## Contents

# I. Introduction

## 1. Introduction

It is well documented that learners' L1 languages affect their second language acquisition, when they produce and perceive the L2 language. Two representative theoretical frameworks, the Speech Learning Model (Flege, 1987, 1995) and the Perceptual Assimilation Model (Best, 1995; Best, McRoberts & Goodell, 2001) focus on this claim. Flege's Speech Learning Model (SLM) focuses on the role of the learner's native language in the second language learning process. This model states that a learner's ability to produce the segment of a non-native speech is affected by the perceptual similarity of target segments to those in their native language. If L1 and L2 segments are similar, the learner can approximate the L2 segment more quickly than less similar segments. However, the less similar, or "new", segments are acquired more accurately gradually, as the learner builds an L2 category from scratch. On the contrary, the more similar segments are strongly entangled with existing L1 categories and thus some aspects of the L1 categories are retained throughout the acquisition process. The Perceptual Assimilation Model (PAM) created by Best and her team postulates that the initial ability of a listener to discriminate sounds is influenced by the similarity in perception between the objects in an L2 and an L1. PAM research conducts a task that involves mapping similarity across languages where they focus on opposing pairs of segments in an unknown language using cross-language similarity mapping task. Based on the results of these advanced level tasks, it categorizes the specific pairs according to different situations. Using this classification, researchers predict the listener's ability to distinguish sounds at a lower level. According to this research, discrimination abilities are influenced by familiarity with sounds which are resembling those in the L2 as well as the extent to which the L1 categories of the listeners

prevent them from being able to differentiate between sounds in terms of their linguistic categories.

Complicating this picture considerably is the fact that the mapping between the languages is specifically perceptual, and perceptual on the basis of the learner. It is well-documented that the native language of a person strongly affects their ability to attend to different, potentially contrastive aspects of speech (Francis et al., 2008; Iverson et al., 2005; Toscano et al., 2010). Specifically, if a language does not use aspects of the signal for contrastive purposes, then native speakers will tend not to attend to this aspect of the signal, and if this aspect of the signal is used for contrastive purposes in the new language, these contrasts will be especially likely to yield 1-to-2 mapping and will prove exceptionally difficult. This thesis explores an example of this sort of case. Contemporary Korean vowels do not exhibit extensive variation in duration for the purposes of lexical contrast. While there were historical vowel-length contrasts in Korean, these have been lost in current Seoul varieties. By contrast, English is known to use timing differences in the vowel for various lexical purposes, specifically to mark voicing of a post-vocalic consonant, and as a direct correlate of the Tense-Lax distinction in the vowels themselves.

## 2. Durational Cues for Consonant Voicing

As suggested by these two models, it is expected that L2 language learners will have difficulty perceiving and producing L2 segments depending on the extent of the perceptual similarity between L2 segments and those in their native language. Among related research, a number of studies in acoustic phonetics have shown that the context in which a vowel occurs affects its acoustic properties. Particularly significant is the change in the duration of the vowel influenced by the voicing of the neighboring consonants. Previous research on vowel duration in the English language consistently indicates that a vowel followed by a voiced consonant is

longer than that followed by a voiceless consonant in native English productions (Denes, 1955; House & Fairbanks, 1953; House, 1961; Klatt, 1973; Peterson & Lehiste, 1960; Summers, 1987). Further, some perception studies have shown that vowel duration plays an important role for English native listeners to discriminate English final consonant voicing contrasts in their English perception (Denes, 1955; Port & Dalby, 1982; Raphael, 1972). On the other hand, studies of Korean indicate that the voicing-related vowel effect is weaker in Korean speakers' English production than that of English native speakers (Chen, 1970; Choi et al., 2016; Kim, 2011; Ko, 1998; Koo, 1998; Kwag, 2019; Oh, 2012; Peterson & Lehiste, 1960; Shin & Sa, 2005; Yoon, 2002).

In addition to the smaller vowel duration differences in Korean, there is some reason to believe that Korean militates against post-vocalic contrasts in general. Korean speakers' weaker voicing-related vowel effect seems to be due to Korean coda neutralization. More specifically, the three-way distinction in Korean consonants, which includes voiceless unaspirated consonants, voiceless aspirated consonants, and voiceless tensed consonants, is neutralized in the syllable coda, in favor of a voiceless unaspirated stop series [p, t, k]. Therefore, assuming L1 transfer effects on L2 English learning, it is expected that Korean speakers would have difficulty in distinguishing English consonant voicing contrasts specifically in the monosyllabic structure but not in the disyllabic structure in English. This is because underlying consonants remain unneutralized intervocalically in Korean (Kim & Jongman, 1996). More specifically, the stops /p, t, k/ are fully voiced between vowels, and thus they are pronounced as [b, d, g] (Major & Faudree, 1996). Further, Jun (1995) claimed that when a word final lenis stop is preceding a vowel, it becomes voiced in resyllabified onset position, which makes the following vowel duration longer. Those studies suggest that Korean speakers might transfer this dependency of voicing contrasts on different prosodic positions from their L1 while producing and perceiving English.

## 3. Durational Contrasts in English Vowels

Considering English tense/lax contrasts in vowel duration modulation, studies of English vowels suggest that English vowels appear to have inherent or intrinsic durational difference from one another in addition to differences in spectral characteristics (Bohn & Flege, 1992; Cebrian, 2006; Klatt, 1976). Perceptually, English native speakers primarily rely on the spectral properties to distinguish tense vowels from lax vowels in productions (Bohn & Flege, 1992; Yu, 2023). On the other hand, second language learners whose native language is without a tense/lax distinction have been shown to have more problems with the English tense/lax vowel pairs than those from languages with the distinction. Evidence that Korean adults have difficulty in producing and perceptually discriminating English tense-lax vowel contrasts has been provided by a number of previous works (Cho & Jeong, 2013; Flege, Bohn & Jang, 1997; Hong, 2012; Ingram & Park, 1997; Jun & Cowie, 1994; Kim, 2010; Kim, 2016; Kim, 2006; Koo, 2000; Lee, 2008; Lee & Cho, 2013; Lee & Cho, 2015; Lee & Lee, 2011; Lee et al., 2017; Tsukada et al., 2005; Yang, 1996; Yang, 2013). This seems to be due to the fact that there is no tense/lax contrast in Korean language, and more generally, no durationally based contrasts in vowels. Based on many previous studies, Korean learners of English as a second or foreign language must consciously learn the vowel length difference depending on either voicing effects or vowel quantity effects. It is then worth questioning whether Korean learners of English would show improvement on these two different contrasts that they do not have in their native language as they gain more L2 English experience.

## 4. Perception vs. Production

Learning a second language phonological system involves the disparate skills necessary to be a speaker and a hearer. The acquisition of production and perception requires learning these different skills, and when there is a difference between L1 and L2, it is expected to have

an influence on each skill in different ways due to the distinct nature of the skills related to each modality. Therefore, it is crucial to consider both production and perception skills in studies of the acquisition of phonological systems. Several researchers have claimed that production and perception are two distinct skills that cannot be compared (Cho & Jeong, 2013; Flege, Bohn & Jang, 1997; Kim, 2010; Lee, 2016). For example, Lee (2016) found that an asymmetry exists between perception and production of Korean language in her study, such that perceptual neutralization happens when laryngeal contrasts with evident acoustic differences are not perceived as distinct sounds. Korean speakers produced two different sounds referring to differences consistently detected through sound analysis, but they could not perceive these distinctions. Cho and Jeong (2013) examined English vowel production and perception by Korean speakers, and they also suggested that perception and production are not related, and that the development of perception and production happens independently. Flege, Bohn and Jang (1997) examined English vowel production and perception by non-native speakers differing in L2 English experience. They found that non-native speakers' accuracy in production and perception of the same English vowels improved together as they gained L2 English experience, but the exact nature of the relation between perception and production is uncertain.

As for the role of exposure to L2 English, the claim that amount of L2 experience affects both perception and production is generally agreed upon (Flege, Bohn & Jang, 1997; Ingram & Park, 1997; Jun & Cowie, 1994). Flege, Bohn and Jang (1997) reported that the non-native subjects with more English experience have been shown to have more accurate production and perception of English vowels than did the subjects with relatively less English experience. Also, particularly pertinent to the current examination of the tense-lax vowel contrast, Korean speakers with more English exposure produced English /ɪ/ more correctly than did the Korean speakers with less exposure to English. Ingram and Park (1997) reported that Korean learners

with relatively more English exposure performed better on Australian English contrast /e-æ/ than did the relatively less English experienced Korean learners.

**5. Structure of the Experiments in this Thesis**

This study aims to investigate whether Korean learners of English can produce the coda consonant contrasts in vowel duration and perceptually distinguish these same coda consonant contrasts. Also, it examines whether Korean coda neutralization affects Korean speakers' production and perception of these English coda consonant contrasts, by including an environment where neutralization happens in Korean (in final, coda position) and where it does not occur (in medial, inter-vocalic position). Further, it concurrently investigates whether intrinsic vowel durational differences in English tense-lax vowel pairs would be produced or perceptually contrasted by the same Korean learners of English. In addition, spectral properties which are also important acoustic information inherent in English target vowels were examined to see whether Korean speakers with more English exposure are able to make use of the cue more effectively than those with less English experience do.

Finally, an important feature of the design of these experiments is that they collect both production and perception data from the same participants. Thus, by cross-subject analysis, it examines whether there is a relationship between production and perception of two English contrasts by Korean learners of English. In the production experiment, preceding vowel duration was measured from the release of the stop burst to the constriction of the following stop or fricative. Also, the first two formant frequencies (F1 and F2) of target vowels were measured at the temporal midpoint of each vowel. In an ABX discrimination task, participants were asked to identify the last stimulus as either the first stimulus (First word) or the second stimulus (Second word) by clicking either "First word" or "Second word" on the computer screen. In a separate identification task, participants were asked to identify the correct

underlying form in each stimulus from the four alternatives by clicking appropriate word on the computer screen.

## 6. Structure of this Thesis

The organization of the dissertation is as follows. The research topic and an overview of the study are introduced in Chapter 1. Prior studies on the production of the final consonant voicing along with tense/lax vowel distinction are reviewed and the production experiment that includes the methodology, acoustic analyses and results are discussed in Chapter 2. Previous literature related to English consonant and vowel contrasts is reviewed and perception experiments that investigate Korean listeners' discrimination and identification of both English consonant contrasts and vowel contrasts are presented in Chapter 3. Concluding remarks that summarize the main findings of the current study, their implications for our understanding of second language acquisition are discussed. Additionally, suggestions for future research directions emphasizing potential areas for further investigation are proposed in Chapter 4.

## II. Production

### 1. Introduction

*Vowel duration modulation by consonant voicing*

Studies of vowel duration in the English language consistently indicate that a vowel followed by a voiced consonant is longer than that followed by a voiceless consonant in native English productions (Denes, 1955; House & Fairbanks, 1953; House, 1961; Klatt, 1973; Peterson & Lehiste, 1960; Summers, 1987). On the contrary, studies of other languages and some studies of ESL learners illustrate that learners from various language backgrounds do not show this large English-like pattern in vowel duration distinction followed by a different consonant voicing. More specifically, non-native learners of English showed a much smaller contrast in vowel duration due to consonant voicing than English native speakers (de Jong & Zawaydeh, 2002; Flege & Hillenbrand, 1986; Flege & Port, 1981; Mack, 1982; Mitleb, 1984a). For example, Mack (1982) examined vowel lengthening depending on the following stop voicing in English production spoken by French and English speakers. Mack found that French-English bilingual speakers showed weaker voicing effects, which were more like those of French monolinguals than those of English native speakers. More specifically, French-English bilingual speakers' vowel duration difference depending on the following coda stop voicing was smaller than that of English native speakers. Mack concluded that French speakers learning English did not utilize the English phonetic rule to distinguish voicing contrasts in their English production. Similarly in Arabic, Mitleb (1984a) claimed that voicing does not specify relative vowel durations in Arabic whereas it does in English. In the subsequent study of de Jong and Zawaydeh (2002), they examined voicing and quantity effects on vowel duration and quality in Arabic speakers' production in Arabic. Long and short vowels were produced before coronal consonants (/t/ and /d/) differing in voicing in the tri-syllabic words. They found that Arabic speakers showed a small difference between vowel duration before voiced consonant and that

before voiceless consonant, which indicates that temporal cues (vowel duration) were not effectively used for the coda voicing contrast. On the other hand, temporal cues were systematically used for vowel quantity contrasts whereby long vowels were significantly longer than short vowels duration-wise, which are phonemically specified in Arabic. As for vowel quality, vowel lengthening due to quantity effects was accompanied by F1 increase whereas that due to coda voicing was accompanied by F1 decrease.

The current study examines the durational patterns in English as acquired by speakers of Korean. Previous research on Korean indicates that the voicing-related vowel effect was weaker in Korean speakers' English production than that of English native speakers (Chen, 1970; Choi et al., 2016; Kim, 2011; Ko, 1998; Koo, 1998; Kwag, 2019; Oh, 2012; Peterson & Lehiste, 1960; Shin & Sa, 2005; Yoon, 2002). For example, Ko (1998) examined vowel lengthening depending on the following stop voicing in English production spoken by Korean speakers and English native speakers. The English front vowel /æ/ was produced before coda stops differing in voicing (/p, t, k/ and /b, d, g/) in monosyllabic words. Ko reported that vowel duration in the voiced context was significantly longer than that in the voiceless context. However, the ratio of vowel duration in the voiceless context to that in the voiced context was 82-98% for Korean speakers whereas that of English native speakers was 56-65%. This indicates that voicing effects on vowel duration were much smaller for Korean speakers than English native speakers. Yoon (2002) compared Korean speakers from Chonnam province with English native speakers in terms of vowel duration depending on either intrinsic vowel quality or following consonant voicing in English production. English vowels /i, ɪ, u, ʊ/ were used to examine durational differences between tense and lax vowels. As for voicing effects, English vowels /i/ and /u/ were produced before coda stops differing in voicing (/t/ and /d/) in monosyllabic words. As a result, Korean speakers showed similar patterns to English native speakers in terms of durational difference depending on vowel quality (tense vs. lax). More

specifically, Korean speakers and English native speakers produced tense vowels 1.5 times longer than lax vowels. On the contrary, Korean speakers showed weaker voicing effects on vowel duration than did English native speakers. More specifically, Korean speakers produced vowel duration in the voiced context 1.5 times longer than that in the voiceless context whereas English native speakers produced vowel duration before voiced consonants 1.8 times longer than that before voiceless consonants. Shin and Sa (2005) compared Korean speakers and English bilingual speakers with English native speakers in terms of voicing effects on vowel duration in English production. English vowels /i, ɪ, ɛ, æ, ʌ, ə/ were produced before either coda fricatives (/f, ɵ, s, ʃ/ and /v, ð, z, ʒ/) or coda affricates (/ʧ/ and /ʤ/) differing in voicing in monosyllabic words. Korean speakers showed weaker voicing effects than did English bilinguals and English native speakers. More specifically, the ratio of vowel duration followed by voiceless consonants to that followed by voiced consonants was 81% for Korean speakers, 69% for English bilingual speakers and 67% for English native speakers. Kim (2011) investigated vowel lengthening depending on the following stop voicing in English production spoken by Korean speakers. English front vowels /i, ɪ, ɛ, æ/ were produced before coda stops differing in voicing (/p, t, k/ and /b, d, g/) in monosyllabic words. According to Kim, Korean speakers showed weaker voicing effects than did English native speakers. More specifically, ratio of vowel duration in the voiceless context to that in the voiced context was 86% for Korean speakers, which was larger than that of English native speakers that was 61% (Chen, 1960) or 2/3 (Peterson & Lehiste, 1960). This indicates that Korean speakers' vowel duration difference depending on the following coda stop voicing was smaller than that of English native speakers. Based on these studies, then, we would expect Korean learners of English to reflect small durational modulation effects, similar to what is found, e.g., in the earlier studies of French L2 speakers.

There is an additional complication in the Korean case, however. In the case of the Korean consonant inventory, the three-way distinction in Korean consonants, which includes voiceless unaspirated consonants /p, t, k, ʧ/, voiceless aspirated consonants /$p^h$, $t^h$, $k^h$, $ʧ^h$/, and voiceless tensed consonants /p', t', k', ʧ', s'/, are neutralized in the syllable coda, in favor of a voiceless unaspirated stop series [p, t, k] as indicated in (1) (Sohn, 1999). Korean has no words with /p', t', ʧ'/ before a consonant or before a word boundary. As for coda neutralization in Korean, it affects three aspects of the consonantal system, including laryngeal, manner and palatal contrasts. First, laryngeal neutralization merges laryngeal contrasts (i.e., /p vs. $p^h$/, /t vs. $t^h$/, /k vs. $k^h$ vs. k'/) into lenis obstruents [p], [t], [k]. Second, neutralization removes manner of articulation differences (e.g., fricative – stop, affricate – stop, fricative – affricate), also merging them into lenis stop [t]. Lastly, palatal neutralization involves neutralization of laryngeal contrast between palatal affricates /ʧ vs. $ʧ^h$/ also into lenis stop [t]. In Korean, no fricatives can take place in coda position possibly because they are inherently released (Davis & Cho, 2006). Further, in final position, neutralization rules have an influence on all obstruents in Korean, and thus all laryngeal contrasts disappear and fricatives manifest as stops (de Jong & Cho, 2012). The effect of all this is that laryngeal contrasts do not appear in post-vocalic environments, and thus, even though Korean, like English, has laryngeal contrasts, they should not be a factor in vowel duration. Therefore, assuming transfer effects, it is expected that Korean coda neutralization would have an influence on vowel duration followed by different consonant voicing especially in the monosyllabic structure in English.

as /ɪ/ in 37%, /ɪ/ productions were heard as /i/ in 17% and /ɛ/ productions were misheard as /æ/ in 66% to a native English listener. Also, Korean speakers' vowel formant frequency contrast between /ɛ/ and /æ/ was much smaller than that of English native speakers.

Choi et al. (2016) compared Korean groups differing in English proficiency (advanced vs. intermediate groups) with an English native group in terms of use of the phonetic cues to English segment contrast. English vowels /ɛ/ and /æ/ were produced in monosyllabic words. The advanced Korean speaker group exploited not only the temporal cues (vowel duration) but also the spectral cues (F1, F2 formant frequencies) for English vowel contrast whereas the intermediate Korean speaker group failed to utilize the spectral cues to distinguish tense/lax contrast in their English production. Jun and Cowie (1994) similarly compared Korean speaker groups differing in L2 English experience in their English vowel production including /i, ɪ, u, ʊ, ʌ/. Korean speakers with more English experience showed similar patterns to those of English native speakers in terms of the formant difference between tense and lax vowels. On the contrary, Korean speakers with less English experience showed a much smaller difference between tense and lax vowels than did English native speakers. Lee and Lee (2011) examined English vowel production by Korean elementary school students differing in English proficiency level. English vowels including /i, ɪ, u, ʊ/ were produced by Korean speakers and these vowels were assessed by English native listeners. They reported that Korean speakers' English lax vowels were more intelligible to English native listeners than the corresponding tense vowels. This was due to the fact that Korean speakers' English tense and lax vowels were both more similar to English native speakers' lax vowels in terms of duration, which was the cue used by English native listeners to distinguish English tense vowels from the corresponding lax vowels spoken by Korean speakers. Koo (2000) examined Korean speakers' English vowel production including /i, ɪ, ɛ, æ, ʌ, ɝ, u, ʊ, ɔ, ɑ/ and the formant values of each vowel were measured. Koo reported that Korean speakers hardly showed a distinctive difference between

tense and lax vowels, especially for /i/-/ɪ/, /ɛ/-/æ/, and /u/-/ʊ/ vowel pairs in terms of formant values whereas tense-lax vowel pairs spoken by English native speakers were quite distinctive in the vowel space. Lee and Cho (2013) compared Korean speakers with English native speakers in terms of vowel duration and formant values of English vowel production including /i, ɪ, eɪ, ɛ, æ, ɑ, ʌ, ɔ, oʊ, ʊ, u/ in monosyllabic words. They found that there was no credible difference between tense and lax vowels in Korean speakers' English production, especially for /i/-/ɪ/, /ɛ/-/æ/, and /u/-/ʊ/ vowel pairs in terms of formant frequencies (F1, F2). On the contrary, Korean speakers showed a quite distinctive difference between tense and lax vowel pairs in terms of vowel duration, which indicates that Korean speakers solely relied on temporal cues to distinguish English tense-lax vowel contrasts in the production. On the contrary, English native speakers showed a credible difference between tense and lax vowels in terms of vowel duration as well as formant values.

The difficulties that Korean learners of English have may be accounted for by Best's (1993, 1995) model that when contrastive L2 vowels are mapped onto a single L1 vowel, it will be relatively difficult for L2 learners to distinguish them; whereas discrimination of contrastive L2 vowels will be more accurate when they are identified as two different L1 vowel categories. For example, Korean vowel length is not reliably contrastive in contemporary varieties while in the speech of earlier generations vowel length was distinctive, predominantly in the first syllable of words, such as *mal* 'horse' vs. *maal* 'speech', *nun* 'eye' vs. *nuun* 'snow', which are well-known minimal pairs (Ahn, 1998; Lee & Ramsey, 2000; Park, 1994). Park (1994) reported that 21 younger metro speakers did not distinguish long vowels from short vowels in a questionnaire. In their responses, 56.8% of long vowels were marked as short and 28.2% of short vowels were marked as long. Also, the recent merger of /e/ and /ɛ/ in Korean may have caused the difficulty for Korean speakers to distinguish the English vowel pairs including /ɛ/-/æ/ in terms of vowel duration as well as vowel quality. This is due to the fact that

there is only one phoneme in the region of mid-front vowel in Korean vowel space due to the merger of /e/ and /ɛ/ (Flege, Bohn & Jang, 1997; Ingram & Park, 1997; Kim, 2010; Lee & Ramsey, 2000). The inventory of Korean vowel phonemes is represented in Figure 1 (Kim, 1999).

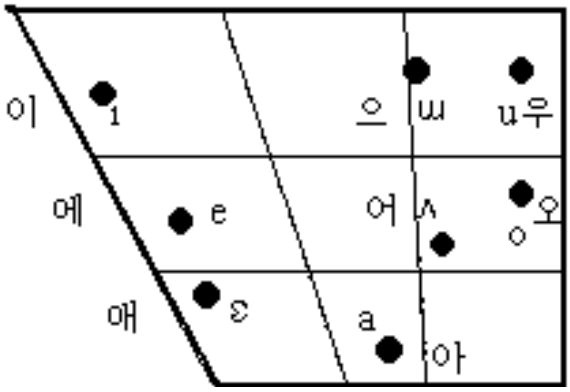

Figure 1. Korean vowel phonemes

*The current study*

This study aims to investigate whether Korean learners of English could produce the coda consonant contrasts in vowel duration. Also, it concurrently investigates whether intrinsic vowel durational differences in English would be shown by the same Korean learners of English. In addition, spectral properties which are also important acoustic information inherent in English target vowels were examined to see whether Korean speakers with more English exposure are able to make use of the cue more effectively than those with less English experience do. The consonants chosen for the experiment in this study consist of sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/) and fricatives (/f, v/, /s, z/). The vowels selected for the experiment include English tense/lax vowel pairs /i, ɪ, e, ɛ, u, ʊ/ which cause high confusion to Korean learners of English because Korean has no tense/lax distinction.

McAllister, Flege, and Piske (2002) claimed that L2 phonetic features will be more difficult to acquire when they are not used as cues for phonological contrasts in learners' L1. Furthermore, according to Bohn's (1995) Desensitization Hypothesis, L2 learners' sensitivity to spectral differences will be lesser when they have only one vowel in the L1. Therefore, it is

expected that Korean learners of English would have difficulty in discriminating English tense/lax pairs since each contrast may fall within a single category in Korean. For example, English /i/ vs. /ɪ/ would be assimilated to Korean single category /i/, /ɛ/ vs. /æ/ would be assimilated to a merged /e-ɛ/ category in Korean and /u/ vs. /ʊ/ would be assimilated to Korean vowel /u/.

In this study, we are examining two different durational effects which include voicing effects (voiced vs. voiceless) and vowel quantity effects (tense vs. lax) and L1 transfer effects on them. Their durational mechanisms are different and thus the L1 transfer effect might be expected to be different from one another. First, vowel quantity effects cover through a whole part of the vowel whereas voicing effects have an influence on the end of the vowel (Summers, 1987). Second, voicing effects are redundant and allophonic whereas vowel quantity effects are phonemic. Third, vowel quantity effects have correlated cues which are vowel quality effects whereas voicing effects are very small and different from the tense/lax differences. Fourth, voicing effects are due to similar L2 phones to Korean learners whereas vowel quantity effects are new L2 phones to Korean learners. According to Flege (1987), similar L2 phones have acoustically different manner from phones in L1 whereas new L2 phones have no counterpart in L1. More specifically, English stops show a 2-way laryngeal contrast whereas Korean stops exhibit a 3-way laryngeal contrast, which is reduced to plain stops in the postvocalic final position due to neutralization processes. As for vowel quantity effects as new L2 phones, English tense/lax contrasts do not have counterpart in Korean, and thus English lax vowels /ɪ, ʊ, ɛ/ are new phones for Korean speakers.

Based on the previous research, the hypotheses for this study are as follows:

**Hypothesis 1**: In production, due to the lack of a coda distinction in obstruent voicing in Korean, Korean L1 speakers would show an attenuated difference due to following obstruent voicing in monosyllabic words. Also, coda neutralization rules in their L1 might be a factor in

their English productions. However, they would show a larger vowel durational difference in disyllabic words since the Korean contrast is not neutralized in medial position.

**Hypothesis 2**: Korean subjects would exhibit different patterns with respect to laryngeal neutralization and manner neutralization in their English production. This is because Lee (2016) reported that laryngeal neutralization was phonetically incomplete in her study whereas manner neutralization was complete in the study of Kim and Jongman (1996).

**Hypothesis 3**: Due to the lack of an intrinsic vowel quantity distinction in Korean, Korean speakers would have difficulty in production of English tense/lax vowel contrasts.

**Hypothesis 4**: Korean subjects would exhibit different L1 transfer effects on two different durational effects which include voicing effects (voiced vs. voiceless) and vowel quantity effects (tense vs. lax).

## 2. Methods

### 2.1 Subjects

32 native Korean speakers who were from various regions in Korea participated in the experiment. 16 of them were recruited from Korea and had less than a week of exposure to an English-speaking country (4 males, 12 females, age: 30 – 40, mean: 36 years old). The other 16 were recruited from Indiana University and they had more than three years of exposure to an English dominant environment (6 males, 10 females, age: 30 – 37, mean: 32 years old). The subjects' average duration of studying English at school in Korea was around 10 years. The subjects consisted of 10 male and 22 female speakers. Two English native speakers were recruited from Indiana University. The speakers consisted of 1 male and 1 female speaker who were from the American Midwest and their age was in their twenties.

## 2.2 Stimuli

The stimuli for production tasks consist of the monosyllabic structure /bVC/ and disyllabic structure /bVCa/ using English vowels including /i, ɪ, e, ɛ, u, ʊ/ and sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/) and fricatives (/f, v/, /s, z/) as a coda for each vowel. Many of the forms constituted nonce words. The same onset consonant /b/ was used for all the items. The monosyllabic and disyllabic tokens were spoken in a carrier sentence, which is "say _ soon". In the disyllabic structure, the final consonant precedes a vowel /a/. This feature indicates that the consonant which follows the vowel is probably syllabified with the following vowel rather than with the target vowel. Chen's (1970) study of English included both cross-syllable and syllable internal voicing effects and both appeared to be approximately of the same magnitude. Therefore, we expect that the consonant affects the duration of the preceding vowel in both monosyllabic and disyllabic words which are spoken by native English speakers. The subjects were given the target nonce words that they were asked to produce and also were given English samples of real words with the same vowel at the right side of the target words that they could refer to. For example, at the right side of the target word /bip/ which was transcribed in IPA, there were real English words such as *'leaf', 'feed', 'eat', 'seat'*, and '*sheep*' which they could refer to.

## 2.3 Procedure

Each subject was recorded digitally using *Audacity* implemented on a laptop in a sound-dampened room, individually. 16 Korean speakers who were recruited from Korea were recorded in a quiet room in the public study room near their place or office and the other 16 who were recruited from Indiana University were recorded in a sound-dampened booth at the Indiana University Phonetics Lab. The target English nonce words including 60 monosyllabic words and 60 disyllabic words (2 structures x 6 vowels x 10 consonants = 120 tokens) were

randomized and presented to the subjects in a reading list. The participants were asked to read the target words in a carrier sentence, “say _ soon” once.

### 2.4 Measurements & Analysis

Preceding vowel duration was measured from the release of the stop burst to the constriction of the following stop or fricative using *Praat*. Also, the first two formant frequencies (F1 and F2) of target vowels were measured at the temporal midpoint of each vowel.

First, repeated measures three-way ANOVAs were conducted using the R statistical package on vowel durations. Statistical analyses were conducted separately for each durational effect. First, as for voicing effects on vowel durations, the vowel durations were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) and final consonant voicing (voiced vs. voiceless) were within-subject factors. Second, as for manner effects on vowel durations, the vowel durations were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) and manner (fricatives vs. stops) were within-subject factors. Third, as for vowel identity effects on vowel durations, the vowel durations were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) and vowel identity (tense vs. lax) were within-subject factors.

