KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

Korean ESL Learners' Production of English Vowel Contrasts: Developmental Variations in L2 Sound Learning*

Mi Sun Park · Joo-Kyeong Lee (University of Seoul)

Park, Mi Sun (First author)
Visiting Professor, Division of General English
University of Seoul
163 Seoulsiripdae-ro, Dongdaemun-gu, Seoul, Korea
Tel: 02) 6490-5226
Email: msp2136@uos.ac.kr

Lee, Joo-Kyeong
(Corresponding author)
Professor, Department of English Language and Literature
University of Seoul
Tel: 02) 6490-2519
Email: jookyeong@uos.ac.kr

ABSTRACT

Park, Mi Sun and Joo-Kyeong Lee. 2024. Korean ESL learners' production of English vowel contrasts: Developmental variations in L2 sound learning. *Korean Journal of English Language and Linguistics* 24, 1318-1332.

The present study elucidates the developmental progression of English vowel production in Korean learners' L2 speech. The study focused on the first and second formant frequencies and duration values of seven American English monophthongs (excluding /u, ʊ, ə/) produced by Korean adults classified into three groups according to their proficiency in pronunciation (the degree of foreign accent): low, intermediate, and high. Then we compared them with those produced by native speakers. Results showed between-group differences in that high ESL learners produced English vowel contrasts by using spectral and/or temporal cues in a way resembling the native speakers', while intermediate and low groups were relatively more limited in using these cues. The study also observed differences in the production of front and back vowels among ESL learners. Specifically, high learners primarily relied on temporal rather than spectral characteristics in producing back vowel contrasts, whereas intermediate and low learners showed no significant distinctions in temporal or spectral dimensions of back vowels. These results highlight the complex developmental processes involved in second language (L2) vowel learning.

KEYWORDS
Korean ESL learners, English vowel contrasts, production, sound learning

1. Introduction

Speaking without a foreign accent is a challenging task for second language (L2) learners. To address this challenge, many previous studies on L2 speech have focused on the various sources of pronunciation difficulties. In general, cross-linguistic differences in the sound systems of the learners' native language (L1) and L2 cause pronunciation difficulties for L2 learners who are likely to end up transferring the features and rules from their L1 phonological system (Ioup 2008). Speech Learning Model (SLM, Flege 1995) claimed that L2 sounds that are "equivalent" (or similar) to those in L1, but not quite the "same" as those in L1 are the most difficult to acquire. These sounds place a heavy learning load on L2 learners and cause them to fail to make distinctions between L2 sounds and, eventually, speak with foreign accents. On a similar note, Perceptual Assimilation Model (Best 1995) explained that two L2 sounds that are assimilated to one category in the L1 (i.e., single-category assimilation) are more difficult to perceive than the ones assimilated to two categories (i.e., two-category assimilation). Now it is of importance in L2 speech research to study how such difficulties remain or decrease in the course of L2 development. The present study aims to investigate segmental errors in L2 vowels, with a focus on how phonemic errors manifest in L2 learners' foreign accent, by examining between-group differences of Korean ESL learners with distinct accentedness and comparing their production of L2 English vowels with that of native English speakers.

While consonants are articulated with a local constriction at a clear-cut articulatory position, vowel articulation is often considered to involve complex and subtle movements of the tongue (e.g., height and frontness/backness) and lip rounding that require scrutiny. Moreover, English vowels also include the tense-lax distinction involving spectral and temporal differences (Flege et al. 1997, Yang 2008), which would be challenging for ESL learners whose L1 does not make such contrast, like Korean; in such a case, the tense-lax properties could be realized as in the articulatory habits that are established to realize the phonemes in L1 (Flege 1995). To examine this limitation, previous research has investigated L2 learners' perception of vowel contrasts (Hwang and Lee 2012). In the studies on L2 vowel production, the researchers have probed the vowel quality (in terms of the first and second formant frequencies) as well as the vowel duration in association with cross-linguistic phonological differences and L2 pronunciation development (Hong 2013, Munro 1993).

Flege et al. (1997) compared groups of nonnative adults of different L1s (Korean, Spanish, German, and Mandarin Chinese) and varying amounts of L2 experience in the U.S. (1 vs. 7 years), and the study focused on their production and perception of the front vowels /i, ɪ, ɛ, æ/ in English. In one of the experiments, the authors acoustically analyzed the vowels by measuring vowel durations and frequencies to estimate vowel height and frontness/backness. They found that L1 Korean speakers did not produce a significant spectral difference between /i/-/ɪ/ or between /ɛ/-/æ/, unlike native English speakers. Instead, they were seen to over-rely on vowel duration to distinguish English /i/-/ɪ/, but such a difference in duration was smaller and non-significant in the L1 Korean speakers with greater L2 English experience. The authors claimed that this might be taken as evidence of L2 vowel learning: As they gain experience in their L2, the Koreans might have learned to stop treating the vowel contrast as a phonemic length distinction or become more sensitive to spectral differences between the vowels.

Baker et al. (2002) investigated the effects of age and L2 experience on the English vowel learning of Korean children and adults and compared their performance to age-matched native English (NE) speakers. One of the experiments tested the Korean participants' perceptual similarity of eight English vowels /i, ɪ, ɛ, æ, u, ʊ, ʌ, ɑ / and revealed three types of cross-language (English-Korean) relationships: (1) English vowel contrast /ɑ/-/ʌ/, each of whose members was identified with a separate Korean vowel, /a/ and /ʌ/, respectively; (2) a contrast highly confusable across the two languages (English /æ/-/ɛ/); and (3) English contrasts both of whose members were

identified with a single Korean vowel (English /i/-/ɪ/ and /u/-/ʊ/). The study also assessed the perception and production accuracy of Korean adults and children differing in amount of L2 English experience (1 vs. 9 years of residence in America). Korean children in general outperformed Korean adults in production of the target English vowels including the sounds that were highly dissimilar across the two languages (e.g., English /ɪ/ and /ʊ/), even with a minimal amount of experience (about 1 year), even though Korean adults had been exposed to English in Korea as part of foreign language instruction in secondary schools. The authors concluded that the finding supports the hypothesis of SLM: Korean children's still-developing phonetic system of L1 is attributable to their being less likely to perceptually identify L2 sounds with L1 sound categories. One notable finding in this study was that Korean adults of 9 years of L2 experience were only more accurate than those of less L2 experience in the perception and production of English /ɑ/ and /ʌ/, which were initially perceived as similar, yet relatively non-confusable (each mapping onto a different Korean vowel), across the two languages.

