

Differential Contribution of Prosodic Cues
in the Native and Non-Native Segmentation of French Speech

Annie Tremblay¹, Caitlin E. Coughlin¹, Carly Bahler², & Stéphanie Gaillard¹
University of Illinois at Urbana-Champaign¹, Indiana University²

Contact Information:

2090 Foreign Language Building, 707 S. Mathews Avenue, Urbana, IL 61801 USA

Phone: (217) 265-0764; Email: atrembla@illinois.edu

*This research was supported in part by the Department of French and the College of Liberal Arts and Sciences. Many thanks to Sarah Brown-Schmidt, Zsuzsanna Fagyal, Scott Frauendorf, Angie Isaacs, Tuan Lam, Eun-Kyung Lee, Molly Lewis, Nathan Owens, Rachel Ryskin, Kristin Tooley, Alison Trude, and Duane Watson for their valuable comments on this work

Abstract

This study investigates the use of prosodic information in the segmentation of French speech by mid-level and high-level English second/foreign language (L2) learners of French and native French listeners. The results of two word-monitoring tasks, one with natural stimuli and one with resynthesized stimuli, show that as L2 learners become more proficient in French, they go from parsing accented syllables as word-initial to parsing them as word-final, but unlike native listeners, they use duration increases but not fundamental frequency (F0) rises as a cue to word-final boundaries. These results are attributed to: (i) the L2 learners' native language, in which F0 rise is a reliable cue to word-initial boundaries but not word-final boundaries; (ii) the co-occurrence of F0 and duration cues in word-final syllables in French, rendering L2 learners' use of F0 rise unnecessary for locating word-final boundaries; and (iii) the marking of word-initial boundaries by F0 cues in French, thus making it difficult for non-native listeners to tease the two types of F0 rise apart. We argue that these factors prevent English listeners from attending to F0 rise as a cue to word-final boundaries in French, irrespective of their proficiency in French.

1. Introduction

Unlike written language, where word boundaries are often denoted by blank spaces (e.g., *le_chat* ‘the_cat’), for spoken language, no single device allows for the reliable identification of word boundaries: words are typically uttered without a pause between them, and phonological processes alter the phonetic realization of words and further blur word boundaries. One of the crucial challenges for second/foreign language (L2) learners is that the cues to word boundaries differ across languages. Thus, an English speaker’s experience with her native language may prove misleading when attempting to segment speech in a new language such as French. L2 learners often can identify words in writing or when spoken in isolation, and yet they may fail to recognize them in continuous speech. To segment language into words, it is not sufficient for non-native listeners to know the word forms uttered by the speaker; they must also know the phonological processes that apply within and across words and the factors that regulate the application of these processes.

One factor that has been shown to play an important role in speech segmentation is prosodic prominence. For example, from the age of 7.5 months, English-acquiring infants use accented syllables to identify word-initial boundaries (e.g., Jusczyk and Aslin 1995; Jusczyk, Houston, and Newsome 1999). Young infants are also sensitive to higher-level prosodic information: in English, they can detect disruptions in intonational phrases at the age of 4.5 months (e.g., Hirsh-Pasek et al. 1987; Jusczyk et al. 1992) and disruptions in phonological phrases at the age of 9 months (e.g., Kemler-Nelson et al. 1989; Gerken, Jusczyk, and Mandel 1994); and in French, they can detect phonological phrase boundaries at the age of 13 months (Christophe et al. 2003). Adults have similarly been shown to use both accentual cues (e.g., in English: Cutler and Butterfield 1992; McQueen, Norris, and Cutler 1994; Cooper, Cutler, and Wales 2002; Mattys 2004; in French: Banel and Bacri 1994) and higher-level prosodic cues (e.g., in French: Christophe et al. 2004; Welby 2006; in Korean: Kim 2004; Kim and Cho 2009; in Japanese: Warner, Otake, and Arai 2010) to segment speech into words.

By comparison, relatively few studies have examined non-native listeners’ use of prosodic information in word recognition and speech segmentation (e.g., Cooper et al. 2002; Sanders, Neville, and Woldorff 2002; Mettouchi, Lacheret-Dujour, Silber-Varod, and Izre’el, 2007; Tremblay 2008; Tyler and Cutler 2009; White, Melhorn, and Mattys 2010; Kim,

Broersma, and Cho to appear). Doing so becomes particularly interesting when the native and target languages differ not only in their prosodic structure, but also in the primary acoustic cues associated with word boundaries. In such cases, L2 learners must map these cues to the correct word boundaries before their segmentation of words from speech can be rapid and efficient. Failing to establish the correct mapping between these cues and word boundaries can, at best, make the word recognition process difficult, and at worst, impede L2 learners from making higher-level prosodic generalizations (e.g., Tremblay and Owens 2010; for discussion, see Carroll 2004).

The present study investigates whether English speakers at mid and high proficiencies in French can learn to use prosodic information for segmenting French speech into words, and how their use of this information unfolds as their proficiency in French increases. French and English differ both in their prosodic structure and in the primary acoustic cues associated with word boundaries. This pairing of native and target languages thus provides a crucial window for examining the processes that underlie L2 learners' extraction and use of prosodic cues in the speech signal. We will present the results of two word-monitoring experiments, one with natural stimuli and one with resynthesized stimuli, that shed further light on these processes.

We begin with an overview of the prosodic structure of French and the cues it provides for locating word boundaries; we then review research on the use of prosodic information in the segmentation of French and English speech, and make predictions for the L2 learners in the present study; finally, we present the method and results of our two experiments, and discuss their implications for understanding the development of L2 speech segmentation.

2. Prosodic Structure of French

Several models have been proposed to explain the structure of the intonational system of French (e.g., Fonagy 1980; Hirst and Di Cristo 1984; Mertens, 1987; Padeloup 1990; Di Cristo and Hirst 1993; Vaissière, 1997; Di Cristo, 2000; Post 2000; Jun and Fougeron 2000, 2002; Astésano 2001; Welby, 2003, 2006; Vassière and Michaud 2006; Astésano, Bard, and Turk 2007). Of particular importance for this study are the different prosodic cues that have been proposed to coincide with word edges in French. To the extent that different intonational models propose similar cues to word edges, they can make similar predictions for the segmentation task

that English listeners face when learning to segment French speech. However, we limit our review to proposals articulated within the autosegmental-metrical framework, namely Jun and Fougeron's (2002) and Welby's (2003, 2006) models (for a more complete literature review, see Jun and Fougeron 2000). We chose models within this framework for ease of cross-linguistic comparisons, as several autosegmental-metrical models have been proposed for explaining the intonational structure of various languages (Jun 2006).

Jun and Fougeron (2002) and Welby (2003, 2006) propose that the domain of prominence in French is the Accentual Phrase (AP), which corresponds roughly to Verluyten's (1982) Accentual Group and Di Cristo & Hirst's (1993) Rhythmic Unit but which does not always coincide with Nespor and Vogel's (1986) syntactically defined Phonological Phrase (PP) (see also Post 2000; Michelas and D'Imperio 2011). Each AP contains a pitch accent (*) at its right edge, as shown in (1). This pitch accent is bitonal, consisting of a low (L) tone and a high (H) tone, except in final position of declarative sentences, where only the L tone surfaces before a falling Intonational Phrase (IP) boundary (L%). In neutral (i.e., non-contrastive) prosodies, the final tone of the pitch accent is predictably anchored to the last (usually non-schwa) syllable of the last word in the AP; it thus coincides with a word-final boundary, and functions as an edge tone. The acoustic correlates of accented syllables include both an increase in duration and, in non-sentence-final position, an increase in fundamental frequency (F0), with the F0 peaking at the offset of the metrically prominent syllable (i.e., its vowel or its coda consonant; Rolland and Løevenbruck 2002; Welby and Løevenbruck 2006; see also Astésano, 2001).

- L H* L L%
- (1) *[[Il aime]_{AP} [le gâteau]_{AP}]_{IP}.*
 'He likes cake.'

Additionally, APs can have a bitonal phrase-initial accent (LHi) at their left edge, as illustrated in (2). The phrase-initial accent is structurally different from the pitch accent in that it is optional (in short APs, it is either absent or not fully realized), it is not associated with a metrically prominent syllable, and its L tone delineates the left edge of the AP (Jun and Fougeron 2002; Welby 2003, 2006). The Hi tone, when present, typically occurs on content

words rather than on function words, and its location varies as a function of the number and length of the content words inside the AP, with a tendency for it to occur on the first syllable of the first content word in the AP (Jun and Fougeron 2000, 2002; Welby 2003, 2006). The primary acoustic correlate of the Hi tone is an increase in F0 (e.g., Rolland and Løevenbruck 2002; Welby 2006; Welby and Løevenbruck 2006; see also Astésano et al. 2007). The phrase-initial accent may also be accompanied by strengthening of the onset consonant of the first content word in the AP (e.g., Padeloup 1990; Di Cristo 1998; Fougeron 2001). When an AP contains both an LH* pitch accent and an LHi phrase-initial accent, the F0 peak of the pitch accent is always higher than the F0 peak of the phrase-initial accent, despite any downdrift effects (e.g., Rolland and Løevenbruck 2002; Welby 2003, 2007). In short APs, it has been suggested that it is the L tone of the phrase-initial accent and the H* tone of the pitch accent that surface, creating a hybrid pitch rise (e.g., Jun and Fougeron 2000, 2002; Welby 2003, 2006; Welby and Løevenbruck 2006). This hybrid rise is exemplified in the first AP of (1).

- L Hi L H* L Hi L L%
- (2) *[[Il aime beaucoup]_{AP} [le chocolat noir]_{AP}]_{IP}.*
 ‘He very much likes dark chocolate.’

Given their alignment with (respectively) the right and left edges of APs, pitch accents and phrase-initial accents can provide cues to, respectively, right and left word boundaries in French. Let us now turn to studies showing that French listeners indeed exploit prosodic cues to pitch and phrase-initial accents when recognizing words in continuous speech.

3. Use of Prosodic Cues in the Segmentation of French and English Speech

Because pitch accents predictably fall on the last syllable of APs in French, they are reliable cues to phrase-final, and thus word-final, boundaries. A number of studies have indeed shown that native French listeners use prosodic cues to pitch accents for segmenting the speech stream into words. For example, Banel and Bacri (1994) found that French speakers who listened to phonemically ambiguous sequences (e.g., /bagaʒ/) were more likely to hear a single disyllabic word (e.g., *bagage* ‘luggage’) if the duration of the second syllable had been increased, and they

were more likely to hear two monosyllabic words (e.g., *bas gage* ‘low pledge’) if the duration of the first syllable had been increased, as it would be if a pitch accent was located on, respectively, the second or first syllable. Bagou, Fougeron, and Frauenfelder (2002) similarly showed that French listeners used both increased duration and F0 rise to segment an artificial language into words, with the latter yielding slightly more accurate segmentation than the former (see also Bagou and Frauenfelder 2006).

In a series of online experiments, Christophe et al. (2004) provided further evidence that phrase-final prosodic boundaries (and pitch accents) mediate lexical access in French. In particular, they examined whether native French listeners segment word-final boundaries at the offset of PPs (that coincided with APs). They found that monosyllabic words such as *chat* ‘cat’ were recognized more slowly when they were temporarily ambiguous at the phonemic level with a competitor word created between the monosyllabic word and the first syllable of the word following it (e.g., *chagrin* /ʃagʁɛ̃/ ‘heartache’ in [*d’un chat grincheux*]_{PP} ‘of a cranky cat’) than when they were not temporarily ambiguous with such a competitor (e.g., [*d’un chat drogué*]_{PP} ‘of a drugged cat’); however, if the monosyllabic word was at a PP- (and AP-) final boundary and thus received a pitch accent (e.g., [*le gros chat*]_{PP} [*grimpait aux arbres*]_{PP} ‘the big cat was climbing trees’), the target word was no longer recognized more slowly when it was temporarily ambiguous with a competitor than when it was not (e.g., [*le gros chat*]_{PP} [*dressait l’oreille*]_{PP} ‘the big cat was sticking up his ears’). These findings suggest that PP-final boundaries, which here coincided with AP-final boundaries and thus were marked with a pitch accent, acted as filter and constrained lexical access. Using a similar task, Michéas and D’Imperio (2010) replicated Christophe et al.’s findings for AP-final boundaries that did not coincide with PP-final boundaries (e.g., [*des pins*]_{AP} [*somptueux*]_{AP} ‘sumptuous pine trees’; competitor word: *pinson* ‘finch’), but for the two types of boundaries, they found an effect of prosody both when the target word overlapped with a potential disyllabic competitor and when it did not overlap with such a competitor. They attributed their results to the fact that their experiment did not reinforce lexical ambiguities—it did not contain true lexical ambiguities where, for example, the monosyllabic target word would have been embedded within a real disyllabic word (e.g., *chat* in *chapeau* ‘hat’).