Second, repeated measures three-way ANOVAs were performed using R statistical package on vowel formant frequencies (F1 and F2). Statistical analyses were conducted separately for each durational effect. First, as for voicing effects on vowel formant frequencies (F1 and F2), the vowel formant frequencies (F1 and F2) were submitted to a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) and final consonant voicing (voiced vs. voiceless) were within-subject factors. If there was a significant interaction effect between independent variables, a follow-up analysis was performed by a series of ANOVAs on each dependent variable. Second, as for vowel identity effects on vowel formant frequencies, the first two formant frequencies (F1 and F2) of target vowels were submitted to a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) and vowel identity (tense vs. lax) were within-subject factors. If there was a significant interaction effect between independent variables, a follow-up analysis was performed by a series of ANOVAs on each dependent variable.

## 3. Results

### 3.1 Vowel duration

#### 3.1.1 Voicing contrasts (Voiced vs. Voiceless)

In this study, vowel durations followed by voiced consonants and voiceless consonants in different prosodic positions were examined to see whether the Korean laryngeal neutralization rules affected Korean speakers' English productions. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more

experienced Korean group vs. English native control group) was a between-subject factor and final consonant voicing (voiced vs. voiceless) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for vowel duration. This analysis indicated that there was a significant main effect of voicing contrast [$F(1, 31) = 203.151$, $p < 0.001$, $\eta^2 = 0.02$, [0.01, 1.00]] where vowel length in the voiced context was longer than that in the voiceless context. There was a significant main effect of prosodic structure [$F(1, 31) = 31.173$, $p < 0.001$, $\eta^2 = 1.22e\text{-}03$, [0.00, 1.00]] where vowel length in the monosyllabic structure was longer than that in the disyllabic structure. However, there was no significant main effect of group [$F(2, 31) = 0.17$, $p > 0.05$, $\eta^2 = 1.00$, [0.00, 1.00]] showing that vowel length of the inexperienced Korean group was not significantly different from that of the experienced Korean group. Previously, it was expected that vowel length of the inexperienced Korean speakers would be longer than that of the experienced Korean speakers but there was not meaningful difference between them.

Prosodic structure interacted with group [$F(2, 31) = 5.371$, $p < 0.05$, $\eta^2 = 8.13e\text{-}03$, [0.00, 1.00]] showing that Korean speakers did not show as extreme of a structure effect on vowel durations as English native speakers did. However, there was no such interaction effect between voicing contrast and group [$F(2, 31) = 1.214$, $p > 0.05$, $\eta^2 = 1.12e\text{-}03$, [0.00, 1.00]] and no interaction effect between voicing contrast and prosodic structure [$F(1, 31) = 3.166$, $p > 0.05$, $\eta^2 = 4.92e\text{-}05$, [0.00, 1.00]]. On the other hand, there was a significant three-way interaction effect of voicing contrast x prosodic structure x group [$F(2, 31) = 26.587$, $p < 0.001$, $\eta^2 = 4.09e\text{-}03$, [0.00, 1.00]] where there was a larger voicing effect for Korean speaker groups than English native group in the disyllabic structure whereas the opposite pattern occurred in the monosyllabic structure.

A series of paired t-tests was run on different voicing conditions. According to the t-tests, mean vowel durations in the voiced context were significantly longer than those in the

voiceless context in the monosyllabic structure for all the groups. The differences remained statistically significant after adjustment for multiple comparisons using a Bonferroni correction for all the groups (the inexperienced Korean group: $t$ = 16.77, $df$ = 479, $p$ < 0.001; the experienced Korean group: $t$ = 15.226, $df$ = 479, $p$ < 0.001; the English native control group: $t$ = 24.804, $df$ = 59, $p$ < 0.001). In addition, in the disyllabic structure, all the groups showed a credible difference between vowels preceding voiced consonants and those followed by voiceless consonants. The differences remained statistically significant after adjustment for multiple comparisons using a Bonferroni correction for all the groups (the inexperienced Korean group: $t$ = 16.012, $df$ = 479, $p$ < 0.001; the experienced Korean group: $t$ = 19.171, $df$ = 479, $p$ < 0.001; the English native control group: $t$ = 12.23, $df$ = 59, $p$ < 0.001).

The voicing effect was obvious for all the groups not only in the monosyllabic structure but also in the disyllabic structure. However, the magnitude of difference between vowel length preceding voiced consonants and that preceding voiceless consonant was different depending on the speaker group and prosodic structure. More specifically, Korean speaker groups (the inexperienced Korean group: 38.9 ms; the experienced Korean group: 35.1 ms) showed a smaller voicing effect than English native speakers (88 ms) in the monosyllabic structure. On the contrary, Korean speaker groups (the inexperienced Korean group: 37.5 ms; the experienced Korean group: 36.7 ms) showed a larger voicing effect than English native speakers (20.3 ms) in the disyllabic structure. These results are plotted in the left panel of Figure 3. There, we can see the range of the vowel length difference depending on the different voicing contexts has a different pattern based on the structure for each group. More specifically, Korean speakers showed a similar range of vowel length difference affected by the following consonant voicing in both monosyllabic and disyllabic structures whereas the English native control group showed a much smaller range of vowel length difference depending on the voicing context in the disyllabic structure than that in the monosyllabic structure as in the left panel of Figure 3.

The relationship between voicing effect size in monosyllabic words and that in disyllabic words was examined to see whether the Korean coda laryngeal neutralization rules affected Korean speakers' English productions. So, for each participant, the average size of the vowel duration difference for voicing in different prosodic position was calculated and the voicing effect size in intervocalic position was plotted against in final position. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 2 below, there is a strong positive association between vowel length differences depending on the final consonant voicing in monosyllabic words and those in disyllabic words for Korean speaker groups. More specifically, the inexperienced Korean speakers showed a strong ($r^2 = 0.58$), and statistically significant ($p < 0.001$) positive association between two variables. Also, the experienced Korean speakers showed a strong ($r^2 = 0.48$), and statistically significant ($p < 0.05$) positive association between two variables. This strong correlation across the two variables demonstrates that the Korean coda laryngeal neutralization rules were not an important factor in Korean speakers' English production, and it is more about just connecting voicing and vowel duration across the board.

To summarize, there was a voicing effect as we expected where vowel length in the voiced context was longer than that in the voiceless context. Korean speakers showed laryngeal effects not only in non-final position but in final position. There was a larger voicing effect for Korean speaker groups than the English native group in the disyllabic structure whereas the opposite pattern occurred in the monosyllabic structure. However, the strong correlation between voicing effect size in monosyllabic words and that in disyllabic words by Korean speakers demonstrates that the Korean coda laryngeal neutralization rules do not have a considerable impact on Korean speakers' English production. Korean speakers with more

English exposure did not show larger laryngeal effects compared to those with less English exposure.

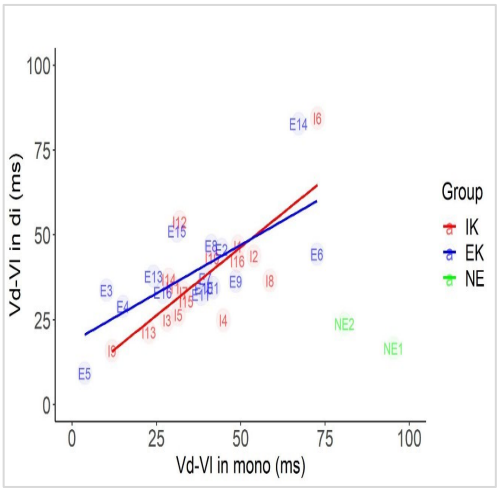


Figure 2. Correlation between vowel duration differences depending on the following consonant voicing in two prosodic structures (monosyllabic vs. disyllabic) in Korean speakers and English speakers' English productions.

### 3.1.2 Manner contrasts (Fricative vs. Stop)

In this study, vowel durations followed by fricatives and stops in different prosodic positions were examined to see whether the Korean manner neutralization rule affected Korean speakers' English productions. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and manner (fricatives vs. stops) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for vowel duration. According to the analysis, there was a significant main effect of manner [$F(1, 31) = 86.69$, $p < 0.001$, $\eta^2 = 4.39e\text{-}03$, [0.00, 1.00]] where vowel length preceding fricatives was longer than that preceding stops. There was a significant main effect of prosodic structure [$F(1, 31) = 32.50$, $p < 0.001$, $\eta^2 = 1.08e\text{-}03$, [0.00, 1.00]] where vowel length in the monosyllabic structure was longer than that in the disyllabic structure. However, there was no significant main effect of group [$F(2, 31) = 0.212$, $p > 0.05$, $\eta^2 = 1.00$, [0.001.00]] showing

that vowel length spoken by the inexperienced Korean group was not significantly different from that of the experienced Korean group.

Prosodic structure interacted with group [$F(2, 31) = 5.83$, $p < 0.05$, $\eta^2 = 7.99e\text{-}03$, [0.00, 1.00]] showing that Korean speakers did not show as extreme of a structure effect on vowel durations as English native speakers did. However, there was no interaction effect between manner and prosodic structure [$F(1, 31) = 3.066$, $p > 0.05$, $\eta^2 = 1.89e\text{-}05$, [0.00, 1.00]], nor between manner and group [$F(2, 31) = 0.13$, $p > 0.05$, $\eta^2 = 2.65e\text{-}04$, [0.00, 1.00]] where durational difference between vowel length preceding fricatives and that preceding stops was not significantly different between experienced and inexperienced Korean groups. There also was no interaction effect between manner, prosodic structure and group [$F(2, 31) = 0.934$, $p > 0.05$, $\eta^2 = 1.62e\text{-}04$, [0.00, 1.00]].

A series of paired t-tests was run on different manner of consonant conditions. According to the t-tests, mean vowel durations preceding fricatives were significantly longer than those preceding stops in the monosyllabic structure for all the groups. The differences remained statistically significant after adjustment for multiple comparisons using a Bonferroni correction for all the groups (the inexperienced Korean group: $t = 8.6744$, $df = 383$, $p < 0.001$; the experienced Korean group: $t = 8.9642$, $df = 383$, $p < 0.001$; the English native control group: $t = 9.6154$, $df = 47$, $p < 0.001$). However, in the disyllabic structure, Korean speaker groups showed a credible difference between vowels preceding fricatives and those preceding stops whereas English native speakers did not exhibit a credible difference between them. After correcting for multiple testing using the Bonferroni correction, there was a significant difference only for Korean speaker groups (the inexperienced Korean group: $t = 6.0681$, $df = 383$, $p < 0.001$; the experienced Korean group: $t = 7.2798$, $df = 383$, $p < 0.001$; the English native control group: $t = 5.1252$, $df = 47$, $p > 0.05$).

Previously, it was expected that Korean speakers would show a significant difference between vowel length preceding fricatives and that preceding stops only in the disyllabic structure and not in the monosyllabic structure because manner neutralization only applies in the postvocalic coda position. However, the results did not follow this expectation such that vowel durations followed by fricatives were significantly longer than those followed by stops not only in the monosyllabic structure but also in the disyllabic structure for all the Korean speaker groups. These results and interactions are plotted in the middle panel of Figure 3. As we can see in the middle panel of Figure 3, the vowel length of the inexperienced Korean group was almost the same as that of the experienced Korean group in the monosyllabic structure and there was only a marginal durational difference between inexperienced and experienced Korean groups in the disyllabic structure.

To summarize, there were manner effects where vowel length preceding fricatives was longer than that preceding stops. Korean speakers showed manner effects not only in non-final position but in final position. There were no larger manner effects for Korean subjects with more English exposure than those with less English exposure.

### 3.1.3 Vowel identity (Tense vs. Lax)

In this study, vowel durations depending on vowel identity were examined to see whether Korean speakers would distinguish tense vowels from lax vowels by duration in their English productions. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and vowel identity (tense vs. lax) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for vowel duration. This analysis indicated that there was a significant main effect of vowel identity [$F(1, 31) = 124.275$, $p < 0.001$, $\eta^2 = 0.02$, [0.01, 1.00]] where tense vowels were longer than lax vowels.

There was a significant main effect of prosodic structure [$F(1, 31) = 31.173$, $p < 0.001$, $\eta^2 =$ 1.17e-03, [0.00, 1.00]] where vowel length in the monosyllabic structure was longer than that in the disyllabic structure. However, there was no main effect of group [$F(2, 31) = 0.17$, $p >$ 0.05, $\eta^2 = 1.00$, [0.00, 1.00]] showing that vowel length of the inexperienced Korean group was not significantly different from that of the experienced Korean group.

Prosodic structure interacted with group [$F(2, 31) = 5.371$, $p < 0.05$, $\eta^2 = 7.82$e-03, [0.00, 1.00]] showing that Korean speakers did not show as extreme of a structure effect on vowel durations as English native speakers did. However, there was no interaction effect between vowel identity and prosodic structure [$F(1, 31) = 0.166$, $p > 0.05$, $\eta^2 = 3.01$e-05, [0.00, 1.00]]. No interaction effect occurred between vowel identity and group [$F(2, 31) = 1.719$, $p >$ 0.05, $\eta^2 = 2.48$e-03, [0.00, 1.00]] showing that the difference between tense vowels and lax vowels was not larger for Korean speakers with more English exposure than those with less English exposure. It was expected that the experienced Korean speakers would show a larger vowel identity effect than the inexperienced Korean speakers, but this was not true. There was no interaction effect between vowel identity, prosodic structure and group [$F(2, 31) = 1.213$, $p > 0.05$, $\eta^2 = 5.52$e-04, [0.00, 1.00]].

A series of paired t-tests was run on tense/lax pairs of English vowels. According to the t-tests, mean vowel durations for tense vowels were significantly longer than those for lax vowels in the monosyllabic structure for all the groups. The differences remained statistically significant after adjustment for multiple comparisons using a Bonferroni correction for all the groups (the inexperienced Korean group: $t = 11.529$, $df = 479$, $p < 0.001$; the experienced Korean group: $t = 9.9875$, $df = 479$, $p < 0.001$; the English native control group: $t = 8.7775$, $df = 59$, $p < 0.001$). In the disyllabic structure, however, Korean speaker groups showed a credible durational difference between tense vowels and lax vowels whereas English native speakers did not show a credible difference. The differences remained statistically significant after

adjustment for multiple comparisons using a Bonferroni correction for Korean speaker groups (the inexperienced Korean group: $t = 11.942$, $df = 479$, $p < 0.001$; the experienced Korean group: $t = 11.038$, $df = 479$, $p < 0.001$; the English native control group: $t = 8.9448$, $df = 59$, $p > 0.05$). Tense vowels were credibly longer in duration than lax vowels for all the speakers in the monosyllabic structure but in the disyllabic structure only Korean speakers showed a credible difference between tense and lax vowels. These results are plotted in the right panel of Figure 3.

To summarize, tense vowels were longer than lax vowels and vowel length in the monosyllabic structure was longer than that in the disyllabic structure for all the speakers. There was no larger vowel identity effect for Korean subjects with more English experience than those with less English experience.

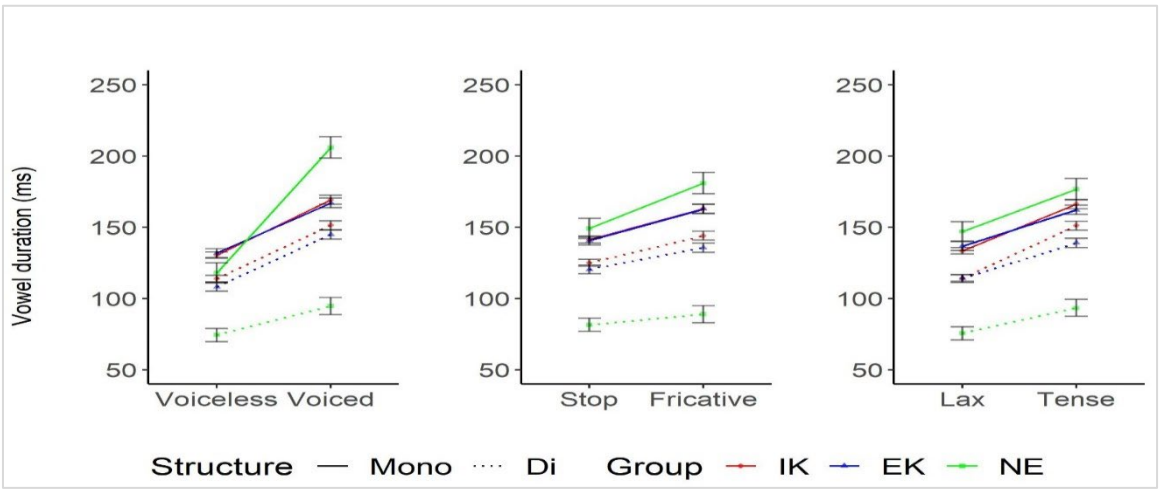


Figure 3. Vowel durations depending on the following consonant voicing, manner, and vowel quantity in two prosodic structures (monosyllabic vs. disyllabic) in Korean speakers and English speakers' English productions. Error bars demonstrate standard errors.

### 3.1.4 Correlation between the voicing effects and vowel identity effects

The relationship between voicing effect size and vowel identity effect size was examined to see whether there is a correlation between two different durational effects in Korean speakers' English production. So, for each participant, the average size of the vowel duration difference for voicing and vowel identity was calculated and the voicing effect size

was plotted against vowel identity effect size. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. There was hardly association between vowel identity effect and voicing effect for Korean speaker groups as in Figure 4. The inexperienced Korean speakers showed a weak ($r^2$ = 0.05), and statistically not significant ($p > 0.05$) negative association between vowel identity effect and voicing effect. The experienced Korean speakers showed very weak ($r^2$ = 0.00), and statistically not significant ($p > 0.05$) positive association between two variables. This weak correlation across the two variables demonstrates that there were different L1 transfer effects on two different durational effects, such that Korean speakers who distinguish English voicing contrasts may or may not distinguish English tense/lax contrasts in their English production.

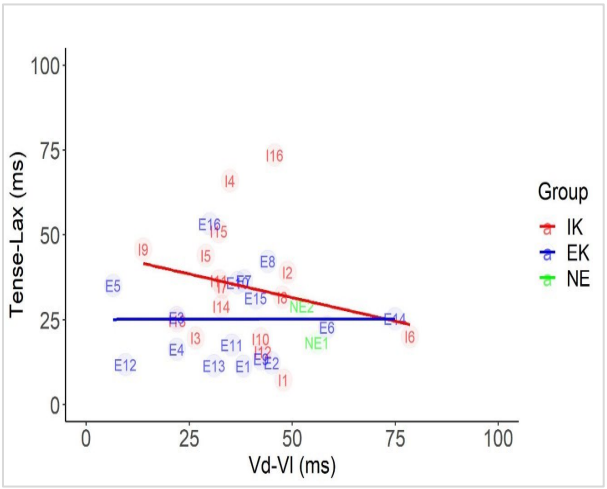


Figure 4. Correlation between the voicing effects and vowel identity effects in Korean speakers and English speakers' English productions.

### 3.2 Vowel formant frequencies

#### 3.2.1 Voicing contrasts (Voiced vs. Voiceless)

Vowel formant frequencies depending on the following consonant voicing were examined to see whether Korean speakers used the spectral properties to distinguish voiced consonants from voiceless counterparts in their English productions.

First, a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and final consonant voicing (voiced vs. voiceless) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for front vowel formant frequencies (F1 and F2). The analysis based on the test of Pillai's trace indicated that there were significant main effects of voicing contrast [$F(2, 30) = 24.6741$, $p < 0.001$], prosodic structure [$F(2, 30) = 56.476$, $p < 0.001$], and group [$F(4, 62) = 3.6553$, $p < 0.05$]. As for interaction effects, there were significant interactions between prosodic structure and group [$F(4, 62) = 2.909$, $p < 0.05$], voicing contrast and group [$F(4, 62) = 5.4583$, $p < 0.05$], and prosodic structure and voicing contrast [$F(2, 30) = 5.6998$, $p < 0.05$] on the dependent variables including F1 and F2.

To figure out the significant interaction effects, a series of ANOVAs was conducted on each formant frequency, separately as a follow-up analysis. First, as for F1 of front vowels, there was a significant main effect of voicing contrast [$F(1, 31) = 38.927$, $p < 0.001$, $\eta^2 = 6.84e-04$, [0.00, 1.00]] where F1 was lower in the voiced context than in the voiceless context. There was a significant main effect of prosodic structure [$F(1, 31) = 108.960$, $p < 0.001$, $\eta^2 = 4.10e-03$, [0.00, 1.00]], but there was no significant main effect of group [$F(2, 31) = 0.225$, $p > 0.05$, $\eta^2 = 1.00$, [0.00, 1.00]]. Voicing contrast interacted with group [$F(2, 31) = 5.717$, $p < 0.05$, $\eta^2 = 7.30e-04$, [0.00, 1.00]] where F1 values of front vowels followed by voiced consonants were lower than those followed by voiceless consonants for all the groups, but the difference between them spoken by Korean groups was not as clearly distinctive as that of English native speakers in most cases. Laryngeal effects on F1 values seemed less pronounced for Korean speakers with more English experience compared to those with less English experience. In general, English native speakers showed much clearer distinction between the F1 of front vowels preceding voiced consonants and those preceding voiceless consonants, and the

difference was more distinctive for high front lax vowel and mid front lax vowel as in Figure 5. There was a significant interaction between voicing contrast and prosodic structure [$F(1, 31) = 7.471$, $p < 0.05$, $\eta^2 = 8.44\text{e-}05$, [0.00, 1.00]]. On the other hand, there was no significant interaction effect between prosodic structure and group [$F(2, 31) = 2.322$, $p > 0.05$, $\eta^2 = 7.12\text{e-}04$, [0.00, 1.00]].

A series of paired t-tests was run on different voicing conditions. According to the t-tests, the F1 of front vowels preceding a voiceless consonant in the monosyllabic structure were significantly higher than those preceding a voiced consonant only for the inexperienced Korean group and English native speaker group. However, after correcting for multiple testing using the Bonferroni correction, the difference was statistically significant only for the inexperienced Korean speaker group (the inexperienced Korean group: $t = -5.5201$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = -2.9416$, $df = 319$, $p > 0.05$; the native English control group: $t = -6.1072$, $df = 39$, $p < 0.05$ (adjusted $p > 0.0083$)). On the contrary, F1 of front vowels preceding a voiceless consonant in the disyllabic structure were significantly higher than those preceding a voiced consonant for all the speaker groups. After correcting for multiple testing using the Bonferroni correction, the difference was statistically significant for all the groups (the inexperienced Korean group: $t = -2.6917$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = -1.891$, $df = 319$, $p < 0.001$; the native English control group: $t = -4.384$, $df = 39$, $p < 0.001$).

As for F2 of front vowels, there was a significant main effect of voicing contrast [$F(1, 31) = 43.306$, $p < 0.001$, $\eta^2 = 9.44\text{e-}04$, [0.00, 1.00]] where F2 was higher in the voiced context than in the voiceless context. There was a significant main effect of prosodic structure [$F(1, 31) = 36.867$, $p < 0.001$, $\eta^2 = 3.47\text{e-}03$, [0.00, 1.00]], and group [$F(2, 31) = 5.246$, $p < 0.05$, $\eta^2 = 1.00$, [0.00, 1.00]]. There was a significant interaction effect between prosodic structure and group [$F(2, 31) = 6.213$, $p < 0.05$, $\eta^2 = 1.93\text{e-}03$, [0.00, 1.00]] where F2 in the monosyllabic

structure was higher than in the disyllabic structure for English native speakers whereas the Korean speaker groups showed the opposite pattern. Voicing contrast interacted neither with group [$F(2, 31) = 1.447$, $p > 0.05$, $\eta^2 = 5.47e\text{-}05$, [0.00, 1.00]], nor with prosodic structure [$F(1, 31) = 2.827$, $p > 0.05$, $\eta^2 = 1.05e\text{-}04$, [0.00, 1.00]].

A series of paired t-tests was run on different voicing conditions. According to the t-tests, the F2 of front vowels preceding a voiceless consonant in the monosyllabic structure were significantly higher than those preceding a voiced consonant for Korean speaker groups. Similarly, the English native speaker group showed a credible difference depending on voicing contrast and the F2 of front vowels preceding a voiceless consonant were lower than those preceding voiced consonant. However, after correcting for multiple testing using the Bonferroni correction, the difference was statistically significant only for Korean speaker groups (the inexperienced Korean group: $t = 2.2317$, $df = 319$, $p < 0.05$; the experienced Korean group: $t = 2.8237$, $df = 319$, $p < 0.05$; the native English control group: $t = -0.04001$, $df = 39$, $p < 0.05$ (adjusted $p > 0.0083$)). As for F2 of front vowels, those preceding a voiceless consonant in the disyllabic structure were significantly higher than those preceding a voiced consonant for the inexperienced Korean speakers and English native speakers. After correcting for multiple testing using the Bonferroni correction, the difference was statistically significant for the inexperienced Korean speaker group and the English native control group (the inexperienced Korean group: $t = 4.7913$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = 3.7693$, $df = 319$, $p > 0.05$.; the native English control group: $t = 0.21438$, $df = 39$, $p < 0.001$).

Next, a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and final consonant voicing (voiced vs. voiceless) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for back vowel formant frequencies (F1 and F2). The analysis based on the test of Pillai's trace indicated

that there were significant main effects of voicing contrast [$F(2, 30) = 17.6799$, $p < 0.001$], and prosodic structure [$F(2, 30) = 10.1936$, $p < 0.05$]. There were no significant interaction effects between factors.

To figure out the significant main effects, a series of ANOVAs was conducted on each formant frequency, separately as a follow-up analysis. First, as for F1 of back vowels, there was a significant main effect of voicing contrast [$F(1, 31) = 36.520$, $p < 0.001$, $\eta^2 = 3.02e\text{-}03$, [0.00, 1.00]] where F1 was lower in the voiced context than in the voiceless context. There was a significant main effect of prosodic structure [$F(1, 31) = 18.637$, $p < 0.05$, $\eta^2 = 7.38e\text{-}03$, [0.00, 1.00]]. As for F2 of back vowels, however, there was no significant main effect of voicing contrast [$F(1, 31) = 1.417$, $p > 0.05$, $\eta^2 = 1.80e\text{-}03$, [0.00, 1.00]] and prosodic structure [$F(1, 31) = 0.777$, $p > 0.05$, $\eta^2 = 4.06e\text{-}04$, [0.00, 1.00]]. These results are plotted in Figure 6.

A series of paired t-tests was run on different voicing conditions. According to the t-tests, the F1 of back vowels preceding a voiceless consonant in the monosyllabic structure were significantly higher than those preceding a voiced consonant for all the groups. However, after correcting for multiple testing using the Bonferroni correction, the difference was statistically significant only for the inexperienced Korean speaker group (the inexperienced Korean group: $t = -3.6471$, $df = 159$, $p < 0.05$; the experienced Korean group: $t = -2.0638$, $df = 159$, $p < 0.05$ (adjusted $p > 0.0083$); the native English control group: $t = -3.3161$, $df = 19$, $p < 0.05$ (adjusted $p > 0.0083$)). On the contrary, F1 of back vowels preceding a voiceless consonant in the disyllabic structure were significantly higher than those preceding a voiced consonant for English native control group and Korean speaker group with more English exposure. After correcting for multiple testing using the Bonferroni correction, the difference was statistically significant for English native control group and Korean speaker group with more English exposure (the inexperienced Korean group: $t = -1.8555$, $df = 159$, $p > 0.05$; the experienced

Korean group: $t$ = -2.5943, $df$ = 159, $p < 0.05$.; the native English control group: $t$ = -5.1067, $df$ = 19, $p < 0.001$).

To summarize, Korean speakers somehow showed laryngeal effects on vowel formants. Korean speakers with more English exposure did not show larger laryngeal effects on formants compared to those with less English exposure. In conclusion, Korean speakers somehow seemed to be able to use spectral cues to distinguish final consonant voicing contrasts in their English productions, and the amount of the English experience did not enhance their performance.

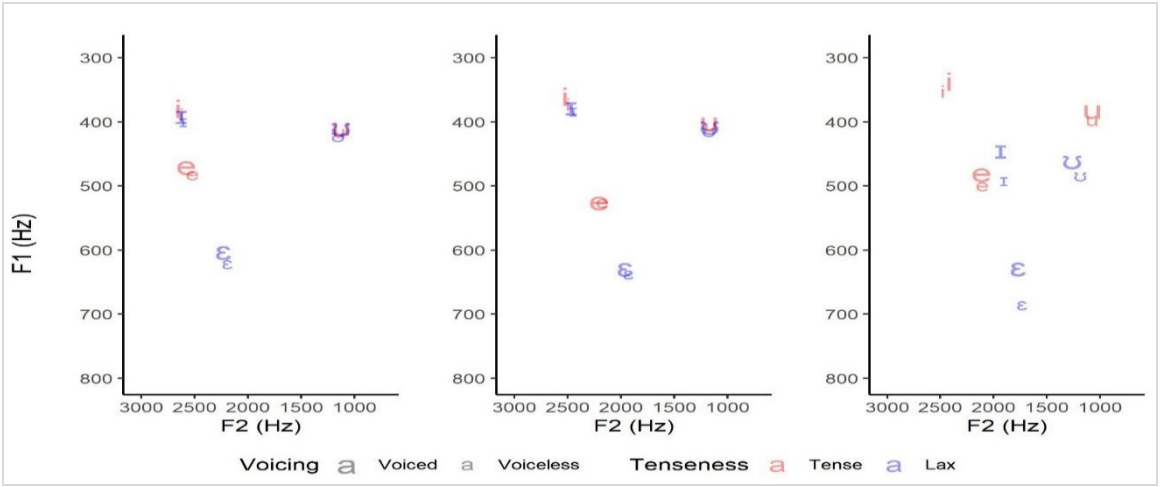


Figure 5. Vowel formant frequencies (F1, F2) depending on the final consonant voicing in Korean speakers and English speakers' English productions, left panel: IK, middle panel: EK, right panel: NE.