The study of Tsukada e al. (2005) employed a similar research design, where the native Korean (NK) adults and children took tasks for identification and discrimination of pairs of contrastive English vowels and an elicitation task assessed via phonetic classification and acoustic analysis of the target English vowels /i, ɪ, eⁱ, ɛ, æ, ɑ, ʌ/. The NK adults were found to produce English vowels less intelligibly and make smaller acoustic contrasts (measured as the magnitude of acoustic distance, i.e., the F1 and F2 contrast) for /ɛ/-/æ/ and for /ɑ/-/ʌ/ than NE speakers or NK children. The NK children discriminated the English vowels more accurately and produced significantly larger between-vowel contrasts than NK adults; however, they did not differ significantly from NE children in producing English vowels (either in terms of intelligibility or the magnitude of between-vowel differences in F1 and F2). Based on the results, the authors speculated that NK adults, probably more fully developed in L1 categories than NK children at the onset of L2 learning, might block the formation of new L2 vowel (and consonant) categories. The authors also commented that the actual span of L2 learning would be greater than that of the NK participants observed in their study (3–5 years), as in the case of NK adults in Baker et al. (2002), who could produce and perceive vowel contrasts more accurately after 9 years of L2 experience.

Such advanced L2 English learners, namely Korean-English bilinguals (i.e., Koreans born and raised in the U.S.), were examined in the study of Oh (2013), along with late Korean learners of English (NK) and native speakers of American English (NE). The author acoustically analyzed the three speaker groups' reading of monosyllabic words containing English [i, ɪ, ɛ, æ] to identify dynamic spectral patterns. The Korean-English bilinguals (KE) showed F1, F2, and F3 transition shapes and overall formant values of the vowels similar to the NE group in the target vowels. The spectral trajectories of the target vowels produced by the KE and NE groups were separate and did not overlap, which indicated that these speakers could make sufficient discrimination among these vowels. The NK group, on the other hand, showed distinct formant ranges and vowel trajectories in the vowels, which suggested that they had difficulty discriminating these vowels. In addition, the three groups showed a similar pattern in vowel durations (higher vowels < lower vowels; lax vowels < tense vowels), while there was also a difference in vowel durations across the groups (NK < NE < KE). The findings of the study indicated that language experiences affect the learning of dynamic patterns of L2 vowels, but Oh noted that certain phonetic aspects (e.g., the formant movement direction of [i]) seemed particularly difficult even for bilingual speakers.

Recent studies on Korean learners of English continued to delve into the spectral and temporal features in the English vowels and their impacts on L2 pronunciation accuracy rating. Lee and Rhee (2019) examined adult Korean EFL learners in distinct proficiency levels (lower, middle, and higher) based on their performance in reading English texts and reported that higher-rated learners were better able to distinguish the contrasts in the front vowel pairs (/i/-/ɪ/, /ɛ/-/æ/) but not in the central and back vowel pairs /u/-/ʊ/ and /ʌ/-/ɔ/. The learners in general showed difficulty distinguishing adjacent vowels using spectral cues (F1 and F2), and the lower-rated

learners were found to use temporal cues less for the tense and lax vowel pairs, which was not consistent with the previous findings that inexperienced ESL learners tend to emphasize vowel duration (Flege et al. 1997).

Lee and Lee (2011) examined the perception and intelligibility of the production of English high tense and lax vowels /i, ɪ, u, ʊ/ in young Korean EFL learners (in Grade 6). They observed that Korean learners have difficulty identifying English /i/ and /ɪ/ and are likely to neutralize these vowels to the lax one, which is similar to the Korean vowel /i/ in terms of duration, but not in spectral properties (F1 and F2), as a result of negative L1 transfer. The authors interpreted that such spectral and durational conflict might confuse Koreans in identifying the vowels /i, ɪ/. On the other hand, the Korean learners could identify the high back lax vowel /ʊ/ significantly better than the tense vowel /u/. Furthermore, the learners with higher English proficiency (as assessed using a diagnostic test) outperformed the learners of lower proficiency in producing more intelligible contrast between the high tense and lax vowels of English, which indicated a dynamic developmental pattern in the acquisition of these vowels.

The studies reviewed so far show that Korean learners of English, especially those in adulthood, make L2 vowel distinctions in a way different from native speakers. There is also room for development (or improvement) as their L2 learning experiences accumulate: as asserted by Flege (1988), L2 speech will improve or get less accented as learning progresses even for late learners. Meanwhile, previous studies were often limited in that they mainly examined the vowels in certain spaces (e.g., front vowels only, as in Flege et al. 1997, Oh 2013) without comprehensive analyses of the vowels in the entire target language system. Lack of consistency in the methods and results has also made it difficult to capture what underlies the developmental progression of English vowels in Korean learners. Taking account of the L2 learning environment is also crucial, since ESL learners who reside in an immersive environment may show differences in the articulation of vowels or consonants from EFL learners who learn English in a limited L2 environment like a classroom.

A desired next step in research would be to investigate the developmental trajectories of L2 sound (here, vowels) learning, and the developmental patterns over time can be investigated by comparing the articulatory patterns (as manifested by the acoustic parameters, F1, F2, and duration) of L2 learners with different levels of English proficiency existing contemporaneously. Thus, the present cross-sectional study aims to compare and document the developmental variations in L2 vowels and vowel contrasts made by Korean ESL learners of different accentedness as extensively as possible, instead of focusing on vowels in a particular vowel space. By elucidating the developmental progression of English vowel production in Korean learners' L2 speech—an area that has been previously underexplored—the findings of the study are expected to help clarify key aspects that should be addressed and emphasized in L2 instruction and assessment at different learning stages.

