Production and Perception of Korean Word-level Prominence by Older and Younger Korean Speakers

Goun Lee
Sungkyunkwan University

Allard Jongman
University of Kansas

Abstract

Prominence refers to the relative emphasis that may be given to a syllable in a word (word-level prominence) or to one or more words in a phrase (phrase-level prominence). Korean has been claimed to have both word-level (Ko, 2013) and phrase-level (Jun, 1996) prominence, with the former realized mainly with duration and the latter with F0 height. However, given the claim that younger Korean speakers have lost duration as the main cue expressing word-level prominence (Kim & Han, 1998; Magen & Blumstein, 1993), it is not clear if and how younger Korean speakers produce word-level prominence. Thus, the current study aims to investigate whether Korean still has word-level prominence. In the acoustic study of the production of Korean word-level prominence (Experiment 1), measurements of duration, intensity, and F0 on (so-called) Korean stress minimal pairs by older and younger Korean speakers revealed that only duration distinguishes Korean word-level prominence. A perception study on word-level prominence in Korean (Experiment 2) revealed that both older and younger Korean listeners weighted the duration cue most heavily in identifying minimal pairs when two of the suprasegmental cues were orthogonally manipulated in each syllable. Interestingly, this perceptual weighting was only observed in the first syllable: none of the listeners changed their perception when cues were signaling second-syllable stress. Based on these acoustic and perceptual findings we conclude that Korean does not have word-level prominence, but only has a phonemic vowel length distinction.

Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen & Mette Hjortshøj Sørensen (Eds.), ***A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*** (pp. 271-301). Dept. of English, School of Communication & Culture, Aarhus University.

1. Introduction

Linguistic prominence is comprised of two levels of prosodic cues – word-level prominence and phrasal-level prominence. Word-level prominence creates lexical contrasts based on the acoustic manifestation of at least one cue, while phrasal-level prominence is conveyed by F0 peaks or valleys that express context-dependent pitch accents, which distinguish a prosodic boundary between words (e.g., Beckman, 1986; Cooper, Eady, & Mueller, 1985; Fry, 1958; Shport & Redford, 2014). For languages with word-level prominence, lexical prosody expresses whether certain syllables are more prominent than neighboring syllables within the same word. The prominence can be realized by multiple suprasegmental cues such as duration, pitch, and intensity. However, there is no absolute value that determines a prominent syllable: rather, the concept of strong-weak is abstract and relative to the adjacent syllables.

Languages selectively pick and choose which cues to use in expressing prominence. In stress languages like English, primary stress may be expressed by more than one acoustic cue, including increased F0, longer duration, and higher intensity compared to unstressed syllables (Fry, 1955; Fry, 1958; Gay, 1978). Segmental cues like vowel reduction can also express lexical stress (Gay, 1978; Koopmans-Van Beinum, 1980). For languages with phrasal-level prominence, pitch is used to group prosodic structure together that is determined by the domain of the accentual phrase. For example, in Japanese, which has both word-level prominence as well as phrasal-level prominence, a low boundary tone occurs at the beginning of every utterance and at the AP (Accentual Phrase)-final boundary. Thus, when this low tone occurs within a sentence, listeners interpret it as belonging to the phrasal boundary (Beckman & Pierrehumbert, 1986).

However, distinguishing the cues that are used to mark word-level prominence from those used to express phrasal-level prominence might be difficult, because the same acoustic correlates that are used to indicate word-level prominence in stress languages – F0 and duration – are utilized to indicate prosodic prominence as well. Cross-linguistically, syllables in sentence-final position are lengthened, and pitch is raised at the non-sentence-final phrasal boundary (Beckman, 1986; Tyler & Cutler, 2009; see Japanese for a low boundary tone at non-sentence-final phrasal-boundary position).

Regarding Korean, it is sometimes claimed that Korean has both word-level prominence (i.e., stress) and phrasal-level prominence. Previous research has claimed that Korean stress is mainly realized in terms

of duration (e.g., Ko, 2013), and phrasal-level prominence realized with pitch (at Accentual Phrase (AP)-level), intensity (at AP-initial position), and duration (at Intonational Phrase (IP)-final position) (Jun, 1993; 1998). While duration is claimed to be a cue to stress, the same cue can also mark the phrasal boundary in expressing phrasal-level prominence in Korean. In the following section, we will first briefly review Korean phrasal-level prominence and word-level prominence, and then discuss the cues that are used to indicate prominence at different levels.

1.1. Phrasal-level prominence in Korean

Phrasal-level prominence plays a crucial role in speech segmentation and production by marking the phrasal boundary in terms of F0 or duration (e.g., Beckman, 1986). Phrasal boundary tones are marked with a raised F0 (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980), and phrase-final position is marked with an increased duration (Klatt, 1975; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). Listeners use these higher-level prosodic cues in segmenting ambiguous segmental information both in L1 and L2 speech (Cho et al., 2007; Christophe et al., 2004; Coughlin & Tremblay, 2012; Kim, 2004; Kim & Cho, 2009; Tremblay, Coughlin, Bahler, & Gaillard, 2012).

The Accentual Phrase (AP), an intonationally defined unit, can mark a phrasal boundary in Korean. The hierarchical structure of prosodic boundaries consists of the syllable, the Phonological Word (PW), the Accentual Phrase (AP), and the Intonational Phrase (IP). The edge of the larger unit always coincides with the edge of the smaller unit: the edge of IP always coincides with the edge of AP, and the edge of AP always coincides with the edge of PW (Selkirk, 1984). Jun (1993; 1998) proposed in her Accentual Phrase (AP) theory that Korean has intonationally defined units (AP) that pattern independently from the word-level prosody.

In the Korean AP system, the initial boundary of the prosodic domain is always delimited with a low tone and the final boundary with a high tone (i.e., #LHLH#; # refers to an AP boundary; each syllable is associated with a tone ; Jun 1993; 1998). This LHLH tone pattern occurs when at least 4 syllables exist in one AP domain. When there are less than 4 syllables in an AP, 2 or 3 surface tone patterns appear by undershooting the initial two tones. For example, when an AP has 3 syllables, two different tone patterns can appear: a #LH# (or #HH#) pattern when the first two syllables are undershot, and a #LHH# (or #HHH#) pattern when

only the first syllable is undershot. When an AP has 2 syllables, only the #LH# (or #HH#) pattern can appear. When the domain-initial syllable is either aspirated or tense, the pitch is raised on the first syllable, bearing #HHLH# intonational pattern (Jun, 2000; Kim, 2004; Kim & Cho, 2009).

The IP-final boundary is also characterized by different tonal patterns such as L%, H%, LH%, and HL% (% refers to an IP boundary). When the AP boundary coincides with the IP boundary, the AP-final tone (L#) is overridden by the IP-boundary tone. At the IP-boundary, final lengthening also occurs along with the IP-boundary tone.

In addition to F0, previous studies have found that other cues, such as duration and amplitude, can also characterize the phrasal-level prominence. With respect to duration, phrase-final lengthening can mark IP boundaries in Korean. Jun (1993) and Chung et al., (1996) found that final lengthening does not occur at the AP level, but at the IP level. However, Cho and Keating (2001) and Oh (1998) found a small but significant AP-final lengthening effect compared to non-AP-final words. Although these studies are not consistent regarding AP-final lengthening in Korean, there is a strong consensus at least that phrase-final lengthening exists in IP-final position (Cho & Keating, 2001; Chung et al., 1996; Jun, 1993; 2000) .

Amplitude can also mark both the AP-initial and -final boundary in Korean. Jun (1995) found that the amplitude of the first syllable was greater than that of the second syllable when a trisyllabic reiterative word like ‘mamama’ was embedded in sentence-medial position. The amplitude of the first syllable was comparable with that of the third syllable, but the third syllable was also marked with low F0 because it was in AP-final position.

1.2. Word-level prominence in Korean

Historically, Korean has been claimed to have a vowel length distinction. The long vowels only appear in the first syllable (Heo, 1965) and are realized with a rising tone. Although it had been widely accepted that Korean had a vowel length distinction for pairs with identical vowel quality (see the IPA manual, 1999, p. 44), most younger speakers have lost this distinction (Kim, 2001; Kim & Han, 1998; Magen & Blumstein, 1993) and only speakers from a few dialects like Chonnam (Ko, 2013) and North Kyungsang (Kenstowicz & Park, 2006) preserve the distinction.

This vowel length distinction has been argued to influence lexical stress[1] in Korean, which is realized as rhythmic shortening or lengthening. The traditional vowel shortening rule takes the long vowel as the underlying form, and posits that the long vowel undergoes vowel shortening, since the realization of the long vowel is only limited to the first syllable. When a monosyllabic word with a short vowel is combined with a monosyllabic word with a long vowel, the long vowel in the second syllable is shortened in the compound word. This vowel shortening also occurs when the long vowel is attached to vowel-initial suffixes (Kim-Renaud, 1974; B.-G. Lee, 1978), but optionally occurs when being attached to consonant-initial suffixes (Ko, 2002, 2013).