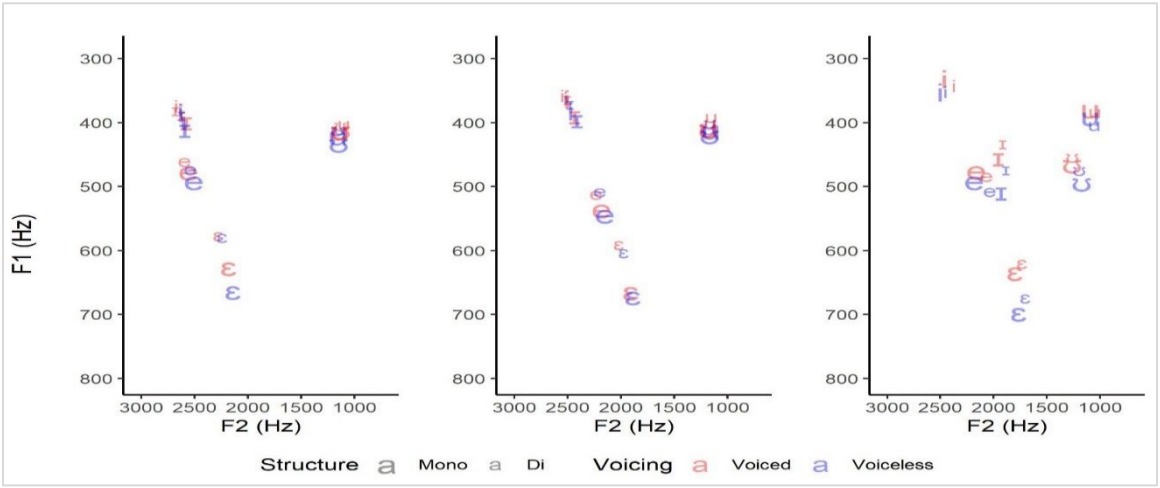


Figure 6. Vowel formant frequencies (F1, F2) depending on the final consonant voicing and the different prosodic structure in Korean speakers and English speakers' English productions, left panel: IK, middle panel: EK, right panel: NE.

### 3.2.2 Vowel identity (Tense vs. Lax)

In this study, vowel formant frequencies depending on vowel identity were examined to see whether Korean speakers used the spectral properties to distinguish tense vowels from lax vowels in their English productions.

First, a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and vowel identity (tense vs. lax) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for front vowel formant frequencies (F1 and F2). The analysis based on the test of Pillai's trace indicated that there were significant main effects of vowel identity [$F(2, 30) = 63.770$, $p < 0.001$], prosodic structure [$F(2, 30) = 56.476$, $p < 0.001$], and group [$F(4, 62) = 3.6553$, $p < 0.05$]. As for interaction effects, there were significant interactions between prosodic structure and group [$F(4, 62) = 2.909, p < 0.05$] , vowel identity and group [$F(4, 62) = 3.388, p < 0.05$], and prosodic structure and vowel identity [$F(2, 30) = 14.4100$, $p < 0.001$] on the dependent variables including F1 and F2.

To figure out the significant interaction effects, a series of ANOVAs was conducted on each formant frequency, separately as a follow-up analysis. First, as for F1 of front vowels, there were significant main effects of vowel identity [$F(1, 31) = 120.909, p < 0.001, \eta^2 = 0.01$, [0.01, 1.00]], prosodic structure [$F(1, 31) = 108.960$, $p < 0.001$, $\eta^2 = 4.46e\text{-}03$, [0.00, 1.00]], but there was no significant main effect of group [$F(2, 31) = 0.225, p > 0.05, \eta^2 = 1.00$, [0.00, 1.00]].

There was a significant interaction between vowel identity and group [$F(2, 31) = 4.474$, $p < 0.05$, $\eta^2 = 6.44e\text{-}03$, [0.00, 1.00]] where both Korean speaker groups showed greatly overlapping areas between high front tense vowels and corresponding lax vowels in the vowel space. On the contrary, in general, English native speakers showed clear distinction between

tense vowels and lax vowels. Vowel identity interacted with prosodic structure [$F(1, 31) = 29.769$, $p < 0.001$, $\eta^2 = 7.23e\text{-}04$, [0.00, 1.00]] where F1 difference between tense vowels and lax vowels was larger in the monosyllabic structure than in the disyllabic structure for all the groups. Korean speakers showed much greater F1 difference between tense vowels and lax vowels for the mid front vowel pair than for the high front vowel pair as in Figure 8. There was no significant interaction between prosodic structure and group [$F(2, 31) = 2.322$, $p > 0.05$, $\eta^2 = 7.76e\text{-}04$, [0.00, 1.00]].

A series of paired t-tests was run on tense/lax pairs of English vowels. According to the t-tests, the F1 of front lax vowels in the monosyllabic structure were significantly higher than those of corresponding tense vowels for all the groups. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the groups (the inexperienced Korean group: $t = -14.23$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = -12.398$, $df = 319$, $p < 0.001$; the native English control group: $t = -16.35$, $df = 39$, $p < 0.001$). Also, F1 of front lax vowels in the disyllabic structure were significantly higher than those of corresponding tense vowels for all the groups. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the groups (the inexperienced Korean group: $t = -14.031$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = -10.055$, $df = 319$, $p < 0.001$; the native English control group: $t = -11.413$, $df = 39$, $p < 0.001$).

As for the effect of English experience on the Korean speakers' performance, Korean speakers with less English experience showed more clear distinction in terms of F1 values between mid front tense vowels and corresponding lax vowels than those with more English experience. More specifically, more advanced Korean speakers' F1 vowel spaces occupied by mid front tense vowels were often as high as corresponding lax vowels and thus they had a lot of overlapping spaces. Further, mid front tense vowels spoken by more advanced Korean

speakers had more variation in terms of F1 values than did less advanced Korean speakers. However, the experienced Korean speaker group resembled English native speakers more in terms of general spectral patterns as in Figure 7.

As for F2 of front vowels, there were significant main effects of vowel identity [$F(1, 31) = 116.283$, $p < 0.001$, $\eta^2 = 0.01$, [0.01, 1.00]], prosodic structure [$F(1, 31) = 36.867$, $p < 0.001$, $\eta^2 = 3.76e\text{-}03$, [0.00, 1.00]], and group [$F(2, 31) = 5.246$, $p < 0.05$, $\eta^2 = 1.00$, [0.00, 1.00]]. Vowel identity interacted with group [$F(2, 31) = 8.164$, $p < 0.05$, $\eta^2 = 9.58e\text{-}03$, [0.00, 1.00]] where both Korean speaker groups showed greatly overlapping spaces between high front tense vowels and corresponding lax vowels in the vowel space. On the contrary, in general, English native speakers showed clear distinction between tense vowels and lax vowels.

There was a significant interaction between vowel identity and prosodic structure [$F(1, 31) = 7.758$, $p < 0.05$, $\eta^2 = 3.28e\text{-}04$, [0.00, 1.00]] where F2 difference between tense vowels and lax vowels was a little bit larger in the monosyllabic structure than in the disyllabic structure for all the groups. Korean speakers showed a greater F2 difference between tense vowels and lax vowels for the mid front vowel pair than for the high front vowel pair as in Figure 8. There was a significant interaction between prosodic structure and group [$F(2, 31) = 6.213$, $p < 0.05$, $\eta^2 = 2.09e\text{-}03$, [0.00, 1.00]] where F2 values in the monosyllabic structure were lower than those in the disyllabic structure for both Korean groups whereas English native speakers showed the opposite pattern.

A series of paired t-tests was run on tense/lax pairs of English vowels. According to the t-tests, the F2 of front tense vowels in the monosyllabic structure were significantly higher than those of corresponding lax vowels for all the groups. However, after correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant only for the English native speaker group and the inexperienced Korean speaker group (the inexperienced Korean group: $t = 13.046$, $df = 319$, $p < 0.001$; the experienced Korean group: $t$

= 11.239, *df* = 319, $p < 0.05$ (adjusted $p > 0.0083$); the native English control group: $t = 18.472$, *df* = 39, $p < 0.001$). F2 of front tense vowels in the disyllabic structure were significantly higher than those of corresponding lax vowels for all the groups. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the groups (the inexperienced Korean group: $t = 11.583$, $df = 319$, $p < 0.001$; the experienced Korean group: $t = 10.04$, $df = 319$, $p < 0.001$; the native English control group: $t = 14.76$, $df = 39$, $p < 0.001$).

Korean speakers with less English experience, in general, showed higher F2 values than those with more English experience and they showed more clear distinction in terms of F2 values between mid front tense vowels and corresponding lax vowels than those with more English experience. More specifically, more advanced Korean speakers' F2 vowel spaces for mid front tense vowels were often as low as corresponding lax vowels and thus they had a lot of overlapping spaces. Mid front tense vowels spoken by more advanced Korean speakers had more variation in terms of F2 values than did less advanced Korean speakers. More experienced Korean speaker group resembled English native speakers more in terms of general spectral patterns as in Figure 7.

Next, a MANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and vowel identity (tense vs. lax) and prosodic structure (monosyllabic vs. disyllabic) were within-subject factors was run for back vowel formant frequencies (F1 and F2). The analysis based on the test of Pillai's trace indicated that there were significant main effects of vowel identity [$F(2, 30) = 7.3583$, $p < 0.05$], and prosodic structure [$F(2, 30) = 10.1936$, $p < 0.05$]. As for interaction effects, there was a significant interaction between vowel identity and group [$F(4, 62) = 7.0926$, $p < 0.001$] on the dependent variables including F1 and F2.

To figure out the significant interaction effects, a series of ANOVAs was conducted on each formant frequency, separately as a follow-up analysis. First, as for F1 of back vowels, there were significant main effects of vowel identity [$F(1, 31) = 14.64$, $p < 0.05$, $\eta^2 = 2.47\text{e-}04$, [0.00, 1.00]], and prosodic structure [$F(1, 31) = 18.637$, $p < 0.05$, $\eta^2 = 7.69\text{e-}03$, [0.00, 1.00]]. Vowel identity interacted with group [$F(2, 31) = 24.69$, $p < 0.001$, $\eta^2 = 0.03$, [0.02, 1.00]] where both Korean speaker groups showed great overlapping spaces between high back tense vowels and corresponding lax vowels in the vowel space. On the contrary, in general, English native speakers showed clear distinction between tense vowels and lax vowels.

A series of paired t-tests was run on tense/lax pairs of English vowels. According to the t-tests, the F1 of back lax vowels in the monosyllabic structure were significantly higher than those of corresponding tense vowels only for the experienced Korean speaker group. However, after correcting for multiple testing using the Bonferroni correction, none of the paired t-tests reached significance (the inexperienced Korean group: $t$ = -1.3787, $df$ = 159, $p > 0.05$; the experienced Korean group: $t$ = -1.3963, $df$ = 159, $p < 0.05$ (adjusted $p > 0.0083$); the native English control group: $t$ = -7.785, $df$ = 19, $p > 0.05$). On the other hand, English native speakers and Korean speakers showed significantly higher F1 for back lax vowels than for corresponding tense vowels in the disyllabic structure. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the groups (the inexperienced Korean group: $t$ = -0.52549, $df$ = 159, $p < 0.001$; the experienced Korean group: $t$ = -2.9323, $df$ = 159, $p < 0.001$; the native English control group: $t$ = -5.844, $df$ = 19, $p < 0.001$). High back tense vowels spoken by more advanced Korean speakers had more variation in terms of F1 values than did less advanced Korean speakers. More experienced Korean speaker group resembled English native speakers a little bit more in terms of general spectral patterns as in Figure 7.

As for F2 of back vowels, there was no significant main effect of vowel identity [$F(1, 31) = 1.916$, $p > 0.05$, $\eta^2 = 3.05e\text{-}05$, [0.00, 1.00]], or prosodic structure [$F(1, 31) = 0.777$, $p > 0.05$, $\eta^2 = 4.10e\text{-}04$, [0.00, 1.00]]. However, there was a significant interaction between vowel identity and group [$F(2, 31) = 3.760$, $p < 0.05$, $\eta^2 = 8.01e\text{-}03$, [0.00, 1.00]] where both Korean speaker groups showed greatly overlapping spaces between high back tense vowels and corresponding lax vowels in the vowel space. On the contrary, English native speakers showed clear distinction between tense vowels and lax vowels.

A series of paired t-tests was run on tense/lax pairs of English vowels. According to the t-tests, F2 of back lax vowels in the monosyllabic structure were not significantly higher than those of corresponding tense vowels for all the groups (the inexperienced Korean group: $t = 0.44596$, $df = 159$, $p > 0.05$; the experienced Korean group: $t = -0.40322$, $df = 159$, $p > 0.05$; the native English control group: $t = -3.5722$, $df = 19$, $p > 0.05$). On the contrary F2 of back lax vowels in the disyllabic structure were significantly higher than those of corresponding tense vowels for all the groups. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the groups (the inexperienced Korean group: $t = -1.1083$, $df = 159$, $p < 0.001$; the experienced Korean group: $t = -0.75899$, $df = 159$, $p < 0.001$; the native English control group: $t = -4.8291$, $df = 19$, $p < 0.001$). High back tense vowels spoken by less advanced Korean speakers had more variation in terms of F2 values than did more advanced Korean speakers. These results are plotted in Figure 7 and 8.

To summarize, both Korean groups showed different patterns from those of English native speakers showing a greatly overlapping area between tense vowels and lax vowels in the formant space. More experienced Korean speaker group resembled English native speakers more in terms of general spectral patterns. However, there seemed no larger vowel identity effect for Korean subjects with more English experience than those with less English experience. As for the structure effect on the vowel identity, Korean speakers showed the

opposite pattern from that of English native speakers, indicating lower F1 and F2 values in the monosyllabic structure than in the disyllabic structure. In conclusion, Korean speakers seemed to use spectral cues for English front tense/lax vowel contrast whereas they failed to exploit the spectral cues effectively for English back tense/lax vowel contrast in their English production.

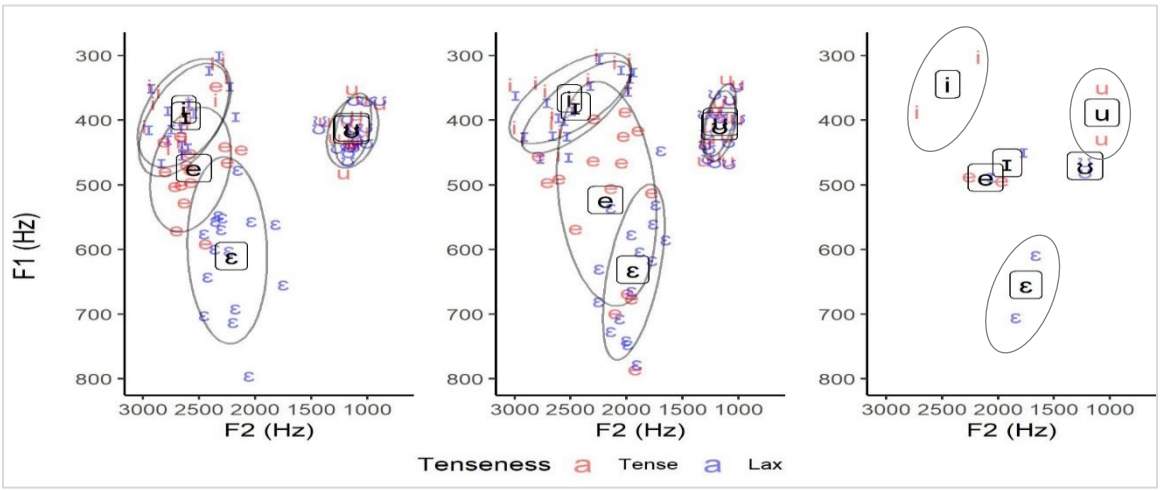


Figure 7. Vowel formant frequencies (F1, F2) depending on the vowel identity in Korean speakers and English speakers' English productions, left panel: IK, middle panel: EK, right panel: NE.

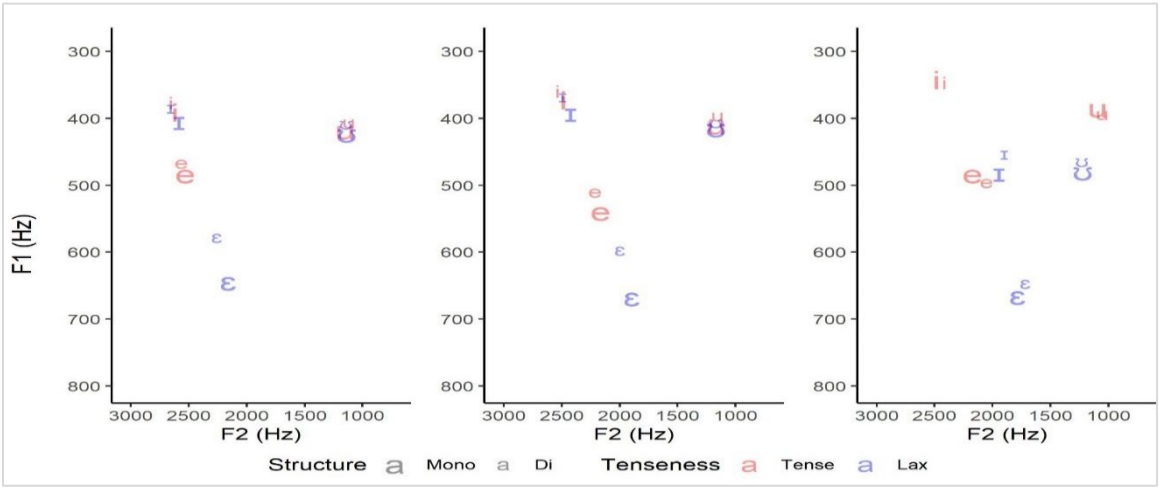


Figure 8. Vowel formant frequencies (F1, F2) depending on the vowel identity and the different prosodic structure in Korean speakers and English speakers' English productions, left panel: IK, middle panel: EK, right panel: NE.

## 4. Discussion

This study investigated whether Korean learners of English produce the coda consonant contrasts in vowel duration. Also, it concurrently investigated whether intrinsic vowel durational differences in English would be shown by the same Korean learners of English. In addition, spectral properties which are also important acoustic information inherent in English target vowels were examined to see whether Korean speakers with more English exposure utilize the cue more effectively than those with less English experience do. The consonants chosen for the experiments in this study consist of sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/) and fricatives (/f, v/, /s, z/). The vowels selected for the experiment include English tense/lax vowel pairs /i, ɪ, e, ɛ, u, ʊ/ which cause high confusion to Korean learners of English because Korean has no tense/lax distinction.

First, it was expected that the Korean L1 speakers would show an attenuated difference in their English production associated with the following obstruent voicing in monosyllabic words because of the lack of a coda distinction in obstruent voicing in Korean. On the contrary, they were expected to show a larger vowel durational difference in disyllabic words since the Korean contrast is not neutralized in medial position. According to the results, however, Korean speakers showed laryngeal effects not only in non-final position but in final position, where vowel length in the voiced context was significantly longer than that in the voiceless context in final position as well as in non-final position. In addition, Korean speakers did not show a larger vowel durational difference depending on the following consonant voicing in disyllabic words than in monosyllabic words. In monosyllabic words, Korean speaker groups showed a smaller vowel durational difference than did English native speaker group (the inexperienced Korean group's mean difference: 38.9 ms; the experienced Korean group's mean difference: 35.1 ms; the native English control group's mean difference: 88 ms). On the contrary, the Korean speaker group showed a larger vowel durational difference than did

English native speaker group in disyllabic words (the inexperienced Korean group's mean difference: 37.5 ms; the experienced Korean group's mean difference: 36.7 ms; the native English control group's mean difference: 20.3 ms).

To account for the typical durational differences between learners and non-learners, the duration differences were also expressed as ratios. When it comes to the ratio of vowel duration in the voiceless context to that in the voiced context, Korean speakers showed weaker voicing effects on vowel duration than did English native speakers in monosyllabic words. More specifically, the ratio of vowel duration in the voiceless context to that in the voiced context in monosyllabic words was 78% for Korean speakers (the inexperienced Korean group's ratio: 77%; the experienced Korean group's ratio: 79%) whereas that of English native speakers was 57.3%. On the contrary, in disyllabic words, Korean speaker groups resembled the English native group. More specifically, the ratio of vowel duration in the voiceless context to that in the voiced context in disyllabic words was 75% for Korean speakers (the inexperienced Korean group's ratio: 75.2%; the experienced Korean group's ratio: 74.7%) and that of English native speakers was 78.6%. This suggests that they have not quite yet acquired voicing effects to distinguish voicing contrasts in English.

Korean coda laryngeal neutralization rules seemed not to have an impact on vowel duration depending on the following obstruent voicing in Korean speakers' English production. More specifically, the ratio of vowel duration in the voiceless context to that in the voiced context demonstrated that Korean speakers showed weaker voicing effects on vowel duration than did English native speakers in monosyllabic words, and Korean speakers resembled the English native group in terms of the ratio in disyllabic words where underlying consonants remain unneutralized intervocalically in Korean. However, there was a strong positive association between voicing effect size in non-final position and that in final position, which

indicates that Korean coda laryngeal neutralization rules are not an important factor for Korean speakers' English production.

Second, it was expected that Korean subjects would exhibit different patterns with respect to laryngeal neutralization and manner neutralization in their English production. This is because Lee (2016) reported that laryngeal neutralization was phonetically incomplete in her study whereas manner neutralization was complete in the study of Kim and Jongman (1996). However, in this study Korean speakers showed manner effects where vowel length preceding fricatives was longer than that preceding stops, not only in non-final position (the inexperienced Korean group's mean difference: 19 ms; the experienced Korean group's mean difference: 15.2 ms; the native English control group's mean difference: 7.6 ms) but in final position (the inexperienced Korean group's mean difference: 21.6 ms; the experienced Korean group's mean difference: 22.2 ms; the native English control group's mean difference: 31.7 ms). In monosyllabic words, Korean speaker groups showed a smaller vowel durational difference than did English native speaker group whereas the opposite pattern occurred in disyllabic words. However, Korean speakers did not show a credibly larger vowel durational difference depending on the following consonant manner in disyllabic words than in monosyllabic words.

When it comes to the ratio of vowel duration preceding stops to that preceding fricatives, Korean speaker groups showed different patterns from what was shown in laryngeal effects. More specifically, Korean speakers showed weaker voicing effects on vowel duration than did English native speakers in monosyllabic words. However, Korean speaker groups resembled English native group in terms of the ratio of vowel duration preceding stops to that preceding fricatives in monosyllabic words (the inexperienced Korean group's ratio: 86.7%; the experienced Korean group's ratio: 86.4%; the native English control group's ratio: 82.5%) as well as in disyllabic words (the inexperienced Korean group's ratio: 86.8%; the experienced

Korean group's ratio: 88.8%; the native English control group's ratio: 91.5%). According to the results, there were no larger manner effects for Korean subjects with more English exposure than those with less English exposure. Thus, it appears that the mechanism that produces the manner effects on duration may be a more-or-less automatic effect of producing the manner contrasts, and so is evident in both native and non-native speakers alike. This differs from the voicing effects which are larger in the native speakers. Note also that the manner effects are much smaller in magnitude than the voicing effects, which might also be due to the manner effects not being magnified as a cue to the following consonant.

Third, it was expected that Korean speakers would have difficulty in production of English tense/lax vowel contrasts due to the lack of an intrinsic vowel quantity distinction in Korean. In this study, however, tense vowels were significantly longer than lax vowels for all the groups (the inexperienced Korean group's mean difference: 34.8 ms; the experienced Korean group's mean difference: 25.1 ms; the native English control group's mean difference: 23.7 ms). In addition, Korean speaker groups resembled English native group in terms of the ratio of lax vowels to tense vowels (the inexperienced Korean group's ratio: 78.1%; the experienced Korean group's ratio: 83.3%; the native English control group's ratio: 82.5%). Korean speakers seemed to exploit temporal cues effectively for conveying English tense/lax contrasts in their English production, which was different from Flege, Bohn and Jang (1997)'s claim. However, there were no larger vowel identity effects for Korean subjects with more English experience than those with less English experience.

Lastly, it was expected that Korean subjects would exhibit different L1 transfer effects on two different durational effects which include voicing effects (voiced vs. voiceless) and vowel quantity effects (tense vs. lax). As for vowel duration, Korean speakers showed laryngeal effects not only in non-final position but in final position, where vowel length in the voiced context was significantly longer than that in the voiceless context. However, when it comes to

the ratio of vowel duration in the voiceless context to that in the voiced context, voicing effects seemed smaller for Korean speakers than for English native speakers. On the contrary, tense vowels were significantly longer than lax vowels for both Korean groups and they resembled English native group in terms of the ratio of lax vowels to tense vowels. Thus, it appears that the Korean learners were more effective at harnessing durational differences for the tense/lax distinction than as a cue for following obstruent voicing.

Reflecting on how the durational effects play out in the quality of the vowels, the Korean speaker groups showed different patterns depending on two different durational effects in their English production. More specifically, for voicing effects, the Korean speaker group with less exposure to English speaking countries somehow utilized spectral cues to distinguish final consonant voicing contrasts in their English production, and the amount of the English experience did not enhance their performance. As for vowel identity effects, however, for the front vowels, both Korean speaker groups seemed to use spectral cues effectively to distinguish tense vowels from lax vowels in their English production. However, for the back vowels, they did not exploit spectral cues effectively for the tense/lax contrast and the amount of the English experience did not enhance their performance. Korean speakers' failure of using spectral cues to distinguish English tense/lax contrast was consistent with Bohn's (1995) Desensitization hypothesis that L2 learners' sensitivity to spectral differences will be lesser when they have only one vowel in the L1. More specifically, difficulty might be imposed for Korean speaker groups because the high English vowel pair /i/ vs. /ɪ/ would assimilate to the high front vowel /i/ in Korean. Also, there is only one high back vowel /u/ in Korean, which ranges over the entire high back region. Therefore, their L1 experience might impose difficulty for Korean speakers to distinguish between English tense and lax contrast. It appears that the reason why Korean speakers used spectral cues more effectively for English front tense/lax vowel contrast

than for back tense/lax vowel contrast seems to be due to the fact that the front vowels have quality differences which are easier for the Korean learners to acquire than the back vowels.

## 5. Conclusion

Based on the results and analyses of the current study, the statistical analysis supported that Korean learners of English utilized temporal cues for the English contrasts in their English production.

Korean speakers showed laryngeal effects not only in non-final position but in final position, where vowel length in the voiced context was significantly longer than that in the voiceless context. However, unlike the native speakers, Korean speakers did not show a large English-like pattern in vowel duration distinction followed by a different consonant voicing in English production. Manner effects differed from the voicing effects in that Korean speakers resembled English native speakers in vowel duration, regardless of L2 English experience. Based on these results, manner effects on vowel duration seemed just to be characteristic of manner contrasts and thus are readily revealed in L2 production.

When it comes to vowel quantity effects, tense vowels were significantly longer than lax vowels for both Korean groups and they resembled English native group in terms of the vowel duration patterns. Thus, it appears that the Korean learners were more effective at harnessing durational differences for the tense/lax distinction than following obstruent voicing contrast.

When it comes to spectral cues, Korean speakers failed to utilize spectral cues successfully to distinguish English contrasts in their English production. Korean speakers with less exposure to English speaking countries seemed to use spectral cues for English consonant voicing contrasts whereas both Korean speaker groups failed to use spectral cues for the English tense/lax distinction, especially in back vowels. This seems to be due to the fact that

temporal cues are universally more salient and thus it is easier for Korean learners to access whereas it is not easy for Korean speakers to be sensitive to small spectral changes in vowel quality due to a smaller vowel inventory in Korean than in English.

Finally, in general, Korean speakers with more English exposure did not make use of the temporal cues nor spectral cues more effectively than did those with less English exposure to distinguish English contrasts in their English production. This seems to be due to the limitation of the current study such that the difference of English experience between two Korean speaker groups was not large enough to enhance their English production. More specifically, the inexperienced Korean speaker group has less than a week whereas the experienced Korean speaker group has around five years of exposure to an English dominant country, and this five-year difference might not have a significant impact on developing their English production skills. In this sense, it is difficult to interpret that the amount of English experience by itself does not indicate how good English production is in this study.

# III. Perception

## 1. Introduction

*Vowel duration modulation by consonant voicing*

The production experiment in Chapter 2 demonstrated that Korean speakers have not quite yet acquired native-like voicing effects to distinguish voicing contrasts in English production, even though there was a distinctive laryngeal effect not only in non-final position but in final position. The Korean speakers' ratio of vowel duration in the voiceless context to that in the voiced context was relatively greater than that of English native speakers. Also, Korean coda laryngeal neutralization rules appear not to be an important factor for Korean speakers' English production.

