Korean Journal of English Language and Linguistics, Vol 26, May 2026, pp. 651-670.
DOI: 10.15738/kjell.26..202605.651

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Intra-Speaker Variability and Perceptual Sensitivity of English Vowels in Advanced Korean EFL Learners

Seyeon Choe · Hyunkee Ahn (Seoul National University)

Received: March 10, 2026
Revised: March 28, 2026
Accepted: April 20, 2026

Choe, Seyeon (First author)
Ph.D. candidate, Department of English Language Education
Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul, Korea
Tel: +82-2-880-7670
Email: caomei214@snu.ac.kr

Ahn, Hyunkee (Corresponding author)
Professor, Department of English Language Education
Seoul National University
Tel: +82-2-880-7673
Email: ahnhk@snu.ac.kr

ABSTRACT

Choe, Seyeon and Hyunkee Ahn. 2026. Intra-speaker variability and perceptual sensitivity of English vowels in advanced Korean EFL learners. *Korean Journal of English Language and Linguistics 26*, 651-670.

The present study investigates Korean EFL learners' intra-speaker variability and perceptual sensitivity in producing three American English vowel contrasts (/i–ɪ/, /ɛ–æ/, and /u–ʊ/). Thirty advanced Korean learners and eight native English speakers produced target vowels embedded in carrier phrases, and acoustic analyses were conducted using normalized Bark difference measures. Perceptual sensitivity was assessed using the A' score to examine how perception relates to production variability. Paired-samples *t*-tests revealed no statistically significant differences in intra-speaker variability between *new* vowels (/ɪ, æ, ʊ/) and their corresponding *similar* vowels (/i, ɛ, u/) in vowel height. A significant difference was observed for the /i–ɪ/ contrast in vowel backness. Within the four-scenario framework of Smith et al. (2019), 81% of tokens fell into Scenario 1 (similar formants and similar variability) and 14% into Scenario 3 (different formants but similar variability), indicating that approximately 95% of productions showed native-like variability. Notably, the distribution of *similar* and *new* vowels across scenarios was the reverse of that reported by Smith et al. (2019), suggesting that learning environment (EFL vs. ESL) may modulate the formation of L2 phonetic categories predicted by the Speech Learning Model. Despite these production patterns, learners differed in perceptual sensitivity, and those sensitive to all three contrasts showed clearer separation among vowel categories than those who were not. These findings suggest that learners' acoustic production patterns should be considered alongside perceptual sensitivity in characterizing L2 speech.

KEYWORDS

intra-speaker variability, perceptual sensitivity, American English vowel contrasts, acoustic analysis

1. Introduction

Studies have revealed that the impact of segmental errors on communication can be substantial in determining speech intelligibility (Bent et al. 2007). Among these errors, vowels play a more pivotal role than consonants in English intelligibility (Bent et al. 2007, Fogerty and Kewley-Port 2009, Levis 2018). This is because vowels convey not only their own phonological information but also coarticulatory cues from adjacent consonants, even in a simple English word such as *pat*. In this sense, vowels carry more information than consonants (Bent et al. 2007, Levis 2018). Despite their importance, accurate perception of English vowel contrasts is often challenging for second language (L2) learners, particularly when the contrasts do not exist in the learners' native phonological system (Best 1995, Flege 1995).

Research has shown that Korean learners tend to assimilate three English vowel pairs—/i–ɪ/, /ɛ–æ/, and /u–ʊ/—into the single Korean vowels /i/, /ɛ/, and /u/, respectively (Hong 2007, 2012, Ingram and Park 1997, Silva 2004, Tsukada et al. 2005, Yang 1996). Thus, the three English vowel pairs /i–ɪ/, /ɛ–æ/, and /u–ʊ/ have consistently been reported as particularly difficult for Korean learners (Hong 2007, Ingram and Park 1997, Kim and Kim 2003, Silva 2004, Tsukada et al. 2005). These learning difficulties can be understood through two major frameworks of L2 speech learning: the Perceptual Assimilation Model (PAM, Best 1995) and the Speech Learning Model (SLM, Flege 1995). While PAM primarily accounts for how non-native contrasts are initially assimilated to native categories based on perceptual similarity, SLM explains how such contrasts may either result in equivalence classification or lead to the formation of new phonetic categories over time with continued L2 experience (Flege 1995). According to SLM, learners are more likely to form separate phonetic categories for *new* sounds (/ɪ, æ, ʊ/) than for *similar* ones (/i, ɛ, u/) as learning progresses (Flege 1995, Lee and Cho 2015). While much is relatively known about how Korean learners perceive and categorize English vowels, less is understood about how consistently they produce these sounds across tokens or repetitions. Intra-speaker variability, in this regard, may offer a useful window into the stability of phonetic categories, although its relationship to learning processes remains to be fully established.

A defining feature of learner language is *variability* (Verspoor et al. 2008). Yet few studies have explored intra-speaker phonetic variability in L2 vowel production (Smith et al. 2019). The concept of variability itself remains vague, partly because it is difficult to distinguish normal developmental variation from instability in speech production (Bley-Vroman 1983). To address this gap, Smith et al. (2019) sought to define intra-speaker variability empirically, describing it as "how consistent a given L2 speaker is in achieving a particular production pattern relative to the 'stability' shown by native speakers" (p. 159). They compared the intra-speaker variability of English vowel pairs (/i–ɪ/, /e-ɛ/, and /u–ʊ/) produced by Korean, Mandarin, and Spanish learners of English with that of native English speakers. Using a four-scenario framework, they examined the relationship between vowel formant accuracy and production stability (Smith et al. 2019, p. 142):

Scenario 1. L2 speakers similar to L1 speakers in both vowel formants and variability;
Scenario 2. L2 speakers similar in formants but more variable;
Scenario 3. L2 speakers different in formants but equally stable;
Scenario 4. L2 speakers different in both formants and variability.

The first and third scenarios indicate native-like degrees of intra-speaker variability. When a speaker's vowel formants and coefficient of variation (CV) fell within ± 2 standard deviations of the native mean, their productions were deemed native-like. Remarkably, the result showed that 79% of non-native speakers exhibited native-like intra-speaker variability regardless of whether their formant values were native-like. This finding provided empirical evidence that non-native speakers are not necessarily more variable than native speakers, even if their productions deviate acoustically from native norms.

Despite this contribution, several aspects of Smith et al.'s (2019) study merit further investigation. First, the target vowels in their study were selected without considering learners' L1 backgrounds, even though the relative difficulty of non-native contrasts depends on the L1 (Best and Tyler 2007). Second, they did not assess learners' perception of the vowel contrasts, leaving a potential key source of production deviation unexplored. Third, they did not account for the 21% of participants who showed non-native-like intra-speaker variability, which may relate to differences in L2 proficiency (e.g., Saito et al. 2016). Building on this foundation, the present study aims to replicate and extend Smith et al. (2019) by investigating advanced Korean EFL learners' intra-speaker variability and perceptual sensitivity in producing three American English vowel pairs /i–ɪ/, /ɛ–æ/, and /u–ʊ/. Focusing on advanced learners allows us to examine how L2 vowel categories are realized at later stages of learning. To this end, Korean learners' vowel productions were acoustically analyzed, and their perceptual sensitivity was assessed (A′ score). Accordingly, the present study addresses the following research questions:

1. Do advanced Korean EFL learners' productions of the *new* English vowels (/ɪ, æ, ʊ/) show a comparable degree of intra-speaker variability to their productions of the *similar* vowels (/i, ɛ, u/)?
2. How are advanced Korean EFL learners distributed across the four scenarios when both production and perceptual sensitivity to each vowel contrast are considered, and how does this distribution compare to Smith et al. (2019)?