Based on the optional vowel shortening, Ko (2002, 2010, 2013) proposed that vowel shortening occurs to avoid accent clash.[2] Ko (2002) defines 'stress' to refer to "the metrical head physically realized on the surface", and 'accent' to refer to "the underlying specification for prominence on a syllable" (Ko, 2002, p. 81, lines 27-29). In other words, 'stress' is the actual location at which physical correlates of word-level prominence are realized with acoustic features such as duration, F0, and amplitude, while 'accent' is the potential location of stress. Ko (2013) claimed that the long vowel is realized with the stress on the syllable, and when a syllable is not realized with stress, the vowel remains as a short vowel. Therefore, when a suffix that is carrying an accent is attached to a monosyllabic long vowel verb stem, the stem vowel is shortened in order to avoid the accent clash. However, when the accent-triggering suffix is attached to a disyllabic verb stem, the long vowel does not need to undergo vowel shortening. The accent of these suffixes is never realized because stress in Korean needs to fall on the first syllable.

1.3. Phonetic evidence for lexical stress in Korean

Ko (2013) examined whether other acoustic correlates of lexical stress are realized along with vowel duration in two dialects of Korean. Since, unlike Seoul Korean, the Chonnam Korean dialect still preserves the vowel length distinction, Ko (2013) hypothesized that the Chonnam dialect might be more conservative in preserving lexical stress and therefore, the

[1] As lexical stress realized by a vowel length distinction is argued to be word-level prominence in Korea, we use 'lexical stress' and 'word-level prominence' interchangeably in this study.

[2] In her analysis, both 'stress' and 'accent' are used in reference to word-level prominence.

manifestation of the four acoustic correlates of stress will be more apparent compared to Seoul Korean. Two different age groups across two dialects (younger Chonnam speakers vs. older Seoul speakers) were chosen, because both the current Chonnam dialect and the traditional Seoul dialect (older Seoul speakers) still preserve the vowel length distinction. Four young Chonnam speakers (1 male, mean age = 34) and four old Seoul Korean speakers (2 males, mean age = 69) recorded 17 stress minimal pairs (e.g., *sákwa* 'apology' vs. *sakwá* 'apple') embedded in a contextually related sentence (e.g., As for **apples**, Taegu is famous for it) and in a contextually neutral sentence (e.g., 'Please pronounce apple clearly').[3] Three acoustic parameters – duration (ms), intensity (dB), and F0 (semitone) – were examined for the first and second vowel of the target word from the neutral sentence. Measurements averaged across the entire vowel from the stressed syllables (*sá from sákwa* 'apology') were compared to those of the unstressed syllables (*sa* from *sakwá* 'apple').

A series of paired t-tests found that young Chonnam speakers use duration, intensity, and F0 to distinguish the vowel-length minimal pairs. The vowel in the 'stressed first syllables' was 77.7 ms longer, 0.64 semitone higher, and 2.6 dB more intense than that in the 'unstressed first syllables'. The vowel in the stressed second syllables was 35.72 ms longer, 1.51 semitone higher, and 1.8 dB more intense than that in the 'unstressed second syllables'. On the other hand, older Seoul Korean speakers only used vowel duration in the first syllable and intensity in the second syllable to distinguish vowel-length minimal pairs. For the older Seoul speakers, the vowel in stressed first syllables was 96.71 ms longer than that in unstressed first syllables, and the vowel in stressed second syllables was 1.02 dB more intense than that in unstressed second syllables.

Additionally, results of two separate models of a mixed effect logistic regression indicated that, in the Chonnam dialect, all three correlates showed a significant effect on predicting stress on the first syllable, whereas only vowel duration and F0 showed a significant effect on predicting stress on the second syllable. On the other hand, in the Seoul dialect, vowel duration was found to be the only factor to predict stress on the first syllable, whereas both vowel duration and F0 showed a significant effect in predicting the stress on the second syllable.

---

[3] The claims regarding word-level prominence in Korean were made based on cases where the phonological process of vowel shortening occurs, whereas Ko (2013) used words with phonemically long and short vowels for the acoustic analysis.

Based on these results, Ko (2013) concluded that Chonnam uses vowel duration, F0, and intensity to express lexical stress, while the Seoul dialect is exhibiting a diachronic change from a stress language to a phrasal-accent language, as supported by the limited expression of word-level prominence. Ko (2013) argued that Seoul Korean had a "duration-based prominence system based on a very limited window of initial syllable" (p. 108, line 10), but the stress has eventually been lost in contemporary Seoul Korean. This raises the question how diachronic change in Seoul speakers' word-level prominence has affected the production and perception of lexical prominence.

1.4. A few potential problems

There are a number of issues in the design and interpretation of the Ko (2013) study that warrant a closer look at the notion of stress/word-level prominence in Korean. First, the participants in Ko (2013)'s study were limited to four older Seoul Korean speakers who came to the USA almost 40 years ago. Their exposure to English for this long period could have affected their production of word-level prominence in Korean, and also, their production in L1 might not reflect contemporary Seoul Korean, especially with respect to the ongoing language changes. It has been found that the use of VOT and F0 in indicating the three-way laryngeal distinction among stops in Korean has changed (Kang & Guion, 2008; Lee & Jongman, 2015; Lee, Politzer-Ahles, & Jongman, 2013; Perkins & Lee, 2010; Silva, 2006; Wright, 2007) and more importantly, the vowel length distinction has been claimed to have disappeared among younger Korean speakers (Kim, 2001; Kim & Han, 1998; Magen & Blumstein, 1993). Considering that, while living in the USA, Ko's speakers did not get as much L1 input as Korean residents, their productions might not be representative of Seoul Korean speakers.

Second, in order to determine the effect of the phrasal boundary on lexical stress, we need to examine productions in two different contexts. Ko (2013) recorded the tokens produced in a carrier sentence, where the target words were embedded in sentence-medial position. However, the carrier sentence that Ko used is unnatural due to the absence of the case marker after the target word. Ko first used a contextually-related sentence in order to prompt the intended word, and then asked the participants to read the target word embedded in a contextually neutral sentence, 'clearly ***apple*** pronounce' [t'o.bak. t'o.bak. **sa.gwa.** par.ɨm.ha.se.jo.]. However, she

deliberately omitted the case marker after the target word, *apple*, since the case marker is an allomorph which will appear as two syllables following an open syllable (e.g., [**sa.gwa**. ra.go.]) and as three syllables following a closed syllable (e.g., [**si.ʤaŋ.** i.ra.go.]). However, the sentence without a case marker sounds extremely unnatural, which might make participants produce the words with unnatural F0 patterns. In fact, Ko explained in her earlier study that the same sentence can be used to express two different prosodic frames, depending on how the sentence is parsed, as illustrated below. (Ko, 2002; p 144). Thus, Ko's stimuli had the potential to attract prosodic focus, introducing another level of prominence.

(1) Two possible ways of phrasing the frame sentence from Ko (2002)

a. Two independent prosodic domains
{t'obak t'obak} {**sa:gwa**} {par**ɨ**mhasejo}
'clearly apple say'

b. A single prosodic domain from the VP
{t'obak t'obak} {**sa:gwa** par**ɨ**mhasejo}
'clearly apple say'

Moreover, Ko instructed her speakers to produce the sentences with a falling intonation, which could also result in unnatural prosody. Ko (2013) explained that this was done in order to prompt the speakers to read the target words in a citation form and also to avoid a list effect. However, Ko did not provide a clear motivation, or references, to clarify how this procedure would achieve natural speech.

Also, Ko only measured raw values for each syllable and compared the difference between the values from the stressed syllables and unstressed syllables in their respective positions. However, the obtained difference might be misleading, because the same difference can also be found from vowel length minimal pairs. Moreover, a direct comparison between the first syllables of the stress minimal pairs does not provide insight into the relative differences between the syllables within a word. In addition, if the speakers claimed that they pronounced the stress pairs as homophones, Ko eliminated those tokens from the analysis. Thus, it is unclear whether the difference found in Ko (2013) is a fair representation of lexical stress, given the fact that the recording procedure was problematic and the data was subjectively selected.

