



# Intonational Cues to Segmental Contrasts in the Native Language Facilitate the Processing of Intonational Cues to Lexical Stress in the Second Language

Hyoju Kim\* and Annie Tremblay

Department of Linguistics, The University of Kansas, Lawrence, KS, United States

## OPEN ACCESS

### Edited by:

William Choi,  
The University of Hong Kong,  
Hong Kong SAR, China

### Reviewed by:

Martin Ho Kwan Ip,  
University of Pennsylvania,  
United States  
Seth Goss,  
Emory University, United States

### \*Correspondence:

Hyoju Kim  
kimhj@ku.edu

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Communication

**Received:** 29 December 2021

**Accepted:** 22 March 2022

**Published:** 14 April 2022

### Citation:

Kim H and Tremblay A (2022)  
Intonational Cues to Segmental  
Contrasts in the Native Language  
Facilitate the Processing of  
Intonational Cues to Lexical Stress in  
the Second Language.  
*Front. Commun.* 7:845430.  
doi: 10.3389/fcomm.2022.845430

This study examines whether second language (L2) learners' processing of an intonationally cued lexical contrast is facilitated when intonational cues signal a segmental contrast in the native language (L1). It does so by investigating Seoul Korean and French listeners' processing of intonationally cued lexical-stress contrasts in English. Neither Seoul Korean nor French has lexical stress; instead, the two languages have similar intonational systems where prominence is realized at the level of the Accentual Phrase. A critical difference between the two systems is that French has only one tonal pattern underlying the realization of the Accentual Phrase, whereas Korean has two underlying tonal patterns that depend on the laryngeal feature of the phrase-initial segment. The L and H tonal cues thus serve to distinguish segments at the lexical level in Korean but not in French; Seoul Korean listeners are thus hypothesized to outperform French listeners when processing English lexical stress realized only with (only) tonal cues (H\* on the stressed syllable). Seoul Korean and French listeners completed a sequence-recall task with four-item sequences of English words that differed in intonationally cued lexical stress (experimental condition) or in word-initial segment (control condition). The results showed higher accuracy for Seoul Korean listeners than for French listeners only when processing English lexical stress, suggesting that the processing of an intonationally cued lexical contrast in the L2 is facilitated when intonational cues signal a segmental contrast in the L1. These results are interpreted within the scope of the cue-based transfer approach to L2 prosodic processing.

**Keywords:** speech perception, spoken word recognition, second language acquisition, Korean learners of English, French learners of English, English lexical stress

## INTRODUCTION

In the domain of speech perception and spoken word recognition, a growing number of studies have begun to examine how second/foreign language (L2) learners perceive non-native suprasegmental contrasts (e.g., Dupoux et al., 2008; Zhang and Francis, 2010; Shport, 2015; Qin et al., 2017, 2019; Connell et al., 2018; Chan and Chang, 2019; Kim and Tremblay, 2021; Tremblay et al., 2021). One influential theoretical approach that seeks to explain the influence of the native

language (L1) on the perception of L2 sound contrasts is the cue-weighting theory of speech perception (e.g., Francis et al., 2000; Francis and Nusbaum, 2002; Holt and Lotto, 2006). This theory emphasizes that speech perception is multidimensional, and acoustic cues are weighted differently not only across categories, but also across languages: Listeners from different language backgrounds hear the same acoustic stimuli differently because of the different weighting of acoustic cues in their L1. Accordingly, the cue-weighting theory stipulates that the contribution of individual acoustic cues that distinguish among phonetic categories transfers from the L1 to the L2.

For prosodic contrasts, the cue-weighting approach has focused on the functional weight of suprasegmental cues for signaling lexical information—that is, how listeners weight suprasegmental cues to lexical contrasts in the L1, and how this weighting affects the perception and processing of suprasegmental cues to prosodic contrasts in the L2. If a particular suprasegmental cue is thought to play an important role in processing lexical contrasts in the L1, it should be used to process prosodic categories in the L2; the more important a cue is in the L1, the more it is predicted to be used in the perception and processing of L2 prosodic contrasts (e.g., Qin et al., 2017; Kim and Tremblay, 2021; see also Tremblay et al., 2018). The present study further investigates how the L1 influences L2 learners' perception of prosodic contrasts, focusing on lexical stress. More specifically, this study aims to address whether listeners' use of intonational cues to a *segmental* contrast in the L1 can facilitate the processing of an intonationally cued *lexical stress* contrast in the L2.

A non-trivial body of research has found that L2 learners' perception and processing of lexical stress in English is influenced by the weighting of suprasegmental cues to lexical contrasts in the L1 (e.g., Cooper et al., 2002; Cutler et al., 2007; Zhang and Francis, 2010; Chrabaszc et al., 2014; Lin et al., 2014; Qin et al., 2017; Connell et al., 2018; Kim and Tremblay, 2021; Tremblay et al., 2021). For instance, when processing acoustic cues to lexical stress in English, English and Mandarin listeners were reported to rely more on fundamental frequency (F0) cues than on duration or intensity cues, whereas Russian listeners relied more on duration cues than on F0 cues (Chrabaszc et al., 2014). Russian listeners' weaker reliance on F0 and greater reliance on duration (compared to English and Mandarin listeners) were attributed to the importance of duration cues to stress contrasts in their L1. Dutch L2 learners of English also showed evidence of L1-to-L2 cue-weighting transfer: Dutch L2 learners of English were found to put greater weight on suprasegmental cues to process lexical stress compared to native English listeners, a finding that was attributed to the lower weight of vowel quality cues to lexical stress in Dutch compared to English (e.g., Cooper et al., 2002; Cutler et al., 2007; Tremblay et al., 2021). These results suggest that the weighting of suprasegmental cues to lexical stress transfers from the L1 to the L2.

Some studies have also provided evidence that listeners can transfer the use of suprasegmental cues to lexical contrasts from one type of prosodic contrast in the L1 to another in the L2 (e.g., Braun et al., 2014: perception of lexical tones by German, French, and Japanese listeners; Choi et al., 2019; Choi, 2022: perception of English lexical stress by Cantonese L2 learners of

English; Kim and Tremblay, 2021: perception of English lexical stress by Gyeongsang Korean and Seoul Korean L2 learners of English; Tremblay et al., 2018: perception of intonational cues to French word-final boundaries by English and Dutch L2 learners of French; Wiener and Goss, 2019: perception of Japanese pitch accent by naïve Mandarin listeners and English L2 learners of Japanese). These studies provide preliminary evidence that those suprasegmental cues that serve important lexical functions in the L1 can be used to process different prosodic categories in the L2. To illustrate, Kim and Tremblay (2021) investigated whether Korean-speaking L2 learners of English would transfer the use of suprasegmental cues from the processing of lexical pitch accents in Korean to the processing of lexical stress in English. Gyeongsang Korean is a tonal dialect of Korean that does not have lexical stress but has lexical pitch accents, whereas Seoul Korean has neither. Gyeongsang Korean listeners were hypothesized to be more sensitive to F0 as a cue to lexical contrasts compared to Seoul Korean listeners. The results showed that Gyeongsang Korean L2 learners of English had an advantage over Seoul Korean L2 learners of English when processing intonationally cued lexical stress in English words, with duration and intensity cues not further enhancing perception in either group. Gyeongsang Korean listeners' ability to process English lexical stress was attributed to their use of F0 cues from the processing of lexical pitch accents in their L1 dialect, suggesting that suprasegmental cues that are important for distinguishing words in the L1 (i.e., F0) are used to process words in the L2. These results suggest that L2 learners whose L1 dialect does not have lexical stress can transfer the use of a suprasegmental cue (here, F0) from a different prosodic category (e.g., lexical pitch accent contrasts) to lexical stress in the L2.

An important question that arises from this research is the scope of cue weighting transfer. The cue-weighting theory of speech perception proposes that the underlying mechanism for learning speech categories or contrasts in both the L1 and the L2 is listeners' selective attention to specific acoustic dimensions, assuming that a phonetic category consists of a multidimensional structure where each dimension corresponds to a feature of the phonetic category (e.g., Iverson and Kuhl, 1995; Kuhl and Iverson, 1995; Francis and Nusbaum, 2002; Francis et al., 2008). Accordingly, the cue-based transfer approach stipulates that L2 learners' ability to attend to a particular cue in the L2 and associate it with a function that differs from that in the L1 depends on how much weight the cue has in the L1. Thus, one prediction of the theory is that the weight of a cue in the L1 will determine whether L2 learners would rely on the cue in the L2, regardless of its actual function in the L2.