Based on the production results, a question arises about the acquisition of voicing effects to distinguish voicing contrasts in English perception by the same Korean speakers. Therefore, the current study examines whether Korean learners of English are able to perceive vowel duration difference due to the following consonant voicing and whether L2 English experience enhances their English perception. Numerous studies have indicated that vowel duration cues played an important role for English native listeners to discriminate English final consonant voicing contrasts in their English perception (Denes, 1955; Port & Dalby, 1982; Raphael, 1972, and see Flege, 1993, with respect to acquisition). For example, Denes (1955) examined the perception of English coda consonant voicing by English listeners. Participants were asked to label a final fricative whether it was heard as voiced fricative /z/ or its voiceless counterpart /s/ in monosyllabic words in synthetic speech. Denes reported that the final fricative was heard as voiceless when the consonant duration was shorter than that of the preceding vowel whereas fricative was heard as voiceless when preceding vowel duration was longer or about the same as that of postvocalic fricative. This result suggests that English native listeners use two different cues, preceding vowel duration and duration of postvocalic fricative, to

discriminate English final fricative voicing in their perception. Port and Dalby (1982) similarly examined the discrimination of English intervocalic stop voicing contrasts in synthetic speech by English listeners. Participants were asked to label whether the target stimuli heard as 'rabid' or 'rapid' for labial stop pairs, and as 'buggy' or 'bucky' for velar stop pairs. Authors reported that consonant/vowel ratio was a primary cue to decide on whether English medial consonant was voiced or voiceless in their perception. Raphael (1972) investigated the perception of voicing feature of English coda consonants in synthetic speech by English native listeners. The English vowel durations were varied before final consonants including stops, fricatives, and consonant clusters in monosyllabic words. After the voiced series were created, voiceless counterpart series were made by removing the final 50 ms F1 transition from the voiced series. Raphael reported that listeners labeled the target consonant as voiceless when preceding vowel duration was short whereas it was labeled as voiced when duration of preceding vowel was long, irrespective of voicing cue. This indicates that the preceding vowel duration is more effective cue than the voicing cue in perceptual categorization of English final consonant voicing contrasts.

Research on Korean indicates that the voicing-related vowel duration differences can play an important role for Korean speakers in discriminating the coda consonant contrasts in their English perception (Kim & Shin, 2014; Ko, 1997; Ko et al., 1997; Ko, 1998; Lee, 2016; Yang 2020). For example, Ko et al. (1997) examined the discrimination of English coda consonant contrasts in English perception by Korean listeners and Canadian English listeners. The English words 'cab' and 'cap' were produced by an English native speaker and the vowel length for the words was manipulated by shortening various parts of the target vowel. The subjects were asked to choose either 'cab' or 'cap' on a response sheet. The authors reported that Korean listeners heard a greater number of the target stimuli as voiceless as the preceding vowel length became shorter, whereas manipulated vowel length did not affect Canadian

English listeners' response on English final consonant voicing. This implies that vowel duration played an important role in classification of English coda consonant voicing contrasts for Korean listeners whereas coda consonant voicing played more important role for Canadian English native listeners. Ko (1997, 1998) examined the perception of English final consonant voicing by Korean listeners and English native listeners. Sentences including English front vowel /æ/ before coda stops differing in voicing (/p, t, k/ and /b, d, g/) in monosyllabic words were used as stimuli. The target vowels were manipulated in many ways and then presented aurally to listeners. According to these results, closure voicing of final consonants played an important role in the perception of final consonant voicing for English listeners. On the other hand, preceding vowel duration differences played a more important role for some Korean listeners whereas closure voicing of final consonant was a sufficient cue for other Korean listeners in their perception of English final consonant voicing.

Kim and Shin (2014) examined the production and perception of preceding vowel length difference depending on the following consonant voicing in English by Korean speakers. Participants consisted of a Korean speaker group speaking standard Korean and the other group speaking the Kyungbuk dialect. The latter group was divided into two groups where one group distinguished Korean long vowels from short vowels while the other group did not show vowel length contrasts. English vowels /i, ɪ, e, æ, aɪ, oʊ, u, ʌ/ before sets of consonants contrasting in voicing (/p, b/, /t, d/, /k, g/, /f, v/, /ɵ, ð /, and /s, z/) in monosyllabic words were aurally presented to the listeners as stimuli. Participants were asked to discriminate English final consonant voicing contrasts in an AX discrimination task. According to the results, Korean listeners' perception scores ranged from 77% to 95%. The authors claimed that the perception preceded production in their study, because the ratio of vowel duration in the voiceless context to that in the voiced context for Korean speakers was greater than that of English native speakers. This indicates that Korean speakers performed better in perception than in production. In addition,

authors suggest that L1 transfer effects were stronger in production than in perception, because Kyungbuk dialect speakers who distinguished Korean long vowels from short vowels performed best among all the groups in terms of vowel length difference in production. Also, this group resembled the duration pattern of English native speakers more than other Korean groups.

Focusing more on the perceptual effect of laryngeal neutralization, in Lee's (2016) study Korean college students classified the correct underlying words which had laryngeally neutralized final consonants in CVC syllables in Korean. Lee reported that Korean participant's average accuracy rate for the identification test was around chance level, suggesting that it was hard for Korean listeners to differentiate underlying voiced final consonants from voiceless counterparts in their Korean perception. Korean listeners tended to identify the target stimuli as the words including voiceless consonants irrespective of the status of underlying coda consonants as vowel duration was shorter. However, when the underlying coda consonant of target stimuli was voiceless, vowel duration did not contribute to Korean listeners' classification of intended coda consonant.

Finally, Yang (2020) examined whether Korean high school students differing in English test level perceive preceding vowel length difference depending on the following stop voicing in their English perception. The English front high to low vowels /ɪ, ɛ, æ, ɑ, ʌ/ before coda stops differing in voicing (/p, t, k/ and /b, d, g/) in monosyllabic words were recorded by English native speakers and then the final consonants were deleted. Korean listeners heard the stimuli and were asked to choose whether the target stimuli included voiced final consonant or voiceless final consonant. After the test, Korean listeners answered the questionnaire asking about which factor affected their response the most among five alternatives, including duration of onset consonant, vowel duration, final consonant duration, stress, and not sure. According to the results, there was a strong correlation between English test level and accuracy rates.

Students with higher English test level performed better at perception of preceding vowel length difference before voiced or voiceless stop than did those with lower English test level. Most listeners chose the duration of final consonant as the most important cue to perception of English final consonant voicing. This implies that preceding vowel duration has information of the following consonant voicing, and thus it is a sufficient cue in the perception of English final consonant voicing. Further, listeners who chose vowel duration as the most important factor for their response gained higher accuracy rates compared to those who chose other factors. Based on previous research, it is expected that Korean speakers would be able to perceive vowel length difference depending on the following consonant voicing in English perception.

*Vowel duration differences due to vowel identity*

Regarding vowel duration modulation due to vowel identity, the production experiment reported in Chapter 2 of this thesis demonstrated that Korean speakers seemed to exploit temporal cues for conveying English tense/lax contrasts in their English production. Korean speakers' tense vowels were significantly longer than lax vowels and Korean speakers resembled English native speakers in terms of the ratio of lax vowels to tense vowels. On the contrary, Korean speakers failed to utilize spectral cues effectively for English tense/lax contrasts in their English production because they showed a greatly overlapping area between tense vowels and lax vowels in the formant space.

Based on the production results, a question arises about the perception of English tense/lax contrasts by the same Korean speakers. Therefore, the current study examines whether Korean learners of English are able to perceive English tense/lax vowel contrasts and whether L2 English experience enhances their English perception. It is expected that Korean speakers would rely more on temporal cues than spectral cues for English tense/lax contrasts

in their English perception. Previous research on Korean learners of English reported that they had difficulty in the discrimination of the English tense/lax vowel contrasts in their English perception (Cho & Jeong, 2013; Flege, Bohn & Jang, 1997; Hong, 2012; Ingram & Park, 1997; Kim, 2010; Kim, 2016; Kim, 2006; Lee, 2008; Lee & Cho, 2015; Tsukada et al., 2005). For example, Flege, Bohn and Jang (1997) examined English vowel perception including /i, ɪ, ɛ, æ/ by non-native listeners differing in L2 English experience and English native listeners. Participants were asked to define the identity of target vowels in /i/-/ɪ/ and /ɛ/-/æ/ synthetic continua where the target vowel was manipulated spectrally and temporally. They found that Korean speakers relied more on temporal cues than spectral cues for discriminating English tense/lax vowel contrasts whereas English native listeners preferred spectral cues over temporal cues in their perception. L2 English experience played a marginal role in Korean listeners' responses. Ingram and Park (1997) also investigated the perception of Australian English vowels including /i:, ɪ, e, æ, a:/ spoken by two Australian male speakers by Korean learners of English differing in English experience. They reported that Korean learners with relatively more English exposure performed better on Australian English monophthong vowel contrast /e-æ/ than did the relatively less English experienced Korean learners. On the contrary, L2 English experience did not have an influence on discriminating /i:-ɪ/ vowel contrast. This seems to be due to the fact that Korean listeners' identification of English vowels was strongly affected by speaker variation in terms of absolute duration of target words. More specifically, Korean listeners transcribed English target words in Korean orthography and as a result, 30% of speaker A's English /ɪ/ was assimilated to a Korean long vowel. On the other hand, all of speaker B's English /ɪ/ were identified as a Korean short vowel.

In the study of Tsukada et al. (2005), Korean adult listeners classified English vowels including /i, ɪ, e, ɛ, æ, ɑ, ʌ, u/ with respect to Korean vowel categories. English vowels including /ɪ, e, ɛ, ʌ, u/ were assimilated to multiple instances of Korean vowel categories

whereas over 86% of /i/, /æ/, and /ɑ/ were classified as Korean /i/, /a/, and /a/, respectively. In the following perception test, Korean child and adult listeners differing in L2 English experience discriminated English vowel contrasts including /i/-/ɪ/, /e/-/ɛ/, /ɛ/-/æ/, and /ɑ/-/ʌ/. According to the results, Korean child listeners performed better than did Korean adult listeners and L2 English experience only affected Korean child listeners' discrimination. Korean child and adult listeners were both worst at discriminating English /ɛ/-/æ/ contrast and Korean adult listeners' overall discrimination of English vowel contrasts was poor.

In Lee's (2008) study, Korean speakers and English native speakers discriminated English high tense-lax vowels including /i/-/ɪ/ and /u/-/ʊ/ pairs in monosyllabic words where vowel duration was modified. As a result, Korean listeners showed an accuracy rate of 62% in distinguishing English high tense vowels from lax vowels in the perception task whereas English native listeners perfectly discriminated English tense-lax vowel pairs. More specifically, Korean listeners showed more confusions for English vowel pairs consisting of lengthened lax vowels and shortened tense vowels. This result suggests that Korean listeners rely on duration cues to distinguish English tense vowels from corresponding lax vowels in their perception.

Kim (2006) examined the perception of English high vowels including /i, ɪ, u, ʊ/ in the monosyllabic words by Korean elementary school students. The students were grouped as either low-level speakers or high-level speakers depending on the scores of speaking performance assessment. Korean listeners were asked to define the identity of the target vowels from stimuli out of four alternatives where no orthographic cues were provided. Kim reported that Korean speakers' accuracy rates were below chance level and they were good at discriminating English /u/-/ʊ/ pair whereas they were not able to distinguish English /i/ from /ɪ/. Kim suggests that the reason why Korean listeners performed worse at discriminating English /i/-/ɪ/ pair in their perception is related to the result of production test where Korean

speakers' vowel duration as well as spectral difference for this English pair were marginal in their production.

Kim (2010) investigated Korean college students' discrimination of English front vowels including /i/-/ɪ/ and /ɛ/-/æ/ pairs in hVd words. Korean listeners were supposed to choose a response from four alternatives. According to the results, Korean speakers performed well for English /i/-/ɪ/ distinction whereas their discrimination rates for English /ɛ/-/æ/ pair were below chance level. Kim claims that Korean speakers showed greater vowel length difference for English /i/-/ɪ/ pair than /ɛ/-/æ/ pair in their production and thus this larger vowel duration difference seemed to play an important role as a temporal cue for English /i/-/ɪ/ distinction in their English perception.

Hong (2012) examined the identification of English vowels including /i, ɪ, ɛ, æ, ɑ, ɔ, u, ʊ, ʌ/ in hVd words by Korean college students. They were asked to choose a response from nine alternatives. Hong calculated perceptual easiness scale for each vowel pair in various ways. First, Hong calculated perceptual easiness scale based on the accuracy rates for each English pair including /i/-/ɪ/, /ɛ/-/æ/, and /u/-/ʊ/. According to the results, English /i/-/ɪ/ pair was easier to discriminate than /ɛ/-/æ/ and /u/-/ʊ/ pairs whereas there was no significant difference between /ɛ/-/æ/ and /u/-/ʊ/ in terms of perceptual easiness. Second, Hong calculated perceptual easiness scale based on d-prime values which are relatively free from bias. As a result, English /i/-/ɪ/ pair was credibly easier to discriminate than /u/-/ʊ/ pair and in turn, /u/-/ʊ/ pair was significantly perceptually easier to distinguish than /ɛ/-/æ/ pair. Lastly, Hong calculated perceptual easiness scale for each vowel using F-score which is more independent of bias. Hong reported that English vowel /i/ was easier to perceive than /ɪ/ and similarly /ɛ/ was perceptually easier than /æ/. On the contrary, there was no significant difference between English /u/ and /ʊ/ in terms of perceptual easiness. As for perceptual easiness for each pair, the result was similar to that of d-prime indicating that English /i/-/ɪ/ pair was easier to discriminate

than /u/-/ʊ/ pair and in turn, /u/-/ʊ/ pair was perceptually easier to distinguish than /ɛ/-/æ/ pair. Hong suggests that vowel duration difference would not be an important cue for English vowel perception and Korean listeners' perception skills might have been affected by height of English vowels.

Cho and Jeong (2013) examined the perception of 11 English vowels including /i, ɪ, eɪ, ɛ, æ, ɑ, ɔ, oʊ, u, ʊ, ʌ/ in monosyllabic words by Korean college students. All the vowels except /ɑ, ʊ/ were spoken in bVt frame whereas /ɑ, ʊ/ were spoken in pVt frame by English native speakers and presented to Korean listeners aurally. Korean listeners were asked to choose the identity of the vowels for the target words out of eleven alternatives. As a result, Korean listeners showed 31%-89% of average accuracy rates for the identification test and they showed lots of bidirectional confusions in addition to some unidirectional confusions. For example, 36% of English tense /i/ was misheard as a counterpart /ɪ/ and 23% of /ɪ/ was misperceived as /i/. Similarly, Korean listeners misheard 56% of English tense /u/ as /ʊ/ and 29% of /ʊ/ as /u/. As for diphthong, /eɪ/ was misheard either as /i/ (21%) or as /ɪ/ (10%). Korean listeners performed better at labeling diphthongs and tense vowels than monophthongs and lax vowels except /i/-/ɪ/ pair. This result suggests that Korean listeners relied on temporal cues for their English vowel perception.

Lee and Cho (2015) investigated the perception of six English vowels including /i, ɪ, ɛ, æ, u, ʊ/ in hVd word by Korean college students. In the mapping test, Korean students were asked to categorize English vowels in terms of Korean vowels from sixteen alternatives. According to the results, English high front vowels including /i/ and /ɪ/ were mapped into a single Korean vowel /i/. Similarly, English high back vowels including /u/ and /ʊ/ were categorized as a single Korean vowel /u/. As for English mid front vowels including /ɛ/ and /æ/, they were both mapped into either Korean /e/ or Korean /ɛ/. In the identification test, Korean listeners were asked to identify the target English vowels from fifteen alternatives

consisting of English vowels. As a result, English vowels including /ɪ, æ, ʊ/ were much more accurately classified than vowels including /i, ɛ, u/. Authors suggest that English vowels such as /ɪ, æ, ʊ/ have no corresponding categories in Korean vowel system and thus a new phonetic category for those vowels need to be created. Therefore, in the process of generating this new category, vowels including /i, ɛ, u/ which already have been considered similar to Korean vowels were extremely replaced with those new vowels, and this process caused the identification rates for these similar vowels to be lower than new vowels.

Finally, Kim (2016) examined Korean high school students' discrimination of English front vowels including /i/-/ɪ/, /e/-/ɛ/, and /ɛ/-/æ/ pairs. Korean speakers were asked to choose one response from two alternatives. According to the results, Korean listeners' average accuracy rate was 77% ranging from 70% to 93%. More specifically, Korean listeners performed poor at discriminating /i/-/ɪ/ and /ɛ/-/æ/ pairs showing around chance level accuracies for each pair. This suggests that Korean speakers have a hard time in distinguishing English front tense-lax vowel pairs.

In sum, based on this previous research, it is expected that Korean speakers would have difficulty in perceiving English tense/lax contrasts in their English perception. Also, it is expected that Korean listeners would rely more on temporal cues than spectral cues for English tense/lax contrasts.

*The current study*

This study aims to examine whether Korean learners of English could discriminate and classify the coda consonant contrasts by exploiting vowel duration cues in their English perception. Also, it concurrently investigates whether intrinsic vowel durational differences in English would be perceived by the same Korean learners of English. In addition, it examines whether Korean learners with more English exposure performed better than those with less English

experience did in their English perception. In this study, we are also investigating whether Korean learners of English could perceive the coda consonant contrasts by only listening to the transitions out of the previous vowel when everything except the transition from the previous vowel is removed. The consonants chosen for the experiment in this study consist of sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/). The vowels selected for the experiment include English tense/lax vowel pairs /i, ɪ, e, ɛ, u, ʊ/ which cause high confusion to Korean learners of English because Korean has no tense/lax distinction.

Based on the previous research, the hypotheses for this study are made as follows:

**Hypothesis 1**: In perception, Korean L1 speakers would be able to perceive vowel durational difference depending on the following obstruent voicing not only in monosyllabic structure but also in disyllabic structure. This is because they showed laryngeal effects not only in final position but in non-final position in their English production. Alternatively, they would show better discrimination or identification of vowel durational difference due to the following obstruent voicing in disyllabic structure than in monosyllabic structure. This is because a coda distinction in obstruent voicing is neutralized in final position whereas the Korean contrast is preserved in medial position.

**Hypothesis 2**: Due to the lack of an intrinsic vowel quantity distinction in Seoul Korean, these Korean speakers would have difficulty in perception of English tense/lax vowel contrasts.

**Hypothesis 3**: Korean subjects would exhibit different L1 transfer effects on two different durational effects which include voicing effects (voiced vs. voiceless) and vowel quantity effects (tense vs. lax). This is because Korean learners harnessed durational differences as a cue more effectively for the tense/lax distinction than for following obstruent voicing in their English production. Alternatively, Korean listeners would be better at distinguishing consonant voicing contrasts than tense/lax vowel contrasts in their perception,

because the Korean participants failed to use spectral cues to distinguish English tense vowels from lax vowels in their productions.

## 2. Methods

### 2.1 Subjects

32 native Korean speakers who were from various regions in Korea participated in the experiment. 16 of them were recruited from Korea and they had less than a week of exposure to an English-speaking country (4 males, 12 females, age: 30 – 40, mean: 36 years old). The other 16 were recruited from Indiana University and they had more than three years of exposure to an English dominant environment (6 males, 10 females, age: 30 – 37, mean: 32 years old). The subjects' average duration of studying English at school in Korea was around 10 years. The subjects consisted of 10 male and 22 female speakers. As for English native control group, eight English native speakers were recruited from Indiana University. The listeners consisted of 6 male and 2 female speakers who were from the American Midwest and their age was in their late twenties.

### 2.2 Stimuli

The stimuli for perception tasks consist of the monosyllabic structure /bVC/ and disyllabic structure /bVCa/ using English vowels including /i, ɪ, e, ɛ, u, ʊ/ and sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/) as a coda for each vowel. To avoid lexical frequency effects, nonce words that were described in IPA were used to circumvent association with the orthographic properties of real words. It is worth noting that a subset of the nonce words closely resemble the orthography of actual lexical words, consisting of 24 instances. The same onset consonant /b/ was used for all the nonce words. The monosyllabic and disyllabic tokens were spoken in a carrier sentence, which is "say _ soon" by two English native speakers

and their voice was recorded in a quiet place individually. Two English native speakers were recruited from Indiana University. The speakers consisted of 1 male and 1 female speaker who were from the American Midwest and their average age was in their twenties. The male speaker in our study hailed from Columbia, Missouri, while the other female speaker originated from South Bend, Indiana. After the recording, the target English nonce words including 36 monosyllabic words and 36 disyllabic words (2 structures x 6 vowels x 6 consonants = 72 tokens) were excised from the carrier sentences and then presented to the subjects aurally. The English vowel duration depending on the following stop voicing and vowel formant frequencies (F1 and F2) of each target vowel spoken by English native speakers were presented in Table 1 and Table 2.

Table 1. Vowel duration

| Structure | Male speaker | | Female speaker | |
|---|---|---|---|---|
| | Voiced | Voiceless | Voiced | Voiceless |
| Monosyllabic structure | 213.3 ms | 118.0 ms | 198.7 ms | 118.0 ms |
| Disyllabic structure | 84.2 ms | 67.6 ms | 105.3 ms | 81.4 ms |

Table 2. Vowel formant frequencies

| Target vowel | Male speaker | | Female speaker | |
|---|---|---|---|---|
| | F1 | F2 | F1 | F2 |
| /i/ | 301.6 | 2175.7 | 387.1 | 2723.2 |
| /ɪ/ | 447.3 | 1774.1 | 485.3 | 2063.8 |
| /e/ | 492.0 | 1966.7 | 485.3 | 2254.4 |
| /ɛ/ | 606.9 | 1660.4 | 702.9 | 1842.5 |
| /u/ | 349.0 | 1069.7 | 428.0 | 1075.8 |
| /ʊ/ | 479.6 | 1224.1 | 461.9 | 1224.2 |

### 2.3 Procedure

#### 2.3.1 Discrimination task

In the ABX discrimination task, the subjects were instructed to indicate their response based on whether the last word matched the first or second word. Specifically, they were asked to press "First word" if the last word matched the first word and to press "Second word" if the last word matched the second word by clicking either “First word” or “Second word” on the

computer screen. The target words were excised from the carrier sentences. The perception task consists of two sessions.

In the first session, the excised target words were presented to subjects aurally. First two tokens were spoken by one speaker and the third token was spoken by another speaker as in Table 3. In the first block, it was investigated whether Korean subjects are able to discriminate vowel duration depending on the following consonant voicing. The second block examines whether Korean listeners are able to discriminate English tense vowels from lax vowels.

In the second session, the final consonant release was deleted for CVC words and the second syllable was deleted for CVCV words. As for CVC words, the release burst and transition of the final consonant was cut out whereas the release burst and transition of the word medial consonant was cut out for CVCV words. After the editing, neither a final stop for CVC nor a word medial stop and vowel were perceptible. The final consonants in parentheses indicate that they were removed as in Table 3. In this experiment, we aim to investigate whether Korean learners of English could perceive the coda consonant contrasts by only listening to the transitions out of the previous vowel when everything except the transition from the previous vowel is removed. The inter-stimulus interval was 1.0s and inter-trial interval was 5s. The stimuli were presented on the computer screen using Praat. Each of the triple pairs was presented once to the listeners. They were asked to identify the last stimulus as either the first stimulus (First word) or the second stimulus (Second word) by clicking either “First word” or “Second word” on the computer screen. The number of correct responses was counted.

Table 3. Discrimination task

1) First session
(1) First block (final consonant voicing contrasts)

| First word | Second word | Target word |
|---|---|---|
| bip | bib | bip |
| Male speaker | | Female speaker |

(2) Second block (vowel quantity contrasts)

| First word | Second word | Target word |
|---|---|---|
| bip | bɪp | bip |
| Male speaker | | Female speaker |

2) Second session
(1) First block (final consonant voicing contrasts)

| First word | Second word | Target word |
|---|---|---|
| bi(p) | bi(b) | bi(p) |
| Male speaker | | Female speaker |

(2) Second block (vowel quantity contrasts)

| First word | Second word | Target word |
|---|---|---|
| bi(p) | bɪ(p) | bi(p) |
| Male speaker | | Female speaker |

### 2.3.2 Identification task

In a separate identification task, Korean subjects were asked to identify the correct underlying form in each stimulus in terms of one of the four alternatives (including tense and lax vowels crossed by final consonant voicing) by clicking the appropriate word on the computer screen as in Table 4. The subjects were given the target nonce words that they were asked to identify and also were given English samples of real words with the same vowel under the target words. The IPA symbols representing each target vowel were included alongside the accompanying real word samples, allowing participants to make reference to them. For example, under the target word /bip/ there was a real English word such as *'eat',* which they could refer to. Each of the target word was aurally presented to the listeners once. The stimuli were excised from the carrier sentence.

In one block, the excised target words were presented and then the stimuli that were cut after the initial vowel were presented in another block. The final consonants in parentheses indicate that they were removed as in Table 4. The inter-trial interval was 7s. The number of correct responses was counted.

Table 4. Identification task

(1) First block

| Target word | Alternatives | | | |
|---|---|---|---|---|
| bɪb | bip | bib | bɪp | bɪb |
| | /i/ (eat) | /i/ (eat) | /ɪ/ (kid) | /ɪ/ (kid) |

(2) Second block

| Target word | Alternatives | | | |
|---|---|---|---|---|
| bɪ(b) | bi(p) | bi(b) | bɪ(p) | bɪ(b) |
| | /i/ (eat) | /i/ (eat) | /ɪ/ (kid) | /ɪ/ (kid) |

**2.4 Measurements & Analysis**

As for the ABX discrimination task, the number of correct responses was counted and each correct response was summed over and converted to percentages. The listeners' response scores for the discrimination task were converted into arcsine transformation measures and then analyzed. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

Repeated measures three-way ANOVAs were conducted using the R statistical package on the listeners' response scores. Statistical analyses were conducted separately for each task and each contrast. First, the listeners' response scores for the ABX discrimination task were analyzed. As for the discrimination of the final consonant voicing contrasts, the listeners' response scores which were converted into arcsine transformation measures were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor. As for the discrimination of the vowel quantity contrasts, the listeners' response

scores which were converted into arcsine transformation measures were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor.

Second, the listeners' response scores for the identification task were analyzed. As for the identification of the underlying final consonant, the listeners' response scores which were converted into d-prime measures were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor. As for the identification of the underlying vowel identity, the listeners' response scores which were converted into d-prime measures were submitted to an ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor.

A regression was run to examine the correlation between the discrimination of the final consonant voicing contrasts and the vowel quantity contrasts as well as the identification of the underlying final consonant and the underlying vowel identity of each subject. Further, the correlation between vowel length ratio in the production and the perception scores of each subject was investigated.

## 3. Results

### 3.1 Discrimination task

#### 3.1.1 Voicing contrasts (Voiced vs. Voiceless)

In this study, the discrimination of the final consonant voicing contrasts in the ABX discrimination task in different prosodic positions was examined to deterimine whether the Korean subjects are able to discriminate the final consonant voicing contrast in their English perception. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed.

First, the perception scores for the non-cut version of the ABX discrimination task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for the arcsine transformed perception scores. This analysis indicated that there was no significant main effect of group [$F(2, 37) = 0.739$, $p > 0.05$, $\eta^2 = 0.04$, [0.00, 1.00]] showing that the perception scores of each group were not significantly different from each other. There was no significant main effect of prosodic structure [$F(1, 37) = 1.365$, $p > 0.05$, $\eta^2 = 0.04$, [0.00, 1.00]] where the perception scores in the monosyllabic structure were not significantly different from those in the disyllabic structure. Prosodic structure interacted with group [$F(2, 37) = 4.181$, $p < 0.05$, $\eta^2 = 0.18$, [0.02, 1.00]] showing that Korean listeners did not show as extreme of a structure effect on the perception scores as English native listeners did. These results are plotted in Figure 9.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the English native listener group. After correcting for multiple testing using the Bonferroni correction, the difference remained statistically

significant only for the English native listener group (the inexperienced Korean group: $t$ = 1.0952, $df$ = 15, $p > 0.05$; the experienced Korean group: $t$ = -1.1839, $df$ = 15, $p > 0.05$; the native English control group: $t$ = 3.4419, $df$ = 7, $p < 0.001$).

Second, the perception scores for the cut version of the ABX discrimination task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for the arcsine transformed perception scores. This analysis indicated that there was no significant main effect of group [$F(2, 37) = 1.951$, $p > 0.05$, $\eta^2 = 0.10$, [0.00, 1.00]] showing that the perception scores of each group were not significantly different from each other. On the other hand, there was a significant main effect of prosodic structure [$F(1, 37) = 275$, $p < 0.001$, $\eta^2 = 0.64$, [0.48, 1.00]] where the perception scores in the monosyllabic structure were significantly higher than those in the disyllabic structure. Prosodic structure did not interact with group [$F(2, 37) = 2.605$, $p > 0.05$, $\eta^2 = 0.12$, [0.00, 1.00]]. These results are plotted in Figure 10.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words for all the groups. However, after correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant only for the English native listener group and the experienced Korean speaker group (the inexperienced Korean group: $t$ = 12.828, $df$ = 15, $p < 0.05$ (adjusted $p > 0.0167$); the experienced Korean group: $t$ = 9.4025, $df$ = 15, $p < 0.001$; the native English control group: $t$ = 7.8685, $df$ = 7, $p < 0.001$).

To summarize, Korean listeners discriminated English consonant voicing contrasts well not only in the monosyllabic structure but in the disyllabic structure in the non-cut version of

the discrimination task. There was an effect of prosodic structure only for English native listeners. On the contrary, all the listeners' discrimination scores were lower in the cut version of the discrimination task and the experienced Korean speaker group and the English native control group showed a significant effect of prosodic structure where the discrimination scores of English consonant voicing contrasts in final position were credibly higher than those in non-final position. Korean listeners with more English exposure did not show the credibly higher perception scores compared to those with less English exposure.

