

For two speakers, L2 alignment was affected by rate. For Speaker 1, L2 was aligned later in the normal rate than in the fast rate (begS3toL2:  $t = 4.11, df = 4, p < 0.05$ , begV3toL2:  $t = 3.82, df = 4, p < 0.05$ ). The same pattern was found for this speaker in the sentence corpus, although the values do not seem comparable across corpora. For Speaker 5, we found the reverse pattern: L2 was aligned later in the fast rate than in the normal rate (begS3toL2:  $t = 2.67, df = 15, p < 0.05$ ; begV3toL2:  $t = 2.65, df = 15, p < 0.05$ ). This pattern was not found in the sentence corpus. Similar patterns were found in the independent t-tests analyses, with significant differences for Speaker 1 (begS3toL2 normal rate: 70 ms, fast rate: 15 ms;  $t = 4.756, df = 19, p < 0.001$ ; begV3toL2 normal rate: -5 ms, fast rate: -44 ms  $t = 3.83, df = 19, p < 0.01$ ) and a marginally significant difference for Speaker 5 (begV3toL2 normal

S	Normal	Fast	Normal	Fast	N (pairs)
	begS3	begS3	begV3	begV3	
	toL2	toL2	toL2	toL2	
1	80	15*	4	-44*	5
	(12)	(7)	(10)	(9)	
2	55	53	-25	-20	10
	(7)	(7)	(7)	(6)	
3	63	37	1	-14	7
	(14)	(18)	(9)	(16)	
4	80	80	3	10	9
	(12)	(13)	(14)	(13)	
5	54	68*	-14	3*	16
	(6)	(7)	(6)	(6)	

Table 7: Alignment of L2 (in ms) relative to the beginning of syllable 3 and relative to the beginning of the syllable 3 vowel, paragraph corpus (paired comparisons). Standard error is shown in parentheses.  $*p < 0.05$ .

rate:  $-14.18$  ms, fast rate:  $1.88$  ms;  $t = 2.013$ ,  $df = 32$ ,  $p = 0.053$ ).

We note that although Speaker 6 showed rate effect in the sentence corpus, her data for the paragraph corpus were too sparse to examine.

### 3.2.6 H2 alignment

As with the sentence corpus, plots of H2 by speaker, rate and syllable structure revealed apparent alignment differences based on syllable structure, but no apparent rate effect. We therefore combined rates in the analyses of the influence of syllable structure on H2 alignment. For the three endV3toH2 comparisons, alpha was set to  $p < 0.017$ , and for the two endSonRhymetoH2 comparisons, alpha was set to  $p < 0.025$ .

Mean alignment latencies of H2 with respect to the end of V3 (endV3toH2) and the end of the sonorant rhyme (endSonRhymetoH2) are shown in Table 8. The same general patterns of H2 alignment were found for the paragraph corpus as for the sentence corpus (see Table 4). First, H2 was generally aligned later for syllables closed by sonorant codas (CVCson) than in syllables with other rhyme structures. Second, for CVCobs (CVCfr and CVCst) and CV structures, H2 was realized within 15 ms of the end of the vowel (although the precise details of H2 alignment varied somewhat across the two corpora). Third, H2 in CVCson syllables was generally realized before the end of the coda consonant (as shown by the negative endSonRhyme means). Finally, standard error tended to be larger for the CVCson structure, indicating greater variability in H2 alignment. As in the sentence corpus, H2s in the paragraph corpus were sometimes realized with a following high plateau. In the paragraph corpus, plateaux averaged 42 ms in length and were all realized in syllables with the

CVCson structure.

The target syllables in the CVCobs condition were not homogeneous; they included in the critical codas a voiceless stop (/t/ in *devinette* ‘riddle’), a voiceless fricative (/s/ in *mérinos* ‘merino (sheep)’), two voiced stops (/d/ in *pyramide* ‘pyramid’ and *marinade* ‘marinade’), and two voiced fricatives (/z/ in *manganèse* ‘manganese’ and *Népalaises* ‘Nepalese women’). While we might have expected very different patterns to emerge for voiced and voiceless codas, this was not the case. One reason for this is that these phonemically voiced coda obstruents were always at least partially devoiced. This may be consequence of regressive assimilation due to the following context, which was always voiceless [k] (*qui, qu’il . . .*) (see Snoeren & Segui 2003 and references therein). Another possibility is that this devoicing is linked to the AP-final position of the segments involved. As Hambye (2005: 137, 138) writes:

<petit>

La “contrainte phonologique de marque” qui exclut les obstruantes voisées en *finale* d’unité prosodique est motivée sur le plan phonétique par la difficulté de programmer à la fois le maintien du voisement dans la consonne et la chute de l’énergie articulatoire liée à cette position finale. Ceci explique donc que plus la frontière prosodique est forte, plus la baisse d’énergie est importante, plus le dévoisement est probable. . . Ce caractère marqué des consonnes sonores finales est renforcé dans le cas des fricatives sonores par le fait que, lors de la production de ces consonnes, le voisement entre en compétition avec d’autres gestes articulatoires; ce qui explique que ces consonnes soient plus sujettes à l’assourdissement (v. Smith 1997; Di Cristo 1985: 335).

The “phonological markedness constraint” that excludes voiced obstruents in the final position of a prosodic unit is motivated from a phonetic point of view by the difficulty of simultaneously programming the maintenance of voicing in the consonant and the fall in articulatory energy linked to this final position. This therefore explains that the stronger the prosodic boundary, the larger the drop in energy, and the more likely it is that there will be devoicing. . . . This marked character of final voiced consonants is reinforced in the case of voiced fricatives by the fact that, during the production of these consonants, voicing competes with other articulatory gestures; this explains that these consonants are more often subjected to devoicing (see Smith 1997; Di Cristo 1985: 335).