2. Literature Review

2.1 Cross-Linguistic Differences in the English and Korean Vowel Systems

American English vowels can be characterized as consisting of eleven non-rhotic segments: five front unrounded vowels (/i, ɪ, e, ɛ, æ/), two back unrounded vowels (/ʌ, ɑ/), and four back rounded vowels (/u, ʊ, o, ɔ/) (Strange et al. 2007). Notably, the mid front and mid back vowels (/e, o/) are typically realized as diphthongized [eɪ] and [oʊ], respectively (Strange et al. 2007). Phonemically, English vowels can be described in terms of tongue height, tongue backness, tenseness, and lip rounding (Giegerich 1992). Tongue height divides vowels into roughly high, mid, and low categories within both the front and back series (Giegerich 1992, Strange et al. 2007). In addition, vowels contrast in *tenseness*, which is accompanied by a systematic difference in duration: tense vowels (/i, e, u, o/) are intrinsically longer than their lax counterparts (Strange et al. 2007). Nevertheless, vowel length is regarded as a redundant feature in English, as no absolute durational criterion exists and listeners primarily rely on vowel quality rather than quantity when distinguishing tense–lax contrasts (Giegerich 1992). Finally, lip rounding occurs only in back vowels (/u, ʊ, o, ɔ/) (Giegerich 1992, Ladefoged and Johnson 2015).

In contrast, Korean has a more restricted vowel inventory that can be analyzed as either eight vowels—/i, e, ɛ, ɑ, ʌ, o, ɯ, u/ (Shin 2015, Tsukada et al. 2005)—or ten when the front rounded vowels /y/ and /ø/ are included (Lee and Ramsey 2000, Yang 1996). However, this inventory is undergoing notable diachronic change. First, the front mid vowels /e/ and /ɛ/ are merging due to the progressive raising of /ɛ/ (Ingram and Park 1997, Lee and Ramsey 2000, Shin 2015, Yang 1996). Second, the historical phonemic vowel-length contrast (e.g., /i/ vs. /iː/) has virtually disappeared (Lee and Ramsey 2000). Consequently, modern Korean tends toward a streamlined front-vowel system in which /i/ and /e/ predominate, while /ɛ/ and long-vowel counterparts are diminishing in use. The most crucial cross-linguistic distinction, therefore, lies in the absence of an English-like tense–lax contrast in Korean. These structural differences between the two vowel systems are summarized in Table 1, which highlights the major phonological contrasts relevant to the present study.

Table 1. Comparison of English and Korean Vowels (adapted from Ladefoged and Johnson 2015, Shin 2015)

	Front		Central		Back		
Height	English	Korean	English	Korean	English	Korean	
High	i*	i (ㅣ)			u*	ɯ** (ㅡ)	u (ㅜ)
	ɪ				ʊ		
Mid	e*	ɛ (ㅔ/ ㅐ)	ʌ		o*	ʌ** (ㅓ)	o (ㅗ)
	ɛ				ɔ		
Low	æ				ɑ*	ɑ** (ㅏ)	

Notes. * denotes English tense vowels, ** denotes Korean unrounded back vowels.

An acoustic comparison of American English and Korean vowels by Yang (1996) shows that the two languages exhibit different vowel-space configurations even after speaker-related variation has been normalized. Yang (1996) shows that the Korean vowel space is characterized as wedge-shaped, with [i, a, u] forming its peripheral anchors, whereas the American English vowel space is comparatively more rectangular. These cross-linguistic differences can be interpreted in terms of sufficient perceptual contrast: in American English, lax vowels are positioned more internally, thereby increasing perceptual distance from adjacent vowels, while in Korean the high-vowel region is more expanded and the low-vowel region less internally differentiated (Yang 1996). This structural difference is particularly evident in the comparison of English tense–lax vowel pairs and their Korean counterparts in Yang (1996). For the high front vowels, the English contrast /i–ɪ/ involves a peripheral [i] and a more centralized [ɪ], whereas Korean has only a single high front vowel /i/, resulting in a broader acoustic space corresponding to both English vowels. Similarly, in the low front vowel region, the English contrast /ɛ–æ/ reflects an expanded low-vowel space in which [æ] occupies a more extreme position to maintain perceptual distinctiveness, while Korean /e/ (or /ɛ/) tends toward reduced differentiation and shows relatively less internal differentiation. In the high back vowel region, the English pair /u–ʊ/ contrasts a tense [u] with a more centralized lax [ʊ], whereas Korean again has a single vowel /u/, whose relatively low F2 value can be understood as maintaining sufficient perceptual distance from neighboring high vowels such as /ɨ/. Taken together, these patterns indicate that American English can be interpreted in terms of tense–lax differentiation that increases internal differentiation within the vowel space, whereas Korean relies on a smaller set of more widely dispersed vowel categories. This asymmetry in vowel-system organization provides a phonetic basis for the well-documented difficulties Korean learners experience in perceiving and producing English vowel contrasts.

2.2 English Vowels for Korean Learners

For Korean learners of English, vowels present one of the most persistent challenges in second language (L2) speech perception and production. Research has consistently shown that many English vowel contrasts are assimilated into a smaller set of Korean vowel categories, leading to what is often termed single-category (SC) assimilation (Lee and Cho 2015, Lee and Cho 2018, Yun 2014) in Perceptual Assimilation Model (PAM, Best 1995). As illustrated in previous mapping studies, the English vowel pairs /i–ɪ/, /ɛ–æ/, and /u–ʊ/ are typically perceived as corresponding to the single Korean vowels /i/, /ɛ/, and /u/, respectively. This pattern reflects the limited vowel inventory of Korean compared to English, as well as ongoing phonological change in Korean itself. In particular, the merger of the front mid vowels /e/ and /ɛ/ has been widely reported (Ingram and Park 1997, Lee and Ramsey 2000, Shin 2015, Yang 1996), further reducing the perceptual distinctiveness of English /ɛ–æ/ for Korean listeners.

According to the PAM (Best 1995), the difficulty of discriminating non-native phoneme contrasts depends on how the sounds are perceptually assimilated to native phonological categories. Contrasts assimilated to two different native categories (TC) are predicted to be discriminated most accurately, followed by category-goodness contrasts (CG), while single-category contrasts (SC) are predicted to be the most difficult to discriminate (Best 1995, Best et al. 2001). Empirical studies using identification and discrimination tasks have consistently demonstrated differential difficulty across English vowel contrasts for Korean learners. Hong (2012), employing a forced-choice identification test with *hVd* stimuli, reported that the /i–ɪ/ pair was easiest to identify, followed by /u–ʊ/, whereas /ɛ–æ/ posed the greatest difficulty. A similar pattern was observed in Suh (2020), who used both assimilation and discrimination tests and found that Korean learners discriminated /i–ɪ/ most accurately, outperforming both /ɛ–æ/ and /u–ʊ/. Notably, Suh (2020) classified the English vowel pair /i–ɪ/ as a Category Goodness (CG) assimilation, while the other pairs (/ɛ–æ/ and /u–ʊ/), which did not show statistically significant differences in goodness ratings, were categorized as Single Category (SC) assimilations. Given the PAM hierarchy (TC > CG > SC), where categories become increasingly easier to discriminate toward the TC, the /i–ɪ/ contrast, categorized as a CG assimilation, is expected to yield higher discrimination accuracy than the SC pairs /ɛ–æ/ and /u–ʊ/.

Beyond the classification of vowel pairs in terms of PAM assimilation types, it is also necessary to consider the status of individual vowels within each contrast, particularly in terms of whether they function as *similar* or *new* sounds for Korean learners. While PAM provides a framework for predicting discrimination difficulty based on perceptual assimilation patterns (e.g., TC, CG, SC), it does not explicitly address how individual L2 sounds are phonologically represented or produced. Importantly, the distinction between new and similar sounds in SLM is grounded in perceived L1–L2 similarity, such that sufficiently dissimilar sounds tend to form new phonetic categories, whereas similar sounds are assimilated into existing L1 categories (Aoyama et al. 2004, Flege 1995). To account for this dimension, insights from L2 speech learning models such as the SLM (Flege 1995) become relevant. Previous studies (Han et al. 2011, Lee and Cho 2015, Lee and Cho 2018) have classified English vowels such as /i/, /ɛ/, and /u/ as *similar* vowels, given their perceptual overlap with Korean categories, and /ɪ/, /æ/, and /ʊ/ as *new* vowels that require the formation of novel phonetic categories. Within this perspective, L2 speech learning involves two distinct processes: the creation of a new category for *new* sounds and the adjustment of an existing category for *similar* sounds (Escudero 2009). Empirical evidence suggests that L2 learners tend to develop more stable phonetic representations for *new* vowels than for *similar* ones, which remain influenced by L1 category structure (Flege 1995, 2003).