Lastly, Ko (2013) made claims about the use of perceptual cues on the basis of her acoustic findings. Without any direct perception data, it is hard to conclude which cue(s) Korean listeners use in their perception. To our knowledge, no study has been conducted examining cue weighting for Korean word-level prominence. Thus, by conducting an acoustic study as well as a perception study, the present research aims to provide evidence regarding whether Korean indeed has lexical stress or simply has a vowel length distinction. In the acoustic study (Experiment 1), we examine whether we can replicate the findings from Ko in two contexts (i.e., at the sentence level and in words in isolation) with speakers of two generations. In the perception study (Experiment 2), we examine which acoustic cue(s) Korean listeners weight in identifying Korean stress pairs.

Given that it is still unclear whether contemporary Seoul Korean has word-level prominence, the present study aims to investigate whether younger Korean speakers produce word-level prominence. If there is word-level prominence in Korean and younger speakers produce it, will younger speakers also perceive it? If not, do they transfer the use of higher-level prosodic cues to the perception of word-level prominence?

2. Experiment 1: Production experiment

2.1. Methods

2.1.1. Participants

We recorded 21 male native speakers of Korean (10 older and 11 younger speakers). All subjects were born and raised in Seoul or Suwon, Kyunggi area where the standard Korean dialect is spoken. The mean age was 71.9 years (*SD* = 1.52) for the older speakers and 23.5 years (*SD* = 3) for the younger speakers. None of the subjects lived in any other region where a different dialect is spoken, except for the older Korean speakers during the Korean war from 1951-1953. All subjects were literate in Korean, and none of the subjects reported any speech or hearing disorder.

2.1.2. Stimuli

Seventeen minimal pairs that were used by Ko (2013) were adopted for the production study. These word pairs are traditionally considered as minimal pairs contrasting in vowel length; Ko (2013) treated them as minimal pairs in terms of stress. For example, for the minimal pair /sa:kwa/ ‘apology’ and /sakwa/ ‘apple’, Ko (2013) treated /sa:kwa/ ‘apology’ as having first-

syllable stress and /sakwa/ 'apple' as having second-syllable stress. These word pairs were first embedded in contextually related sentences in order to cue the semantic meaning of the target word to the participants, and then presented in a contextually neutral sentence as well as in isolation. The number of syllables of the contextually related sentences was balanced. Examples of semantically-related sentences and neutral sentences for the word pair /sakwa/ are as follows:

(1) Examples for 'apology'

a. Semantically-related sentence for '*apology*'
[ʤal.mo.sɨl. ha.mjən. **sa:.gwa**.ha.go. mʌn.ʤʌ. joŋ.sʌ.lɨl. pin.da.] (16 syllables)
'If you do wrong, you should give an **apology** first and ask for forgiveness.'

b. Semantically neutral sentence for '*apology*'
[i. dan.ʌ.nɨn. **sa:.gwa.** im.ni.da.]
'This word is **apology**'

(2) Examples for 'apple'

c. Semantically-related sentence for '*apple*'
[ʤɛ.sa. gwa.il.lo. **sa.gwa**.wa. pɛ.ga. ʤa.ʤu. sa.joŋ.dwɛn.da.] (16 syllables)
'For fruits to use at ancestor veneration ceremonies, **apples** and pears are often used.'

d. Semantically neutral sentence for '*apple*'
[i. dan.ʌ.nɨn. **sa.gwa.** im.ni.da.]
'This word is **apple**'

Only the tokens that were produced in neutral sentences (e.g., critical words produced in AP-initial position) and in isolation were examined. All stimuli were presented in Korean orthography in a randomized order, without any indication of the vowel length or stress location. In total, 714 tokens were recorded in a contextually neutral sentence (17 pairs x 2 repetitions x 21 speakers), and 357 tokens in word isolation (17 pairs x 21 speakers).

2.1.3. Procedure

The recordings were conducted in Suwon and Seoul. For the older Korean speakers, the recording was made in a quiet room in a local hotel, and for the younger Korean speakers, the recording was made in a seminar room at Sungkyunkwan University in Seoul, Korea. A Marantz Digital Recorder (PMD 671) and a Shure head-mounted microphone were used for the recording of both groups. The subjects were asked to read the stimulus sentences where the target words were embedded in different carrier sentences. First, the subjects read the target words embedded in contextually related sentences. Immediately after that, the subjects read the same target words embedded in the contextually neutral sentences with two repetitions. Then, the speakers read the same target word in isolation with one repetition. The sampling rate of the recording was 22,050 Hz and these recordings were analyzed using the speech analysis program Praat (version 5.4.03) (Boersma & Weenink, 2018).

2.1.4. Measurements

Duration, intensity, and F0 values were measured for each vowel from the first and second syllable of the target words. Duration was measured from the onset of F1 to the offset of F2 of each syllable.[4] Tokens that exhibited devoicing of high vowels between two voiceless consonants were eliminated from the analysis. This applied to both members of the minimal pairs. A total of 102 tokens (47 from the older speakers' productions) were eliminated from the productions recorded at the sentence level, and 58 tokens (28 from older speakers' productions) were eliminated from the productions recorded in isolation. The intensity values were averaged over each vowel. For F0, the F0 values from 20% to 80% of the duration of each vowel were averaged to avoid perturbation effects from the preceding consonant.[5] Thus, a total of , measurements (612 tokens x 5 measurements) were taken from the tokens produced in contextually neutral sentences, and 1,495 measurements (299 tokens x 5 measurements) were taken from the tokens produced in isolation.

To control for differences across speakers in terms of duration, F0, and amplitude of the target syllable, second-to-first syllable ratios for these measurements were used, following Beckman (1986)'s formulas:

---

4 We used duration ratio in order to control for any variations in speech rate.

5 We used average F0 values to be able to directly compare our measurements to those of previous studies.

F0 ratio (in semitones) = 17.31 ln[Hz(S2)/Hz(S1)]
Average intensity ratio = dB (S2) - dB (S1)
Log duration ratio = ln[ms(S2)/ms(S1)].

Thus, it is expected that first-syllable stressed words (e.g., [sa:gwa] 'apology') have a negative value of each ratio, and second-syllable stressed words (e.g., [sagwa] 'apple') will result in a positive value.

2.1.5. Data analysis

Factorial repeated measures ANOVAs were conducted with the second-to-first-syllable ratio values for intensity, duration, and F0 as dependent variables, and Stress (first syllable vs. second syllable), Age Group (older vs. younger), and Phrasal Condition (sentence vs. isolation) as independent variables.

2.2. Results

2.2.1. Duration

When examining second-to-first syllable duration ratios, the results showed main effects of Stress [$F(1, 19)=11.92$, $p=.003$], Age Group [$F(1, 19)=4.51$, $p=.005$], and Phrasal Condition [$F(1, 19)=10.60$, $p=.004$]. These results indicate that first-syllable stressed words had smaller duration ratio values (0.21) than second-syllable stressed words (0.34), and older Korean speakers produced stress pairs with smaller ratio values (0.19) than the younger Korean speakers (0.37). Also, speakers produced the stress pairs with smaller duration ratio values at the sentence level (0.19) than in isolation (0.36).

We also found a marginally significant three-way interaction among Stress, Age Group, and Phrasal Condition [$F(1, 19)=3.78$, $p=.067$], indicating that the duration ratio by stress between the two speaker groups was marginally affected by Phrasal Condition. The duration ratio difference between the first- and second-syllable stressed words as a function of Phrasal Condition was greater for the older speakers (sentence: 0.24, isolation: 0.1), than for the younger speakers (sentence: 0.07, isolation: 0.08).

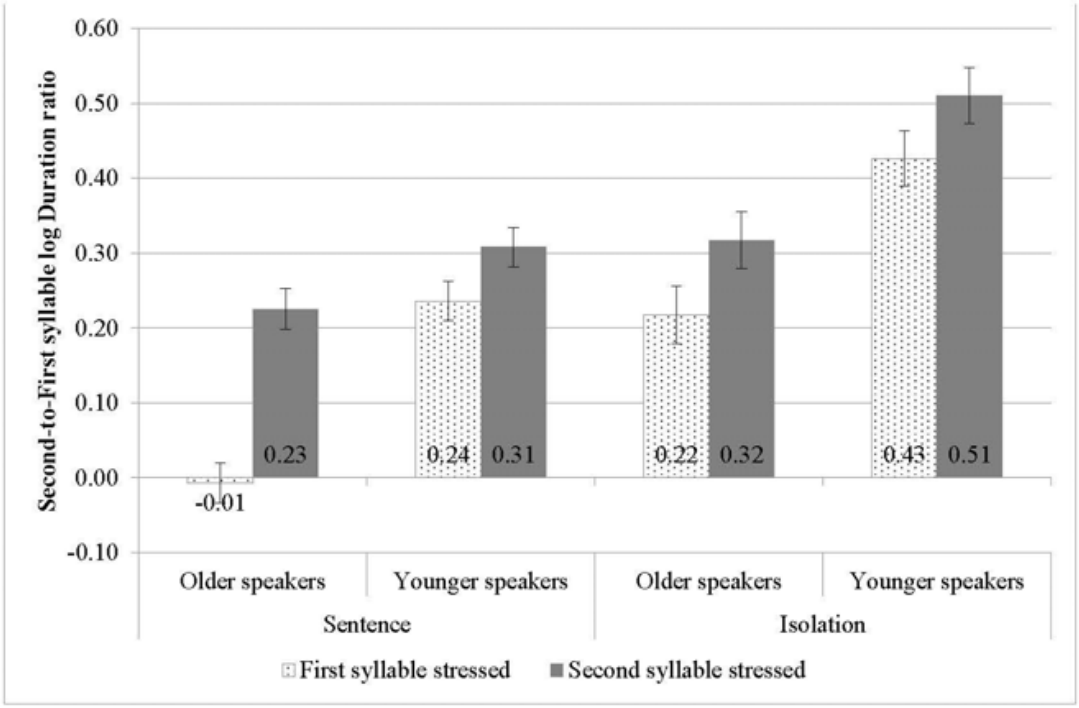

Figure 1. Second-to-first syllable log duration ratio values of the first- and second-syllable stressed words in two different contexts between two speaker groups.