Fricatives in our corpus did show a higher degree of devoicing than stops (with voicing throughout an average of 16% of the fricative noise for phonemically voiced fricatives *vs.* voicing throughout an average of 69% of the closure for phonemically voiced stops), in line with the results of earlier studies (see, for example, Bauvois 2000 on devoicing in Belgian French, cited in Hambye 2005). We note that perceptually, the fricatives in our study seem voiced; many other factors play a role in the perception of voicing (consonant duration, vowel length, intensity of frication, etc., see Hambye 2005 and references therein). We further address the realization of H2 in voiced obstruent coda consonants in the discussion section.

For Speaker 1, endV3toH2 was significantly different for syllables with the CVCson structure than those with the CV structure ( $t = 2.74, df = 11, p < 0.05$ ). EndSon-RhymetoH2 was significantly different for syllables with the CVCson structure than for those with the other two structures (CVCson *vs.* CV:  $t = 8.14, df = 11, p < 0.001$ ,

Speaker	S3 structure	endV3	endSonRhyme	N
		toH2	toH2	
1	CVCobs	8 (7)		8
	CVCson	15 (8)	-71(7)	6
	CV	-8 (3)		7
2	CVCobs	-9 (4)		8
	CVCson	31 (10)	-37(4)	8
	CV	-14 (3)		8
3	CVCobs	-9 (3)		6
	CVCson	22 (9)	-45(8)	6
	CV	-12 (2)		7
4	CVCobs	-11 (3)		10
	CVCson	38 (6)	-50(3)	8
	CV	-10 (5)		7
5	CVCobs	-13 (2)		11
	CVCson	34 (5)	-54(7)	12
	CV	-15 (2)		

Table 8: Alignment of H2 (in ms) relative to the end of the syllable 3 vowel and relative to the end of the sonorant rhyme, paragraph corpus. Note that for CVCson syllables, endV3toH2 and endSonRhymetoH2 are equivalent. Standard error is shown in parentheses.

CVCson *vs.* CVCobs:  $t = 7.47, df = 12, p < 0.001$ ). There were no other significant differences.

For all other speakers, all comparisons between CVCson and the other two other syllable structures (and only those comparisons) were significant (Speaker 2, endV3toH2: CVCson *vs.* CV:  $t = 4.51, df = 14, p < 0.001$ ; CVCson *vs.* CVCobs:  $t = 3.81, df = 14, p < 0.01$ ; endSonRhymetoH2: CVCson *vs.* CV:  $t = 4.85, df = 14, p < 0.001$ , CVCson *vs.* CVCobs:  $t = 4.94, df = 14, p < 0.001$ ; Speaker 3, endV3toH2: CVCson *vs.* CV:  $t = 3.75, df = 11, p < 0.025$ , CVCson *vs.* CVCobs:  $t = 3.40, df = 10, p < 0.025$ , endSonRhymetoH2: CVCson *vs.* CV:  $t = 4.22, df = 11, p < 0.01$ , CVCson *vs.* CVCobs:  $t = 4.48, df = 10, p < 0.01$ ; Speaker 4, endV3toH2: CVCson *vs.* CV:  $t = 5.89, df = 13, p < 0.001$ , CVCson *vs.* CVCobs:  $t = 7.56, df = 16, p < 0.001$ , endSonRhymetoH2: CVCson *vs.* CV:  $t = 6.24, df = 13, p < 0.001$ , CVCson *vs.* CVCobs:  $t = 8.83, df = 16, p < 0.001$ ; Speaker 5, endV3toH2: CVCson *vs.* CV:  $t = 8.87, df = 21, p < 0.001$ , CVCson *vs.* CVCobs:  $t = 8.75, df = 21, p < 0.001$ , endSonRhymetoH2: CVCson *vs.* CV:  $t = 5.38, df = 21, p < 0.001$ , CVCson *vs.* CVCobs:  $t = 5.70, df = 21, p < 0.001$ ).

#### 4 Discussion

The analyses of  $F_0$  excursion, rise time, and slope analyses showed a great deal of inter-speaker variability. We found some significant rate effects on each of the three parameters, but no patterns that held for all (or even most) speakers. In addition, comparisons between the two corpora also reveal intra-speaker variability. For example, for the paragraph corpus, Speaker 2 had significantly smaller  $F_0$  and shorter rise times in the fast rate than in the normal rate, but she did not show this

pattern in the sentence corpus. There were almost no significant results for a given speaker that held across both corpora.<sup>22</sup>

We found no overwhelming support for the constant slope hypothesis. It was not generally the case that rise time was correlated with  $F_0$  excursion. A correlation was found for three speakers, Speaker 2 in the paragraph corpus, Speaker 5 in the sentence corpus, and Speaker 3 in both corpora (a rare case in which the pattern held across corpora). Significant rate effects on  $F_0$  excursion were uncommon, and of the three significant differences found, one showed the opposite pattern of that predicted by the constant slope hypothesis: in the sentence corpus, Speaker 1 had larger  $F_0$  excursions in the fast rate than in the normal rate. Despite the lack of clear evidence in favor of the constant slope hypothesis, given the apparent similarity in mean slope for some speakers in the two rates, we cannot entirely discount the possibility that speakers have at least preferences with respect to the steepness of  $F_0$  rises.

The variability found in the  $F_0$  excursion, rise time, and slope results and the lack of consistent evidence in support of a single hypothesis are unsurprising. Similar results were found for French in Fougeron & Jun (1998) and Welby (2006). Although differences in the scaling of the individual tones are not relevant to the questions examined here, we note that some significant differences were found, although no consistent pattern, in line with the results of Fougeron & Jun (1998) and Welby (2006).

We now turn to our analysis of the alignment of L2, the low starting point of the early rise. The alignment patterns observed in the current study were somewhat unexpected, given what we had observed in earlier studies. L2 was almost always realized in the last syllable of the AP, only rarely in the penultimate syllable. By

contrast, in our earlier studies (Welby 2002, 2003, 2006), while L2 was most commonly found in the last syllable, for some speakers, it was also often found in the penultimate syllable.

In both the current study and in our earlier studies, however, L2 demonstrated clear inter-speaker alignment differences. The studies also have in common that no clear candidate emerged as an alignment landmark for L2. We do not have an explanation for these inter-study differences, but they are not altogether surprising given reports in the literature about the variability in alignment of the start of the late rise (see discussion in Post 2000).

