Bilingualism: Language and Cognition

cambridge.org/bil

Re-examining the effect of phonological similarity between the native- and second-language intonational systems in second-language speech segmentation

Annie Tremblay[1], Sahyang Kim[2], Seulgi Shin[1] and Taehong Cho[3]

[1]University of Kansas; [2]Hongik University and [3]Hanyang University

Research Article

Cite this article: Tremblay A, Kim S, Shin S, Cho T (2021). Re-examining the effect of phonological similarity between the native- and second-language intonational systems in second-language speech segmentation. *Bilingualism: Language and Cognition* **24**, 401–413. https://doi.org/10.1017/S136672892000053X

Received: 16 October 2019
Revised: 17 August 2020
Accepted: 17 August 2020
First published online: 4 November 2020

Keywords:
intonation; tonal cues; second-language speech segmentation; Korean

Address for correspondence: Annie Tremblay, Department of Linguistics, 1541 Lilac Lane, Blake Hall Rm 426, Lawrence, KS 66045, E-mail atrembla@ku.edu

Abstract

This study investigates how phonological and phonetic aspects of the native-language (L1) intonation modulate the use of tonal cues in second-language (L2) speech segmentation. Previous research suggested that prosodic learning is more difficult if the L1 and L2 intonations are phonologically similar but phonetically different (French–Korean) than if they are phonologically different (English–French/Korean) (Prosodic-Learning Interference Hypothesis; Tremblay, Broersma, Coughlin & Choi, 2016). This study provides another test of this hypothesis. Korean listeners and French-speaking and English-speaking L2 learners of Korean in Korea completed an eye-tracking experiment investigating the effects of phrase tones in Korean. All groups patterned similarly with the phrase-final tone, but, unlike Korean and French listeners, English listeners showed early benefits from the phrase-initial tone (signaling word-initial boundaries in English). Importantly, French listeners patterned like Korean listeners with both tones. The Prosodic-Learning Interference Hypothesis is refined to suggest that prosodic learning difficulties may not be persistent for immersed L2 learners.

1. Introduction

Intonation plays a vital role in speech communication, from basic functions such as grouping words into phrases for marking prosodic junctures (DELIMITATIVE functions) and marking relative prominence among prosodic constituents (CULMINATIVE function) to more complex functions such as signaling the structure of discourse information and communicating attitudes and emotions. The processes that underlie speech perception and comprehension and the resulting outcomes are indeed modulated by intonation in native listeners (e.g., Brown, Salverda, Dilley & Tanenhaus, 2011, 2015a; Brown, Salverda, Gunlogson & Tanenhaus, 2015b; Christophe, Peperkamp, Pallier, Block & Mehler, 2004; Ito, Jincho, Minai, Yamane & Mazuka, 2012; Ito & Speer, 2008; Kim & Cho, 2009; Kim, Mitterer & Cho, 2018; Salverda, Dahan & McQueen, 2003; Salverda, Dahan, Tanenhaus, Crosswhite, Masharov & McDonough, 2007; Spinelli, Grimault, Meunier & Welby, 2010; Steffman, 2019). Yet, the learning of intonational cues in second/foreign-language (L2) speech perception and comprehension has received little attention in research, at least compared to the learning of segmental (i.e., consonant, vowel) information (for examples of L2 studies on the perception and/or comprehension of intonation, see Mok, Yin, Setter & Nayan, 2016; Ortega-Llebaria & Colantoni, 2014; Ortega-Llebaria, Nemogá & Presson, 2015; Ortega-Llebaria, Olson & Tuninetti, 2018; Puga, Fuchs, Setter & Mok, 2017; Tremblay, Broersma & Coughlin, 2018; Tremblay et al., 2016; Tremblay, Coughlin, Bahler & Gaillard, 2012). Accordingly, the theories and models that have been proposed to explain L2 speech perception and comprehension have often focused on the learning of segmental categories (e.g., Best & Tyler, 2007; Flege, 1995; van Leussen & Escudero, 2015) and in general have little to say about the learning of intonation.

To illustrate, at the segmental level, theories of L2 speech learning indicate that the perceptual learning of L2 sounds and the use of L2 sounds in speech comprehension are affected by how L2 sounds relate acoustically and perceptually to existing sounds in the native (i.e., first) language (L1) and whether L2 and corresponding L1 sounds form segmental and lexical contrasts (e.g., Best & Tyler, 2007; Flege, 1995; van Leussen & Escudero, 2015). Specifically, L2 listeners have been proposed to perceptually assimilate L2 sounds to L1 sounds when the L2 and L1 sounds are realized similarly (from an acoustic and/or articulatory perspective), leading L2 listeners to have perceptual difficulties when two contrastive L2 sounds assimilate to the same L1 sound. However, since intonation can serve different non-lexical functions (e.g., as briefly mentioned above, intonation serves delimitative and culminative functions that are largely independent from segmental and lexical information), and given that its acoustic realization is much more variable than that of segmental information, L2 listeners' learning and use of intonation may differ in important ways from their learning and use of segmental

© The Author(s), 2020. Published by Cambridge University Press

CAMBRIDGE UNIVERSITY PRESS

information. It is thus unclear whether the mechanisms proposed to influence the perceptual learning of segmental categories similarly explain the perceptual learning of intonation in L2 listeners.

Another important difference between segmental and intonational information relates to how "phonological" differences among languages are operationalized. At the segmental level, the term "phonological" refers to segmental changes that create lexical contrasts among words composed of contrastive sounds in a given language. At the intonational level, on the other hand, the term "phonological" usually refers to how different tones (e.g., high and low) are combined to generate a particular tune in a given language (e.g., Pierrehumbert, 1980). Crucially, like segmental information, the intonational systems of languages also differ at the phonetic level (e.g., how phonologically similar tonal make-ups may differ in terms of their precise timing in a prosodic domain; cf. Ladd, 2012). In light of the differences between segmental and intonational information, several issues have yet to be resolved – for instance, how phonological and phonetic aspects of the L1 and L2 intonations modulate the learning of intonational cues in L2 speech perception and comprehension.

With the goal of shedding some light on these questions, the present study investigates how phonological and phonetic aspects of the L1 intonation modulate the use of tonal cues in L2 speech segmentation. Tremblay et al. (2016) have proposed that if the L1 and L2 intonations are phonologically similar yet phonetically different, listeners may perceive the L2 intonation as identical to the L1 intonation and not learn the fine-grained differences between them, a learnability problem they referred to as the Prosodic-Learning Interference Hypothesis (Tremblay et al., 2016). The authors suggest that listeners may perceptually assimilate L2 tonal categories to L1 tonal categories, as has been proposed for the learning of segmental information (e.g., Best & Tyler, 2007; Flege, 1995; van Leussen & Escudero, 2015), and that L2 tonal categories should be sufficiently distant from L1 tonal categories for L2 listeners to be able to learn these new categories. Tremblay et al. (2016)'s proposal thus entails that the mechanisms underlying the learning of segmental and intonational information do not differ in significant ways.

The present study elaborates on these possibilities by examining whether the learning of intonational cues to word boundaries in L2 speech segmentation is indeed more difficult if the L1 and L2 have phonologically similar but phonetically different intonations than if they have phonologically different intonations. More precisely, this study re-examines Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis by investigating the use of tonal information in the segmentation of Seoul Korean (thereafter, Korean) in three listening groups: native Korean listeners, French-speaking L2 learners of Korean, and English-speaking L2 learners of Korean. Korean and French (discussed below) have phonologically similar intonational patterns but their exact phonetic realization differs between the two languages. Investigating the use of tonal cues to word boundaries in French-speaking L2 learners of Korean will illuminate the degree to which intonational cues for detecting word boundaries can be learned when the intonations of the L1 and L2 are phonologically similar but phonetically different. Furthermore, since the intonation of English (discussed below) differs phonologically from that of Korean, comparing the performance of English-speaking L2 learners of Korean with that of proficiency-matched French-speaking L2 learners of Korean will elucidate the extent to which phonological (dis)similarity influences the learning of intonational cues to word boundaries in L2 speech segmentation, thus providing an ideal test of Tremblay et al. (2016)'s hypothesis.

