



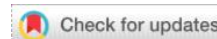
pISSN 2005-8063
eISSN 2586-5854
2024. 9. 30.
Vol.16 No.3
pp. 33-38

말소리와 음성과학

Phonetics and Speech Sciences

한국음성학회지

<https://doi.org/10.13064/KSSS.2024.16.3.033>



Exploring stress encoding cues in English by Korean L2 speakers*

Goun Lee**

Department of General Education, Dankook University, Yongin, Korea

Abstract

The present study investigated the perceptual cues utilized by Korean L2 learners of English in recognizing lexical stress in English nonwords, with a focus on the roles of fundamental frequency (F0) and duration. Twenty-three Korean learners of English participated in a sequence recall task involving nonword stimuli under five different conditions: (1) the naturally-produced stimuli, (2) the duration-only condition, (3) the F0-only condition, (4) the duration-F0 matching condition, and (5) the duration-F0 conflicting condition. The results demonstrate that F0 is the primary cue for stress perception among Korean L2 learners, whereas duration acts as a secondary cue, particularly when F0 is unreliable or absent. These findings highlight the influence of L1 prosodic structures on L2 perception and suggest that Korean L2 learners adapt their perceptual weighting of stress based on cue availability. This study contributes to the understanding of the role of cue weighting in L2 prosodic acquisition.

Keywords: lexical stress acquisition, cue weighting, cue transfer, sequence recalling task, nonword stimuli, L2 stress perception

1. Introduction

Stress is one of the prosodic features produced at the segmental and suprasegmental level to indicate the prominence of a syllable within a word. Cross-linguistically, languages generally realize stress through three mechanisms: free stress, fixed stress, and no stress. In free stress languages, such as English, stress can fall on different syllables across words without a fixed pattern. English stress is characterized by increased duration, higher pitch, and greater intensity (Fry, 1955, 1958; Gay, 1978), accompanied by stronger gestural movements compared to unstressed syllables (Xu & Xu, 2005). In contrast, fixed stress languages exhibit predictable stress placement on a specific syllable (e.g., Finnish, Spanish,

Polish), consistently using one or more of these cues. However, no stress languages (e.g., Seoul Korean, French) do not utilize stress to distinguish between words.

Based on cross-linguistic differences in how languages express stress—free stress language, fixed stress language, or no stress—a plethora of research in the field of second language (L2) acquisition has explored how L2 learners acquire and process a new prosodic feature when their native language does not utilize such a feature. For example, Lin et al. (2014) investigated how Mandarin (claimed to have lexical stress in L1) and Korean learners (no stress language) of English process English stress patterns, focusing on the influence of their native languages' prosodic features. The study examined this by using sequence recalling task and lexical decision

* The present research was supported by the research fund of Dankook University in 2024–2025 (No. R-2024-00653).

** gounlee@dankook.ac.kr, Corresponding author

Received 18 August 2024; Revised 19 September 2024; Accepted 19 September 2024

© Copyright 2024 Korean Society of Speech Sciences. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

tasks, comparing the performance of Mandarin, Korean, and English speakers. Results showed that Korean learners, whose language lacks stress patterns, struggled more with English stress contrasts than Mandarin learners, whose language includes tonal and stress features. The findings support the Stress Parameter Model, indicating that native prosody impacts second language stress perception and processing.

Based on this finding, Qin et al. (2017) further explored whether cross-linguistic differences within the same native language can influence the acquisition of stress in L2. Chinese language use F0 to indicate different tones, Mandarin Chinese is claimed to have lexical stress while Taiwanese Chinese lacks lexical tone in its dialect. Based on this claim, Qin et al. (2017) specifically examined whether Mandarin-speaking learners of English, whose first language (L1) features lexical stress, process English word stress differently from Taiwan Mandarin speakers, whose L1 lacks lexical stress, particularly in their ability to utilize duration and fundamental frequency (F0) cues in a sequence recall test with English non-words. Results showed that while both groups can use F0 to encode stress, Standard Mandarin speakers are more adept at using duration cues than Taiwan Mandarin speakers, suggesting that L1 dialectal differences influence L2 stress processing.