2. Experiment

2.1 Participants

The study initially had a cohort of 40 Korean adults speaking English as their second language (ESL) and living in the United States (28 females and 12 males, mean age: 27.8). None of these learners had learned languages other than their L1 (Korean) before age 12, but their L2 English experience (namely, length of residence in the United States) varied (range: 0.5 to 18 years). Twenty-three native listeners of American English participated as raters of foreign accentedness for the 40 ESL learners' recordings. They had the experience of teaching English as a second or foreign language. On average, they had taught English for 5.48 years in a second language setting (range: 0.5–27 years) or for 1.72 years in a foreign language setting (range: 0.5–10 years).

2.2 Stimuli

In order to collect speech stimuli, 40 Korean ESL participants read an English passage consisting of 69 words. It was an excerpt from the Speech Accent Archive (Weinberger 2015, see the Appendix), where speech corpus featuring diverse English segments and the scripts are available. The passage contained most English monophthongs except the high back lax vowel [ʊ]. Since the current work focuses on whether Korean ESL learners successfully maintain the contrast of confusing vowels in their production, the high back vowel contrast of [u] and [ʊ] was not able to be considered in the current work. The participants were asked to read the passage silently for 30 seconds. If there was an unfamiliar word, the learner pointed at it, and the first author read the word just once without providing further explanation. After 30 seconds of silent reading, the learners read the text while being audio-recorded. The recording samples were about 30 seconds long on average, and the intensity was normalized (75 dB) using Praat (version 6.4.14, Boersma and Weenink 2024).

2.3 Procedure

The ESL learners' accentedness was judged by the native listeners on a 9-point scale (1: heavy foreign accent, 9: no foreign accent) on an online rating form created using Qualtrics. At the beginning of the form was detailed instruction on rating, where the raters could listen to two baseline samples—one produced by a native English speaker and the other by a native Korean speaker who was not included in the 40 participants in this study. The 40 speech samples were presented in a randomized order, and the listeners could assess the samples at their own pace and replay the samples multiple times if needed. The raters' inter-rater consistency (Cronbach's alpha) was 0.98 ($p<0.05$), and the mean accentedness score was computed for each Korean ESL learner.

Based on the mean accentedness scores, 23 of the ESL learners were selected and divided into three proficiency groups[1]: low (score range: 1–3, n=5), intermediate (score range: 4–6, n=12), and high (score range: 7–9, n=6), as summarized in Table 1. The learners falling in the middle ranges (i.e., 3–4 and 6–7) were excluded from the analysis, in order to make an interval between proficiency levels. Recording samples of the same text, produced by 10 native speakers of American English, were retrieved from the Speech Accent Archive and compared with Korean groups for analysis.

Table 1. Korean ESL and Native Groups

Speaker group	Low (n=5)	Intermediate (n=12)	High (n=6)	Native (n=10)
Gender	4 F / 1 M	9 F / 3 M	3 F / 3 M	5 F / 5 M
Min. of the mean accentedness score	2.23	4.30	7.13	Not Applicable
Max. of the mean accentedness score	2.93	5.70	8.50	Not Applicable
Mean years of residence in the U.S. (SD)	1.6 (1.52)	2.2 (2.85)	9.8 (6.28)	Not Applicable

---

[1] In this study, the term *proficiency* refers to proficiency in the domain of *pronunciation*, distinct from proficiency in general that encompasses the domains beyond phonology (e.g., grammar and vocabulary). Proficiency in this study may, or may not necessarily, be related to the ESL learners' age of onset of L2 learning or length of L2 experience; however, such relationships are beyond the scope of the present study.

2.4 Analysis

We analyzed the acoustic features of seven English monophthongs—four front vowels /i, ɪ, ɛ, æ/ and three back vowels /ɔ, ʌ, ɑ/ (see Table 2)—that appeared in the 23 speech samples selected in 2.2. Acoustic data of F1, F2, and duration were collected using Online Forced Aligner (Pevnick n.d.) and Praat (Boersma and Weenink 2024). We extracted the midpoint values of F1 and F2 and calculated the normalized duration of the target vowels in each of the speech samples.[2] Statistical analyses of mixed-effect linear regression were performed using the lme4 function in R to examine the phonemic distinctiveness and/or differences between the confusing vowels across proficiency groups. Following the previous literature (Baker et al. 2002, Lee and Rhee 2019), five pairs of vowels were taken into consideration to address Korean ESL learners' difficulties and confusions in making spectral and temporal contrasts: two pairs of front vowels such as the pair of /i/ and /ɪ/ and the pair of /ɛ/ and /æ/ and three pairs of back vowels like /ɔ/ and /ʌ/, /ʌ/ and /ɑ/, and /ɑ/ and /ɔ/. F1 and F2 values and normalized durations were employed as dependent variables. Front vowels and back vowels were analyzed in separate regression models for each of Korean and native groups. This study attempted to avoid direct comparisons of three Korean groups with the native group because contrast maintenance should be analyzed within a group. Korean learners' individual measures of F1 or F2 are not necessarily similar to those of native speakers, but they can be interpreted as producing vowel contrasts successfully when their statistical pattering is statistically similar to that of native speakers.

Table 2. Target Vowels and Tokens

Target vowels	Tokens
/i/	*please, these(*2), peas, cheese, need, three, meet*
/ɪ/	*bring, things(*2), six, thick, big, kids,*
/ɛ/	*Stella, fresh, red, Wednesday*
/æ/	*ask, slabs, snack, plastic, bags*
/ɔ/	*call, small, store, also*
/ʌ/	*brother*
/ɑ/	*Bob*

2.5 Results

Figure 1 shows native speakers' results of acoustic measurements. F1 values increase as the front vowels are from high to low (/i/ to /æ/) because F1 reflects tongue height. F2 values range from around 1700 Hz to 2300 Hz; higher F2 is the characteristic of front vowels. For normalized duration, high tense vowel /i/ and the low vowel /æ/ were dominantly longer than /ɪ/ and /ɛ/, respectively. The statistical results of the front vowels /i, ɪ, ɛ, æ/ are presented in Table 3. The native speakers maximally used all of the acoustic features to distinguish /i/ and /ɪ/ in their utterances. The F1 and F2 values were significantly different between /i/ and /ɪ/ (β=130.73, p<0.005 for F1, and β=-375.86, p<0.005 for F2), and the normalized duration of /ɪ/ was significantly shorter than that of /i/ (β=-0.329, p=0.015). All of three acoustic features are good predictors for the distinction between /i/ and /ɪ/ in native speakers' production. On the other hand, the contrast between /ɛ/ and /æ/ seems to be produced merely by F1 and duration. F2 played no significant role in making articulatory differences between them. Articulatorily, the tongue height of /ɛ/ is higher than that of /æ/, which realized a significantly higher F1 for /æ/ than /ɛ/. In addition, native