The intonation of Korean has been analyzed as having an Accentual Phrase (AP) with the default tonal pattern of L(HL)H when the AP-initial segment is lenis (Jun, 1998, 2000). In Korean, the AP-final H is LOOSELY anchored to the AP-final syllable and the (subsequent) AP-initial L is CLOSELY anchored to the (following) AP-initial syllable; this means that the peak of the F0 rise in the AP-final syllable can vary, but it is sufficiently EARLY to allow the F0 to be low and reach its target L tone early on in the (subsequent) AP-initial syllable (Jun, 2000, example (9)). French is similar to Korean in that it also has an AP with the default tonal pattern of L(HL)H (Jun & Fougeron, 2000, 2002). In French, however, the AP-final H tone is CLOSELY anchored to the AP-final syllable and the (subsequent) AP-initial L tone is LOOSELY anchored to the (following) AP-initial syllable; consequently, the last (non-reduced) syllable of AP-final words has an F0 rise that peaks towards the end of the AP, and the F0 lowers in the (subsequent) AP-initial syllable and reaches its target L tone somewhat later (e.g., Jun & Fougeron, 2002, Figure 8a; Welby, 2006). Hence, Korean and French are phonologically similar but differ in the phonetic realization of the tonal categories that signal the beginning and end of phrases (and thus words). Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis therefore predicts that French-speaking L2 learners of Korean and Korean-speaking L2 learners of French should have difficulty learning the fine-grained phonetic differences between, respectively, the Korean and the French intonations.

In contrast to both Korean and French, the intonational realization of words in English is not dependent on the position of the word in the phrase, but rather on the distribution and types of pitch accents and on the types of boundary tones (i.e., tones that occur at the end of a large phrase) in relation to discourse information (e.g., Beckman & Pierrehumbert, 1986; Ladd, 2012; Pierrehumbert, 1980; Pierrehumbert & Hirschberg, 1990). Moreover, unlike Korean and French, English has lexical stress, with words being statistically more likely to be stressed on the first syllable than elsewhere in the word (e.g., Clopper, 2002; Cutler & Carter, 1987). When English words receive a neutral (i.e., H) pitch accent, the F0 rise associated with the pitch accent occurs on the stressed (often word-initial) syllable (Beckman & Elam, 1997). English is thus phonologically (and phonetically) different from both Korean and French. Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis therefore predicts that English-speaking L2 learners of Korean and English-speaking L2 learners of French should be successful in learning the phonological (and phonetic) differences between the English and, respectively, the Korean and the French intonations.

In a visual-world eye-tracking study, Tremblay et al. (2016) investigated the use of fundamental frequency (F0) rise as a cue to the end of phrase-final French words by native French listeners, Korean-speaking L2 learners of French, and English-speaking L2 learners of French. The two groups of L2 learners of French did not differ in their proficiency in or experience with French. Participants heard stimuli where the monosyllabic target word and the first syllable of the following adjective (e.g., *chat lépreux* 'leprous cat') were temporarily ambiguous with a disyllabic competitor word (e.g., *chalet* 'cabin'). The monosyllabic target word was manipulated so that it would contain or not contain an H tone, realized as an F0 rise that peaked at the offset of the monosyllabic word. Native French listeners showed a greater

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

target-over-competitor fixation advantage when the F0 signaled the end of words in phrase-final position than when it did not, with this effect emerging early on in the word recognition process. English-speaking L2 learners of French also showed this effect but later in the word recognition process, indicating that they needed time to integrate this cue. By contrast, Korean-speaking L2 learners of French did not show a facilitative effect of the F0 rise cue, indicating that they were unable to use the phrase-final F0 rise to locate word boundaries in French (e.g., Kim & Cho, 2009; Tremblay, Cho, Kim & Shin, 2019, discussed below). It was on the basis of these results that Tremblay et al. (2016) proposed the Prosodic-Learning Interference Hypothesis. The authors suggested that Korean listeners could not use the F0 rise to locate the end of phrase-final words in French because the F0 of the AP-final H tone peaked too late in the AP-final syllable and did not reach its subsequent L target sufficiently early in the following AP. Importantly, Tremblay et al. (2016) proposed that the similarity between the French and Korean intonations posed a learnability problem for Korean-speaking L2 learners of French, who did not show evidence of having learned the fine-grained differences between the two systems.

The present study provides a further test of Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis by investigating Korean, French, and English listeners' use of F0 cues to word boundaries in the segmentation of Korean speech. The prosodic structure of Korean proposed by Jun (1998, 2000) is such that the AP-final H tone should provide an intonational cue to the end of words in phrase-final position, and the AP-initial L tone should provide an intonational cue to the beginning of words in phrase-initial position. Native Korean listeners have indeed been found to use these two intonational cues to segment Korean speech into words. Using word-spotting experiments, Kim and Cho (2009) showed that Korean listeners' speech segmentation was enhanced by the AP-final H tone in the presence of a subsequent AP-initial L tone and by the AP-initial L tone in the presence of a preceding AP-final H tone, suggesting that the contrast between the H and L tones helped Korean listeners break the speech signal down into individual words. Similarly, using an artificial-language learning paradigm, Kim, Broersma, and Cho (2012) showed that Korean listeners' learning of statistical dependencies among syllables in a continuous artificial speech stream was enhanced by the presence of both an H tone on the word-final syllable and an L tone on the word-initial syllable (compared to when tonal information did not signal word boundaries). Finally, also in an artificial-language learning paradigm, Tremblay et al. (2019) found that Korean listeners' learning of statistical dependencies among syllables benefited from the presence of a word-final H tone if the scaling of word-initial L tone was sufficiently low. These results suggest that Korean listeners may pay closer attention to the phonetic realization of the AP-initial L tone than to that of the AP-final H tone, possibly explaining Korean listeners' inability to use a late-peaking F0 rise to locate the end of phrase-final words in French (in Tremblay et al., 2016, this rise was not followed by an AP-initial L tone closely aligned with the beginning of the AP).

The findings reported for Korean listeners' use of intonational cues to word boundaries differ from those reported for native French listeners. While French listeners' speech segmentation has been found to benefit from an AP-final H tone (e.g., Christophe et al., 2004; Michelas & D'Imperio, 2010; Tremblay et al., 2016; Tremblay et al., 2012), it does not appear to be contingent on the presence of a subsequent AP-initial L tone closely aligned with the beginning of the AP boundary. In fact, Spinelli et al. (2010) and Welby (2007) have shown that French listeners are more likely to perceive a word-initial boundary as the F0 at the onset of a segmentally ambiguous string INCREASES (e.g., they are more likely to perceive /lafiʃ/ as *l'affiche* 'the poster' and less likely to perceive it as *la fiche* 'the index card' as the F0 on /la/ increases; Spinelli et al., 2010). The lack of relationship between the AP-initial L tone and the beginning of phrase-initial words in French may be explained in two ways: (i) the AP-initial L tone is not closely aligned with the AP-initial syllable in French due to the late alignment of the AP-final H tone (unlike Korean); and (ii) APs are more likely to begin with a determiner in French than in Korean, thus making the AP-initial L tone less useful for locating the beginning of phrase-initial words in French.

Given these similarities and differences between Korean and French, we might expect French-speaking L2 learners of Korean to be more successful in their use of the AP-final H tone as a cue to the end of phrase-final words than in their use of the AP-initial L tone as a cue to the beginning of phrase-initial words in Korean. Importantly, Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis predicts that the phonological similarities between the Korean and French intonations should prevent French listeners from learning the fine-grained phonetic differences between French and Korean. While this is true for both the AP-final H tone and the AP-initial L tone, French listeners' predicted that the non-target-like phonetic representation of the AP-final H tone is not expected to impair their segmentation of Korean, as this H tone would still occur in the final syllable of Korean APs and thus would still signal the end of phrase-final words in Korean (like French). By contrast, French listeners' predicted non-target-like phonetic representation of the AP-initial L tone would prevent them from learning that the AP-initial L tone can signal the beginning of phrase-initial words in Korean (unlike French), thus impairing their ability to use the AP-initial L tone to locate the beginning of phrase-initial words in Korean.

As briefly mentioned above, tonal cues have also been found to enhance native English listeners' speech segmentation when an F0 rise corresponding to an H tone is heard in word-initial position (e.g., Tyler & Cutler, 2009). This finding was attributed to the statistical tendency for words to be stressed on the initial syllable in English (e.g., Clopper, 2002; Cutler & Carter, 1987), with the H tone being aligned with the stressed syllable (e.g., Beckman & Elam, 1997). Thus, for English listeners, learning to segment Korean speech into words involves learning a markedly different intonation and associating the AP-final H tone and the AP-initial L tone to, respectively, the end and beginning of phrase-final and phrase-initial words. Although both tones need to be learned, one might expect that it will be more difficult for English listeners to use the AP-initial L tone to locate the beginning of phrase-initial words in Korean given that word-initial boundaries are often signaled by an H tone in English. Crucially, the striking phonological differences between the Korean and the English intonations should enable English-speaking L2 learners of Korean to learn the differences between the two systems and use BOTH the AP-final H tone and the AP-initial L tone in the segmentation of Korean speech.

In order to test these predictions, we conducted a visual-world eye-tracking experiment that orthogonally manipulated the AP-final and AP-initial tones in Korean.

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

2. Method

2.1. Participants

A total of 34 adult native Korean listeners (mean age: 25.5, standard deviation (SD): 2.5, 11 women), 24 French-speaking L2 learners of Korean (mean age: 26.6, SD: 4.2, 18 women), and 19 English-speaking L2 learners of Korean (mean age: 29.1, SD: 6.4, 9 women) participated in this study.[1] All participants were tested at a university in Seoul, South Korea, and none reported having speech impairments.