Recent research has shown that native French listeners can also use phrase- (i.e., AP-) initial accents to identify word-initial boundaries and access words. Welby (2003, 2007) reported that native French listeners were more likely to hear nonce phonemic sequences (e.g., [melamōdin]) as two words (*mes lamondines*) than as one word (e.g., *mélamondines*) if the F0 of the second syllable had been increased. Spinelli, Welby, and Schaegis (2007) similarly found that French listeners could use prosodic information to discriminate between and identify lexical words such as *l'affiche* 'the poster' and *la fiche* 'the sheet': because the Hi tone in the phrase-initial accent is typically aligned with the first syllable of content words, it occurs on the syllable /la/ in *l'affiche* but on the syllable /fiʃ/ in *la fiche*. French listeners could use this acoustic information (among others) not only in offline tasks, but also in online ones, with target words being activated more when they matched the intended segmentation than when they did not. In a follow-up study, Spinelli et al. (2010) demonstrated that raising the F0 of /la/ in *la fiche* resulted in the greater selection and easier recognition of vowel-initial words (e.g., *affiche*) than if the F0 had not been manipulated. In their natural stimuli, the disyllabic sequences contained the L tone of the phrase-initial accent (aligned with the offset of /la/ in *la fiche* and with the onset of /la/ in *l'affiche*) and the H* tone of the pitch accent (aligned with the offset of the vowel in /fiʃ/), indicating that increasing the F0 immediately preceding or following the L tone resulted in more segmentation of word-initial boundaries. This suggests that the prosodic information in phrase-initial accents has an immediate effect on French listeners' word recognition and mediates lexical access.

Not all studies have shown such an effect, however. Bagou and Frauenfelder (2006) report that French listeners exposed to an artificial language benefit from an increase in F0 in word-initial syllables only when the word also contained an increase in both F0 and duration in its final syllable. This discrepancy may be due to the fact that the artificial language in Bagou and Frauenfelder (2006) did not contain function words, thus making it difficult to examine the effect of phrase-initial F0 cues, which tend to occur at the boundaries between function and content words (for discussion, see Welby 2007; Spinelli et al. 2010). Alternatively, these results might indicate that pitch accents are more reliable cues to word-final boundaries than phrase-initial accents are to word-initial boundaries in French.

English speakers also use prosodic information to recognize words in continuous speech. Unlike French, English has two levels of prominence: one at the lexical level (i.e., stress, whose placement within words varies based on factors such as syllable weight, word class, and affixation; for discussion, see Halle and Vergnaud 1987; Hammond 1995), and one at the phrase level (i.e., accent, whose placement within sentences interacts with factors such as information structure; for discussion, see Ladd 1996). Pitch accents in English are generally aligned with stressed syllables (e.g., Beckman and Elam 1997), which, statistically, tend to be word-initial rather than word-final, especially in disyllabic and trisyllabic nouns (e.g., Cutler and Carter, 1987; Clopper 2002). Hence, accented syllables can potentially provide a reliable cue to word-initial boundaries in English.

Many studies have shown that native English listeners can indeed use stress to locate word-initial boundaries in continuous speech. In a juncture perception task where stimuli were barely audible, Cutler and Butterfield (1992) found that English listeners tend to hear word-initial boundaries at the onset of stressed syllables (e.g., *achieve her way instead* was perceived as *a cheaper way to stay*). Similar findings were reported in online studies: McQueen et al. (1994) showed that English words were more easily detected when they were the second syllable of nonsense iambic sequences (e.g., *mess* in /nə'mɛs/) than when they were the first syllable of nonsense trochaic sequences (e.g., *mess* in /'mɛstəm/), because a word-initial boundary can be detected at the onset of the stressed syllable in the former, but no word-final boundary is detected at the offset of the stressed syllable in the latter. Mattys (2004) also found that when stimuli were presented with background noise, stress-initial disyllabic primes embedded in quadrisyllabic nonsense sequences (e.g., /'diplə'nəʊtə/) facilitated the activation of matching trisyllabic targets (e.g., *notable*) more than did stress-final disyllabic embedded primes (e.g., /'dipləmɪ'kæ/) for matching trisyllabic targets (e.g., *mechanic*). These findings suggest that prosodic information may be particularly useful for word segmentation in English, especially when lexical information is degraded or absent (see also Mattys, White, and Melhorn 2005).

In clear speech contexts, however, prosodic information does not appear fully constrain lexical access in English, perhaps because prosodic cues are highly redundant with segmental cues (in particular, vowel reduction). For example, Cutler and Clifton (1984) and Fear, Cutler, and Butterfield (1995) found that stress had a much smaller effect on word recognition in the

absence of vowel reduction; Cutler (1986) reported that segmentally near-identical word pairs differing in stress placement (e.g., *forbear*) primed each other; and Small, Simon, and Goldberg (1988) found that mis-stressing had little effect on the recognition of noun-verb homographs (e.g., *convert*). These findings contrast with those of studies in other languages such as Dutch and Spanish, where prosodic cues, which are not as redundant with segmental cues, have been found to constrain lexical access even in clear speech contexts (e.g., in Dutch: Cutler and Donselaar 2001; Donselaar, Koster, and Cutler 2005; in Spanish: Soto-Faraco, Sebastián-Gallés, and Cutler 2001).

The primary acoustic correlates of prosodic prominence in English are duration, F0, and intensity (e.g., Lieberman 1960; Beckman 1986), and the importance of each correlate depends in part on the location of the stressed syllable in the word, with F0 rise being a strong cue to word-initial stress but with increased duration being a stronger cue to stress in non-word-initial positions (Tremblay and Owens 2010). The close relationship between prosodic information and vowel reduction in English further confirms the importance of duration as a cue to prominence. In French, on the other hand, both F0 rise and increased duration are strong acoustic correlates of pitch accents in word-final position. This makes the prediction, then, that English listeners should use duration increase, but not F0 rise, as a cue to word-final boundaries in French, at least at low levels of proficiency.

Using artificial language-learning experiments, Tyler and Cutler (2009) found that both English and French listeners used duration to locate word-final boundaries. They attributed these findings to the universality of duration as a cue to word-final boundaries across languages rather than to the relationship between prosodic information and vowel reduction in English (see also Vaissière 1983; Hayes 1995; Saffran, Newport, and Aslin 1996). Tyler and Cutler also found, however, that only the French listeners used F0 rise to locate word-final boundaries; English listeners instead used F0 rise to locate word-initial boundaries. In a similar study, Kim et al. (to appear) reported that both Dutch and Korean listeners could use duration as a cue to word-final boundaries, but only the Korean listeners initially used F0 rise as a cue to word-final boundaries, and the Dutch listeners were able to do so only after sufficient exposure to the artificial language. Dutch and English have similar prosodic structures, in that they have two levels of prominence (one at the word level and one at the phrase level), and accented syllables tend to be word-initial

in both languages. Conversely, the prosodic structure of Korean is similar to that of French, with prominence being at the phrase level and with a H(*) tone coinciding with word-final boundaries in both languages. The findings of Tyler and Cutler (2009) and Kim et al. (to appear) thus suggest that listeners use F0 cues to word boundaries as a function of the mapping between these cues and word boundaries in the native language.

Since the alignment of F0 rise with word edges varies cross-linguistically, L2 learners must learn the particular prosodic configuration of the target language in order to use this information successfully in speech segmentation. For the present study, this means that English L2 learners of French must not only learn to parse accented syllables as word-final (rather than word-initial), but also to use F0 rise as a cue to word-final boundaries. This can potentially be difficult given that the F0 rise in phrase-initial accents, when present, can also provide a cue to word-initial boundaries in French. Hence, in APs containing both a pitch accent and a phrase-initial accent, English L2 learners of French (and native French listeners) must also distinguish pitch accents from phrase-initial accents and use them to locate, respectively, word-final and word-initial boundaries in continuous speech.

Few studies have examined L2 learners' use of prosodic information in word recognition and speech segmentation. Cooper et al. (2002) showed that when hearing phonemically identical stressed and unstressed fragments (e.g., /'kæm/ vs. /kæm/), Dutch L2 learners of English were in fact better than native English listeners at identifying the word to which the fragment belonged (e.g., respectively, *campus* and *campaign*). These findings were attributed to the fact that prosodic information is less correlated with segmental information in Dutch than it is in English. Using a similar task, Tremblay (2008) found that French L2 learners of English had much more difficulty than native English listeners in identifying the word to which the fragment belonged. While prosodic information can signal word boundaries in French, for multisyllabic words that begin (or end) with phonemically identical syllables, it does not distinguish between different lexical competitors, thus making it difficult for French L2 learners of English to use this information in lexical access.

Non-native listeners can use prosodic information to segment speech into words, but like native listeners, their reliance on this information depends on the extent to which lexical information is available. Sanders et al. (2002) showed that Spanish and Japanese L2 learners of

English had less difficulty detecting phonemes when they occurred in the onset of word-initial stressed syllables than when they occurred in the onset of word-medial stressed syllables. Importantly, this effect was larger when lexical information was missing from the stimuli than when it was present in the stimuli. On the other hand, White et al. (2010) found that Hungarian L2 learners of English recognized trisyllabic words (e.g., *corridor*) more rapidly when the disyllabic prime they heard was preceded by a lexical word (e.g., *anythingcorri*) than when it was preceded by a non-word (e.g., *imoshingcorri*), irrespective of whether the target word and prime were stressed on the first or second syllable (e.g., *confusing* for *anythingconfu* and *imoshingconfu*). These L2 learners' non-reliance on prosodic information, which did not vary as a function of their proficiency in English, is somewhat surprising given that stress is word-initial in Hungarian and thus provides an excellent cue to word-initial boundaries.

The present study examines English listeners' use of prosodic information in the recognition of French words. More precisely, it investigates whether English L2 learners of French can learn to use prosodic cues to word-final boundaries for segmenting the speech stream into words, and how their use of prosodic cues changes as their proficiency in French increases. It does so in two experiments, one containing natural stimuli (Experiment 1) and one containing resynthesized stimuli (Experiment 2). Both experiments use an experimental paradigm adapted from Christophe et al. (2004) and Michelas and D'Imperio (2010), but instead of detecting words that were present in the sentence, the participants were asked to monitor words that were created at the phonemic level between a monosyllabic word and the syllable following it (e.g., *chalet* 'cabin' in *chat lépreux* 'leprous cat').¹ In one condition, the monosyllabic word received a pitch accent, and thus the target word crossed an AP boundary that did not coincide with a PP boundary (e.g., *[[Le chat]_{AP} [lépreux et légendaire]_{AP}]_{PP} s'endort doucement* 'The leprous and legendary cat is slowly falling asleep'); in the other condition, the monosyllabic word did not receive a pitch accent, and thus the target word was located within an AP (and PP) (e.g., *[[Le chat lépreux]_{AP}]_{PP} s'endort doucement* 'The leprous cat is slowly falling asleep'). Unlike some of the previous studies, lexical information was not degraded, but the participants were put under time pressure so as to see whether prosodic information would guide their speech segmentation. Since pitch accents fall on word-final syllables in French and phrase-final boundaries are aligned with word-final boundaries, if the participants use prosodic information in natural speech, they

should make fewer incorrect detections of the target word in the across-AP condition than in the within-AP condition. Note that since the participants were asked to monitor the same word in the two prosodic conditions, this experimental paradigm makes it possible to examine the effects of prosodic information independently of the effects of lexical information. In Experiment 1, the acoustic cues to pitch accents (and AP-final boundaries) were not manipulated explicitly, but the relationship between these cues in the stimuli and the participants' proportion of false alarms were examined.

4. Experiment 1

4.1. Method

4.1.1. Participants

Eleven native French listeners (mean age: 27.9, standard error: 1.2) and 28 native English speakers (mean age: 23.0, standard error: 0.4) at mid and high proficiencies in French participated in this study. They were undergraduate and graduate students at a Midwestern university. The participants had normal or corrected-to-normal vision and did not report hearing impairment. They received candy in return for their participation.

The L2 learners had completed at least four semesters of French at the time of the study, and most of them had little exposure to French before the onset of puberty. Their proficiency in French was identified with the help of a cloze (i.e., fill-in-the-blank) test independently shown to provide a reliable estimate of proficiency in French (Tremblay and Garrison 2010; Tremblay 2011). The participants were evenly divided into two proficiency groups (mid, high) on the basis of their cloze test scores. Most of the participants in the mid-level group were third-year undergraduate students in French, whereas most of the participants in the high-level group were graduate students in French (who teach first- and second-year French classes).

The participants also completed a language background questionnaire in which they specified relevant biographical information. For L2 learners, this information included their age of first exposure to French, their number of years of instruction on French, the number of months they spent in a French-speaking environment, and their percent weekly use of French. The L2 learners' cloze test scores and their biographical information are provided for each proficiency group in Table 1.