Regarding the Korean L2 learners of English, Kim & Tremblay (2022) compared Seoul Korean learners of English to French L2 learners of English to investigate the perceptual sensitivity to lexical stress is influenced by the use of suprasegmental cues in their native languages. According to the Cue Weighting Hypothesis (Chang, 2018; Chrabaszcz et al., 2014; Qin et al., 2017; Schertz et al., 2015; Tremblay et al., 2018; Zhang & Francis, 2010), listeners prioritize certain cues based on the functional importance of suprasegmental features in their L1. If a particular cue is heavily used in the L1, its functional load will likely affect the perception and processing of similar prosodic cues, such as lexical stress, in the L2.

Given these differences in L1 cue weighting, Kim & Tremblay (2022) hypothesized that Korean and French learners would differ in their sensitivity to F0 when perceiving English lexical stress. While Korean and French have similar Accentual Phrase (AP) system expressed as LHLH pitch pattern, the AP in Korean is a key intonational unit, with tonal patterns triggered by the type of segment (Jun, 1998; Jun & Fougeron, 2000). Specifically, when the first consonant of the AP is lenis, the first tone is denoted as low tone, while the aspirated and fortis stops are expressed with high tone (Jun & Fougeron, 2000, 2002). In contrast, the AP in French does not interact with F0, resulting in F0 to be only as the secondary cue in contrasting stop (Kirby & Ladd, 2015; Serniclaes, 1987). Based on this, Kim & Tremblay (2022) tested whether Seoul Korean-speaking learners of English would be more sensitive in using F0 than French speakers in perceiving English lexical stress by using stress sequence test. Their study found that Seoul Korean L2 learners of English outperformed French L2 learners of English in processing intonationally cued lexical stress in English words. This supports that L2 learners whose L1 uses a suprasegmental cue, such as fundamental frequency (F0), to distinguish segmental features can transfer that cue from segmental contrasts in the L1 to suprasegmental contrasts in the L2.

However, an important question remains unanswered. While the Korean participants in Lin et al. (2014) demonstrated an accuracy of approximately 25% in recalling four-word stimuli, the participants in Kim & Tremblay (2021, 2022) showed significantly higher

accuracy (89%). This disparity may be attributed to the use of real words in Kim & Tremblay (2021, 2022), as opposed to nonwords in Lin et al. (2014). Real words are easier to encode and retain in phonological short-term memory due to the presence of existing lexical representations in long-term memory (Hulme et al., 1991). Nonwords, lacking long-term representations and being encoded only at the form level, are more challenging to retain in short-term memory. Furthermore, the use of varying recall sequences (2-sequence, 3-sequence, 4-sequence) in Lin et al. (2014) may have increased the task's difficulty, as participants could not predict the number of words to recall in the stimuli, unlike in Kim & Tremblay (2021, 2022). Additionally, the use of a single type of stimuli (i.e., a balanced distribution of stress patterns within the four-word sequence, such as equal numbers of first- and second-syllable stresses) might have facilitated the use of F0 cues in the participants' stress perception, as it increases the predictability of the stimuli type.

Also, Kim & Tremblay (2022) manipulated their stimuli to neutralize the intensity and duration cues, allowing participants to rely solely on F0 cues to perceive single word stress patterns in four-word sequences. They justified this approach by referencing their previous study (Kim & Tremblay, 2021), which found no difference in perceptual sensitivity between Gyeongsang Korean and Seoul Korean speakers when comparing naturally produced stimuli (where duration, intensity, and F0 together signal stress) to manipulated stimuli (where only F0 signals stress). However, other studies have reported different results. For instance, in a study where the English word 'object' was orthogonally manipulated in duration, intensity, F0, and vowel reduction, Lee (2022) found that Korean learners of English were most sensitive to the vowel reduction cue when perceiving English words. Regarding suprasegmental cues, Korean listeners exhibited similar sensitivity to F0 and intensity.

Additional research has shown that Korean learners utilize multiple cues in perceiving English stress. Kang & Kim (2019) investigated how segmental (vowel reduction) and suprasegmental cues (F0, intensity, duration) affect Korean listeners' perception of English stress in nonword stimuli. Their study manipulated the acoustic stimuli in five incremental steps for suprasegmental cues and three steps for segmental cues. The results indicated that while all these cues play a crucial role in identifying English stress, higher proficiency learners relied more heavily on vowel reduction, whereas lower proficiency learners relied more on suprasegmental cues, particularly intensity. The results of these two studies contrast with the findings of Kim & Tremblay (2021), who concluded that Koreans show no sensitivity to either duration or intensity when perceiving lexical stress.