All Korean listeners identified Seoul Korean as their native dialect; all had parents who spoke Korean as their L1, and all spent their childhood and adolescence in Korea. All French listeners had parents who spoke French as their L1, and all lived in a French-speaking country during their childhood and adolescence (20 spent their childhood and adolescence in France and 2 in Belgium; 2 spent some of their childhood in Côte d'Ivoire and their adolescence in France). All English listeners had parents who spoke English as their L1, and all lived in an English-speaking country during their childhood and adolescence (13 spent their childhood and adolescence in the US, 2 in Canada, 1 in the UK, 1 in Ireland, 1 in Australia, and 1 in Australia and Singapore).

The L2 learners completed a mock reading test of the Test of Proficiency in Korean II (60th TOPIK II – Reading – Mock Test; https://www.topikguide.com/topik-ii-reading-mock-test-1/). TOPIK II targets Levels 3–6 (intermediate to advanced) in Korean. The mock test is out of 50, and the passing score for Level 3 (lower intermediate) is 20.[2] The L2 learners' proficiency scores are presented in Table 1, together with their age of first listening exposure to Korean, number of years of Korean instruction, and months spent in Korea. As can be seen from the proficiency scores in Table 1, English listeners reached the passing score for Level 3 (lower intermediate), but French listeners did not quite do so. However, an independent-samples *t*-test assuming unequal variances revealed that the two L2 groups did not differ significantly from each other in their proficiency in Korean ($t[32] = -1.45$, $p = .079$). An independent-samples *t*-test assuming unequal variances also found the two L2 groups not to differ from each other in their number of years of Korean instruction ($t[32] = -1.43$, $p = .081$). An independent-samples *t*-test assuming equal variances further showed that the two L2 groups did not differ from each other in their age of first listening exposure to Korean ($t<|1|$). However, an independent-samples *t*-test assuming unequal variances revealed that the two L2 groups differed from each other in the number of months they have spent in Korea ($t[21] = -2.7$, $p < .007$), with English listeners having spent more time in Korea compared to French listeners. Thus, other than for the amount of time they had spent in Korea, the two L2 groups were comparable in their proficiency in and experience with Korean. Although the observed numerical trends (proficiency in Korean, years of Korean instruction) and significant difference (months spent in Korea) would be problematic given the predictions formulated in the preceding section, they are not given the results that were ultimately found (for details, see the Results and Discussion sections).[3]

2.2. Materials

Experimental sentences were created that contained a temporary lexical ambiguity between a disyllabic target word in post-boundary AP-initial position and a disyllabic competitor word spanning the AP boundary. The experimental sentences had the structure $[\text{Adverb}]_{AP}$ $[\text{Subject} + \text{case-marker}]_{AP}$ $[\text{Object} + \text{case-marker}]_{AP}$ $[\text{Verb}]_{AP}$. In these sentences, the subject always ended with the case marker *–ka, –i,* or *–to*. The target word was the following disyllabic object (e.g., *masul* 'magic'), and the competitor word was a disyllabic word that began with the same syllable as the case marker and ended with the first syllable of the target word (e.g., *kama* 'palanquin'). The segmental competitor was thus heard before the target word was heard. All sentential contexts preceding the target word were semantically compatible with both the target and competitor words (e.g., *paŋkim sɛsinpu-ka masul-il…* 'just-now the-new-bride-subj magic-obj …' vs. *paŋkim sɛsinpu-ka kama…* 'just now the new-bride's palanquin…'). The experiment contained 36 experimental sentences, with the above three case markers each appearing in 12 sentences.

The 36 experimental sentences were interspersed with 108 filler sentences, 8 of which were used in the practice session. Of these filler sentences, 36 had the same sentence structure and the target word in the same position as the experimental sentences, but the subject did not contain a case marker, and the target word instead began with the syllables *ka–, i–,* or *to–,* each in 12 sentences. In these filler sentences, the competitor word began with the second syllable of the target word (e.g., target: *kaʧi* 'eggplant'; competitor: *ʧiʧin* 'earthquake'). The remaining filler sentences had a similar sentence structure but the location of the adverb varied across sentences. Of these sentences, 40 had the target word in subject position and 32 sentences had the target word in the adverb position. For these sentences, the competitor word overlapped with the target word in its first syllable (e.g., target: *namu* 'tree'; competitor: *napi* 'butterfly').

The visual display contained orthographic representations of the target and competitor words and of two distracter words (for a validation of the use of orthography in visual-world eye-tracking experiments, see McQueen & Viebahn, 2007). The distracters were phonologically and semantically unrelated to the target and competitor, and showed the same type of segmental overlap with each other as did the target and competitor (e.g., for the experimental items, the first syllable of one distractor was the second syllable of the other distracter). All the auditory sentences and visual words used in the experiment can be found in Appendix A of the Supplementary Materials (Supplementary Materials).

Three repetitions of the sentences were recorded by a female native speaker of Korean. The experimental sentences were then resynthesized in order to create four tonal boundary conditions: H#L, H#H, L#L, and L#H, where # is the AP boundary. In the natural productions, the subject ended with the AP-final H tone and the object began with the AP-initial L tone. For each sentence, the stimulus in the H#L condition was created by mimicking the H#L tones from a different recording of the same sentence; the stimulus in the H#H condition was created by extending the AP-final H tone of the resynthesized H#L stimulus to the following AP-initial syllable, with a slight decline over the AP-initial H tone so that the rest of the contour would sound natural; the stimulus in the L#L condition was created by extending the AP-initial L tone of the resynthesized H#L stimulus to the previous AP-final syllable; and the stimulus in the L#H condition was created by reversing the AP-final H tone and the AP-initial L tone of the resynthesized H#L stimulus, with a slight decline over the AP-initial H tone so that the rest of the contour would sound natural. The filler sentences were similarly resynthesized so that the experimental sentences would not stand out. The resynthesis was

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

Table 1. L2 Learners' Proficiency and Language Background Information

	Proficiency Score (/50)	Age of First Exposure to Korean (Years)	Duration of Korean Instruction (Years)	Time in a Korean-speaking Environment (Months)
French listeners	19.0 (8.6)	19.5 (4.9)	2.5 (2.2)	22.5 (19.7)
English listeners	23.9 (11.7)	20.4 (4.3)	3.7 (3.1)	60.4 (56.8)

Note. Mean (SD)

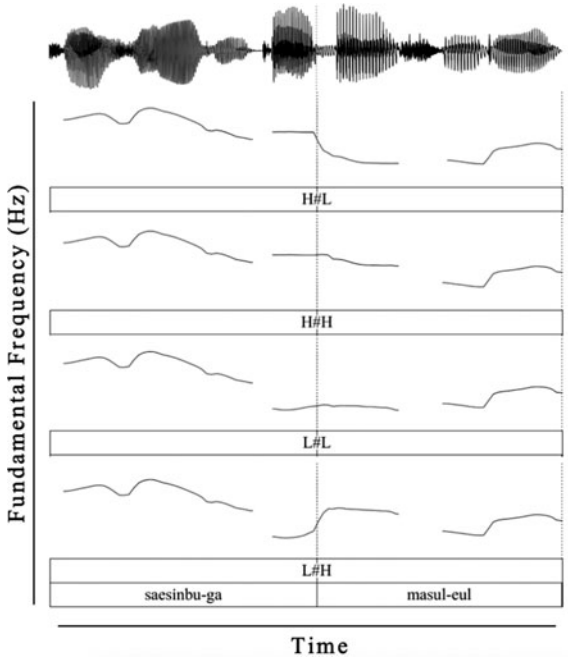


Fig. 1. Subject-object phrases from an example sentence in all four tonal boundary conditions

done manually using the PSOLA function in Praat (Boersma & Weenink, 2017). Figure 1 shows the four pitch contours created for an example experimental sentence.

The experimental items were distributed in four lists, with participants hearing each sentence in only one tonal condition. The four tonal conditions were counterbalanced across lists and interspersed with the 108 filler sentences.

2.3. Procedures

Experiment Builder software (SR Research) was used to create and administer the eye-tracking experiment, and Eyelink software (SR Research) was used to monitor participants' eye movements. Eye movements were recorded at a sampling rate of 250 Hz using the head-mounted EyeLink II eye-tracker. The stimuli were heard with an ASIO-compatible sound card, ensuring accurate audio timing in relation to the recording of eye movements.

In each trial, participants saw four orthographic words in a (non-displayed) 2 x 2 grid for 3,000 milliseconds before a fixation cross appeared in the middle of the screen for 500 milliseconds; as the fixation cross disappeared, the four words reappeared on the screen in their original position and the auditory stimulus was heard (synchronously) over headphones. Participants were instructed to click on the target word with the mouse as soon as they heard the target in the stimulus. Their eye movements were recorded from the onset of the case marker (e.g., *–ka ma…*). The trial ended with the participants' response, followed by an interval of 1,000 milliseconds.

The eye-tracking experiment began with the practice session (8 trials) followed by the main experiment (136 trials). The 36 experimental trials and 100 filler trials from the main experiment were presented in four blocks (34 trials per block), with each block containing 9 experimental trials. The order of the experimental and filler trials within a block and the order of blocks were randomized across participants. The participants were offered to take a break after the second block. The experiment lasted approximately 20 minutes.

After completing the experiment, the L2 learners filled out a word familiarity questionnaire that contained the target and competitor words used in the experiment. They rated each word on a scale from 0 ("I have never seen/heard this word") to 4 ("I have frequently seen/heard this word, I know what it means, and I can provide a definition for it"). The two L2 groups did not differ in their ratings of the words (French: mean = 2.18, SD = 0.6, English: mean = 2.21, SD = 0.6, $t<|1|$).[4]

2.4. Data analysis

Experimental trials that received distracter responses or no response, or for which eye movements could not reliably be tracked, were excluded from the analyses. This resulted in the exclusion of 5.7% of all the trials (1.2% of the Korean listeners' data, 8.9% of the French listeners' data, and 9.5% of the English listeners' data).

For the remaining trials, we analyzed participants' eye movements in the four regions of interest, corresponding to the four orthographic words on the screen. Eye movements were analyzed from the onset of the case-marker (e.g., *–ka ma…*) to examine the effect of the AP-final tone, and from the onset of the target word (e.g., *masul*) to examine the effect of the subsequent AP-initial tone. Proportions of fixations to the target, competitor, and distracter words were extracted in 8-ms time windows from the specified onset to 1,400 ms post-onset for the purpose of initial data visualization. Statistical analyses were conducted on the difference between the empirical-logit-transformed proportions of target and competitor fixations over smoothed 50-ms time windows (for a similar analysis of visual-world eye-tracking data, see Creel, 2014). We will refer to this dependent variable as listeners' target-over-competitor fixation advantage.

Listeners' target-over-competitor fixation advantage was modeled using growth curve analysis (Mirman, 2014; Mirman, Dixon

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

& Magnuson, 2008). The growth curve analyses were run using the lme4 package in R (Bates, Maechler, Bolker & Walker, 2015) for fitting linear mixed-effects models. Initial analyses compared Korean listeners' fixations to French listeners' and English listeners' fixations, and additional analyses compared French listeners' fixations to English listeners' fixations. These analyses included the AP-final or AP-initial tone (high, low), listeners' L1 (Korean [if applicable], French, English), orthogonally derived time polynomials (linear, quadratic, cubic), and their interactions as fixed effects, with Korean listeners' performance in the high tone condition as baseline for the initial analyses and with French listeners' performance in the high tone condition as baseline for the additional analyses. A backward-fitting function from the package LMERConvenienceFunctions (Tremblay & Ransijn, 2015) was then used to identify the model that accounted for significantly more of the variance than all simpler models, as determined by log-likelihood ratio tests. Only the results of the model with the best fit are presented (for a discussion of this approach, see Mirman, 2014), with p values being calculated using the lmerTest package in R (Kuznetsova, Brockhoff & Christensen, 2016). All analyses included participant as random intercept and the time polynomials as random slopes, thus modeling a different curve for each participant. The first set of analyses focuses on the effect of the AP-final tone (across AP-initial tones) and the second set of analyses focuses on the effect of the AP-initial tone (across AP-final tones).

If tonal information modulates lexical access, in addition to showing a significant effect of tonal information, listeners' fixation curves should show a significant interaction between the AP-final and/or AP-initial tone and at least one of the time polynomials. More specifically, we would expect listeners' fixation curves to be less ascending (linear time polynomial), more 'U'-shaped (quadratic time polynomial), and/or less 'S'-shaped (cubic time polynomial) in the non-enhancing tone condition (L for the AP-final tone and H for the AP-initial tone) than in the enhancing tone condition (H for the AP-final tone and L for the AP-initial tone). The shallower and/or more 'U'-shaped characteristics of the fixation curve would be due to greater interference from the competitor word, and the less 'S'-shaped characteristic of the fixation curve would be due to listeners' target fixations reaching an asymptote in the enhancing tone condition but not in the non-enhancing tone condition.

We expect Korean listeners' fixation curves to show an inhibiting effect of the AP-final L tone compared to the AP-final H tone and an enhancing effect of the AP-initial L tone compared to the AP-initial H tone. English listeners' fixation curves are predicted to show the same pattern of results, though it may take these listeners more time to integrate the use of these tonal cues given the phonological (and phonetic) differences between the Korean and English intonations (for such results, see Tremblay et al., 2016). This may be especially true of the AP-initial tone, since word beginnings are often signaled by an H tone in English. Importantly, following Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis, we predict that French listeners would show the same pattern of results for the AP-final tone but not for the AP-initial tone. As indicated earlier, French listeners' predicted non-target-like phonetic representation of the AP-initial L tone is expected to inhibit their learning that the AP-initial L tone can signal the beginning of phrase-initial words in Korean (unlike French), thus impairing their ability to use the AP-initial L tone to locate the beginning of phrase-initial words in Korean. In other words, the phonological similarities between the Korean and French intonations are predicted to pose a learnability problem for French listeners' use of the AP-initial L tone in Korean speech segmentation.

3. Results

Korean, French, and English listeners' raw proportions of target, competitor, and distracter fixations over 8-ms time windows from the onset of the manipulated tonal conditions (i.e., from the onset of the case marker) to 1,400 ms post onset can be found in Figure B1 of Appendix B of the Supplementary Materials (Supplementary Materials).

3.1. AP-final tone

Korean, French, and English listeners' empirical-logit-transformed target-over-competitor fixation advantage from the onset of the AP-final tone (i.e., from the onset of the case marker, which has an average duration of 114 ms.) is shown in Figure 2. Values above 0 mean that listeners fixated the target word more than the competitor word. The solid line represents listeners' actual data, whereas the dashed line represents listeners' predicted target-over-competitor fixation advantage according to the growth curve analysis with the best fit (presented next, see Table 2). The shading represents one standard error above and below the mean. Recall that the phonologically canonical AP-final tone is H, so the directionality of the predicted tonal effect for the AP-final tone is H > L.

The growth-curve analysis with the best fit on all listeners' target-over-competitor fixation advantage from the onset of the AP-final tone to 1,400 ms post onset is presented in Table 2. The model included the following fixed effects: the simple effects of time (linear, quadratic, cubic), AP-final tone, and L1; the two-way interactions between time (linear, quadratic) and AP-final tone, between time (linear, quadratic, cubic) and L1, and between AP-final tone and L1; and the three-way interactions between time (linear, quadratic), AP-final tone, and L1.

The relevant significant effects from Table 2 can be summarized as follows. The negative estimate for the significant simple effect of AP-final tone indicates that Korean listeners showed lower proportions of fixations when the AP-final tone was L than when it was H. The negative and positive estimates for the significant interactions between, respectively, the linear and quadratic time polynomials and the AP-final tone mean that Korean listeners' fixation curve became less ascending and more 'U'-shaped when the AP-final tone was L than when it was H. Additionally, the significant three-way interactions between the linear and quadratic time polynomials, the AP-final tone, and L1 (both French and English) indicate that both French and English listeners differed from Korean listeners in the effect of AP-final tone as a function of time. The negative estimates for the interactions with the linear time polynomial suggest that French and English listeners showed a stronger interaction (in the same direction) between the linear time polynomial and the AP-final tone compared to Korean listeners (whose corresponding interaction also had a negative estimate), and the negative estimates for the interactions with the quadratic time polynomial suggests that French and English listeners showed a weaker or reverse interaction between the quadratic time polynomial and the AP-final tone compared to Korean listeners (whose corresponding interaction had a positive estimate).

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

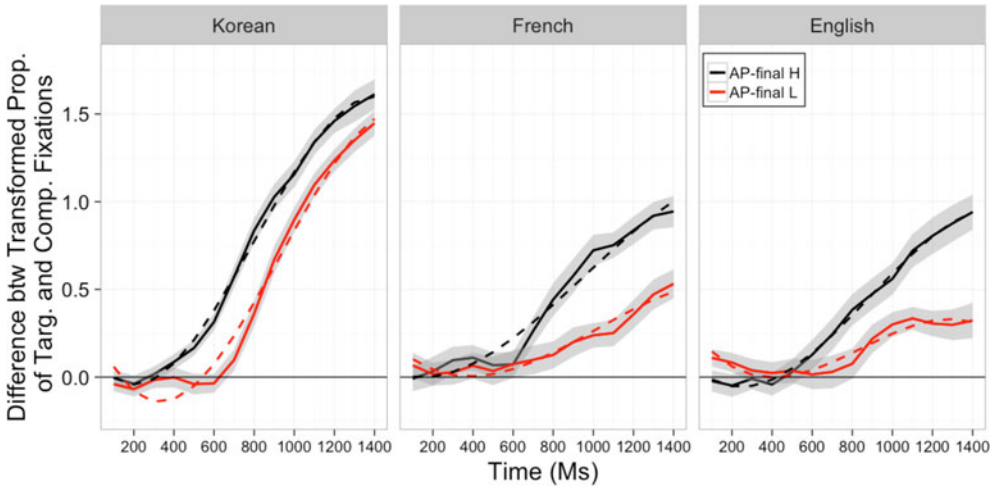


Fig. 2. Difference between listeners' transformed proportions of target and competitor fixations from the onset of the AP-final tone; the black and red lines represent listeners' fixations when the AP-final tone was, respectively, canonical (H) and non-canonical (L); the solid lines represent listeners' data, and the dashed lines represent listeners' predicted data based on the model with the best fit (Table 2); the shading represents one standard error above and below the mean

These results are as predicted for Korean listeners, whose target-over-competitor fixation advantage was inhibited when the AP-final tone changed from a phonologically canonical H tone to a non-canonical L tone (across AP-initial tones). This confirms that Korean listeners' speech segmentation benefits from an AP-final H tone. Importantly, French and English listeners' fixation curves differed from Korean listeners' fixation curves in the strength and possibly directionality of the effect of AP-final tone as a function of time. Note that the lack of significant interaction between the AP-final tone and L1 for either L2 group suggests that French and English listeners did not differ from Korean listeners in the overall effect of AP-final tone, with the L tone inhibiting the three groups' target-over-competitor fixation advantage across time, as predicted; the L2 learners differed from native listeners only in the precise shape of their fixation curves.

To understand how L2 learners' fixation curves differed from those of Korean listeners, we ran an additional analysis that directly compared French and English listeners' fixation curves; this analysis allowed us to examine the strength and directionality of the interactions between time and AP-final tone for the two L2 groups while also comparing the L2 groups to each other. The growth-curve analysis with the best fit on L2 listeners' target-over-competitor fixation advantage from the onset of the AP-final tone to 1,400 ms post onset is presented in Table 3. The fixed effects included in the model were the simple effects of time (linear, quadratic, cubic) and AP-final tone, and the two-way interaction between time (linear) and AP-final tone. For this analysis, the alpha level was adjusted to .025.

The important significant effects from Table 3 can be described as follows. The negative estimate for the significant simple effect of AP-final tone indicates that French listeners showed lower proportions of fixations when the AP-final tone was L than when it was H. The negative estimate for the significant interaction between the linear time polynomial and the AP-final tone indicates that French listeners' fixation curve was less ascending when the AP-final tone was L than when it was H, with the estimate for this effect being larger in size than the corresponding effect observed for Korean listeners (see Table 2). Crucially, the model with the best fit did not include any effect of L1 or any interaction between L1 and AP-final tone or between L1, AP-final tone, and the time polynomials. This means that the simple effect of AP-final tone and the two-way interaction between the linear time polynomial and the AP-final tone was true of both French and English listeners.

These results confirm the stronger interaction between the linear time polynomial and the AP-final tone for French and English listeners compared to Korean listeners (hypothesized from Table 2), suggesting that L2 learners had more difficulty recovering from the misleading AP-final L tone. It appears that, because of this difficulty, French and English listeners' fixation curves in the AP-final L tone condition did not have more of a 'U'-shape than their fixation curves in the AP-final H tone condition, unlike Korean listeners' fixation curves. Thus, although the AP-final tone had different effects on the precise shape of native and L2 listeners' fixation curves, the consequence of these effects was similar – to inhibit listeners' target-over-competitor fixation advantage when the AP-final tone deviated from the phonologically canonical H tone, as predicted. Importantly, the L2 groups did not differ from each other in their ability to use the AP-final tone as a cue to the end of phrase-final words in Korean.

3.2. AP-initial tone

Korean, French, and English listeners' empirical-logit-transformed target-over-competitor fixation advantage from the onset of the AP-initial tone (i.e., from the offset of the case marker and onset of the target word) is shown in Figure 3. Again, values above 0 mean that listeners fixated the target word more than the competitor word. The solid line represents listeners' actual data, and the dashed line represents listeners' predicted target-over-competitor fixation advantage according to the growth curve analysis with the best fit (presented next, see Table 4). The shading represents one standard error above and below the mean. Recall that the phonologically canonical AP-initial tone is L in Korean, so the directionality of the predicted tonal effect for the AP-initial tone is L > H.

The growth-curve analysis with the best fit on all listeners' target-over-competitor fixation advantage from the onset of the AP-initial tone to 1,400 ms post onset is presented in Table 4.

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

Table 2. Growth Curve Analysis with Best Fit on All Listeners' Target-over-Competitor Fixation Advantage from the Onset of the AP-Final Tone

Effect	Estimate	Std. Error	*t*	*p*
(Intercept)	0.720	0.029	24.760	<.001
Time				
Linear	2.239	0.109	20.590	<.001
Quadratic	0.159	0.075	2.120	.036
Cubic	−0.355	0.050	−7.110	<.001
AP-Final Tone (L)	−0.224	0.022	−10.270	<.001
L1 (French)	−0.304	0.045	−6.730	<.001
L1 (English)	−0.361	0.049	−7.420	<.001
Time x AP-Final Tone (L)				
Linear	−0.216	0.081	−2.660	.008
Quadratic	0.404	0.081	4.960	<.001
Time x L1 (French)				
Linear	−0.971	0.169	−5.740	<.001
Quadratic	0.020	0.117	<|1|	>.1
Cubic	0.254	0.078	3.280	.002
Time x L1 (English)				
Linear	−0.940	0.182	−5.170	<.001
Quadratic	0.056	0.125	<|1|	>.1
Cubic	0.174	0.083	2.090	.040
AP-Final Tone (L) x L1 (French)	−0.010	0.034	<|1|	>.1
AP-Final Tone (L) x L1 (English)	0.020	0.036	<|1|	>.1
Time x AP-Final Tone (L) x L1 (French)				
Linear	−0.497	0.127	−3.930	<.001
Quadratic	−0.351	0.127	−2.770	.006
Time x AP-Final Tone (L) x L1 (English)				
Linear	−0.696	0.136	−5.110	<.001
Quadratic	−0.467	0.136	−3.430	.001

Table 3. Growth Curve Analysis with Best Fit on L2 Listeners' Target-over-Competitor Fixation Advantage from the Onset of the AP-Final Tone

Effect	Estimate	Std. Error	*t*	*p*
(Intercept)	0.391	0.025	15.900	<.001
Time				
Linear	1.282	0.087	14.690	<.001
Quadratic	0.196	0.061	3.240	.002
Cubic	−0.136	0.044	−3.130	.003
AP-Final Tone (L)	−0.220	0.020	−11.180	<.001
Time x AP-Final Tone (L)				
Linear	−0.801	0.074	−10.870	<.001

The model included the following fixed effects: the simple effects of time (linear, quadratic, cubic), AP-initial tone, and L1; the two-way interactions between time (linear, quadratic) and AP-initial tone, between time (linear, quadratic) and L1, and between AP-initial tone and L1; and the three-way interactions between time (linear, quadratic), AP-initial tone, and L1.

The relevant significant effects from Table 4 are the following. The positive estimate for the significant effect of AP-initial tone means that Korean listeners' overall proportion of fixations was higher when the AP-initial tone was L than when it was H. The negative estimate for the significant interaction between the quadratic time polynomial and the AP-initial tone indicates that Korean listeners' fixation curve had less of a 'U' shape when the AP-initial tone was L than when it was H. The negative estimate for the significant interaction between AP-initial tone and L1 for English listeners means that English listeners' overall fixations showed a weaker or reverse effect of AP-initial tone compared to Korean listeners' fixations (whose corresponding effect had a positive estimate). Moreover, the significant three-way interactions between the linear and quadratic time polynomials, the AP-initial tone, and L1 (both French and English) indicate that both French and English listeners differed from Korean listeners in the effect of AP-initial tone they showed as a function of time. The positive estimates for the interactions with the linear time polynomial suggest that French and English listeners showed a stronger interaction (in the same direction) between the linear time polynomial and the AP-initial tone compared to Korean listeners (whose corresponding interaction also had a positive estimate), and the positive estimates for the interactions with the quadratic time polynomial suggests that French and English listeners showed a weaker or reverse interaction between the quadratic time polynomial and the AP-initial tone compared to Korean listeners (whose corresponding interaction had a negative estimate).

These results are again as predicted for Korean listeners, whose target-over-competitor fixation advantage was enhanced when the AP-initial tone changed from a phonologically non-canonical H to a canonical L (across AP-final tones). This confirms that an AP-initial L tone helps Korean listeners locate the beginning of phrase-initial words in continuous speech. The English listeners' weaker effect of AP-initial tone compared to Korean listeners is likely due to the reversal of the AP-initial tone effect from 200 to 700 ms in English listeners' fixations. As with the AP-final tone, French and English listeners' fixation curves differed from Korean listeners' fixation curve in the strength and possibly directionality of the effect of the AP-initial tone as a function of time.

To understand these differences, we again ran an additional analysis that directly compared French and English listeners' fixation curves; as with our analysis for the AP-final tone, this analysis allowed us to examine the strength and directionality of the interactions between time and the AP-initial tone for the two L2 groups while also comparing the L2 groups to each other. The growth-curve analysis with the best fit on L2 listeners' target-over-competitor fixation advantage from the onset of the AP-initial tone to 1,400 ms post onset is presented in Table 5. The fixed effects included in the model were: the simple effects time (linear, cubic), AP-initial tone, and L1; and the two-way interactions between time (linear, cubic) and AP-initial tone, and between AP-initial tone and L1. The alpha level for this analysis was adjusted to .025.

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

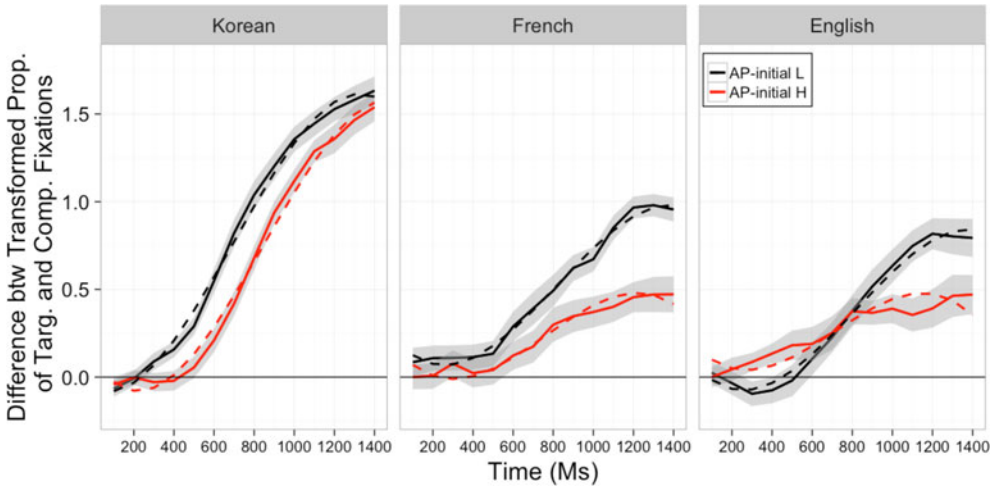


Fig. 3. Difference between listeners' transformed proportions of target and competitor fixations from the onset of the AP-initial tone; the black and red lines represent listeners' fixations when the AP-initial tone was, respectively, canonical (L) and non-canonical (H); the solid lines represent listeners' data, and the dashed lines represent listeners' predicted data based on the model with the best fit (Table 4); the shading represents one standard error above and below the mean

Table 4. Growth Curve Analysis with Best Fit on All Listeners' Target-over-Competitor Fixation Advantage from the Onset of the AP-Initial Tone

Effect	Estimate	Std. Error	*t*	*p*
(Intercept)	0.623	0.031	20.230	<.001
Time				
Linear	2.082	0.108	19.360	<.001
Quadratic	0.275	0.077	3.580	<.001
Cubic	−0.276	0.036	−7.650	<.001
AP-Initial Tone (L)	0.189	0.023	8.310	<.001
L1 (French)	−0.376	0.047	−7.960	<.001
L1 (English)	−0.347	0.051	−6.840	<.001
Time x AP-Initial Tone (L)				
Linear	0.096	0.085	1.130	>.1
Quadratic	−0.414	0.085	−4.860	<.001
Time x L1 (French)				
Linear	−1.335	0.159	−8.400	<.001
Quadratic	−0.262	0.119	−2.200	.030
Time x L1 (English)				
Linear	−1.450	0.171	−8.490	<.001
Quadratic	−0.361	0.128	−2.810	.006
AP-Initial Tone (L) x L1 (French)	0.062	0.035	1.760	.078
AP-Initial Tone (L) x L1 (English)	−0.109	0.038	−2.870	.004
Time x AP-Initial Tone (L) x L1 (French)				
Linear	0.495	0.132	3.750	<.001
Quadratic	0.537	0.132	4.070	<.001
Time x AP-Initial Tone (L) x L1 (English)				
Linear	0.604	0.142	4.260	<.001
Quadratic	0.635	0.142	4.480	<.001

Table 5. Growth Curve Analysis with Best Fit on L2 Listeners' Target-over-Competitor Fixation Advantage from the Onset of the AP-Initial Tone

Effect	Estimate	Std. Error	*t*	*p*
(Intercept)	0.240	0.034	7.160	<.001
Time				
Linear	0.581	0.090	6.430	<.001
Cubic	−0.075	0.059	−1.270	>.1
AP-Initial Tone (L)	0.251	0.028	9.100	<.001
L1 (English)	0.028	0.050	<|1|	>.1
Time x AP-Initial Tone (L)				
Linear	0.640	0.077	8.310	<.001
Cubic	−0.214	0.077	−2.780	.006
AP-Initial Tone (L) x L1 (English)	−0.171	0.041	−4.120	<.001

The important significant effects in Table 5 can be summarized as follows. The positive estimate for the significant simple effect of AP-initial tone indicates that French listeners had higher proportions of fixations when the AP-initial tone was L than when it was H. The positive estimate for the significant interaction between the linear time polynomial and the AP-initial tone indicates that French listeners' fixation curve was more ascending when the AP-initial tone was L than when it was H, with the estimate for this effect being much larger in size than the corresponding effect observed for Korean listeners (see Table 4). The negative estimate for the significant interaction between the cubic time polynomial and the AP-initial tone indicates that French listeners' fixation curve was more 'S' shaped when the AP-initial tone was L than when it was H. (Note, however, that the AP-initial tone was not found to interact with both the cubic time polynomial and L1 in the previous analysis; see Table 4.) Importantly, the negative estimate for the two-way interaction between the AP-initial tone and L1 suggests that English

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

listeners' overall fixations show a weaker or reverse effect of AP-initial tone compared to French listeners. To further examine the effect of AP-initial tone in English listeners' fixations, the model in Table 5 was releveled with English listeners as baseline. The releveled model revealed a significant effect of AP-initial tone with a positive estimate ($\beta = 0.080$, SE = 0.031, $t = 2.585$, $p < .009$), indicating that English listeners' overall target-over-competitor fixation advantage is indeed greater when the beginning of phrase-final words is signaled by an AP-initial L tone. The lack of three-way interaction between the time polynomials, AP-initial tone, and L1 in the analysis of L2 learners' data suggests that the two L2 groups showed similar fixation curves despite the advantage for the AP-initial H tone early on in English listeners' fixations.

These results confirm the stronger interaction between the linear time polynomial and the AP-initial tone for French and English listeners compared to Korean listeners (hypothesized from Table 4), suggesting that L2 learners had more difficulty recovering from the misleading (i.e., non-canonical) AP-initial H tone. This may have caused French and English listeners' fixation curves in the AP-initial L tone condition not to have less of a 'U'-shape than their fixation curves in the AP-initial H tone condition, unlike Korean listeners' fixation curves. Hence, although the AP-initial L tone had different effects on the exact shape of native and L2 listeners' fixation curves, the consequence of these effects was similar – to enhance listeners' target-over-competitor fixation advantage. Crucially, English listeners differed from both Korean listeners and French listeners in the size of the effect of the AP-initial tone due to a reversal of the effect early on in the trial. These L2 results are not expected: French listeners did not experience the predicted difficulty in using the AP-initial tone to locate the beginning of phrase-initial words in Korean, and English listeners' weaker effect of AP-initial tone caused by the early advantage for words beginning with an AP-initial H tone suggests that they experienced more difficulty in using this tone than French listeners.

4. Discussion

The present study investigated whether the learning of intonational cues to word boundaries in speech segmentation would be more difficult if the L1 and L2 had phonologically similar but phonetically different intonations than if they had phonologically different intonations, thus providing another testbed for evaluating Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis. It examined how native Korean listeners, French-speaking L2 learners of Korean, and English-speaking L2 learners of Korean use tonal information in the segmentation of Korean speech. Participants completed a visual-world eye-tracking experiment that tested whether listeners' speech segmentation would be inhibited by a phonologically non-canonical AP-final L tone (relative to a canonical AP-final H tone) and enhanced by a phonologically canonical AP-initial L tone (relative to a non-canonical AP-initial H tone). Both Korean and English listeners were predicted to show such effects, with English listeners integrating the use of tonal cues later on in the word recognition process due to the great phonological differences between the L1 and L2 intonation systems. French listeners were expected to show an inhibitory effect of the phonologically non-canonical AP-final L tone despite their predicted non-target-like phonetic representation of the AP-final H tone, because this H tone occurs in the final syllable of Korean APs and thus still signals the end of phrase-final words in Korean (like French). However, French listeners were not expected to show an enhancing effect of the canonical AP-initial L tone because their predicted non-target-like phonetic representation of the AP-initial L tone would inhibit their learning that the AP-initial L tone can signal the beginning of phrase-initial words in Korean (unlike French). This predicted difficulty was hypothesized to come from the phonological similarity but fine-grained phonetic differences between the Korean and French intonations.

As expected, the results first showed that Korean listeners' speech segmentation was modulated by both the AP-final and the AP-initial tones, with listeners' target-over-competitor fixation advantage being greater in the phonologically canonical tone conditions (AP-final H tone, AP-initial L tone) than in the phonologically non-canonical tone conditions (AP-final L tone, AP-initial H tone), as predicted. These tonal effects manifested themselves early on in the word-recognition process, suggesting that listeners make rapid use of this tonal information to distinguish between the target and competitor words. These results replicate the finding of earlier research that the AP-final H tone and the AP-initial L tone enhance Korean listeners' speech segmentation (Kim et al., 2012; Kim & Cho, 2009; Tremblay et al., 2019). These findings suggest that listeners compute the prosodic structure of a given utterance by exploiting tonal patterns before and after a hypothesized lexical boundary, with the AP-final and AP-initial tonal cues affecting listeners' target-over-competitor fixation and modulating lexical access.

Contrary to the predictions made on the basis of Tremblay et al. (2016)'s Prosodic-Learning Interference Hypothesis, however, the results also showed that French listeners' speech segmentation displayed a target-like effect of tonal information for both the AP-final tone and the AP-initial tone, with listeners' target-over-competitor fixation advantage being greater in the phonologically canonical tone conditions (AP-final H tone, AP-initial L tone) than in the phonologically non-canonical tone conditions (AP-final L tone, AP-initial H tone). These results were *not* predicted by the Prosodic-Learning Interference Hypothesis, as French employs an AP-initial L tone that is phonologically similar to, yet phonetically different from, the corresponding AP-initial L tone in Korean, with the French tone not being closely anchored to the AP-initial syllable, unlike that in Korean. Therefore, it is not necessarily the case that the learning of intonational cues to word boundaries in speech segmentation is difficult to achieve if the L1 and L2 have phonologically similar but phonetically different intonations, contrary to Tremblay et al. (2016)'s proposed Prosodic-Learning Interference Hypothesis. These results are particularly striking given that the French participants did not even reach the lower intermediate proficiency threshold in the Korean reading proficiency test that they completed.

One important difference between the present study and that of Tremblay et al. (2016), however, is that the current L2 learners were tested in an environment where the target language was spoken and had spent more time in an L2-speaking environment. In other words, although the participants did not have advanced proficiency in Korean, they had daily exposure to Korean. If the Prosodic-Learning Interference Hypothesis is more likely to affect L2 learners who have less exposure to the L2, it may have made it difficult for the Korean-speaking L2 learners of French in Tremblay et al. (2016) to learn the properties of French intonation (relative to English-speaking L2 learners of French). Crucially, perceptual assimilation difficulties may not be persistent for the learning of intonational information, with immersed L2 learners

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

being able to perceive subtle L1-L2 intonational differences, as shown in the present study. This prediction, if correct, would suggest that the mechanisms that underlie the learning of intonational and segmental information are similar, but the nature of the information to be learned differs such that learning difficulties are less persistent for intonational information than for segmental information. More precisely, it may be the case that non-proficient L2 learners assimilate phonologically similar L2 tones to existing L1 tones like they do with phonologically similar L2 and L1 segments (e.g., Best & Tyler, 2007; Flege, 1995; van Leussen & Escudero, 2015), but that, with sufficient exposure to the L2, they become more sensitive to the fine-grained phonetic details of intonation compared to those of segments. This difference may be due to the less categorical and non-lexical nature of intonational information compared to that of segmental information, and possibly to the greater ease with which intonational information can be extracted from the speech signal compared to segmental information (after all, newborn infants already show a preference for the prosody of the birth mother's language; e.g., Mehler & Dupoux, 1994). Future research should determine whether learning difficulties are indeed less persistent for intonational information than for segmental information.

The results also revealed that English listeners did not differ from French listeners in their use of the AP-final tone in speech segmentation but showed a weaker effect of AP-initial tone due to a reversal of the tonal effect early on in the word recognition process. This reversal, with English listeners' target-over-competitor fixation advantage being greater when the AP-initial tone was H than when it was L, is likely a transfer effect from an L1-based speech segmentation routine, with English listeners' early word activation benefiting from a word-initial H tone before showing the effect in the correct direction for Korean. One straightforward explanation of these results is that the learning of intonational cues may be more difficult if it requires listeners to suppress an L1-based relationship between a cue and a word edge than if it requires them to learn a new, L2-based relationship between a cue and a word edge. Functionally monolingual English listeners have been shown to use F0 rise as a cue to word-initial boundaries but not as a cue to word-final boundaries (e.g., Tremblay, Namjoshi, Spinelli, Broersma, Cho, Kim, Martínez-García & Connell, 2017; Tyler & Cutler, 2009). For English listeners, therefore, suppressing the association between an H tone and word-initial boundaries is more difficult to accomplish than learning a new association between an H tone and the last syllable of a phrase-final word. Previous studies on the use of segmental cues in speech segmentation have shown that L2 learners can learn new, L2-based speech segmentation routines but have difficulty suppressing L1-based segmentation routines (e.g., Tremblay & Spinelli, 2014; Weber & Cutler, 2006). The current results suggest that this may also be the case for the learning of intonational information.

5. Conclusion

The present study investigated how French-speaking and English-speaking L2 learners of Korean use intonational cues to locate word boundaries in Korean, with focus on whether the learning of intonational cues to word boundaries in speech segmentation is more difficult if the L1 and L2 have phonologically similar but phonetically different intonations than if they have phonologically different intonations (Tremblay et al., 2016). The results of an eye-tracking speech segmentation experiment with native Korean listeners and French-speaking and English-speaking L2 learners of Korean yielded the following findings: (i) both native and L2 listeners exploited intonational cues to word boundaries when segmenting Korean speech into words; (ii) the learning of L2 intonational cues that differ in subtle ways from L1 intonational cues is not necessarily difficult if L2 learners have had sufficient exposure to the L2 (French-speaking L2 learners of Korean); and (iii) it is more difficult to suppress segmentation routines that are based on L1 intonational cues than to learn segmentation routines that are based on L2 intonational cues (English-speaking L2 learners of Korean) (see also Tremblay & Spinelli, 2014; Weber & Cutler, 2006). More broadly, the current study provides a strong incentive for future research to examine the effect of linguistic experience on L2 learners' use of intonational cues in parallel with their use of segmental information (cf. Cho, McQueen & Cox, 2007) in speech segmentation.

Acknowledgements. This research is based upon work supported in part by the National Science Foundation (BCS-1423905, awarded to the first author) and by the Global Research Network Program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A2A2912410, awarded to the second and third authors).

Supplementary Material. Supplementary material can be found online at https://doi.org/10.1017/S136672892000053X

Notes

[1] Three additional native Korean listeners participated but could not be included in the analysis due to data loss; one additional French-speaking L2 learner of English participated but could not be included in the analysis due to data loss; two additional English-speaking L2 learners of Korean participated but were excluded from the analysis because one or both of their parents did not speak English as L1.

[2] The complete TOPIK II is composed of three subtests: a listening test, a reading test, and a writing test (100 points each; https://www.topikguide.com/topik-overview/). The passing score for Level 3 is 120/300. We have inferred that this passing score would correspond to 20/50 on the mock reading test that our participants completed.

[3] We do not think the English and French L2 learners of Korean differed in their Korean reading abilities for at least three reasons: (i) the two L2 groups did not differ in their Korean proficiency (the Korean proficiency test was administered in Korean, using the Korean writing system); (ii) the two L2 groups shared a the same alphabetic system, not placing either group at an advantage when learning to read and write in Korean; and (iii) the writing system of Korean is phonetically transparent and easy to learn: "The Korean alphabet is so simple that its sixteen totally distinct letters can be learned in minutes with the aid of the hangul-in-a-hurry chart" (Grant, 1979, p. 12; see also Kim, 2006; Park, 2008; Taylor & Taylor, 1995). It is thus very unlikely that the English and French L2 learners of Korean differed in their Korean reading abilities.

[4] We do not think that the two groups differed in their knowledge of the target and competitor words, as they did not differ in their familiarity with the target and competitor words. Crucially, even if they had, the orthographic presentation of the word ensures that participants can complete the task on the basis of phonetic information alone, which they master reading at a very early stage of learning (see Note 3). Word knowledge and/or familiarity is thus not essential to participants' successful completion of the task.

References

Bates, D, Maechler, B, Bolker, B and Walker, S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48.

Beckman, ME and Elam, GA (1997) *Guidelines for ToBI labeling*. The Ohio State University Research Foundation.

Beckman, ME and Pierrehumbert, J (1986) Intonational structure in English and Japanese. *Phonology Yearbook* **3**, 255–310.

Best, CT and Tyler, MD (2007) Nonnative and second-language speech perception. In MJ Munro & O-S Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). Amsterdam: John Benjamins.

Boersma, P and Weenink, D (2017) Doing phonetics by computer (Version 6.0.36). Retrieved from http://www.praat.org

Brown, M, Salverda, AP, Dilley, LC and Tanenhaus, MK (2011) Expectations from preceding prosody influence segmentation in online sentence processing. *Psychon Bull Rev* **18**, 1189–1196.

Brown, M, Salverda, AP, Dilley, LC and Tanenhaus, MK (2015a) Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance* **41**, 306–323.

Brown, M, Salverda, AP, Gunlogson, C and Tanenhaus, MK (2015b) Interpreting prosodic cues in discourse context. *Language, Cognition, and Neuroscience* **30**, 149–166.

Cho, T, McQueen, JM and Cox, EA (2007) Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* **35**, 210–243.

Christophe, A, Peperkamp, S, Pallier, C, Block, E and Mehler, J (2004) Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language* **51**, 523–547.

Clopper, CG (2002) Frequency of stress patterns in English: A computational analysis. *Indiana University Linguistics Club Working Papers Online, 2*. Retrieved from http://www.indiana.edu/iulcwp

Creel, SC (2014) Tipping the scales: auditory cue weighting changes over development. *Journal of Experimental Psychology: Human Perception and Performance* **40**, 1146–1160.

Cutler, A and Carter, DM (1987) The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language* **2**, 133–142.

Flege, JE (1995) Second language speech learning: Theory, findings, and problems. In W Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–273). Timonium, MD: York Press.

Grant, BK (1979) *Guide to Korean characters*. Seoul, South Korea: Hollym International Corporation.

Ito, K, Jincho, N, Minai, U, Yamane, N and Mazuka, R (2012) Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language* **66**, 265–284.

Ito, K and Speer, SR (2008) Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language* **58**, 541–573.

Jun, S-A (1998) The Accentual Phrase in the Korean prosodic hierarchy. *Phonology* **15**, 189–226.

Jun, S-A (2000) K-ToBI (Korean ToBI) labeling conventions. *UCLA Working Papers in Phonetics* **99**, 149–173.

Jun, S-A and Fougeron, C (2000) A phonological model of French intonation. In A Botinis (Ed.), *Intonation: Analysis, Modeling and Technology* (pp. 209–242). Dordrecht: Kluwer Academic Publishers.

Jun, S-A and Fougeron, C (2002) Realizations of accentual phrase in French intonation. *Probus* **14**, 147–172.

Kim, S (2006) Hangual and teaching pronunciation to beginners. *The Education of Korean Language* **18**, 217–244.

Kim, S, Broersma, M and Cho, T (2012) The use of prosodic cues in learning new words in an unfamiliar language. *Studies in Second Language Acquisition* **34**, 415–444.

Kim, S and Cho, T (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.

Kim, S, Mitterer, H and Cho, T (2018) A time course of prosodic modulation in phonological inferencing: The case of Korean post-obstruent tensing. *PLoS One* **13**, e0202912, 0202911-0202928.

Kuznetsova, A, Brockhoff, B and Christensen, H (2016) Tests in linear mixed effects models. Version 2.0.32. Retrieved from https://cran.r-project.org/web/packages/lmerTest/index.html

Ladd, DR (2012) *Intonational Phonology* Cambridge: Cambridge University Press.

McQueen, JM and Viebahn, MC (2007) Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology* **60**, 661–671.

Mehler, J and Dupoux, E (1994) *What infants know: The new cognitive science of early development*. Cambride, MA: Blackwell.

Michelas, A and D'Imperio, M (2010) Accentual phrase boundaries and lexical access in French. *Proceedings of Speech Prosody 2010*. Retrieved from http://speechprosody2010.illinois.edu/papers/100882.pdf

Mirman, D (2014) *Growth curve analysis and visualization using R*. Boca Raton, FL: Taylor & Francis.

Mirman, D, Dixon, JA and Magnuson, JS (2008) Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language* **59**, 475–494.

Mok, P, Yin, Y, Setter, J and Nayan, NM (2016) Assessing knowledge of English intonation patterns by L2 speakers. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings of the 2016 Speech Prosody Conference* (pp. 543–547). Boston.

Ortega-Llebaria, M and Colantoni, L (2014) L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition* **36**, 331–353.

Ortega-Llebaria, M, Nemogá, M and Presson, N (2015) Long-term experience with a tonal language shapes the perception of intonation in English words: How Chinese-English bilinguals perceive "Rose?" vs. "Rose". *Bilingualism: Language and Cognition* **20**, 367–383.

Ortega-Llebaria, M, Olson, DJ and Tuninetti, A (2018) Explaining Cross-Language Asymmetries in Prosodic Processing: The Cue-Driven Window Length Hypothesis. *Lang Speech*, 23830918808823.

Park, EC (2008) Literacy experience in Korean: Implications for learning to read in a second language. In K Koda & AM Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 201–221). New York: Routledge.

Pierrehumbert, J (1980) *The phonology and phonetics of English intonation*. Unpublished dissertation, MIT.

Pierrehumbert, J and Hirschberg, J (1990) The meaning of intonational contours in the interpretation of discourse. In P Cohen, J Morgan and M Pollack (Eds.), *Intentions in communication* (pp. P271–311). Cambridge, MA: MIT Press.

Puga, K, Fuchs, R, Setter, J and Mok, P (2017) *The Perception of English Intonation Patterns by German L2 Speakers of English*. Paper presented at the Interspeech 2017.

Salverda, AP, Dahan, D and McQueen, JM (2003) The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition* **90**, 51–89.

Salverda, AP, Dahan, D, Tanenhaus, MK, Crosswhite, K, Masharov, M and McDonough, J (2007) Effects of prosodically modulated sub-phonetic variation on lexical competition. *Cognition* **105**, 466–476.

Spinelli, E, Grimault, N, Meunier, F and Welby, P (2010) An intonational cue to word segmentation in phonemically identical sequences. *Atten Percept Psychophys* **72**, 775–787.

Steffman, J (2019) Intonational structure mediates speech rate normalization in the perception of segmental categories. *Journal of Phonetics* **74**, 114–129.

Taylor, I and Taylor, M (1995) *Writing and literacy in Chinese, Korean, and Japanese*. Amsterdam: Benjamins.

Tremblay, A, Broersma, M and Coughlin, CE (2018) The functional weight of a prosodic cue in the native language predicts the learning of speech segmentation in a second language. *Bilingualism: Language and Cognition* **21**, 640–652.

Tremblay, A, Broersma, M, Coughlin, CE and Choi, J (2016) Effects of the native language on the learning of fundamental frequency in second-language speech segmentation. *Frontiers in Psychology* **7**, 985.

Tremblay, A, Cho, T, Kim, S and Shin, S (2019) Phonetic and phonological effects of tonal information in the segmentation of Korean speech: An artificial-language segmentation study. *Applied Psycholinguistics* **40**, 1221–1240.

Tremblay, A, Coughlin, CE, Bahler, C and Gaillard, S (2012) Differential contribution of prosodic cues in the native and non-native segmentation of French speech. *Laboratory Phonology* **3**, 385–423.

Tremblay, A, Namjoshi, J, Spinelli, E, Broersma, M, Cho, T, Kim, S, Martínez-García, MT and Connell, K (2017) Experience with a second language affects the use of fundamental frequency in speech segmentation. *PLoS One* **12**, e0181709.

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press

Tremblay, A and Ransijn, J (2015) Model selection and post-hoc analysis for (G)LMER models. Retrieved from https://cran.r-project.org/web/packages/LMERConvenienceFunctions/

Tremblay, A and Spinelli, E (2014) English listeners' use of distributional and acoustic-phonetic cues to liaison in French: Evidence from eye movements. *Language and Speech* **57**, 310–337.

Tyler, MD and Cutler, A (2009) Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America* **126**, 367–376.

van Leussen, J-W and Escudero, P (2015) Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology* **6**.

Weber, A and Cutler, A (2006) First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America* **119**, 597.

Welby, P (2006) French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* **34**, 343–371.

Welby, P (2007) The role of early fundamental frequency rises and elbows in French word segmentation. *Speech Communication* **49**, 28–48.

https://doi.org/10.1017/S136672892000053X Published online by Cambridge University Press