Our plots of the data revealed potential rate-dependent differences in L2 alignment for a few speakers. These differences were confirmed for Speaker 6 in the sentence corpus, Speaker 5 in the paragraph corpus, and Speaker 1 in both corpora. These differences are unexpected in a segmental anchoring account, which predicts that the start (like the end) of a rise should be stably anchored with respect to a segmental landmark and that this stability should be preserved under time pressure. We also note that a number of speakers who did not show significant rate effects exhibited considerable variability in L2 alignment (see, for example, the standard error values for Speaker 2 in the sentence corpus or Speakers 3 and 4 in the paragraph corpus).<sup>23</sup>

One might be tempted to argue that L2 is aligned to the beginning of the last syllable of the AP, based on the observation that it is often realized either in the onset or early in the vowel of this syllable. Yet, a consideration of all the available evidence, from the current study as well as from our previous work and from reports in the literature suggests that such a conclusion would be premature. It would not account for cases in the current study in which L2 is realized fairly late in the last

syllable (as in Figure 9, for example) or for cases in earlier studies in which L2 is realized in the penultimate syllable. And the rate effects and alignment variability observed are inconsistent with the stability found for start and end points in other languages.

Further evidence against an alignment of L2 to the beginning of the last syllable of the AP comes from the scatterplot in Figure 8. Considering that the peak of the late rise, H2, was often aligned to the end of the last syllable of the AP, if the start of the rise, L2, were aligned to the beginning of this syllable, we would expect there to be a strong correlation between the duration of the last syllable (in our two corpora, syllable 3) and rise time. There was, however, never even a weak correlation between the two. This lack of correlation is apparent in Figure 8 and also holds for the paragraph corpus.

As noted earlier, Figure 8 also shows that Hypothesis L2-B, the hypothesis of L2 as a leading tone of H2 in the traditional AM understanding of leading tones, is unsupported by the data. Studies for several other languages have also cast doubt on the AM concept of leading and trailing tones linked to associated tones by a fixed temporal interval (Arvaniti et al. 1998, 2000; Grice et al. 2000; Frota 2002). Nevertheless, the fact that both the start and the end of the French late rise (L2 and H2) are realized near the end of the AP gives us reason to believe that the two tones form a cohesive unit, as proposed in Hypothesis L2-C. We never observe an alignment pattern like the one in Figure 12, in which L2 is realized toward the middle of the AP rather than toward the end. We have never observed this pattern, which is physically possible, in our earlier studies, nor has the pattern been reported in the literature.

In contrast to the patterns found for L2, in our analysis of the alignment of H2,

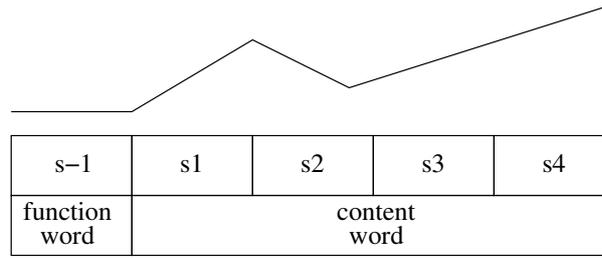


Figure 12: Schematic example of an unattested alignment of L2.

we found a consistent pattern of alignment across speakers. For only one speaker in one corpus was there a significant difference in the alignment patterns for syllable structures CV and CVCobs. By contrast, for all speakers, the alignment pattern for CVCson syllables was significantly different both from that for CV syllables and from that for CVobs syllables. For CV and CVCobs (CVCst and CVCfr) syllables, H2 was aligned at the very end of the vowel, or within about 20 ms of the end. For CVCson syllables, H2 was aligned either at the end of the vowel or somewhere in the coda consonant, or rarely at the very end of this consonant.

This pattern of results does not match the predictions of any of our hypotheses. Hypothesis H2-A is disconfirmed since H2 was not consistently realized with respect to the end of the vowel in CVCson syllables and we found significant alignment differences between CVCson syllables and syllables with other structures. Hypotheses H2-B and H2-C are disconfirmed because H2 was not consistently realized with respect to the end of the sonorant coda (end of the syllable) in CVCson. Further evidence against Hypothesis H2-C comes from the lack of consistent evidence of undershoot of H2 in CVCobs targets.

The patterns observed for French are similar to those found in studies of English, Dutch, and Neapolitan Italian, in which, as mentioned earlier, voiced or sonorant

codas pulled accent peaks rightwards (van Santen & Hirschberg 1994; Rietveld & Gussenhoven 1995; D’Imperio 2000). They are different from those observed for Mandarin, for which Xu (1998) found no influence of syllable structure (CV vs. CVN) on the alignment of tones. In our data, we also found some evidence of speaker-dependent alignment differences. Recall that Speaker 6 tended to align H2 near the end of the vowel, even for CVCson syllables, while Speaker 1, for example, aligned H2 well into the coda consonant in CVCson syllables (see Table 4). Similar variability may be present in other languages. (Note that Van Santen & Hirschberg 1994 examined a single speaker, and Rietveld & Gussenhoven 1995 is a perception study using synthetic speech.)

To account for the pattern of results observed in our data, we could postulate different anchor points depending on syllable structure. The anchor point for H2 in CV and CVCobs syllables would be the end of the vowel. The anchor point for CVCson syllables would be harder to define, given the variability of the  $F_0$  peak in this condition, but we could propose the end of the syllable. Yet, differences in phonetic alignment based on differences in segmental composition may mask underlying phonological regularities (see, for example, Arvaniti et al. 1998 on the alignment of the peak of the Greek prenuclear rise and Ladd et al. 2000 on the alignment of the peak of the Dutch prenuclear rise).<sup>24</sup> We therefore looked to provide a unified account.