To address the unsolved questions raised by previous studies, including the discrepancy in recall accuracy between Korean participants in Kim & Tremblay (2021, 2022) and Lin et al. (2014), the current study investigated which suprasegmental cues, specifically F0 and duration, Korean L2 learners of English with Seoul dialect utilize in their perception of lexical stress. This study aimed to determine whether Korean L2 learners' ability to perceive stress patterns in nonword stimuli differs from their performance with real words from previous studies and to assess which suprasegmental cue between F0 and duration they weight more in perceiving English stress. The experimental design was adapted from Qin et al. (2017), utilizing nonword stimuli in which F0 and

duration cues were resynthesized to signal stress patterns. The research questions for this study are as follows:

- (1) Will Korean L2 learners of English perceive stress patterns in nonword stimuli?
- (2) Will Korean L2 learners of English be able to perceive lexical stress by relying on only one cue?
- (3) Will Korean L2 learners of English be facilitated by the use of duration cues in addition to F0 cues in their perception of lexical stress?
- (4) How will Korean L2 learners of English perceive stress when two cues (duration and F0) are in conflict?

2. Method

2.1. Subjects

A total of 23 Korean L2 learners of English (5 males, 18 females) from Seoul district participated in this study. All participants were born and raised in Gyeonggi and Seoul areas in Korea with Seoul accent. The average age of the subjects was 21.43 years ($SD=1.97$), and the average age at which they began learning English was 7.3 years ($SD=2.03$). Participants reported their English proficiency as ranging from intermediate to advanced levels. Before the experiment, they completed a proficiency test (Michigan Proficiency Test; Briggs et al., 2003), where they listened to English sentences and selected the most appropriate reply from provided statements. The average score on the Michigan proficiency cloze test was 40.88 ($SD=3.02$) out of 45, indicating that all participants' proficiency levels were in the upper-intermediate to advanced range. None of the participants reported any hearing or speaking disorders.

2.2. Stimuli

2.2.1. Naturally produced stimuli

In this study, the same type of English nonwords used in Qin et al. (2017) were adopted for the sequence recalling test. The stimuli were possible English stress minimal pairs constructed with a consonant-vowel (CV) $C^1V^1C^2V^2$ structure. Three types of the vowels - /ɪ/, /ʊ/, and /ʌ/ - were used in the first syllable and [i] was used in the second syllable to prohibit vowel reduction. To ensure that consonants did not provide segmental cues to stress, only fricatives and voiced stops were used in the C1 and C2 positions (e.g., Cho & Keating, 2001; Tremblay, 2009). Thus, a total of twelve experimental nonwords (/bɪsi/, /brɪvi/, /dʊθi/, /dʊzi/, /gʌfi/, /gʌði/, /sɪvi/, /zʊθi/, /fʌði/, /vɪsi/, /θʊzi/, /hʌfi/) were utilized.

A female speaker of Midwestern accent produced these nonword stimuli in a carrier sentence, "Say ___ again" with four repetitions, and two tokens that represent the best for the stress pattern were chosen for the perception experiment. The stimuli set for the naturally produced stimuli were 48 tokens (12 nonwords × 2 stress patterns × 2 tokens).

This study adopted the same stimuli and experimental design as Qin et al. (2016). Their study found significant differences in the first-to-second syllable ratios of F0 and duration for stimuli, with pairwise *t*-tests ($p < .05$). Additionally, they noted that, regardless of the stress position, the final syllable (second syllable) tended to be longer, which they attributed to the phenomenon of word-final lengthening.

2.2.2. Manipulated stimuli

In order to investigate which suprasegmental cues Korean L2 learners of English attune to perceive the lexical stress, three suprasegmental cues-intensity, F0, and duration-were manipulated. For the manipulation, 4 stimuli (/sɪvi/, /zʊθi/, /fʌði/, and /hʌfi/) with initial fricatives were chosen as voiced stops and fricatives might have potential differences in expressing contrastive stress. The stimuli set consisted of 16 tokens in total (4 segmental nonwords × 2 stress patterns × 2 repetitions). Duration and F0 cues in the stimuli of the testing phase 2 were manipulated in different conditions: Duration-only condition (duration alone signal stress), F0-only condition (F0 alone signal stress), Duration-F0 matching condition (F0 and duration cues congruently signal stress), and Duration-F0 conflicting condition (F0 and duration are incongruent in signaling stress). For the stimuli used in the conflicting condition, for example, when one cue (e.g., F0) signaled stressed syllable with higher F0 values, the other cue (e.g., duration) signaled unstressed syllable with shorter duration, making conflicting condition to stress. Examining the participants' correct responses in conflicting condition will enable us to see which cues Korean L2 learners of English would utilize as the primary cue in perceiving the lexical stress of English nonwords.