Table 1. *L2 Learners' Cloze Test Scores and Biographical Information, Experiment 1*

Group	Cloze (/45)	Age of first Exposure	Years of instruction	Months of immersion	% weekly use
Mid (n=14)	19.4 (0.9)	11.2 (0.7)	8.0 (0.7)	6.0 (5.1)	8.8 (1.8)
High (n=14)	32.7 (0.8)	12.5 (0.9)	11.4 (1.1)	15.4 (3.0)	27.7 (3.6)

Note. Mean (standard error)

A multivariate analysis of variance (MANOVA) conducted on the L2 learners' language background information, with proficiency level as between-subject variable, reveals a significant effect of proficiency ($F[4,22]=6.38, p<.001$), with the effect reaching significance for years of instruction on French ($F[1,25]=6.94, p<.014$) and percent weekly use of French ($F[1,25]=21.37, p<.001$), but not for months of residence in a French-speaking environment ($F[1,25]=3.19, p=.086$) or age of first exposure to French ($F[1,25]=1.36, p=.255$). The overall results of the MANOVA, and the significant differences between the two groups in their years of instruction on French and their percent weekly use of French, provide some validation of the cloze test scores.²

4.1.2. *Materials*

The participants heard sentences in which a target word was created at the phonemic level between a monosyllabic word and the first syllable of the disyllabic adjective following it (e.g., *chalet* 'cabin' in *chat lépreux* 'leprous cat').³ In the across-AP condition, the monosyllabic word received a pitch accent, and the target word crossed an AP boundary (e.g., $[[Le \underline{chat}]_{AP} [lépreux \text{ et } légendaire]_{AP}]_{PP} s'endort doucement$ 'The leprous and legendary cat is slowly falling asleep'). The first AP contained an LH* tonal pattern, with the L tone belonging to either a phrase-initial accent or a pitch accent and the H* tone belonging to a pitch accent.⁴ In the within-AP condition, the pitch accent instead fell on the last syllable of the post-nominal adjective (e.g., $[[Le \underline{chat} \underline{lépreux}]_{AP}]_{PP} s'endort doucement$ 'The leprous cat is slowly falling asleep'). The AP in this condition contained an LLH* tonal pattern, with the first L tone belonging to a phrase-initial accent and the LH* tones belonging to a pitch accent. The sentences in the two conditions shared the first three words (article, monosyllabic word, and adjective). In the across-AP

condition, an additional modifier was added in the second AP so that the prosodic boundary after the monosyllabic word would sound natural. The first four syllables of the two experimental conditions thus differed only in their prosody, with *chat* receiving an H* tone in the across-AP condition and an L tone in the within-AP condition. The F0 movement of the first four syllables of the stimuli in the two conditions is exemplified in Figure 1.

The experiment also included a control condition in which the target word was in the sentence; prosodically, this condition was identical to the within-AP condition (e.g., [*Le chalet suisse*]_{AP} *est agréable* ‘the Swiss cabin is pleasant’). The target words in the experimental and control sentences were all the subject of the sentence. A total of 36 (experimental and control) triplets were included in the task. The participants were assigned to one of three lists and saw each experimental or control item in only one condition (total: 12 items per condition). A complete list of the critical determiner-noun-adjective(-adjective) sequences is provided in the Appendix. The experimental and control items were interspersed with 72 distracter items, of which 10 were practice items. The target words in the distracter items were located in different syntactic positions (e.g., subject, object). Half of the trials in the experiment (including both the critical and distracter items) contained the target word, and half did not.

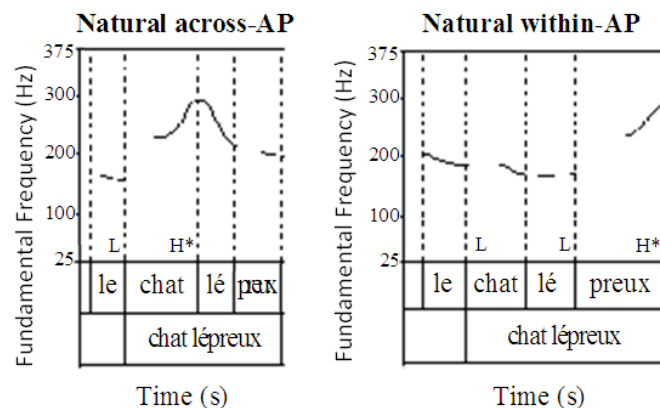


Figure 1. Example of natural stimuli (target: chalet ‘cabin’), Experiment 1

The auditory stimuli were recorded by a female native speaker of French from Bordeaux (France) using a Marantz PMD 750 solid state recorder and head-mounted condenser microphone. The speaker was trained to produce the stimuli such that an H* tone would fall on the monosyllabic noun in the across-AP condition but on the last syllable of the post-nominal

adjective in the within-AP and control conditions. In both experimental conditions, the peak F0 of the H* tone was aligned with the AP-final boundary. The H* tone produced on the monosyllabic noun in the across-AP condition was not followed by a pause so that the disyllabic target word could be erroneously detected.

The recordings were then normalized for intensity, and acoustic analyses of the first two syllables in the stimuli (e.g., *le chat* in the experimental conditions and *le cha-* in the control condition) were performed in Praat (Boersma & Weenink, 2007). For each syllable, the duration and intensity were measured, and for each vowel within these syllables, the average F0 and peak F0 were measured. Table 2 provides the duration and amplitude measurements for the syllables and the F0 measurements for the vowels.

Table 2. *Acoustic Measurements of Stimuli, Experiment 1*

		Across-AP	Within-AP	Control
Duration (ms)	Syllable 1	170 (6)	166 (6)	157 (6)
	Syllable 2	231 (9)	193 (6)	177 (5)
Intensity (dB)	Syllable 1	66.4 (0.3)	66.9 (0.4)	66.6 (0.4)
	Syllable 2	66.4 (0.2)	65.7 (0.3)	65.9 (0.4)
Mean F0 (Hz)	Vowel 1	197 (3)	189 (3)	192 (3)
	Vowel 2	253 (5)	188 (3)	192 (2)
Peak F0 (Hz)	Vowel 1	207 (4)	194 (3)	200 (4)
	Vowel 2	259 (7)	217 (5)	221 (4)

Note. Mean (standard error)

As can be seen from the acoustic measurements, Syllable 2 and Vowel 2 (e.g., *chat* and *cha-*) were acoustically more prominent in the across-AP condition than in the within-AP and control ones, whereas Syllable 1 and Vowel 1 (e.g., *le*) were similar in the three conditions. Paired-samples *t*-tests were performed on the acoustic measurements of the first two syllables in the three conditions, with the alpha level adjusted to .008 (Bonferroni correction for six comparisons—Syllable/Vowel 1 and Syllable/Vowel 2 in the across-AP, within-AP, and control conditions). The results of these statistical analyses are reported in Table 3.

Table 3. Paired-Samples *t*-Tests on Acoustic Analyses of Stimuli, Experiment 1

	Comparison	Paired-samples <i>t</i> -test	
Duration	Across-AP-Within AP	$t < 1 $	
	Syllable 1	Across-AP-Control	$t(35)=2.78, p<.009$
		Within-AP-Control	$t(35)=-2.27, p<.03$
	Syllable 2	Across-AP-Within AP	$t(35)=4.56, p<.001$
		Across-AP-Control	$t(35)=6.05, p<.001$
		Within-AP-Control	$t(35)=-4.7, p<.001$
Intensity	Across-AP-Within AP	$t(35)=-1.83, p<.075$	
	Syllable 1	Across-AP-Control	$t < 1 $
		Within-AP-Control	$t < 1 $
	Syllable 2	Across-AP-Within AP	$t(35)=2.72, p<.010$
		Across-AP-Control	$t(35)=1.35, p<.186$
		Within-AP-Control	$t < 1 $
Mean F0	Across-AP-Within AP	$t(35)=1.88, p<.068$	
	Syllable 1	Across-AP-Control	$t(35)=1.43, p<.162$
		Within-AP-Control	$t < 1 $
	Syllable 2	Across-AP-Within AP	$t(35)=14.64, p<.001$
		Across-AP-Control	$t(35)=16.52, p<.001$
		Within-AP-Control	$t(35)=1.93, p<.062$
Peak F0	Across-AP-Within AP	$t(35)=3.13, p<.004$	
	Syllable 1	Across-AP-Control	$t(35)=1.61, p<.116$
		Within-AP-Control	$t(35)=2.43, p<.020$
	Syllable 2	Across-AP-Within AP	$t(35)=5.42, p<.001$
		Across-AP-Control	$t(35)=4.84, p<.001$
		Within-AP-Control	$t < 1 $

For syllable duration, the statistical analyses revealed significant differences between the second syllables of all three conditions; for mean F0, they yielded a significant difference between the second vowels of the across-AP and within-AP conditions; and for peak F0, they

revealed a significant difference between the first vowels of the across-AP and within-AP conditions, and a significant difference between the second vowels of the across-AP and both the within-AP and control conditions. These results indicate that the acoustic cues distinguishing the two experimental conditions were signaled mostly by Syllable/Vowel 2, which had a significantly longer duration, higher mean F0, and higher peak F0 in the across-AP condition than in the within-AP condition.⁵

Another prosodic cue that could influence the participants' false alarm rates is the possible presence of a Hi tone from a phrase-initial accent at the left edge of APs, here potentially on the first syllable of the adjective in the across-AP condition and on the monosyllabic noun in the within-AP condition. The intonational patterns of the stimuli were thus examined closely to determine whether such a Hi tone had unintentionally been produced. For the stimuli in the across-AP condition, this inspection revealed a steep F0 fall on the first syllable of the adjective, and no evidence of a second rise in the adjective. Phrase-initial accents in French typically begin with an L tone, and the fall following the Hi tone is usually gradual rather than steep (Jun & Fougeron 2002). This suggests that this F0 fall on the first syllable of the adjective in the across-AP condition was due to the H* tone from the pitch accent on the previous syllable rather than to a Hi tone from a phrase-initial accent. For the stimuli in the within-AP condition, the mean F0 of the vowel in the monosyllabic noun was similar to that of the vowel in the article preceding it (as shown in Table 2). This suggests that the monosyllabic word in the within-AP condition did not contain a Hi tone from a phrase-initial accent. We are therefore confident that the main cue to prosodic boundaries in our stimuli was the H* tone from the pitch accent on the monosyllabic noun in the across-AP condition.

The present experimental design can potentially introduce a confounded variable, that of speech rate: the presence of a pitch accent on the monosyllabic noun in the across-AP condition can result in slower speech rate, potentially leading the participants to have fewer false alarms in the across-AP condition than in the within-AP one. To determine whether this was indeed the case, we measured the duration of the noun-adjective sequences in the across-AP and within-AP conditions. These additional analyses revealed that the noun-adjective sequences were in fact shorter in the across-AP condition (M: 667 ms., SD: 14) than in the within-AP condition (M: 701 ms., SD: 13), a difference which is statistically significant ($t[35] = -2.71, p < .01$). This difference

was likely due to the fact that the two multisyllabic adjectives in the second AP had to be articulated reasonably fast in order to be contained within the same AP. Hence, if speech rate influenced the results, it would likely be in the opposite direction to what is predicted from the experimental manipulation.

4.1.3. *Procedures*

The experiment was administered with E-Prime (Psychology Software Tools, Inc.; Schneider, Eschman, and Zuccolotto 2002). In each trial, the participants saw a target word printed in the center of the computer display. Five-hundred milliseconds after the onset of the visual presentation of the word, they heard (through headphones) a sentence containing or not containing that word. They were asked to press “o” (for *oui* ‘yes’) if they heard the word in the display and do nothing if they did not hear the word. Their accuracy rates were measured. In order to increase the likelihood that the participants would incorrectly detect the target words that were not in the sentences in the experimental conditions, the participants were put under time pressure when completing the task. Fourteen of the distracter items that contained the target word were followed by a reminder screen that the participants should try to respond faster. The order of item type (e.g., experimental, control, distracter) was the same for all the participants, but the particular test item appearing under each type was randomized across participants.

4.1.4. *Data analysis and predictions*

Two experimental triplets were excluded, because the wrong recordings had inadvertently been used as stimuli in one of the two experimental conditions. This resulted in the exclusion of 5.5% of the data.

The results will be reported as accuracy rates for the control condition and as false alarm rates for the experimental conditions. Since the former were at ceiling, no statistical analyses were performed on them. For the latter, linear mixed models were performed in SPSS, with prosodic information (across AP, within AP) as fixed within-subject variable and with participants and test items as random variables. The native and non-native listeners’ data were analyzed separately due to the uneven number of participants in each group. For the non-native group, proficiency (mid, high) also entered the model as fixed between-subject variable. To

examine the relationship between the participants' ability to use prosodic information and the acoustic cues in the stimuli, linear regression analyses were also performed on the participants' proportion of false alarms with these cues as predictors.

If the participants use prosodic information to segment French speech into words, they should show significantly lower false-alarm rates in the across-AP condition than in the within-AP condition. Given the findings of previous studies (e.g., Tyler and Cutler 2009), we might also expect to find significant relationships between the native French listeners' proportion of false alarms and the average F0 in the critical monosyllabic noun, and between the English listeners' proportion of false alarms and the duration of the critical monosyllabic noun. Note that the English listeners' use of duration as a cue to word-final boundaries would not necessarily indicate that they perceive the accented syllables as prominent, since duration is a cue to word-final boundaries (irrespective of stress placement) in English; by contrast, given that F0 rises and intensity increases do not signal word-final boundaries in English, the English listeners' use of these cues would suggest that they indeed perceive the accented syllables as prominent.