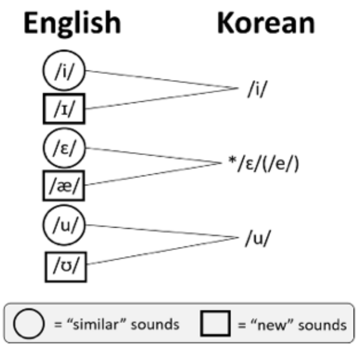

Figure 1. Perceptual Mappings of English Vowels to Korean Vowels

Note. * Merging of the two Korean sounds is underway.

In short, English vowels represent an area of systematic asymmetry between the two languages. Korean learners' perceptual mapping of multiple English vowels onto fewer Korean categories, together with the ongoing /e–ɛ/ merger in Korean, leads to inherent challenges in acquiring the English tense–lax distinctions. These cross-linguistic relationships and assimilation patterns are summarized in Figure 1, which illustrates how differences between the English and Korean vowel systems underlie Korean learners' difficulties in English vowel perception and production. Building on this distinction, the present study examines whether production patterns differ systematically between *new* and *similar* vowels across vowel contrasts categorized under PAM, thereby linking perceptual assimilation with similarity-based phonetic category formation in the SLM and its production outcomes.

2.3 Major Differences from the Previous Smith et al.'s Study

Smith et al. (2019) examined the intra-speaker variability of ESL learners' production of American English tense and lax vowels. Thirty L2 speakers from Korean, Mandarin, and Spanish backgrounds and ten native English speakers produced CVC words containing the vowel pairs /i–ɪ/, /e–ɛ/, and /u–ʊ/. Acoustic measurements of F1 and F2 formants were analyzed using coefficients of variation to assess production variability. The results indicated that while L2 learners often produced vowel formants that differed from native speakers, their production variability was rarely greater than that of native speakers. Most learners showed similar levels of variability regardless of whether their vowels were native-like or not, suggesting that stable phonetic targets can exist even when the acoustic realization differs from native norms.

The present study extends and refines the work of Smith et al. (2019) in several important respects. First, while Smith et al. (2019) examined learners in an English as a Second Language (ESL) environment, the current study focuses on learners in an English as a Foreign Language (EFL) context. Second, the participant pool was narrowed to include only native English (NE) and native Korean (NK) speakers, thereby allowing for a more controlled cross-linguistic comparison. Third, dialectal variation was strictly controlled in the present study, unlike in Smith et al. (2019). Fourth, participants' English proficiency level was considered to account for potential effects of individual differences on vowel production stability. Fifth, the target vowels were refined to three difficult contrasts for Korean learners—/i–ɪ/, /ɛ–æ/, and /u–ʊ/—with an additional control pair (/i–ɑ/) for the perception task. The original /e–ɛ/ contrast in Smith et al. (2019) was replaced with /ɛ–æ/, as tense vowel /e/ generally surfaces as a diphthong [eɪ], increasing the perceptual salience and reducing the difficulty of the contrast for Korean listeners (Tsukada et al. 2005). Sixth, the present study employed *hVd* stimuli instead of *CVd* tokens, providing a more standardized phonetic frame for analyzing vowel formants. Finally, unlike Smith et al. (2019), this study incorporated a perceptual component, enabling a direct comparison between production variability and perceptual sensitivity.

3. Methodology

3.1 Participants

Native English speakers were recruited through personal referral, whereas Korean EFL learners were recruited via online and offline advertisements. Eight native speakers of American English (NE: 4 females, 4 males) participated, recruited via snowball sampling. Their ages ranged from 23 to 35 years (M = 29, SD = 4.21), and all but one had lived in Korea for more than one year. In addition, 30 Korean EFL learners (NK: 15 females, 15 males) participated, with a mean age of 24 years (SD = 2.9). All participants were classified as advanced based on TOEIC listening comprehension (LC) scores (max = 495, M = 467, SD = 26) and had resided in Seoul or Kyung-gi, speaking Standard Korean. Although the study was initially designed to include learners across multiple proficiency levels, the final sample consisted only of advanced learners due to recruitment constraints. As a result, the findings of the present study should be interpreted as reflecting performance within a relatively homogeneous, high-proficiency group. Eleven participants reported study-abroad experience in English-speaking countries (length of residence: M = 1.6 years, SD = 1.9). This study was approved by the Institutional Review Board of Seoul National University (IRB No. 2007/003-022).

3.2 Speaking Stimuli

Four English vowel pairs were selected for this study. Three target pairs (/i–ɪ/, /ɛ–æ/, /u–ʊ/), which are known to be difficult for most native Korean learners of English to discriminate, were primarily analyzed. An additional pair (/i–ɑ/) was included as a control to verify the validity of the perceptual task (Tsukada et al. 2005). Seven *hVd* words were chosen to elicit the seven English vowels in sequence: *heed, hid, had, head, who'd, hood,* and *hod*. Unlike Smith et al. (2019), all target words were restricted to the *hVd* form in order to minimize contextual coarticulatory influences, as phonetic contexts can inflate variability (Xie and Jaeger 2020). The *hVd* frame was deliberately selected for its onset and coda. The glottal fricative /h/ was chosen because, like vowels, it involves the entire vocal tract in articulation and is expected to exert minimal influence on the following vowel (Ladefoged and Johnson 2015). The alveolar stop /d/ was used as the coda, since it lengthens the preceding vowel more than its voiceless counterpart /t/, thereby increasing the likelihood of capturing a longer steady state in vowel production, even if the difference is relatively small (Ladefoged and Johnson 2015). To compensate for the limited phonetic contexts, the number of sentence readings was increased from three as in the previous study by Smith et al. (2019) to five. A carrier sentence, *"I like to say ______ some of the time"* (Smith et al. 2019), was employed to help speakers maintain a consistent speech rate (Munro 1993). Phonemic symbols for each target word were provided right below the orthographic forms.

3.3 Perceptual Sensitivity Task

A categorical discrimination task (A′ score, Snodgrass et al. 1985) was used to assess Korean listeners' perceptual sensitivity to three English vowel contrasts (/i–ɪ/, /ɛ–æ/, /u–ʊ/). A control pair, /i–ɑ/, known to be relatively easy for Koreans to discriminate (Tsukada et al. 2005), was included to verify task validity and participants' ability to complete the trials. Each trial contained three tokens produced by different native English speakers with balanced gender. Two trial types were constructed: *change trials*, which contained one odd item (e.g., hid_5 hid_3 $heed_7$), and *no-change trials*, which contained three productions of the same vowel by different speakers (e.g., $heed_3$ $heed_5$ $heed_8$).

Stimuli were derived from *hVd* words segmented in *Praat* (Version 6.1.20 beta, Boersma and Weenink 2020) and compiled in *Audacity* (Version 1.3.4-beta, Audacity Team 2008). Each token was separated by 800 ms of silence, with 500 ms of silence at the beginning and end of each trial. Sixty-four items in total (16 per vowel pair) were created, divided equally into change and no-change trials. The items were presented in the following order: control pair (/i–ɑ/), followed by the three target pairs (/u–ʊ/, /ɛ–æ/, and /i–ɪ/). For all items, participants were asked the same question: "which one is distinct from the others?" Correct responses were options 1–3 for *change trials* and option 4 for *no-change trials*. The A′ score ranges from 0 to 1, where a value of 1 indicates perfect perceptual sensitivity. Scores of 0.50 or below indicate a lack of phonetic sensitivity, reflecting chance-level discrimination (Flege et al. 1999a, Tsukada et al. 2005).

3.4 Procedure

When recording was carried out, a single-blind design was adopted to reduce bias by preventing participants from consciously focusing on the articulation of target vowels. To this end, participants were recruited under the cover title, *Korean EFL Learners' Perception and Production of American English Words*. At the end of the experiments, they were fully debriefed regarding the purpose of the study, and all consented to the use of their data.

Native English speakers participated in two sessions six months apart. In the first, their speech was recorded to create listening stimuli for the perceptual sensitivity task; in the second, six of the eight participants returned to complete the perceptual sensitivity task. Korean participants attended a single session in which they recorded the same English sentences and completed the perceptual sensitivity task. Honoraria were provided after each session as an incentive for participation.