For the manipulation, the intensity values of the experimental stimuli were first normalized to 70 dB. Subsequently, the duration and F0 values were adjusted to match the average values of the naturally produced tokens using the PSOLA function in Praat (Boersma & Weenink, 2012). The average duration of the stressed and unstressed syllables in the naturally produced tokens was 249 ms, which was used as the baseline for manipulation. This baseline token was then manipulated for F0 and duration to indicate the desired stress patterns. The duration was manipulated to reflect unstressed (292 ms for σ1; 176 ms for σ2,) and stressed (212 ms for σ1; 317 ms for σ2) syllables. For the F0 values, the average F0 of the baseline token for each syllable was 189 Hz, and F0 values were similarly adjusted to indicate unstressed (161 Hz for σ1; 175 Hz for σ2) and stressed (238 Hz for σ1; 193 Hz for σ2) syllables.

In the duration-only condition, where only duration cues indicated stress, the F0 values for both syllables were kept constant at 189 Hz. In the F0-only condition, where only F0 cues indicated stress, the syllable durations for both positions were maintained at 249 ms. In the duration-F0 matching condition, both F0 and duration cues congruently signaled stressed (σ1: 212 ms & 238 Hz; σ2: 317 ms & 193 Hz) or unstressed syllables (σ1: 292 ms & 161 Hz; σ2: 176 ms & 175 Hz) based on mean value of the naturally produced stimuli. In the duration-F0 conflicting condition, the stressed and unstressed syllables were mismatched for the first and second syllables. For example, when the first syllable's duration was set at 212 ms to indicate stress, its F0 was set at 161 Hz to indicate an unstressed syllable. Similarly, when the second syllable's duration was 317 ms to indicate stress, its F0 was manipulated to 193 Hz, indicating an unstressed syllable. In this way, four possible tokens for the experiment nonword of duration-F0 conflicting condition were generated.

2.3. Procedure

The sequence-recalling task comprised three stages: a familiarization phase, testing phrase 1 (naturally produced stimuli), and testing phase 2 (manipulation stimuli). During the familiarization phase, participants were trained to associate the numbers 1 and 2 on

a keyboard with first-syllable or second-syllable stressed words. The familiarization phrase was conducted with a stress minimal pair of English real word (e.g., “trusty” vs. “trustee”). On each trial, participants were given the feedback on whether their responses were correct or not. Participants completed a practice session of 12 sequences to ensure comprehension of the task before beginning the actual experiment. Following 18 trials of practice, participants then moved onto the familiarization test to correctly identify the stress pattern of the auditorily presented stimuli, which was required to reach an accuracy of 95% or higher to proceed to the next tests (testing phase 1 & 2). Those who did not meet this criterion had to repeat the familiarization task, up to two more times. The familiarization phase took between 10 to 20 minutes, depending on how quickly the accuracy criterion was met.

In the testing phase, participants were asked to recall sequences of four tokens by pressing the keys 1 (first-syllable stressed) and 2 (second syllable stressed) in the correct order. For the testing phase stimuli, each sequence included two tokens with word-initial stress and two with word-final stress (e.g., [fʌ'di] [fʌ'di] [fʌ'di] [fʌ'di]) among the nonword stimuli mentioned in 2.2. Following previous studies (Kim & Tremblay 2021, 2022; Qin et al., 2017), only six different sequence types ([1122], [2211], [1212], [1221], [2121], [2112]) were employed to balance the number of nonwords with initial and final stress. The order of sequences and tokens within each sequence was randomized for each participant. Thus, the experiment of testing phases comprised 72 experimental trials (12 nonwords×6 orders).

On each trial, participants saw a visual prompt of “next trial.”, followed by four auditory presented nonword sequence with a 50 ms interstimulus interval, following previous studies (e.g., Dupoux et al., 2001; Kim & Tremblay 2021, 2022; Qin et al., 2016). The final interstimulus interval was followed by an auditory prompt “OK” in a different female voice to prevent reliance on echoic memory. The intertrial interval was 1,500 ms. The entire experiment took between 20 to 30 minutes to complete.