4.2. Results

The proportion of correct responses in the control condition indicated that all three groups were successful at detecting the target word when it was in the stimuli (mid-level L2 learners: .95, SD: .09; high-level L2 learners: .95, SD: .08; native listeners: .98, SD: .04). This suggests that the L2 learners' listening skills were sufficiently good to detect French words in continuous speech.

Figure 2 shows the native and non-native listeners' mean proportion of false alarms (and standard errors)—that is, their proportion of detection of the target word (e.g., *chalet*)—in the across-AP and within-AP conditions. Recall that this target word was not in the sentences they heard, but was created at the phonemic level between a monosyllabic noun and the first syllable of the following disyllabic adjective (e.g., *chat lépreux*). As can be seen from the results in Figure 2, the participants incorrectly detected the target word fewer times in the across-AP condition than in the within-AP condition, and this difference tended to increase with improved proficiency in French. The L2 learners' false alarms, especially those of the mid-level L2

learners, might have been exacerbated by the time pressure that was imposed on them, which was necessary to ensure they would incorrectly detect the target word.

A linear mixed model conducted on native speakers' proportion of alarm rates revealed a significant effect of prosodic information ($F[1,236]=14.51, p<.001$). A similar linear mixed model on L2 learners' proportion of false alarms also revealed a significant effect of prosodic information ($F[1,605]=19.67, p<.001$) and a significant effect of proficiency ($F[1,26]=5.97, p<.022$), but no interaction between prosodic information and proficiency ($F[1,605]=1.92, p=.167$), indicating that the high-level L2 learners did not show a significantly larger effect of prosodic information than the mid-level L2 learners. These results suggest that both the native and non-native French listeners used prosodic cues in the predicted direction for detecting word-final boundaries.

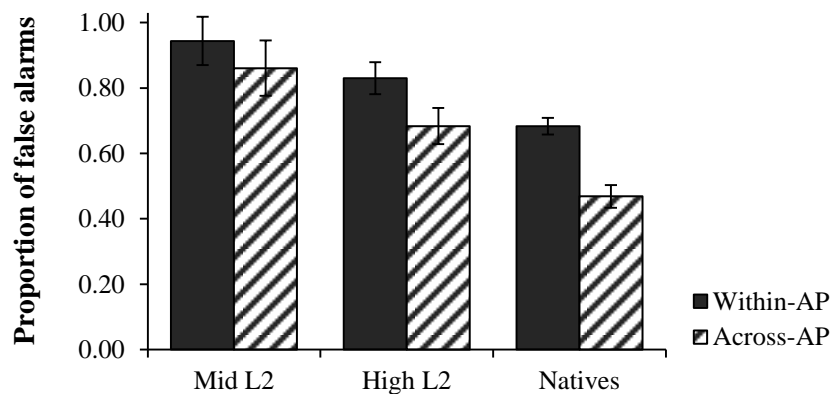


Figure 2. *Proportion of false alarms in the experimental conditions, Experiment 1*

In order to investigate the relationship between the participants' proportion of false alarms and the acoustic cues in the monosyllabic nouns, three linear regressions were performed on the proportion of false alarms, with acoustic cue (duration and intensity of the monosyllabic noun, and mean F0 of the vowel in the monosyllabic noun), group (mid L2, high L2, natives), and the interaction between the prosodic cue and group as predictors.⁶ For the three models, the only predictor that reached significance is the acoustic cue \times group interaction (duration: $r^2=.31, p<.001$; intensity: $r^2=.21, p<.001$; F0: $r^2=.28, p<.001$). Subsequent linear regressions were thus computed separately for each group, with duration, intensity, and F0 values entering the models in a stepwise fashion. These linear regressions are plotted in Figure 3.

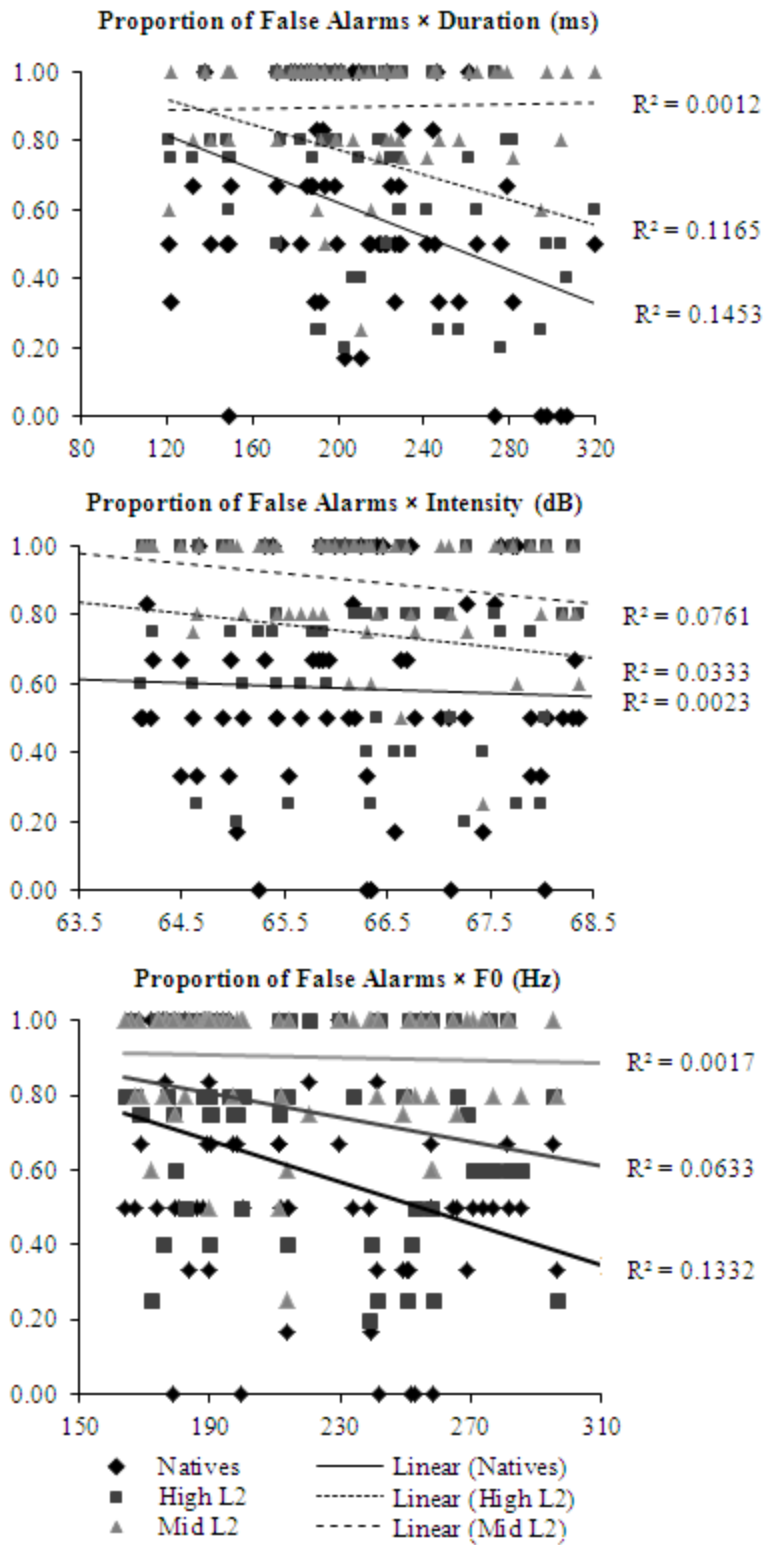


Figure 3. Relationship between the false alarm rates and acoustic cues, Experiment 1

These analyses yielded significant relationships between the native listeners' and high-level L2 learners' proportions of false alarms and the duration values (respectively, $r^2=.15$, $p<.001$; $r^2=.12$, $p<.004$), between the mid-level L2 learners' proportions of false alarms and the intensity values ($r^2=.08$, $p<.023$), and between the native listeners' proportion of false alarms and the F0 values ($r^2=.13$, $p<.002$).

These results suggest that whereas the native French listeners used both duration increase and F0 rise to identify word-final boundaries in continuous speech, the high-level L2 learners used only duration increase, and the mid-level L2 learners used only intensity increase. It is unclear why the mid-level L2 learners did not also rely on duration, given that it is correlated with prominence in English and it is an important cue to word-final boundaries cross-linguistically. One possibility is that they had poorer listening skills than the high-level L2 learners, and thus had more difficulty detecting syllable boundaries in continuous speech. This could have led them to rely on intensity as a cue to accented syllables rather than on duration as a cue to word-final boundaries (in this case, only the latter required listeners to detect syllable boundaries).

4.3. *Summary*

The results of Experiment 1 showed that the L2 learners used prosodic information to recognize words in continuous speech, indicating that they were able to associate prosodic cues with word-final boundaries in French. Yet, unlike native French listeners, they did not appear to rely on F0 rise to detect word-final syllables. Because both duration and intensity cues coincided with F0 cues in our stimuli, the L2 learners did not need to rely on F0 rise to hear word-final boundaries in French. One might argue on the basis of our results, then, that the high-level L2 learners did not necessarily perceive French accented syllables as such. Since duration is a reliable cue to word-final boundaries in English, irrespective of whether the word-final syllable is stressed, the high-level L2 learners might have heard the accented monosyllabic noun as a word-final syllable simply because of its longer duration. Notice that this cannot be true of the mid-level L2 learners, however, as increased intensity is not a cue to word-final boundaries in English. This means that the mid-level L2 learners must have perceived the accented

monosyllabic nouns as prominent. Since these learners were clearly less advanced in French than the high-level L2 learners (see Table 1), we believe it is unlikely that the high-level L2 learners did not perceive the accented syllables as prominent. What both groups have yet to accomplish is to use the correct mapping between F0 rise and prosodic prominence when the latter is word-final.

Experiment 1 only provides indirect evidence that English listeners cannot use F0 rise to locate word-final boundaries in French, however. It remains to be seen whether L2 learners would be able to use F0 rise in resynthesized stimuli where it is the only cue to word-final boundaries. In the absence of other cues, it is possible that L2 learners' speech segmentation would benefit from F0 rise. On the other hand, since increased duration and F0 rise both characterize AP-final syllables in French, L2 learners may not ever attend to F0 as a cue to word-final boundaries, instead relying on duration. Experiment 2 sheds further light on these issues by examining whether English L2 learners of French can use F0 cues irrespective of duration cues. The stimuli from the across-AP and within-AP conditions of Experiment 1 were manipulated such that the pitch contour of the first four syllables (e.g., *le chat lépreux*) in the across-AP and within-AP conditions were swapped. This made it possible to examine the effect of F0 rise independently of duration increase.

5. Experiment 2

5.1. Method

5.1.1. Participants

Twelve native French listeners (mean age: 29.2, standard error: 3.7) and 24 native English speakers (mean age: 21.6, SD: 2.7) at mid and high proficiencies in French and participated in this study. They were undergraduate and graduate students at a Midwestern university. The participants had normal or corrected-to-normal vision and did not report hearing impairment. They received financial compensation in return for their participation.

The L2 learners had also completed at least four semesters of French at the time of the study, and most of them had little exposure to French before the onset of puberty. Their proficiency in French was identified with the help of the same cloze test as in Experiment 1. The participants were evenly divided into two proficiency groups (mid, high) on the basis of their

cloze test scores. Again, most of the participants in the mid-level group were third-year undergraduate students in French, whereas most of the participants in the high-level group were graduate students in French (who teach first- and second-year French classes).

The participants also completed a language background questionnaire in which they specified relevant biographical information. For L2 learners, this information again included their age of first exposure to French, their number of years of instruction on French, the number of months they spent in a French-speaking environment, and their percent weekly use of French. The L2 learners' cloze test scores and their biographical information are provided for each proficiency group in Table 4.

Table 4. *L2 Learners' cloze test scores and biographical information, Experiment 2*

Group	Cloze (/45)	Age of first exposure	Years of instruction	Months of immersion	% weekly use
Mid (n=12)	20.1 (1.3)	9.6 (1.3)	7.6 (0.7)	1.9 (0.8)	7.0 (1.1)
High (n=12)	29.6 (1.1)	10.1 (0.7)	9.4 (0.9)	10.8 (3.1)	15.8 (2.4)

Note. Mean (standard error)

A MANOVA conducted on the L2 learners' language background information, with proficiency level as between-subject variable, reveals a significant effect of proficiency ($F[4,18]=5.73, p<.004$), with the effect reaching significance for months of residence in a French-speaking environment ($F[1,21]=8.65, p<.008$) and percent weekly use of French ($F[1,21]=10.54, p<.004$). The effect did not reach significance for years of instruction on French ($F[1,21]=2.27, p=.147$) or age of first exposure to French ($F<1$). These results again indicate that the cloze test results were more or less aligned with the L2 learners' language background information.