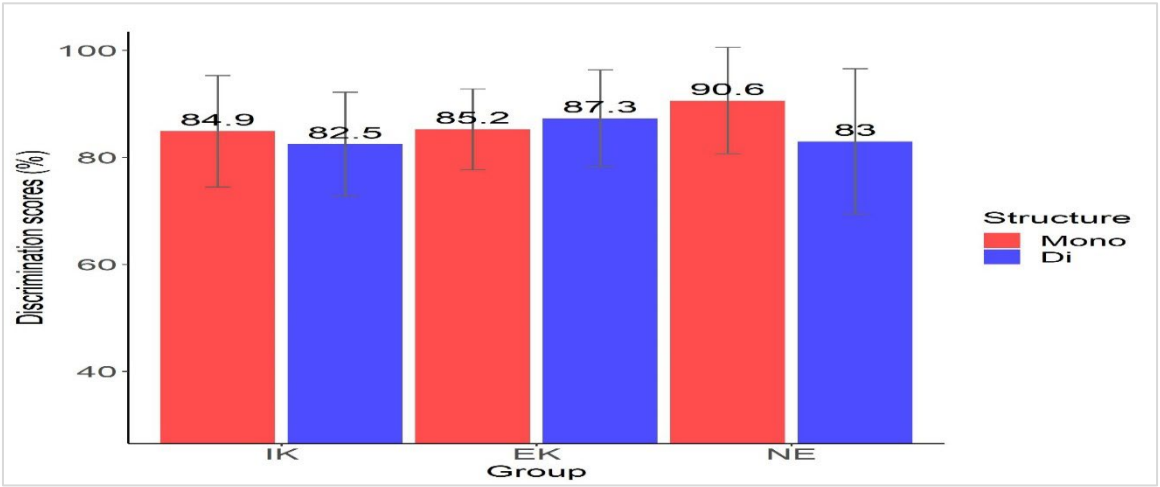


Figure 9. The perception scores for the non-cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

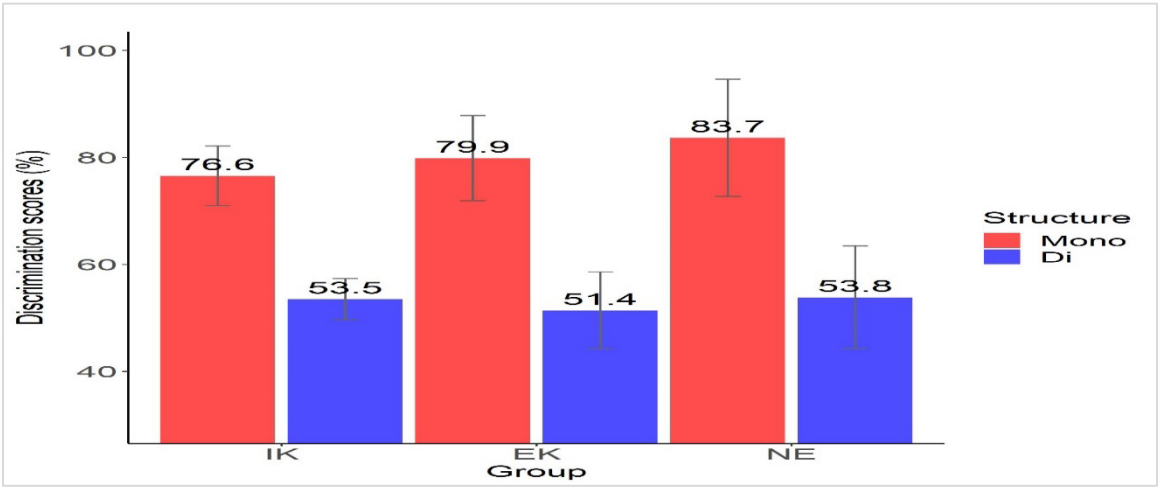


Figure 10. The perception scores for the cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.1.2 Vowel identity (Tense vs. Lax)

In this study, the discrimination of the vowel quantity contrasts in the ABX discrimination task in different prosodic positions was examined to see whether the Korean subjects are able to discriminate the vowel quantity contrast in their English perception. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed.

First, the perception scores for the non-cut version of the ABX discrimination task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for the arcsine transformed perception scores. This analysis indicated that there was a significant main effect of group [$F(2, 37) = 27$, $p < 0.001$, $\eta^2 = 0.59$, [0.41, 1.00]] showing that the perception scores of the English native listener group were significantly different from those of the Korean groups. There was no significant main effect of prosodic structure [$F(1, 37) = 3.672$, $p > 0.05$, $\eta^2 = 0.09$, [0.00, 1.00]] where the perception scores in the monosyllabic structure were not significantly different from those in the disyllabic structure. Prosodic structure did not interact with group [$F(2, 37) = 3.213$, $p > 0.05$, $\eta^2 = 0.15$, [0.00, 1.00]]. The results are plotted in Figure 11.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the English native listener group. After correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant for the English native listener group (the inexperienced Korean group: $t = 0.5093$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = 0.34218$, $df = 15$, $p > 0.05$; the native English control group: $t = 2.8427$, $df = 7$, $p < 0.001$).

Second, the perception scores for the cut version of the ABX discrimination task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for the arcsine transformed perception scores. This analysis indicated that there was a significant main effect of group [$F(2, 37) = 23.73$, $p < 0.001$, $\eta^2 = 0.56$, [0.37, 1.00]] showing that the perception scores of the English native listener group were significantly higher than those of the Korean groups. There was a significant main effect of prosodic structure [$F(1, 37) = 32.666$, $p < 0.001$, $\eta^2 = 0.47$, [0.27, 1.00]] where the perception scores in the monosyllabic structure were significantly higher than those in the disyllabic structure. Prosodic structure interacted with group [$F(2, 37) = 3.376$, $p < 0.05$, $\eta^2 = 0.15$, [0.00, 1.00]]. The results are plotted in Figure 12.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the inexperienced Korean listener group and the English native control group. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant only for the English native listener group and the inexperienced Korean speaker group (the inexperienced Korean group: $t = 5.0522$, $df = 15$, $p < 0.001$; the experienced Korean group: $t = 2.119$, $df = 15$, $p > 0.05$; the native English control group: $t = 2.3557$, $df = 7$, $p < 0.05$).

To summarize, Korean listeners did not discriminate English vowel quantity contrasts well not only in monosyllabic structure but in disyllabic structure in the non-cut version of the discrimination task. There was an effect of prosodic structure only for English native listeners. On the contrary, all the listeners' discrimination scores were slightly higher in monosyllabic structure in the cut version of the discrimination task. Only the Korean listener group with less

English exposure and the English native control group showed a significant effect of prosodic structure where the discrimination scores of English vowel contrasts in final position were credibly higher than those in non-final position. Korean listeners with more English exposure did not show the credibly higher perception scores compared to those with less English exposure.

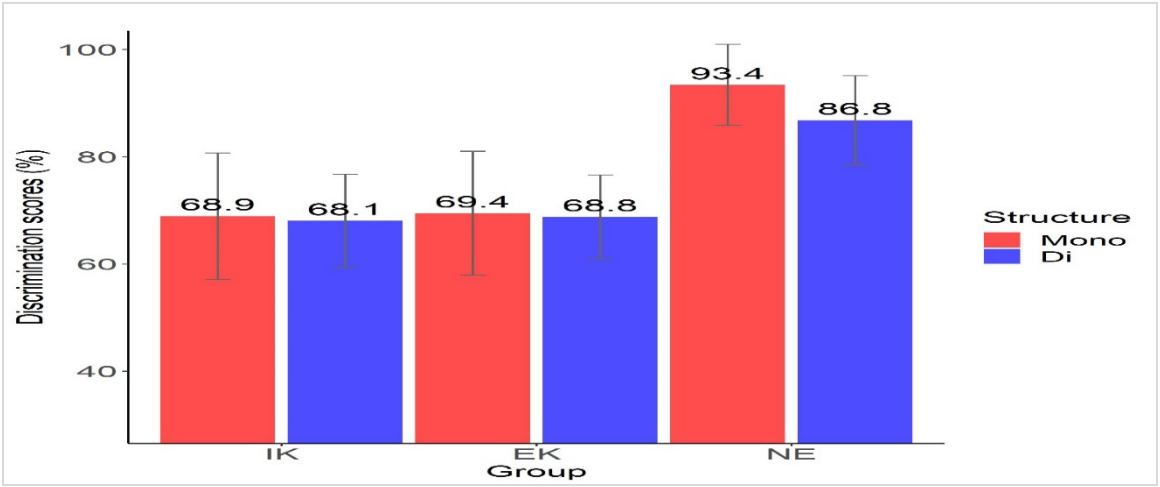


Figure 11. The perception scores for the non-cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

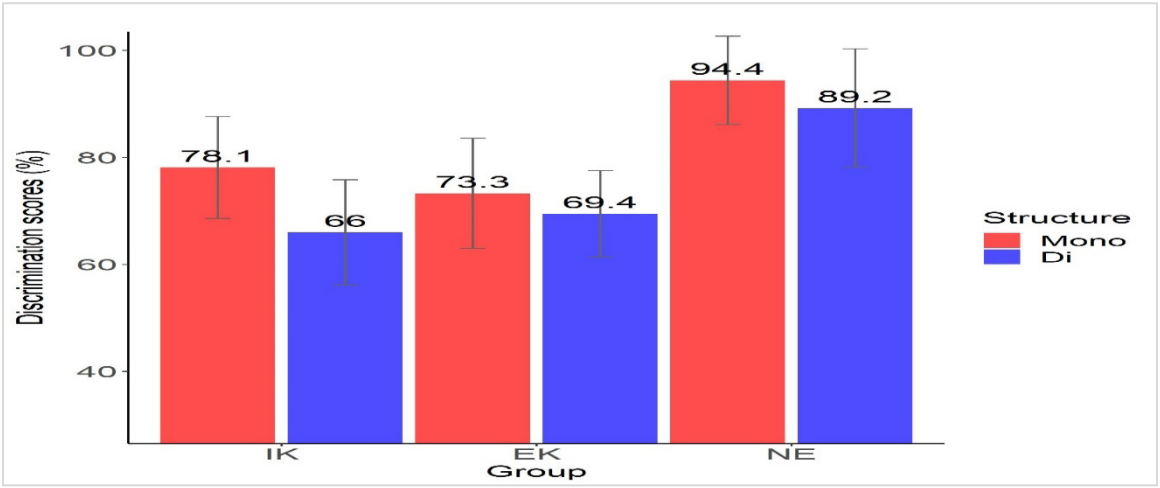


Figure 12. The perception scores for the cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

#### 3.1.2.1 Vowel identity (High front vowels: Tense vs. Lax)

In this section, the discrimination of the vowel quantity contrasts for high front vowels in the ABX discrimination task in different prosodic positions was examined to see whether the Korean subjects are able to discriminate the vowel quantity contrasts for high front vowels in their English perception. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed.

First, the perception scores for the non-cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the difference between mean perception scores in the monosyllabic words and those in the disyllabic words reached significance only for the English native control group. However, after correcting for multiple testing using the Bonferroni correction, none of the paired t-tests reached significance (the inexperienced Korean group: $t$ = 4.9648, $df$ = 15, $p > 0.05$; the experienced Korean group: $t$ = 0.95643, $df$ = 15, $p > 0.05$; the native English control group: $t$ = 1.9148, $df$ = 7, $p < 0.05$ (adjusted $p > 0.0167$)). The results are plotted in Figure 13.

Second, the perception scores for the cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the inexperienced Korean listener group and the English native control group. However, after correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant only for the inexperienced Korean speaker group (the inexperienced Korean group: $t$ = -0.72137, $df$ = 15, $p < 0.001$; the experienced Korean group: $t$ = 0.48435, $df$ = 15, $p > 0.05$; the native English control group: $t$ = 1.7552, $df$ = 7, $p < 0.05$ (adjusted $p > 0.0167$)). The results are plotted in Figure 14.

To summarize, Korean listeners did not discriminate English vowel quantity contrasts well not only in monosyllabic structure but in disyllabic structure in the non-cut version of the

discrimination task. There was no effect of prosodic structure for all the groups. On the contrary, Korean speakers' discrimination scores were slightly higher in monosyllabic structure in the cut version of the discrimination task. Only the Korean listener group with less English exposure showed a significant effect of prosodic structure where the discrimination score of English vowel contrasts in final position was credibly higher than that in non-final position. Korean listeners with more English exposure did not show the credibly higher perception scores compared to those with less English exposure.

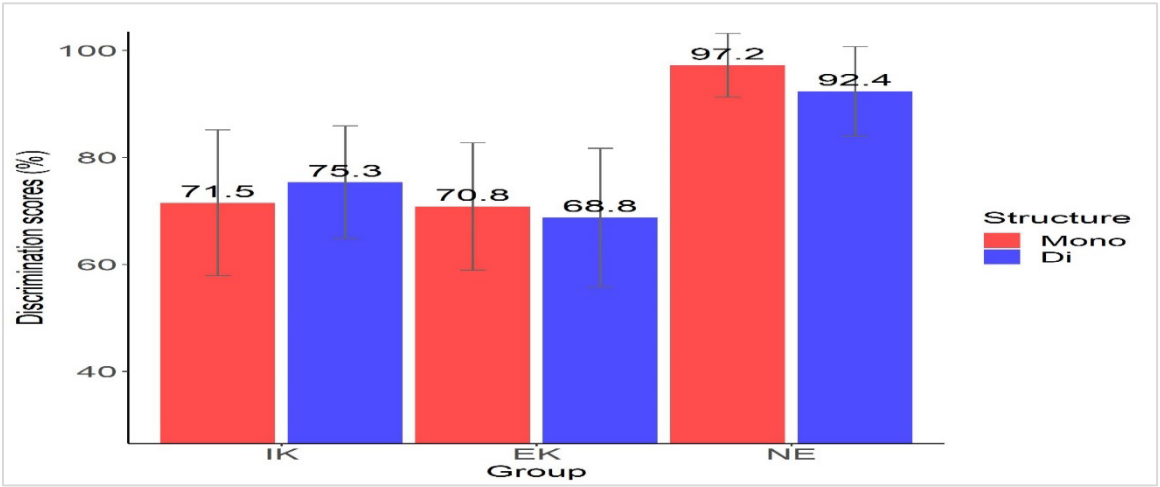


Figure 13. The perception scores for the non-cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

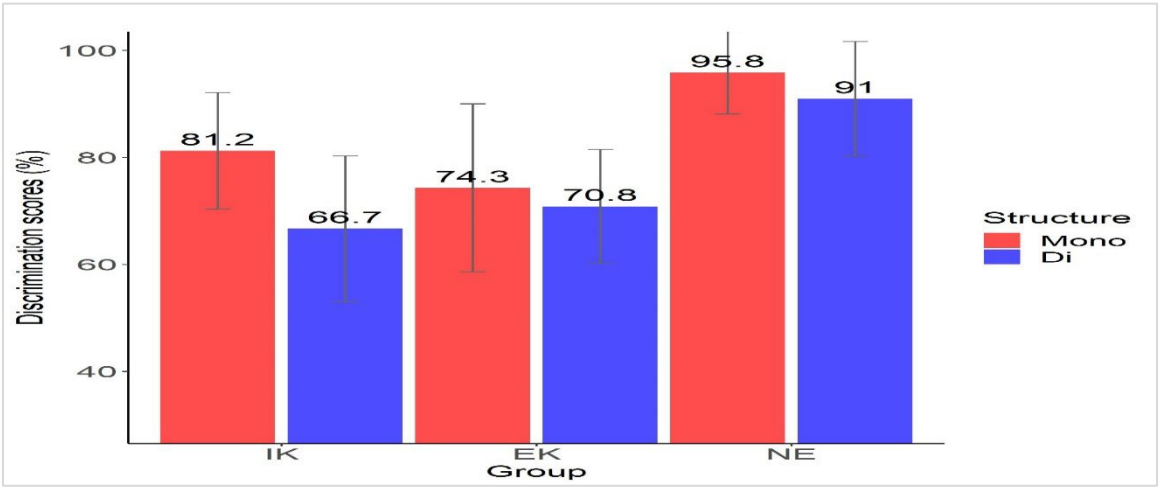


Figure 14. The perception scores for the cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.1.2.2 Vowel identity (Mid front vowels: Tense vs. Lax)

In this section, the discrimination of the vowel quantity contrasts for mid front vowels in the ABX discrimination task in different prosodic positions was examined to see whether the Korean subjects are able to discriminate the vowel quantity contrasts for mid front vowels in their English perception. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed.

First, the perception scores for the non-cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the English native listener group. After correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant for the English native listener group (the inexperienced Korean group: $t = 1.6563$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = 0.17696$, $df = 15$, $p > 0.05$; the native English control group: $t = 3.9898$, $df = 7$, $p < 0.001$). The results are plotted in Figure 15.

Second, the perception scores for the cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were not significantly higher than those in the disyllabic words for all the groups. After correcting for multiple testing using the Bonferroni correction, none of the paired t-tests reached significance (the inexperienced Korean group: $t = 1.3296$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = 1.6914$, $df = 15$, $p > 0.05$; the native English control group: $t = 0.69554$, $df = 7$, $p > 0.05$). The results are plotted in Figure 16.

To summarize, Korean listeners discriminated English vowel quantity contrasts relatively more successfully in monosyllabic structure than in disyllabic structure in the non-cut version of the discrimination task. There was an effect of prosodic structure only for English

native listeners. Similarly, all the listeners seemed to successfully discriminate vowel quantity contrasts in both prosodic structures in the cut version of the discrimination task. There was no significant effect of prosodic structure on perception scores for all the speaker groups. Korean listeners with more English exposure did not show the credibly higher perception scores compared to those with less English exposure.

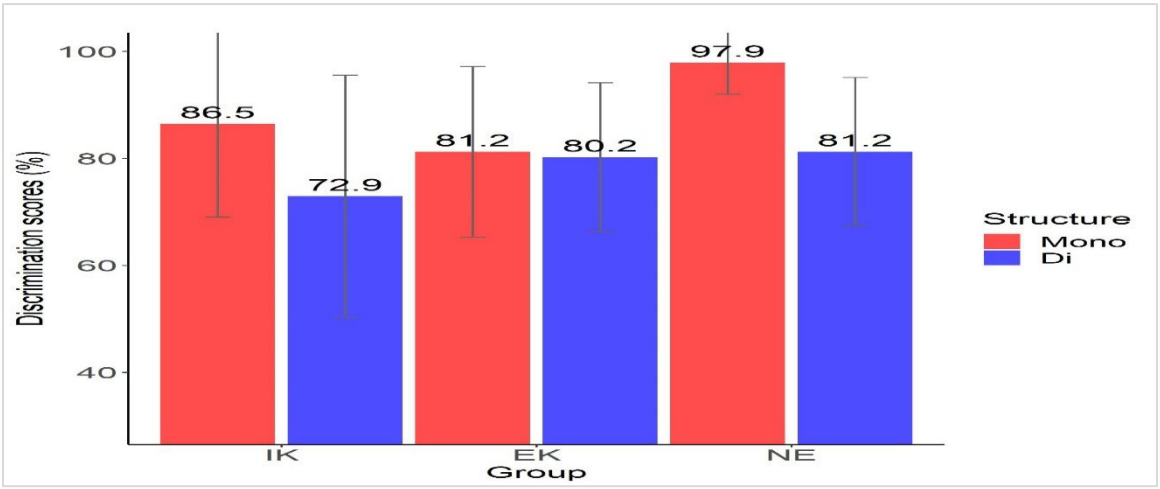


Figure 15. The perception scores for the non-cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

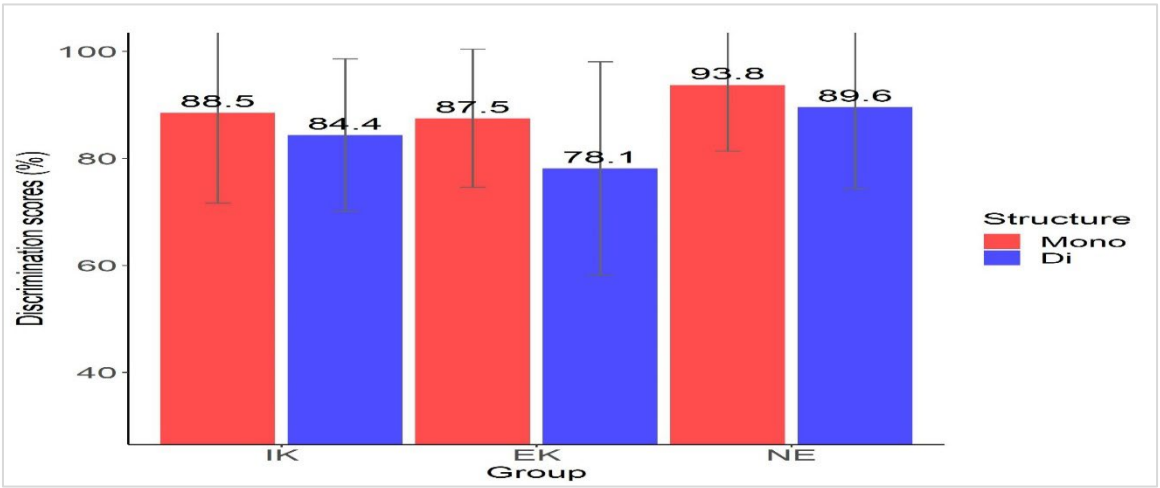


Figure 16. The perception scores for the cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.1.2.3 Vowel identity (High back vowels: Tense vs. Lax)

In this section, the discrimination of the vowel quantity contrasts for high back vowels in the ABX discrimination task in different prosodic positions was examined to see whether

the Korean subjects are able to discriminate the vowel quantity contrasts for high back vowels in their English perception. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed.

First, the perception scores for the non-cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the difference between the mean perception scores in the monosyllabic words and those in the disyllabic words did not reach significance for all the speaker groups. After correcting for multiple testing using the Bonferroni correction, none of the paired t-tests reached significance (the inexperienced Korean group: $t = 0.71248$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = -0.22791$, $df = 15$, $p > 0.05$; the native English control group: $t = 1.0077$, $df = 7$, $p > 0.05$). The results are plotted in Figure 17.

Second, the perception scores for the cut version of the ABX discrimination task were analyzed. A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the inexperienced Korean listener group and the English native control group. However, after correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant only for the English native listener group (the inexperienced Korean group: $t = 2.8667$, $df = 15$, $p < 0.05$ (adjusted $p > 0.0167$); the experienced Korean group: $t = 0.58405$, $df = 15$, $p > 0.05$; the native English control group: $t = 2.8504$, $df = 7$, $p < 0.001$). The results are plotted in Figure 18.

To summarize, Korean listeners did not discriminate English vowel quantity contrasts well not only in monosyllabic structure but in disyllabic structure in the non-cut version of the discrimination task. There was no significant effect of prosodic structure on the perception scores for all the speaker groups. Similarly, only the English native control group successfully discriminated English vowel quantity contrasts and only the English native control group

showed a significant effect of prosodic structure where the discrimination score of English vowel contrasts in final position was credibly higher than that in non-final position. Korean listeners with more English exposure did not show the credibly higher perception scores compared to those with less English exposure.

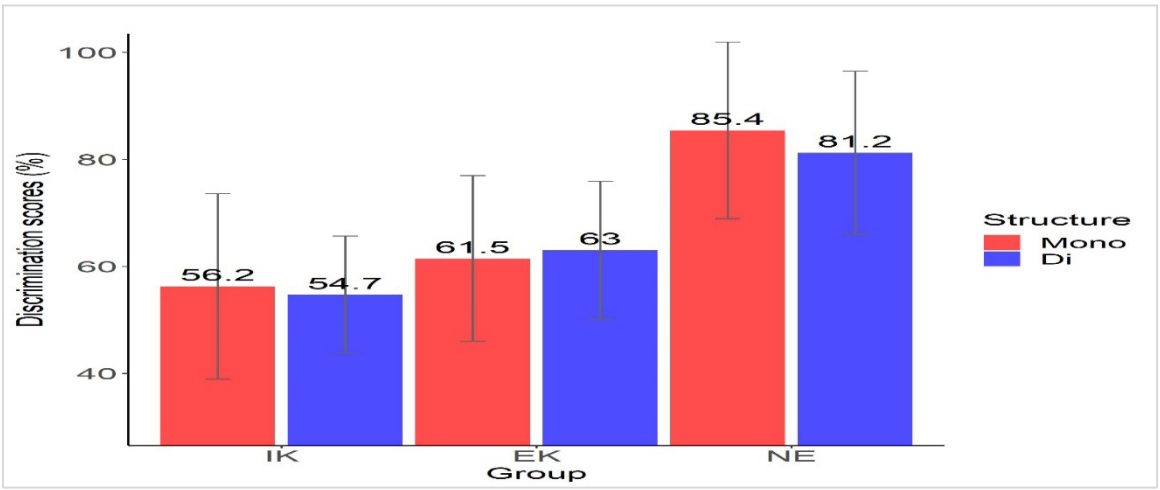


Figure 17. The perception scores for the non-cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

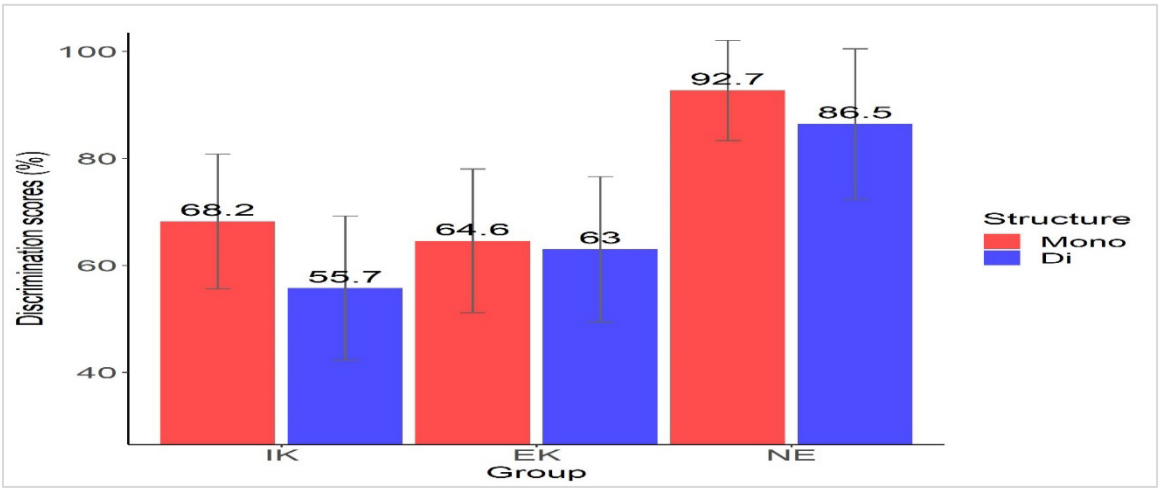


Figure 18. The perception scores for the cut version of the ABX discrimination task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.2 Identification task

#### 3.2.1 Voicing contrasts (Voiced vs. Voiceless)

In this study, the identification of the underlying final consonant in the identification task in different prosodic positions was examined to see whether the Korean listeners identify

the correct underlying final consonant in each stimulus. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

First, the perception scores for the non-cut version of the identification task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for d-prime measures. This analysis indicated that there were a significant main effect of group [$F(2, 37) = 7.958$, $p < 0.05$, $\eta^2 = 0.30$, [0.09, 1.00]] and a significant main effect of prosodic structure [$F(1, 37) = 11.937$, $p < 0.05$, $\eta^2 = 0.24$, [0.07, 1.00]]. Prosodic structure interacted with group [$F(2, 37) = 4.594$, $p < 0.05$, $\eta^2 = 0.20$, [0.02, 1.00]] showing that Korean speakers showed higher accuracy rates in disyllabic structure than in monosyllabic structure whereas the English native control group showed the opposite pattern. The results are plotted in Figure 19.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the difference between the mean perception scores in the disyllabic words and those in the monosyllabic words reached significance only for the experienced Korean listener group and the English native control group. However, after correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant only for the experienced Korean speaker group (the inexperienced Korean group: $t = -1.5751$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = -4.7675$, $df = 15$, $p < 0.001$; the native English control group: $t = 1.9459$, $df = 7$, $p < 0.05$ (adjusted $p > 0.0167$)).

Second, the perception scores for the cut version of the identification task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was

a within-subject factor was run for d-prime measures. This analysis indicated that there were a significant main effect of group [$F(2, 37) = 15.29$, $p < 0.001$, $\eta^2 = 0.45$, [0.24, 1.00]] and a significant main effect of prosodic structure [$F(1, 37) = 98.624$, $p < 0.001$, $\eta^2 = 0.73$, [0.60, 1.00]] where the perception scores in the monosyllabic structure were significantly higher than those in the disyllabic structure. Prosodic structure interacted with group [$F(2, 37) = 5.872$, $p < 0.05$, $\eta^2 = 0.24$, [0.05, 1.00]] showing that Korean speakers did not show as extreme of a structure effect on identification scores as English native speakers did. The results are plotted in Figure 20.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words for all the listener groups. After correcting for multiple testing using the Bonferroni correction, the differences remained statistically significant for all the speaker groups (the inexperienced Korean group: $t = 3.7456$, $df = 15$, $p < 0.001$; the experienced Korean group: $t = 6.7451$, $df = 15$, $p < 0.001$; the native English control group: $t = 8.6555$, $df = 7$, $p < 0.001$).