2.4. Data Analysis

The results were analyzed in logistic mixed regressions using a generalized linear mixed effect model (GLMER) from the lme4 package (Bates et al., 2015) in R (R Core Team, 2021). The models analyzed the accuracy of the sequence recalling test as the dependent variable (1=correct, 0=incorrect). A correct response was recorded when participants correctly identified the stress position of all four consecutive tokens in a sequence (e.g., 1121=first-syllable stressed, first-syllable stressed, second-syllable stressed, first-syllable stressed). Thus, the percentage of correct responses was used as the dependent variable, and stimulus types (Naturally Produced Stimuli vs. Duration-only, Pitch-only, Duration-pitch Conflicting, Duration-pitch Matching Stimuli) were entered as independent variables, with Subject and Trial as random effects. The Naturally Produced Stimuli were set as the reference level for the independent variable, and the other stimulus conditions were compared against this baseline.

3. Results

The generalized mixed-effects model revealed significant main effects for Duration-only stimuli, Duration-pitch Conflicting stimuli, and Duration-pitch Matching stimuli ($p<.01$). The negative estimate

values for the Duration-only stimuli indicate that these stimuli significantly impair sequence recall accuracy compared to the Naturally Produced stimuli. When stress was signaled solely by pitch, participants did not exhibit any significant difference in their recall performance ($p>.05$). However, when the duration and pitch cues conflicted, these conflicting cues had a detrimental effect on sequence recall accuracy compared to the Naturally Produced stimuli, as indicated by the negative estimate of -2.73 . When duration and pitch cues were congruent and matched to signal stress, these matching cues facilitated better sequence recall performance, with the result approaching statistical significance ($p=.05$, Estimates= 0.37). Thus, the results of the analysis suggest that F0 (fundamental frequency) is the primary cue for perceiving stress patterns among Koreans, while duration serves as a secondary cue for stress perception among Korean L2 learners of English. The detailed results of the generalized mixed-effects model are presented in Table 1.

Table 1. Summary of results of the logistic regression

Variable	Estimates (SE)	Z	p-value
(Intercept)	1.15 (0.28)	4.05	<.01
Duration-only	-2.70 (0.20)	-13.62	<.01
F0-only	0.20 (0.19)	1.06	.29
Duration-F0 Conflicting	-2.73 (0.20)	-13.71	<.01
Duration-F0 Matching	0.37 (0.19)	1.92	.05

Figure 1 represents the correct rate of sequence recalling test as a function of stimuli condition (Natural stimuli, Pitch-only, Duration-only, Duration-pitch Conflicting, Duration-pitch Matching).

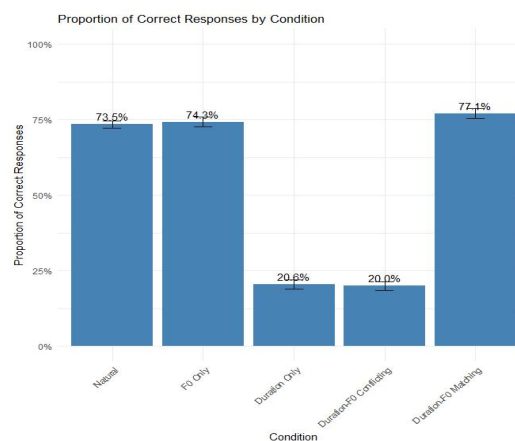


Figure 1. Correct rate of sequence recalling test as a function of stimuli condition.

4. Discussion

The present study aimed to investigate which suprasegmental cues Korean L2 learners of English rely on when perceiving lexical stress in nonword stimuli, specifically focusing on the roles of duration and fundamental frequency (F0). Firstly, the result indicated that Korean L2 learners primarily rely on F0 as the dominant cue for stress perception, as shown from the non-significant difference between processing naturally produced stimuli

and stimuli that only containing F0 cues. This result aligns with previous findings (Kim & Tremblay, 2021, 2022) which highlight the importance of F0 cues in Korean prosody. This reliance on F0 over duration, particularly when these cues were conflicting, suggests that F0 serves as a primary cue in the perception of English nonword stress pattern for Korean learners. This may be due to the prominence of F0 in Korean's intonational structure, where pitch patterns play a crucial role in demarcating prosodic boundaries (Jun, 1998; Jun & Fougeron, 2002).