Tremblay et al. (2018) provided empirical evidence that acoustic cues that serve one function in the L1 can indeed be reallocated to a different function in the L2. The authors tested whether English and Dutch L2 learners of French would differ in their use of F0 cues to word-final boundaries in the segmentation of French speech. Both English and Dutch have lexical stress contrasts, but the functional weight of F0 cues for signaling lexical identity is higher in Dutch than in English due to the lower weight of vowel quality cues to lexical stress in Dutch. Thus, it was hypothesized that Dutch listeners would show greater

reliance on F0 cues than English listeners when locating word-final boundaries in French. In other words, Dutch listeners were predicted to transfer the higher functional weight of F0 cues from the processing of lexical information in the L1 (i.e., lexical stress contrasts) to the detection of word-final boundaries in the segmentation of French speech. The results of an eye-tracking experiment revealed that Dutch listeners showed greater reliance on F0 cues than English listeners when locating word-final boundaries in French. This suggests that acoustic cues that serve one function in the L1 (i.e., signaling lexical information) can be transferred to a different function in the L2 (i.e., signaling word boundaries).

Further probing the question of L1-based cue transfer, one may also ask whether cues can transfer across different *types* of linguistic contrasts (e.g., from intonationally cued segmental contrasts to intonationally cued lexical stress contrasts). If cues that have a similar function (e.g., to signal lexical information) can transfer from one prosodic contrast to another (e.g., Kim and Tremblay, 2021: from Gyeongsang Korean lexical pitch accents to English lexical stress; Qin et al., 2017: from Mandarin lexical tones to English lexical stress), then we should also expect cues to transfer from the perception of intonationally cued *segmental* contrasts to the perception of intonationally cued *lexical stress* contrasts, as these two types of contrast serve a similar function—to signal lexical information. In other words, from a cue-weighting perspective, there is no reason not to expect L1-based cue transfer to occur. Some research has provided evidence for the transfer of F0 cues across different types of contrasts. Francis and Nusbaum (2002), for example, showed that English listeners can learn to use F0 as a cue to the Korean stop contrast after short-term identification training in a laboratory environment. This could be taken as evidence for the transfer of F0 cues from intonationally cued lexical stress contrasts to segmental contrasts, as F0 plays some role in the perception of lexical stress in English. However, since F0 also covaries with VOT in English stops, it remains unclear whether English listeners' ability to process F0 cues in L2 segmental contrasts was caused by their use of F0 cues to segmental contrasts or by their use of F0 cues to intonationally cued lexical stress contrasts (or both). The present study will shed further light on this question by investigating whether F0 cues can transfer from the perception of intonationally cued segmental (i.e., stop) contrasts in the L1 to the perception of intonationally cued lexical stress contrasts in the L2. To do so, two groups of L2 learners—Seoul Korean and French L2 learners of English—will be compared. By addressing this question, the present study will clarify the scope of cue-weighting transfer in L2 prosodic processing.

Korean has a three-way laryngeal stop contrast, which is typologically rare. Prior studies have described Korean as having a short Voice Onset Time (VOT) and high F0 for fortis stops, an intermediate VOT and low F0 for lenis stops, and a long VOT and high F0 for aspirated stops in word-initial position (e.g., Lisker and Abramson, 1964; Cho, 1996). It has also been documented that, in Seoul Korean, the VOT of lenis and aspirated stops has gradually merged over time, with the contrast now depending on the F0 of the following vowel (e.g., Silva, 2006; Kang and Guion, 2008; Kang, 2014). The realization of stops in Seoul

Korean is dependent on the prosodic position in which these stops occur, such as the Accentual Phrase (Silva, 2006). More specifically, in trisyllabic Korean words, a low F0 (L) and upward F0 trajectory are observed if the word-initial segment is a lenis stop, and a high F0 (H) and downward F0 trajectory is observed if the initial segment is a non-lenis stop (i.e., fortis and aspirated stops). In other words, the consonant-induced F0 distinction in Korean extends far beyond the initial portion of the immediately following vowel (Jun, 1996; Silva, 2006). Korean listeners have also been found to use F0 cues in the perception of stop contrasts: Lee et al. (2013) and Schertz et al. (2015) demonstrated that Seoul Korean listeners used F0 as a primary cue and VOT as a secondary cue to perceive the lenis-aspirated stop contrast and both F0 and VOT as primary cues to perceive the fortis-lenis stop contrast. Thus, F0 plays an important role in distinguishing stop contrasts for Seoul Korean listeners<sup>1</sup>.

What remains unclear is whether Seoul Korean listeners' reliance on F0 for processing segmental distinctions in the L1 can contribute to enhancing their processing of intonationally cued lexical stress contrasts in the L2. Korean listeners have been shown to have more difficulty than Mandarin listeners in the processing of English lexical stress (Lin et al., 2014). It is therefore unlikely that Korean listeners' use of F0 cues to stop contrasts in the L1 would completely overcome any difficulty they may have in the processing of lexical stress in the L2. However, since F0 cues have an extremely high functional weight in Mandarin due to the importance of lexical tones, the Korean-Mandarin group comparison is not one that can determine whether intonational cues to segmental contrasts in the L1 can provide at least some help in the perception of intonationally cued lexical stress in the L2.

To answer this question, the present study compares Seoul Korean L2 learners of English and proficiency-matched Metropolitan French L2 learners of English in the processing of intonationally cued lexical stress. An important intonational unit in Korean is the Accentual Phrase (AP): If the AP has four or more syllables and its initial segment has the feature of [–stiff vocal folds] (e.g., lenis stops or sonorants), the AP has an LHLH tonal pattern; if the phrase-initial segment is [+stiff vocal folds] (e.g., aspirated and fortis stops, coronal fricatives, and /h/), the AP has an HHLH tonal pattern (for details, see Jun, 1998, 2000). The cue to the segmental contrast that is hypothesized to transfer from the L1 to the L2 is associated with a tone that is triggered by the type of segment (i.e., L for lenis stops or H for aspirated and fortis stops). It is predicted that this intonational cue to the segmental contrast may help Seoul Korean listeners when processing intonationally cued English lexical stress.

Despite the typological differences between the two languages, French has a very similar prosodic system to that of Korean, with the AP also being an important intonational unit in French. If the AP has four or more syllables, it has a LHILH\* tonal pattern, where Hi indicates the secondary or initial phrasal prominence

<sup>1</sup>Most perception studies used only one place of articulation (usually bilabial stops) because they assumed that a general cue-weighting pattern would be consistent across places of articulation. In production, however, Broersma (2010) showed the cue-weighting to be generalizable across all three places of articulation.

and H\* indicates the primary or final pitch accents (for details, see Jun and Fougeron, 2000, 2002). Crucially, French has only one underlying tonal pattern (i.e., LHiLH\*), with the pattern not varying on the basis of the phrase-initial segment. This difference in the underlying tonal patterns of the two languages allows us to investigate whether Seoul Korean L2 learners of English can transfer the use of F0 cues from the processing of segmental contrasts in the L1 to the processing of intonationally cued lexical stress in the L2.

Previous research conducted by Dupoux and colleagues (Dupoux et al., 2001, 2008, 2010) has shown that French monolinguals and French L2 learners of Spanish performed much more poorly than Spanish monolinguals on tasks that required them to process phonetically variable stress under a memory load (sequence recall task), a difficulty that the researchers termed stress “deafness” and attributed to the absence of lexical stress in French. In principle, the stress processing difficulties found in French listeners should be replicated in Seoul Korean listeners, given that Seoul Korean, like French, does not have lexically contrastive stress.