Sessions were conducted individually. While some recordings initially took place face-to-face in a soundproof laboratory, most were later conducted online via *Zoom* (Version 5.1.2, Zoom Video Communications, Inc. 2020) due to the COVID-19 pandemic. Face-to-face sessions were recorded in a soundproof lab using *Praat* or, when necessary, a portable recorder (Tascam DR-100MKII). Online participants met with the first author via *Zoom*, recorded speech using their own devices, and subsequently sent the audio files along with signed consent forms to the author by email. Although many participants used *Praat*, others relied on built-in applications on smartphones, tablets, or personal computers. Consequently, the collected file formats included *mp3* (1), *flac* (1), *m4a* (12), and *wav* (26). Since m4a files cannot be processed in *Praat* (Version 6.1.20 beta), all recordings in this format were batch-converted into *wav* files (PCM, 16-bit mono, 44,100 Hz) (Styler 2021) using *GoldWave* (Version 6.15, GoldWave Inc. 2015). Finally, perceptual sensitivity data were collected using Google Forms, administered in Korean or English according to each participant's native language.

3.5 Data Analysis

3.5.1 Acoustic analysis of vowels

A total of 1,140 tokens were analyzed, produced by 38 speakers (8 native English and 30 Korean participants), with each producing six vowels five times. Sound files were coded to indicate participant group (NE/NK), gender (F/M), ID number, and target word (e.g., NKF01had). Tokens were grouped by front/back vowels and gender to account for acoustic variation.

All vowel tokens were manually segmented in *Praat* by comparing spectrograms with the recordings. The first three formants (F1, F2, F3) were extracted at the vowel midpoint using an open-source script (Crosswhite, n.d.) in order to minimize coarticulatory effects. When atypical formant values were detected, tokens were cross-checked both spectrographically and auditorily and adjusted accordingly. To minimize lip rounding effects on F2, vowel backness was represented by the difference between F1 and F2 (i.e., F2−F1), the smaller the distance, the more back the vowel (Ladefoged and Johnson 2015). Formant measurements employed the Burg method (50 Hz pre-emphasis, 0.025 s window length), and the maximum number of formants and ceiling values were adjusted by gender and vowel type (Yang 2019): 4.0 with a ceiling of 5,500 Hz for female front vowels, 5.0 with a ceiling of 5,500 Hz for female back vowels, 4.5 with a ceiling of 5,000 Hz for male front vowels, and 5.0 with a ceiling of 5,000 Hz for male back vowels.

To control for gender-related differences, Bark Difference Metric normalization (Syrdal and Gopal 1986) was applied using the *NORM Vowel Normalization Suite 1.1* (Thomas and Kendall 2007). The normalized height dimension is shown as Z_3-Z_1 (Bark) while the normalized front-back dimension is displayed as Z_3-Z_2 (Bark).

3.5.2 Intra-speaker variability

Coefficient of variation (CV) (Kesteven 1946) was used to observe the relative intra-speaker variability of the six target vowel productions by 30 Korean participants. The CV is calculated based on the standard deviation (*sd*) and the mean formant frequency ($\bar{x}$) of five repetitions of each vowel of each speaker utilizing the formula, $sd/\bar{x}$ (Smith et al. 2019).

3.5.3 Perceptual sensitivity task

The A' score, calculated for each vowel pair, is based on the proportion of hits (correct response in change trials) and false alarms (incorrect responses in no-change trials) (Tsukada et al. 2005). When H is equal to FA, A′ was set at 0.5. If H is greater than FA, A′ was calculated as A′ = 0.5 + ((H−FA)*(1+H−FA)) / ((4*H)*(1−FA)), if H is less than FA, A′ = 0.5 − ((FA−H)*(1+FA−H)) / ((4*FA)*(1−H)) (Tsukada et al. 2005, pp. 272-273). A′ ranges from 0 (no sensitivity) to 1 (perfect sensitivity), with 0.5 representing chance performance (Flege et al. 1999a, Tsukada et al. 2005). Scores were rounded to two decimal places, and values above 0.50 were interpreted as evidence of perceptual sensitivity.

4. Results and Discussion

4.1 Intra-Speaker Variability: *New* vs. *Similar* Sounds

To address the first research question, paired-samples t-tests were conducted to examine whether intra-speaker variability differed between the *new* vowels (/ɪ, æ, ʊ/) and their corresponding *similar* vowels (/i, ɛ, u/) in terms of vowel height and vowel backness. Overall, the results indicated no statistically significant differences between the new and similar sounds within each vowel pair.

Table 2 presents the results of the paired-samples t-tests comparing intra-speaker variability in vowel height between the new and similar vowels, while Table 3 presents the results for vowel backness. For the /i–ɪ/ pair, no significant difference was observed in vowel height between the two sounds, $t(29) = -0.35$, $p = .73$, $d = -0.06$, 95% CI [−0.42, 0.30]. However, a significant difference was found in CV for vowel backness, $t(29) = -3.38$, $p = .002$, $d = -0.62$, 95% CI [−1.00, −0.22]. For the /ɛ–æ/ pair, neither vowel height nor vowel backness showed a significant difference between the *new* and *similar* sounds. For vowel height, $t(29) = -0.66$, $p = .52$, $d = -0.12$, 95% CI [−0.48, 0.24]. Similarly, vowel backness did not differ significantly between the two sounds, $t(29) = -0.59$, $p = .56$, $d = -0.11$, 95% CI [−0.47, 0.25]. A similar pattern was observed for the /u–ʊ/ pair. No statistically significant differences were found for either vowel height, $t(29) = 1.23$, $p = .23$, $d = 0.22$, 95% CI [−0.14, 0.58], or vowel backness, $t(29) = 1.21$, $p = .24$, $d = 0.22$, 95% CI [−0.14, 0.58].

Figure 2 illustrates the normalized Bark difference vowel plots with ellipses representing one standard deviation for each vowel category produced by native English (NE) and native Korean (NK) speakers, generated using the *NORM Vowel Normalization Suite 1.1* (Thomas and Kendall 2007). As shown in the figure, the dispersion of the *new* and *similar* vowels appears comparable within each vowel pair, visually supporting the statistical results indicating no reliable differences in intra-speaker variability between the two vowel types. Overall, the new and similar sounds did not differ significantly with respect to intra-speaker variability, with one exception: the /i–ɪ/ contrast in vowel backness.

Table 2. A Summary of a Paired Sample t-test about the Comparison of Intra-Speaker Variability (CV) in Vowel Height (Z_3-Z_1 (Bark)) between *New* English Vowels and *Similar* English Vowels for Korean Participants

					95% CI	
Vowel Pair	*t*	*df*	*p-value*	*Cohen's d*	Lower	Upper
/i–ɪ/	−0.35	29	0.73	−0.06	−0.42	0.30
/ɛ–æ/	−0.66	29	0.52	−0.12	−0.48	0.24
/u–ʊ/	1.23	29	0.23	0.22	−0.14	0.58

Table 3. A Summary of a Paired Sample t-test about the Comparison of Intra-Speaker Variability (CV) in Vowel Backness (Z_3-Z_2 (Bark)) between *New* English Vowels and *Similar* English Vowels for Korean Participants

					95% CI	
Vowel Pair	*t*	*df*	*p-value*	*Cohen's d*	Lower	Upper
/i–ɪ/	−3.38	29	0.002*	−0.62	−1.00	−0.22
/ɛ–æ/	−0.59	29	0.56	−0.11	−0.47	0.25
/u–ʊ/	1.21	29	0.24	0.22	−0.14	0.58

Note. * $p < .05$

Perceptual sensitivity to the vowel contrasts was also examined, given its close relationship with L2 speech production (Flege 1995). The average A' scores for the control vowel pair (/i–ɑ/) were near ceiling for both native English speakers (*M* = 0.98, *SD* = 0.05) and Korean participants (*M* = 0.92, *SD* = 0.15). This indicates that participants were able to reliably discriminate the control contrast, suggesting that the perception task functioned as intended. Native English speakers (*n* = 6) showed near-ceiling perceptual sensitivity across all three target vowel contrasts (*M* = 0.83–0.98). In contrast, Korean participants exhibited moderate perceptual sensitivity for the front vowel contrasts /i–ɪ/ (*M* = 0.61, *SD* = 0.15) and /ɛ–æ/ (*M* = 0.51, *SD* = 0.15), both exceeding the threshold value of 0.5. However, perceptual sensitivity for the back vowel contrast /u–ʊ/ fell below this threshold (*M* = 0.47, *SD* = 0.12), indicating reduced ability to discriminate this pair. This pattern is broadly consistent with the PAM prediction (Best 1995, Best et al. 2001), with the CG contrast /i–ɪ/ yielding higher perceptual sensitivity than the SC contrasts.