In the account we propose, the differences in alignment patterns fall out from physical constraints on voicing that differ between syllable structures. Essential to the account is the concept of a segmental “anchorage.” We call this paper “Anchored down in Anchorage” not because we are Michelle Shocked fans (although PW is; see

Shocked 1988), but because the metaphor of an ANCHORAGE, an area where boats can drop anchor, more aptly describes the case of the peak of the French late rise than the more narrow metaphor of an immovable anchor. Our imaginary anchorage might consist of stretch of seabed bounded by underwater reefs or rocky crags; prudent sailors will drop anchor only where their anchors can sink into the sandy bottom, but never beyond the reefs. In our account, the segmental anchorage for H2 is the region stretching from approximately 20 ms before the end of the vowel to the end of the AP. We believe that the righthand boundary of the anchorage region is the end of the AP and not the end of the last full syllable because of our observations of alignment patterns in rare cases where the two do not coincide. In cases in which the target AP is produced with a final schwa, the H2 peak can appear in the vowel or coda consonant of the last full syllable, but also in the schwa syllable (contrary to reports in the literature, for example, Dell 1984: 69). An example, taken from the corpus, is given in Figure 13.<sup>25</sup>

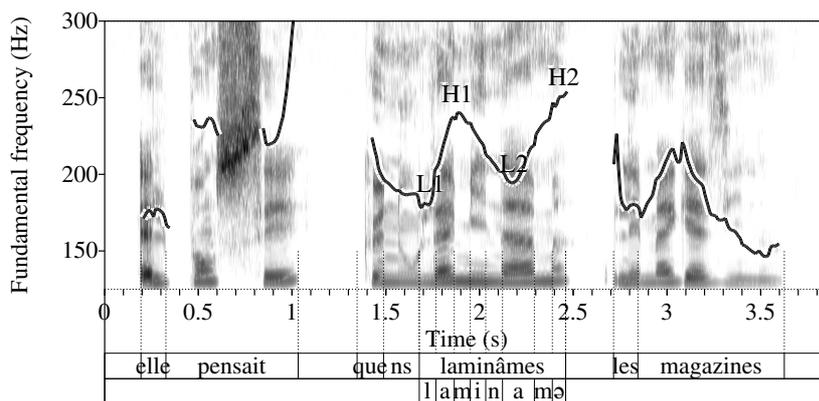


Figure 13: Illustration of H2 (late rise peak) realized in a schwa syllable (Speaker 1, normal rate). The gloss is ‘She thought that we had laminated the magazines’.

For items with the syllable structure CV, H2 can thus be realized at the end of the vowel or just beyond it (in cases in which voicing continues slightly beyond the end of the vowel/clear formant structure). For items with the syllable structure CVCson, H2 can be realized at the end of the vowel or in the coda consonant. For items with the syllable structure CVCobs, H2 can be realized at the end of the vowel or, in principle, in the voiced part of the consonant. In the current data, we observed, however, that it was rare for H2 to be realized more than a few milliseconds into a voiced obstruent coda. To verify that this pattern also holds for CVCobs items with fully voiced codas, we examined a subset of the data from Welby (2003, 2006). The corpus in this study contained one target item with a CVCobs in the final syllable, the word *limonade* ‘lemonade’ ([li.mo.nad]) in the sentence *La limonade a été versée par Anna* ‘The lemonade was poured by Anna’. In all of the tokens, the [d] was fully voiced, yet in 88% items in which the target AP *la limonade* was produced with a late rise, the  $F_0$  peak of the rise (H2) was realized in the vowel. We propose that the surface alignment difference between these two types of CVC syllables, CVCson, on the one hand, in which H2 is often realized in the coda consonant, and CVCobs, on the other, in which the realization of H2 in the coda consonant is much less common, is due to physical constraints on voicing rather than a phonological difference.

From the point of view of production, maintaining voicing during an obstruent requires articulatory effort to compensate for competing aerodynamic constraints (for discussion, see, for example, Maddieson 1997). For stops, since the oral closure causes supraglottal pressure to rise, the speaker must do something active to maintain the pressure differential across the glottis (such as lowering the larynx to increase the size of the oral cavity) and prevent the vocal folds from stalling. For fricatives, high

volume velocity of airflow is needed to produce frication at the constriction site, but the vibration of the vocal folds slows airflow. Simply maintaining voicing requires effort on the part of the speaker. To realize an  $F_0$  peak in a voiced obstruent, the speaker needs not only to maintain vocal fold vibration for voicing, but also to increase the rate of vibration to achieve an  $F_0$  peak. In the case of the French H2, it may simply not be worth expending the effort to realize the peak in a voiced obstruent, since the peak can be realized with no extra effort at the end of the vowel. These two alignment choices do not differ phonologically. Further evidence for this interpretation is the fact that there was no consistent difference in  $F_0$  height for H2 in syllables with sonorant rhymes versus those with obstruent rhymes; there was no evidence of undershoot.

Additional evidence that the target for H2 is not limited to the end of the vowel but extends to the end of the AP comes from the fact that we sometimes observed not a simple  $F_0$  but a high plateau (or a slightly falling plateau). These plateaux were often quite clear in CVCson syllables, but we also observed smaller plateaux in the CVCobs syllables. At the risk of torturing our metaphor, we imagine the anchor dragged along the seabed in high wind, though never beyond the reefs.

From the point of view of perception, the alignment of H2 after the vowel in CVCson syllables may reflect an attempt by the speaker to enhance this target. Plateau realizations of H2 may fulfill a similar purpose, providing a longer window for temporal integration. But the question of when and why H2 (or H1) is realized as a simple peak or as a plateau remains to be investigated. In addition, realizing H2 in the vowel may also preserve cues to the identity of the voiced coda consonant, insofar as voiced consonants have been claimed to have intrinsically lower  $F_0$  (Maddieson

1997, and references therein). In addition, H2 realized during the closure of a voiced stop would not be very salient, since this is a very low amplitude portion of the signal.

Although we did not directly examine rate effects on H2 alignment, we note that unlike our plots of L2 data, our plots of the H2 data did not show any obvious effects. Rather, there was a good deal of overlap between fast and normal rate tokens. We would expect further studies to confirm this claim.