Figure 1 illustrates the duration ratio values of the first- and second-syllable stressed words in two different contexts between the two speaker groups.

2.2.2. Intensity

When examining second-to-first syllable intensity ratios, the results showed main effects of Stress [$F(1, 19)=10.64$, $p<.001$] and Phrasal Condition [$F(1, 19)=46.50$, $p<.001$]. These results indicate that the first-syllable stressed words had a smaller intensity ratio (-0.19 dB) than the second-syllable stressed words (0.37 dB), and the intensity ratio values were greater for the productions from the sentence level (2.10 dB) than those from isolation (-1.92 dB). We also found a two-way interaction between Stress and Age Group [$F(1, 19)=4.85$, $p<.001$], indicating that the intensity ratio difference between the stress pairs was greater for the older speakers than the younger speakers. Figure 2 illustrates the second-to-first syllable intensity ratio values of the first- and second-syllable stressed words in two different contexts between the two speaker groups.

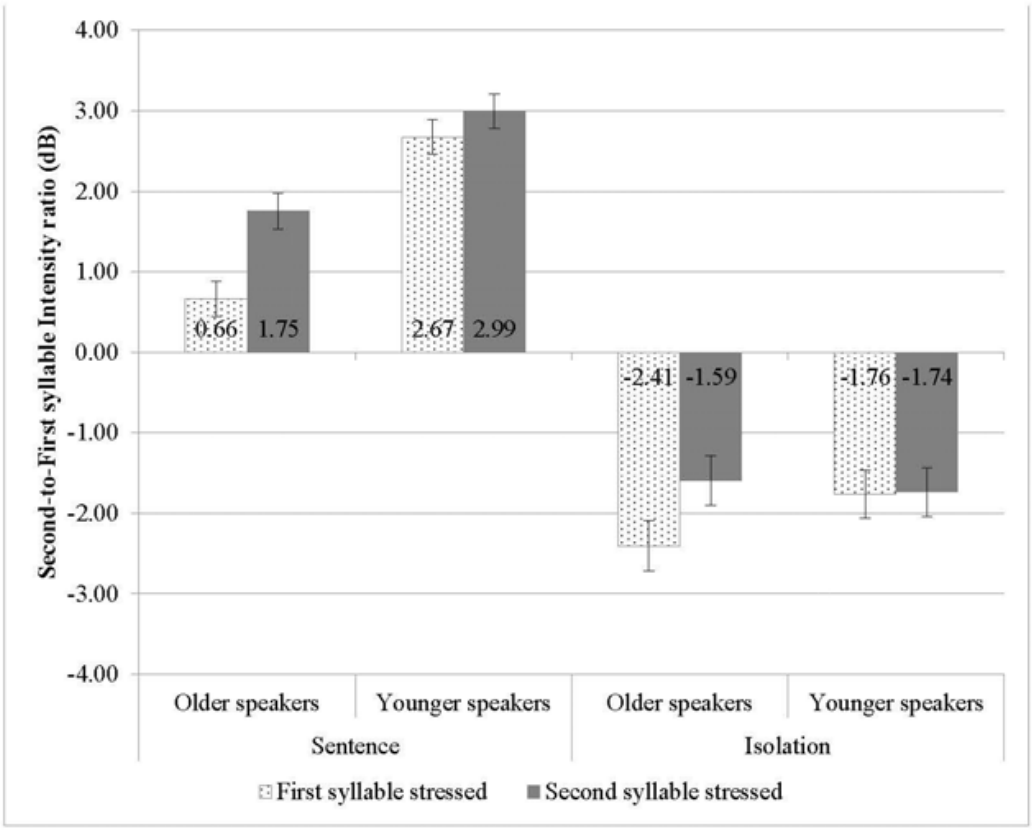

Figure 2. Second-to-first syllable intensity ratio values of the first- and second-syllable stressed words in two different contexts between two speaker groups.

2.2.3. F0

When examining second-to-first syllable F0 ratios, only a main effect of Phrasal Condition [$F(1, 19)=81.50$, $p<.001$] was found, indicating that the F0 ratio was greater in the productions at the sentence level (0.98) than in the productions in isolation (-2.21). We also found a significant interaction between Phrasal Condition and Age Group [$F(1, 19)=6.23$, $p<.001$], indicating that the F0 ratio difference between the two contexts was greater for the younger speakers (1.85) than the older speakers (1.62). Figure 3 illustrates the second-to-first syllable F0 ratio values of the first- and second-syllable stressed words in two different contexts between the two speaker groups.

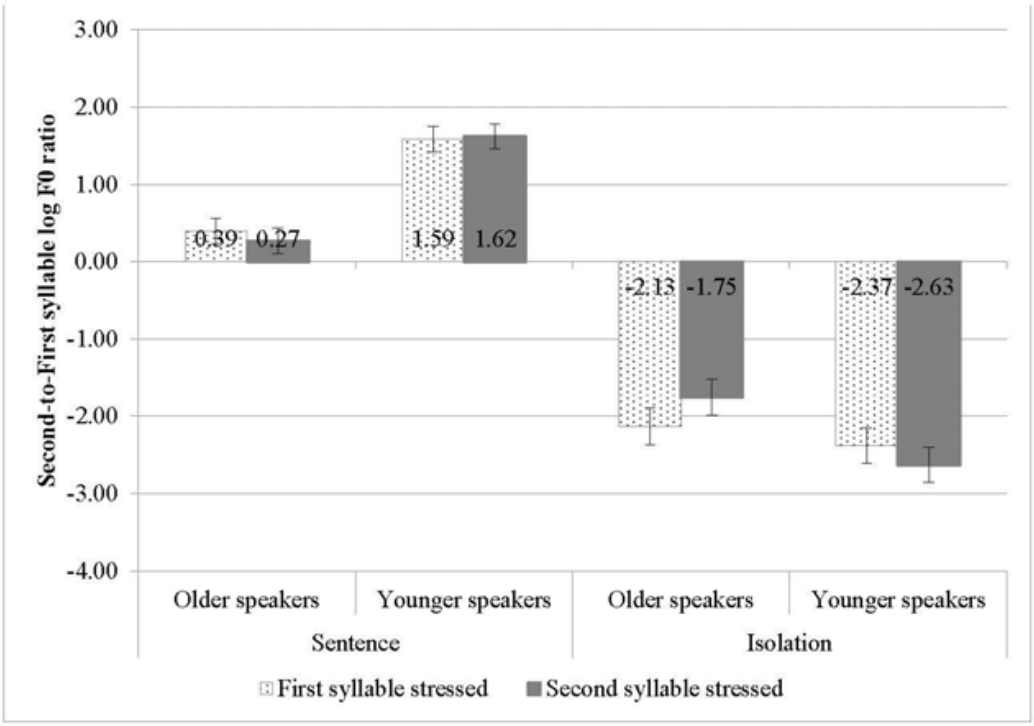

Figure 3. Second-to-first syllable F0 ratio values of the first- and second-syllable stressed words in two different contexts between two speaker groups.

2.3. Discussion and conclusion

A comparison of the ratio values in the two different contexts revealed several interesting facts. First, we found that both older and younger Korean speakers use durational difference in their production of Korean stress, as supported by the main effect of Stress on the second-to-first syllable duration ratio. We also observed a generational difference in the way the two groups of speakers used duration. Younger speakers always produced a second syllable that was longer than the first syllable, regardless of the stress position. In contrast, older speakers produced a longer first syllable at the sentence level, as shown by the negative values of the second-to-first syllable duration ratio. We also found that different phrasal levels affect the durational difference, as shown by the main effect of Phrasal Condition, suggesting that the final lengthening effect on the second syllable in isolation partially neutralized the effect of the vowel length distinction in the first syllable. This final-lengthening effect on stress seems to influence older speakers more than younger speakers. This is not surprising, considering

that previous studies have claimed that younger speakers have been losing the durational differences in contemporary Seoul Korean (Kim, 2001; Kim & Han, 1998; Magen & Blumstein, 1993).