However, and crucially, the cue-based transfer approach would additionally predict that Seoul Korean listeners would outperform French listeners in the processing of intonationally cued English lexical stress because Seoul Korean listeners would transfer the use of F0 cues from the processing of the laryngeal segmental contrasts in the L1 to the processing of lexical stress in the L2. One may ask whether French listeners could, to some degree, transfer the use of F0 as a secondary cue to stop contrasts from the perception of French stops to the perception of intonationally cued English lexical stress. This is unlikely, as Serniclaes (1987) reported that VOT is the dominant cue to the voicing contrast in French; other cues (e.g., F1 onset frequency, duration of formant transition, initial F0, F0 contour, rise time, and burst energy) come into play only when VOT is ambiguous (in perception as well as in production; see Kirby and Ladd, 2015). In other words, VOT provides the major perceptual cue to stop contrasts in French, all the other cues being secondary (e.g., Serniclaes, 1987, cited in Saerens et al., 1989). Since F0 is not a primary cue to stop contrasts in French, but it is for the lenis-aspirated and the fortis-lenis contrast in Seoul Korean, the cue-based transfer approach would predict French listeners to have more difficulty processing intonationally cued English lexical stress compared to Seoul Korean listeners<sup>2</sup>.

<sup>2</sup>The approach adopted here is that listeners attend to acoustic cues that have a high functional weight (e.g., that serve to distinguish words) in the L1. This approach assumes that whether or not the L1 has lexical stress does not have much bearing on whether listeners can perceive an intonationally cued lexical stress contrast in the L2, as long as the cue that signals stress is important for distinguishing words in the L1. For example, Taiwan Mandarin listeners, who do not have lexical stress in their L1, have no difficulty perceiving English lexical stress when it is cued with F0, a cue that is very important to the perception of Mandarin lexical tones (e.g., Qin et al., 2017). Some researchers have made phonologically driven predictions for the processing of lexical stress—that is, predictions that are contingent on the higher-level patterning of stress in the L1 (e.g., Peperkamp and Dupoux, 2002; Peperkamp et al., 2010). While the predictions of the two different approaches may coincide in some cases, we believe it is not necessary to make reference to the phonological patterning of stress in the L1 to predict what listeners do when processing lexical stress in the L2.

This hypothesis was tested using a sequence-recall task similar to those used in previous research on the perception of lexical stress contrasts (e.g., Dupoux et al., 2001, 2008, 2010; Lin et al., 2014; Qin et al., 2017; Kim and Tremblay, 2021). In the association phase of a sequence-recall task, participants are trained to associate words that differ in stress with different keys of a computer keyboard. Then, in the testing phase, participants hear auditory word sequences and attempt to recall them by using the same keys in the corresponding order. The auditory words in each sequence are produced by different talkers and thus are acoustically variable. Because of the short-term memory load that this task imposes, listeners must be able to process lexical stress in a phonologically abstract way in order to recall the sequences accurately; processing lexical stress in an acoustic way would impose too high of a demand on short-term memory for the listener to be able to recall the sequence accurately. Because listeners may vary in their short-term memory capacity, a control condition in which listeners hear a phonological contrast that exists in the L1 is also used as the baseline. Hence, this type of task provides a robust method for investigating the phonological processing of lexical stress, and it discourages response strategies given the memory load it imposes on listeners.

The experiment used in the present study manipulated auditory stimuli in which the lexical stress contrast was conveyed only by F0 cues (with duration and intensity being neutralized), as Seoul Korean and French listeners are expected to differ only in the use of F0 cues. Since this study focuses on the processing of intonationally cued lexical stress, the stimuli did not involve vowel reduction cues, which play an important role in native English listeners' perception of lexical stress in English words (e.g., Cooper et al., 2002; Cutler et al., 2007; Zhang and Francis, 2010; Chrabaszczyk et al., 2014). Under a cue-weighting transfer view, it is only in the use of F0 cues to lexical stress that Korean and French listeners are expected to show disparities.

Additionally, the current study controlled for Seoul Korean participants' knowledge of other tonal dialects and languages based on a quantitative assessment of their language experience, unlike previous studies that investigated Korean listeners' perception of English lexical stress (e.g., Lin et al., 2014)<sup>3</sup>. In doing so, we can assure that any potential advantage from Seoul Korean listeners in the perception of English lexical stress does not stem from their experience with tonal dialects of Korean (e.g., Gyeongsang Korean) or other tonal languages (e.g., Mandarin, Japanese).

Unlike previous studies, the present experiments used real English words to ensure that participants processed the words in the language in which they were intended (i.e., English). If participants hear non-words, we cannot determine with certainty whether they processed the non-words in English

<sup>3</sup>Altmann (2006)'s study on the perception of English stress tested a variety of L1 groups, including “Seoul” Korean listeners. However, due to the lack of detailed information about the participants' language background, it remains unclear whether the Korean participants she tested had any knowledge of tonal dialects of Korean (e.g., Gyeongsang Korean). Controlling for Korean listeners' Korean dialect is critical because experience with other Korean dialects may change how listeners weight F0 cue when processing the Korean stop contrast (for more detail, see Lee and Jongman, 2019; Kim and Jongman, 2021).

mode, thus processing suprasegmental cues as if they belonged to intonationally cued English lexical stress. Using real English words solves this issue.

## METHOD

### Participants

The experiment targeted 50 Seoul Korean L2 learners of English, 50 French L2 learners of English, and 50 native English listeners as a control group; Korean participants who did not speak Gyeongsang Korean and were not regularly exposed to it were recruited from universities in Seoul; French participants were recruited from universities in Aix-en-Provence, Grenoble, and Paris; and English participants were recruited from a Midwestern American university. The participants did not have speech or hearing impairments or learning disabilities. The participants were tested *via* a web-based survey design software, Qualtrics (Qualtrics LLC, 2020). Each participant completed three tasks: (1) a language background questionnaire; (2) a sequence-recall task with English stimuli; and (3) a lexical decision task to assess their lexical proficiency in English (LexTALE; Lemhöfer and Broersma, 2012). The complete session took approximately 45 min. Korean and French participants received financial compensation for participating in the experiment. English participants received extra credits for one of the introductory courses in Linguistics.

The English proficiency of Seoul Korean and French L2 learners of English was controlled based on the information obtained in the language background questionnaire and their LexTALE scores. After specific exclusion criteria were applied (see Section Data Analysis for detail), the present study included 42 Seoul Korean L2 learners of English (25 female), 35 French L2 learners of English (15 female), and 32 native English listeners (19 female). **Table 1** summarizes the relevant language background information for all three groups and the English proficiency data for the Seoul Korean and French L2 learners of English. Statistical analyses revealed that the Seoul Korean and French listeners did not show a significant difference in any of the variables reported in **Table 1** except for their self-reported percentage of daily English usage [ $t_{(75)} = -6, p < 0.001$ ]. Since the significant difference is in a direction that is not confounded with the predictions, it is not problematic for the interpretation of the results.

### Materials

The lexical items used in this study were identical to those of Kim and Tremblay (2021). The lexical stress contrast stimuli for the experimental condition were a minimal pair of English words that differed in their stress pattern (*TRUS*ty vs. *trus*TEE for the practice phase; *OFF*set vs. *off*SET for the test phase)<sup>4</sup>.

<sup>4</sup>Since most of the previous studies that conducted sequence-recall experiments used nonword pairs as auditory stimuli, we did not test listeners' knowledge of the English word pairs we used; listeners do not need to know the words in the sequence to be able to classify them as stressed on the first or second syllable based on the auditory information they hear. Note also that we used a single minimal pair in each condition to follow the method developed by Dupoux and colleagues for the sequence recall task. Although this may limit the generalizability of the results,

**TABLE 1** | Participants' language background information.

	Korean	French	English
Age (years)	26.2 (5.1)	25.8 (5)	20.4 (3.5)
LexTALE (/100)	70.6 (12)	74.7 (17)	92 (5.6)
AOE (years)	9.9 (2.3)	12.3 (4.4)	—
LOR (months)	1.5 (3.4)	8.9 (18)	—
LOE (years)	13.8 (4.7)	10.4 (4.3)	—
Daily English usage (%)	10 (9)	34.4 (24.1)	—
Self-rated English proficiency score (1–5)	2.6 (0.8)	3.1 (0.8)	—
Self-rated English accent score (1–10)	5.3 (2.4)	6.1 (1.8)	—

Values are means (standard deviations). AOE, Age of first exposure to English; LOR, Length of residence in an English-speaking country; LOE, Length of English education.

These stimuli did not involve vowel reduction cues to lexical stress, the contrast being signaled only by suprasegmental cues. The segmental contrast stimuli for the control condition were minimal pair of English words that differed only in the place of articulation of the word-initial segment (*taller* vs. *caller* for the practice phase; *table* vs. *cable* for the test phase). Since English, Korean, and French all have a contrast between the aspirated alveolar stop and the aspirated velar stop, all listeners should be able to perceive this segmental difference. This control condition thus also serves as a test to determine whether participants attended to the task.

The lexical items were recorded by one female and one male native speaker of American English to increase phonetic variability, as was done in previous studies (e.g., Dupoux et al., 2001, 2008, 2010; Lin et al., 2014; Kim and Tremblay, 2021). The speakers recorded each lexical item five times in the carrier sentence, *Say \_\_\_\_\_ again*, using a microphone (Electro Voice N/D 767a) and a digital recorder (Marantz PMD 671) at a sampling rate of 44.1 kHz. From the five repetitions of each stress pattern from each speaker, three best tokens were selected, yielding a total of 24 tokens: 12 experimental tokens (3 tokens × 2 words × 2 speakers) and 12 control tokens (3 tokens × 2 words × 2 speakers). Since it is only in the use of F0 cues to lexical stress that Seoul Korean and French listeners were expected to differ, duration and intensity cues to lexical stress were neutralized. The intensity of all words was first normalized to 70 dB based on the root-mean-square (RMS) amplitude. Each syllable was then manipulated such that its duration and intensity would be that of the average across the two stress patterns. All manipulation procedures were implemented in Praat (Boersma and Weenink, 2019). We used the Pitch Synchronous Overlap and Add (PSOLA) function for duration manipulation, and the Multiply function for intensity. Acoustic measurements of the manipulated stimuli are summarized in **Table 2**.

### Procedure

The task was built using web-based survey design software, Qualtrics (Qualtrics LLC, 2020). The sequence-recall task

we believe the task would be more difficult (possibly too difficult) if listeners had to categorize different words in different trials.

**TABLE 2 | Mean F0, duration, and intensity** of the English critical stimuli (standard deviation).

F0 (Hz)	Word-initial stress			Word-final stress		
	$\sigma 1$	$\sigma 2$	Ratio	$\sigma 1$	$\sigma 2$	Ratio
Offset (M)	139 (9.4)	89 (5.3)	1.6	115 (2.4)	160 (12.1)	0.7
Offset (F)	226 (8.7)	181 (3.1)	1.3	173 (3.8)	229 (7.5)	0.8
<b>Duration (ms)</b>	<b><math>\sigma 1</math></b>	<b><math>\sigma 2</math></b>	<b>Ratio</b>	<b><math>\sigma 1</math></b>	<b><math>\sigma 2</math></b>	<b>Ratio</b>
Offset (M)	269 (25.2)	476 (11.8)	0.6	262 (2.9)	467 (12.7)	0.6
Offset (F)	314 (1.7)	484 (8.7)	0.6	291 (1.4)	513 (2.9)	0.6
<b>Intensity (dB)</b>	<b><math>\sigma 1</math></b>	<b><math>\sigma 2</math></b>	<b>Ratio</b>	<b><math>\sigma 1</math></b>	<b><math>\sigma 2</math></b>	<b>Ratio</b>
Offset (M)	73 (0.4)	72 (0.2)	1.0	73 (0.2)	72 (0.0)	1.0
Offset (F)	70 (0.2)	69 (0.1)	1.0	70 (0.1)	69 (0.0)	1.0

$\sigma 1$  and  $\sigma 2$  indicate the first and second syllable, respectively; the letters "M" and "F" in parentheses stand for male and female speakers, respectively.

consisted of two tiers. The first tier tested listeners' processing of lexical stress contrasts, and the second tier tested listeners' processing of phonemic contrasts. Each tier consisted of four blocks. The first two blocks of the tier formed the practice phase, and the last two blocks of the tier formed the testing phase.

**Figure 1** illustrates the structure of the experiment.

The practice phase involved an association block and a practice block, each with feedback. In the association block of the practice phase, the participants were trained to associate 1 and 2 on a computer keyboard with the two English words that differed in their stress (*TRUS**ty* vs. *trus**TEE*) or in their initial segment (*taller* vs. *caller*). For each contrast type, there were a total of ten association trials (2 stimuli  $\times$  5 repetitions). On each trial, immediate feedback was provided to the participants as to whether they associated the stimulus with the correct button.

In the practice block of the practice phase, participants were asked to recall sequences of the stimuli they learned to associate with 1 and 2 in the association block. For example, if a participant heard the sequence of *TRUS**ty*—*trus**TEE*—*trustee*—*TRUS**ty*, they would need to enter [1221] as a response. The segmental contrast condition had the same logic. Six different orders of four-item sequences (i.e., [1122], [2211], [1212], [2121], [2112], [1221]) were used for the practice block of each contrast type. There were thus six trials (1 repetition  $\times$  6 orders) for each contrast type, and the participants received immediate feedback on the accuracy of their responses (i.e., correct or incorrect). Within a four-item sequence, each item was separated by an inter-stimulus interval of 50 ms, as in previous studies (e.g., Dupoux et al., 2001, 2008; Qin et al., 2017; Kim and Tremblay, 2021). The last item in the sequence was followed by a pure tone to prevent participants from using echoic memory to recall the sequences. On each trial, participants had 5 s to respond after they heard the sequence. After 5 s, the next trial automatically began with an inter-trial interval of 1,500 ms. **Figure 2** illustrates the composition of a trial for the four-item sequence-recall task.

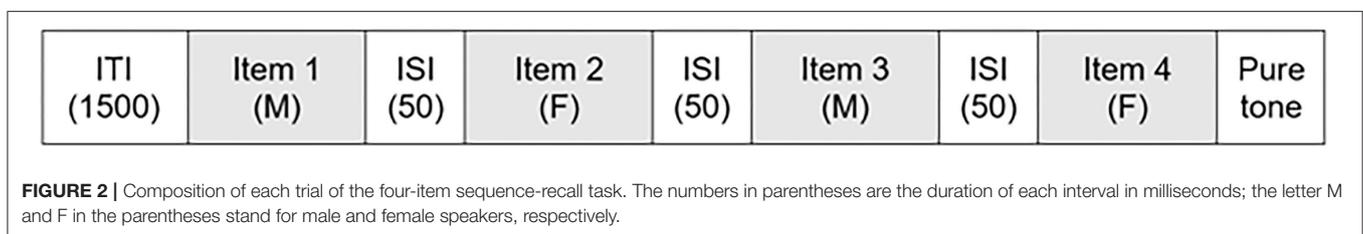
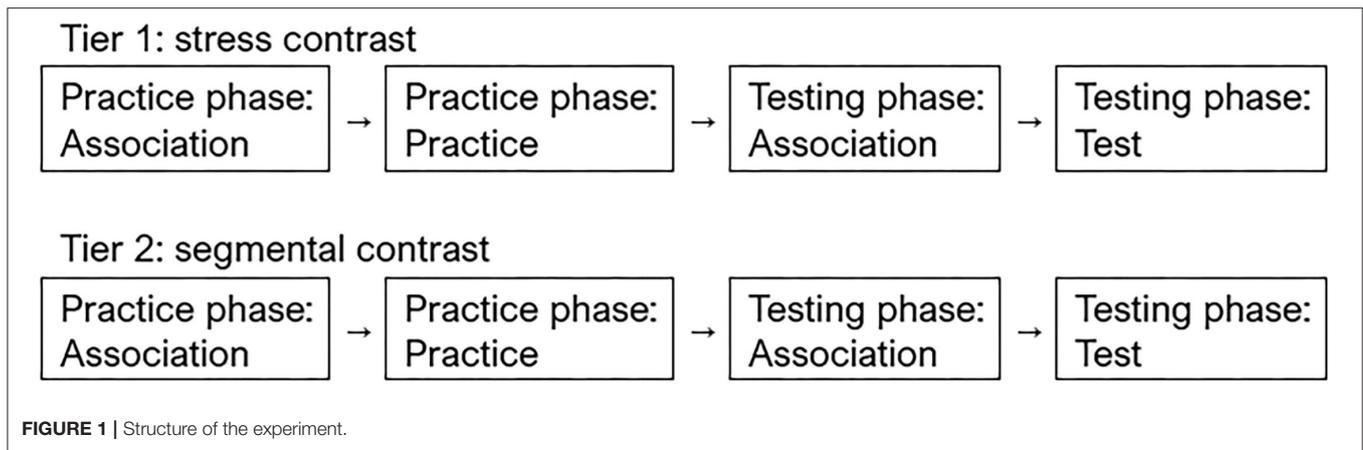
The testing phase contained an association block with feedback and a test block without feedback. In the association block, participants were trained to associate 1 and 2 on a keyboard with two other English words that differed in their lexical stress or in their initial segment, this time with the stimuli

*OFF**set* vs. *off**SET* for the stress contrast and *table* vs. *cable* for the segmental contrast. There was a total of ten association trials (2 stimuli  $\times$  5 repetitions) for each contrast type. In the association block, the participants received immediate feedback on their accuracy in each trial. In the test block, participants were asked to recall the four-item sequences of English words that differed in stress or segment. Ten different token orders (i.e., [1121], [1122], [1211], [1212], [1221], [2112], [2121], [2122], [2211], [2212]) were used in each of the test blocks. Thus, for each contrast type, the test block included 30 trials (3 tokens  $\times$  10 orders). Participants did not receive feedback on the accuracy of their responses in this block. The trials within a block were randomized across participants.

## Data Analysis

The data of participants who did not reach a 75% (22/30) accuracy rate on the segmental block (5 Korean listeners, 15 French listeners, 18 English listeners) were excluded from the analyses, under the assumption that they likely did not focus on the task (which is more likely to happen in web-based experiments). Among the remaining Korean participants, 3 participants who self-reported being able to speak Gyeongsang Korean fluently (despite our attempt not to recruit such participants) were also excluded from the analyses. These filtering processes left 42 Seoul Korean, 35 French, and 32 English listeners in the data analyses.

Mixed-effects logistic regression models were conducted on the participants' sequence-recall accuracy. The data were fitted into the model using the *glmer* function of the *lme4* package (Bates et al., 2015) of the statistical software R and R studio (R Development Core Team, 2019). The model focused on the participants' accuracy in the segmental and lexical stress contrast conditions by participants' L1. The dependent variable of the model was ACCURACY, which is a binary response of correct or incorrect. The participants' response on each trial was coded as correct if they correctly recalled the complete sequence of four items and as incorrect if the sequence was incorrectly recalled. The fixed effects in the model were L1 (English vs. Seoul Korean vs. French), CONTRAST TYPE (segmental contrast vs. stress contrast, baseline: stress contrast), and their interactions.



Since the effect of L1 has three levels, the model was run once with Seoul Korean listeners as a baseline and once with English listeners as a baseline. Random intercepts included participants, test items, and sequence orders. The best model was automatically selected using the backward fitting function of the *LMERConvenienceFunctions* package (Tremblay and Ransijn, 2015).

If Seoul Korean listeners transfer the use of F0 cues from the processing of intonational cues to segmental contrasts in Seoul Korean to the processing of intonationally cued lexical stress contrasts in English, they should be more accurate than French listeners at processing the lexical stress contrasts in English. If this prediction is correct, we should find a significant interaction between L1 (French) and CONTRAST TYPE in the model with Korean listeners' accuracy on the stress contrasts as a baseline.

## RESULTS

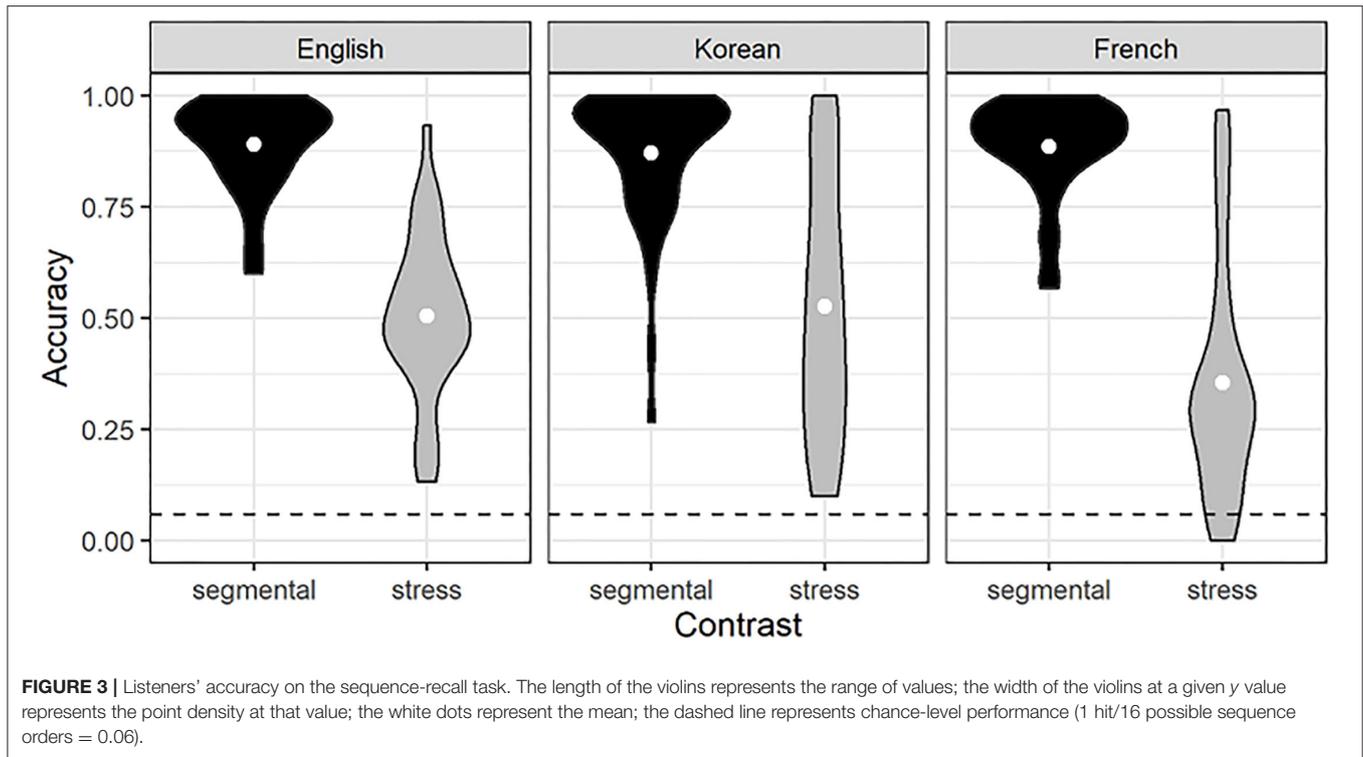
Listeners' accuracy on the sequence-recall task is provided in **Figure 3**, and **Table 3** summarizes the fixed-effect coefficients in the mixed-effects logistic regression model.

The model with Korean listeners' accuracy in the stress contrast condition as a baseline (**Table 3A**) revealed that Seoul Korean listeners outperformed French listeners but not English listeners in the stress contrast condition, as evidenced by the significant simple effect of L1 for French listeners but not for English listeners. Seoul Korean listeners' accuracy in the segmental contrast condition was significantly higher than that in the stress contrast condition, as evidenced by the simple effect of CONTRAST TYPE. Additionally, there

was a significant interaction between L1 and CONTRAST TYPE for the French group but not for the English group, indicating that Seoul Korean listeners' accuracy showed a smaller difference between the segmental contrast and the stress contrast compared to French listeners. The simple effect of L1 and the interaction effect confirm that French listeners showed greater difficulty processing English lexical stress contrasts than Seoul Korean listeners. Because the segmental condition served as control condition and because the results yielded a significant interaction between L1 and CONTRAST TYPE, the effect of L1 on the stress contrast condition cannot be attributed to short-term memory capacity differences between the two L1 groups.

The model with English listeners' accuracy in the stress contrast condition as a baseline (**Table 3B**) showed that English listeners outperformed French listeners in the stress contrast condition, as evidenced by the significant simple effect of L1. The simple effect of CONTRAST TYPE indicates that English listeners' accuracy in the segmental contrast condition was significantly higher than that in the stress contrast condition. There was a significant interaction effect between L1 and CONTRAST TYPE for the French group, meaning that French listeners' accuracy showed a greater difference between the segmental contrast and the stress contrast compared to English listeners. The simple effect of L1 and the interaction effect indicate that French listeners showed greater difficulty processing lexical stress contrasts than English listeners, a result that again cannot be attributed to short-term memory capacity differences between the two L1 groups.

Additional *post-hoc* analyses showed that Seoul Korean and French listeners' accuracy in the stress contrast condition was



**TABLE 3 |** Summary of fixed-effect coefficients in the mixed-effects logistic regression model on listeners' accuracy on the sequence-recall task.

Fixed effects	Est.	SE	z-value	Pr(> z )
<b>(A) Model with Korean listeners' accuracy in the stress contrast condition as baseline</b>				
(Intercept)	0.203	0.224	< 1	0.365
L1 (English)	-0.183	0.279	< 1	0.512
L1 (French)	-0.934	0.276	-3.388	<0.001
Contrast type (segmental)	2.229	0.126	17.69	<0.001
L1 (English) × Contrast type (segmental)	0.113	0.175	< 1	0.519
L1 (French) × Contrast type (segmental)	1.007	0.179	5.631	<0.001
<b>(B) Model with English listeners' accuracy in the stress contrast condition as baseline</b>				
(Intercept)	0.019	0.244	< 1	0.937
L1 (Korean)	0.183	0.279	< 1	0.512
L1 (French)	-0.750	0.292	-2.568	<0.05
Contrast type (segmental)	2.342	0.140	16.68	<0.001
L1 (Korean) × Contrast type (segmental)	-0.113	0.175	< 1	0.519
L1 (French) × Contrast type (segmental)	0.895	0.189	4.712	<0.001

not correlated with demographical factors such as L2 learners' self-rated English proficiency score (Korean:  $r = 0.24, p = 0.13$ ; French:  $r = 0.09, p = 0.58$ ), self-rated English accent score (Korean:  $r = 0.25, p = 0.11$ ; French:  $r = 0.27, p = 0.12$ ), LexTALE score (Korean:  $r = 0.21, p = 0.18$ ; French:  $r = 0.12, p = 0.49$ ), length of residence in an English-speaking country (Korean:  $r = -0.17, p = 0.3$ ; French:  $r = -0.11, p = 0.54$ ), length of English education (Korean:  $r = 0.16, p = 0.32$ ; French:  $r = 0.29, p = 0.09$ ), or age of first exposure to English (Korean:  $r = 0.11, p = 0.5$ ; French:  $r = 0.2, p = 0.24$ ). Thus, L2

learners' performance in the stress contrast condition could not be attributed to their proficiency in or familiarity with English. Additionally, Seoul Korean listeners did not show a significant correlation between their accuracy on the task and their degree of exposure to Gyeongsang Korean ( $r = 0.001, p = 0.99$ ) or between their accuracy and their self-rated Gyeongsang Korean speaking score ( $r = 0.14, p = 0.37$ ), suggesting that Seoul Korean listeners' performance in the stress contrast condition is not related to their knowledge of the tonal dialect of Korean.

## DISCUSSION

This study investigated whether listeners transfer the use of intonational cues from the perception of segmental contrasts in the L1 to the perception of intonationally cued lexical stress in the L2. The results showed that Seoul Korean L2 learners of English had an advantage over proficiency-matched French L2 learners of English when processing intonationally cued lexical stress in English words. These results provide support for the hypothesis that L2 learners whose L1 uses a suprasegmental cue (F0) to distinguish segmental features can transfer the use of that cue from one contrast type (i.e., segmental) in the L1 to another (i.e., suprasegmental contrasts) in the L2.

The results provide important evidence on how the use of F0 cues in the L1 can modulate the processing of lexical stress in the L2. Seoul Korean listeners' accuracy on the stress contrast condition was significantly higher than that of French listeners. This suggests that Seoul Korean L2 learners of English, who do not have lexical stress contrasts in their L1, can transfer the use of F0 cues from the processing of intonational cues to the laryngeal stop distinction in Seoul Korean to the processing of intonationally cued lexical stress in English. In other words, the processing of an intonationally cued lexical contrast in the L2 is facilitated when intonational cues signal a segmental contrast in the L1 compared to when they do not.

Interestingly, English listeners' accuracy in the stress contrast condition was on par with that of Seoul Korean listeners. This may be due to the absence of vowel quality cues to lexical stress in the stimuli. English listeners have been shown to use vowel quality as the most important cue when processing lexical stress contrasts, followed by pitch, duration, and intensity cues (e.g., Zhang and Francis, 2010; Chrabaszcz et al., 2014; Tremblay et al., 2021). The unavailability of vowel reduction cues in the present experiment is likely an important factor in explaining English listeners' difficulty in the processing of lexical stress (see also Experiment 2 of Kim and Tremblay, 2021).

For French listeners, the results of this study are consistent with those of previous studies on the processing of lexical stress by French listeners (Dupoux et al., 2001, 2008, 2010). For instance, in Dupoux et al. (2008), the results of French L2 learners of Spanish, who completed two- and four-item sequence-recall tasks with Spanish-like non-words (e.g., *MIpa* vs. *miPA*), are comparable to those in the present study, with a mean accuracy of 28.3% on the Spanish stress contrast condition. Thus, our results provide additional evidence that French listeners have difficulty processing lexical stress contrasts regardless of the L2 that they process.

The current findings clarify the scope of the cue-based transfer approach to L2 lexical stress processing by showing that the use of intonational cues can transfer *across* contrast types. We attribute Seoul Korean listeners' ability to process English lexical stress to their ability to use F0 cues when processing the laryngeal stop contrasts in their L1. One cannot preclude the possibility that French listeners transfer the use of F0 as a secondary cue to stop contrasts from the perception of French stops to the perception of English lexical stress. However, since F0 has a marginal effect on the perception of stop contrasts in French, the amount of transfer taking place is likely limited, whereas F0

is a primary cue to the lenis-aspirated and the fortis-lenis stop contrast in Seoul Korean, resulting in Seoul Korean listeners' superior performance compared to French listeners.

The present results are interesting to compare to those of Kim and Tremblay (2021). Using a similar sequence-recall task, Kim and Tremblay (2021) found that Gyeongsang Korean listeners outperformed Seoul Korean listeners in the processing of English lexical stress, a finding that was attributed to the transfer of F0 cues from lexical pitch accents in Gyeongsang Korean to lexical stress in English. One important implication from their findings and from ours is that cue transfer from the L1 to the L2 is *relative* and depends on the functional weight of the cue in the L1—specifically, how important the cue is for distinguishing lexical candidates (for discussion, see Tremblay et al., 2018). Taken together, the findings of these two studies suggest that F0 has a greater functional weight in Gyeongsang Korean than in Seoul Korean, and it has a greater functional weight in Seoul Korean than in French.

As mentioned in the introduction, speech perception is multidimensional, and acoustic cues do not equally contribute to signaling a sound contrast. This is also true of lexical stress in English. The present study neutralized duration and intensity cues to lexical stress, as Seoul Korean and French listeners were not necessarily predicted to differ in their use of these two cues. It would be interesting to investigate how Seoul Korean and French listeners weight suprasegmental cues to lexical stress in English when all three cues can potentially signal stress. The results of Kim and Tremblay (2021) suggest that Seoul Korean listeners do not benefit from the addition of duration and intensity cues to auditory stimuli that contrast in intonationally cued English lexical stress (unlike English listeners). Further research should compare Seoul Korean and French listeners on the weighting of all three suprasegmental cues to English lexical stress to determine if French listeners show greater reliance on duration and intensity cues to English stress than Seoul Korean listeners as a compensation strategy for their difficulty in using F0 cues.

From a theoretical perspective, the present findings have important implications. The cue-weighting theory of speech perception proposes that the underlying mechanism for learning speech categories or contrasts in both the L1 and the L2 is listeners' selective attention to specific acoustic dimensions, assuming that a phonetic category consists of a multidimensional structure where each dimension corresponds to a feature of the phonetic category (e.g., Iverson and Kuhl, 1995; Kuhl and Iverson, 1995; Francis and Nusbaum, 2002; Francis et al., 2008). Accordingly, the cue-based transfer approach stipulates that L2 learners' ability to attend to a particular cue in the L2 and associate it with a contrast or function that differs from that in the L1 depends on how much weight the cue has in the L1. The findings of this study indicate that intonational cues that have a similar function (e.g., to signal lexical information) can transfer from one type of contrast in the L1 (e.g., segmental contrast) to another type of contrast in the L2 (e.g., suprasegmental contrast). Thus, the results of the present study extend the scope of the cue-based transfer approach to the processing of L2 lexical stress in showing that L1-based cue transfer is not limited by the type of contrast signaled in the L1 and L2.

One important question that arises from the current findings is whether there are limits or constraints on L1-based cue transfer. A cue-based approach conjectures that phonetic learning involves cross-talker, cross-context, and cross-language generalization. Hence, there is no a priori reason to expect cues not to transfer across prosodic categories, contrast types, or functions. However, since what is important for this approach is the relative weight of cues, and because listeners focus their attention on the cues that have been deemed to have the greatest weight (i.e., primary cues; e.g., Francis and Nusbaum, 2002; Kondaurova and Francis, 2010), it is possible that listeners show transfer effects only for primary cues, and not for cues that have a weaker weight (i.e., secondary cues). It may thus be that the limits or constraints of L1-based cue transfer depend not on the prosodic category, contrast type, or function that the cues serve to signal or perform in the L1 and L2, but on the relative importance of specific cues across languages. In other words, cues may be more likely to transfer or have a noticeable effect on L2 speech perception if they are primary cues insofar as listeners are more likely to attend to these cues, and not as much if they are secondary cues, regardless of types of contrasts or functions.

From this perspective, the results of the present study may be interpreted as French listeners having more difficulty increasing the weight given to F0 in the perception of intonationally cued English lexical stress compared to Seoul Korean listeners because F0 is a comparatively less important lexical cue in French than in Seoul Korean. In other words, Seoul Korean listeners' ability to attend to F0 in the L2 may be explained by their relatively more extensive experience attending to this acoustic cue in the L1.

In a similar vein, it would be interesting to investigate whether Seoul Korean and French listeners differ in the use of vowel quality cues to English stress. French does not have lexical stress, but it has a reduced vowel, the schwa, which is never accented intonationally: If a phrase ends with a schwa, the previous syllable receives the phrase-final pitch accent, and the schwa can be pronounced or deleted depending on the context and/or the French dialect (e.g., Jun and Fougeron, 2002; Welby, 2006; Meunier and Espesser, 2011). Even within a phrase, the schwa can be deleted depending on the phonetic context in which it occurs (e.g., Jun and Fougeron, 2002). By contrast, Korean does not have a reduced vowel, but it has vowels that can assimilate to reduced vowels in English, with these vowels not having any relationship with accenting. The cue-based transfer approach would predict that Seoul Korean and French listeners would not necessarily differ in the processing of stress when the unstressed syllable is reduced, unlike what was predicted for the present study: French listeners would be able to transfer their use of vowel quality cues in the L1 to the processing of English reduced vowels in unstressed syllables, and Korean listeners would also be expected to process vowel quality cues to English stress but for a different reason—although Korean does not have vowel reduction, Korean listeners would be able to process vowel quality cues to English stress by assimilating full and reduced vowels to different Korean vowels (for such a proposal, see Connell et al., 2018).

## CONCLUSION

The present study investigated whether the use of intonational cues can transfer from the processing of segmental contrasts in the L1 to the processing of intonationally cued lexical stress in the L2. A comparison of Seoul Korean and proficiency-matched French L2 learners of English showed that Seoul Korean listeners transferred their use of F0 cues from the processing of the laryngeal stop contrasts in Korean to the processing of lexical stress in English, as evidenced by their greater ability to process English stress compared to French listeners. From a theoretical perspective, this study further specified the scope of the cue-based transfer hypothesis, suggesting that listeners can transfer the use of intonational cues from the processing of segmental contrasts to the processing of lexical stress. Further research and more empirical data are needed to better understand the nature of the limits or constraints on cue-weighting transfer.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors upon request.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Human Research Protection Program, The University of Kansas (IRB ID: STUDY00145019). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

HK created and recorded the stimuli, and analyzed the data. All authors contributed to conception, design of the study, wrote the first draft of the manuscript, contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work was supported by a Linguistics Research Scholarship awarded to HK from the Department of Linguistics at the University of Kansas. The article processing charges related to the publication of this article were supported in part by The University of Kansas (KU) One University Open Access Author Fund sponsored jointly by the KU Provost, KU Vice Chancellor for Research & Graduate Studies, and KUMC Vice Chancellor for Research and managed jointly by the Libraries at the Medical Center and KU - Lawrence.

## ACKNOWLEDGMENTS

We thank the reviewers for their invaluable comments and suggestions. We also thank Dr. Allard Jongman, Dr. Joan Sereno, Dr. Jie Zhang, and the student members of LING 850 at the

University of Kansas for their insightful feedback on this work. We would also like to thank Dr. Elsa Spinelli and Dr. Pauline

Welby for their help with advertising the study in France. Any misunderstandings are, of course, our own.

## REFERENCES

- Altmann, H. (2006). *The perception and production of second language stress: a cross-linguistic experimental study* (dissertation). The University of Delaware, Newark, DE, United States.
- Bates, D., Maechler, B., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Boersma, P., and Weenink, D. (2019). *Doing Phonetics by Computer (Version 6.0.46)*. Retrieved from: <http://www.praat.org> (accessed January 28, 2019).
- Braun, B., Galts, T., and Kabak, B. (2014). Lexical encoding of L2 tones: The role of L1 stress, pitch accent and intonation. *Second Lang. Res.* 30, 323–350. doi: 10.1177/0267658315110926
- Broersma, M. (2010). “Korean lenis, fortis, and aspirated stops: effect of place of articulation on acoustic realization,” in *Proceedings of the 11th Annual Conference of the International Speech Communication Association* (Makuhari), 941–944. doi: 10.21437/Interspeech.2010-317
- Chan, I. L., and Chang, C. B. (2019). Perception of nonnative tonal contrasts by Mandarin-English and English-Mandarin sequential bilinguals. *J. Acoust. Soc. Am.* 146, 956–972. doi: 10.1121/1.5120522
- Cho, T. (1996). *Vowel correlates to consonant phonation: an acoustic-perceptual study of Korean obstruents* (master's thesis). The University of Texas at Arlington, Arlington, TX, United States.
- Choi, W. (2022). Theorizing positive transfer in cross-linguistic speech perception: the Acoustic-Attentional-Contextual hypothesis. *J. Phonet.* 91, 101135. doi: 10.1016/j.wocn.2022.101135
- Choi, W., Tong, X., and Samuel, A. G. (2019). Better than native: tone language experience enhances English lexical stress discrimination in Cantonese-English bilingual listeners. *Cognition* 189, 188–192. doi: 10.1016/j.cognition.2019.04.004
- Chrabaszcz, A., Winn, M., Lin, C. Y., and Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *J. Speech Lang. Hear. Res.* 57, 1468–1479. doi: 10.1044/2014\_JSLHR-L-13-0279
- Connell, K., Hüls, S., Martínez-García, M. T., Qin, Z., Shin, S., Yan, H., et al. (2018). English learners' use of segmental and suprasegmental cues to stress in lexical access: an eye-tracking study. *Lang. Learn.* 68, 635–668. doi: 10.1111/lang.12288
- Cooper, N., Cutler, A., and Wales, R. (2002). Constraints of lexical stress on lexical access in English: evidence from native and nonnative listeners. *Lang. Speech* 45, 207–228. doi: 10.1177/00238309020450030101
- Cutler, A., Wales, R., Cooper, N., and Janssen, J. (2007). “Dutch listeners' use of suprasegmental cues to English stress,” in *Proceedings of the 16th International Congress for Phonetic Sciences*, eds J. Trouvain and W. J. Barry (Dudweiler: Pirrot) 1913–1916.
- Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (2001). A robust method to study stress ‘deafness’. *J. Acoust. Soc. Am.* 110, 1606–1618. doi: 10.1121/1.1380437
- Dupoux, E., Peperkamp, S., and Sebastián-Gallés, N. (2010). Limits on bilingualism revisited: stress ‘deafness’ in simultaneous French-Spanish bilinguals. *Cognition* 114, 266–275. doi: 10.1016/j.cognition.2009.10.001
- Dupoux, E., Sebastián-Gallés, N., Navarrete, E., and Peperkamp, S. (2008). Persistent stress ‘deafness’: the case of French learners of Spanish. *Cognition* 106, 682–706. doi: 10.1016/j.cognition.2007.04.001
- Francis, A. L., Baldwin, K., and Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Percept. Psychophys.* 62, 1668–1680. doi: 10.3758/BF03212164
- Francis, A. L., Kaganovich, N., and Driscoll-Huber, C. (2008). Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *J. Acoust. Soc. Am.* 124, 1234–1251. doi: 10.1121/1.2945161
- Francis, A. L., and Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 349–366. doi: 10.1037/0096-1523.28.2.349
- Holt, L. L., and Lotto, A. J. (2006). Cue weighting in auditory categorization: implication for first and second language acquisition. *J. Acoust. Soc. Am.* 119, 3059–3071. doi: 10.1121/1.2188377
- Iverson, P., and Kuhl, P. K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *J. Acoust. Soc. Am.* 97, 553–562. doi: 10.1121/1.412280
- Jun, S.-A. (1996). The influence of the microprosody on the macroprosody: a case of phrase initial strengthening. *UCLA Work. Pap. Phonet.* 92, 97–116.
- Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology* 15, 189–226. doi: 10.1017/S0952675798003571
- Jun, S.-A. (2000). K-ToBI (Korean ToBI) labeling conventions. *UCLA Work. Pap. Phonet.* 99, 149–173.
- Jun, S.-A., and Fougeron, C. (2000). “A phonological model of French intonation,” in *Intonation: Analysis, Modeling and Technology*, ed A. Botinis (Dordrecht: Kluwer Academic Publishers), 209–242. doi: 10.1007/978-94-011-4317-2\_10
- Jun, S.-A., and Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus* 14, 147–172. doi: 10.1515/prbs.2002.002
- Kang, K. H., and Guion, S. G. (2008). Clear speech production of Korean stops: changing phonetic targets and enhancement strategies. *J. Acoust. Soc. Am.* 124, 3909–3917. doi: 10.1121/1.2988292
- Kang, Y. (2014). Voice Onset Time merger and development of tonal contrast in Seoul Korean stops: a corpus study. *J. Phonet.* 45, 76–90. doi: 10.1016/j.wocn.2014.03.005
- Kim, H., and Jongman, A. (2021). The influence of inter-dialect contact on the Korean three-way laryngeal distinction: an acoustic comparison among Seoul Korean speakers and Gyeongsang speakers with limited and extended residence in Seoul. *Lang. Speech*. doi: 10.1177/00238309211037720
- Kim, H., and Tremblay, A. (2021). Korean listeners' processing of suprasegmental lexical contrasts in Korean and English: a cue-based transfer approach. *J. Phonet.* 87, 101059. doi: 10.1016/j.wocn.2021.101059
- Kirby, J. P., and Ladd, D. R. (2015). “Stop voicing and F0 perturbations: evidence from French and Italian,” in *Proceedings of the 18th International Congress of Phonetic Sciences* (Glasgow: The University of Glasgow).
- Kondaurova, M., and Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: comparison of three training methods. *J. Phonet.* 38, 569–587. doi: 10.1016/j.wocn.2010.08.003
- Kuhl, P., and Iverson, P. (1995). “Linguistic experience and the ‘perceptual magnet effect’,” in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed W. Strange (Baltimore, MD: York Press), 121–154.
- Lee, H., and Jongman, A. (2019). Effects of sound change on the weighting of acoustic cues to the three-way laryngeal stop contrast in Korean: diachronic and dialectal comparisons. *Lang. Speech* 62, 509–530. doi: 10.1177/0023830918786305
- Lee, H., Politzer-Ahles, S., and Jongman, A. (2013). Speakers of tonal and non-tonal Korean dialects use different cue weightings in the perception of the three-way laryngeal stop contrast. *J. Phonet.* 41, 117–132. doi: 10.1016/j.wocn.2012.12.002
- Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: a quick and valid Lexical Test for Advanced Learners of English. *Behav. Res. Methods* 44, 325–343. doi: 10.3758/s13428-011-0146-0
- Lin, C. Y., Wang, M., Idsardi, W. J., and Xu, Y. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingual. Lang. Cogn.* 17, 316–346. doi: 10.1017/S1366728913000333
- Lisker, L., and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 384–422. doi: 10.1080/00437956.1964.11659830
- Meunier, C., and Espesser, R. (2011). Vowel reduction in conversational speech in French: the role of lexical factors. *J. Phonet.* 39, 271–278. doi: 10.1016/j.wocn.2010.11.008
- Peperkamp, S., and Dupoux, E. (2002). A typological study of stress ‘deafness’. *Labo. Phonol.* 7, 203–240. doi: 10.1515/9783110197105.203
- Peperkamp, S., Vendelin, I., and Dupoux, E. (2010). Perception of predictable stress: a cross-linguistic investigation. *J. Phonet.* 38, 422–430. doi: 10.1016/j.wocn.2010.04.001

- Qin, Z., Chien, Y. F., and Tremblay, A. (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Appl. Psycholinguist.* 38, 541–570. doi: 10.1017/S0142716416000321
- Qin, Z., Tremblay, A., and Zhang, J. (2019). Influence of within-category tonal information in the recognition of Mandarin-Chinese words by native and non-native listeners: an eye-tracking study. *J. Phonet.* 73, 144–157. doi: 10.1016/j.wocn.2019.01.002
- Qualtrics, LLC (2020). *Qualtrics [Computer Program]*. Qualtrics. Retrieved from: <https://www.qualtrics.com> (accessed January 1, 2020).
- R Development Core Team (2019). *R: A Language and Environment for Statistical Computing (Version Version 1.2.1335)*. Vienna: R Foundation for Statistical Computing. Retrieved from: <http://www.rproject.org>
- Saerens, M., Serniclaes, W., and Beeckmans, R. (1989). Acoustic versus contextual factors in stop voicing perception in spontaneous French. *Lang. Speech* 32, 291–314. doi: 10.1177/002383098903200401
- Schertz, J., Cho, T., Lotto, A., and Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *J. Phonet.* 52, 183–204. doi: 10.1016/j.wocn.2015.07.003
- Serniclaes, W. (1987). *Etude expérimentale de la perception du trait de voisement des occlusives du français* (dissertation). Université Libre de Bruxelles, Brussels, Belgium.
- Shport, I. A. (2015). Perception of acoustic cues to Tokyo Japanese pitch-accent contrasts in native Japanese and naive English listeners. *J. Acoust. Soc. Am.* 138, 307–318. doi: 10.1121/1.4922468
- Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23, 287–308. doi: 10.1017/S0952675706000911
- Tremblay, A., Broersma, M., and Coughlin, C. E. (2018). The functional weight of a prosodic cue in the native language predicts the learning of speech segmentation in a second language. *Bilingual. Lang. Cogn.* 21, 640–652. doi: 10.1017/S136672891700030X
- Tremblay, A., Broersma, M., Zeng, Y., Kim, H., Lee, J., and Shin, S. (2021). Dutch listeners' perception of English lexical stress: a cue-weighting approach. *J. Acoust. Soc. Am.* 149, 3703–3714. doi: 10.1121/10.0005086
- Tremblay, A., and Ransijn, J. (2015). *Package 'LMERConvenienceFunctions'*.
- Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *J. Phonetics.* 34, 343–371.
- Wiener, S., and Goss, S. (2019). Second and third language learners' sensitivity to Japanese pitch accent is additive: an information-based model of pitch perception. *Stud. Second Lang. Acquisit.* 41, 897–910. doi: 10.1017/S0272263119000068
- Zhang, Y., and Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *J. Phonet.* 38, 260–271. 6

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Kim and Tremblay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.