Two other MANOVAs were conducted on the cloze test scores and language background information of the L2 learners in Experiments 1 and 2 (one for the mid-level L2 learners and one for the high-level L2 learners), with experiment as between-subject variable, to determine whether the mid- and high-level participants in the two experiments had significantly different proficiencies. For these analyses, the alpha level was adjusted to .025 (Bonferroni correction for

two comparisons—one for the mid-level L2 learners and one for the high-level L2 learners). For the mid-level learners, the analysis did not reach significance ($F < 1$). For the high-level learners, the analysis revealed a marginally significant difference between the two groups ($F[5,19]=2.69$, $p=.055$), with the effect being marginally significant for the L2 learners' cloze test scores ($F[1,23]=3.79$, $p=.065$) and age of first exposure to French ($F[1,23]=4.49$, $p=.046$) and significant only for their percent weekly use of French ($F[1,23]=7.06$, $p < .014$). These results indicate that whereas the mid-level L2 learners in the two experiments were at similar proficiencies, the high-level L2 learners in Experiment 2 used French less often than those in Experiment 1, and they showed a tendency toward having a lower proficiency than those in Experiment 1, even though they had first been exposed to French at a slightly earlier age. These differences will be considered when appropriate in the interpretation of the results.

5.1.2. *Materials*

The participants again heard sentences in which a target word was created at the phonemic level between a monosyllabic word and the first syllable of the disyllabic adjective following it (e.g., *chalet* 'cabin' in *chat lépreux* 'leprous cat'). As in Experiment 1, in the natural across-AP condition, the monosyllabic word received a pitch accent, and the target word crossed an AP boundary (e.g., $[[Le \textit{chat}]_{AP} [\textit{lépreux et légendaire}]_{AP}]_{PPS}$ 'endort doucement' 'The leprous and legendary cat is slowly falling asleep'). The first AP contained an LH* tonal pattern, with the L tone belonging to either a phrase-initial accent or a pitch accent and the H* tone belonging to a pitch accent. In the within-AP condition, the pitch accent instead fell on the last syllable of the post-nominal adjective (e.g., $[[Le \textit{chat lépreux}]_{AP}]_{PPS}$ 'endort doucement' 'The leprous cat is slowly falling asleep'). The AP in this condition contained an LLH* tonal pattern, with the first L tone belonging to a phrase-initial accent and the LH* tones belonging to a pitch accent. Two additional conditions were created by resynthesizing the speech stimuli such that the F0 contour of the first four syllables in the stimuli would be swapped between the two experimental conditions. The first four syllables of the resynthesized across-AP sentences thus contained the F0 contour of the corresponding syllables in the within-AP condition, and the first four syllables of the resynthesized within-AP sentences contained the F0 contour of the corresponding syllables in the across-AP condition. This manipulation resulted in stimuli that

exhibited a four-way contrast in the cues they provided to word-final boundaries: (i) the natural within-AP stimuli did not contain F0 or durational cues to word-final boundaries (as compared to the other conditions); (ii) the resynthesized within-AP stimuli contained F0 cues to word-final boundaries; (iii) the resynthesized across-AP stimuli contained durational cues to word-final boundaries; and (iv) the natural across-AP stimuli contained both F0 and durational cues to word-final boundaries. Since the participants were at ceiling on the control condition in Experiment 1, Experiment 2 included only the natural and resynthesized across-AP and within-AP conditions.

Thirty-two of the 36 test items from Experiment 1 were used in Experiment 2. The participants were assigned to one of four lists and saw each experimental item in only one condition (total: 8 items per condition). The experimental items were interspersed with 45 distracter items from Experiment 1, of which 5 were practice items. Half of the trials in the experiment (including both experimental and distracter items) contained the target word, and half did not.⁷ The experimental stimuli used from Experiment 1 were resynthesized using close-copy stylization (e.g., Pijper 1983). The first four syllables of the experimental items were divided into 20 segments each, and the average F0 of each segment was extracted. The existing pitch points in each segment were then dragged vertically using the Pitch Synchronous OverLap-Add (PSOLA) method in Praat (Boersma and Weenink 2004) so that they would approximate the value of the extracted average in the corresponding segment of the opposite condition. After the initial resynthesis, the pitch contour of the natural and resynthesized conditions were examined very closely, and resynthesized contours that were judged not to be sufficiently similar to the natural contours of the opposite condition were altered so that they would approximate them. Once the contours were judged to be satisfactory, a stop Hann-band filter from 500 to 1000 Hertz with a smoothing of 100 Hertz was applied to all the stimuli to mask the occasionally robotic sound that resulted from the F0 manipulation. This filter did not significantly affect the segmental quality of the stimuli. Figure 4 shows an example of natural and resynthesized stimulus in the across-AP and within-AP conditions.

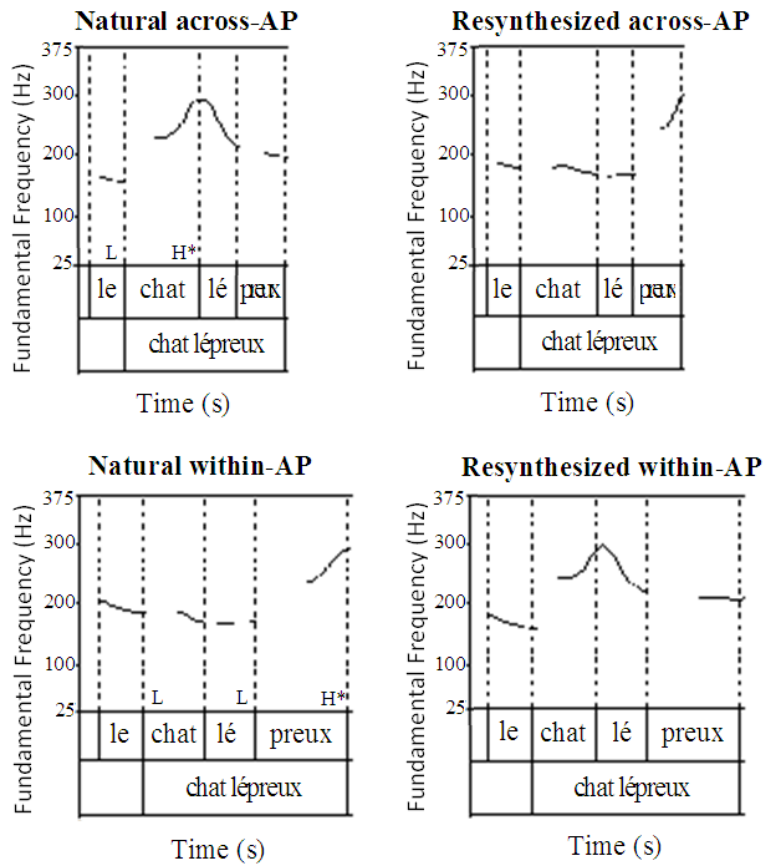


Figure 4. Example of natural and resynthesized stimuli (target: chalet 'cabin'), Experiment 2

Again, acoustic analyses of the first two syllables in the stimuli (e.g., *le chat* in the experimental conditions) were performed in Praat (Boersma & Weenink, 2007). For each syllable, the duration and intensity were measured, and for each vowel within these syllables, the average and peak F0 were measured. Table 5 provides the duration and amplitude measurements for the syllables and the F0 measurements for the vowels in these syllables.⁸

Table 5. *Acoustic Measurements of Stimuli, Experiment 2*

		Natural		Resynthesized	
		Across-AP	Within-AP	Across-AP	Within-AP
Duration (ms)	Syllable 1	171 (6)	166 (6)	171 (6)	166 (6)
	Syllable 2	234 (9)	193 (5)	234 (9)	193 (5)
Intensity (dB)	Syllable 1	66.2 (0.3)	66.8 (0.4)	66 (0.3)	66.6 (0.4)
	Syllable 2	65 (0.3)	64.8 (0.2)	63 (0.3)	64.5 (0.3)
Mean F0 (Hz)	Vowel 1	196 (3)	188 (3)	190 (2)	195 (3)
	Vowel 2	255 (4)	186 (3)	188 (3)	255 (5)
Peak F0 (Hz)	Vowel 1	203 (4)	192 (3)	194 (2)	199 (3)
	Vowel 2	280 (5)	190 (3)	194 (3)	275 (5)

Note. Mean (standard error)

Repeated-measures ANOVAs were performed on the acoustic measurements of the first two syllables in the four conditions, with prosody (across-AP, within-AP) and resynthesis (no, yes) as within-item variables. In these analyses, the alpha level was adjusted to .025 (Bonferroni correction for two comparisons—Syllable/Vowel 1 and Syllable/Vowel 2). For duration, the analyses revealed a significant effect of prosody only on Syllable 2 ($F[1,31]=22.18$, $p<.001$; all other F 's <1); for the intensity of Syllable 1, the analyses revealed a significant effect of resynthesis ($F[1,31]=31.58$, $p<.001$; prosody: $F[1,31]=4.225$, $p=.048$; prosody \times resynthesis: $F<1$), and for the intensity of Syllable 2, they revealed a significant effect of resynthesis ($F[1,31]=87.6$, $p<.001$) and a significant prosody \times resynthesis interaction ($F[1,31]=52.92$, $p<.001$; prosody: $F[1,31]=2.83$, $p=.102$); for the mean F0 of Vowel 1, the analyses revealed a marginally significant prosody \times resynthesis interaction ($F[1,31]=4.8$, $p<.036$; other F 's <1), and for the mean F0 of Vowel 2, they revealed a significant prosody \times resynthesis interaction ($F[1,31]=262.02$, $p<.001$; other F 's <1); finally, for the peak F0 of Vowel 1, the analyses revealed a significant prosody \times resynthesis interaction ($F[1,31]=6.86$, $p<.014$; prosody: $F[1,31]=1.08$, $p<.307$; resynthesis: $F<1$), and for the peak F0 of Vowel 2, they revealed a significant effect of prosody ($F[1,31]=16.07$, $p<.001$) and a significant prosody \times resynthesis interaction ($F[1,31]=312.92$, $p<.001$; resynthesis: $F<1$).

Subsequent paired-samples *t*-tests were performed on the intensity and F0 values for which significant interactions were found. For these comparisons, the alpha level was adjusted to 0.008 (Bonferroni correction on alpha 0.025 for four post-hoc comparisons—the natural vs. resynthesized across-AP conditions, the natural vs. resynthesized within-AP conditions, the natural across-AP vs. within-AP conditions, and the resynthesized across-AP vs. within-AP conditions). The results, presented in Table 6, revealed effects of prosody and resynthesis on the mean F0 and peak F0 values of Syllable 2 for all pairs, an effect of resynthesis on the peak F0 value of Syllable 1 and on the intensity value of Syllable 2 in the across-AP conditions, and an effect of prosody on the intensity values of Syllable 2 in the resynthesized conditions.⁹ These analyses indicate that both duration and F0 distinguished the naturally recorded across-AP and within-AP conditions, but only intensity and F0 distinguished the natural conditions from the resynthesized ones, and this acoustic information was signaled largely by Syllable/Vowel 2.

Table 6. *Paired-Samples t-Tests on Acoustic Analyses of Stimuli, Experiment 2*

	Natural	Resynthesized	
Intensity, Syllable 2	Across-AP	Across-AP	$t(31)=12.66, p<.001$
	Within-AP	Within-AP	$t(31)= -1.55, p=.131$
	$t< 1 $	$t(31)= -3.45, p<.002$	
Mean F0, Vowel 1	Across-AP	Across-AP	$t(31)=2.68, p=.012$
	Within-AP	Within-AP	$t(31)= -1.62, p=.116$
	$t(31)=1.73, p=.094$	$t(31)= -2.23, p=.033$	
Mean F0, Vowel 2	Across-AP	Across-AP	$t(31)=15.83, p<.001$
	Within-AP	Within-AP	$t(31)= -15.87, p<.001$
	$t(31)=16.85, p<.001$	$t(31)= -14.87, p<.001$	
Peak F0, Vowel 1	Across-AP	Across-AP	$t(31)=3.65, p<.001$
	Within-AP	Within-AP	$t(31)= -1.56, p=.129$
	$t(31)=2.11, p=.043$	$t(31)= -2.33, p=.026$	
Peak F0, Vowel 2	Across-AP	Across-AP	$t(31)=16.74, p<.001$
	Within-AP	Within-AP	$t(31)= -18.13, p<.001$
	$t(31)=16.84, p<.001$	$t(31)= -17.68, p<.001$	

5.1.3. *Procedures*

The procedures used in Experiment 2 were exactly the same as those used in Experiment 1.

5.1.4. *Data analysis and predictions*

Again, the results will be reported as false alarm rates. Linear mixed models were performed in SPSS, with prosodic information (across AP, within AP) and resynthesis (no, yes) as fixed within-subject variables and with participants and test items as random variables. The native and non-native listeners' data were analyzed separately due to the uneven number of participants in each group. For the non-native group, proficiency also entered the model as fixed between-subject variable.