Second, intensity is used as a cue to stress only by older speakers, as supported by a two-way interaction between Stress and Age Group. The ratio analysis for each condition revealed that intensity ratio only cues stress for older speakers at the sentence level. However, how intensity expresses prominence within a word varied as a function of context: intensity ratio values in isolation consistently showed negative values, while intensity ratio values at the sentence level consistently had positive values. This indicates that at the sentence level, the second syllable had higher intensity values than the first syllable, whereas in isolation, the first syllable had higher intensity values than the second syllable. The different intensity pattern between the two syllables across different contexts suggests that intensity is strongly affected by the phrasal-boundary effect, resulting in higher intensity at the boundary-initial position in isolation than at the boundary-initial position in sentence-medial position.

With respect to F0, no effect of stress was found for either speaker group, indicating that F0 is not a reliable parameter to indicate Korean stress. The two-way interaction between Phrasal Condition and Age Group indicates that F0 clearly marks the phrasal domain, especially in younger Korean speakers' productions.

The question then arises whether duration was ever used as a primary cue to lexical stress. Although we predicted that the duration effect would be weakened in isolation due to final lengthening, if Korean indeed had lexical stress, multiple cues in the first syllable should still indicate lexical stress. While both duration and intensity were significant cues to stress, inconsistent ratio patterns were observed as a function of Phrasal Condition. If Korean had lexical stress expressed with duration and intensity, as found by Ko (2013), both cues should pattern similarly in the two contexts. However, such a pattern was not found. Note that while the duration ratio values showed positive values in both contexts except for older speakers' productions at the sentence level, intensity ratios were either positive (at the sentence) or negative (in isolation) for both speaker groups. Therefore, the absence of consistent results across two contexts between the two speaker groups seems to suggest that duration could not be a cue to stress, and the effect of intensity was due to the effect of phrasal-level prominence.

However, there is still a possibility that lexical stress has diminished only in production but still exists in perception. Thus, the next issue to investigate is whether Korean speakers will only be sensitive to duration in differentiating stress in perception, or whether they will use other cues as well. For example, will older Korean speakers use intensity in addition to duration in distinguishing Korean stress, as we found in their production? Will younger listeners be only sensitive to duration or will they also be sensitive to other cues like intensity and F0? Considering that the second-to-first syllable duration ratio that we found in Experiment 1 was rather small, albeit significant, younger listeners may not rely on duration at all in perception. These issues will be addressed in the following perception study.

3. Experiment 2: Perception

Stress is realized by multiple cues (e.g., Ladefoged, Draper, & Whitteridge, 1958; Lehiste & Peterson, 1959; Lieberman, 1960) and the intrinsic characteristics of stressed syllables – longer duration, higher F0, and greater intensity – are strongly correlated with the perception of stress. When one of these parameters is not in the predicted direction, there is always a trade-off effect with other cues (e.g., Lieberman, 1960). That is, when one acoustic parameter (e.g., F0) does not contribute to the perception of stress, other cues (e.g., amplitude) may compensate and take on a greater role. This pattern was found not only in free-stress languages, such as English and Dutch, but also in fixed-stress languages such as Spanish (Llisterri, Machuca, de la Mota, Riera, & Ríos, 2003) and Arabic (de Jong & Zawaydeh, 1999), suggesting that stress is conveyed by multiple cues, and different languages use these acoustic cues with different degrees of saliency in indicating stress.

Based on this, we will use a perception study to re-examine whether Korean has a truly lexical stress that is mainly realized with duration or whether Korean only has a phonemic vowel length distinction. If Korean has both lexical stress and phonemic vowel length, Korean listeners should be sensitive to duration in the first syllable and weight intensity in the second syllable in distinguishing Korean word-level prominence. If Korean has only phonemic vowel length, but not lexical stress, then Korean listeners may only be sensitive to duration in the first syllable.

In addition, we investigate if older and younger Korean listeners differ in their use of acoustic cues in processing stress contrasts in Korean.

Even if younger Korean speakers have lost the vowel length contrast (or lexical stress) in their production, they might still have a perceptual distinction, since they are exposed to the duration distinction in the speech of their elders. Moreover, we also aim to investigate whether in their perception Korean listeners use the same cue(s) they used in their production to distinguish the stress pairs.

3.1. Methods

3.1.1. Participants

The same 10 older Korean speakers and 12 younger Korean speakers who participated in Experiment 1 took part in the perception study on Korean stress pairs.

3.1.2. Stimuli

3.1.2.1. Original base token

The Korean minimal stress pair 'sakwa' was chosen as the stimulus because both members of the pair had a similar frequency of occurrence (frequency of first-syllable stressed /sa:kwa/ 'apology': 48 per 3 million words; frequency of second-syllable stressed /sakwa/ 'apple' : 63 per 3 million) as provided by the National Institute of Korean Corpus (2002). These tokens were produced by an older Korean male speaker (Speaker A, age 68) who did not participate in the production study. We selected the production of the first-syllable stressed word, /sa:kwa/ 'apology', as the baseline token in order to preserve possible acoustic information in the long vowel and also to minimize any possibility of losing acoustic information by lengthening the short vowel to a long vowel. The manipulation range was based on the minimum and maximum value of three acoustic parameters – duration, F0, and intensity – of both younger and older speakers' productions of the /sakwa/ pair from Experiment 1.

3.1.2.2. Stimulus manipulation

All stimuli were manipulated from the single token /sa:kwa/, so that we could control any unintended changes in phonation type or vowel quality. The stimuli were first produced in semantically-related sentences (e.g., *'If you do wrong, you should give an* ***apology*** *first and ask for forgiveness.'*), and then produced in a semantically neutral carrier sentence (i.e., [i. dan.ʌ.nin. **sa:.gwa.** im.ni.da.] '*This word is apology*'). The token that was produced in the neutral sentence was used as the baseline

token. For the stimulus manipulation, the maximum and minimum values of F0, intensity, and duration for the two sets of stimuli across all the older and younger speakers were used as endpoints. For each condition, two parameters (e.g., duration x F0) were orthogonally manipulated to signal the stress pattern while the other cue (e.g., intensity) was controlled to be ambiguous (step 3 in Table 1). Each cue had 5 steps from unstressed to stressed syllable based on the acoustic data collected in Experiment 1. We also made sure the step size was greater than the Just Noticeable Difference (JND) (Flanagan, 1955; Flanagan & Saslow, 1958; Klatt, 1973; Fujisaki, Nakamura, & Imoto, 1975; Klatt & Cooper, 1975; Nishinuma, Di Cristo, & Espesser, 1983; Turk & Sawusch, 1996) for each cue, so that the listeners could perceive the difference between successive steps. When the first syllable was manipulated, the second syllable was controlled to be ambiguous (at step 3) between a stressed and an unstressed syllable, and the first syllable was controlled to be ambiguous (at step 3) when the second syllable was manipulated. These manipulated tokens were then embedded in a semantically neutral carrier sentence (e.g., This word is ___ . [i. dan.ʌ.nin. _____ im.ni.da.]) produced by Speaker A, and presented as the auditory stimuli in the perception experiment.

Duration, F0, and intensity were all manipulated using Praat (Boersma & Weenink, 2018). Table 1 represents the manipulation values of the five steps for the first and second syllables.

	First syllable					Second syllable				
	Unstressed				Stressed	Unstressed				Stressed
Step	**1**	2	3	4	**5**	**1**	2	3	4	**5**
Duration (ms)	56	85	114	143	172	87	89.25	91.5	93.75	96
F0 (Hz)	98.95	103.32	107.69	112.06	116.43	112.63	117.02	121.41	128.80	130.19
Intensity (dB)	60.00	62.75	65.5	68.25	71.00	60.00	61.25	62.50	63.75	65.00

Table 1. Five steps of manipulation values of first and second syllable for duration, F0, and intensity.