Although we performed statistical analyses for only the items with the LHLH pattern in the target sequence, the pattern of alignment for the 40 items with the LLH pattern (15 in the sentence corpus, 25 in the paragraph corpus) was in line with that found for the LHLH patterns. In almost 90% of the CV and CVCobs items in the two corpora, H2 was produced before the end of V3 (12 of 13 for the sentence corpus, 14 of 16 for the paragraph corpus). In contrast, in the two CVCson items in the sentence corpus and seven of the nine CVCson items in the paragraph corpus, H2 was realized in the coda consonant. There may be tonal alignment differences between the LLH and LHLH patterns: Welby (2003, 2006) found that H2 was aligned slightly later in LHLH patterns than in other patterns (including LLH), possibly due to tonal crowding. But the difference in surface alignment between CVCson and other structures remains.

The results of the current study have a number of implications for the segmental anchoring hypothesis. First, we did not find convincing evidence that both the start and the end of the French late rise were stably anchored to segmental landmarks. We failed to find a plausible segmental anchor for the start of the rise (L2): it showed a great deal of variability and for some speakers, its alignment varied across rates.<sup>26</sup> In addition, rise times were not consistently shorter in the fast rate, as expected if

both the start and the end of the rise are anchored. We did find that the peak of the late rise (H2) was consistently aligned with respect to segmental landmarks, but in our segmental anchorage account we propose a different type of consistency than that described in traditional segmental anchoring accounts. Unlike those accounts, however, segmental anchorage provides a unified account of alignment for all syllable structures, despite apparent surface differences. The results of the current study and those of earlier studies call into question the segmental anchoring assumption that  $F_0$  excursions will remain stable across changes in speaking rate. Without a consistent effect of rate on rise time and a consistent lack of effect on  $F_0$  excursion, rate effects on slope in the direction predicted by segmental anchoring cannot be interpreted as support for the hypothesis. Finally, the inability of segmental anchoring to account for the alignment and scaling facts of French clearly casts doubt on the idea of segmental anchoring as a spoken language universal.

Although this is not the main focus of our paper, our results also pose a challenge to the Xu model, in which  $F_0$  movements are aligned with respect to “host units” Xu & Wang (2001) and in particular, the syllable. In the latest version of the model Xu and Liu (this volume) argue that “the syllable is the basic time structure that specifies the alignment of consonants, vowels, tones and phonation registers.” Recall that Xu (1998) interpreted the lack of variance in tone alignment across syllable structures (CV vs. CVNasal) as evidence that “the syllable is the proper domain for tone implementation” for Mandarin (p. 179). Similarly, Xu (1998) found no influence of speaking rate on tone alignment and interpreted this stability as support for the model. The syllable structure-dependent variation found across different syllable structures in French thus undermines the idea of the syllable as the host unit for

French tonal movements, as do the rate effects found for L2 alignment. Further challenging the idea that the entire late rise is aligned with respect to a syllable is the fact that L2 can be realized in the penultimate or in the final syllable of the AP. A similar problem is posed by the early rise, whose H1 peak shows no evidence of being associated to a syllable (see Welby 2003, 2006 for details).

We stress that we do not conceive of segmental anchorage, in which a region rather than a specific point is the “docking site” for the start or end of an intonational rise or fall, as a spoken language universal. While we have used the concept to explain the alignment of the peak of the French late rise (H2), we do not claim that it will account for alignment patterns found in all other languages. The segmental anchoring literature demonstrates convincingly that for some languages (e.g., Greek and German), starting and end points of rises cluster tightly around a single anchor point. Indeed, even for French, segmental anchorage may not provide an appropriate account for all alignment patterns. What we do claim, however, is that alignment patterns of the type found for the French H2 will be found for other languages. That is, the segmental anchorage alignment pattern is one of a range of alignment patterns possible in the world’s spoken languages.

## *5 Acknowledgments*

We thank Pascal Perrier for helpful comments, Barbara Gili Fivela and Cristel Portes for their useful comments on an earlier version of the manuscript, and Allison Blodgett, Christian Bulfone, Nino Medves, and W. Detmar Meurers for technical assistance. We are also grateful to our speakers for their participation. Pauline Welby’s work on this project was supported by a Marie Curie International Fellowship within

the 6th European Community Framework Programme.

## Notes

<sup>1</sup>The figure shows a schematic representation of the segmental anchoring relationship between the starting and end points of an  $F_0$  rise and segmental landmarks. It should not be interpreted to suggest that the two are necessarily synchronized.

<sup>2</sup>These units do not always exactly correspond, although we will not detail the differences here.

<sup>3</sup>The tonal alignment facts make it clear that this pattern involves a rise from L1 to H2 (for details, see Welby 2003).

<sup>4</sup>The French of metropolitan France, *l'Hexagone*.

<sup>5</sup>In earlier work, we referred to the starting point of the early and late rise as *e-el* (for EARLY ELBOW) and *l-el* (LATE ELBOW), respectively. For greater transparency, we refer to them here as *L1* and *L2*.

<sup>6</sup>“Sonorant rhyme” refers to the syllable nucleus and any following sonorant coda consonant. Following the classic definition of *rhyme*, it excludes onset consonants. Note that the term is used differently in van Santen and Hirschberg (1994).

<sup>7</sup>Note that Figures 4, 5, and 6 schematize hypothesized alignments of H2. They should not be interpreted to suggest that H2 will be exactly synchronized to a given anchor point, nor should they be interpreted as making claims about the alignment of other tones.

<sup>8</sup>Note that Hypothesis H2-C does not distinguish between voiceless and voiced

obstruents. Realization of a peak in a voiced consonant is possible, but requires increased articulatory effort. We therefore reasoned that voiceless and voiced obstruent codas might pattern together with respect to alignment. One of the corpora examined contained target items with voiced coda consonants in the final syllable, allowing us to examine this question.

<sup>9</sup>In principle, other factors might come into play in the scaling of H2 in CVCobs syllables. For example, H2 could be scaled at least in part with respect to H1. However, in our earlier work we have found that the H2 is often, though crucially not always, higher than H1 (Rolland & Løevenbruck 2002; Welby 2003). An undershot H2 should therefore be possible.