To summarize, Korean speakers' perception scores were high not only in monosyllabic structure but in disyllabic structure in the non-cut version of the identification task. The difference depending on the structure was significant only for Korean listeners with more English experience where the identification score of underlying English consonant in non-final position was credibly higher than that in non-final position. On the other hand, all the listener groups showed higher perception scores in monosyllabic structure than in disyllabic structure in the cut version of the identification task. In this case, the identification scores of English final consonants in final position were credibly higher than those in non-final position for all the listener groups. According to the results, English experience seemed not to enhance Korean

listeners' identification performance significantly not only in the non-cut version of the task but in the cut version of the task.

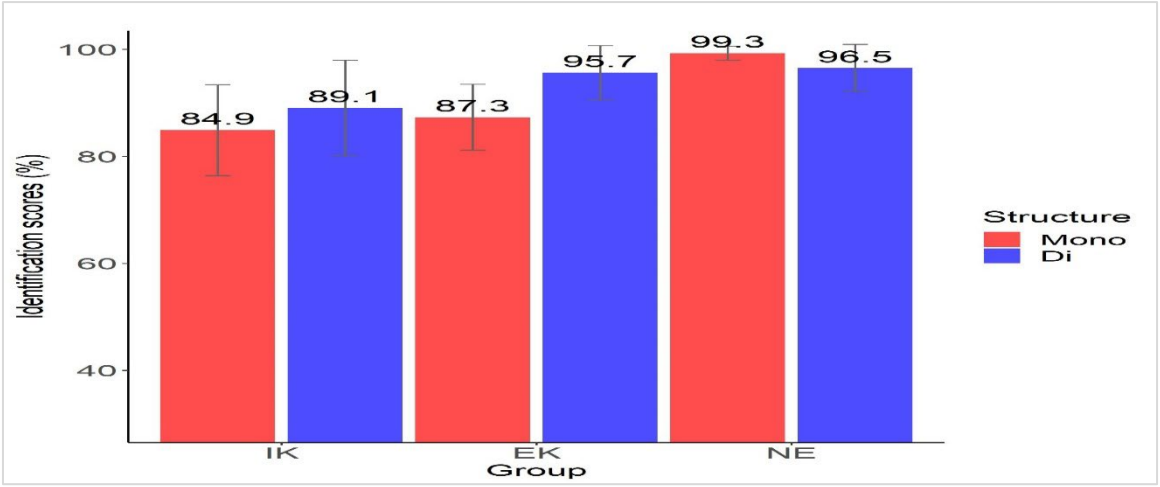


Figure 19. The perception scores for the non-cut version of the identification task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

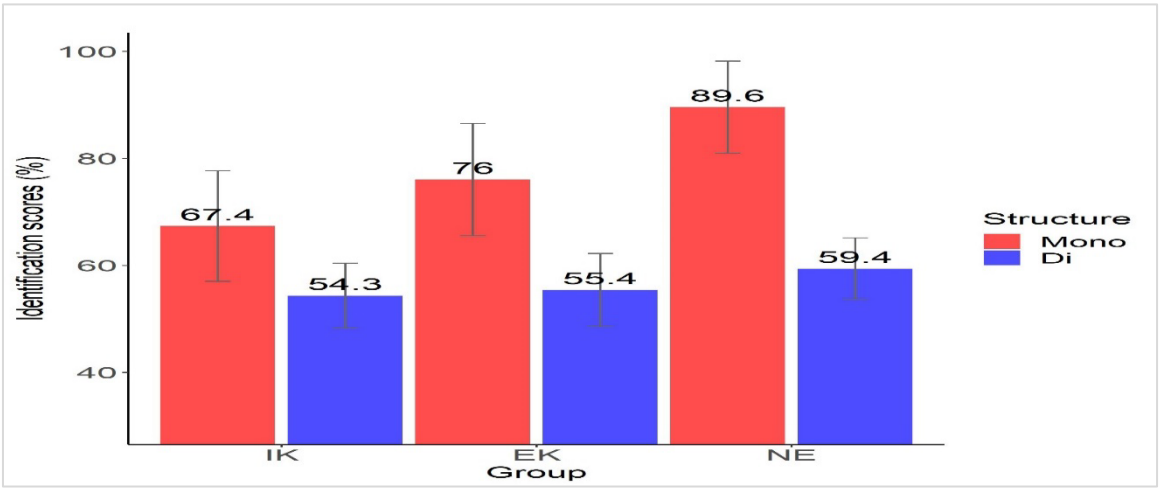


Figure 20. The perception scores for the cut version of the identification task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.2.2 Vowel identity (Tense vs. Lax)

In this study, the identification of the underlying vowel identity in the identification task in different prosodic positions was examined to see whether the Korean listeners identify the correct underlying vowel identity in each stimulus. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

First, the perception scores for the non-cut version of the identification task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for d-prime measures. This analysis indicated that there was a significant main effect of group [$F(2, 37) = 16.77$, $p < 0.001$, $\eta^2 = 0.48$, [0.27, 1.00]] showing that the perception scores of the English native listener group were significantly higher than those of the Korean groups. There was a significant main effect of prosodic structure [$F(1, 37) = 8.765$, $p < 0.05$, $\eta^2 = 0.19$, [0.04, 1.00]] where the perception scores in the monosyllabic structure were significantly higher than those in the disyllabic structure. Prosodic structure did not interact with group [$F(2, 37) = 0.390$, $p > 0.05$, $\eta^2 = 0.02$, [0.00, 1.00]]. The results are plotted in Figure 21.

A series of paired t-tests was run on different prosodic structures. After correcting for multiple testing using the Bonferroni correction, none of the paired t-tests reached significance (the inexperienced Korean group: $t = 2.7589$, $df = 15$, $p < 0.05$ (adjusted $p > 0.0167$); the experienced Korean group: $t = 1.6541$, $df = 15$, $p > 0.05$; the native English control group: $t = 0.6681$, $df = 7$, $p > 0.05$).

Second, the perception scores for the cut version of the identification task were analyzed. An ANOVA in which the amount of L2 exposure to English speaking country (less experienced Korean group vs. more experienced Korean group vs. English native control group) was a between-subject factor and prosodic structure (monosyllabic vs. disyllabic) was a within-subject factor was run for d-prime measure. This analysis indicated that there was a significant main effect of group [$F(2, 37) = 16.15$, $p < 0.001$, $\eta^2 = 0.47$, [0.26, 1.00]] showing that perception scores of the English native listener group were significantly higher than those of the Korean groups. There was a significant main effect of prosodic structure [$F(1, 37) =$

12.714, $p < 0.05$, $\eta^2 = 0.26$, [0.08, 1.00]] where the perception scores in the monosyllabic structure were significantly higher than those in the disyllabic structure. Prosodic structure did not interact with group [$F(2, 37) = 0.012$, $p > 0.05$, $\eta^2 = 6.55e\text{-}04$, [0.00, 1.00]]. The results are plotted in Figure 22.

A series of paired t-tests was run on different prosodic structures. According to the t-tests, the mean perception scores in the monosyllabic words were significantly higher than those in the disyllabic words only for the experienced Korean listener group and the English native control group. However, after correcting for multiple testing using the Bonferroni correction, the difference remained statistically significant only for the experienced Korean speaker group (the inexperienced Korean group: $t = 1.6179$, $df = 15$, $p > 0.05$; the experienced Korean group: $t = 4.4951$, $df = 15$, $p < 0.001$; the English native control group: $t = 1.9321$, $df = 7$, $p < 0.05$ (adjusted $p > 0.0167$)).

To summarize, Korean listeners did not identify underlying English vowel identity well not only in monosyllabic structure but in disyllabic structure in the non-cut version of the task as well as in the cut version of the task. All the speakers' perception scores were higher in monosyllabic structure than in disyllabic structure in both versions of the identification task. However, the difference depending on the structure was not significant for all the listeners in the non-cut version of the identification task. On the other hand, the identification score of English vowel identity in final position was credibly higher than that in non-final position only for Korean listeners with more English experience in the cut version of the identification task. According to the results, English experience seemed not to enhance Korean listeners identification performance not only in the non-cut version but in the cut version of the task.

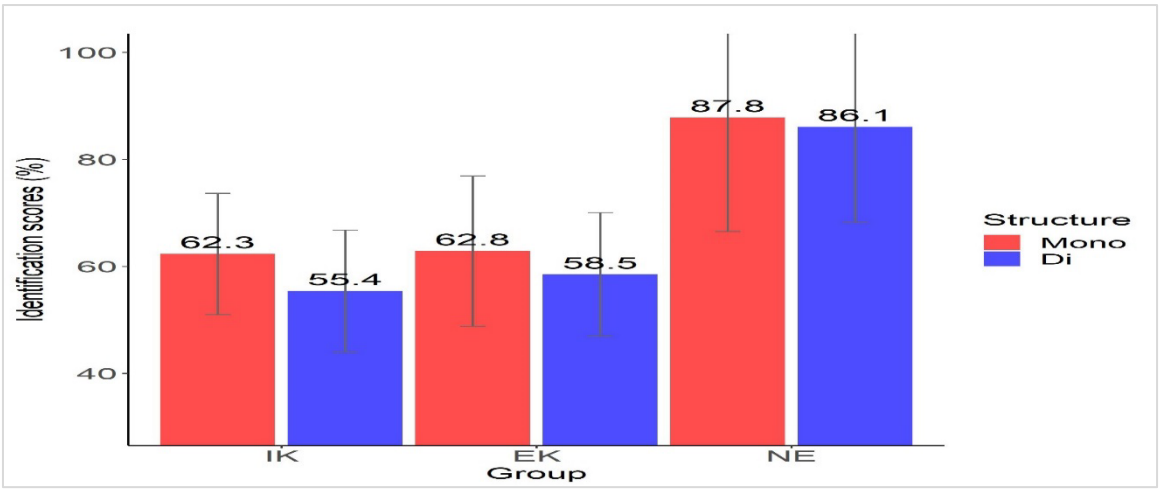


Figure 21. The perception scores for the non-cut version of the identification task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

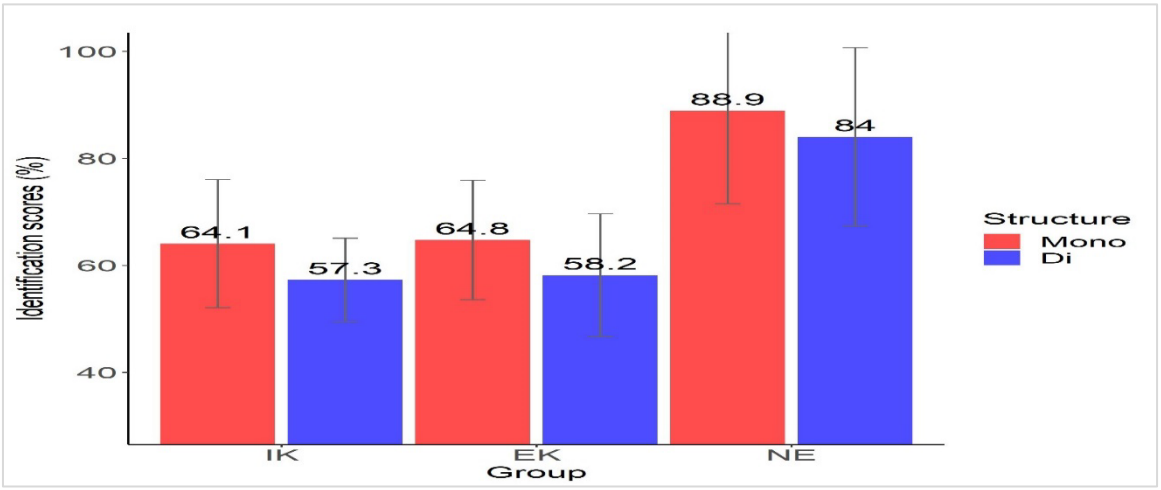


Figure 22. The perception scores for the cut version of the identification task in two prosodic structures (monosyllabic vs. disyllabic) by Korean listeners (IK and EK) and English listeners (NE). Error bars demonstrate standard deviation.

### 3.3 Correlation between accuracy in voicing contrasts vs. accuracy in vowel identity contrasts

In this section, the correlation between the discrimination scores of the final consonant voicing contrasts and those of the vowel quantity contrasts was examined. Further, the relationship between the identification scores of the underlying final consonant and those of the underlying vowel identity was examined. We aimed to see whether there is a certain relationship between these two perception scores by individual listeners. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation

measures and then analyzed. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

### 3.3.1 Discrimination scores

The relationship between the discrimination abilities of the final consonant voicing contrasts and those of the vowel quantity contrasts was examined to see whether there is a correlation between two different perception scores. For each participant, the average perception scores of the discrimination of each contrast in the ABX discrimination task were calculated and plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 23 below, there was little association between the discrimination scores of the final consonant voicing contrasts and those of the vowel quantity contrasts for the Korean groups. Inexperienced Korean speakers showed a weak ($r^2 = 0.30$), but statistically significant ($p < 0.05$) positive association between the discrimination accuracy rates of the final consonant voicing contrasts and those of the vowel quantity contrasts. Experienced Korean speakers showed a weaker ($r^2 = 0.08$), and statistically not significant ($p > 0.05$) positive association between two variables. On the contrary, the English native listener group showed a strong ($r^2 = 0.68$), and statistically significant ($p < 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables demonstrates that there were different L1 transfer effects on two different durational effects, such that Korean listeners who distinguish English consonant voicing contrasts may or may not distinguish English tense/lax contrasts in their English perception.

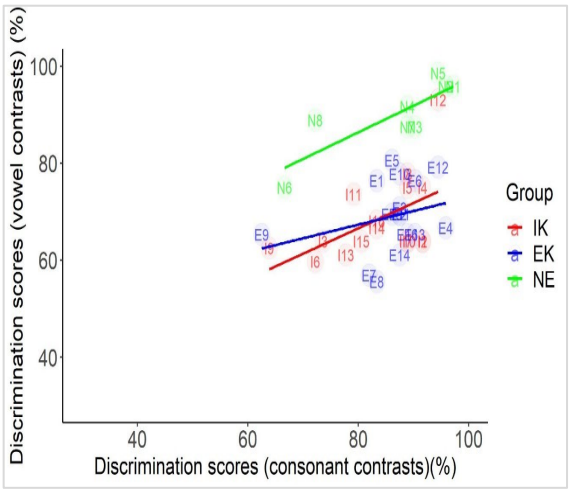

Figure 23. The correlation between the discrimination of the final consonant voicing contrasts and that of the vowel quantity contrasts by Korean listeners (IK and EK) and the English native control group (NE).

### 3.3.2 Identification scores

The relationship between the identification scores of the underlying final consonant and those of the underlying vowel identity was examined to see whether there is a correlation between two different perception scores. For each participant, the average perception scores of the identification of the underlying final consonant and those of underlying vowel identity in the identification task were calculated and plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 24 below, there was little association between the perception scores of the identification of the underlying final consonant and those of the underlying vowel identity for Korean groups. More specifically, inexperienced Korean listeners showed a weak ($r^2 = 0.10$), and statistically not significant ($p > 0.005$) positive association between two variables. Also, experienced Korean listeners showed a weak ($r^2 = 0.16$), and statistically not significant ($p > 0.05$) positive association between two variables. The English native listener group, however, showed a strong ($r^2 = 0.76$), and statistically significant ($p < 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables demonstrates that there were different L1 transfer effects on two different durational effects, such that Korean listeners who identify the correct underlying final consonant may or may not identify the correct underlying vowel identity in their English perception.

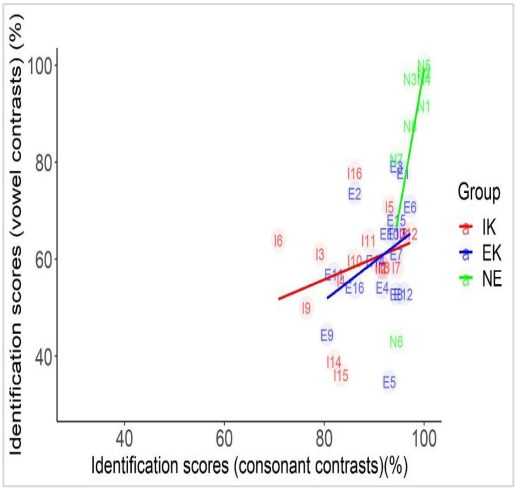


Figure 24. The correlation between the identification of the underlying final consonant and that of the underlying vowel identity by Korean listeners (IK and EK) and the English native control group (NE).

### 3.4 Correlation between accuracy in the discrimination task vs. accuracy in the identification task

In this study, the correlation between the discrimination accuracy and the identification accuracy of each subject was examined to see whether there is a certain relationship between these two different abilities by individual listeners. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

#### 3.4.1 Voicing contrasts (Voiced vs. Voiceless)

The relationship between the discrimination of the English final consonant voicing contrasts and the identification of the English underlying final consonant was examined to see

whether there is a correlation between two different perception scores. For each participant, the average perception scores of the discrimination of the final consonant voicing contrasts and those of the identification task were calculated and plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 25 below, the Korean groups showed the different patterns depending on their English experience. For example, inexperienced Korean speakers showed a strong ($r^2 = 0.53$), and statistically significant ($p < 0.05$) positive association between the discrimination of the final consonant voicing contrasts and the identification of the underlying final consonant. On the other hand, experienced Korean speakers showed a weak ($r^2 = 0.13$), and statistically not significant ($p > 0.05$) positive association between two variables. The English native listener group showed a strong ($r^2 = 0.58$), and statistically significant ($p < 0.05$) positive association between two variables.

The inexperienced Korean groups' strong correlation across the two variables demonstrates that there was a certain relationship between two different abilities, such that Korean listeners who discriminate English consonant voicing contrasts may identify the correct underlying English consonant in their English perception. Similarly, even though experienced Korean listeners showed a weak correlation across the two variables, it seems likely that Korean listeners who discriminated English consonant voicing contrasts identified the correct underlying English consonant in their English perception.

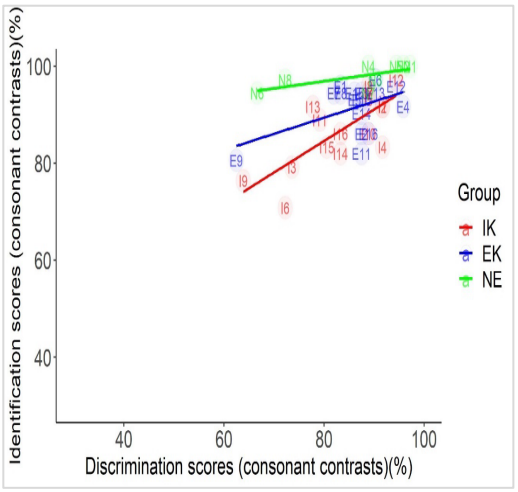

Figure 25. The correlation between the perception scores in the discrimination task and the perception scores in the identification task by Korean listeners (IK and EK) and the English native control group (NE).

### 3.4.2 Vowel identity (Tense vs. Lax)

The relationship between the discrimination of the English vowel quantity contrasts and the identification of the English underlying vowel was examined to see whether there is a correlation between two different perception abilities. For each participant, the average perception scores of the discrimination of the vowel quantity contrasts and those of the identification task were calculated and plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. According to Figure 26 below, there was little association between the perception scores of the discrimination of the vowel quantity contrasts and those of the identification of the underlying vowel for the Korean groups. Inexperienced Korean speakers showed a weak ($r^2 = 0.09$), and statistically not significant ($p > 0.05$) positive association between the discrimination of the vowel quantity contrasts and the identification of the underlying vowel. Experienced Korean speakers showed a weaker ($r^2 = 0.00$), and statistically not significant ($p > 0.05$) positive association between two variables. On the contrary, the English native listener group showed a strong ($r^2 = 0.67$), and statistically significant ($p < 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables demonstrates that there was no demonstrable relationship between two different abilities, such that Korean listeners who discriminate English vowel quantity contrasts may or may not identify the correct underlying English vowel in their English perception.

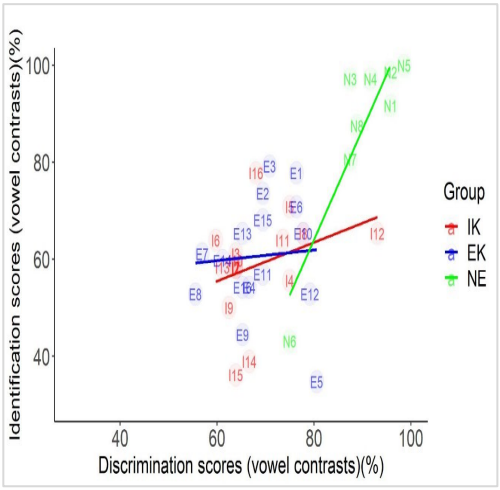


Figure 26. The correlation between the perception scores in the discrimination task and the perception scores in the identification task by Korean listeners (IK and EK) and the English native control group (NE).

### 3.5 Confusion matrix

Confusion matrices were made to examine the relationship between English voicing contrasts and English vowel quantity contrasts in the listeners' responses. The listeners' correct responses were summed over and converted to percentages.

#### 3.5.1 Identification task (Non-cut version)

Table 5 displays the mean scores of classifying English nonce words as one of the four nonce word alternatives of English depending on prosodic structure.

First, in the monosyllabic structure, all the words were identified as multiple words by Korean listeners. On the contrary, English native listeners often showed bidirectional confusion and only the word including a tense vowel in the voiceless context was identified as multiple words. When the target word has a tense vowel which is inherently long, all the listeners

showed higher accuracy rates for the word in the voiced context than that in the voiceless context. On the other hand, all the listeners showed higher accuracy rates for the word preceding the voiceless consonants than that preceding the voiced consonants when the target word has a lax vowel which is inherently short. This seems to be due to the fact that the word including a tense vowel in the voiced context and that with a lax vowel in the voiceless context are either extremely long or short, and thus they are less confused with other words, duration-wise. On the contrary, the word with a tense vowel in the voiceless context and that including a lax vowel in the voiced context were often classified as multiple words by Korean listeners. Korean listeners had a hard time identifying a tense vowel when the vowel duration is short due to the following voiceless consonants and they had a hard time identifying a lax vowel when the vowel duration is long due to the following voiced consonants. This suggests that English voicing contrasts and English vowel quantity contrasts have a certain relationship, such that changing the following consonant voicing induces more errors in identifying tense or lax for the target word. On the other hand, English voicing contrasts and English vowel quantity contrasts are independent due to the fact that the number of errors on the non-reverse diagonal were smaller than that on the reverse diagonal. More specifically, we have two-feature mistakes on the non-reverse diagonal where both features are misclassified. We have one-feature mistakes on the reverse diagonal where either English voicing identification or English vowel quantity classification is wrong. As we can see in Table 5, the number of two-feature mistakes is smaller than that of one feature mistakes, suggesting that two different features do not interact with each other.

As for the accuracy of each feature, the scores of all the listeners were lower for English vowel quantity identification than for English consonant voicing classification. This suggests that listeners had a harder time identifying tense or lax than voiced or voiceless for the perception task as in Table 6 and 7. As for the response bias for each feature, all the listeners

showed a tendency to identify the target vowel as tense rather than as lax for English vowel quantity contrasts and they exhibited a tendency to classify the target consonant voicing as voiced rather than as voiceless for English consonant voicing contrasts in the perception task as in Table 6 and 7.

Second, in the disyllabic structure, the scores of the listeners were generally lower than those in the monosyllabic structure and all the words were identified as multiple words by all the listeners. When the target word has a tense vowel which is inherently long, all the listeners showed higher accuracy rates for the word in the voiced context than that in the voiceless context. On the other hand, all the listeners showed higher accuracy rates for the word preceding the voiceless consonants than that preceding the voiced consonants when the target word has a lax vowel, which is inherently short. This seems to be due to the fact that the word including a tense vowel in the voiced context and that with a lax vowel in the voiceless context are either extremely long or short, and thus they are less confused with other words, duration-wise. On the contrary, the word with a tense vowel in the voiceless context and that including a lax vowel in the voiced context were often classified as multiple words by Korean listeners. Korean listeners had a hard time identifying a tense vowel when the vowel duration is short due to the following voiceless consonants and they had a hard time identifying a lax vowel when the vowel duration is long due to the following voiced consonants. This suggests that English voicing contrasts and English vowel quantity contrasts have a certain relationship, such that changing the following consonant voicing induces more errors in identifying tense or lax for the target word. On the other hand, English voicing contrasts and English vowel quantity contrasts are independent due to the fact that the number of errors on the non-reverse diagonal were smaller than that on the reverse diagonal. As we can see in Table 5, the number of one-feature mistakes is greater than that of two-feature mistakes, suggesting that two different features do not interact with each other for the mistakes.

As for the accuracy of each feature, the scores of all the listeners were lower for English vowel quantity identification than for English consonant voicing classification, suggesting that listeners had a harder time identifying tense or lax than voiced or voiceless for the perception task as in Table 6 and 7. As for the response bias for each feature, Korean listeners showed a tendency to identify the target vowel as lax than as tense for English vowel quantity contrasts whereas English native listeners showed the same pattern as in the monosyllabic structure, where the tendency to identify the target vowel as tense was greater for English vowel quantity contrasts. All the listeners exhibited a tendency to classify the target consonant voicing as voiceless rather than as voiced for English consonant voicing contrasts in the perception task as in Table 6 and 7.

Table 5. The mean scores of classifying English nonce words as one of the four nonce words of English (Non-cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | | | Di | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T+ | T- | L+ | L- | T+ | T- | L+ | L- |
| Tense+VD | IK | **61.8** | 10.4 | 21.5 | 6.3 | **46.5** | 2.8 | 46.5 | 4.2 |
| | EK | **67.4** | 7.6 | 20.8 | 4.2 | **51.0** | 3.4 | 44.8 | 0.7 |
| | NE | **91.7** | 0.0 | 8.3 | 0.0 | **93.2** | 1.4 | 4.1 | 1.4 |
| Tense+VL | IK | 9.0 | **47.2** | 9.0 | 34.7 | 6.9 | **43.8** | 4.9 | 44.4 |
| | EK | 14.6 | **52.8** | 3.5 | 29.2 | 1.4 | **45.8** | 4.2 | 48.6 |
| | NE | 1.4 | **86.1** | 1.4 | 11.1 | 4.2 | **76.4** | 0.0 | 19.4 |
| Lax+VD | IK | 44.4 | 6.3 | **43.1** | 6.3 | 35.4 | 10.4 | **43.8** | 10.4 |
| | EK | 52.8 | 6.3 | **35.4** | 5.6 | 37.1 | 1.4 | **58.0** | 3.5 |
| | NE | 20.8 | 0.0 | **79.2** | 0.0 | 21.1 | 1.4 | **74.6** | 2.8 |
| Lax+VL | IK | 4.9 | 23.6 | 8.3 | **63.2** | 2.1 | 30.6 | 2.1 | **65.3** |
| | EK | 4.9 | 27.1 | 4.2 | **63.9** | 1.4 | 27.8 | 1.4 | **69.4** |
| | NE | 0.0 | 6.9 | 0.0 | **93.1** | 1.4 | 6.9 | 1.4 | **90.3** |

Table 6. The mean scores of classifying English Tense/Lax vowels as intended (Non-cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | Stimuli | Group | Di | |
|---|---|---|---|---|---|---|---|
| | | T | L | | | T | L |
| T | IK | **32.1** | 17.9 | T | IK | **25.0** | 25.0 |
| | EK | **35.6** | 14.4 | | EK | **25.4** | 24.6 |
| | NE | **44.8** | 5.2 | | NE | **43.8** | 6.2 |
| L | IK | 19.8 | **30.2** | L | IK | 19.6 | **30.4** |
| | EK | 22.7 | **27.3** | | EK | 16.9 | **33.1** |
| | NE | 6.9 | **43.1** | | NE | 7.7 | **42.3** |
| Bias | IK | 26.0 | 24.0 | Bias | IK | 22.3 | 27.7 |
| | EK | 29.2 | 20.8 | | EK | 21.2 | 28.8 |
| | NE | 25.9 | 24.1 | | NE | 25.7 | 24.3 |

Table 7. The mean scores of classifying English consonant voicing as intended (Non-cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | Stimuli | Group | Di | |
|---|---|---|---|---|---|---|---|
| | | VD | VL | | | VD | VL |
| VD | IK | **42.7** | 7.3 | VD | IK | **43.1** | 6.9 |
| | EK | **44.1** | 5.9 | | EK | **47.7** | 2.3 |
| | NE | **50.0** | 0.0 | | NE | **48.3** | 1.7 |
| VL | IK | 7.8 | **42.2** | VL | IK | 4.0 | **46.0** |
| | EK | 6.8 | **43.2** | | EK | 2.1 | **47.9** |
| | NE | 0.7 | **49.3** | | NE | 1.7 | **48.3** |
| Bias | IK | 25.3 | 24.7 | Bias | IK | 23.5 | 26.5 |
| | EK | 25.4 | 24.6 | | EK | 24.9 | 25.1 |
| | NE | 25.3 | 24.7 | | NE | 25.0 | 25.0 |

### 3.5.2 Identification task (Cut version)

Table 8 displays the mean scores of classifying English nonce words as one of the four nonce word alternatives of English depending on prosodic structure. The patterns of the responses in the cut version of the task were different from those in the non-cut version.