Interestingly, noticeable production variability was observed even for the vowel contrast that Korean participants perceived most accurately, namely /i–ɪ/. As shown in Figure 2, the distribution of /i/ tokens extends toward the lower-right region of the vowel space for some speakers. Two speakers in particular (NKF01 and NKM05) showed this pattern most clearly. To illustrate this pattern more closely, Figure 3 presents the vowel spaces of these two speakers, generated using *Visible Vowels* (Heeringa and Van de Velde 2018). In both cases, the /i/ category shows a broader distribution than /ɪ/, suggesting greater intra-speaker variability in the /i/ category. Notably, this pattern appeared in both speakers despite differences in age of onset (4 vs. 11 years) and reported English exposure (1 vs. 3 hours per week), suggesting that factors beyond the amount of English available in EFL contexts may contribute to the observed variability.

One possible explanation is developmental variability in L2 speech production. Notably, these two speakers (NKF01 and NKM05), who showed perceptual sensitivity to the /i–ɪ/ contrast (A' score > 0.5), exhibited greater variability in /i/, which may reflect an ongoing developmental process. Even when learners demonstrate perceptual sensitivity to /i–ɪ/, their production categories may still be undergoing stabilization, during which they experiment with different phonetic realizations before establishing a consistent articulatory target (Bley-Vroman 1983, Verspoor et al. 2008). Welch's *t*-tests revealed no statistically significant differences in intra-speaker variability between native English speakers and Korean participants across all vowels in both vowel height (Z_3–Z_1) and vowel backness (Z_3–Z_2). Thus, the observed difference for the /i–ɪ/ contrast in vowel backness may reflect subtle developmental patterns that are not captured by group-level comparisons between native and non-native speakers. This pattern of the /i–ɪ/ contrast may also be interpreted within the framework of the SLM (Flege 1995). According to the SLM, L2 sounds perceived as similar to existing L1 sounds (in this case, /i/) may be assimilated to the corresponding L1 category through equivalence classification. In such cases, learners *modify* an existing category rather than *create* a new one (Escudero 2009), which may limit the development of a distinct category for the *similar* L2 sounds. This, in turn, may lead to less stable production targets, resulting in increased phonetic variability, as observed in the present data.

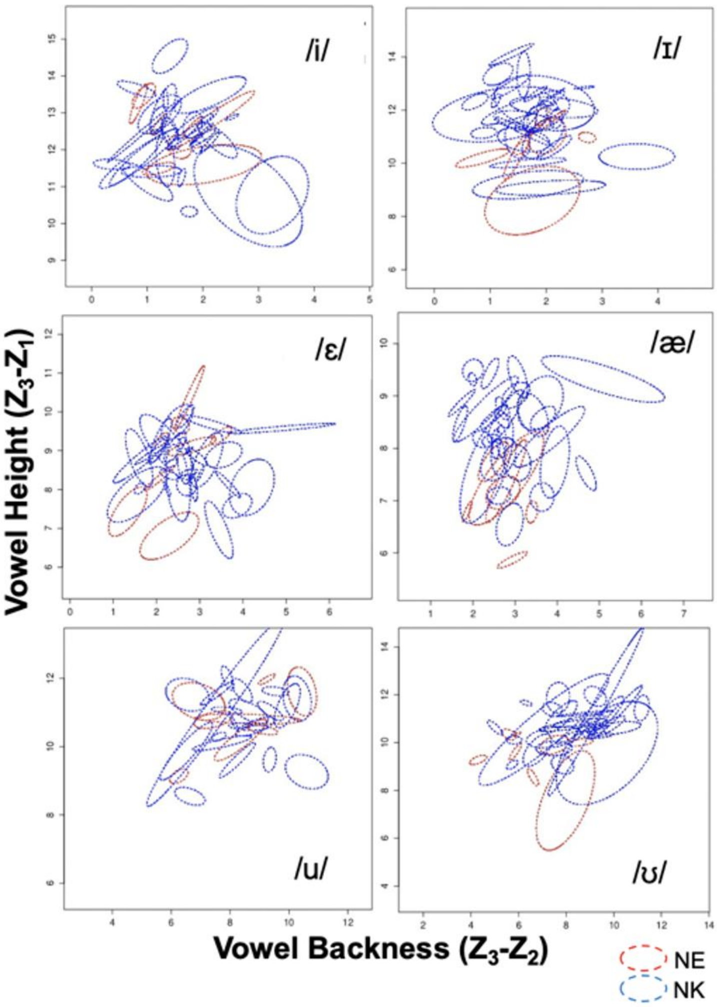

Figure 2. Bark Difference Normalized Vowel Plots with Ellipses by Native English speakers (NE) and Korean speakers (NK)

Note. Ellipses indicate one standard deviation.

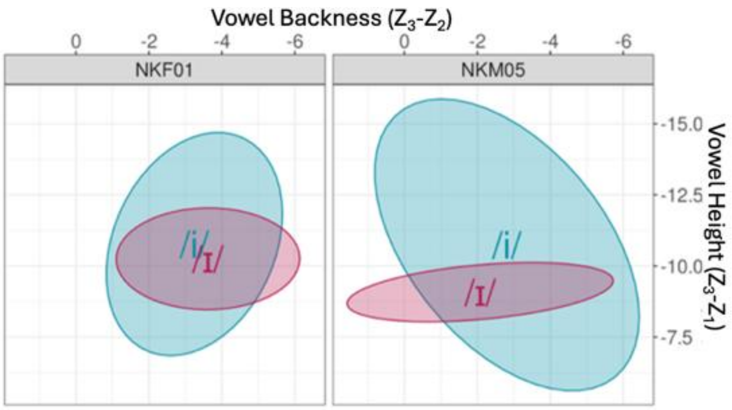

Figure 3. Formant Plots Illustrating Intra-Speaker Variability in the Production of /i/ and /ɪ/ by Two Korean Participants with Perceptual Sensitivity to the Contrast

4.2 Advanced Korean EFL Learners' Productions in the Four-Scenario Framework

Table 4 and Table 5 summarize the distribution of vowel productions across the four scenarios defined in the adapted framework from Smith et al. (2019), based on similarities in vowel formants and intra-speaker variability. Following Smith et al. (2019), productions were classified based on whether their acoustic values and coefficient of variation (CV) fell within ±2 standard deviations of the native speaker mean. In the present study, normalized vowel space dimensions were used, with height represented as Z_3–Z_1 (Bark) and front-back position as Z_3–Z_2 (Bark), instead of raw formant values. Productions were then categorized into four scenarios depending on whether acoustic values and variability fell within or outside this range.

Table 4 summarizes the similarities and differences in vowel height (Z_3–Z_1 Bark) and vowel backness (Z_3–Z_2 Bark), as well as intra-speaker variability (CV), for the six vowels produced by the Korean participants, based on the four-scenario framework proposed by Smith et al. (2019). According to the scenario classification, the similar vowels /i/, /ɛ/, and /u/ corresponded to Scenario 1 (similar formants and similar variability), while the new vowels /ɪ/, /æ/, and /ʊ/ corresponded to Scenario 3 (different formants but similar variability). This pattern is the reverse of what Smith et al. (2019) reported. In their data, the tense vowels /i/, /e/, and /u/—which would correspond to similar vowels under SLM terminology—patterned with Scenario 3 (different formants from native speakers), whereas the lax vowels /ɪ/, /ɛ/, and /ʊ/—corresponding to new vowels under SLM—patterned with Scenario 1 (native-like formants). Although Smith et al. (2019) did not interpret their findings in SLM terms, their results would be consistent with the SLM prediction (Flege 1995) that L2 learners more readily form distinct phonetic categories for new sounds, while similar sounds remain influenced by L1 categories. The present findings do not fit this prediction in a straightforward manner. The advanced Korean EFL learners in this study showed native-like formants for *similar* vowels but not for *new* vowels—the opposite of what SLM would predict. Several factors may contribute to this unexpected pattern. First, learning environment may matter: Smith et al.'s (2019) participants were ESL learners residing in an English-speaking country, whereas the present participants were EFL learners with limited L2 immersion. Reduced L2 input in EFL contexts may constrain the formation of new phonetic categories even at advanced proficiency. Second, methodological differences between the two studies—including Bark normalization versus raw formant comparisons, the hVd stimulus frame, and the specific contrasts examined (/ɛ–æ/ vs. /e–ɛ/)—may also contribute to the divergent pattern. These interpretations remain tentative and warrant further investigation directly comparing ESL and EFL learners under matched conditions.