<sup>10</sup>Our hypotheses do not make different predictions based on whether the voiceless obstruent coda is a stop or a fricative. We decided, however, to include both syllables ending in /t/ and those ending in /s/, since the verb paradigms of the language offered the possibility.

<sup>11</sup> The verb forms used are chiefly literary forms that are not typically found in everyday speech. They are, however, pronounced in contexts in which one reads aloud. Although a number of speakers noted that the forms were “unusual,” none had any apparent difficulty in pronouncing them. Francophone readers will note that *démunis* is also the present subjunctive, and that *démunit* /de.my.ni/, the simple past, and *démunît*, the imperfect subjunctive are phonemically identical (both /de.my.ni/).

<sup>12</sup>We had hoped to include target verbs in *-re*, some of which have the vowel /y/ in

the simple past and the imperfect subjunctive (e.g., *moudre* ‘to mill, grind’, *moulût* /mu.ly/). But a search of the French database LEXIQUE did not reveal any verbs corresponding to our criteria.

<sup>13</sup>We are aware that there may be intonational differences between restrictive and non-restrictive (appositive) relative clauses, as claimed by Philippe Martin (personal communication), for example, and that a falling pattern across the phrase containing the head noun is possible. In many cases in the paragraph corpus, the relative clause could be interpreted as restrictive or non-restrictive. For example, in the second target sentence in (4), it is possible to imagine both a non-restrictive reading and a restrictive reading, as shown below.

Les vitamines qu’ils recèlent naturellement sont un prétexte pour les gourmands.

NON-RESTRICTIVE: ‘The vitamins, which they naturally contain, are just an excuse for people who like to eat.’

RESTRICTIVE: ‘The vitamins that they naturally contain are just an excuse for people who like to eat. (But the vitamins that they contain because they are genetically modified are another story.)’

The relatives were never set off by commas, which may have biased readers to a restrictive interpretation, although lack of commas does not force a restrictive interpretation. In any case, an investigation of the claimed intonational difference is beyond the scope of the current study. In our analyses, we considered only targets

that contained late rises (which were far more common than falls).

<sup>14</sup>The lack of exact minimal triplets in the corpus was probably an advantage: the target words were well “hidden” in the paragraphs, and speakers were largely unaware of the experimental manipulation.

<sup>15</sup>The script looked for a maximum in the region of the first and second syllables of the AP for H1, and a maximum in the region of the last syllable of the AP and the immediately following syllable for H2.

<sup>16</sup>Reliable hand measurements may be possible for some languages or some corpora, as Lickley et al. (2005) note, citing Arvaniti (personal communication). For our own data, for example, if we examined L2 only in LHLH patterns,  $F_0$  minima identified by eye would probably have corresponded to those points found by the automatic procedure. But for other cases, including our LLH patterns, a procedure using line-fitting or the identification of an acceleration maximum (used in Xu 1998) is appropriate.

<sup>17</sup>This is the crucial comparison, since it allows us to be sure that there was not only a global rate increase, but also an increase in the target region. Note that including pauses in the rate calculations had no effect on the rates calculated for target words, since there was never a pause within a target AP.

<sup>18</sup>We are confident that our target phrases were produced as accentual phrases, not intonation phrases (a higher level phrase). This distinction is important since different degrees of final syllable lengthening have been attributed to the two levels of phrasing; this would pose a clear problem if some of our target phrases were APs and some IPs. However, we observe downdrift across the target utterances, with no

evidence of the pitch reset we might expect at the beginning of a new IP. For example, in Figures 3a, 9, and 10, the H2 of the target AP is lower than the H2 of the preceding AP. Moreover, within each corpus, the target sentences are all quite similar, and the level of phrasing produced across utterances and across speakers is judged by author HL, a native speaker, to be comparable. Finally, it would be unusual for a speaker to introduce an intonation phrase boundary (that is, a very strong boundary) at the end of the target in either of our corpora. In the first corpus, the target phrase (e.g., *que tu démunisses* in Figures 3a) ends at the middle of a verb phrase, between the verb (*démunisses*) and its complement (*les malheureux*); in the second, the target phrase ends in the middle of the noun phrase subject of the sentence (e.g., *Les vitamines qu'ils recèlent naturellement* in (4)). Another argument for a given level of phrasing might come from a comparison of syllable lengthening between the last syllable of the target phrase (argued to be the last syllable of an AP) and the last syllable of the utterance (uncontroversially the last syllable of an intonation phrase). Such a comparison is not possible for the current corpus, however, since the syllables in the two conditions are not segmentally balanced. For example, in Figure 3, the target phrase *les vitamines* ends with a high vowel ([i]) in a closed syllable, while the last phrase of the utterance ends with a mid-high vowel ([ø]) in an open syllable (*malheureux*).

<sup>19</sup>The fact that we observed far fewer LHLH patterns than expected is likely to be due to the rate at which participants read the sentence corpus; rates were quite high, even in the normal rate condition. In retrospect, we could have interspersed filler sentences or sentences for another experiment with the critical sentences. Including sentences of different structures might have made the reading task slightly more chal-

lenging, and prevented speakers from falling into a rhythm that allowed them to read so quickly. As we will see, speaking rates were slower for paragraph corpus; by its nature, a paragraph will contain sentences of different structures and lengths. Similar observations are made by Lickley et al. (2005), who argue for the usefulness of read lab speech in studying intonation patterns. These authors note the importance of varying the structure of items in a read corpus to avoid “monotonous” productions of very similar items.

<sup>20</sup>We note that the pronunciation of word final schwas is relatively common in the speech of Parisians (for discussion of schwa in Parisian French, see for example, Adda-Dekker & Lamel 1999). This left a total of 147 items for the subsequent analyses.

<sup>21</sup>The formula used for the ERB calculation was the formula given in Hermes & van Gestel (1991):  $E = 16.7 \log_{10}(1 + f/165.4)$ , where  $E$  is the ERB-rate in ERB and  $f$  is frequency in hertz. The formula used for the semitone calculation was that given in 't Hart et al. (1990):  $D = 12 \log_2 \frac{f_1}{f_2} = \frac{12}{\log_{10} 2} \log_{10} \frac{f_1}{f_2}$ , where  $D$  is the distance in semitones between two frequencies  $f_1$  and  $f_2$  in hertz.