First, in the monosyllabic structure, all the words were identified as multiple words by all the listeners. When the target word has a tense vowel which is inherently long, all the listeners showed higher accuracy rates for the word in the voiced context than that in the voiceless context. In addition, all the listeners showed higher accuracy rates for the word preceding the voiceless consonants than that preceding the voiced consonants when the target

word has a lax vowel which is inherently short. This seems to be due to the fact that the word including a tense vowel in the voiced context and that with a lax vowel in the voiceless context are either extremely long or short, and thus they are less confused with other words, duration-wise. On the contrary, the word with a tense vowel in the voiceless context and that including a lax vowel in the voiced context were highly classified as multiple words by Korean listeners. Korean listeners tended to identify the target word including a tense vowel in a voiceless context as well as that including a lax vowel in a voiced context as the word including a tense vowel in the voiced context. All the listeners seemed to pick the most frequent items when the stimulus was cut in half by guessing. English voicing contrasts and English vowel quantity contrasts are independent due to the fact that the number of errors on the non-reverse diagonal were smaller than that on the reverse diagonal. As we can see in Table 8, the number of one-feature mistakes is greater than that of two-feature mistakes, suggesting that two different features do not interact with each other for the mistakes.

As for the accuracy of each feature, the scores of all the listeners were generally lower for English vowel quantity identification than for English consonant voicing classification, suggesting that listeners had a harder time identifying tense or lax than voiced or voiceless for the perception task as in Table 9 and 10. As for the response bias for each feature, all the listeners showed a tendency to identify the target vowel as tense rather than as lax for English vowel quantity contrasts and Korean listeners exhibited a tendency to classify the target consonant voicing as voiced rather than as voiceless for English consonant voicing contrasts. On the other hand, English native listeners exhibited a tendency to classify the target consonant voicing as voiceless rather than as voiced for English consonant voicing contrasts in the perception task as in Table 9 and 10.

Second, the scores of the listeners in the disyllabic structure were much lower for the words in the voiced context than those in the monosyllabic structure and all the words were

identified as multiple words by all the listeners. The listeners showed higher accuracy rates for the target words in the voiceless context than those in the voiced context. For example, all the listeners showed the highest accuracy rates for the word including a lax vowel preceding the voiceless consonants and this seems to be due to the fact that as the target word is extremely short it is less confused with other words, duration-wise. On the contrary, other words were often classified as multiple words, especially those in the voiced context were mainly identified as the word preceding the voiceless consonants. This demonstrates that the listeners tended to pick a voiceless consonant when the end of the target word was cut off which caused voicing errors, and thus appeared to induce more errors in identifying tense or lax for the target word. This suggests that English voicing contrasts and English vowel quantity contrasts have a certain relationship, such that changing the following consonant voicing induces more errors in identifying tense or lax for the target word. On the other hand, English voicing contrasts and English vowel quantity contrasts are independent due to the fact that the number of errors on the non-reverse diagonal were smaller than that on the reverse diagonal except the word including a tense vowel in the voiced context. As we can see in Table 8, the number of one-feature mistakes is greater than that of two-feature mistakes, suggesting that two different features do not interact with each other for the mistakes.

As for the accuracy of each feature, the scores of all the listeners for English vowel quantity identification were not that different than those of English consonant voicing classification. This suggests that listeners had difficulty identifying tense or lax and voiced or voiceless for the perception task as in Table 9 and 10. As for the response bias for each feature, Korean listeners showed a tendency to identify the target vowel as lax than as tense for English vowel quantity contrasts whereas English native listeners showed an opposite pattern. All the listeners exhibited a tendency to classify the target consonant voicing as voiceless than as voiced for English consonant voicing contrasts in the perception task as in Table 9 and 10.

Table 8. The mean scores of classifying English nonce words as one of the four nonce words of English (Cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | | | Di | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | T+ | T- | L+ | L- | T+ | T- | L+ | L- |
| Tense+VD | IK | **52.8** | 20.1 | 16.7 | 10.4 | **25.7** | 28.5 | 13.2 | 32.6 |
| | EK | **65.3** | 11.8 | 17.4 | 5.6 | **21.4** | 23.4 | 18.6 | 36.6 |
| | NE | **88.9** | 9.7 | 0.0 | 1.4 | **21.9** | 64.4 | 2.7 | 11.0 |
| Tense+VL | IK | 31.9 | **37.5** | 13.2 | 17.4 | 17.4 | **33.3** | 13.2 | 36.1 |
| | EK | 25.7 | **39.6** | 14.6 | 20.1 | 16.0 | **26.4** | 11.8 | 45.8 |
| | NE | 16.7 | **73.6** | 0.0 | 9.7 | 6.9 | **77.8** | 0.0 | 15.3 |
| Lax+VD | IK | 39.6 | 12.5 | **33.3** | 14.6 | 15.3 | 24.3 | **17.4** | 43.1 |
| | EK | 52.1 | 8.3 | **28.5** | 11.1 | 9.8 | 19.6 | **14.7** | 55.9 |
| | NE | 15.3 | 5.6 | **72.2** | 6.9 | 2.8 | 21.1 | **18.3** | 57.7 |
| Lax+VL | IK | 13.9 | 20.1 | 13.9 | **52.1** | 13.2 | 22.9 | 10.4 | **53.5** |
| | EK | 6.9 | 16.0 | 11.8 | **65.3** | 9.0 | 16.0 | 6.3 | **68.8** |
| | NE | 1.4 | 11.1 | 0.0 | **87.5** | 0.0 | 11.1 | 1.4 | **87.5** |

Table 9. The mean scores of classifying English Tense/Lax vowels as intended (Cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | Stimuli | Group | Di | |
|---|---|---|---|---|---|---|---|
| | | T | L | | | T | L |
| T | IK | **35.6** | 14.4 | T | IK | **26.2** | 23.8 |
| | EK | **35.6** | 14.4 | | EK | **21.8** | 28.2 |
| | NE | **47.2** | 2.8 | | NE | **42.8** | 7.2 |
| L | IK | 21.5 | **28.5** | L | IK | 18.9 | **31.1** |
| | EK | 20.8 | **29.2** | | EK | 13.6 | **36.4** |
| | NE | 8.3 | **41.7** | | NE | 8.8 | **41.2** |
| Bias | IK | 28.6 | 21.4 | Bias | IK | 22.6 | 27.4 |
| | EK | 28.2 | 21.8 | | EK | 17.7 | 32.3 |
| | NE | 27.8 | 22.2 | | NE | 25.8 | 24.2 |

Table 10. The mean scores of classifying English consonant voicing as intended (Cut version). The target words heard as intended have boldface.

| Stimuli | Group | Mono | | Stimuli | Group | Di | |
|---|---|---|---|---|---|---|---|
| | | VD | VL | | | VD | VL |
| VD | IK | **35.6** | 14.4 | VD | IK | **17.9** | 32.1 |
| | EK | **40.8** | 9.2 | | EK | **16.1** | 33.9 |
| | NE | **44.1** | 5.9 | | NE | **11.4** | 38.6 |
| VL | IK | 18.2 | **31.8** | VL | IK | 13.5 | **36.5** |
| | EK | 14.8 | **35.2** | | EK | 10.8 | **39.2** |
| | NE | 4.5 | **45.5** | | NE | 2.1 | **47.9** |
| Bias | IK | 26.9 | 23.1 | Bias | IK | 15.7 | 34.3 |
| | EK | 27.8 | 22.2 | | EK | 13.4 | 36.6 |
| | NE | 24.3 | 25.7 | | NE | 6.8 | 43.2 |

## 4. Discussion

This study examined whether Korean learners of English could discriminate and identify the English coda consonant voicing contrasts in their English perception. Also, it concurrently investigates whether intrinsic vowel durational differences in English would be perceived by the same Korean learners of English. In addition, it examines whether Korean learners with more English exposure performed better than those with less English experience did in their English perception. The consonants chosen for the experiment in this study consist of sets of plosives contrasting in voicing (/p, b/, /t, d/, and /k, g/). The vowels selected for the experiment include English tense/lax vowel pairs /i, ɪ, e, ɛ, u, ʊ/ which cause high confusion to Korean learners of English because Korean has no tense/lax distinction.

First, it was expected that the Korean L1 speakers would be able to discriminate or identify vowel durational differences depending on the following obstruent voicing not only in monosyllabic structures but also in disyllabic structures. This is because they showed laryngeal effects not only in final position but also in non-final position in their English productions. Alternatively, it was expected that they would show better discrimination or identification of vowel durational differences due to the following obstruent voicing in disyllabic structure than in monosyllabic structure. This is because a coda distinction in obstruent voicing is neutralized in final position; whereas the Korean contrast is preserved in medial position.

According to the results, Korean listeners successfully discriminated and identified obstruent voicing not only in monosyllabic structures but also in disyllabic structures. Korean listeners' accuracy rates in disyllabic words were not significantly higher than those in monosyllabic words in the discrimination task or in the identification task. On the other hand, English native speakers consistently showed higher accuracy rates in monosyllabic structures than in disyllabic structures in both tasks. Korean speakers resembled English native speakers in terms of the discrimination and identification performance.

When the portion of the acoustics including the consonant release and following material was removed, the perception scores of all the listeners decreased in both monosyllabic and disyllabic structures. Especially those in disyllabic structures went down notably and were at an approximately chance level for all the speaker groups. The reason why the accuracy rates in disyllabic structure greatly went down in both tasks seems to be because allophonic variation of English obstruents in non-final position and in final position was disturbed as disyllabic words were edited into monosyllabic words artificially with the release taken out. It seems likely that all the listeners paid more attention to the release of consonants in the following syllable rather than the vowel duration preceding obstruents to distinguish English consonant voicing contrasts in their perception. Also it is possible that listeners were relying on the duration of the obstruent closure, which was also made inaudible by removing the consonant releases. And thus, as the release was removed, they seemed to have difficulty in attending to the signal where the surface form and the prosodic location were jointly interpreted.

English experience seemed not to enhance Korean listeners' performance significantly not only in the discrimination task but in the identification task.

It suggests that Korean speakers utilized the release of obstruents more effectively than vowel duration preceding obstruents to distinguish English obstruent voicing contrasts in their perception. It appeared that vowel duration difference depending on the following obstruent voicing produced by Korean speakers in their production did not help them to discriminate and identify English consonant voicing contrasts in their perception. When it comes to Korean coda neutralization rules, they seemed not to have an influence on Korean listeners' discrimination and identification of obstruent voicing in their English perception. Korean speakers performed well in both prosodic structures not only in the discrimination task but in the identification task. Further, when the release was removed, Korean speakers' perception scores went down more dramatically in disyllabic structures than in monosyllabic structures. As a result, their accuracy

rates in monosyllabic structures became significantly higher than those in disyllabic structures. This issue seems not unique to non-native speakers because Korean speakers resembled English native speakers in terms of perception performance patterns. More specifically, when the release was removed, English native speakers' perception scores in both prosodic structures went down. However, their perception scores in disyllabic structures decreased more sharply than those in monosyllabic structures. As a result, their accuracy rates in monosyllabic structures became significantly higher than those in disyllabic structures.

Second, it was expected that Korean speakers would have difficulty in perception of English tense/lax vowel contrasts due to the lack of an intrinsic vowel quantity distinction in Korean. According to the results, as expected, Korean listeners showed poor performance not only in the discrimination task but in the identification task; whereas English native speakers performed well in both tasks. Korean listeners seemed to have difficulty in discriminating and identifying vowel durational differences depending on the vowel quantity contrasts not only in monosyllabic structures but also in disyllabic structures. When comparing the discrimination scores of each vowel pair, it was observed that the discrimination scores for the high back tense/lax vowel pair were relatively lower than those of the high front and mid front tense/lax vowel pairs. This pattern implies that the relatively lower frequency of the /u/ phoneme may have contributed to the lower discrimination scores for the high back tense/lax vowel pair. Further, Korean listeners might have failed to utilize spectral cues to perceptually distinguish English tense/lax vowel pairs, which are preferred over vowel durational cues by English native speakers for English tense/lax vowel distinction (Bohn & Flege, 1992). Additionally, it is worth noting that the high back tense/lax vowel pair has a smaller acoustic space between the two vowels than other vowel pairs. This reduced acoustic distance may have made it more challenging for Korean listeners to effectively exploit spectral cues to successfully discriminate these pairs. As a result, Korean listeners might have over-relied on vowel durational cues rather

than spectral cues, which seemed not to help Korean speakers to discriminate or identify English tense/lax vowel contrasts in their perception. It appeared that vowel duration differences depending on the vowel quantity contrasts produced by Korean speakers in their production did not help them to discriminate and identify English tense/lax vowel contrasts in their perception.

When removing the release, Korean listeners showed different patterns depending on the task. In the discrimination task, Korean listeners' accuracy rates increased for monosyllabic structure whereas those in disyllabic structure remained almost the same. It appeared that taking out the release and following material helped Korean listeners concentrate more on vowel itself in the first syllable compared to when there was additional information in the second syllable. That information might have rather prevented Korean speakers from focusing on the vowel to distinguish English vowel contrasts. Further, systematically longer vowel durations in monosyllabic words than in disyllabic words might have played a role for Korean speakers to discriminate English tense vowels from lax vowels better in monosyllabic words than in disyllabic words. On the contrary, in the identification task, taking out the release seemed not affect Korean speakers' perception performance, not just in monosyllabic structures but in disyllabic structures, as well. There was no larger vowel identity effect for Korean subjects with more English experience than those with less English experience.

Third, it was expected that Korean subjects would exhibit different L1 transfer effects on the two different durational effects, voicing effects (voiced vs. voiceless) and vowel quantity effects (tense vs. lax). This is because Korean learners harnessed durational differences more effectively for the tense/lax distinction than for following obstruent voicing in their English production. Alternatively, Korean listeners would be better at distinguishing consonant voicing contrasts than tense/lax vowel contrasts in their perception, because Korean participants failed to use spectral cues to distinguish English tense vowels from lax vowels in their productions.

According to the perceptual results, Korean listeners showed different patterns for the two different durational effects. When it comes to the consonant voicing contrasts, Korean listeners performed well and resembled the patterns of the English native speaker group in the discrimination task as well as in the identification task. On the other hand, Korean listeners seemed to have difficulty in discriminating and identifying English tense/lax vowel contrasts in their perception. They showed much lower perception scores than did English native speakers for English vowel distinction. This seems to be due to the fact that Korean listeners might have failed to use spectral cues and possibly over-relied on vowel durational cues to distinguish English tense/lax vowel contrasts in their perception, which seemed not help Korean learners to discriminate and identify English vowel contrasts. English native speakers, however, utilized spectral cues effectively for English vowel contrasts in perception.

Removing the consonant release seems to have a different influence on the two different durational effects. Discrimination scores for consonant voicing contrasts went down whereas accuracy rates for tense/lax vowel contrasts went up in monosyllabic structures when the release was cut off. It seems that cutting off the release diminished Korean speakers' discrimination ability for English obstruent voicing contrasts in their perception. On the other hand, it seems to enhance Korean speakers' concentration on the vowel portion for discriminating English tense/lax vowel contrasts.

When it comes to the effect of L2 English experience on the perception scores, English experience marginally enhanced Korean speakers' performance for English consonant voicing contrasts; whereas there seems to be no influence of L2 English exposure on the performance for English tense/lax vowel contrasts.

Overall, higher perception scores for English consonant voicing contrasts than for English tense/lax vowel contrasts by Korean speakers indicate that Korean learners were better at distinguishing English consonant voicing contrasts than English vowel contrasts in their

perception. The results imply that Korean speakers have acquired discrimination and identification skills for English consonant voicing contrasts whereas they have not yet acquired those skills for English tense/lax vowel contrasts in their English perception. This seems to be because it might have been easier for Korean speakers to utilize the temporal cues, which is the duration of the obstruent closure to distinguish English consonant voicing contrasts. On the other hand, it might have been difficult to use spectral cues which play an important role in addition to vowel duration cues for distinguishing English vowel contrasts. Korean does not have tense/lax vowel distinction and thus difficulty might be imposed for Korean speaker groups. More specifically, the high English vowel pair /i/ vs. /ɪ/ would assimilate to the high front vowel /i/ in Korean. Also, there is only one high back vowel /u/ in Korean, which ranges over the entire high back region. Therefore, their L1 experience might impose difficulty for Korean speakers to distinguish between English tense and lax contrasts. Korean speakers' failure of using spectral cues to distinguish English tense/lax contrasts was consistent with Bohn's (1995) Desensitization hypothesis that L2 learners' sensitivity to spectral differences will be lesser when they have only one vowel in the L1. Further, Choi et al. (2016) claimed that the reason why Korean learners of English relied more on temporal cues in their study was due to the fact that they were universally more salient and that Korean speakers with a smaller vowel space were less sensitive to small spectral changes than English native speakers with a larger vowel space.

## 5. Correlation between production and perception

With respect to production and perception, a number of researchers have suggested that perceptual development tends to precede segmental production development whereby L1 acquisition is recapitulated by L2 acquisition (Aoyama, 2003; Flege, 1989, 1995; Kim, 2016; Kim & Shin, 2014; Polka, 1991). Ingram and Park (1997) expected that L2 learners' foreign

vowel productions will mirror perception of those vowels in terms of cross-language differences, claiming that learners with better discriminability of L2 vowels are expected to have production accuracy in those vowels. On the contrary, some researchers suggested that L2 segmental perception follows L2 segmental production (Tsukada et al., 2005). However, perception and production are inherently incommensurable, and thus it is not easy to directly compare the attainment extent of native-like L2 production and perception in the data (Cho & Jeong, 2013; Flege, Bohn & Jang, 1997; Kim, 2010).

In this study, the relationship between the production and either the discrimination or the identification of the two different contrasts including final consonant voicing contrasts and the vowel quantity contrasts was examined to see whether there is a certain relationship between these two different dynamics conducted by Korean participants. The listeners' response scores for the ABX discrimination task were converted into arcsine transformation measures and then analyzed. The listeners' response scores for the identification task were converted into d-prime measures and then analyzed.

## 5.1 Discrimination task

### 5.1.1 Voicing contrasts (Voiced vs. Voiceless)

The relationship between the discrimination of the final consonant voicing contrasts and the ratio of vowel length in the voiced context to that in the voiceless context by Korean listeners was examined to see whether there is a correlation between two different dynamics. So, for each participant, the average perception scores of the discrimination of the final consonant voicing contrasts and the ratio of vowel length in the voiced context to that in the voiceless context were calculated and they were plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 27 below, there was

little association between the perception scores of the discrimination of the final consonant voicing contrasts and the ratio of vowel length in the voiced context to that in the voiceless context for Korean listeners. Inexperienced Korean speakers showed a weak ($r^2 = 0.07$), and statistically not significant ($p > 0.05$) positive association between the discrimination of the final consonant voicing contrasts and the ratio of vowel length in the voiced context to that in the voiceless context. Experienced Korean speakers showed a weaker ($r^2 = 0.01$), and statistically not significant ($p > 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables does not demonstrate a relationship between the discrimination of English voicing contrasts and the production of vowel duration difference depending on the following consonant voicing, such that Korean listeners who distinguish English consonant voicing contrasts in their English perception may or may not distinguish English consonant voicing contrasts in their English production.

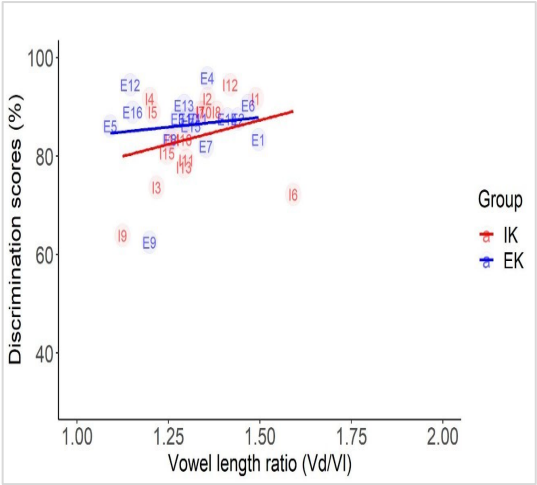


Figure 27. The correlation between the discrimination of the final consonant voicing contrasts and the ratio of vowel length in the voiced context to that in the voiceless context by Korean listeners.

### 5.1.2 Vowel identity (Tense vs. Lax)

The relationship between the discrimination of the vowel quantity contrasts and the ratio of vowel length of tense to that of lax by Korean listeners was examined to see whether there is a correlation between two different dynamics. So, for each participant, the average

perception scores of the discrimination of the vowel quantity contrasts and the ratio of vowel length of tense to that of lax were calculated and they were plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 28 below, there was little association between the perception scores of the discrimination of the vowel quantity contrasts and the ratio of vowel length of tense to that of lax for Korean listeners. Inexperienced Korean speakers showed a weak ($r^2$ = 0.01), and statistically not significant ($p > 0.05$) negative association between the discrimination of the vowel quantity contrasts and the ratio of vowel length of tense to that of lax. Experienced Korean speakers showed a weak ($r^2$ = 0.07), and statistically not significant ($p > 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables does not demonstrate that there is a relationship between the discrimination of English vowel quantity contrasts and the production of vowel duration difference depending on the vowel quantity, such that Korean listeners who distinguish English vowel quantity contrasts in their English perception may or may not distinguish English vowel quantity contrasts in their English production.

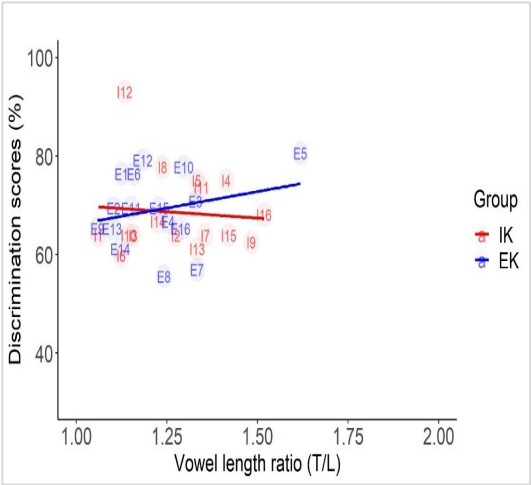


Figure 28. The correlation between the discrimination of the vowel quantity contrasts and the ratio of vowel length of tense to that of lax by Korean listeners.

## 5.2 Identification task

### 5.2.1 Voicing contrasts (Voiced vs. Voiceless)

The relationship between the identification of the underlying final consonant and the ratio of vowel length in the voiced context to that in the voiceless context by Korean listeners was examined to see whether there is a correlation between two different dynamics. So, for each participant, the average perception scores of the identification of the underlying final consonant and the ratio of vowel length in the voiced context to that in the voiceless context were calculated and they were plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 29 below, there was little association between the perception scores of the identification of the underlying final consonant and the ratio of vowel length in the voiced context to that in the voiceless context for Korean listeners. Inexperienced Korean speakers showed a weak ($r^2 = 0.02$), and statistically not significant ($p > 0.05$) positive association between the identification of the underlying final consonant and the ratio of vowel length in the voiced context to that in the voiceless context. Similarly, experienced Korean speakers showed a weak ($r^2 = 0.06$), and statistically not significant ($p > 0.05$) positive association between two variables.

Korean listeners' weak correlation across the two variables does not demonstrate that there is a relationship between the identification of the underlying final consonant and the production of vowel duration difference depending on the following consonant voicing, such that Korean listeners who identify the underlying final consonant in their English perception may or may not distinguish English consonant voicing contrasts in their English production.

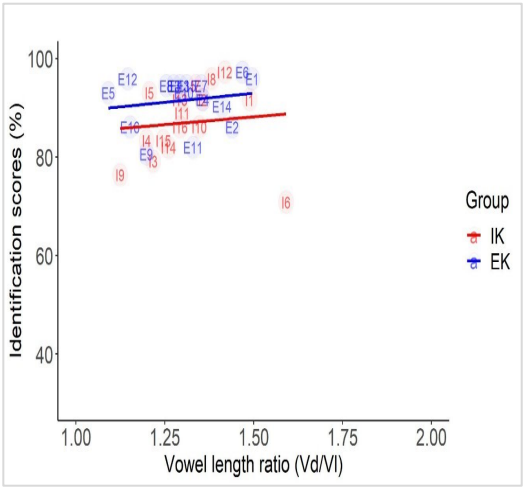

Figure 29. The correlation between the identification of the underlying final consonant and the ratio of vowel length in the voiced context to that in the voiceless context by Korean listeners.

### 5.2.2 Vowel identity (Tense vs. Lax)

The relationship between the identification of the underlying vowel identity and the ratio of vowel length of tense to that of lax by Korean listeners was examined to see whether there is a correlation between two different dynamics. So, for each participant, the average perception scores of the identification of the underlying vowel identity and the ratio of vowel length of tense to that of lax were calculated and they were plotted against each other. The correlation coefficient and the p-value for it for each group were calculated to determine whether there is a statistically meaningful correlation across two variables. As in Figure 30 below, there was little association between the perception scores of the identification of the underlying vowel identity and the ratio of vowel length of tense to that of lax for Korean listeners. Inexperienced Korean speakers showed an extremely weak ($r^2$ = 0.00), and statistically not significant ($p > 0.05$) negative association between the identification of the underlying vowel identity and the ratio of vowel length of tense to that of lax. Experienced Korean speakers showed a weak ($r^2 = 0.14$), and statistically not significant ($p > 0.05$) negative association between two variables.

Korean listeners' weak correlation across the two variables does not demonstrate that there is any relationship between the identification of the underlying vowel identity and the

production of vowel duration difference depending on the vowel quantity, such that Korean listeners who identify English vowel identity in their English perception may or may not distinguish English vowel quantity contrasts in their English production.

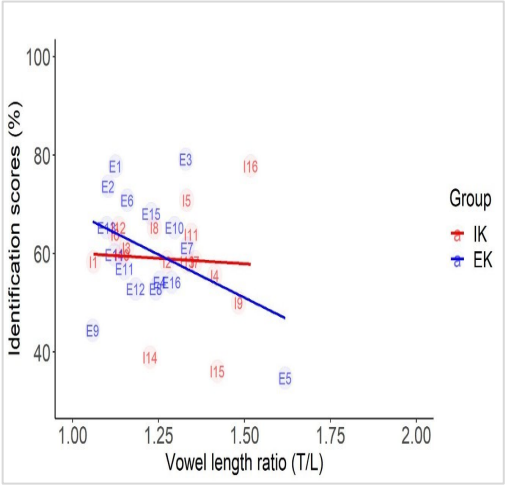


Figure 30. The correlation between the identification of the underlying vowel identity and the ratio of vowel length of tense to that of lax by Korean listeners.

## 6. General discussion

In this study, multiple pairs of variables were analyzed to investigate potential correlations between them across the different Korean participants.

First, the relationship between the participants' performance on the two different contrasts, final consonant voicing and the tense/lax distinction in vowels either in the discrimination task or in the identification task was investigated across the different Korean participants.

Beginning with the discrimination task, it seemed that there are a lot of variation within the two different Korean speaker groups for both contrasts. However, discrimination scores for final consonant voicing contrasts by individual Korean speakers seem not to be related to those for the tense/lax distinction in vowels of each speaker in both groups. Inexperienced Korean speakers showed a marginal correlation across the two variables whereas the experienced Korean group showed a weak correlation across them. Korean listeners' marginal or weak correlation across the two variables does not demonstrate any relationship between the

discrimination of English consonant voicing contrasts and the discrimination of English vowel contrasts for both groups. This result implies that Korean listeners who distinguish English consonant voicing contrasts may or may not distinguish English vowel contrasts in their perception. Similarly, as for the identification task, it seemed that there are a lot of variation within the two different Korean speaker groups for both contrasts. However, identification scores for final consonant voicing contrasts by individual Korean speakers seem not related to those for the tense/lax distinction in vowels of each speaker in both groups. Korean listeners' weak correlation across the two variables does not demonstrate any relationship between the identification of English consonant contrasts and the identification of English vowel contrasts. This result implies that Korean listeners who identify English consonant contrasts may or may not identify English vowel contrasts in their perception.

Second, the relationship between the participants' performance in the two different tasks for each contrast, the discrimination accuracy and the identification accuracy either for final consonant voicing or for the tense/lax distinction in vowels was investigated across the different Korean participants.

Beginning with the consonant voicing contrasts, there seemed a lot of variation within the two different Korean speaker groups in both tasks. When it comes to the inexperienced Korean group, discrimination scores by individual Korean speakers seem related to identification scores of each speaker. Similarly, discrimination scores by individual Korean speakers appear to be related to identification scores of each speaker in the experienced Korean speaker group, except one speaker (E9). This participant performed worse than other members of the same group in both tasks and he was much better at identifying English consonant voicing contrasts than discriminating them. Inexperienced Korean speakers showed a strong correlation across the two variables whereas the experienced Korean group showed a weak correlation across them, possibly due to one speaker (E9). Therefore, it seemed that there is a

relationship between the discrimination of English consonant contrasts and the identification of English consonant contrasts for both groups, if one speaker (E9) is excluded. This result implies that Korean listeners who distinguish English consonant voicing contrasts may identify English consonant voicing contrasts in their perception. As for tense/lax vowel contrasts, it seemed that there is more variation within the two different Korean speaker groups in both tasks. However, discrimination scores by individual Korean speakers seem not related to identification scores of each speaker in both groups in both tasks. Korean listeners' weak correlation across the two variables does not demonstrate any relationship between the discrimination of English vowel contrasts and the identification of English vowel contrasts. This result implies that Korean listeners who distinguish English vowel contrasts may or may not identify English vowel contrasts in their perception.