For vowel backness, statistically significant differences were observed only for /u/ and /ʊ/, whereas no significant differences were found for /i/, /ɪ/, /ɛ/, and /æ/. As with vowel height, intra-speaker variability did not differ significantly for any vowel.[1] Consequently, /i/, /ɪ/, /ɛ/, and /æ/ were classified as Scenario 1, while /u/ and /ʊ/ were classified as Scenario 3. Overall, Scenario 1 occurred seven times (58%), whereas Scenario 3 occurred five times (42%). No instances of Scenario 2 or Scenario 4 were observed in the present data. Accordingly, the present analysis focuses primarily on vowel height, as vowel backness did not show meaningful differences across the vowel pairs, consistent with the analytical approach adopted by Smith et al. (2019).

Table 4. A Summary of Vowel Height (Z_3–Z_1 (Bark)) and Vowel Backness (Z_3–Z_2 (Bark)) Similarities and Differences for Six Vowels and Intra-Speaker Variability (CV) of Korean Participants and the Scenarios They Reflect

	Vowel Height: Statistical difference?			Vowel Backness: Statistical difference?		
Vowel	Formant	CV	Scenario	Formant	CV	Scenario
/i/	No	No	1	No	No	1
/ɪ/	Yes	No	3	No	No	1
/ɛ/	No	No	1	No	No	1
/æ/	Yes	No	3	No	No	1
/u/	No	No	1	Yes	No	3
/ʊ/	Yes	No	3	Yes	No	3

Table 5. Distribution of Korean Speakers Across Four Scenarios Based on Vowel Height (Z_3–Z_1 (Bark)) Deviation from Native English Speaker Baseline (±2 *SD*)

Vowel	Scenario 1 Similar formants Similar variability	Scenario 2 Similar formants Different variability	Scenario 3 Different formants Similar variability	Scenario 4 Different formants Different variability
/i/	24	2	2	2
/ɪ/	24	0	6	0
/ɛ/	29	1	0	0
/æ/	19	1	10	0
/u/	28	1	1	0
/ʊ/	22	1	7	0
Total	146 (81%)	6 (3%)	26 (14%)	2 (1%)

As can be seen in Table 5, the majority of productions fell into Scenario 1, characterized by similar formants and similar variability, accounting for 146 cases (81%) of the total observations. In contrast, Scenario 2 (similar formants but different variability) was relatively rare, with only 6 cases (3%). Scenario 3, representing different formants but similar variability, accounted for 26 cases (14%). Scenario 4, which reflects both different formants and different variability, was observed in only 2 cases (1%). A chi-square test indicated that the distribution of productions across the four scenarios was statistically significant, $\chi^2(15, N = 180) = 32.70$, $p = .005$, with a small-

[1] Note that these between-group formant comparisons differ from the within-group CV comparisons reported in Table 3: while the /i–ɪ/ pair showed greater within-NK CV asymmetry in backness (Table 3), NK and NE speakers did not differ significantly in CV for any vowel (Welch's *t*-tests).

to-moderate effect size (*Cramer's V* = 0.25). This result indicates that vowel productions were distributed unevenly across the four scenarios, with a clear dominance of Scenario 1. When considering intra-speaker variability alone, most tokens fell into Scenario 1 (81%) and Scenario 3 (14%), both of which represent patterns with native-like variability. Taken together, this indicates that approximately 95% of productions exhibited variability patterns similar to those of native speakers. In contrast, Smith et al. (2019) reported that 21% of their participants showed non-native-like intra-speaker variability. This difference may suggest differences in sampling criteria, as Smith et al. (2019) did not control for learner proficiency, whereas the present study restricted the sample to advanced learners.

Although most productions of advanced Korean learners fell into Scenario 1, speakers within this scenario differed in their perceptual sensitivity to the vowel contrasts. Because the number of participants in this case was very small, statistical comparisons of formant differences across individual vowel contrasts were not conducted. Instead, the vowel space distributions are presented descriptively to illustrate general patterns of overlap and separation across speakers. Figure 4 presents representative vowel space distributions for Scenario 1 speakers with and without perceptual sensitivity to the three target English vowel contrasts, generated using *Visible Vowels* (Heeringa and Van de Velde 2018). The figure shows that perceptual sensitivity to each vowel contrast is closely related to the degree of separation in vowel production. NKF02 (AOA: age 5, LOR 6 years) and NKF14 (AOA: age 6, LOR: 4.5 years) were the only Korean participants who demonstrated perceptual sensitivity to all three vowel contrasts, with mean A' scores approaching the perfect score. Their vowel distributions showed relatively little overlap and were broadly consistent with the patterns observed for native English speakers. On the other hand, the other Korean participants who showed lack of perceptual sensitivity to all three vowel pairs displayed considerable overlap across contrasts. This pattern is consistent with the close relationship between perception and production proposed in the SLM (Flege 1995), which posits that accurate production of L2 sounds depends on learners' ability to perceive them accurately. In acoustic phonetics, vowel categories are typically interpreted relative to a speaker's overall vowel space rather than in terms of fixed absolute formant values (Ladefoged and Johnson 2015, Yang 1996). In this context, successful L2 production may not necessarily require exact replication of native-speaker acoustic targets, but rather the maintenance of sufficient relative contrasts within a speaker's vowel space, consistent with Liljencrants and Lindblom (1972), who argue that vowel systems are structured to maximize perceptual contrast within the available acoustic space. Accordingly, an important question in L2 vowel production may be whether learners produce sufficiently distinct L2 vowel categories within their own vowel system. Moreover, evidence from speech perception research suggests that listeners can rapidly adapt to systematic deviation in foreign-accented speech (e.g., Xie et al. 2018). This perceptual flexibility indicates that L2 speech intelligibility may not depend strictly on native-like acoustic accuracy. Instead, maintaining distinct contrasts between vowel categories may be more critical, as it helps prevent the neutralization of L2 phonological distinctions that are crucial for distinguishing meaning. In this respect, frameworks that evaluate L2 learners' productions primarily in comparison with native-speaker acoustic targets, such as the four-scenario framework proposed by Smith et al. (2019), may fail to capture learners' ability to preserve functional vowel contrasts within their own vowel space.

The vowel space distributions shown in Figure 4 also provide evidence that the establishment of distinct L2 phonological categories depends not only on age of acquisition but also on the extent of immersion in an L2 environment. Speakers who arrived in English-speaking environments at an early age and had extended exposure to English ((b) NKF02 and NKF14) exhibited clearly separated vowel categories for all three contrasts (/i–ɪ/, /ɛ–æ/, and /u–ʊ/), similar to those observed in (a) native speakers. In contrast, learners with limited perceptual sensitivity exhibited substantial overlap between vowel categories. This group included speakers with short-term