Thus, 25 tokens were created for each manipulation condition (e.g., 5 steps of duration x 5 steps of intensity; 5 steps of intensity by 5 steps of F0; 5 steps of F0 by 5 steps of duration) for each syllable. Three pairs of cues were manipulated: F0 & intensity; F0 & duration; duration & intensity. The 5 steps of both cues in each pair were crossed to form 25 stimuli, while the third cue was controlled to be neutral at step 3. For example, we created 25 stimuli by manipulating F0 and duration, while keeping intensity at step 3. In all, we defined 75 stimuli in this way. However, this procedure resulted in repeating conditions with two cues at level 3 twice, and three cues at level 3 three times for a total of 14 repetitions. Therefore, 61 unique stimuli were created with this procedure at each syllable level for a total of 122 stimuli. These were each repeated 3 times for a total of 366 tokens for each subject.

3.1.3. Procedure

Before the stress perception test, we examined the older Korean listeners' hearing sensitivity using a pure tone threshold test. Using an up-5dB, down-10dB procedure, 6 octaves were tested in total: 250, 500, 1000, 2000, 4000, and 8000 Hz. Normal-hearing thresholds were defined as thresholds which are better than 20 dB. All older listeners passed the threshold of the hearing acuity test.

Next, a word identification task was employed to examine which suprasegmental cue(s) Korean listeners are sensitive to in perceiving stress contrasts, and whether there is a generational difference between older and younger Korean listeners. First, participants were presented with two pictures of objects denoting the target words on a computer screen, associated with either number key [1] (first-syllable stressed word, / sa:kwa/) or [0] (second-syllable stressed word, /sakwa/) on the keyboard. Next, they were asked to identify the auditorily presented word by clicking either the [1] or [0] key. The position of the pictures and the numbers associated with them were counterbalanced across participants.

The experiment was conducted with three different blocks of 122 trials in randomized order. The intertrial interval (ITI) was 1500 ms. A practice session with 12 trials was conducted before the main experiment to ensure that the participants were familiar with the task. The subjects were allowed to take a short break between blocks.

3.1.4. Data analysis

We conducted a binomial logistic regression to examine the effect of three acoustic parameters (duration, intensity, and F0) on the perception of the first and second syllable between the two listener groups, using the lme4 package (Bates, 2005; Bates & Maechler, 2010) in the R statistical environment (R development Core Team, 2012, Version 3.1.2). The model had Response (/sa:kwa/ 'apology' vs. /sakwa/ 'apple') as a dependent variable, and Age Group (younger vs. older), Syllable (first vs. second), Intensity manipulation (steps 1-5), Duration manipulation (steps 1-5), and F0 manipulation (steps 1-5) as fixed effects and Participants as random effect. The model tested main effects of the independent variables, two-way interactions between the cues (Duration by Intensity, Intensity by F0, F0 by Duration), two-way interactions between Age Group and Syllable, and three-way interactions among Age Group and two of the cues (e.g., Age Group by Duration by Intensity). When there was a significant interaction between the independent variables, we stratified the data by Syllable and Age Group to probe the interaction between the variables. The older speaker group was used as the baseline to determine whether younger Korean listeners are less sensitive to Korean stress. Thus, the baseline in the model was the older group's performance on words with second syllable prominence (e.g., /sakwa/ 'apple') with step 1 values of intensity, duration, and F0.

3.2. Results

A linear mixed-effects model fitted on all participants' responses in identifying the Korean stress pairs. A series of fitted mixed-effects regression models were tested in a stepwise analysis to find the most parsimonious model. Table 2 presents the result of the logistic regression on both syllables. Only significant results are reported here. The results revealed significant main effects of Age Group (p=.05), Syllable (p<.01), and Duration (p<.01). These results indicate that older listeners' responses were more biased toward the first-syllable stressed word /sa:kwa/ (68 %) than those of younger listeners (57 %), and the listeners' response was biased toward the first-syllable stressed word (74%) when the second syllable was manipulated, as compared to the tokens for which the first syllable was manipulated (50%). Also, the probability that listeners perceived the second syllable as stressed was 53 % when the duration step was at 1, and decreased to 45%, 37%, 30%, and 27% when the duration step was at 2, 3,

4, and 5, respectively. We also found significant interactions between Age Group and Duration (p=.05), and Syllable and Duration (p<.01). In order to better understand these interactions, we stratified the data by Syllable, and then ran two separate models at each syllable level.

Variable	Estimate (*SE*)	*Z*	*p*
(intercept)	7.02 (1.82)	3.85	<.01
Age Group_young	-4.54 (2.29)	-1.98	=.05
Syllable	-4.03 (1.15)	-3.50	<.01
Duration	-2.25 (0.51)	-4.43	<.01
Age Group_young:Duration	1.24 (0.63)	1.98	=.05
Syllable:Duration	1.18 (0.32)	3.73	<.01

Table 2. Summary of results of the optimal model from the logistic regression examining responses at both syllable levels.

In the *post-hoc* analysis, we ran separate linear mixed-effects models examining all participants' responses to the tokens for which the first syllable was manipulated. The main effects of Age Group (p<.01) and Duration (p<.01) indicate that older listeners gave more first-syllable stressed responses (54 %) than younger listeners (47 %), and as the duration of the first syllable increased, listeners' responses shifted from words with second-syllable prominence to first-syllable prominence. The second-syllable stressed response rate was 78 % when the duration step was at 1, and decreased to 65%, 48%, 34%, and 27% when the duration step was at 2, 3, 4, and 5, respectively. Table 3 presents a summary of results of the model at the level of the first syllable.

Variable	Estimate (*SE*)	*Z*	*p*
(intercept)	2.80 (0.28)	9.85	<.01
Age Group_young	-1.40 (0.37)	-3.76	<.01
Duration	-1.00 (0.57)	-17.83	<.01
Age Group_young:Duration	0.59 (0.07)	8.50	<.01

Table 3. Summary of results of the logistic regression examining responses at the first syllable level

Figure 4 presents the probability of second-syllable stressed responses between the two groups for the tokens for which three acoustic cues of the first syllable were manipulated, showing the different use of the duration cue (red lines) between the two listener groups (old: dashed lines; younger: solid lines) in perceiving Korean stress pairs.

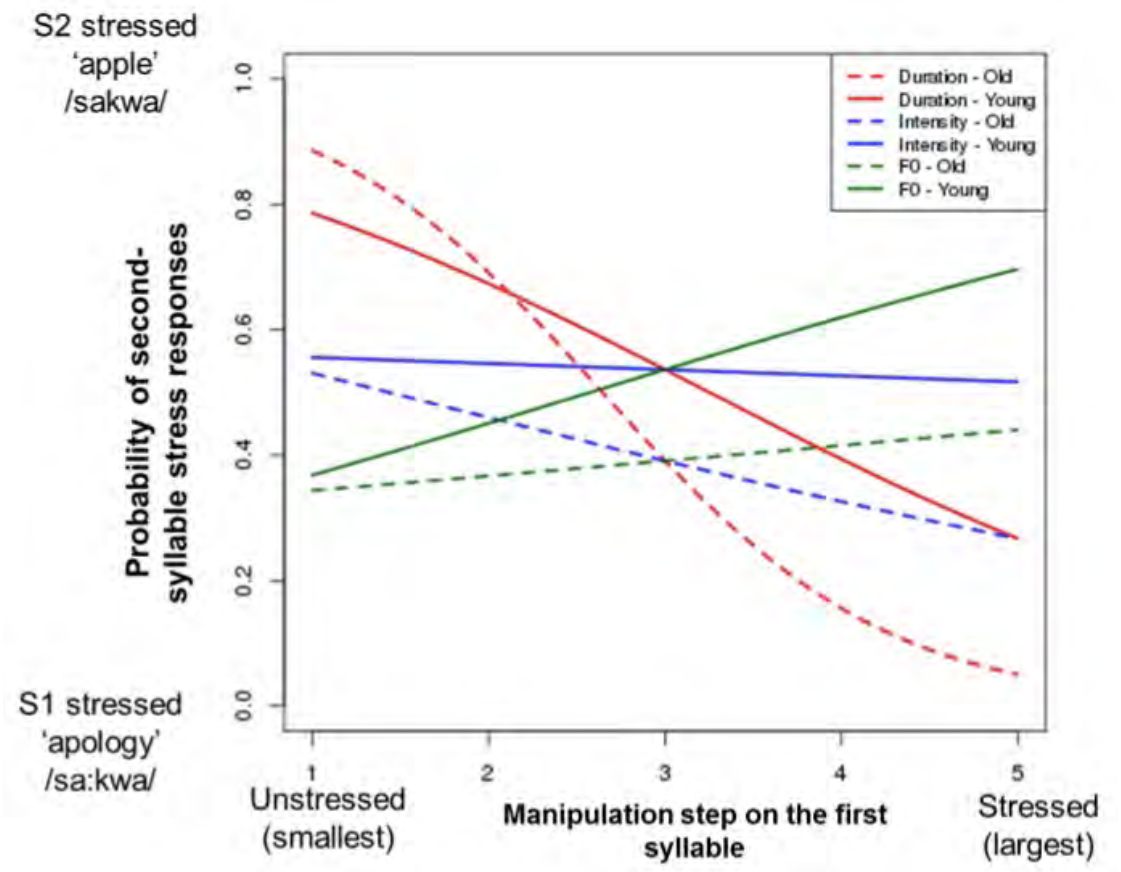

Figure 4. Probability of second-syllable stressed responses, /sakwa/, between the two listener groups. X-axis indicates the manipulated steps of each cue. 1 indicates that duration values were at the minimum endpoint, expressing first-syllable unstressed; and 5 indicates that duration values were at the maximum endpoint, expressing first-syllable stressed. Dotted lines indicate older Korean listeners' responses, and solid lines indicate younger Korean listeners' responses. Listeners' responses for each cue are illustrated with different colors: red, blue, green lines indicate listeners' responses for duration, intensity, and F0, respectively.