If the participants parse accented syllables as word-final, they should show fewer false alarms in the across-AP condition than in the within-AP condition, as they did in Experiment 1. Furthermore, if they are sensitive to F0 rise as a cue to word-final boundaries, they should show fewer false alarms in the resynthesized within-AP condition, where the first four syllables contained the pitch contour of the across-AP condition, than in the natural within-AP condition. Since the resynthesized across-AP condition contained durational cues to word-final boundaries, they may not generate higher false alarms than the natural across-AP condition, which contained both durational and F0 cues to word-final boundaries. Note that since the two natural conditions were not resynthesized, they may generate higher accuracy rates (thus fewer false alarms) than the resynthesized conditions. Finding an interaction between prosodic information and resynthesis, with fewer false alarms in the resynthesized within-AP condition than in the natural one, is thus crucial for the resynthesis manipulation to demonstrate the predicted effect of F0 rise on speech segmentation.

5.2. *Results*

Figure 5 presents the native and non-native listeners' proportion of false alarms in each of the conditions. As can be seen from these results, the native listeners and high-level L2 learners incorrectly detected the target word fewer times in the across-AP condition than in the within-AP

condition, but the mid-level L2 learners showed the reverse pattern. Furthermore, only the native listeners showed an effect of resynthesis, and they did so only in the within-AP condition.

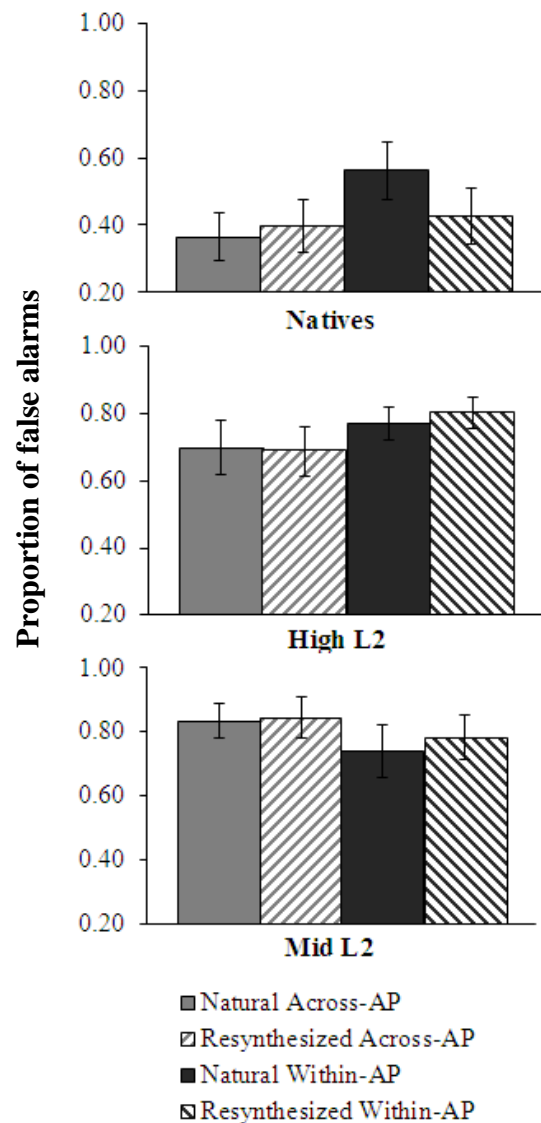


Figure 5. *Proportion of false alarms in the experimental conditions, Experiment 2*

Linear mixed models on the native listeners' false alarm rates revealed a significant effect of prosody ($F[1,369]=6.5, p<.011$) and a marginally significant interaction between prosody and resynthesis ($F[1,369]=3.4, p<.065$; resynthesis: $F[1,369]=1.3, p<.248$). Separate linear mixed models were thus performed on the across-AP and within-AP conditions. For these comparisons, the alpha level was adjusted to .025 (Bonferroni correction for two comparisons—the across-AP

and within-AP conditions). These analyses revealed a marginally significant effect of resynthesis only in the within-AP conditions ($F[1,179]=4.6, p=.034$). These results indicate that native listeners made fewer incorrect detections of the target word when it crossed an AP boundary than when it did not, and in the within-AP conditions, they made fewer incorrect detections of the target word when the stimuli had the F0 of the across-AP condition (where the monosyllabic word was accented) than when they had their natural (unaccented) prosody.

Linear mixed models on L2 learners' false alarm rates, with prosody, resynthesis, and proficiency (mid, high) as fixed variables and with participant and item and random variables, revealed a significant interaction between prosody and proficiency ($F[1,738]=9.7, p<.002$; all other F 's <1). Separate linear mixed models were thus performed on each group. For these comparisons, the alpha level was adjusted to .025 (Bonferroni correction for two comparisons—the mid-level and high-level L2 learners). These analyses revealed significant effects of prosody for both the mid-level L2 learners ($F[1,369]=4.8, p<.029$) and the high-level L2 learners ($F[1,369]=4.9, p<.027$; other F 's <1), with the only the high-level L2 learners showing the effect in the predicted direction. These results indicate that unlike the native listeners and high-level L2 learners, the mid-level L2 learners made more incorrect detections of the target word when it crossed an AP boundary than when it did not, and neither L2 group was influenced by the resynthesis manipulation.

5.3. *Summary*

The results of Experiment 2 confirm that native French listeners identify word-final boundaries at the offset of accented syllables, and they can use F0 rise alone as a cue to word-final boundaries. The flattening of the pitch contour in the resynthesized across-AP stimuli did not increase the false alarm rates, probably because the duration of the accented syllable in these stimuli had not been altered, and thus provided a cue to word-final boundaries that native listeners were also able to exploit. On the other hand, neither L2 group was able to use F0 rise alone as a cue to word-final boundaries in French. Furthermore, whereas the results of the high-level L2 learners in Experiment 2 patterned with those of the high-level L2 learners in Experiment 1, the results of the mid-level L2 learners in the two experiments were in opposite directions, with the mid-level learners in Experiment 2 showing more false alarms in the natural

across-AP condition than in the natural within-AP condition. These results are somewhat surprising in light of the fact that the two mid-level L2 groups were not found to have significantly different proficiencies, either in their cloze test scores or in their language experience in French.

Let us now turn to a discussion of these results and their implications for understanding the development of L2 speech segmentation.

6. General discussion

Non-native listeners exposed to a new language are faced with the task of tracking the linguistic and probabilistic cues that coincide with word-initial and word-final boundaries. Segmenting speech into words cannot be only a by-product of lexical access, as the sound processes that blur word boundaries would otherwise make word recognition much more difficult than we know it to be in native listeners. To parse a language rapidly and efficiently, non-native listeners need to exploit information from various sources in order to predict the likelihood that a given word will be uttered, and use this information to facilitate lexical access.

The questions that motivated the present study are whether English listeners can use prosodic information when recognizing words in French, and how their use of this information unfolds as their proficiency in French increases. Not only are English and French different in their placement of accented syllables in the word (which tend to be word-initial in English but phrase-, and thus word-, final in French), but the acoustic cues to word boundaries also differ between the two languages, such that duration increase is a reliable cue to word-final boundaries in both languages but F0 rise is a reliable cue to word-final boundaries only in French. English L2 learners of French must therefore learn to parse the F0 rise associated with pitch accents as occurring on word-final syllables. At least two additional factors can make this process difficult: because duration is a reliable cue to word-final boundaries in French, English L2 learners of French may not need to attend to F0 rise as another cue to word-final boundaries in French; and F0 rise can also be indicative of phrase-initial accents, and thus of word-initial boundaries, in French (Jun and Fougeron 2002; Welby 2006), potentially making it challenging for L2 learners to distinguish between the two types of accents.

The present study has shown that when a French syllable is accented, and thus has both a higher F0 (coupled with higher intensity) and a longer duration, it is less likely to be parsed by native listeners and high-level L2 learners as a word-initial syllable than when it is unaccented. This indicates that both native listeners and high-level English L2 learners of French use prosodic information to segment French into words. The results of the mid-level L2 learners in Experiment 2, though not those of the corresponding learners in Experiment 1, suggest that before this seemingly native-like use of prosodic information takes place, English L2 learners of French go through a stage during which they parse accented syllables as word-initial rather than as word-final, as we might predict they would do on the basis of their native language (e.g., Cutler and Butterfield 1992; McQueen et al. 1994; Mattys 2004). This means that in earlier stages of development, these L2 learners indeed use F0 rises, but as a cue to word-initial boundaries rather than word-final boundaries.

What remains unclear is why the mid-level L2 learners in the two experiments patterned differently. On the one hand, although cloze tests are an excellent measure of global proficiency (for discussion, see Tremblay 2011), they are completed in the written mode, unlike the experiments in the present study. This can be considered an advantage, in that cloze tests are not circular with oral experiments, but perhaps such tests do not capture sufficient variability in L2 learners' global proficiency, that which pertains to oral proficiency. On the other hand, language background information such as that collected in the present study relies on self-reported data, which may not always be sufficiently precise. Oral measures of proficiency (e.g., oral proficiency interviews) would have been more difficult to administer, but they might have helped capture proficiency differences between our two mid-level groups that the cloze test and language background information did not capture. Hence, studies investigating speech perception and processing should perhaps use oral (in addition to written) proficiency measures, ones that are nonetheless not circular with the abilities tested by the experimental paradigms.

The results of the present study also indicate that native French listeners and English L2 learners of French rely on different acoustic cues of prominence to locate word-final boundaries in continuous speech: whereas native French listeners rely on both F0 rise and duration increase, as has been found in previous studies (e.g., Banel and Bacri 1994; Bagou et al. 2002; Christophe et al. 2004), English listeners appear to use only duration and, to some extent, intensity, to locate

word-final boundaries in French, as indicated by the regression analyses performed in Experiment 1 and the results of the resynthesized conditions in Experiment 2. These findings are in line with those of Tyler and Cutler (2009), who showed that French listeners, but not English listeners, use F0 rise as a cue to word-final boundaries. Our results also pattern similarly to those of Kim et al. (to appear), who found that upon initial exposure to an artificial language, Korean listeners, but not Dutch listeners, associate F0 rise to word-final boundaries. Of course, assuming that it is more difficult to use new (i.e., L2) cue-to-word-edge mappings than older (i.e., native language) ones, it is also possible that the L2 learners in our study were not able to use F0 rise as a cue to word-final boundaries because of the time pressure that was imposed on them.

Although it remains to be seen whether English L2 learners of French would be able to associate F0 rise to word-final boundaries at more advanced proficiencies or in experimental paradigms where they are not put under time pressure, our findings suggest that these L2 learners may in fact not be able to attain this ability. This may be due to the fact that duration also signals word-final boundaries in French. Carroll (2004) proposes that in order for L2 learners to restructure their prosodic representations, parsing failure must first take place. Using the F0 rise associated with pitch accents as a cue to word-initial boundaries in French would clearly trigger parsing failure, as L2 learners would segment accented syllables as word-initial syllables rather than as word-final syllables. By contrast, repressing the use of this F0 rise as a cue to word-initial boundaries and using only duration increase as a cue to word-final boundaries would be an efficient parsing strategy and yield what appears to be a native-like segmentation of the speech stream, even if L2 learners would not extract the same prosodic cues from the speech signal as native listeners do. In their study on the use of rhythmic cues in speech segmentation, Cutler et al. (1992) proposed that bilinguals may only be able to repress processing routines from the dominant language that are inefficient for parsing the weaker language; in other words, they may not be able to create new processing routines. Our results are also consistent with this proposal: L2 learners can repress the use of F0 rise as a cue to word-initial boundaries, but they are unable to map F0 rise to word-final boundaries. We suggest, however, that this mapping problem is due in part to a lack of functional necessity (since duration can signal word-final boundaries in French) and in part to the lack of reliability of F0 rise as a cue to word-final boundaries (since it can also signal word-initial boundaries in French), rather than to an inherent failure to create new

processing routines. Research focusing on a different pairing of native and target languages, in which a given prosodic cue to word boundaries is used in the target language and does not coincide with another prosodic cue from the native language, should elucidate this issue. For example, one could examine Korean listeners' use of F0 rise as a cue to word-initial boundaries in English, where F0 rise (but not duration increase) is a cue to word-initial boundaries in English but not in Korean.