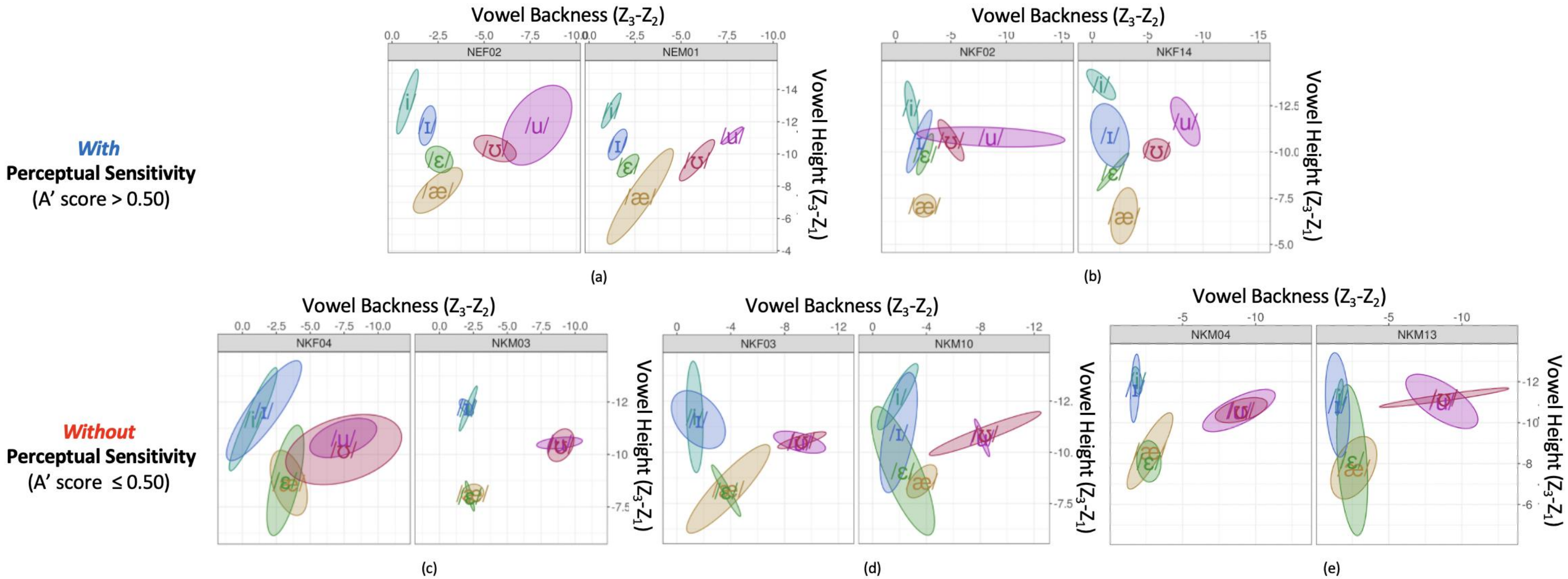

Figure 4. Vowel Space Distributions of English Vowels (/i–ɪ/, /ɛ–æ/, /u–ʊ/) for Representative Speakers in Scenario 1, With and Without Perceptual Sensitivity to All Three Vowel Contrasts

Note. Ellipses indicate one standard deviation. All Korean speakers in panels (b–e) were classified as Scenario 1 (Similar Formants, Similar Variability) in the four-scenario framework (Smith et al., 2019). Despite differences in perceptual sensitivity, they all showed production patterns consistent with Scenario 1.

(a) Native English speakers (one male, one female)

(b) Korean speakers with perceptual sensitivity, early AOA, and extended residence in English-speaking environments (US: AOA = 5, LOR = 6 years; India: AOA = 6, LOR = 4.5 years).

(c–e) Korean speakers without perceptual sensitivity, differing in language experience: (c) short-term residence abroad (US/Australia; LOR = 10 months), (d) early AOA but no experience abroad, and (e) the highest self-reported English listening and speaking study time (12 hours/week) residence in English speaking countries ((c) NKF04 and NKM03, each with a length of residence of 10 months) as well as those whose experience with English was primarily classroom-based ((d) NKF03 and NKM10, early AOA without residence in English speaking countries, (e) NKM04 and NKM13, reporting the highest weekly English study hours—12 hours—for listening and speaking). Despite considerable formal study, these learners ((c) – (e)) showed noticeable overlap between vowel categories in each contrast pair. This pattern suggests that these learners had not fully established distinct phonological representations for the English vowel contrasts examined in the present study. Importantly, the results indicate that an early age of acquisition alone is insufficient for successful phonological category formation. Several speakers who began learning English relatively early but lacked sustained immersion in an English-speaking environment ((d) NKF03 and NKM10) showed vowel distributions characterized by significant category overlap. This may indicate that the quantity of input provided in typical EFL instructional settings is not sufficient to yield perceptual learning outcomes comparable to those observed in more naturalistic environments. Extensive exposure to the statistical distribution of L2 sounds facilitates the development of new phonetic categories, whereas limited exposure can lead learners to assimilate L2 sounds into existing L1 categories. This raises the question of how EFL learners might compensate for limited naturalistic exposure. One pedagogical approach that has been proposed in the literature is high variability phonetic training (HVPT), which provides learners with structured exposure to a wide range of acoustic variability (Choe et al. 2025, Thomson 2018). Whether such training could help EFL learners establish more distinct categories for new English vowels—particularly given the patterns observed in the present study—is a question that warrants future empirical investigation.

These findings are broadly consistent with theoretical accounts of second language speech learning that emphasize the interaction between age and experience. According to the SLM (Flege 1995), the formation of new L2 phonetic categories depends on learners' ability to detect phonetic differences between L1 and L2 sounds and on the amount of experience with L2 input. Similarly, research on L2 perception suggests that experience with L2 input can influence how learners perceive non-native speech contrasts (Best and Tyler 2007, Gorba and Cebrian 2021). The present results further support the view that age-related advantages in L2 phonological acquisition are mediated by immersion and continued interaction with the L2 environment rather than by age alone. In this respect, the findings align with previous work showing that age effects in L2 acquisition interact with experiential variables such as language use and exposure (Flege et al. 1999b). Taken together, the results suggest that early immersion facilitates the development of perceptual sensitivity to L2 phonological contrasts, whereas limited exposure may hinder the establishment of new phonetic categories even when learning begins relatively early. The present findings therefore underscore the importance of sustained L2 input for phonological development and support models of second language acquisition that emphasize the interaction between age of acquisition and linguistic experience.

Nevertheless, caution is warranted when interpreting such patterns solely on the bases of acoustic measures. Kang and Pickering (2013) cautioned against overreliance on acoustic analyses because, unlike human judgments, acoustic measures rely on precise calibration that may not align with listeners' perceptual evaluations. Similarly, Munro and Derwing (2015) emphasized that "without listener data, acoustic measures themselves are of little or no value" (p. 14). These observations suggest that future studies should combine acoustic analyses with listener-based evaluations, treating acoustic measures as complementary to human perceptual judgments. In this respect, the four-scenario framework proposed by Smith et al. (2019) may benefit from incorporating listener perception more explicitly. Future research could further examine how the four-scenario framework relates to listener-based measures such as intelligibility and comprehensibility (e.g., Munro and Derwing 1995). Such investigations would help clarify whether production patterns classified under different scenarios correspond to differences in how speech is actually perceived by listeners.

5. Conclusion

This study investigated intra-speaker variability in the production of English vowels by advanced Korean EFL learners, comparing new (/ɪ, æ, ʊ/) and similar (/i, ɛ, u/) vowels within the four-scenario framework (Smith et al. 2019). Overall, learners showed comparable variability across both vowel types, indicating relatively stable production at advanced proficiency. However, two speakers who showed perceptual sensitivity to the /i–ɪ/ contrast exhibited greater variability in the vowel /i/, which may suggest that their phonological category for this sound may still be under development. Moreover, most productions of Korean speakers were classified as Scenario 1 (81%) or Scenario 3 (14%), indicating largely native-like variability patterns. This distribution contrasts with Smith et al. (2019), likely due to differences in proficiency control. Notably, even within native-like production patterns in the four-scenario framework proposed by Smith et al. (2019), individual differences in perceptual sensitivity were observed, with greater sensitivity associated with more distinct vowel category separation in production.

The present study is not without limitations. First, our original research design aimed to include learners across multiple proficiency levels in order to capture developmental differences more comprehensively. However, the final sample was limited to advanced learners due to difficulties in recruiting intermediate- and lower-proficiency participants, suggesting the need for alternative sampling strategies beyond online and offline advertisements to ensure broader proficiency representation. Future research should include a wider range of proficiency levels to provide a more comprehensive account of L2 vowel perception and production. Second, the lack of randomization in stimulus presentation may have introduced order effects. Future studies should randomize stimulus sequences to minimize this potential bias. Third, the use of a context-reduced *hVd* format may not fully capture natural speech processing, as vowel perception is influenced by phonetic and lexical context. Employing more naturalistic sentence-level stimuli would offer greater ecological validity. Lastly, another limitation of the present study is that participants' English proficiency was controlled using TOEIC LC scores only. While this approach provided a practical and standardized means of ensuring a relatively homogeneous proficiency level, future studies may benefit from incorporating multiple measures of language proficiency, including both receptive and productive abilities, to allow for a more comprehensive characterization of participants. Despite these limitations, this study extends previous research by applying the four-scenario framework to advanced Korean EFL learners and highlighting the role of perceptual sensitivity in shaping L2 vowel production patterns.