In order to explore the interactions between Age Group and Duration, we further conducted a separate linear mixed-effects model examining participants' responses to the tokens for which the first syllable was manipulated as a function of the listener groups (see Table 4).

Variable	Estimate (*SE*)	*Z*	*p*
(intercept)	2.81 (0.31)	8.94	< .01
Duration	-1.01 (0.06)	-17.81	< .01

Table 4. Summary of results of the logistic regression examining responses of older listeners at the level of the first syllable.

For the older listeners, we found a main effect of Duration ($p < .01$), indicating that older listeners' second-syllable stressed responses increased as duration decreased. The second-syllable stressed response rate was 52 % when the duration step was at 1, and decreased to 41 %, 31 %, 22 %, and 17 % when the duration step was at 2, 3, 4, and 5, respectively. For younger listeners, we also found a main effect of Duration ($p<.01$), indicating that the probability of the second-syllable stressed responses increased as the duration decreased (see Table 5). The second-syllable stressed response probability was 54 % when the duration step was at 1, and decreased to 48 %, 42 %, 36 %, and 35 % when the duration step was at 2, 3, 4, and 5, respectively.

Variable	Estimate (*SE*)	*Z*	*p*
(intercept)	1.39 (0.22)	6.46	<.01
Duration	-0.42 (0.04)	-10.72	<.01

Table 5. Summary of results of the logistic regression examining responses of younger listeners at the level of the first syllable

A separate linear mixed-effects model examining all participants' responses to the tokens for which the second syllable was manipulated, found no main effects or interactions, indicating that none of the listener groups were using the duration cue in perceiving Korean stress pairs. Figure 5 represents older and younger listeners' responses as a function of manipulated steps of three cues, showing the lack of effect of acoustic cues on the perception of the Korean stress pairs on the second syllable.

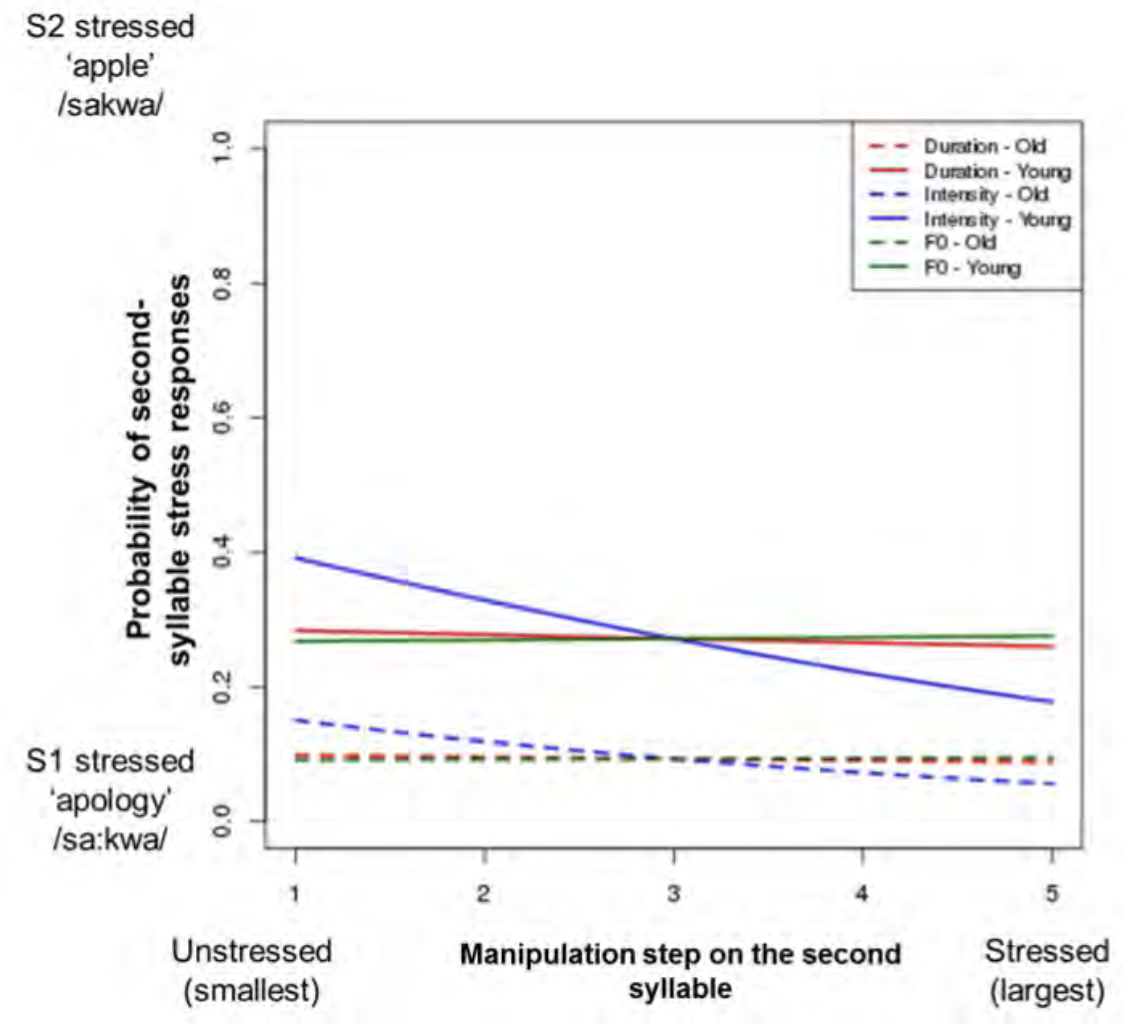

Figure 5. Probability of second-syllable stressed responses, /sakwa/, between the two listener groups. X-axis indicates the manipulated steps of each cue. 1 indicates that duration values were at the minimum endpoint, expressing first-syllable unstressed; and 5 indicates that duration values were at the maximum endpoint, expressing first-syllable stressed. Dotted lines indicate older Korean listeners' responses, and solid lines indicate younger Korean listeners' responses. Listeners' responses for each cue are illustrated with different colors: red, blue, green lines indicate listeners' responses for duration, intensity, and F0, respectively.

3.3. Discussion and Conclusion

Experiment 2 revealed several interesting facts. First, we found that neither older nor younger Korean listeners use intensity and F0 independently from duration in identifying prominence in Korean. Both older and younger Korean listeners used only the duration cue in identifying Korean stress pairs. In addition, we also found that younger Korean listeners still weight duration in their identification of phonemic vowel length, despite

the fact that contemporary Seoul Korean has almost completely lost the vowel length distinction. According to exemplar theory (Goldinger, 1996; Johnson, 1997), listeners mentally store variant details of speech sounds in their episodic memory, while mapping similar tokens into a single abstract category as a large cloud of exemplars. Highly similar tokens are tightly clustered and organized within a category, while dissimilar tokens are far apart and mapped onto two different categories. However, an exemplar-based model predicts that the production of categories may be deviant from perception, since the lexical entries that each listener stores are gathered from different speakers (Bybee, 2001; Johnson, 1997; Pierrehumbert, 2000, 2001). In the present case, the younger Korean speakers may have stored both long and short vowels as poor exemplars of a single category, reflecting the loss of the vowel length distinction in their production. However, the variations in the categorized percept (e.g., older speakers' contrastive production of the vowel length distinction) may have influenced younger listeners' perceptual sensitivity in identifying the vowel length contrasts. In other words, although younger listeners do not have two distinct vowel length categories, since younger listeners have collected both long vowels as in some of exemplars of /sa:kwa/ 'apology' and short vowels for /sakwa/ 'apple', the younger listeners may recognize this pair based on the duration of the vowel. This view, then, is compatible with our finding that younger Korean listeners still are sensitive to phonemic vowel length although the vowel length distinction has all but disappeared in their production.