Another interesting finding of our study is that both native French listeners and English L2 learners of French use prosodic information (albeit differently) even if lexical information is not degraded or absent. This may be due in part to the fact that the participants were put under time pressure while completing the experiment, thus increasing the likelihood that they would rely on non-segmental information for resolving temporary lexical ambiguities in the speech stream. Mattys et al. (2005) proposed that speech segmentation proceeds according to a hierarchy of cues, with cues such as lexical information overriding cues such as stress and prosody; "lower-level" cues have a stronger effect on word recognition when "higher-level" cues are not available, and they reduce (but do not eliminate) the effect of "higher-level" cues if they conflict with them. The efficiency of cues within each level depends on their reliability for identifying word boundaries, which is computed by statistical learning mechanisms (see also Saffran et al. 1996; Saffran 2001). The less variable and more reliable nature of "higher-level" cues as compared to "lower level" cues (among other factors) contributes to explaining the nature of the proposed hierarchy. Increasing evidence in support for it is emerging from the processing literature (e.g., Mattys et al. 2005; Mattys et al. 2007; Mattys & Melhorn 2007; White et al. 2010; see also Norris, McQueen, and Cutler 1995). To the extent that our participants resolved temporary lexical ambiguities with prosody as a result of being put under time pressure, our findings are in line with this hierarchy, and suggest that prosodic information also constrains non-native listeners' speech segmentation. Importantly, our results could not be attributed to the L2 learners' familiarity with the target words or to lexical frequency effects, as the participants monitored the exact same word in the across-AP and within-AP conditions. The paradigm we used was thus excellent for teasing apart prosodic effects from lexical ones.

Of course, this research is not without limitations. Testing more advanced English L2 learners of French would answer some of the questions that the present study was unable to

address. One other direction for this work would be to compare L2 learners' ability to use F0 rises in both pitch accents and phrase-initial accents for segmenting speech into words. Whereas (non-sentence-final) pitch accents have a steep rise and are anchored to specific syllables (i.e., the last non-schwa syllable of the AP), phrase-initial accents have a more gradual F0 rise, which (when present) is not always anchored to the same syllable in the word, although it tends to be word-initial (Jun and Fougeron 2000, 2002; Welby 2003, 2006). It would be quite revealing to examine non-native listeners' ability to use both these rises for locating, respectively, word-final and word-initial boundaries in continuous speech. We might predict that L2 learners would suppress their use of F0 rise as a cue to word-initial boundaries, but only when the rise is steep (as in pitch accents), not when the rise is more gradual (as in phrase-initial accents). This would suggest that L2 learners can use fine-grained prosodic information for locating word boundaries in continuous speech. Further research should establish whether this is indeed the case.

7. Conclusion

The present study investigated English and French listeners' use of prosodic information in the segmentation of French speech. The results of two word-monitoring experiments under time pressure confirmed that both the native and non-native listeners used prosodic information to identify word-final boundaries in French, but they extracted different cues from the speech signal, with the native listeners relying on F0 rise and with the L2 learners relying on duration and intensity increases. We suggested that L2 learners' inability to map F0 rise to word-final boundaries is due to: (i) their native language, in which F0 rise is a reliable cue to word-initial boundaries but not word-final boundaries; (ii) the co-occurrence of F0 and duration cues in word-final syllables in French, rendering L2 learners' use of F0 rise unnecessary for locating word-final boundaries; and (iii) the marking of word-initial boundaries by F0 cues in French, thus making it difficult for non-native listeners to tease the two types of F0 rise apart. Further research should examine whether L2 learners at near-native levels of proficiency can in fact use F0 rise as a cue to word-final boundaries, and whether they treat pitch accents differently from phrase-initial accents when segmenting the speech stream.

Appendix: Experimental Items

Across-AP	Within-AP	Control	Syllables 1-3
<i>ce bal conventionnel et pesant</i> 'this conventional and cumbersome bal'	<i>ce bal conventionnel</i> 'this conventional bal'	<i>ce balcon arrondi</i> 'this rounded balcony'	/səbalkɔ̃/
<i>ces bancs dominicains chrétiens</i> 'these Dominican and Christian benches'	<i>ces bancs dominicains</i> 'these Dominican benches'	<i>ces bandeaux acajous</i> 'these mahogany headbands'	/sebãdo/
<i>ce banc distinct et éloigné</i> 'this distinct and distant bench'	<i>ce banc distinct</i> 'this distinct bench'	<i>ce bandit basque</i> 'this Basque bandit'	/səbãdi/
<i>la boue gisante et huileuse</i> 'the recumbent and oily mud'	<i>la boue gisante</i> 'the recumbent mud'	<i>la bougie blanche</i> 'the white candle'	/labuʒi/
<i>le cerf vorace et majestueux</i> 'the voracious and majestic stag'	<i>le cerf vorace</i> 'the voracious stag'	<i>le cerveau droit</i> 'the right brain'	/ləsɛʁvo/
<i>ce chat grincheux et bedonnant</i> 'this cranky and chubby cat'	<i>ce chat grincheux</i> 'this cranky cat'	<i>ce chagrin fou</i> 'this incredible grief'	/səʒaɡʁɛ̃/
<i>le chat lépreux et légendaire</i> 'the leprous and legendary cat'	<i>le chat lépreux</i> 'the leprous cat'	<i>le chalet suisse</i> 'the Swiss cabin'	/ləʃalɛ/
<i>ce chat pauvre et mal avenant</i> 'this poor and annoying cat'	<i>ce chat pauvre et sale</i> 'this poor and dirty cat'	<i>ce chapeau melon</i> 'this bowler hat'	/səʃapo/
<i>le corps bossu et déformé</i> 'the hunched and misshapen body'	<i>le corps bossu</i> 'the misshapen body'	<i>le corbeau noir</i> 'the black crow'	/ləkɔʁbo/
<i>ces corps végétatifs et ralentis</i> 'these vegetative and idle bodies'	<i>ces corps végétatifs</i> 'these vegetative bodies'	<i>ces corvées ménagères</i> 'these house chores'	/sekɔʁve/
<i>le coût rentable et sécurisé</i> 'the profitable and secure cost'	<i>le coût rentable</i> 'the profitable cost'	<i>le courant fort</i> 'the strong current'	/ləkɔʁã/
<i>ce coup singulier et symbolique</i> 'this peculiar and symbolic hit'	<i>ce coup singulier</i> 'this peculiar hit'	<i>ce coussin en mousse</i> 'this foam cushion'	/səkusɛ̃/
<i>les dents gélatineuses et difformes</i> 'the gelatinous and deformed teeth'	<i>les dents gélatineuses</i> 'the gelatinous teeth'	<i>les dangers nucléaires</i> 'the nuclear dangers'	/ledãʒɛ/

<i>ce phare doré et balisé</i>	<i>ce phare doré</i>	<i>ce fardeau lourd</i>	/səfɑ̃do/
‘this golden and blazed lighthouse’	‘this golden lighthouse’	‘this heavy burden’	
<i>ce fort maléfique et hanté</i>	<i>ce fort maléfique</i>	<i>ce format papier</i>	/səfɔ̃ma/
‘this maleficent and haunted fort’	‘this maleficent fort’	‘this paper format’	
<i>ce fou larmoyant et navrant</i>	<i>ce fou larmoyant</i>	<i>ce foulard marron</i>	/səfulaʁ/
‘this tearful and upsetting madman’	‘this tearful madman’	‘this brown scarf’	
<i>ces fours miniatures et métallisés</i>	<i>ces fours miniatures</i>	<i>ces fourmis oranges</i>	/sɛfɔ̃ʁmi/
‘these miniature and metallic ovens’	‘these miniature ovens’	‘these orange ants’	
<i>le mat tournant et brisé</i>	<i>le mat tournant</i>	<i>le matou rond</i>	/ləmatu/
‘the rotating and broken mast’	‘the rotating mast’	‘the round tomcat’	
<i>cette mort surprenante et mystique</i>	<i>cette mort surprenante</i>	<i>cette morsure profonde</i>	/sɛtmɔʁsyʁ/
‘this surprising and mystic death’	‘this surprising death’	‘this deep bite’	
<i>ces paons flétris et ternis</i>	<i>ces paons flétris</i>	<i>ces pamphlets roses</i>	/sɛpɑ̃flɛ/
‘these faded and dull peacocks’	‘these faded peacocks’	‘these pink pamphlets’	
<i>le père militaire et carriériste</i>	<i>le père militaire</i>	<i>le permis de chasse</i>	/ləpɛʁmi/
‘the military and career-oriented father’	‘the military father’	‘the hunting permit’	
<i>le père silencieux et calme</i>	<i>le père silencieux</i>	<i>le persil chinois</i>	/ləpɛʁsi/
‘the silent and calm father’	‘the silent father’	‘the Chinese parsley’	
<i>le père vertueux et enthousiaste</i>	<i>le père vertueux</i>	<i>le pervers sexuel</i>	/ləpɛʁvɛʁ/
‘the virtuous and enthusiastic father’	‘the virtuous father’	‘the sexual pervert’	
<i>ces pins séduisants et géants</i>	<i>ces pins séduisants</i>	<i>ces pincées brutales</i>	/sɛpɛ̃sɛ/
‘these appealing and giant pine trees’	‘these appealing pine trees’	‘these brutal pinches’	
<i>ce plat fondant et salé</i>	<i>ce plat fondant</i>	<i>ce plafond rond</i>	/sɛplafɔ̃/
‘this delicious and salted dish’	‘this delicious dish’	‘this round ceiling’	
<i>les ports ténébreux et désuets</i>	<i>les ports ténébreux</i>	<i>les portées des chattes</i>	/ləpɔʁtɛ/
‘the obscure and quaint ports’	‘the obscure ports’	‘the kitty litters of the cats’	
<i>ce port trépidant et célèbre</i>	<i>ce port trépidant</i>	<i>ce portrait photo</i>	/sɛpɔʁtyʁ/
‘this vibrant and famous port’	‘this vibrant port’	‘this photo portrait’	
<i>les rangs partiels et désordonnés</i>	<i>les rangs partiels</i>	<i>les remparts rouges</i>	/ləʁɑ̃paʁ/
‘the partial and disorganized lines’	‘the partial lines’	‘the red battlements’	

Native and Non-Native Segmentation of French Speech, 41

<i>le rat tondu et bedonnant</i>	<i>le rat tondu</i>	<i>le raton gris</i>	/ləʁatɔ̃/
‘the shaved and chubby rat’	‘the shaved rat’	‘the grey young-rat’	
<i>ces sols dallés et brillants</i>	<i>ces sols dallés</i>	<i>ces soldats belges</i>	/sesɔlda/
‘these paved and shiny floors’	‘these paved floors’	‘these Belgian soldiers’	
<i>les temps bourgeois et catholiques</i>	<i>les temps bourgeois</i>	<i>les tambours blancs</i>	/letɑ̃buʁ/
‘the bourgeois and catholic times’	‘the bourgeois times’	‘the white drums’	
<i>les tours néogothiques montantes</i>	<i>les tours néogothiques</i>	<i>les tournées artistiques</i>	/letuʁnɛ/
‘the neogothic rising towers’	‘the neogothic towers’	‘the artistic tours’	
<i>les vers gélatineux et luisant</i>	<i>les vers gélatineux</i>	<i>les vergers abondants</i>	/levɛʁʒɛ/
‘the gelatinous and glistening worms’	‘the gelatinous worms’	‘the bountiful orchards’	
<i>les vers séchés et ondulés</i>	<i>les vers séchés</i>	<i>les versets tristes</i>	/levɛʁsɛ/
‘the dry and wavy worms’	‘the dry worms’	‘the sad verses’	
<i>les vies perdues et sacrifiées</i>	<i>les vies perdues</i>	<i>les vipères vertes</i>	/levipɛʁ/
‘the lost and sacrificed lives’	‘the lost lives’	‘the green garden-snakes’	
<i>la vie réjouissante et sereine</i>	<i>la vie réjouissante</i>	<i>la virée nocturne</i>	/lavire/
‘the delightful and serene life’	‘the delightful life’	‘the late-night ride’	

Notes

1. The present design was favored over one in which the participants would be asked to detect the monosyllabic word in the sentence (e.g., *chat*), because the latter would have required an assessment of whether non-native listeners have particular competitor words (e.g., *chalet*) in their lexicon, which could not have been done prior to the experiment and thus would have resulted in the exclusion of test items and possibly participants.
2. The difference between the two proficiency groups in their months of residence in a French-speaking environment may not have reached significance due to the relatively small number of participants. Nonetheless, the results pattern in the direction that we would expect based on the cloze test scores.
3. The native French speaker who recorded the stimuli was from Bordeaux and thus produced vowels in open syllables as tensed in both word-final and non-word-final positions (e.g., *chalet* [ʃale], *chat lépreux* [ʃalepχø]).
4. An anonymous reviewer suggested that the first AP in the across-AP condition must have been recorded with a single hybrid LH* rise, with the L tone belonging to a phrase-initial accent and thus potentially signaling a word-initial boundary. We recognize that this is a possibility. If it did belong to a phrase-initial accent and signaled word-initial boundaries, however, we might expect as many or more incorrect detection of *chalet* in the across-AP condition than in the within-AP condition (where the two consecutive LL tones would not provide a cue to word-initial boundaries), contrary to fact (see the results section).
5. The higher peak F0 on Syllable 1 (e.g., *le*) in the across-AP condition than in the within-AP condition could potentially cue the listeners to the word-initial boundary, but the results again suggest that this is not the case (see Footnote 4).
6. The peak F0 of the vowel in the monosyllabic noun was not included in these analyses, because it did not explain a significant amount of variance in the participants' false alarms.
7. Because Experiment 2 did not include a control condition in which the target word was present, a smaller number of distracter items from Experiment 1 were used so that half of the test items would contain the target word and half would not.