Acknowledgements

An earlier version of this work was presented in the first author's master's thesis (Choe 2022) submitted to Seoul National University.

References

Aoyama, K., J. E. Flege, S. G. Guion, R. Akahane-Yamada and T. Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32(2), 233-250.

Audacity Team. 2008. Audacity (Version 1.3.4-beta) [Computer program]. Available online at http://audacityteam.org/

Bent, T., R. Bradlow and B. L. Smith. 2007. Segmental errors in different word positions and their effects on intelligibility of non-native speech. In O.-S. Bohn and M. J. Munro, eds., *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege,* 331-347. Amsterdam: John Benjamins.

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

Best, C. T. and M. D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro and O.-S. Bohn, eds., *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 13-34. Amsterdam: John Benjamins.

Best, C. T., G. W. McRoberts and E. Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America* 109(2), 775-794.

Bley-Vroman, R. 1983. The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning* 33, 1-17.

Boersma, P. and D. Weenink. 2020. Praat: Doing phonetics by computer (Version 6.1.20 beta) [computer software]. Available online at http://www.praat.org/

Choe, S. 2022. *Korean EFL Learners' Intra-Speaker Variability and Perceptual Sensitivity of Three American English Vowel Pairs: [i–ɪ], [ɛ–æ], [u–ʊ].* Master's thesis, Seoul National University, Seoul, Korea.

Choe, S., H. Park and H. Ahn. 2025. The efficacy of high variability phonetic training for L2 speech perception in EFL contexts: A meta-analytic approach. *Korean Journal of English Language and Linguistics* 25, 1416-1443.

Crosswhite, K. n.d. *Formant-logger*. Praat Script Resources, UCLA Phonetics Lab. Available online at http://phonetics.linguistics.ucla.edu/facilities/acoustic/formant_logging.txt

Escudero, P. 2009. Linguistic perception of "similar" L2 sounds. *Phonology in Perception* 15, 152-190.

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

Flege, J. E. 2003. Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller, eds., *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities* 6, 319-355. Berlin: Mouton de Gruyter.

Flege, J. E., I. R. MacKay and D. Meador. 1999a. Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America* 106(5), 2973-2987.

Flege, J. E., G. H. Yeni-Komshian and S. Liu. 1999b. Age constraints on second-language acquisition. *Journal of Memory and Language* 41(1), 78-104.

Fogerty, D. and D. Kewley-Port. 2009. Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *The Journal of the Acoustical Society of America* 126(2), 847-857.

Giegerich, H. J. 1992. *English phonology: An introduction*. Cambridge, UK: Cambridge University Press.

GoldWave Inc. 2015. Goldwave (Version 6.15) [computer software]. Available online at www.goldwave.com

Gorba, C. and J. Cebrian. 2021. The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. *Journal of Phonetics* 88, 101094.

Han, J. I., T. H. Choi, I. Lim and J. K. Lee. 2011. The interlanguage speech intelligibility benefit for Korean learners of English: Production of English front vowels. *Phonetics and Speech Sciences* 3(2), 53-61.

Heeringa, W. and H. Van de Velde. 2018. Visible vowels: A tool for the visualization of vowel variation. In *Proceedings CLARIN Annual Conference 2018*, 4034-4035. Available online at https://www.visiblevowels.org/

Hong, S. 2007. The characteristics of vowel identification errors of university-level Korean students of American English: HCA. *Language and Linguistics* 39, 257-277.

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. *Studies in Phonetics, Phonology and Morphology* 18(3), 491-511.

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

Kang, O. and L. Pickering. 2013. Acoustic and temporal analysis for assessing speaking. In A. J. Kunnan, ed., *The Companion to Language Assessment*, 1047-1062. Malden, MA: Wiley-Blackwell.

Kesteven, G. L. 1946. The coefficient of variation. *Nature* 158(4015), 520-521.

Kim, G. N. and S. J. Kim. 2003. Korean native speakers' perception of English sounds according to the groupings of phonetic contrasts. *Speech Sciences* 10(1), 59-67.

Ladefoged, P. and K. Johnson. 2015. *A Course in Phonetics* (7th ed.). Boston, MA: Cengage Learning.

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

Lee, K. Y. and M. H. Cho. 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. and M. H. Cho. 2018. Predicting L2 vowel identification accuracy from cross-language mappings between L2 English and L1 Korean. *Language Sciences* 66, 183-198.

Levis, J. M. 2018. *Intelligibility, Oral Communication, and the Teaching of Pronunciation*. Cambridge, UK: Cambridge University Press.

Liljencrants, J. and B. Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48(4), 839-862.

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.

Munro, M. J. and T. M. Derwing. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45(1), 73-97.

Munro, M. J. and T. M. Derwing. 2015. A prospectus for pronunciation research in the 21st century. *Journal of Second Language Pronunciation* 1(1), 11-42.

Saito, K., P. Trofimovich and T. Isaacs. 2016. Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics* 37(2), 217-240.

Shin, J. 2015. Vowels and consonants. In L. Brown and J. Yeon, eds., *The Handbook of Korean Linguistics*, 1-21. New Jersey, NJ: Wiley-Blackwell.

Silva, D. J. 2004. Phonological mappings as dynamic: The evolving contrastive relationship between English and Korean. *Linguistic Research* 21, 57-73.

Smith, B. L., E. Johnson and R. Hayes-Harb. 2019. ESL learners' intra-speaker variability in producing American

English tense and lax vowels. *Journal of Second Language Pronunciation* 5(1), 139-164.

Snodgrass, J. G., G. Levy-Berger and M. Haydon. 1985. *Human Experimental Psychology.* Oxford: Oxford University Press.

Strange, W., A. Weber, E. S. Levy, V. Shafiro, M. Hisagi and K. Nishi. 2007. Acoustic variability within and across German, French, and American English vowels: Phonetic context effects. *The Journal of the Acoustical Society of America* 122(2), 1111-1129.

Styler, W. 2021. *Using Praat for Linguistic Research*. University of Colorado at Boulder, Phonetics Lab. Available online at https://wstyler.ucsd.edu/praat/

Suh, H. 2020. A study on Korean students' perception of English short vowels. *Studies in Linguistics* 55, 265-286.

Syrdal, A. K. and H. S. Gopal. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America* 79(4), 1086-1100.

Thomas, E. R. and T. Kendall. 2007. *NORM: The Vowel Normalization and Plotting Suite.* Available online at https://lingtools.uoregon.edu/norm/norm1.php

Thomson, R. I. 2018. High variability [pronunciation] training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation* 4(2), 208-231.

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

Verspoor, M., W. Lowie and M. Van Dijk. 2008. Variability in second language development from a dynamic systems perspective. *The Modern Language Journal* 92(2), 214-231.

Xie, X. and T. F. Jaeger. 2020. Comparing non-native and native speech: Are L2 productions more variable? *The Journal of the Acoustical Society of America* 147(5), 3322-3347.

Xie, X., K. Weatherholtz, L. Bainton, E. Rowe, Z. Burchill, L. Liu and T. F. Jaeger. 2018. Rapid adaptation to foreign-accented speech and its transfer to an unfamiliar talker. *The Journal of the Acoustical Society of America* 143(4), 2013-2031.

Yang, B. 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. 2019. A comparison of normalized formant trajectories of English vowels produced by American men and women. *Phonetics and Speech Sciences* 11(1), 1-8.

Yun, G. 2014. Korean listeners' perception of L2 English phoneme contrast. *Studies in Phonetics, Phonology and Morphology* 20(2), 161-185.

Zoom Video Communications, Inc. 2020. *Zoom* (Version 5.1.2) [Computer software]. Available online at https://zoom.us

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