<sup>22</sup>We note that cross-corpora comparisons were not possible for Speaker 4 and Speaker 6.

<sup>23</sup>We cannot discount the possibility that early or later alignment of L2 may signal pragmatic differences, with a later L2 conveying greater assertiveness, for example (see discussion in Post 2000: 126). We did not perform formal perception studies to address this question, so in principle, some of the variability in L2 alignment may be due to pragmatic differences. But if these differences do exist, our informal listening

tests suggest that they are quite subtle. For example, it does not seem that speakers who showed rate effects in L2 alignment conveyed a greater degree of affirmativeness in one rate than in the other.

<sup>24</sup>Many alignment studies, however, have used materials with mostly or exclusively sonorant target sequences (Prieto et al. 1995; Ladd et al. 1999, 2000; D’Imperio 2000; Face 2002; Atterer & Ladd 2004, *inter alia*). Thus, for many languages, questions remain on the influence of syllable structure and segmental composition on alignment.

<sup>25</sup>This item, like all those with schwa in the target region, was not used in the analyses. (See § 3.2.3 and § 3.1.2.)

<sup>26</sup>We do not propose a segmental anchorage for L2, because no candidate region with clear, well-defined boundaries emerged from the alignment data.

### *Bibliographical References*

ADDA-DEKKER Martine & Lori LAMEL 1999. Pronunciation variants across system configuration, language and speaking style. *Speech Communication* 29. 83–98.

ARVANITI Amalia & D. Robert LADD 1995. Tonal alignment and the representation of accentual targets. *13th International Congress of Phonetic Sciences (ICPhS)*. 220–223. Stockholm.

ARVANITI Amalia, D. Robert LADD & Ineke MENNEN 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics* 26. 3–25.

ARVANITI Amalia, D. Robert LADD & Ineke MENNEN 2000. What is a starred tone? Evidence from Greek. In BROE Michael B. & Janet B. PIERREHUMBERT (eds.)

2000. *Papers in Laboratory Phonology V: Acquisition and the Lexicon*. Cambridge: Cambridge University Press. 119–131.
- ATTERER Michaela & D. Robert LADD 2004. On the phonetics and phonology of “segmental anchoring” of F0: Evidence from German. *Journal of Phonetics* 32. 177–197.
- BAUVOIS Cécile 2000. Des Belges et de l’assourdissement des sonores. *XXII<sup>e</sup> Congrès International de Linguistique et de Philologie Romanes* 3. 35–43.
- BECKMAN Mary E. & Julia HIRSCHBERG 1994. The ToBI annotation conventions. The Ohio State University & AT&T Bell Telephone Laboratories. Ms.
- BENQUEREL André-Pierre 1971. Duration of French vowels in unemphatic stress. *Language and Speech* 14. 383–391.
- BOERSMA Paul & David WEENINK 2005. PRAAT: Doing phonetics by computer (version 4.3) [Computer program]. Retrieved from <http://www.praat.org>
- DELL François 1984. L’accentuation dans les phrases en français. In DELL, François, Daniel HIRST & Jean-Roger VERGNAUD (eds.) 1984. *La forme sonore du langage*. Paris: Hermann. 65–122.
- DI CRISTO Albert 1976. Indices prosodiques et structure constituante. *Cahiers de Linguistique, d’Orientalisme et de Slavistique* 7. 27–40.
- DI CRISTO Albert 1985. *De la microprosodie à l’intonosyntaxe*. Aix-Marseille: Université de Provence.

- DI CRISTO Albert 1998. Intonation in French. In HIRST Daniel & Albert DI CRISTO (eds.) 1998. *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press. 195–218.
- DI CRISTO Albert 1999. Vers une modélisation de l’accentuation du français: première partie. *French Language Studies* 9. 143–179.
- DI CRISTO Albert 2000. Vers une modélisation de l’accentuation du français: seconde partie. *French Language Studies* 10. 27–44.
- D’IMPERIO Mariapaola 2000. *The Role of Perception in Defining Tonal Targets and their Alignment*. The Ohio State University. PhD dissertation.
- D’IMPERIO Mariapaola, Robert ESPESSER, Hélène LÆVENBRUCK, Caroline MENEZES, Noël NGUYEN & Pauline WELBY To appear. Are tones aligned with articulatory events? Evidence from Italian and French. In COLE Jennifer & José HUALDE (eds.), *Laboratory Phonology 9*. Mouton de Gruyter.
- FACE Timothy L. 2002. When push comes to shove: Tonal crowding in Madrid Spanish. *The Linguistic Association of Korea Journal* 10. 77–100.
- FOUGERON Cécile & Sun-Ah JUN 1998. Rate effects on French intonation: Prosodic organization and phonetic realization. *Journal of Phonetics* 26. 45–69.
- FROTA Sónia 2002. Tonal association and target alignment in European Portuguese nuclear falls. In GUSSENHOVEN Carlos & Natasha WARNER (eds.) 2002. *Laboratory Phonology 7*. Berlin: Mouton de Gruyter. 387–418.
- GRICE Martine, D. Robert LADD & Amalia ARVANITI (2000), On the place of phrase accents in intonational phonology. *Phonology* 17. 143–185.

- HAMBYE Philippe 2005. *La prononciation du français contemporain en Belgique: Variation, normes et identités*. Université Catholique de Louvain. Doctoral dissertation.
- HERMES, Dik J. & Joost C. VAN GESTEL 1991. The frequency scale of speech intonation. *Journal of the Acoustical Society of America* 90. 97–101.
- HIRST Daniel & Albert DI CRISTO 1996. Y a-t-il des unités tonales en français? *XXIèmes Journées d'Étude sur la Parole*. 223–226. Avignon.
- IGARASHI Yosuke 2004. “Segmental Anchoring” of F<sub>0</sub> under changes in speech rate: Evidence from Russian. *Speech Prosody 2004*. 25–28. Nara.
- 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. *Probus* 14. 147–172. (A special issue on the intonation of Romance languages, José HUALDE (ed.)).
- LADD D. R., Dan FAULKNER, Hanneke FAULKNER & Astrid SCHEPMAN 1999. Constant “segmental anchoring” of F<sub>0</sub> movements under changes in speech rate. *Journal of the Acoustical Society of America* 106. 1543–1554.
- LADD D.R., Ineke MENNEN & Astrid SCHEPMAN 2000. Phonological conditioning of peak alignment in rising pitch accents in Dutch. *Journal of the Acoustical Society of America* 107. 2685–2695.

- LICKLEY Robin J., Astrid SCHEPMAN & D. Robert LADD 2005. Alignment of “phrase accent” lows in Dutch falling-rising questions: Theoretical and methodological implications. *Language and Speech* 48. 157–183.
- MADDIESON Ian 1997. Phonetic universals. In HARDCASTLE William J. & John LAVER (eds.) 1997. *The Handbook of Phonetic Sciences*. Oxford & Cambridge, MA: Blackwell. 619–639.
- MARTIN Philippe 1979. Une théorie syntaxique de l’accentuation en français. *Studia Phonetica* 15. 1–12.
- MARTIN Philippe 1987. Prosodic and rhythmic structure in French. *Linguistics* 25. 925–949.
- MARTIN Philippe 2004. L’intonation de la phrase dans les langues romanes: l’exception du français. *Langue Française* 141. 36–55.
- MERTENS Piet 1987. *L’intonation du français. De la description linguistique à la reconnaissance automatique*. Université de Leuven. Doctoral dissertation.
- MERTENS Piet 2002. Synthesizing elaborate intonation contours in text-to-speech for French. *Speech Prosody 2002*. 499–502. Aix-en-Provence.
- MERTENS Piet, Jean-Philippe GOLDMAN, Éric WEHRLI & Arnaud GAUDINAT 2001. La synthèse de l’intonation à partir de structures syntaxiques riches. *Traitement Automatique des Langues* 42. 145–192.
- O’SHAUGHNESSY Douglas 1981. A study of French vowel and consonant durations. *Journal of Phonetics* 9. 385–406.

- PASDELOUP Valérie 1990. *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Institut de Phonétique d'Aix-en-Provence, Université de Provence. Doctoral dissertation.
- PIERREHUMBERT Janet B. 1980. *The Phonology and Phonetics of English Intonation*. Massachusetts Institute of Technology. PhD dissertation. (Distributed 1987 by Indiana University Linguistics Club).
- PIERREHUMBERT Janet B. & Mary E. BECKMAN 1988. *Japanese Tone Structure*. Cambridge, MA: MIT Press.
- PIERREHUMBERT Janet B. & Shirley A. Steele 1989. Categories of tonal alignment in English. *Phonetica* 46. 181–196.
- POST Brechtje 2000. *Tonal and Phrasal Structures in French Intonation*. The Hague: Thesus.
- POST Brechtje 2002. French tonal structures. *Speech Prosody 2002*. 11–13. Aix-en-Provence.
- PRIETO Pilar, Jan VAN SANTEN & Julia HIRSCHBERG 1995. Tonal alignment patterns in Spanish. *Journal of Phonetics* 23. 429–451.
- RIETVELD Toni & Carlos GUSSENHOVEN 1995. Aligning pitch targets in speech synthesis: Effects of syllable structure. *Journal of Phonetics* 23. 375–385.
- ROLLAND Guillaume & Hélène LÆVENBRUCK 2002. Characteristics of the accentual phrase in French: An acoustic, articulatory and perceptual study. *Speech Prosody 2002*. 611–614. Aix-en-Provence.

- ROSSI Mario 1985. L'intonation et l'organisation de l'énoncé. *Phonetica* 42. 135–153.
- VAN SANTEN Jan & Julia HIRSCHBERG 1994. Segmental effects on timing and height of pitch contours. *3rd International Conference on Spoken Language Processing (ICSLP)*. 719–722. Yokohama.
- SHOCKED Michelle 1988. Anchorage. *Short Sharp Shocked*. Mercury. ([http://www.michelleshocked.com/chords\\_anchorage.htm](http://www.michelleshocked.com/chords_anchorage.htm)).
- SMITH Caroline 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. *Journal of Phonetics* 25. 471–500.
- SNOEREN Natalie D. & Juan SEGUI 2003. A voice for the voiceless: Voice assimilation in French. *15th International Congress of Phonetic Sciences (ICPhS)*. 2325–2328. Barcelona.
- 'T HART Johan, René COLLIER & Antonie COHEN 1990. *A Perceptual Study of Intonation*. Cambridge: Cambridge University Press.
- VAISSIÈRE Jacqueline 1992. Rhythm, accentuation and final lengthening in French. In SUNDBERG Johan, Lennart NORD & Rolf CARLSON (eds.) 1992. *Music, Language, Speech and Brain*. (International Symposium Series 59). Stockholm: Wenner-Gren. 108–120.
- WELBY Pauline 2002. The realization of early and late rises in French: A production study. *Speech Prosody 2002*. 695–698. Aix-en-Provence.
- WELBY Pauline 2003. *The Slaying of Lady Mondegreen, being a Study of French Tonal association and Alignment and their Role in Speech Segmentation*. The Ohio State University. PhD dissertation.

WELBY Pauline 2006. French intonational structure: Evidence from tonal alignment.

*Journal of Phonetics* 34. 343–371.

XU Yi 1998. Consistency of tone-syllable alignment across different syllable structures

and speaking rates. *Phonetica* 55. 179–203.

XU Yi & Fang LIU 2006. Tonal alignment, syllable structure and coarticulation:

Toward an integrated model. This volume.

XU Yi & Q. Emily WANG 2001. Pitch targets and their realization: Evidence from

Mandarin Chinese. *Speech Communication* 33. 319–337.

Authors' postal and electronic addresses:

Institut de la Communication Parlée

Institut National Polytechnique de Grenoble

46, avenue Félix Viallet

38031 Grenoble Cedex 1

France

{welby, loeven}@icp.inpg.fr