---

[2] Normalized duration was calculated by dividing the duration of a target vowel by the duration of the vowel *toy* in each participant's speech. The vowel /ɔj/ was chosen because its formant trajectories are distinctive and clear-cut, making it easier to calculate its duration.

speakers of American English produced /æ/ statistically longer than /ɛ/. Both vowels are phonemically lax and expected to have similar temporal durations, though /æ/ might be phonetically slightly longer due to its greater lip aperture accompanied by more time to be taken. However, results showed that /æ/ was significantly longer than /ɛ/. The tongue height gesture alone might not be possibly sufficient to make a full contrast between /ɛ/ and /æ/; thus, an additional feature of duration was involved to strengthen the phonetically longer duration of /æ/.
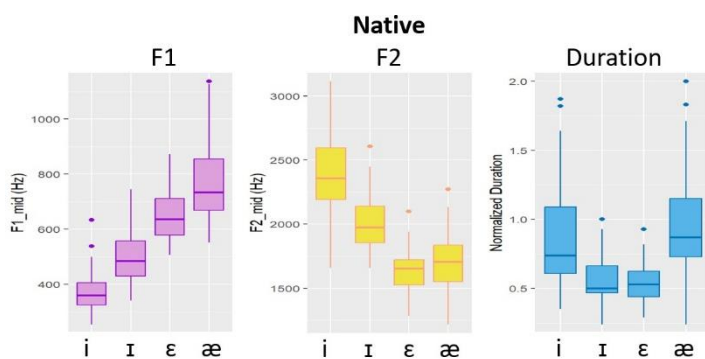

Figure 1. Acoustic Results of Front Vowels for Native Speakers

Table 3. Statistical Results of Front Vowels for Native Speakers

Contrast	F1	F2	Duration
i - ɪ	β = 130.73	β = -375.86	β = -0.329
	p < 0.005	**p < 0.005**	**p = 0.015**
ɛ - æ	β = 126.12	β = 52.99	β = 0.406
	p < 0.005	p = 0.339	**p = 0.020**

Note: Bold values denote statistical significance.

Figure 2 and Table 4 show the results of the high-proficiency ESL learners' group. Their acoustic patterns were very similar to those of native speakers in that they used significantly different F1 (β=98.36, p<0.005), F2 (β=-369.25, p<0.005), and duration (β=-0.271, p<0.005) to articulate the phonemic distinction of the high vowels /i/ and /ɪ/. For the pair of /ɛ/ and /æ/, F1 and duration were the only good predictors, similar to native speakers. F2 values do not seem to predict the distinctiveness between /ɛ/ and /æ/. In short, the ESL learners with high proficiency seemed to be as accurate as native speakers in the production of English front vowels.
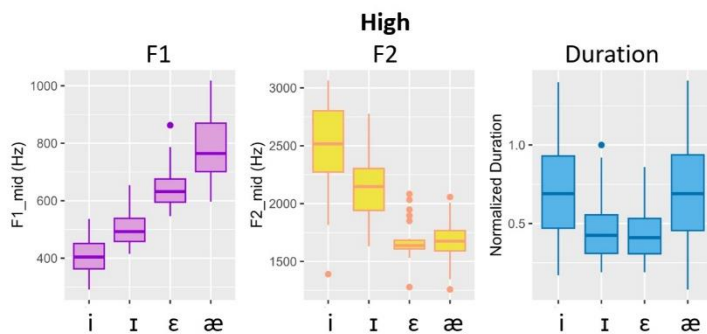

Figure 2. Acoustic Results of Front Vowels for Korean High-Proficiency Learners

Table 4. Statistical Results of Front Vowels for Korean High-Proficiency Learners

Contrast	F1	F2	Duration
i - ɪ	β = 98.36	β = -369.25	β = -0.271
	p < 0.005	**p < 0.005**	**p < 0.005**
ɛ - æ	β = 132.75	β = -52.99	β = -0.299
	p < 0.005	p = 0.339	**p < 0.005**

The intermediate learners, on the other hand, made only temporal distinctions in their production of /i/ and /ɪ/ (β=-0.265, p<0.005) and manipulated merely F1 to make different articulations of /ɛ/ and /æ/ (β=90.06, p<0.005) as displayed in Figure 3 and Table 5. Unlike native speakers and the high-proficiency learners who used three acoustic cues to differentiate high front vowels /i/ and /ɪ/ and two cues of them (F1 and duration) for the pair of /ɛ/ and /æ/, the intermediate-proficiency learners merely used one single cue for each pair of the front vowels, and they did not seem to achieve a full contrast for the front vowel phonemes.
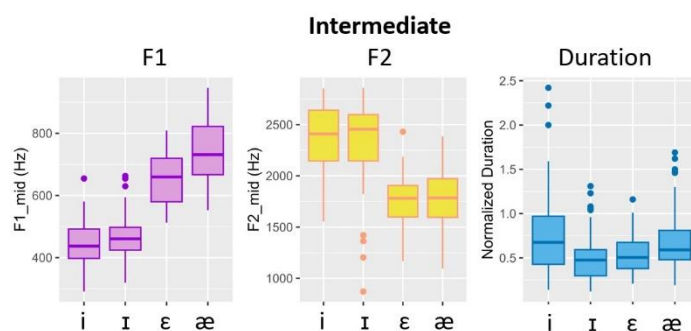

Figure 3. Acoustic Results of Front Vowels for Korean Intermediate-Proficiency Learners

Table 5. Statistical Results of Front Vowels for Korean Intermediate-Proficiency Learners

Contrast	F1	F2	Duration
i - ɪ	β = 20.95	β = -41.54	β = -0.265
	p = 0.239	p = 0.576	**p < 0.005**
ɛ - æ	β = 90.06	β = 9.70	β = 0.156
	p < 0.005	p = 0.92	p = 0.210