However, the results also demonstrated that when duration and F0 cues were congruently signaling stress, learners showed improved accuracy in identifying stress patterns. The improvement in accuracy observed with the duration-F0 matched stimuli, despite the naturally produced stimuli also containing both duration and F0 cues, may be attributed to a practice effect (Fitts & Posner, 1967; Gopher et al., 1989; Hausknecht et al., 2007). Specifically, it is possible that participants became more familiar with the task by the time they reached the second phase of testing, leading to better performance. Since the first phase involved naturally produced stimuli, this initial exposure might have allowed participants to become more adept at the task, resulting in higher accuracy during the second phase with the duration-F0 matched stimuli.

Another possible explanation for this finding is that the fixed intensity level (70 dB) across all experimental stimuli may have reduced the role of intensity as a cue for stress perception, thereby making the duration and F0 cues more salient in the matching condition. Future research could further investigate the impact of intensity variation on stress perception among Korean L2 learners of English to better understand the relative influence of this cue in their perception of lexical stress. In addition to considering the role of intensity, it is also important to note that the findings suggest Korean L2 learners do not rely solely on F0 when identifying lexical stress.

Although F0 emerged as a dominant cue in several conditions, it is important to acknowledge that Korean L2 learners did not rely solely on F0 when identifying lexical stress, contrary to the findings of Kim & Tremblay (2021). The fact that Korean L2 learners of English were able to perform the sequence recalling task in the F0-duration conflicting condition as well as in the duration-only condition at levels exceeding chance (6.25%) indicates that they were not solely dependent on F0. This finding is consistent with prior research demonstrating that listeners tend to rely on secondary cues when the primary cue is either absent or unreliable (Francis et al., 2008; Gordon et al., 1993; Holt & Lotto, 2006 among many others). Similarly, in the current study, when F0 cue is unreliable (F0-duration conflicting condition) or absent (duration-only condition), duration becomes the primary cue in distinguishing stress patterns. If Korean listeners did not utilize duration cues at all in stress perception, we would expect the accuracy levels in these two conditions to be markedly lower, potentially even below chance level. Additionally, this differential weighting between F0 and duration in the perception of lexical stress is further supported by Lee (2022) such that Korean L2 learners weight F0 more heavily than duration cue when perceiving English lexical stress. These results highlight the flexibility of cue integration in speech perception, highlighting the adaptability of L2 learners in navigating complex prosodic environments.

Taken together, the results reveal significant insights into the perceptual sensitivity by Korean L2 learners and contribute to the

broader understanding of how non-native speakers acquire new prosodic features that does not exist in their native language, supporting cue weighting approach (e.g., Francis & Nusbaum, 2002; Francis et al., 2000; Holt & Lotto, 2006). The cue-weighting theory of speech perception suggests that listeners acquire speech categories or contrasts in both their first language (L1) and a second language (L2) by selectively attending to specific acoustic dimensions, based on the assumption that speech perception is inherently multidimensional. Thus, acoustic cues weight differently not only across phonetic categories but also across languages. As a result, listeners from different linguistic backgrounds perceive the same acoustic stimuli differently, shaped by the specific weighting of acoustic cues in their L1. According to the cue-weighting theory, the influence of individual acoustic cues that distinguish phonetic categories in the L1 is transferred to the L2.

Within this framework, the cue-weighting approach emphasizes the functional weight of suprasegmental cues in expressing lexical contrast. Specifically, it examines how listeners prioritize these cues for lexical contrasts in their L1 and how this influences their perception and processing of suprasegmental cues in L2 prosodic contrasts. If a particular suprasegmental cue is used more heavily in the L1, it is likely to be utilized similarly for prosodic categories in the L2. The greater the importance of a cue in the L1, the more it is expected to influence the perception and processing of L2 prosodic contrasts (e.g., Kim & Tremblay, 2021; Lee, 2022; Qin et al., 2017). This study also demonstrated the impact of L1 on L2 learners' perception of prosodic contrasts by showing that Korean learners weighted F0 more heavily than duration cue in processing English lexical stress. The limitation of the current study is the lack of the results from native English speakers, as we are not sure how much reliance L2 listeners would weight on the duration cues when F0 cues are absent, as compared to the native listeners.