Another interesting finding is that only cue manipulation in the first syllable influenced listeners' responses; a perceptual shift from the first-syllable stressed word to the second-syllable stressed word was not found when the second syllable was manipulated in terms of duration, intensity, or F0. Given that stress is defined as the relative difference in prominence between two syllables (Pierrehumbert, 1979; Beckman & Pierrehumbert, 1986), the fact that the cue manipulation in the second syllable did not trigger a change in the perception of the stress location indicates that Korean listeners only put perceptual weight on the first syllable. If Korean had stress, the changes in prominence in the second syllable should also induce the perceptual shift; however, the current study did not find such a pattern. Therefore, taking into account these two pieces of evidence provided by the current perception study, in addition to the acoustic evidence from Experiment 1, we conclude that Korean does not employ lexical stress, and that what has been claimed as stress pairs are actually vowel length contrasts.

Overall, the findings of the current study revealed that Korean does not employ word-level prominence, but that (so-called) Korean lexical stress pairs are only differentiated in terms of vowel length in the initial syllable.

References

Bates, D. (2005). Fitting linear mixed models in R. *R News*, *5*, 27-30.

Bates, D., & Maechler, M. (2010). Matrix: sparse and dense matrix classes and methods. *R package version 0.999375-43, URL http://cran. r-project. org/ package= Matrix.*

Beckman, M. E. (1986). *Stress and non-stress accent*. Vol. 7. Dordrecht: Foris..

Beckman, M. E. & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*(1), 255–309.

Boersma, P. & Weenink. D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.41, retrieved 6 August 2018 from http://www.praat.org/

Bybee, J. (2001). *Phonology and Language Use. Language*. Cambridge, UK: Cambridge University Press.

Cho, T. & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*(2), 155-190.

Cho, T., McQueen, J. M. & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*(2), 210-243.

Christophe, A., Peperkamp, S., Pallier, C., Block, E. & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language*, *51*(4), 523-547.

Chung, K., Chang, S., Choi, J., Nam, S., Lee, M., Chung, S., & Jee, S. (1996). *A study of Korean prosody and discourse for the development of speech synthesis/ recognition system*. Daejon, Korea: KAIST Artificial Intelligence Research Center.

Cooper, W. E., Eady, S. J. & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *The Journal of the Acoustical Society of America*, *77*(6), 2142-2156.

Coughlin, C. E. & Tremblay, A. (2012). Non-native listeners' delayed use of prosodic cues in speech segmentation. In Cox, F., Demuth, K., Lin, S., Miles, K., Palethorpe, S., Shaw, J. & Yuen, I. (Eds.), *Proceedings of the 14th Australasian Conference on Speech Science and Technology* (pp. 189-192). Macquarie University, Sydney, Australia.

de Jong, K., & Zawaydeh. B. A. (1999). Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics, 27*(1), 3-22.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, *26*(1), 138.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1(126), 126-152. Retrieved from http://las.sagepub.com/content/1/2/126.short

Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, *21*(4), 347-353.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(5), 1166-1183.

Heo, W. (1965). Kwuke Umwoonhak [Korean phonology]. Seoul: cengumsa.

International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge, U.K.: Cambridge University Press.

Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. In Johnson, K. & Mullennix, J. W. (Eds.), Talker variability in speech processing (pp. 143-166). New York, NY: Academic Press.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody*. Ph.D. dissertation, The Ohio State University.

Jun, S.-A. (1995). A phonetic study of stress in Korean. *The Journal of the Acoustical Society of America*, *98*(5), 2893.

Jun, S.-A. (1996). *The phonetics and phonology of Korean prosody: intonational phonology and prosodic structure*. New York: Garland Press.

Jun, S.-A. (1998). The Accentual Phrase in the Korean prosodic hierarchy. *Phonology*, *15*(2), 189-226.

Jun, S.-A. (2000). K-ToBI (Korean ToBI) Labelling Conventions - Version 3.1. *The Korean Journal of Speech Science*, *7*(1), 143-169.

Kang, K.-H. & Guion, S. G. (2008). Clear speech production of Korean stops: changing phonetic targets and enhancement strategies. *The Journal of the Acoustical Society of America*, *124*(6), 3909-3917.

Kenstowicz, M. & Park, C. (2006). Laryngeal features and tone in Kyungsang Korean: A phonetic study. *Studies in Phonetics, Phonology and Morphology*, *12*(2), 247-264.

Kim-Renaud, Y.-K. (1974). *Korean consonantal phonology*. Ph.D. dissertation, University of Hawaii.

Kim, H.-S., & Han, J.-I. (1998). Vowel length in Modern Korean: An acoustic analysis. In Park, B.-S. & Yoon, J. H. (Eds.), *Proceedings of the 14th International Conference on Korean Linguistics* (pp. 412-418). Seoul: Hanshin.

Kim, S.-H. (2001). *Hyenday kwuke-euy eumcang* [The Vowel Length of Modern Korean]. Seoul: Yeok-Lak.

Kim, S. (2004). *The role of prosodic phrasing in Korean word segmentation*. Ph.D. dissertation, University of California Los Angeles.

Kim, S. & Cho, T. (2009). The use of phrase-level prosodic information in lexical segmentation: evidence from word-spotting experiments in Korean. *The Journal of the Acoustical Society of America*, *125*(5), 3373-3386.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, *3*(3), 129-140.

Ko, E.-S. (2002). *The phonology and phonetics of word-level prosody and its interaction with phrase-level prosody: A study of Korean in comparison to English*. Ph.D. dissertation, University of Pennsylvania.

Ko E.-S. (2010). Stress and Long Vowels in Korean: Chicken or Egg First? *Japanese Korean Linguistics*, *17*, 377-390.

Ko, E.-S. (2013). A metrical theory of Korean word prosody. *Linguistic Review*, *30*(1), 79-115.

Koopmans-Van Beinum, F. J. (1980). *Vowel contrast reduction: an acoustic and perceptual study of Dutch vowels in various speech conditions*. Ph.D. dissertation, Universiteit van Amsterdam.

Ladefoged, P., Draper, M. H. & Whitteridge, D. (1958). Syllables and stress. Miscellanea. *Phonetica*, *3*, 1-14.

Lee, B.-G. (1978). *Kukeuy cangmoumhwa-wa posangseng* [Vowel lengthening and compensation in Korean]. Seoul: Kukehak 3. (Reprinted from Umwun Hyensange Isseseey Ceyak [Constraints in phonological phenomena], pp. 23-57, 1979, Seoul: Hanshin).

Lee, H. & Jongman, A. (2015). Acoustic evidence for diachronic sound change in Korean prosody: A comparative study of the Seoul and South Kyungsang dialects. *Journal of Phonetics*, *50*, 15-33.

Lee, H., Politzer-Ahles, S. & 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. *Journal of Phonetics*, *41*(2), 117-132.

Lehiste, I., & Peterson, G. E. (1959). Vowel amplitude and phonemic stress in American English. *Journal of the Acoustical Society of America*, *31*(4), 428-435.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, *32*(4), 451-454.

Llisterri, J., Machuca, M., de la Mota, C., Riera, M. & Río, A. (2003). The perception of lexical stress in Spanish. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2023-2026).

Magen, H. S. & Blumstein, S. E. (1993). Effects of speaking rate on the vowel length distinction in Korean. *Journal of Phonetics*, *21*, 387-410.

Oh, M. (1998). The prosodic analysis of intervocalic tense consonant lengthening in Korean. *Japanese Korean Linguistics*, *8*, 317-330.

Perkins, J. & Lee, S. J. (2010). Korean affricates and consonant - tone interaction. In *Proceedings of the 6th workshop on Altaic formal linguistics*. MIT Working Paper of Linguistics (pp. 277-286).

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation*. Ph.D. dissertation, Massachusetts Institute of Technology.

Pierrehumbert, J. B. (2000). What people know about sounds of language. *Studies in Linguistic Sciences*, *29*(2), 111-120.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J. & Hopper, P. (Eds.), *Frequency and the emergence of linguistic structure* (pp. 137-158). Amsterdam: John Benjamins.

Selkirk, E. (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Massachusetts: MIT Press.

Shport, I. A., & Redford, M. A. (2014). Lexical and phrasal prominence patterns in school-aged children's speech. *Journal of Child Language*, *41*(04), 890-912.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, *23*(2), 287-308.

Tremblay, A., Coughlin, C. E., Bahler, C. & Gaillard, S. (2012). Differential contribution of prosodic cues in the native and non-native segmentation of French speech. *Journal of Laboratory Phonology*, *3*(2), 385-423.

Tyler, M. D. & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, *126*(1), 367-376.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M. & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707-1717.

Wright, J. D. (2007). *The phonetic contrast of Korean obstruents*. Ph,D, dissertation, University of Pennsylvania.