Interestingly, the low-proficiency learners' acoustic results seem to be very similar to those of the intermediate learners as illustrated in Figure 4. Such graphical outlooks were convinced by the results of the mixed-effect linear regression analysis as in Table 6. Normalized duration was the only predictor for making the vowels /i/ and /ɪ/ phonemically distinctive (β=-0.228, p=0.0356), and F1 was a mere cue for the pair of /ɛ/ and /æ/ (β=66.07, p=0.0013). For the Korean intermediate and low learners, neither tongue height (F1) nor frontness/backness (F2) was not used as a cue in contrasting English /i/ and /ɪ/ (all p>0.05). They simply produced /i/ longer than /ɪ/. This indicates that learners of moderate to heavy Korean accents tend to produce English /i/ simply as a long vowel and /ɪ/ as a shorter one, while minimally manipulating the tongue position. This is a pattern also observed in previous research (e.g., Flege et al., 1997, Yang, 2008). Concerning the contrast between /ɛ/ and /æ/, native speakers and high learners articulated /æ/ significantly longer than /ɛ/, but duration was not adopted to facilitate the contrast between /ɛ/ and /æ/ for the learners of intermediate and low proficiency.
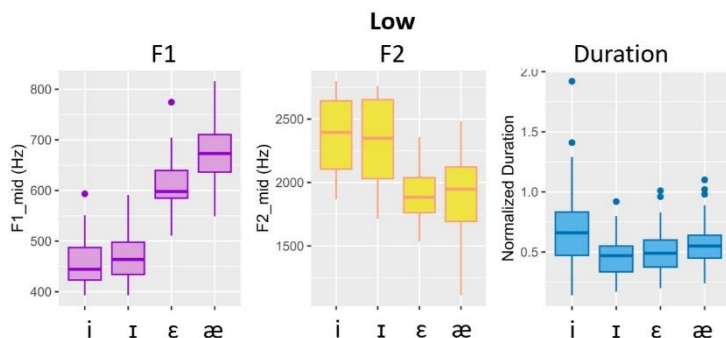
Figure 4. Acoustic Results of Front Vowels for Korean Low-Proficiency Learners

Table 6. Statistical Results of Front Vowels for Korean Low-Proficiency Learners

Contrast	F1	F2	Duration
i - ɪ	β = 15.956	β = -71.596	β = -0.228
	p = 0.272	p = 0.432	**p = 0.0356**
ɛ - æ	β = 66.07	β = -14.26	β = 0.084
	p = 0.0013	p = 0.903	p = 0.529

Moving on to the results of back vowels, Figure 5 shows the results of native speakers. Back vowels are more complicated than front vowels, because there are three confusing vowel pairs such as /ɔ/ vs. /ʌ/, /ʌ/ vs. /ɑ/, and /ɑ/ vs. /ɔ/. The graphical representations of three acoustic measures illustrate in Figure 5 that the vowels /ɔ/ and /ʌ/ pattern together for tongue height articulation as their F1 mean values look similar and that /ʌ/ and /ɑ/ pattern together for tongue frontness/backness as their F2 means appear equivalent.
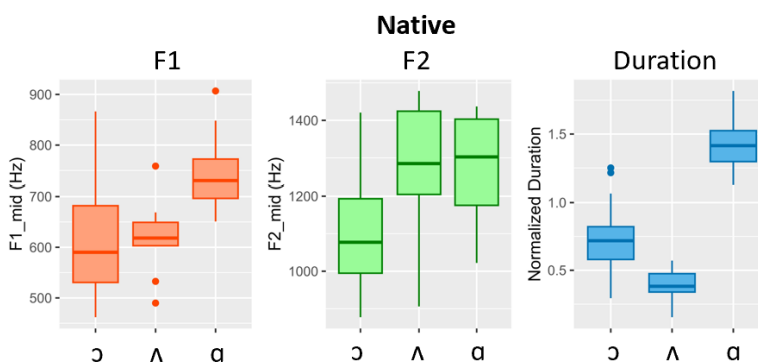

Figure 5. Acoustic Results of Back Vowels for Native Speakers

Table 7 presents the results of the regression analyses for the production of native speakers' back vowels /ɔ, ʌ, ɑ/. It shows that the native speakers used at least two acoustic features to distinguish the confusing vowel pairs of the three back vowels. For example, the low vowel /ɑ/ was produced with significantly higher F1, as compared to mid vowels (β=125.63, p<0.031 for /ʌ/ vs. /ɑ/, and β=123.24, p=0.0058 for /ɑ/ vs. /ɔ/). Since /ɑ/ is the only low vowel and the other two vowels /ɔ/ and /ʌ/ are mid vowels, tongue height lowering was significantly greater for /ɑ/ than for /ɔ/ and /ʌ/. In addition, F2 was significantly lower in the production of /ɔ/, a mere rounded vowel among them, which involved elongation of the front cavity due to lip-rounding. Hence, the difference in F2 values

was significant in the vowel pair /ɔ/ and /ʌ/ (β=176.32, p=0.0041) as well as the pair /ɑ/ and /ɔ/ (β=173.47, p=0.0049). Moreover, the native speakers made clear distinctions in all back vowels (/ɑ/ > /ɔ/ > /ʌ/) in normalized duration. A greater vocal opening for the low vowel /ɑ/ than for the mid vowel /ɔ/ yielded longer duration (β=0.694, p=0.010), while the duration of the lax vowel /ʌ/ was shorter than those of the tense vowels (β=1.033, p=0.0067 for /ʌ/ vs. /ɑ/, and β=0.339, p<0.005 for /ɔ/ vs. /ʌ/).