In summary, this study reinforces the notion that L1 prosodic structure significantly influences L2 prosodic perception, lending support to the Cue Weighting Approach. Additionally, the findings highlight the pivotal role of F0 in the stress perception of Korean L2 learners of English, with duration functioning as a secondary, yet still significant, cue. The ability of Korean learners to shift their perceptual weighting to a secondary cue in the absence of the primary cue suggests the potential for dynamic cue integration. Future research could explore whether Korean L2 learners might eventually adjust their cue weighting to give duration a weight comparable to that of F0 when processing lexical stress.

Acknowledgements

We would like to thank Zhen Qin for providing the acoustic stimuli and experimental setup used in this study.

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Package 'lme4'. *Convergence*, 12(1), 1-129.
- Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer (version 5.3) [Computer program]. Retrieved from <http://www.praat.org/>
- Briggs, S., Dobson, B., & Rohle, T. (2003). The University of Michigan examination for the certificate of proficiency in English:

- Official past papers. Oxford, UK: Oxford University Press.
- Chang, C. B. (2018). Perceptual attention as the locus of transfer to nonnative speech perception. *Journal of Phonetics*, 68, 85-102.
- Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155-190.
- Chrabaszc, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, 57, 1468-1479.
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness." *The Journal of the Acoustical Society of America*, 110(3), 1606-1618.
- Fitts, P. M., & Posner, M. I. (1967). Human performance. Pacific Grove, CA: Brooks/Cole.
- Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, 62(8), 1668-1680.
- Francis, A. L., Kaganovich, N., & Driscoll-Huber, C. (2008). Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *The Journal of the Acoustical Society of America*, 124(2), 1234-1251.
- Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 349-366.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, 27(4), 765-768.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1(2), 126-152.
- Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, 21(4), 347-353.
- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, 71(1-3), 147-177.
- Gordon, P. C., Eberhardt, J. L., & Rueckl, J. G. (1993). Attentional modulation of the phonetic significance of acoustic cues. *Cognitive Psychology*, 25(1), 1-42.
- Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Moriarty Gerrard, M. O. (2007). Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability. *The Journal of Applied Psychology*, 92(2), 373-385.
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119(5), 3059-3071.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30(6), 685-701.
- Jun, S. A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, 15(2), 189-226.
- Jun, S. A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology* (pp. 209-242). Norwell, MA: Kluwer Academic Publishers.
- Jun, S. A., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147-172.
- Kang, H., & Kim, H. J. (2019). Segmental and suprasegmental effects on Korean listeners' English stress perception. *Korean Journal of Linguistics*, 44(4), 721-747.
- Kim, H., & Tremblay, A. (2021). Korean listeners' processing of suprasegmental lexical contrasts in Korean and English: A cue-based transfer approach. *Journal of Phonetics*, 87, 101059.
- Kim, H., & 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. *Frontiers in Communication*, 7, 845430.
- Kirby, J. P., & Ladd, D. R. (2015, August). Stop voicing and F0 perturbations: Evidence from French and Italian. Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK.
- Lee, G. (2022). Perceptual weighting on English lexical stress by Korean learners of English. *Phonetics and Speech Sciences*, 14(4), 19-24.
- Lin, C. Y., Wang, M. I. N., Idsardi, W. J., & Xu, Y. I. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition*, 17, 316-346.
- Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, 52, 183-204.
- Tremblay, A., Broersma, M., & 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. *Bilingualism: Language and Cognition*, 21, 640-652.
- Qin, Z., Chien, Y. F., & Tremblay, A. (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, 38(3), 541-570.
- R Core Team. (2021). R: A language and environment for statistical computing (version 4.1.2) [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Serniclaes, W. (1987). Étude expérimentale de la perception du trait de voisement des occlusives du français (Doctoral dissertation). Université Libre de Bruxelles, Brussels, Belgium.
- Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33(2), 159-197.
- Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, 38, 260-271.

• **Goun Lee**, Corresponding author
 Assistant Professor, Dept. of General Education,
 Liberal Arts College, Dankook University
 152 Jukjeon-ro, Suji-gu, Yongin 16890, Korea
 Tel: +82-31-8005-3943
 Email: gounlee@dankook.ac.kr
 Areas of interest: Acoustic phonetics, L2 acquisition, Speech
 Production and Perception Interface