Lastly, the relationship between the production and either the discrimination or the identification of the two different contrasts, final consonant voicing and the tense/lax distinction in vowels, was examined to see whether there is a relationship between the production of durational differences and these two different dynamic contrasts across the different Korean participants.

Beginning with the consonant voicing contrasts, it appeared that there is a lot of variation within the two different Korean speaker groups in both tasks. However, vowel duration differences produced by individual Korean speakers seem not related to the perception scores of each speaker in both groups. Korean listeners' weak correlation across the two variables does not demonstrate any relationship between the discrimination or identification of English consonant voicing contrasts and the production of vowel duration difference depending on the following consonant voicing. This result implies that Korean listeners who distinguish or identify English consonant voicing contrasts in their perception may or may not distinguish English consonant voicing contrasts in their production. As for tense/lax vowel contrasts,

Korean speakers showed similar patterns as those for consonant voicing contrasts. More specifically, there was a lot more variation across two different Korean speaker groups regarding vowel duration differences in production and perception scores in both tasks. However, vowel duration differences produced by individual Korean speakers seem not related to their perception scores.

The lower discrimination or identification performance observed in the perception experiment may be attributed to the English high back tense/lax vowel pair. Specifically, the high back tense/lax vowel pair showed lower discrimination scores than other vowel pairs, suggesting that they posed greater challenges for Korean listeners during the task.

On the contrary, the discrimination scores for the mid front tense/lax vowel pair were found to be higher than those of other pairs. The observed higher discrimination scores for the mid front tense/lax vowel pair than those of other pairs cannot be solely attributed to the mapping between English and Korean vowels. According to Cho and Jeong (2013), English high front vowels /i/ and /ɪ/ were mapped into a single Korean vowel /i/ and English high back vowels /u/ and /ʊ/ were categorized as a single Korean vowel /u/. As for English mid front vowels /e/ and /ɛ/, they were both mapped into either Korean /e/ or Korean /ɛ/ (Tsukada et al., 2005). Therefore, one possible explanation for this result could be the acoustic characteristics of the English mid front tense vowel, which includes diphthongization and potentially provides additional vowel durational cues. It is likely that participants relied on these temporal cues to distinguish the mid front tense/lax vowel pair, while facing difficulty in distinguishing other pairs that lacked similar temporal cues.

Korean listeners' weaker correlation across the two variables demonstrates that there is no obvious relationship between the discrimination and identification of English vowel quantity contrasts and the production of vowel duration difference between tense and lax vowels. This result implies that Korean listeners who distinguish or identify English tense/lax

vowel contrasts in their perception may or may not distinguish English tense/lax vowel contrasts in their production.

In summary, there was no demonstrated relationship between multiple pairs of variables, including the participants' performance on the two different contrasts in both tasks, the participants' performance in the two different tasks for English vowel contrasts, and vowel duration production skills and perception skills not only for voicing effects but also for vowel quantity effects for Korean speakers regardless of their L2 English experience. On the other hand, there was a positive relationship between the participants' performance in the two different tasks for English consonant contrasts.

## 7. Conclusion

The current study finds evidence that the learning perception skills for different contrasts such as English final consonant voicing contrasts and tense/lax distinction are likely to be developing separately in Korean learners' English acquisition.

For example, Korean speakers were relatively accurate at distinguishing English consonant voicing contrasts not only in non-final position but also in final position in the discrimination task or in the identification task. On the contrary, Korean speakers seemed to have much more difficulty in distinguishing English tense/lax vowel contrasts in both prosodic positions and in both perception tasks. Comparison of responses to stimuli in which material following the post-vocalic consonant was removed and stimuli which retained this information suggests that the reason why Korean speakers performed better at perceptually distinguishing English final consonant voicing contrasts than English tense/lax vowel contrasts may be because it might have been easier for Korean speakers to use the duration of the obstruent closure to distinguish English consonant voicing contrasts. On the other hand, it might have

been difficult to utilize spectral cues which play an important role in addition to vowel duration cues for distinguishing English tense/lax vowel contrasts.

In addition, the results of the current research suggest that the different individuals have different perceptual acquisition processes. For example, when it comes to the relationship between the participants' performance on the two different contrasts, final consonant voicing and the tense/lax distinction in vowels either in the discrimination task or in the identification task, there was no demonstrated relationship between the participants' performance on the two different contrasts in either of the tasks. This suggests that these two different contrasts may involve different perceptual mechanisms, and these mechanisms vary across individual speakers. It is also possible that the factors affecting the acquisition of English consonant voicing contrasts do not necessarily affect the acquisition of English tense/lax vowel distinction. Future research may need to investigate the specific factors and mechanisms to better understand the acquisition of these contrasts.

As for the relationship between the participants' performance in the two different tasks for each contrast, the discrimination accuracy and the identification accuracy both for final consonant voicing or for the tense/lax distinction in vowel, there was no relationship between the participants' performance in the two different tasks for English tense/lax vowel contrasts. On the contrary, there was a strong positive association between the participants' performance in the two different tasks for English consonant voicing contrasts. This suggests that the underlying perceptual mechanisms of these two contrasts are not the same. That is to say, the perception of these different contrasts may be influenced by different factors, and the variability in the relationship between various perceptual tasks may depend on the specific contrast. Further studies may need to investigate the potential factors and the particular features of each contrast that may influence Korean speakers' perception in their L2 English acquisition.

Finally, as for the effect of L2 English experience on the perception scores, English exposure marginally enhanced Korean speakers' performance for English consonant voicing contrasts; whereas there seems to be no influence of L2 English exposure on the performance for English tense/lax vowel contrasts. However, the effect of L2 English exposure seems not significant among the groups examined here. This may be due to the limitation of the current study such that the difference of English experience between two Korean speaker groups was not large enough to differentiate their English perception performance. More specifically, the inexperienced Korean speaker group has less than a week of English experience in an English speaking country whereas the experienced Korean speaker group has around five years of exposure to an English dominant country. This five-year difference might not have a significant impact on developing their English perception skills. Therefore, further research may need to examine longer difference in L2 English exposure, which could have a more significant effect on perception development.

## IV. Conclusion

The current research provides evidence that the learning of production and perception skills for different contrasts such as English final consonant voicing contrasts and tense/lax distinction are developing separately in the process of English acquisition by Korean learners of English.

First, in production it was expected that the Korean L1 speakers would show smaller voicing effects on vowel duration in monosyllabic words whereas they were expected to show a larger vowel durational difference in disyllabic words since a coda distinction in obstruent voicing is neutralized only in final position. In perception, it was expected that the Korean L1 speakers would be able to discriminate or identify vowel durational differences due to the following obstruent voicing not only in monosyllabic structures but also in disyllabic structures. Alternatively, they were expected to show better discrimination or identification in disyllabic structure than in monosyllabic structure. Contrary to expectations, the results of the current study revealed that, with respect to the proportion of vowel duration in the voiceless context to that in the voiced context, Korean speakers showed much weaker voicing effects on vowel duration than did English native speakers in their English production. On the contrary, the perception experiment demonstrated that Korean speakers were accurate at perceptually distinguishing English consonant voicing contrasts not only in non-final position but also in final position in the discrimination task and in the identification task.

Second, Korean speakers were expected to have difficulty in production and perception of English tense/lax vowel contrasts due to the lack of an intrinsic vowel quantity distinction in Korean. In contrast to predictions, Korean speakers resembled English native speakers in terms of the proportion of lax vowel to tense vowel duration, whereas they failed to use spectral cues effectively for those contrasts in their English production. However, Korean speakers had difficulty in perceptually distinguishing English tense/lax vowel contrasts in both prosodic

positions not only in the discrimination task but also in the identification task. These findings taken together suggest that the production and perception skills linked to each contrast are independent from one another. Therefore, it is important to consider both production and perception abilities as separate skills in research on second language acquisition. Moreover, the observed changes in final obstruent voicing contrasts had an impact on the perception of vowel contrasts. This suggests a necessary interdependency or shared processing mechanism between the two contrasts, since they interact in the interpretation of the signal. This would suggest they should be considered as a single skill in learning length cues. Regardless of this, though, the main results of the current study is that variation in skills related to one contrast are generally not strongly related to those for the other skill across the participants in this study.

In addition, the results of the present study support the conclusion that different individuals have different production and perception processes in English language acquisition. For example, when it comes to the relationship between the participants' performance in the two different experiments for each contrast, vowel duration differences spoken in the production experiment and the discrimination or the identification scores in the perception experiment for final consonant voicing and for the tense/lax distinction in vowel, there was no demonstrated relationship between the participants' performance in the two different experiments for either of the contrasts. These results imply that the factors influencing the acquisition of production skills do not necessarily influence the acquisition of perception skills. It is also possible that these two different contrasts may involve distinct mechanisms, and individual speakers show variation in these mechanisms. Future research may need to examine the particular factors and mechanisms to better understand the acquisition of two different abilities for these two contrasts.

As for the effect of L2 English experience on the production and perception performance in the groups examined here, English exposure in the U.S. seems not to enhance

Korean speakers' performance significantly either for English consonant voicing contrasts or for English tense/lax vowel contrasts. In general, the effect of L2 English exposure seems marginal. This may to be due to the limitation of the present research such that the five-year difference of English experience between two Korean speaker groups was not large enough to differentiate their English production and perception performance from one another. Therefore, it seems necessary that further research needs to investigate longer and perhaps more intense L2 English exposure, which could have a more significant effect on production and perception development.

Another possible direction for future studies could be to explore whether the frequency of lexical items associated with the stimuli affects the identification performance by Korean speakers. The IPA representation of the English nonce words can sometimes closely resemble the orthography of real words, which could potentially have had an influence on the identification scores. Examining the corpus systematically shows two things. First, the number of monosyllabic nonce words which correspond to existing lexical words was much higher than that of disyllabic nonce words and second, actual minimal pairs corresponded to specific voicing contrasts. A lexical frequency effect would predict that participants exposed to a larger number of frequent lexical items may show better discrimination or identification performance than those who were exposed to less frequent items. In this study, it was generally observed that identification scores for monosyllabic structures were higher than those for disyllabic structures, although there were some cases where scores in disyllabic structures were also relatively high. This finding might indicate that the abundance of monosyllabic English nonce words which were similar to existing words in the lexicon might have contributed to the participants' improved performance in the monosyllabic structure. The greater availability of phonological patterns resembling real words in the monosyllabic structure could have facilitated better perceptual processing and contributed to the enhanced performance. Further,

the observed better performance in identifying English final consonant voicing contrasts might be attributed to the higher exposure of participants to a greater number of minimal pairs comprising real words as compared to the English tense/lax vowel contrasts.

Analyzing the frequency effect of the data will provide valuable insights into the factors that contributed to the participants' identification performance. By investigating the effect of word frequency on identification performance, we can gain a better understanding of the underlying mechanisms and cognitive processes involved in identification. This analysis will shed light on the specific factors that shape and influence participants' performance in the identification task.

One limitation of the current study is that some of the IPA labels used for the English nonce words used in the experiment closely resembled real words with a different vowel. For example, IPA [i] corresponds to the English orthographic representation of its lax counterpart, e.g. [big] corresponds to *big*. This may have caused potential confusion for participants and influenced their performance. It should be noted that this appears to be the only correspondence of this sort. Future study could consider using a picture naming task that does not involve any potential orthographic effects or use real words with similar frequencies to control frequency effects.

Overall, this study has contributed to the comprehension of the production and perception of English consonant contrasts and vowel contrasts by Korean speakers which is the first large-scale examination of its kind. The current research has established the groundwork for further investigation in second language acquisition research.

## Appendix A. Reading list for the production experiment

| | | |
|---|---|---|
| 1 | Say bʊv soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 2 | Say bɛga soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 3 | Say buv soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 4 | Say bɪp soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 5 | Say bɛb soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 6 | Say buda soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 7 | Say bepa soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 8 | Say bib soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 9 | Say bɛta soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 10 | Say bɛpa soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 11 | Say bʊba soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 12 | Say bif soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 13 | Say bisa soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 14 | Say bev soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 15 | Say big soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 16 | Say bip soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 17 | Say bifa soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 18 | Say bɛg soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 19 | Say bɪba soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 20 | Say biv soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 21 | Say buk soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 22 | Say bʊd soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 23 | Say bup soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 24 | Say beta soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 25 | Say bɪpa soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 26 | Say bud soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |

| | | |
|---|---|---|
| 27 | Say bʊb soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 28 | Say bega soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 29 | Say biz soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 30 | Say bɛf soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 31 | Say buva soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 32 | Say bɛda soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 33 | Say bef soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 34 | Say bɪka soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 35 | Say bɪva soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 36 | Say bɛba soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 37 | Say but soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 38 | Say buga soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 39 | Say bek soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 40 | Say bɛva soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 41 | Say biga soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 42 | Say beka soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 43 | Say beza soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 44 | Say bʊsa soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 45 | Say bɛka soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 46 | Say bɪta soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 47 | Say bipa soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 48 | Say bʊga soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 49 | Say bɛd soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 50 | Say bida soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 51 | Say bʊta soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 52 | Say buz soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 53 | Say bɛp soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |

| | | |
|---|---|---|
| 54 | Say beg soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 55 | Say bɪv soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 56 | Say bɪsa soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 57 | Say bʊza soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 58 | Say bɪfa soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 59 | Say bep soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 60 | Say bɪz soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 61 | Say bʊk soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 62 | Say beva soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 63 | Say buba soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 64 | Say biva soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 65 | Say bɛv soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 66 | Say buka soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 67 | Say biba soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 68 | Say bʊf soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 69 | Say bik soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 70 | Say bes soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 71 | Say bʊva soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 72 | Say beb soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 73 | Say bʊka soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 74 | Say buza soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 75 | Say bʊp soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 76 | Say bɛfa soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 77 | Say bʊs soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 78 | Say bʊda soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 79 | Say bɪs soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 80 | Say besa soon | as in the words 'take', 'play', 'place', 'days', 'eight' |

| | | |
|---|---|---|
| 81 | Say bɪda soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 82 | Say bʊz soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 83 | Say bɛt soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 84 | Say beba soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 85 | Say bid soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 86 | Say bika soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 87 | Say bɛz soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 88 | Say bɪg soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 89 | Say bufa soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 90 | Say buf soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 91 | Say bɛk soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 92 | Say bita soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 93 | Say bub soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 94 | Say bɪk soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 95 | Say bɛs soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 96 | Say bɪf soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 97 | Say buta soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 98 | Say bʊt soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 99 | Say befa soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 100 | Say biza soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 101 | Say bɪga soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 102 | Say bʊg soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 103 | Say beda soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 104 | Say bit soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |
| 105 | Say bɪb soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 106 | Say bupa soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 107 | Say bet soon | as in the words 'take', 'play', 'place', 'days', 'eight' |

| | | |
|---|---|---|
| 108 | Say busa soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 109 | Say bug soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 110 | Say bʊpa soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 111 | Say bus soon | as in the words 'food', 'pool', 'cool', 'two', 'shoot' |
| 112 | Say bɪza soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 113 | Say bɪt soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 114 | Say bed soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 115 | Say bɛsa soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 116 | Say bɛza soon | as in the words 'get', 'pet', 'send', 'rent', 'red' |
| 117 | Say bez soon | as in the words 'take', 'play', 'place', 'days', 'eight' |
| 118 | Say bʊfa soon | as in the words 'good', 'should', 'put', 'cook', 'book' |
| 119 | Say bɪd soon | as in the words 'hit', 'kid', 'Miss', 'ship', 'sit' |
| 120 | Say bis soon | as in the words 'leaf', 'feed', 'eat', 'seat', 'sheep' |

## References


Ahn, S.-C. (1998). *An Introduction to Korean Phonology*. Seoul: Hanshin Publishing Co.

Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year* (pp. 289-304). Dordrecht, The Netherlands: Kluwer Academic.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167-200). Timonium, MD: York.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279-304). Baltimore: York Press.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*(2), 131-158.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics, 34*(3), 372-387.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica, 22*(3), 129-159.

Cho, M.-H. & Jeong, S. (2013). Perception and production of English vowels by Korean learners: A case study. *Studies in phonetics, phonology and morphology, 19*(1), 155-177.

Choi, J., Kim, S., & Cho, T. (2016). Phonetic encoding of consonant voicing contrast under different focus conditions in L1 vs. L2 English. *Frontiers in Psychology, 7*(624), 1-17.

Davis, S., & Cho, M. H. (2006). Phonetics versus phonology: English word final /s/ in Korean loanword phonology. *Lingua, 116*(7), 1008-1023.

De Jong, K. J., & Cho, M.-H. (2012). Loanwords phonology and perceptual mapping: Comparing two corpora of Korean contact with English. *Language, 88*(2), 341-368.

De Jong, K. J., & Park, H. (2012). Vowel epenthesis and segment identity in Korean learners of English. *Studies in Second Language Acquisition, 34*(1), 127-155.

De Jong, K. J., Silbert, N. H., & Park, H. (2009). Generalization across segments in second language consonant identification. *Language Learning, 59*(1), 1-31.

De Jong, K. J., & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics 30*(1), 53-75.

Denes, P. (1955). Effect of duration on the perception of voicing. *The Journal of the Acoustical Society of America, 27*(4), 761-764.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*(1), 47-65.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America, 93*(3), 1589-1608.

Flege, J., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *The Journal of Phonetics, 25*(4), 437-470.

Flege, J. & Hillenbrand, J. (1986). Differential use of temporal cues to the [s-z] contrast by native and non-native speakers of English. *The Journal of the Acoustical Society of America, 79*(2), 508-517.

Flege, J. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*(2), 125-146.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36*(2), 268-294.

Han, J.-I. (2015). *영어 음성학과 발음지도* (*English Phonetics and Pronunciation Guide*). Seoul: Hangukmunhwasa.

Hong, S. (2012). The relative perceptual easiness between perceptually assimilated vowels for university-level Korean learners of American English and measurement bias in an identification test. *The Phonology-Morphology Circle of Korea, 18*(3), 491-511.

House, A. S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America, 33*(9), 1174-1178.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America, 25*(1), 105-113.

Ingram, J., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics, 25*(3), 343-370.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America, 118*(5), 3267-3278.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In Connell, B., & Arvaniti, A. (Eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV* (pp. 235-253). England: Cambridge University Press.

Jun, S.-A., & Cowie, I. (1994). Interference for ‘new’ and ‘similar’ vowels in Korean speakers of English. *Ohio State University Working Papers, 43*, 117-130.

Kim, H., & Jongman, A. (1996). Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics, 24*(3), 295-312.

Kim, H.-K. (1999, June 25). *Korean Phonology for the KAEPS system*. https://www.youngkorean.com/kaeps/data/kor_phon.html

Kim, J.-E. (2010). Perception and production of English front vowels by Korean speakers. *Phonetics and Speech Sciences, 2*(1), 51-58.

Kim, J.-E. (2011). A study on the correlation between English word-final stop and vowel duration produced by speakers of Korean. *Phonetics and Speech Sciences, 3*(1), 15-22.

Kim, J.-Y., & Shin, S.-H. (2014). Transfer of Korean vowel length contrast to English in perceiving and producing English syllable-final VC - Focusing on Gyungbuk and Standard Korean. *The New Studies of English Language & Literature 71*, 227-260.

Kim, S.-h. (2016). *A study of Korean high school students' production and perception of English front vowels*. Master's thesis, Chosun University.

Kim, S.-J. (2006). The perception and production of tense and lax vowels by early English learners and its pedagogical implication. *The New Korean Association of English Language and Literature, 48*(2), 127-145

Klatt, D. H. (1973). Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America, 54*(4), 1102-1104.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America, 59*(5), 1208-1221.

Kluender, K., Diehl, R., & Wright, B. (1988). Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics, 16*(2), 153-169.

Ko, H. (1997). *A study on consonantal voicing-dependent vowel duration in English -through a comparative analysis between Korean and English native speakers with production and perception experiments*. Master's thesis, Wonkwang University.

Ko, H. (1998). A study on consonantal voicing-dependent vowel duration in English. *The Journal of the Acoustical Society of Korea, 17*(1), 399-402.

Ko, H., Lee, M., Jo, M., & Lee, S. (1997). 영어 어말자음 유/무성 대립과 모음길이와의 상관관계 연구 - 한국어 모국어화자 영어 모국어화자의 발성과 인지실험을 통하여 - (A study on the correlation between English final consonant voicing contrast and vowel length -through a comparative analysis between Korean and English native speakers with production and perception experiments-). *In Proceedings of the KSPS conference*, 189-195.

Koo, H.-S. (1998). The influence of consonant environment upon the vowel duration. *Speech sciences, 4*(1), 7-17.

Koo, H.-S. (2000). Characteristics of English vowels spoken by Koreans. *The Korean Journal of Speech Sciences, 7*(3), 99-108.

Kwag, S. (2019). *A study on the relationship between length of the preceding vowel of voiced or voiceless consonants in word final position and English pronunciation assessment: Focusing on the speaking of Elementary English learners*. Master's thesis, Yonsei University.

Lee, I., & Ramsey, R. (2000). *The Korean Language*. Albany, NY: State University of New York Press.

Lee, J.-K., & Lee, Y.-W. (2011). Speech intelligibility of Korean learners' and a native English speaker's production of English tense and lax vowels: Focusing on high vowels. *Studies in Phonetics, Phonology and Morphology, 17*(3), 497-520.

Lee, J.-Y. (2008). Perception of English high vowels: Duration as a cue by Korean speakers of English. *Kansas Working Papers in Linguistics 30*, 195-204.

Lee, K. (2016). *Neutralization of coda obstruents in Korean: Evidence in production and perception*. Doctoral dissertation, Northwestern University.

Lee, K.-Y., & Cho, M.-H. (2013). Production of English vowels by Korean learners. *The Journal of the Korea Contents Association, 13*(9), 495-503.

Lee, K.-Y., & Cho, M.-H. (2015). Perception of English vowels by Korean learners: Comparisons between new and similar L2 vowel categories. *The Journal of the Korea Contents Association, 15*(8), 579-587.

Lee, S., Nam, H., Kang, J., Shin, D.-J., & Kim, Y. S. (2017). The impact of language-learning environments on Korean learners' English vowel production. *Phonetics and Speech Sciences, 9*(2), 69-76.

Mack, M. (1982). Voicing-dependent vowel duration in English and French: Monolingual and bilingual production. *The Journal of the Acoustical Society of America, 71*(1), 173-178.

Major, R., & Faudree, M. C. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition 18*, 69-90.

McAllister, R., Flege, J. & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*(2), 229-258.

Mitleb, F. (1984a). Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics, 12*(1), 23-27.

Oh, E. (2012). Further issues on the duration differences in vowels due to the voicing of the following stops in English. *Phonetics and Speech Sciences, 4*(3), 85-92.

Park, J.-W. (1994). Variation of vowel length in Korean. In Y.-K. Kim-Renaud (Ed.), *Theoretical Issues in Korean Linguistics* (pp. 175-187). Honolulu, HI: University of Hawaii Press.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693-703.

Port, R. F., & Dalby, J. (1982). Consonant/vowel ratio as a cue for voicing in English. *Perception & Psychophysics, 32*(2), 141-152.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America, 51*(4), 1296-1303.

Shin, D.-J., & Sa, J.-J. (2005). An experimental study on vowel duration differences before voiced and voiceless consonants pronounced by Korean learners of English-From fricatives and affricates sounds. *The Korean society of phonetic sciences and speech technology*, 91-95.

Sohn, H.-M. (1999). *The Korean Language*. Cambridge: Cambridge university Press.

Summers, W. V. (1987). Effects of stress and final-consonant voicing on vowel production: Articulatory and acoustic analyses. *The Journal of the Acoustical Society of America, 82*(3), 847-863.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization: electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science, 21*(11), 1532-1540.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics, 33*(3), 263-290.

Yang, B.-G. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*(2), 245-261.

Yang, B.-G. (2013). A comparative study of relative distances among English front vowels produced by Korean and American speakers. *Phonetics and Speech Sciences, 5*(4), 99-107.

Yang, E. (2020). *Korean high school students' perception of English vowel length and analysis of their current status: A focus on length before English stop consonants*. Master's thesis, Korea University.

Yoon, K.-M. (2002). *The effect of phonological environments on English vowel length.* Master's thesis, Chonnam National University.

Yu, L. (2023). *The perceptual cue weighting of English tense-lax vowel contrasts by first language (L1) and second language (L2) speakers of English*. Doctoral dissertation, Indiana University.

# Curriculum Vitae

## Education

**Ph.D. in Linguistics** **2023**
**M.A. in Linguistics in Computational Linguistics track** **2021**
Indiana University, Bloomington, IN

**M.Ed. in English Language Education** **2006**
Kyonggi University, Suwon, Korea

**B.A. in Chinese Language & Literature** **2003**
**B.A. in English Language & Literature**
Sookmyung Women's University, Seoul, Korea

## Professional Experience

**Annotator | Lausanne Business Solution, Philadelphia, PA** **2018 - 2019**

Expertly conducted NER analysis and general annotation of Korean/English texts. Actively engaged in correcting and identifying model mistakes in Korean texts along with noun chunking / key phrase identification of Korean texts. Initiated linguistic data analysis and annotation to clean, transform, and model data to discover useful information. Designed and developed guidelines and instructions for text analyses. Executed in-depth research and discovered significant information about Korean language via corpus building and sourcing of various Korean texts.

**Associate Japanese Instructor | Indiana University, Bloomington, IN** **2016**
J101 Elementary Japanese, Fall 2016

**Associate Korean Instructor | Indiana University, Bloomington, IN** **2013 - 2022**
K101 Elementary Korean, Fall 2013, Fall 2014, Fall 2022(10.24-12.16)
K102 Elementary Korean, Spring 2014, Spring 2015, Spring 2017
K201 Intermediate Korean, Fall 2015, Fall 2017
K202 Intermediate Korean, Spring 2016
K302 Advanced Korean, Spring 2018, Spring 2019

**Research Assistant/TA | Kyonggi University, Suwon, Korea** **2008 - 2010**

Delivered assistance to a professor with grading students' assignments. Actively engaged in data sorting and preprocessing for professor's project.

- Published two papers with a professor as a co-author funded by NRF (National Research Foundation of Korea).

## Activities

**Workshop | Korea University Winter School, Seoul, Korea** **2021**
• Learned Intro to Deep Learning and Natural Language Processing

**Workshop | Korea AI Academy, Seoul, Korea** **2019**
• Made a Chatbot for Learners of Korean using Python.
• Learned how to use Flask / Web Crawling to make a Chatbot.

**Intensive Data Science Course | DS School, Seoul, Korea** **2019**
• Gained hands-on experience of beginner/intermediate level of Python, Pandas, SQL, Scikit-learn, Seabom, Matplotilb

**Volunteer/Performance | Indiana University, Bloomington, IN** **2015 - 2022**
• Volunteer at the Gayageum experience booth at Korean Night, 2019, 2022
• Gayageum performance at Korean Night, 2017
• Samulnori performance at Asian Festival, World Language Festival, Korean Night, 2015

## Publications

• **Universal Dependency Treebank for Xibe, 2020**. Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020), 205-215.
(https://github.com/UniversalDependencies/UD_Xibe-XDT)
Zhou, He, Chung, Juyeon, Kübler, Sandra & Tyers, Francis. M

• **Production and perception of English vowel length depending on the following consonant voicing by Korean learners of English, 2019**. Proceedings of Seoul International Conference on Speech Sciences 2019, 119. Chung, Juyeon

• **Cross-Language Perception by Prosodic Position: Korean Perception of English Posterior Obstruents, 2010**. The Linguistic Society of Korea 57, 83-108.
Cho, Mi-Hui & Chung, Juyeon

• **Orthographic Influence in the Perception and Production of English Intervocalic Consonants: A Pilot Study, 2009.** Journal of the Korea Contents Association 9, 459-466.
Cho, Mi-Hui & Chung, Juyeon

• **The Influence of Orthography on the Pronunciation of English C-spelled Words, 2006.** English Language and Linguistics 22, 93-113.
Cho, Mi-Hui, Lee, Shinsook & Chung, Ju-Yeon

## Conference Presentations

**• Acoustical Society of America, 2022**
Perception of English vowel length modulation due to vowel identity and the identity of the following consonant by Korean learners of English

**• The 4th Workshop on Universal Dependencies (UDW 2020), COLING, 2020**
Universal Dependency treebank for Xibe

**• Acoustical Society of America, 2019**
Production and perception of English vowels preceding voiced and voiceless consonants by Korean learners of English

**• Seoul International Conference on Speech Sciences, 2019**
Production and perception of English vowel length depending on the following consonant voicing by Korean learners of English

**• The 24th Annual Mid-Continental Phonetics & Phonology Conference, 2019**
Production and perception of English vowel length by Korean learners of English

**• Hanyang International Symposium on Phonetics and Cognitive Sciences of Language, 2019**
Production and perception of English vowel length depending on the following consonant voicing by Korean learners of English

**• Acoustical Society of America, 2016**
Production of English vowels preceding voiced and voiceless consonants by Korean learners of English

**• The 20th Annual Mid-Continental Phonetics & Phonology Conference, 2015**
Production of English vowel length depending on the following consonant by Korean L2 learners of English

## Academic Honors and Awards

• Student Conference Travel Grant, Indiana University, 2019
• Student Conference Travel Grant, Indiana University, 2016