Table 7. Statistical Results of Back Vowels for Native Speakers

Contrast	F1	F2	Duration
ɔ - ʌ	β = -2.392 p = 0.975	β = 176.32 **p = 0.0041**	β = 0.339 **p < 0.005**
ʌ - ɑ	β = 125.63 **p = 0.031**	β = -3.149 p = 0.983	β = 1.033 **p = 0.0067**
ɑ - ɔ	β = 123.24 **p = 0.0058**	β = 173.47 **p = 0.0049**	β = 0.694 **p = 0.010**

The emphasis on temporal differences in back vowels was also seen in the data of the high learners (see Figure 6 and Table 8). Similarly to the native speakers, the high ESL learners produced significantly longer /ɑ/ and shorter /ʌ/, with /ɔ/ in the middle, yet distinct (all p<0.005). However, the F1 and F2 values in the high learners' back vowels were not as distinctive as in the native speakers' (all p>0.05). The high-proficiency learners showed the identical patterns to native speakers to produce front vowels, but their tongue height or tongue frontness/backness gestures were not sufficiently distinctive to differentiate the three pairs of back vowels respectively. That is, their production did not seem to result in a complete contrast between each pair.
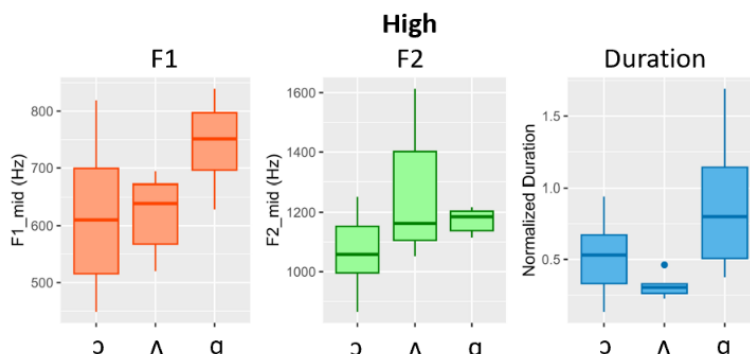

Figure 6. Acoustic Results of Back Vowels for Korean High-Proficiency Learners

Table 8. Statistical Results of Back Vowels for Korean High-Proficiency Learners

Contrast	F1	F2	Duration
ɔ - ʌ	β = 7.540 p = 0.933	β = 193.30 p = 0.08	β = -0.195 **p = 0.037**
ʌ - ɑ	β = 123.69 p = 0.324	β = -87.329 p = 0.435	β = 0.575 **p < 0.005**
ɑ - ɔ	β = 131.239 p = 0.212	β = 105.97 p = 0.262	β = 0.38 **p < 0.005**

Turning our attention to the results of Korean intermediate learners' production of back vowels, Figure 7 presents the mean values of F1, F2, and normalized duration. The back vowels produced by the intermediate

learners were not statistically different in general as shown in Table 9. The only significant difference was in the tongue height for the pair /ɑ/ vs. /ɔ/ (β=190.06, p=0.03). While native speakers used minimally two acoustic cues in order to make the back vowels distinctive, Korean intermediate learners merely used one (i.e., duration) similar to the high-proficiency learners. Again, the intermediate learners did not seem to achieve articulatory contrasts among back vowel phonemes except for the pair of /ɑ/ and /ɔ/.
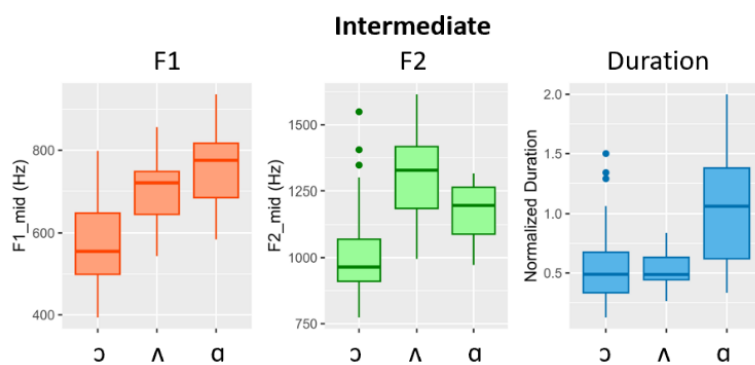

Figure 7. Acoustic Results of Back Vowels for Korean Intermediate-Proficiency Learners

Table 9. Statistical Results of Back Vowels for Korean Intermediate-Proficiency Learners

Contrast	F1	F2	Duration
ɔ - ʌ	β = 127.33	β = 295.61	β = -0.052
	p = 0.09	p = 0.067	p = 0.814
ʌ - ɑ	β = 62.72	β = -126.95	β = 0.529
	p = 0.408	p = 0.066	p = 0.132
ɑ - ɔ	β = 190.06	β = 168.66	β = 0.476
	p = 0.03	p = 0.206	p = 0.102

The mean values of F1, F2, and normalized duration of Korean low-proficiency learners are shown in Figure 8. The results of the mixed-effect linear regression analysis are presented in Table 10. The low learners differentiated F1 values in all three pairs (β=123.43, p=0.0013 for /ɔ/ vs. /ʌ/, β=139.60, p=0.0034 for /ʌ/ vs. /ɑ/, and β=263.03, p<0.005 for /ɑ/ vs. /ɔ/), a pattern different from that of native speakers of English or high or intermediate learners. Such stepwise higher F1 values from /ɔ/ to /ʌ/ and further to /ɑ/ are similar to the F1 values of Korean vowels (/a/ > /ʌ/ > /o/). As /ʌ/ is a low-mid vowel and /o/ is a high-mid vowel in Korean, the F1 value of /ʌ/ is higher than that of /o/. The F2 values were significantly different only in the pair of /ɔ/ and /ʌ/ (β=363.05, p=0.0003), but the other pairs did not show significant differences in F2. The particularly low F2 for /ɔ/ seems to be attributed to the extreme lip-rounding of Korean /o/. The low-proficiency learners attempted to round the lips strongly as if they had been pronouncing the mid back vowel /o/ in Korean; for example, the word *small* was realized as [smol], not [smɔːl], exhibiting a strong Korean accent. In addition, the low learners made non-significant distinctions in duration, unlike high and intermediate learners.
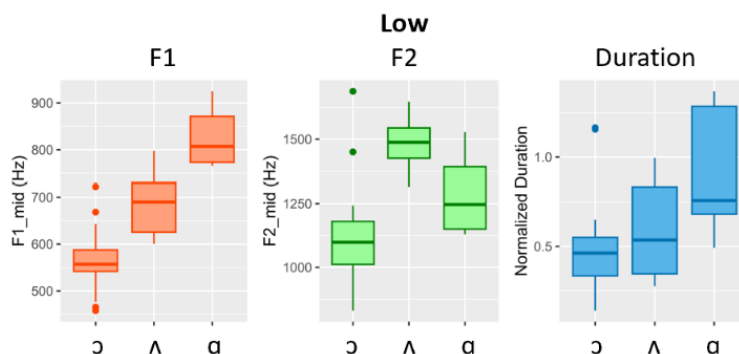
Figure 8. Acoustic Results of Back Vowels for Korean Low-Proficiency Learners