8. The values reported for the natural stimuli in Table 5 are slightly different from those reported for the corresponding stimuli in Table 2, because not all stimuli from Experiment 1 were used in Experiment 2.
9. Again, the values of the statistical analyses reported for the natural stimuli in Table 6 are slightly different from those reported for the corresponding stimuli in Table 3, because not all stimuli from Experiment 1 were used in Experiment 2.

References

- Astsésano, Corine. 2001. *Rythme et accentuation en français. Invariance et variabilité stylistique*. Paris: Editions L'Harmattan.
- Astsésano, Corine, Ellen Gurman Bard, & Alice Turk. 2007. Structural influences on initial accent placement in French. *Language and Speech* 50. 423–446.
- Bagou, Odile, Cécile Fougeron, & Ulrich H. Frauenfelder. 2002. Contribution of prosody to the segmentation and storage of “words” in the acquisition of a new mini-language. In Bernard Bel & Isabelle Marlien (eds.), *Proceedings of Speech Prosody 2002*, 159–162. Aix-en-Provence: Laboratoire Parole et Langage.
- Bagou, Odile, & Ulrich H. Frauenfelder. 2006. Stratégie de segmentation prosodique: Rôle des prééminences initiales et finales dans l'acquisition d'une langue artificielle. In *Actes des XXVIèmes Journées d'Étude sur la Parole* [Proceedings of the XXVIth Day of Study on Speech], 571–574. Dinard, France.
- Banel, Marie-Hélène, & Nicole Bacri. 1994. On metrical patterns and lexical parsing in French. *Speech Communication* 15. 115–126.
- Beckman, Mary E. 1986. *Stress and Non-Stress Accent*. Dordrecht: Foris.
- Beckman, Mary. E., & Gayle A. Elam. 1997. Guidelines for ToBI labeling (version 3.0). Available at www.ling.ohio-state.edu/_tobi/ame_tobi/labelling_guide_v3.pdf.
- Boersma, Paul, & David Weenink. 2004. Praat: Doing phonetics by computer (Version 5.2.11) [Computer program]. Available at www.praat.org.
- Carroll, Susanne E. 2004. Segmentation: Learning how to ‘hear words’ in the L2 speech stream. *Transactions of the Philological Society* 102. 227–254.
- Christophe, Anne, Ariel Gout, Sharon Peperkamp, & James Morgan. 2003. Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics* 31. 585–598.
- Christophe, Anne, Sharon Peperkamp, Christophe Pallier, Eliza Block, & Jacques Mehler. 2004. Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language* 51. 523–547.

- Clopper, Cynthia. 2002. Frequency of stress patterns in English: A computational analysis. *IULC Working Papers Online*, 2. Available at <https://www.indiana.edu/~iulcwp/pdfs/02-clopper02.pdf>
- Cooper, Nicole, Anne Cutler, & Roger Wales. 2002. Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech* 45. 207–228.
- Cutler, Anne. 1986. Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29. 201–220.
- Cutler, Anne, & Sally Butterfield. 1992. Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 31. 218–236.
- Cutler, Anne, & David M. Carter. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133–142.
- Cutler, Anne & Charles E. Clifton Jr. 1984. The use of prosodic information in word recognition. In Herman Bouma & Don G. Bouwhuis (eds.), *Attention and Performance X: Control of Language Processes*, 183–196. Hillsdale, NJ: Erlbaum.
- Cutler, Anne, & Wilma van Donselaar. 2001. *Voornaam* is not a homophone: Lexical prosody and lexical access in Dutch. *Language and Speech* 44. 171–195.
- Cutler, Anne, Jacques Mehler, Dennis Norris, & Juan Segui. 1992. The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology* 24. 381–410.
- Di Cristo, Albert. 1998. Intonation in French. In Daniel Hirst & Albert Di Cristo (eds.), *Intonation Systems: A Survey of Twenty Languages*, 195–218. Cambridge: Cambridge University Press.
- Di Cristo, Albert. 2000. Vers une modélisation de l'accentuation en français. Deuxième partie: Le modèle. *Journal of French Language Studies* 10. 27–44.
- Di Cristo, Albert, & Daniel Hirst. 1993. Rythme syllabique, rythme mélodique et représentation hiérarchique de la prosodie du français. *Travaux de l'Institut de Phonétique d'Aix* 15. 13–23.
- Fónagy, Ivan. 1980. L'accent en français: Accent probabilitaire. In Ivan Fónagy & Pierre R. León (eds.), *L'accent en français contemporain*, 123–133. Ottawa: Didier.
- Donselaar, Wilma van, Mariëtte Koster, & Anne Cutler. 2005. Exploring the role of lexical stress in lexical recognition. *Quarterly Journal of Experimental Psychology* 58A. 251–273.

- Fear, Beverly D., Anne Cutler, & Sally Butterfield. 1995. The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America* 97. 1893–1904.
- Fougeron, Cécile. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29. 109–135.
- Gerken, LouAnn, Peter W. Jusczyk, & Denise R. Mandel. 1994. When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition* 51. 237–265.
- Halle, Morris, & Jean-Roger Vergnaud. 1987. *An Essay on Stress*. Cambridge, MA: MIT Press.
- Hammond, Michael. 1999. *The Phonology of English*. Oxford: Oxford University Press.
- Hirsh-Pasek, Kathryn, Deborah G. Kemler-Nelson, Peter W. Jusczyk, Kimberly W. Cassidy, Benjamin Druss, & Lori Kennedy. 1987. Clauses are perceptual units for young infants. *Cognition* 26. 269–286.
- Jun, Sun-Ah. 2006. *Prosodic Typology*. Oxford: Oxford University Press.
- Jun, Sun-Ah., & Cécile Fougeron. 2002. Realizations of accentual phrase in French intonation. *Probus* 14. 147–172.
- Jusczyk, Peter W., & Richard N. Aslin. 1995. Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology* 29. 1–23.
- Jusczyk, Peter W., Derek M. Houston, & Mary Newsome. 1999. The beginnings of word segmentation in English-learning infants. *Cognitive Psychology* 39. 159–207.
- Jusczyk, Peter W., Deborah G. Kemler-Nelson, Kathryn Hirsh-Pasek, Lori Kennedy, Amanda Woodward, & Julie Piwoz. 1992. Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252–293.
- Kemler-Nelson, Deborah G., Kathryn Hirsh-Pasek, Peter W. Jusczyk, & Kimberly W. Cassidy. 1989. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 16. 55–68.
- Kim, Sahyang. 2004. The role of prosodic cues in word segmentation in Korean. In *Proceedings of the 8th International Conference on Spoken Language Processing*, 3005–3008. Jeju Island, Korea.
- Kim, Sahyang, Mirjam Broersma, & Taehong Cho. (to appear). The use of prosodic cues in processing an unfamiliar language. *Studies in Second Language Acquisition*.

- Kim, Sahyang, & Taehong Cho. 2009. The use of phrase-level prosodic information in lexical segmentation: Evidence from word-spotting experiments in Korean. *Journal of the Acoustical Society of America* 125. 3373–3386.
- Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* 32. 451–454.
- Mattys, Sven L. 2004. Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance* 30. 397–408.
- Mattys, Sven L., & James F. Melhorn. 2007. Sentential, lexical, and acoustic effects on the perception of word boundaries. *Journal of the Acoustical Society of America* 122. 554–567.
- Mattys, Sven L., James F. Melhorn, & Laurence White. 2007. Effects of syntactic expectations on speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance* 33. 960–977.
- Mattys, Sven L., Laurence White, & James F. Melhorn. 2005. Integration of multiple speech segmentation cues. *Journal of Experimental Psychology: General* 134. 477–500.
- McQueen, James M., Dennis Norris, & Anne Cutler. 1994. Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20. 621–638.
- Mertens, Piet. 1987. *L'Intonation du français*. Unpublished PhD dissertation, Katholieke Universiteit Leuven.
- Mettouchi, Amina, Anne Lacheret-Dujour, Vered Silber-Varod, & Shlomo Izre'el. 2007. Only prosody? Perception of speech segmentation in Kabyle and Hebrew. *Nouveaux cahiers de linguistique française* 28. 207–218.
- Michelas, Amandine, & Mariapaola D'Imperio. 2010. Accentual phrase boundaries and lexical access in French. In *Proceedings of the Speech Prosody 2010 Conference, Chicago*. Retrieved from <<http://speechprosody2010.illinois.edu/papers/100882.pdf>>.
- Michelas, Amandine, & Mariapaola D'Imperio. 2011. Uncovering the role of the intermediate phrase in the syntactic parsing of French. In Wai-Sum Lee & Eric Zee (eds.),

- Proceedings of the 17th International Congress for Phonetic Sciences*, 1374–1377. Hong Kong: City University of Hong Kong.
- Nespor, Marina, & Irene Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Norris, Dennis, James M. McQueen, & Anne Cutler. 1995. Competition and segmentation in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21. 1209–1228.
- Pasdeloup, Valérie. 1990. *Modèle de règles rythmiques du français appliqué à la synthèse de la parole*. Aix-en-Provence: Université de Provence.
- Pijper, Jan Roelof de. 1983. *Modeling British English Intonation*. Dordrecht: Foris.
- Post, Brechtje. 2000. *Tonal and Phrasal Structures in French Intonation*. Unpublished PhD dissertation, University of Nijmegen.
- Rolland, Guillaume, & Hélène Løevenbruck. 2002. Characteristics of the accentual phrase in French: An acoustic, articulatory and perceptual study. In *Proceedings of Speech Prosody 2002*, 611–614. Aix-en-Provence, France.
- Saffran, Jenny R. 2001. The use of predictive dependencies in language learning. *Journal of Memory and Language* 44. 493–515.
- Saffran, Jenny R., Elissa L. Newport, & Richard N. Aslin. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35. 606–621.
- Sanders, Lisa D., Helen J. Neville, & Marty G. Woldorff. 2002. Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research* 45. 519–530.
- Schneider, Walter, Amy Eschman, & Anthony Zuccolotto. 2002. *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools Inc.
- Soto-Faraco, Salvador, Núria Sebastián-Gallés, & Anne Cutler. 2001. Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45. 412–432.
- Spinelli, Elsa, Nicolas Grimaud, Fanny Meunier, & Pauline Welby. (2010). An intonational cue to segmentation in phonemically identical sequences. *Attention, Perception, & Psychophysics* 72. 775–787.

- Spinelli, Elsa, Pauline Welby, & Anne-Laure Schaegis. 2007. Fine-grained access to targets and competitors in phonemically identical spoken sequences: The case of French elision. *Language and Cognitive Processes* 22. 828–859.
- Tremblay, Annie. 2008. Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics* 29. 553–584.
- Tremblay, Annie. 2011. Proficiency assessment in second language acquisition research: “Clozing” the gap. *Studies in Second Language Acquisition* 33. 339–372.
- Tremblay, Annie, & Meryl D. Garrison. 2010. Cloze tests: A tool for proficiency assessment in research on L2 French. In Matthew T. Prior, Yukiko Watanabe, & Sang-Ki Lee (eds.), *Selected proceedings of the Second Language Research Forum 2008*, 73–88. Somerville, MA: Cascadilla Press.
- Tremblay, Annie, & Nathan Owens. 2010. The role of acoustic cues in the development of (non-)target-like second-language prosodic representations. *Canadian Journal of Linguistics* 55. 85–114.
- Tyler, Michael D., & Anne Cutler. 2009. Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America* 126. 367–376.
- Vaissière, Jacqueline. 1983. Language-independent prosodic features. In Anne Cutler & Robert D. Ladd (eds.), *Prosody: Models and Measurements*, 53–66. Berlin: Springer.
- Vassière, Jacqueline. 1997. Langues, prosodies et syntaxe. *Revue Traitement Automatique des Langues* 38. 53–82.
- Vassière, Jacqueline, & Alexis Michaud. 2007. Prosodic constituents in French: A data-driven approach. In Iván Fonagy, Yuji Kawaguchi, & Tsunekaku Moriguchi (eds.), *Prosody and Syntax*, 47–64. Amsterdam: John Benjamins.
- Verluyten, S. Paul 1982. Recherches sur la prosodie et la métrique du français [Investigations about prosody and metrics in French]. Unpublished Ph.D. dissertation, University of Antwerpen.
- Warner, Natasha, Takashi Otake, & Takayuki Arai. 2010. Intonational structure as a word-boundary cue in Tokyo Japanese. *Language & Speech* 53. 107–131.

- Welby, Pauline. 2003. *The Slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation*. Unpublished PhD dissertation, Ohio State University.
- Welby, Pauline. 2006. French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34. 343–371.
- Welby, Pauline. 2007. The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication* 49. 28–48.
- Welby, Pauline, & Hélène Løevenbruck. 2006. Anchored down in anchorage: Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics* 18. 74–124.
- White, Laurence, James F. Melhorn, & Sven L. Mattys. 2010. Segmentation by lexical subtraction in Hungarian speakers of second language English. *The Quarterly Journal of Experimental Psychology* 63. 544–554.