Table 10. Statistical Results of Back Vowels for Korean Low-Proficiency Learners

Contrast	F1	F2	Duration
ɔ - ʌ	β = 123.43 **p = 0.0013**	β = 363.05 **p = 0.0003**	β = 0.103 p = 0.58
ʌ - ɑ	β = 139.60 **p = 0.0034**	β = -194.70 p = 0.08	β = 0.320 p = 0.227
ɑ - ɔ	β = 263.03 **p < 0.005**	β = 168.35 p = 0.063	β = 0.423 p = 0.08

3. Discussion

The present study investigated the developmental variations in L2 vowel learning by observing between-group differences in the values of F1 and F2 and the duration of the English vowels /i, ɪ, ɛ, æ, ɔ, ʌ, ɑ/ produced by Korean ESL learners of distinct accentedness. The acoustic analyses revealed spectral and temporal properties of Korean-accented English vowels, and comparisons of the between-group differences showed that Korean ESL learners are in the developmental stage of learning L2 English vowels.

In distinguishing vowel contrasts, the high-proficiency learners resembled the native speakers, especially in their production of front vowels. In the case of back vowels, they manipulated temporal duration in such a way of /ɑ/ > /ɔ/ > /ʌ/, which was exactly the same as native speakers. Both /ɑ/ and /ɔ/ are tense vowels, which should be longer in duration than /ʌ/. Moreover, /ɑ/ accompanies a greater oral aperture than /ɔ/, which will induce temporal elongation, resulting in longer duration. Korean high learners seemed to have temporal contrasts in their articulation of the three back vowels. However, the differences in tongue height or frontness/backness gestures were not statistically significant in the mixed-effect linear regression analysis. In other words, F1 or F2 values were not good predictors for the high learners' articulation of back vowels. This suggests that they successfully accomplished complete contrasts among front vowels but made partial or incomplete articulatory contrasts for back vowels, merely focusing on duration.

In the case of Korean low learners, how they make spectral and temporal distinctions differed from the ways of the native speakers for both front and back vowels. They merely relied on durational differences between /i/ and /ɪ/ and tongue height (F1) distinctions between /ɛ/ and /æ/ for front vowels. Recall that native speakers manipulated tongue height, tongue frontness/backness, and duration for the contrast of /i/ and /ɪ/ and tongue height and duration for the distinction of /ɛ/ and /æ/. The number of articulatory cues was considerably limited in Korean low learners,

which brought about incomplete contrasts among front vowels. In addition, their production of back vowels showed that low learners attempted to contrast the back vowels in different manners from native speakers. While the tongue height gesture reflected by F1 differed in such a way of /ɑ/ > /ʌ/ ≈ /ɔ/ in the production of native speakers, Korean low learners showed /ɑ/ > /ʌ/ > /ɔ/. The F2 values reflecting the tongue frontness gesture also totally differed between native speakers and the low learners in such a way that the low group differentiated the back vowels in /ʌ/ > /ɑ/ ≈ /ɔ/ and that the native speakers did in /ɑ/ ≈ /ʌ/ > /ɔ/. As mentioned in the previous section (2.5. Results), the low learners seemed to attempt to make articulatory contrasts among back vowels, yet incorrectly. The acoustic properties found in low learners' L2 speech could be attributed to the negative transfer from the Korean vowel system. Korean low learners seemed to make *incomplete* contrasts of front vowels and *incorrect* contrasts of back vowels in their production of English vowels.

Korean intermediate learners shared the identical patterning of front vowel production to Korean low learners in that they contrasted /i/ and /ɪ/ only in the temporal dimension and they differentiated /ɛ/ and /æ/ merely in F1. While native speakers took advantage of three features—F1, F2, and duration—to distinguish the phonemic contrast of /i/ and /ɪ/ and F1 and duration for the pair of /ɛ/ and /æ/, Korean intermediate learners used a limited number of features like low learners, suggesting that they produced incomplete contrasts of English front vowels. Concerning their production of back vowels, almost no pairs showed statistical differences in F1, F2, or normalized duration, except F1 between /ɑ/ and /ɔ/. That is, the local differences between back vowels in each pair were not significant enough to induce the articulatory contrast the same way as they did for native speakers. This indicates that Korean intermediate learners did not achieve phonemic contrasts in the articulation of English back vowels.

Korean ESL learners' articulatory patterning of English vowels is summarized in Table 11. It suggests an earlier development of front vowel contrasts relative to back vowels in L2 vowel learning. For example, the Korean high learners who successfully articulated the front vowels in both spectral and temporal dimensions did not even maintain such distinction in back vowels. They merely manipulated temporal differences among them (/ɑ/ > /ɔ/ > /ʌ/), which was not necessarily required, because the duration was shown /ɑ/ ≈ /ɔ/ > /ʌ/ in native speakers' production. In other words, the high learners already completed the contrastive production of front vowels as opposed to the incomplete contrasts for back vowels. The intermediate and low learners did not even use the temporal cues to distinguish the back vowels. They similarly produced incomplete contrasts for front vowels while they made almost no contrasts or even incorrect contrasts for back vowels.

Table 11. Korean ESL Learners' Production of Vowel Contrasts

	High	Intermediate	Low
Front vowels	**Complete** contrasts	**Incomplete** contrasts	**Incomplete** contrasts
Back vowels	**Incomplete** contrasts	**(Almost) No** contrasts	**Incorrect** contrasts

The developmental variation in front vowels vs. back vowels could be due to the acoustic and/or articulatory space in the learners. While there are two confusing contrast pairs (/i/-/ɪ/ and /ɛ/-/æ/) in the front vowel dimension, there are three pairs among back vowels (/ɔ/-/ʌ/, /ʌ/-/ɑ/, and /ɑ/-/ɔ/), in addition to the high back vowel pair /u/-/ʊ/, which was excluded in the present study. Then, the space for back vowels becomes crowded, causing one vowel to crowd into the space of another vowel nearby. The greater density in the English back vowel system thus yields a greater learning load (four back vowel pairs vs. two front vowel pairs) for Korean learners in realizing and producing each pair distinctively, which had been predicted and confirmed in previous studies (Koo 2000).

In addition to the crowded vowel space, the low-proficiency learners' confusion of English and Korean back vowels seemed to increase learning difficulty. In this study, the low group differentiated the back vowels only in F1 (/ɑ/ > /ʌ/ > /ɔ/), which was dissimilar to those of native speakers (/ɑ/ > /ʌ/ ≈ /ɔ/). They seemed to produce /ɔ/

as a similar vowel in their L1 Korean, /o/, whose F1 is lower than /ʌ/. Not only that, their /ɔ/ also involved lip-rounding (as shown by low F2), in a way similar to the production of Korean /o/. Although the present study did not include perception data, the case of similar vowels English /ɔ/ and Korean /o/ suggests that these vowels have been perceived by low-proficiency learners as being part of the same (single) category. Such vowels conform to the prediction of Flege's SLM (1995) that L2 vowels similar to the ones in L1 would be greatly difficult for low learners to distinguish, contributing to their accentedness.

When Korean learners made incomplete contrasts in L2 vowel production as presented in Table 11, they did not use as many articulatory and/or acoustic features as native speakers. For example, the intermediate and low learners showed differences in the production of contrasts between the /i/-/ɪ/ pair and the /ɛ/-/æ/ pair. They contrasted /i/ and /ɪ/ using temporal cues only and /ɛ/-/æ/ with only F1, unlike the native and high learners who contrasted the /i/-/ɪ/ pair using both spectral and temporal features (i.e., F1, F2, and duration) and distinguished the /ɛ/-/æ/ pair in the tongue height (F1) and temporal dimensions. It can be interpreted that they recognize and/or perceive the contrast of English front vowels, but their perception has not been appropriately realized or applied in articulation yet. Such a middle stage of incomplete contrasts, emerging from no contrasts or incorrect contrasts to complete contrasts, can be possibly a step of *covert contrasts* (Eckman et al. 2015) in L2 sound development. This indicates a disparity between learners' perceptual awareness and articulatory proficiency, which is consistent with Flege's (1995) assertion that perception advances production in L2 speech learning. Future studies that collect and analyze both production and perception data from L2 learners will help corroborate the processes involved in the development of L2 vowel contrasts.

4. Conclusion

In conclusion, this study contributes to our understanding of the developmental trajectory of Korean ESL learners' L2 vowel learning (from the contrast /i/-/ɪ/ to the contrast /ɛ/-/æ/, and then the back vowel contrasts) and highlights the complexities involved in the process. The findings of the study provide implications for Korean learners of English who aim to improve their accentedness: An essential step would be to diagnose and evaluate the vowel duration as well as tongue movement in their L2 vowel production and then to help them attend to the necessary acoustic and/or temporal features for target L2 vowels. The findings also extensively suggest the need for distinct pedagogical methods for learning back vowels which learners will need more time to learn than front vowels. Finally, the limitations of the study must be noted. The vowels examined in this study were not evenly distributed (as seen in Table 2) because the original text (Weinberger 2015) had considerably more tokens for front vowels than for back vowels. Further comprehensive studies on L2 vowel learning are thus desired.

References

Baker, W., P. Trofimovich, M. Mack and J. E. Flege. 2002. The effect of perceived phonetic similarity on non-native sound learning by children and adults. In Proceedings of *the 26th Annual Boston University Conference on Language Development,* 36-47.

Best, C. T. 1995. A direct realist view of cross-language speech perception. In W. Strange, ed., *Speech Perception and Linguistic Experience,* 171-204. York Press.

Boersma, P. and D. Weenink. 2024. Praat: doing phonetics by computer (version 6.4.14) [Computer software].

Available online at http://www.praat.org/

Eckman, F., G. Iverson and J. Song. 2015. Overt and covert contrast in L2 phonology. *Journal of Second Language Pronunciation* 1(2), 254-278.

Flege, J. E. 1988. Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84(1), 70-79.

Flege, J. E. 1995. Second-language speech learning: Theory, findings, and problems. In W. Strange, ed., *Speech Perception and Linguistic Experience,* 233-277. York Press.

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

Hong, S. 2013. Korean talker's cue weighting perception strategies in perceiving English /ɑ/, /ɔ/ and /ʌ/ in comparison with American talkers'. *Studies in Phonetics, Phonology, and Morphology* 19(3), 529-554.

Hwang, I. and S. Lee. 2012. Perception of English vowel categories by Korean university students. *Korean Journal of Linguistics* 37(4), 1095-1117.

Ioup, G. 2008. Exploring the role of age in the acquisition of a second language phonology. In J. H. Edwards and M. Zampini, eds., *Phonology and Second Language Acquisition*, 41-62. John Benjamins.

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

Lee, J.-K. and Y.-W. Lee. 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, S. and S. C. Rhee. 2019. The relationship between vowel production and proficiency levels in L2 English produced by Korean EFL learners. *Phonetics and Speech Sciences* 11(2), 1-13.

Munro, M. J. 1993. Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech* 36(1), 39-66.

Oh, E. 2013. Dynamic spectral patterns of American English front monophthong vowels produced by Korean-English bilingual speakers and Korean late learners of English. *Linguistic Research* 30(2), 293-312.

Pevnick, G. n.d. Online Forced Aligner [Web-based tool]. University of Wisconsin-Milwaukee. Available online at http://web.uwm.edu/forced-aligner

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

Weinberger, S. 2015. *Speech Accent Archive*. George Mason University. Available online at http://accent.gmu.edu

Yang, B. 2008. An acoustical comparison of English tense and lax vowels produced by Korean and American males. *Speech Sciences* 15(4), 19-27.

Appendix
Speech Stimuli (from Weinberger 2015)

Please call Stella. Ask her to bring these things with her from the store: six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary