



The Role of Perceptual and Acoustic Similarity in Learners' Perception of L2 Vowels*

Shinsook Lee (Korea University) · Hyunah Baek (Ajou University)



This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: June 30, 2025

Revised: July 23, 2025

Accepted: August 5, 2025

Lee, Shinsook (First and corresponding author)
Professor, Department of English Language Education, Korea University
145 Anam-ro, Seongbuk-gu
Seoul, Korea
Email: leesseng@korea.ac.kr

Baek, Hyunah (Co-author)
Assistant professor, Department of English Language and Literature
Ajou University
206 World cup-ro, Yeongtong-gu
Suwon-si, Gyeonggi-do, Korea
Email: hyunahbaek@ajou.ac.kr

* This research was supported by the College of Education, Korea University Grant in 2025. An earlier version of the paper was presented in the 11th International Symposium on the Acquisition of Second Language Speech, New Sounds 2025 (April 22-25), Toronto, Canada.

ABSTRACT

Lee, Shinsook and Hyunah Baek. 2025. The role of perceptual and acoustic similarity in learners' perception of L2 vowels. *Korean Journal of English Language and Linguistics* 25, 1084-1101.

This study investigates the role of L1-Korean-L2-English vowel similarity and acoustic distance in L1-Korean speakers' perception of L2-English vowels, given that few studies have examined the joint effects on learners' L2 vowel perception of L1-L2 sound similarity/dissimilarity and the acoustic distance of L2 vowel production from native speaker reference. Thirty-three Korean learners of English produced and identified 12 English vowels (e.g., *bet*, *but*, *boat*) and mapped the same English vowels onto close Korean vowels. Acoustic distances of the learners' produced vowels from native English speakers' norms were calculated as Euclidean distance in the F1/F2 space. The learners' vowel identification accuracy was assessed and L1-L2 perceptual vowel similarity/dissimilarity was calculated based on the learners' L2-to-L1 mapping patterns. Results of a linear mixed-effects regression model showed a significant main effect of vowel type (i.e., similar, somewhat similar, dissimilar), indicating that the perception accuracy was significantly higher for L1-similar L2 vowels compared to L1-somewhat similar and L1-dissimilar L2 vowels. The main effect of acoustic distance was also significant, with greater distance associated with lower perception accuracy for L1-similar L2 vowels. However, accuracy rates were varied within each vowel type as the identification accuracies of some L1-somewhat similar or L1-dissimilar L2 vowels tended to be high when they bear salient perceptual cues like vowel duration or r-coloring. These results are only partly in line with the propositions of the Speech Learning Model (SLM/SLM-revised, Flege 1995, Flege and Bohn 2021) concerning the relative difficulty and easiness of the acquisition of L1-similar L2 vowels versus L1-dissimilar L2 vowels, indicating that different L2 vowels develop at different rates.

KEYWORDS

L2 vowel perception, Korean EFL learners, L1-L2 perceptual similarity, acoustic distance, vowel variations

1. Introduction

1.1 Theories of L2 Phonological Acquisition

The concept of cross-language or L1-L2 phonetic similarity plays a crucial role in determining the success or failure of the perception and/or production of L2 sounds. Models of L2 speech learning such as the Speech Learning Model (SLM, Flege 1995)/the SLM-revised (SLM-r, Flege and Bohn 2021) and the Second Language Linguistic Perception (L2LP) model (Escudero 2005, Escudero and Yazawa 2023) postulate that L2 learners are less likely to perceive and/or produce L2 sounds accurately as a function of the perceived or acoustic phonetic similarity between an L2 sound and its nearest L1 counterpart.

In particular, the SLM (Flege 1995)/SLM-r (Flege and Bohn 2021) postulates that the greater the perceived phonetic similarity between an L2 sound and its closest L1 analogue, the harder it is for L2 learners to learn the L2 sound. This further suggests that L2 learners are less likely to form a new L2 phonetic category in such a case. The SLM-r (p. 33) argues for assessing cross-language similarity/dissimilarity perceptually rather than acoustically as L2 learners' perceptual judgment of L2-to-L1 sound similarity/dissimilarity sometimes deviates from acoustic measures (Levy and Strange 2008, Strange 2007). The SLM-r measures the perceived L2-to-L1 phonetic similarity/dissimilarity through a cross-language categorization (or L1 labeling) in a forced-choice task. After labeling an L2 token in terms of L1 categories, L2 learners assess it for degrees of perceived similarity/dissimilarity from the L1 sound category employed to categorize the L2 token. That is, the perceived L2-to-L1 phonetic distance was measured with L1 categorization and rating data and this indicates which L1 sound is the closest counterpart of a target L2 sound in phonetic space (p. 34). For instance, if English /ɪ/ was predominantly mapped onto Korean /i/ (95%) with a high goodness rating (5.0 out of 7) and English /ɔ/ onto two Korean vowels /ʌ/ (59%) and /o/ (31%) with rather low goodness ratings (3.9 and 3.4, respectively) (Lee and Cho 2020, p. 9), the SLM/SLM-r predicts that English /ɪ/, which is similar to Korean /i/, would cause more perceptual and production difficulties than English /ɔ/, which is at least somewhat distinct from any existing Korean vowels.

The L2LP model (Escudero 2005, Escudero and Yazawa 2023) also claims that acoustic-phonetic similarity between L2 learners' native language and target language can predict the level of L2 learners' difficulty in perceiving and/or producing L2 sound categories. However, the model focuses on L2 phonemic contrasts rather than on individual L1 and L2 sounds unlike the SLM/SLM-r. Specifically, the L2LP model considers three learning scenarios exploring L2 phonemic contrasts. First, the SIMILAR scenario predicts less difficulty for L2 learners as two contrastive sounds in L2 are mapped onto two different L1 sound categories; L1-German listeners' perception of the L2 English /i/-/ɪ/ contrast represents this scenario as German also has the /i/-/ɪ/ contrast. Second, the NEW scenario predicts the greatest learning difficulty; it is illustrated by L1-Korean listeners' difficulty in perceiving the L2 English /i/-/ɪ/ contrast because the contrast does not exist in Korean; thus, L1-Korean listeners are more likely to perceive the two English vowels as the single Korean vowel /i/. In this scenario, the challenge arises because an existing L1 category must be split up or a new L2 category must be formed. Third, the SUBSET scenario occurs when two contrastive sounds in L2 are perceived as multiple L1 categories either without overlap (SUBSET EASY) or with overlap (SUBSET DIFFICULT). An illustration of SUBSET DIFFICULT comes from Elvin et al.'s (2021) examination of monolingual Australian English listeners' categorization and discrimination of Brazilian Portuguese vowels. The acoustic similarity of the listeners' L1 vowel productions and Brazilian Portuguese vowels was compared based on linear discriminant analyses. The results for acoustic similarity showed that, for instance, the Brazilian Portuguese /i/-/e/ contrast falls into SUBSET DIFFICULT for the Australian English listeners, as the two vowels were acoustically similar to as well as perceived as the same Australian English

vowels /i:/, /ɪ/, and /ɪə/, and the listeners struggled to notice the vowel contrast.¹

1.2 The Effect of L1 and L2 Sound Similarity/Dissimilarity in L2 Vowel Perception

The existing literature for the influences of the perceived and/or acoustic phonetic similarity of L1 and L2 on L2 sound acquisition showed some mixed results. To illustrate, Flege (1998) looked into accentedness ratings of early and late L1-Italian L2-English bilinguals' production of L1-similar (/i, ε, α/) and L1-dissimilar/new (/ɪ, ʊ, ə/) English vowels. The accentedness ratings of the L1-similar English vowels produced by native English speakers were significantly higher than those of the same English vowels produced by early bilinguals, whereas it was not true for the new English vowels. Jun and Cowie (1994) examined F1 and F2 values of American English vowels /i, ɪ, u, ʊ, ʌ/ produced by Korean learners of English as a second language (ESL) whose length of residences (LORs) in the U.S. were different; more experienced learners with LOR of 26–31 years vs. less experienced learners with LOR of 1.2–5.3 years. The Korean learners' productions of English vowels /i/ and /u/, which have Korean counterparts /i/ and /u/, respectively, were overall less accurate, while the more experienced learners' productions of English vowels /ɪ/ and /ʊ/, which do not have any matching Korean vowels, were more accurate than the less experienced learners' productions, although individual variations were observed among the learners. Similar results were found by Doan and Oh (2023) who investigated Vietnamese listeners' identification of British English vowels. The authors reported that experienced Vietnamese listeners were more accurate in identifying English vowels than inexperienced listeners. Moreover, the experienced listeners' identification accuracies of new English vowels /ɪ/ and /ʊ/, which do not have acoustically similar Vietnamese counterparts, were high (90%), while their accuracies of similar English vowels /æ/ and /ʌ/ were low (20%) or not high (60%), respectively.

In contrast, Munro (1993) reported that L1-Arabic male speakers' production of English /i/, which is spectrally close to the Arabic vowel /i:/, was the highest-rated vowel (i.e., least-accented), whereas their production of English /ɛ/, which is a new vowel category, was the lowest-rated vowel (i.e., most-accented). Llompart and Reinisch (2019) examined the perception and production of the English /i/-/ɪ/ and /ɛ/-/æ/ contrasts by German learners of English. The English /i/-/ɪ/ contrast has a close matching German contrast, whereas the English /ɛ/-/æ/ contrast does not. The German learners completed a vowel categorization test and a word reading test. The results showed that the German learners performed better on the English /i/-/ɪ/ contrast than on the English /ɛ/-/æ/ contrast, which suggests that the perceived phonetic similarity of L2 to L1 did not predict the German learners' English vowel performance with accuracy. Cebrian et al. (2021), who investigated the perception and production of the British English vowels /i, ɪ, ε, ə/ by Spanish/Catalan learners of English, also reported that the learners overall performed quite well on English /ɛ/ across diverse tasks (i.e., vowel discrimination, vowel identification, picture naming, and English-to-Spanish vowel mapping), although English /ɛ/ has a Spanish analog /e/.

1.3 The Relationship Between Perceptual and Acoustic Similarity in L2 Vowel Perception

Alispahic et al. (2017) examined whether acoustic closeness/differences between L1 and L2 vowels could predict non-native vowel perception. Australian English listeners and Peruvian Spanish listeners completed XAB

¹ The Perceptual Assimilation Model for L2 (PAM-L2, Best and Tyler 2007) predicts L2 discrimination patterns based on articulatory-phonetic similarity as well as phonological similarity between learners' L1 and L2. The PAM-L2 and the L2LP model have some commonalities. For instance, the SIMILAR scenario in the L2LP model is analogous to two-category assimilation in the PAM-L2 and the NEW scenario in the L2LP model is similar to single-category assimilation in the PAM-L2. However, the comparison of the two models is not a main interest of the current study.

discrimination and categorization tasks with six Dutch vowel contrasts (/i/, /ɪ/, /a/, /ɑ/, /y/, /ɤ/). The results showed that both listener groups' discrimination was more accurate when the Dutch vowel contrasts were estimated to be perceptually easy with respect to acoustic properties between L1 and L2, while their discrimination was less accurate when the contrasts were estimated to be difficult. For example, Australian English listeners categorized the Dutch /ɪ/-/i/ contrast in terms of multiple L1 categories and their discrimination of the contrast was low because of a higher acoustic overlap between the Dutch and Australian English vowel categories. However, the listeners performed much better on the Dutch /i-y/ contrast because they relied on F2 for backness and F3 for rounding differences to discriminate between the contrast, even though these vowels were also categorized in terms of several L1 categories.

On the other hand, Strange et al. (2004) investigated the extent to which perceptual similarity patterns between North German (NG) and American English (AE) vowels were predicted by context-specific (i.e., citation-form vs. sentence) acoustic (spectral and temporal) similarity between NG vowels and AE vowels. The target vowels were produced in isolated *hVp* (NG) and *hVba* (AE) syllables (study 1) and in *hVC* syllables inserted in a short sentence (study 2) and discriminant analyses were performed to examine cross-language acoustic similarity. Native American English listeners with no experience in German assimilated NG vowels to AE vowels and provided goodness ratings in both citation-form and sentence contexts. The results showed that spectral similarity between NG and AE vowels did not always make accurate predictions of perceptual similarity patterns. Specifically, the results from study 1 (citation-form) showed that "new" NG high and mid front rounded vowels (/y, ø, œ/ except /y:/) were acoustically intermediate between AE front unrounded vowels and back rounded vowels, but the four NG vowels were perceptually more similar to AE back rounded vowels. In sentence contexts (study 2), NG high front rounded vowels showed a more consistent assimilation pattern to AE back rounded vowels, while NG mid and mid-low vowels (except NG /e:/) were judged less similar to their AE counterparts although the NG vowels and their AE counterparts were acoustically more similar in that context. Strange (2007) examined acoustic similarity (F1 and F2) of American English (AE), German, and French vowels in diverse contexts. The results of the similarity in citation-form showed that German and French front rounded vowels were acoustically more similar to AE front unrounded vowels than to back rounded vowels. For instance, French front rounded vowels /y/ and /ø/ were acoustically similar to AE front unrounded vowels /i, e/ and /ɪ/, respectively. However, native AE listeners overwhelmingly judged the same French vowels to be similar to AE back rounded vowels in an L2-to-L1 perceptual assimilation task. This indicates that there is a disparity between acoustic similarity measurements and perceptual mapping comparisons of L2-to-L1 similarity patterns. Based on the results, Strange contended that perceptual assimilation patterns, rather than acoustic similarity measurements, better accounted for the perceptual relationships between the sounds in L1 and L2.

Balas et al. (2023) examined factors deciding perceptual and acoustic similarity between L1- Polish and L2- Norwegian vowels. Fifteen Polish learners of Norwegian completed an L2-to-L1 perceptual assimilation task and provided goodness-of-fit ratings and Euclidean distances between Norwegian and Polish vowels were calculated. The results showed that a model with the Euclidean distance computed in terms of F1 and F2 performs better than a model with the Euclidean distance calculated in terms of F1, F2, and F3, but there was no statistically significant relationship between L2-to-L1 assimilation counts and Euclidean distance measures.

Therefore, the studies reviewed so far indicate that the effect of the perceived and/or acoustic phonetic similarity of L1 and L2 on L2 vowel acquisition yielded some mixed results and there were also disparities between cross-language perceptual assimilation patterns and cross-language acoustic similarities.

1.4 The Present Study

The findings of previous studies about the effects on L2 vowel perception of the perceived phonetic similarity of L1 and L2 vowels and acoustic properties are not without controversy. Moreover, most previous studies have focused on the impact of L1-L2 perceptual similarity or cross-language (i.e., L1-L2) acoustic similarity on L2 vowel perception. Hence, the present study explored the joint effects on L2 vowel perception of L1-L2 perceptual similarity and acoustic distance of L2 vowel production from native speaker reference rather than cross-language acoustic similarity. Specifically, the present study purports to answer the following research questions: First, do L1-L2 perceptual similarity and acoustic distances between L2 vowel productions and L2 native speakers' productions jointly affect L2 learners' L2 vowel perception? If L1-similar L2 vowels are more difficult to be perceived or produced than L1-dissimilar (or new) L2 vowels, as predicted by the SLM/SLM-r, we would expect that, for L1-similar L2 vowels, L2 learners' identification accuracy of L2 vowels will not be high and the acoustic distance between L2 learners' L2 vowel productions and native speaker reference will be great. However, if L1-similar L2 vowels are easier to be perceived or produced than L1-dissimilar L2 vowels, as reported by Llompart and Reinisch (2019) and Cebrian et al. (2021), we would expect a reverse pattern; for L1-similar L2 vowels, L2 learners' L2 vowel identification accuracy will be high and their L2 vowel productions will be close to native speakers' acoustic norms. Or if L1-L2 vowel similarity and acoustic distance measures affect L2 learners' L2 vowel perception differently, we would expect that there will not be any meaningful interaction between the two factors. For instance, for L1-similar L2 vowels, L2 learners' identification accuracy could be high, although their acoustic distance measures are deviated from native speaker reference, or L2 learners' identification accuracy could be low, but their acoustic distance measures are close to native speakers' acoustic norms. The same scenario could be made for L1-dissimilar L2 vowels. Secondly, the present study investigates whether all the L1-similar, L1-somewhat similar or L1-dissimilar L2 vowels exhibit a consistent pattern with respect to vowel identification accuracy and acoustic distance measures. Additionally, the study examines whether L2 learners' L2 proficiency affects their identification accuracy of L2 vowels, as previous studies have documented that L2 proficiency played a role in L2 vowel perception and/or production (Cho and Lee 2023, Flege 1999).

2. Experiment²

2.1 Participants

Thirty-three Korean learners of English as a foreign language (EFL) (9 males and 24 females) were recruited from a private university in Seoul and their mean age was 22.5 years. They were all English-language education majors except some who were English-language education and education double-majors. The learners' mean IBT-TOEFL scores were 102.2 (range: 85~114), indicating that they could be considered advanced or upper-intermediate level learners of English.³ Their initial age of English learning was 8.2 years and none of them had lived in English-speaking countries for more than 6 months.

² The participants' mapping results were presented in Lee and Shin (2015), perception results in Lee et al. (2019), and part of their production data in Lee et al. (2017) and native speakers' acoustic measurements in Lee et al. (2022), in which the presented results and data were analyzed from different perspectives along with different sets of data.

³ Some of the learners reported TOEIC, TEPS, or IELTS scores. In such cases, these test scores were converted to IBT-TOEFL scores for direct comparison.

2.2 Stimuli

Twelve English vowels were used as stimuli couched in a monosyllabic /b/___/t/ word (/i/ *beat*, /ɪ/ *bit*, /eɪ/ *bait*, /ɛ/ *bet*, /æ/ *bat*, /ʌ/ *but*, /ɜ:/ *Burt*, /ɑ:/ *Bart*, /u/ *boot*, /oo/ *boat*, /ɔ/ *bought*, /ɑ/ *bot*).⁴ The whole experiment comprised three tasks: an English vowel production task followed by an English vowel identification task and then an L2 English-to-L1 Korean vowel mapping task. The same stimulus set was employed in all the tasks. Four native speakers (two males and two females) of North American English (NAE) recorded the stimuli used in the English vowel identification test. Two of them were from Ohio and Illinois and the other two were from California in the U.S. (mean age: 25.7 years). The native speakers recorded the target words several times in a carrier framework of “Say _____ again.” Three native speakers’ productions were recorded with a RODE-NT1A microphone using the Speech Filing System in a soundproof room at University College London (UCL) in London, while one native speaker’s productions were recorded with a Sony ECM-MS907 microphone using the Speech Filing System in a sound-attenuated room at a private university in Seoul. The sampling rate of all recordings was 44.1 kHz. Each target word (e.g., *boot*) in the native speakers’ productions was segmented at word boundaries by a phonetically trained researcher. The researcher examined both the waveform and spectrogram. A pronunciation instructor at UCL and two NAE speakers with specialized knowledge of phonetics at UCL inspected the stimuli and only the best stimuli selected by these experts were used in the experiments. Two native speakers who recorded the stimulus set did not make a distinction between the two vowels /ɑ/ (*bot*) and /ɔ/ (*bought*). Hence, if the Korean learners identified these native speakers’ productions of *bot* and *bought* as the opposite, the learners’ responses were considered correct. The four native speakers’ vowel productions were acoustically measured in terms of F1, F2, F3, and vowel duration, and the results are provided in Appendix 1. Formant frequencies of the target vowels were measured at the midpoint of each vowel and these acoustic measures were averaged across target vowels.

2.3 Procedure

First, for the production task, the Korean EFL learners produced the stimuli in a carrier framework of “Say _____ again.” They went over each stimulus word before the test in order to make certain that they were familiar with the stimuli. The stimuli were shown on a computer monitor and the learners produced each stimulus word three times in a sound-attenuated room at a private university in Seoul. The learners recorded the stimuli with a Sony ECM-MS907 microphone using the Speech Filing System and the sampling rate of all recordings was 44.1 kHz. Each target word (e.g., *boot*) in the Korean EFL learners’ productions was segmented at word boundaries by checking both the waveform and spectrogram. Formant frequencies (F1 and F2) of the target vowels were measured at the midpoint of each vowel and these acoustic measures were averaged across target vowels (Appendix 2).

The vowel perception (i.e., identification) task was conducted using Praat (Boersma and Weenink 2014). The four native English speakers’ productions of 12 English words (described in section 2.2) were used as stimuli for the identification task. The Korean learners wore headphones and listened to each stimulus word and then identified the word they heard among the given choices (e.g., *bat*, *boat*, *Burt*) presented on the computer monitor. The learners could play each stimulus word up to three times and the stimuli were randomly presented. Before the vowel perception task, the learners did practice with 12 trials but did not receive any feedback. There were 96

⁴ R-colored vowels (i.e., /ɜ:/, /ɑ:/) are considered monophthongs as r-coloring is regarded as part of the preceding vowel (Edwards 2003, Ladefoged 2006).

stimuli in total in the vowel perception task (12 vowels × 4 speakers × 2 repetitions).

The L2-to-L1 vowel mapping task was also conducted using Praat. One female native English speaker's vowel productions were used as stimuli in this task and there were 24 tokens (12 English vowels × 1 speaker × 2 repetitions). The learners were asked to label L2 English vowels with the most similar L1 Korean vowel. They also assessed goodness-of-fit between L2 English vowels and the chosen L1 Korean vowels on a 7-point Likert scale where "7" indicated "very similar" and "1" "very different." The Korean vowel alternatives provided with Korean orthography were displayed on the computer screen: <이> /i/, <에> /e/, <애> /ɛ/, <으> /u/, <어> /ʌ/, <우> /u/, <오> /o/, <아> /a/, <야> /jʌ/, <얘> /jɛ/, <요> /jo/, <왜> /wɛ/ <유> /ju/, <와> /wa/.⁵ The learners could play each stimulus word up to three times and they did practice with 12 trials before the main task.⁶

2.4 Analyses: L2-to-L1 Vowel Mapping

In order to determine the degree of similarity/dissimilarity between the vowels in English and Korean, we used the results from the L2-to-L1 vowel mapping task, given in full in Appendix 3. According to Tyler (2021, pp. 167-8), categorization consistency needs to be high (95%) when classifying L1-similar vs. L1-dissimilar L2 sounds if "categorization tasks provide a reliable indication of listeners' perceptual assimilation." He further contends that goodness ratings can be interpreted in a straightforward way when the categorization consistency is high. Also, based on Park and de Jong (2008, 2017), Lee and Cho (2020, pp. 8, 18) considered one-to-one/two/three category mapping patterns along with high/somewhat high/somewhat low/low mean goodness ratings to classify whether target L2 vowels were L1-similar or new L2 vowels. Hence, the current study adopted the criteria advocated by these previous studies (Lee and Cho 2020, Park and de Jong 2008, 2017, Tyler 2021). First, L2-English vowels were considered similar to L1-Korean vowels if there was one-to-one mapping and goodness-of-fit ratings were high (i.e., greater than 5 on the 7-point Likert scale).⁷ Accordingly, the English vowels /i/ (*beat*), /ɪ/ (*bit*), /ʌ/ (*but*), and /u/ (*boot*) were categorized as similar to their Korean counterparts. Next, L2-English vowels were regarded as somewhat similar to L1-Korean vowels if there was one-to-one mapping but goodness-of-fit ratings fell between 4 and 5 or if there were one-to-multiple mappings but goodness-of-fit ratings were somewhat high (higher than average rating of 4.5). Hence, five English vowels, /ɜ/ (*Burt*), /ɛ/ (*bet*), /æ/ (*bat*), /ɑ/ (*Bart*), and /ɑ/ (*bot*) were categorized as somewhat similar to their matching Korean vowels.⁸ Further, L2-English vowels were classified as dissimilar (or new) to L1-Korean vowels if there were one-to-multiple mappings or goodness-of-fit ratings were low (lower than 4). Thus, the English vowels /eɪ/ (*bait*), /oʊ/ (*boat*), and /ɔ/ (*bought*) were categorized as dissimilar to Korean vowels. Here note that L1-similar L2 vowels such as /i/ (*beat*), /ʌ/ (*but*), and /u/ (*boot*) cannot be regarded as L1-identical vowels given that these L2 vowels were different from their Korean counterparts <이> /i/, <어> /ʌ/,

⁵ The Korean vowels were composed of eight monophthongs and six onglide plus vowel sequences. The Korean vowels were selected on the basis of a pilot study to include the Korean vowel responses that are most likely to match the whole set of given English vowels. The Korean vowels /e/ and /ɛ/ were also included as some previous studies found that native Korean speakers differentiate between Korean /e/ and /ɛ/ in both perception and production (Lee 2023, Lee and Cho 2020, Lee et al. 2018).

⁶ In the current study, English vowel production, identification, and L2 English-to-L1 Korean vowel mapping tasks were run in that order. As an anonymous reviewer suggested, future study should use a counterbalanced design in order to avoid training effects.

⁷ Considering that there were 12 Korean alternatives, only cases of mapping probability higher than chance level (i.e., $100 \div 12 = 8.3$) were taken into account.

⁸ Notice that English vowels /ɛ/ (*bet*), /æ/ (*bat*) were both mapped onto Korean <애> (/ɛ/) and <에> (/e/) with similar mean percentages and goodness ratings (Appendix 3), but the <에> (/e/) mapping pattern was more prominent in both cases. This suggests that the Korean learners still differentiated between Korean <애> (/ɛ/) and <에> (/e/) to some extent.

and <우> /u/, respectively, with respect to F1 and F2 values (Hillenbrand et al. 1995, Lee et al. 2023, Yang 1996). Further, we distinguished between L1-similar L2 vowels and L1-somewhat similar L2 vowels as previous studies suggested that similarity should be regarded as a gradient property rather than a categorial property (i.e., “new” vs. “similar”) (Park and de Jong 2008, Strange et al. 1998).

3. Results

3.1 Overall Results

In order to analyze the effects of L1-L2 perceptual similarity and acoustic distance on Korean EFL learners' perception of English vowels, a linear mixed-effects regression model was built using the `lmer()` function from the `lme4` package (Bates et al. 2015) for R (R Core Team 2019, 2023). The dependent variable was mean perception accuracies of each participant and fixed effects were vowel type (Similar/Somewhat similar/Dissimilar) and mean acoustic distances of each participant. Importantly, the acoustic distance measured in raw formant values does not take into account individual differences in vowel space and intrinsic characteristics of each vowel. Therefore, we conducted additional analyses using normalized acoustic distance values. To obtain normalized acoustic distance, we first z-score normalized F1 and F2 values of each vowel token by both gender and vowel, following DiCanio et al. (2015) and Lee et al. (2017). Then, we calculated the Euclidean distance of each token produced by the Korean learners from the normalized mean F1 and F2 values of the native speakers of the same gender. Vowel type was treatment-coded with Similar as the reference level. The random effect structure included by-participant random intercept. The *p*-values were obtained using the `lmerTest` package (Kuznetsova et al. 2017). Table 1 summarizes the result of the model.

Table 1. Effects of Vowel Type and Normalized Acoustic Distance on Perception Accuracy

Model formula: Perception accuracy ~ Vowel type * Normalized acoustic distance + (1 subject)						
	<i>Estimate</i>	<i>SE</i>	<i>df</i>	<i>t</i>	<i>Pr(> t)</i>	
(Intercept)	0.9421	0.0588	374.5085	16.0300	0.0000	***
VowelType (Similar vs. Somewhat similar)	-0.2514	0.0732	378.3708	-3.4370	0.0007	***
VowelType (Similar vs. Dissimilar)	-0.2420	0.0833	377.6684	-2.9060	0.0039	**
Normalized acoustic distance	-0.0597	0.0279	388.4619	-2.1390	0.0330	*
VowelType (Similar vs. Somewhat similar) * Normalized acoustic distance	0.0710	0.0373	380.6895	1.9040	0.0577	
VowelType (Similar vs. Dissimilar) * Normalized acoustic distance	0.0789	0.0409	380.6265	1.9270	0.0547	

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

The results of the model showed a significant main effect of vowel type, indicating that the perception accuracy was significantly higher for L1-similar L2 vowels compared to L1-somewhat similar and L1-dissimilar L2 vowels. The main effect of normalized acoustic distance was also significant, with greater distance associated with lower perception accuracy for L1-similar L2 vowels (Figure 1). The interaction between vowel type and normalized acoustic distance was not significant, although there was a numerical trend that, unlike L1-similar L2 vowels, greater normalized acoustic distance was associated with slightly higher perception accuracy for L1-somewhat similar and L1-dissimilar L2 vowels (Figure 1).

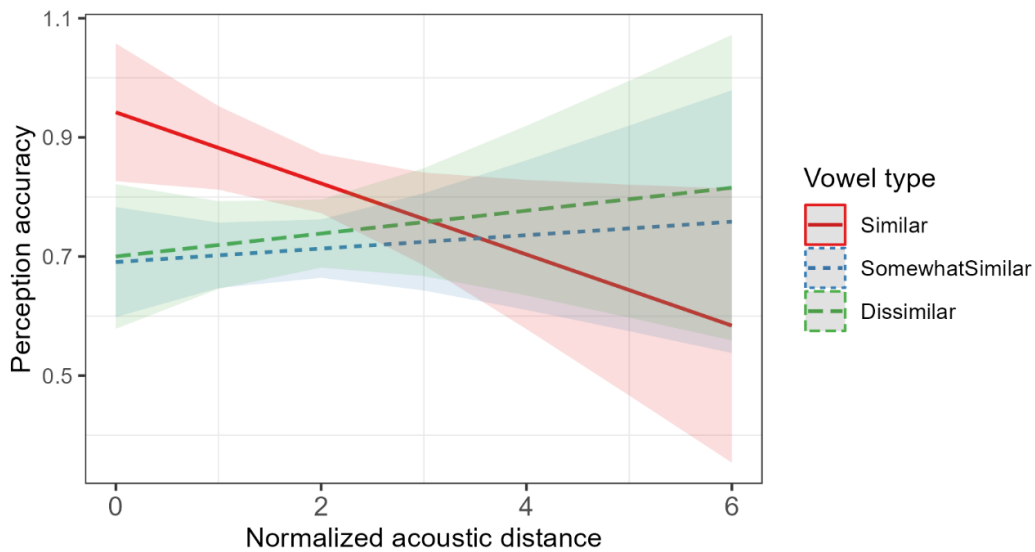


Figure 1. Perception Accuracy by Vowel Type and Normalized Acoustic Distance

Additionally, the present study examined whether the Korean EFL learners' perception accuracy rates were affected by their standard English proficiency test scores. This was because several previous studies have reported that L2 proficiency influenced L2 learners' perception and/or production of L2 vowels (Cho and Lee 2023, Flege 1999) and because the Korean learners' IBT-TOEFL scores ranged from 85 to 114, although they were considered advanced or upper-intermediate level learners of English based on their mean IBT-TOEFL scores. As can be seen in Figure 2, there was no correlation between the learners' perception accuracy rates and their English proficiency test scores ($r=0.040, p=0.839$), indicating that the learners whose English proficiency test scores were high did not necessarily have high accuracy rates. For instance, the top left-hand corner of the scatter plot in Figure 2 exhibits that some learners' accuracy rates were high (above 80%), although their English proficiency test scores were below 100.

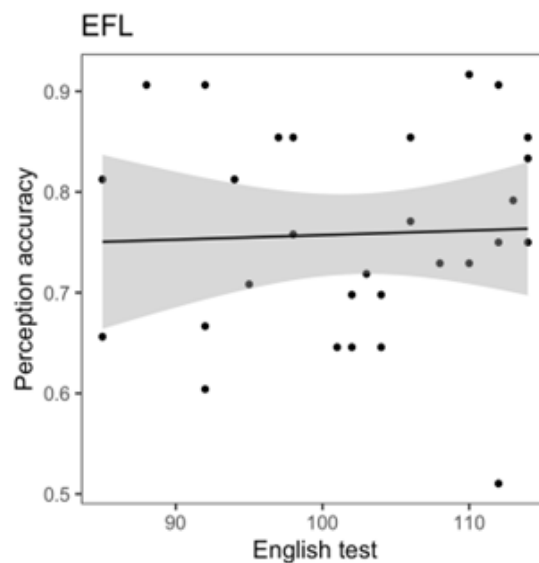


Figure 2. Relationship Between Perception Accuracy and Standard English Proficiency Test Scores

3.2 Variations in Accuracy for Each Vowel Type

The present study further investigated whether all the L1-similar, L1-somewhat similar, and L1-dissimilar L2 English vowels exhibited consistent patterns in accuracy rates for each vowel type. The mean perception accuracy of each vowel is shown in Table 2.

Table 2. Mean Percentage Accuracy of Each Vowel for Each Vowel Type

L1-similar vowels			L1-somewhat similar vowels				L1-dissimilar vowels				
beat	bit	but	boot	bet	bat	Burt	Bart	bot	bought	bait	boat
88 (33)	67 (47)	79 (4)	97 (16)	56 (22)	75 (25)	92 (18)	83 (27)	48 (26)	53 (50)	89 (31)	78 (41)

Note: Standard Deviations are in parentheses.

The results indicated that variations existed within each vowel type. Specifically, among the L1-similar L2 vowels, the learners' perception accuracy of /u/ (*boot*) and /i/ (*beat*) were high while their perception accuracy of /ɪ/ (*bit*) was not high. As for the L1-somewhat similar L2 vowels, the perception accuracy of /ɜ:/ (*Burt*) was high and that for /ɑ:/ (*Bart*) was somewhat high. In contrast, the perception accuracy of /ɛ/ (*bet*) was somewhat low and that for /ɑ/ (*bot*) was the lowest. Among the L1-dissimilar L2 vowels, the perception accuracy of /eɪ/ (*bait*) was high, while that for /ɔ:/ (*bought*) was somewhat low. Moreover, the R function `avov()` indicated that the variations in accuracy rates within each vowel type were statistically significant, as shown in Table 3.

Table 3. Statistical Results of Vowels in Accuracy Rates Within Each Vowel Type

Similar vowels					
	<i>Df</i>	<i>Sum sq</i>	<i>Mean Sq</i>	<i>F</i>	<i>Pr(>F)</i>
Vowel	3	1.627	0.5422	18.77	0.000 ***
Residuals	128	3.697	0.0289		
Somewhat similar vowels					
	<i>Df</i>	<i>Sum sq</i>	<i>Mean Sq</i>	<i>F</i>	<i>Pr(>F)</i>
Vowel	4	4.444	1.111	19.680	0.000 ***
Residuals	160	9.030	0.056		
Dissimilar vowels					
	<i>Df</i>	<i>Sum sq</i>	<i>Mean Sq</i>	<i>F</i>	<i>Pr(>F)</i>
Vowel	2	2.284	1.142	15.51	0.000 ***
Residuals	96	7.068	0.0736		

Note: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

First, for the L1-similar L2 vowels, the accuracy rate of /ɪ/ was much lower than that for other vowels. Statistical analyses (the TukeyHSD post-hoc test) showed that the differences in accuracy rates between the following vowel pairs were significant (all p adjusted $< .05$): *bit* and *beat*, *bit* and *boot*, *bit* and *but*, and *but* and *boot*. Figure 3 plots normalized mean F1 and F2 values of the L1-similar L2 vowels produced by male and female EFL learners with native speaker reference, which shows that greater acoustic distance was associated with lower perception accuracy. For instance, the vowels /i/ and /u/ produced by the male EFL learners were very close to the vowels produced by the male native English speakers while the female EFL learners' productions were not that close to the female native speakers' productions. However, the EFL learners' accuracy rates of these vowels were high

overall. In contrast, the vowel /ɪ/ produced by the EFL learners deviated much from the native reference point and the learners' accuracy of /ɪ/ was lower than the accuracies of other L1-similar L2 vowels.

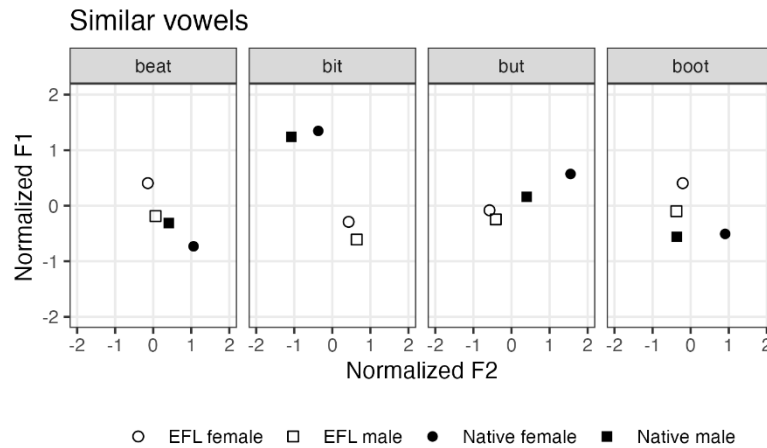


Figure 3. Normalized Mean F1 and F2 Values of the L1-similar L2 Vowels Produced by the EFL Learners and Native English Speakers

Secondly, for the L1-somewhat similar L2 vowels, the differences in accuracy rates between the following vowel pairs were all significant (all p adjusted $<.05$), again showing variations: *bat-bet*, *Burt-bet*, *Bart-bet*, *Burt-bat*, *bot-bat*, *bot-Burt*, *bot-Bart*. Figure 4 plots normalized mean F1 and F2 values of the L1-somewhat similar L2 vowels produced by male and female EFL learners with native speaker reference. As can be seen in Figure 4, the EFL learners' productions of English /ɑ/ and /ɛ/ were close to the native speaker reference (especially the vowel /ɑ/), but the learners' accuracy of these vowels were (somewhat) low. By contrast, the learners' production of English /æ/ diverged from the native speaker reference but its accuracy was higher than the accuracies of /ɑ/ and /ɛ/. As for the r-colored vowels /ɜ/ and /ɑɪ/, the EFL learners' productions of these vowels were not much diverged from/or close to the native speaker reference and the learners' accuracy rates for these vowels were high.

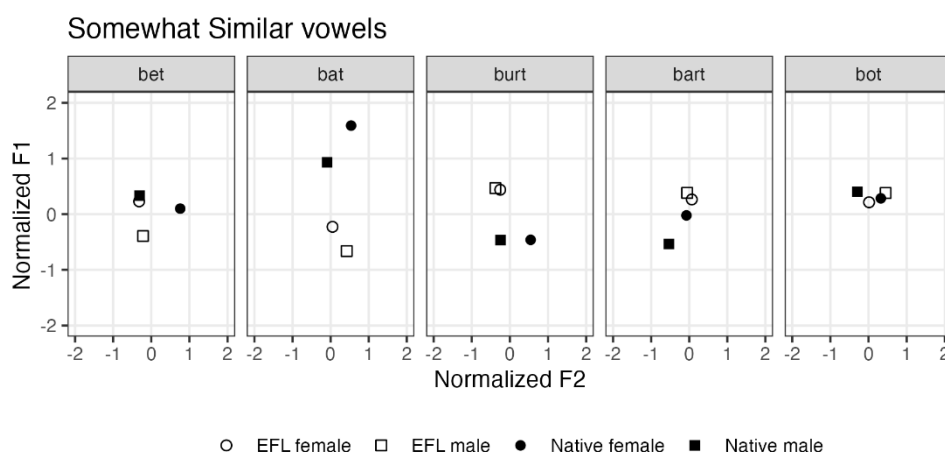


Figure 4. Normalized Mean F1 and F2 Values of the L1-somewhat Similar L2 Vowels Produced by the EFL Learners and Native English Speakers

For the L1-dissimilar L2 vowels, the differences in accuracy rates between the vowels in *bought* and *boat* and *bought* and *bait* were all statistically significant (all p adjusted $<.05$), due to the low accuracy of the vowel /ɔ/. Figure 5 plots normalized mean F1 and F2 values of the L1-dissimilar L2 vowels produced by male and female EFL learners with native speaker reference. As shown in Figure 5, the male EFL learners and the male native English speakers produced the vowel /eɪ/ very closely and the female EFL learners' production was also rather close to the female native speaker reference, and the learners' accuracy of the vowel was high overall. However, the learners' production of /ɔ/ was somewhat deviated from the native speaker reference and the learners' accuracy for the vowel was low. As for the vowel /oʊ/, the learners' production diverged from the native speaker norms to some degree, but the learners' accuracy of the vowel was 78%.

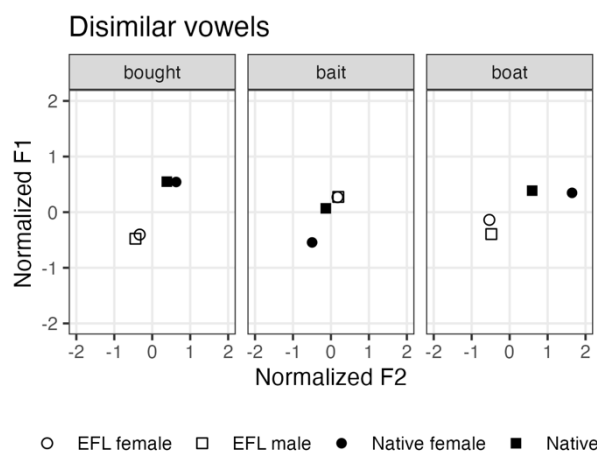


Figure 5. Normalized Mean F1 and F2 Values of the L1-dissimilar L2 Vowels Produced by the EFL Learners and Native English Speakers

Thus, the results indicated that not all L2 vowels were acquired in the same way, showing variations for each vowel type. The results also indicated that L2 learners' L1-L2 perceptual similarity was not necessarily associated with the acoustic distances of L2 learners' L2 vowel productions from native speaker reference, especially for the L1-somewhat similar and L1-dissimilar L2 vowels.

4. General Discussion

The current study investigated whether Korean EFL learners' perception of L2 English vowels was affected by L1-L2 sound similarity/dissimilarity and the acoustic distance of L2 vowel production from native speaker reference. First of all, the results showed that there was a significant main effect of acoustic distance. This was because, for L1-similar L2 English vowels, as the EFL learners' acoustic distance was close to the native speaker reference, their perception accuracy tended to be high. For example, the Korean learners' production of English /u/ (*boot*) was close to the native speaker norms and their perception accuracy was very high, whereas the learners' production of English /ɪ/ was deviated from the native speaker reference and their perception accuracy was not high (refer to Table 2, Figure 3). This seems to indicate that the learners may have established target-close L2 vowel categories at least for some L1-similar L2 vowels (e.g., /u/ and /i/) and their productions of these L2 vowels

were attuned to their established target-close L2 vowel categories, showing a close alignment between perception and production.⁹ As for the L1-somewhat similar L2 vowels, the EFL learners' acoustic distance did not necessarily coincide with their perception accuracy. For instance, the EFL learners' production of English /ɑ/ was very close to the native speaker reference, but their perception accuracy was the lowest, whereas the learners' production of English /æ/ showed a reverse pattern. The EFL learners' production of r-colored vowels /ɜ/ and /ɑɪ/ were not much diverged from/or somewhat close to the native speaker norms and the learners' accuracy rates for these vowels were high. The L1-dissimilar L2 vowels sided with the L1-somewhat similar L2 vowels. For example, the EFL learners' productions of English /ɔ/ and /oo/ were diverged from the native speaker reference to some extent, but the learners' accuracy rates of the vowels were different (53% vs. 78%, respectively). The results from the L1-somewhat similar and L1-dissimilar L2 vowels indicate that acoustic distances of L2 vowel production from native speaker reference did not always make accurate predictions of the EFL learners' perception accuracy.

Secondly, there was a significant effect of vowel type. That is, the perception accuracy of L1-similar L2 English vowels was significantly higher than that for L1-somewhat similar or L1-dissimilar L2 vowels. Nonetheless, it is worth noting that there were significant variations in perception accuracy within each vowel type (refer to Table 2, Figures 3, 4, 5). Among the L1-similar L2 vowels, the perception accuracy of /ɪ/ (*bit*) was not high while that for /i/ (*beat*) was high, indicating that the high front lax vowel /ɪ/ causes more perceptual difficulty for the Korean learners than its tense counterpart. This was partly attributed to the Korean learners' establishment of a target-approximated phonetic category for English /i/. As for the L1-somewhat similar L2 vowels, the vowel /ɑ/ (*bot*) showed the lowest accuracy rate. The lowest accuracy rate of /ɑ/ may partly be ascribable to the merger between the vowels /ɑ/ and /ɔ/ in many parts of the U.S. (e.g., the West, the Midwest and the South), in which speakers produce the vowels in words like *bot* and *bought* and *pond* and *pawned* identically (Boberg 2015). And this sound change was also reflected in the native English speakers' productions in the current study (Appendix 1). The vowels /ɛ/ (*bet*) and /æ/ (*bat*) exhibited bidirectional confusions, but the accuracy rate of /æ/ was much higher than that of /ɛ/ (75% vs. 56%). This may be because the vowel /æ/ tends to be lengthened, providing an additional cue for the perception of the vowel (Ladefoged 2006). The r-colored vowels /ɜ/ (*Burt*) and /ɑɪ/ (*Bart*) showed a high and somewhat high accuracy rate, respectively, which suggests that the r-coloring property of the vowels serves as a prominent perceptual cue (Edwards 2003, Ladefoged 2006). For the L1-dissimilar L2 vowels, the vowel /ɔ/ (*bought*) showed a low accuracy rate, partly because of its merger with /ɑ/. However, the accuracy of /eɪ/ (*bait*) was very high as it contains salient durational cues (Edwards 2003, Ladefoged 2006), while that of /oo/ (*boat*) was not that high despite its durational cues. This may be partly because F1 and F2 converge for the vowel /oo/'s latter half, weakening the acoustic cue of the latter half, whereas F1 and F2 diverge greatly for the vowel /eɪ/'s latter half, strengthening the acoustic cue of the latter half.

Therefore, the findings of the current study are partly in line with the propositions made by the SLM/SLM-r that, as the perceptual distance between the vowels in L1 and L2 becomes larger, L2 learners are more likely to notice the differences between these vowels, which leads to a more successful perception and production of the L2 vowels. That is, the Korean EFL learners in the current study were able to identify L1-somewhat similar L2 vowels or L1-dissimilar L2 vowels more accurately when the vowels bear salient perceptual properties such as r-coloring and vowel duration. Given that these perceptual cues are not available in identifying L1 vowels, these cues were more likely to have assisted the Korean learners in their perception of the target L2 vowels, supporting the proposition made by the SLM/SLM-r. However, the learners were able to perceive L1-similar L2 vowels more

⁹ Note that the high back vowel /u/ tends to be fronted in NAE similar to British English, especially in the pronunciation of young people (Clopper et al. 2005, Evans and Iverson 2007).

successfully when their productions of the vowels approximated native speaker reference more closely, suggesting a potential category formation of the vowels. The result is not in line with the SLM/SLM-r's postulate that L1-similar L2 vowels cause more perceptual and/or production difficulty than L1-dissimilar L2 vowels; rather the results corroborate the findings of Llompart and Reinisch (2019) and Cebrian et al. (2021) concerning the relative easiness of the acquisition of L1-similar L2 vowels.¹⁰ Moreover, the results suggest that different L2 vowels develop at different rates given the variations in perception accuracy for each vowel type.

In addition, the fact that the Korean learners perceived L1-similar L2 vowels more accurately as their productions of the L2 vowels more closely approximated native speaker reference suggests a close alignment between speech perception and production. Casserly and Pisoni (2010, p. 643) reviewed research from different fields that, they contended, indicates a direct link between the two modalities. Their arguments are based on a wide range of experimental evidence from audiovisual speech perception, from phonetic convergence, and from neurobiological observations of "mirror neurons," that is, neurons that fire when we do an action (e.g., production) and when we perceive other people doing the same action, such that "perceiving the speech cause[s] activation of the motor plans that would be used in producing the same speech." Specifically, Casserly and Pisoni suggested that these mirror neurons could provide a basic mechanism for the direct link between the two modalities.

In a similar vein, Yang (2014) examined whether there is a link between Korean EFL learners' perceived and produced English vowel spaces. The Korean learners perceived synthetic English vowel stimuli varied in terms of pitch and formant values and produced the same English vowels (e.g., *heed, head, hood*, etc.). The results overall showed that the learners had difficulty with high front and high back tense and lax vowel distinctions (/i-ɪ/, /u-ʊ/) and also with low front and low back vowel distinctions (/ɛ-æ/, /ɑ-ɔ/) in both perception and production, indicating a potential link between the two modalities.

Finally, there was no correlation between the Korean learners' perception accuracy and their English proficiency test scores, indicating that the learners' English proficiency test scores did not impact their perception accuracy. This suggests that other learner factors such as their auditory acuity and processing, aptitude for L2 learning, and motivation should be considered in the future study in order to investigate whether these learner properties have an impact on L2 learners' perception of L2 vowels (Moyer 1999, Saito and Tierney 2025, Trofimovich et al. 2015). Further, the results from Strange et al. (2004) and Strange (2007) showed that both acoustic and perceptual similarity between the vowels in L1 and L2 varied in response to the prosodic context (i.e., citation-form vs. sentence). Hence, future study should examine cross-language perceptual similarity between L1 and L2 vowels and acoustic similarity between L2 learners' production of L2 vowels and that of native L2 speakers in diverse prosodic contexts. Additionally, the present study did not examine the impact of other acoustic properties such as F3 and duration on L2 vowel perception. Accordingly, future study should explore these acoustic properties, in addition to F1 and F2, to further shed light on our understanding of L2 vowel perception.

5. Conclusion

The study investigated the joint effects on Korean EFL learners' L2 vowel perception of L1-L2 sound similarity/dissimilarity and the acoustic distance of L2 vowel production from native speaker reference. Thirty-three Korean EFL learners performed three tasks: production and identification of English vowels and English-to-

¹⁰ Note that the so-called L1-similar L2 vowels were not acoustically identical to the L2 vowels produced by native English speakers (Appendix 2).

Korean vowel mapping with *bVt* English words. The study first found that, for the L1-similar L2 English vowels, when the learners' acoustic distance was close to the native speaker reference, their perception accuracy tended to be high. In contrast, for the L1-somewhat similar and L1-dissimilar L2 vowels, the EFL learners' acoustic distance did not always make accurate predictions of the learners' perception accuracy of the vowels. The study also found a significant effect of vowel type as the learners showed higher perception accuracy for the L1-similar L2 vowels than for the L1-somewhat similar and L1-dissimilar L2 vowels. However, the accuracy of vowel perception showed variations for each vowel type since the perception accuracies of some L1-somewhat similar and L1-dissimilar L2 vowels were high when the vowels carry noticeable perceptual cues like r-coloring (*Burt*) or vowel duration (*bait*). Therefore, these results are only partly in line with the propositions of the SLM/SLM-r concerning the relative difficulty and easiness of the acquisition of L1-similar L2 vowels versus L1-dissimilar L2 vowels. Also, given that variations existed in perception accuracy within each vowel type, different L2 vowels seem to develop at different rates. The study contributes to advancing L2 phonological acquisition theory with its investigation of the joint effects of L1-L2 perceptual similarity and acoustic distances (rather than cross-language acoustic similarity) on the Korean EFL learners' perception of L2 English vowels.

References

- Alispahic, S., K. E. Mulak and P. Escudero. 2017. Acoustic properties predict perception of unfamiliar Dutch vowels by adult Australian English and Peruvian Spanish listeners. *Frontiers in Psychology* 8(52), 1-11.
- Balas, A., M. Wrembel, K. Kaźmierski and J. Weckwerth. 2023. Factors determining perceptual and acoustic similarity between native and non-native vowels. In *Proceedings of the 20th International Congress of Phonetic Sciences–ICPhS 2023*, 2612-2616.
- Bates, D., M. Mächler, B. M. Bolker and S. C. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1-48.
- Best, C. T. and M. D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In O. S. Bohn and M. J. Munro, eds., *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 13-34. John Benjamins.
- Boberg, C. 2015. North American English. In M. Reed and J. M. Levis, eds., *The Handbook of English Pronunciation*, 229-250. John Wiley & Sons.
- Boersma, P. and D. Weenink. 2014. Praat: doing phonetics by computer [Computer Software] Available online at <http://www.praat.org>
- Casserly, E. and D. Pisoni. 2010. Speech perception and production. *Cognitive Science* (1), 629-647.
- Cebrian, J., C. Gorbil and N. Gavalda. 2021. When the easy becomes difficult: Factors affecting the acquisition of the English /i:/-/ɪ/ contrast. *Frontiers in Communication* 6, 1-17.
- Cho, M-H. and S-S. Lee. 2023. The perception-production link varies with stages of L2 development and vowel properties. *Posnan Studies in Contemporary Linguistics* 59(2), 315-340.
- Clopper, C. G., D. B. Pisoni and K. de Jong. 2005 Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118, 1661-76.
- DiCanio, C., H-S. Nam, J. D. Amith, R. C. García and D. H. Whalen. 2015. Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics* 48, 45-59.
- Doan, A. and E. Oh. 2023. The role of L2 experience on the perceived similarity and identification of British English vowels by Vietnamese speakers. *Linguistic Research* 40 (Special Edition), 127-149.

- Edwards, H. 2003. *Applied Phonetics: The Sounds of American English* (3rd edition). Thomson Delmar Learning.
- Elvin, J., D. Williams, J. A. Shaw, C. T. Best and P. Escudero. 2021. The role of acoustic similarity and non-native categorization in predicting non-native discrimination: Brazilian Portuguese vowels by English vs. Spanish listeners. *Languages* 6(1), 44:1-29.
- Escudero, P. 2005. *Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization*. Doctoral dissertation, Utrecht University.
- Escudero, P. and K. Yazawa. 2023. The second language linguistic perception model (L2LP). In M. Amengual, ed., *The Cambridge Handbook of Bilingual Phonetics and Phonology*, 1-37. Cambridge University Press.
- Evans, B. G. and P. Iverson. 2007. Plasticity in vowel perception and production: A study of accent in young adults. *Journal of the Acoustical Society of America* 121, 3814-26.
- 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*, 233-272. York Press.
- Flege, J. E. 1998. The role of subject and phonetic variables. In K. Gruber, D. Higgins, K. Olsen and T. Wysocki, eds., *Papers from the 34th Annual Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society.
- Flege, J. E. 1999. Age of learning and second language speech. In D. Birdsong, ed., *Second Language Acquisition and the Critical Period Hypothesis*, 101-131. Lawrence Erlbaum Publishers.
- Flege, J. E. and O. S. Bohn. 2021. The revised speech learning model (SLM-r). In R. Wayland, ed., *Second Language Speech Learning: Theoretical and Empirical Progress*, 3-83. Cambridge University Press.
- Hillenbrand, J., L. Getty, M. Clark and K. Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97, 30993-111.
- Jun, S-A. and I. Cowie. 1994. Interference for 'new' and 'similar' vowels in Korean speakers of English. *Ohio State University Working Papers* 43, 117-130.
- Kuznetsova, A., P. B. Brockhoff and R. H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13), 1-26.
- Ladefoged, P. 2006. *A Course in Phonetics* (5th edition). Thomson Wadsworth.
- Lee, K-H., W-J. Kim and S-S. Lee. 2023. Effects of English-learning context on English and Korean vowel productions: Focusing on Korean female learners of English. *Language and Linguistics* 101, 49-80.
- Lee, S-S. and D-J. Shin. 2015. Estimating Korean EFL listeners' perception of English vowels with reference to cross-language labelling. *Studies in Phonetics, Phonology and Morphology* 21, 297-321.
- Lee, S-S., H-S. Nam, J-K. Kang, D-J. Shin and Y-S. Kim. 2017. The impact of language-learning environments on Korean learners' English vowel production. *Phonetics and Speech Sciences* 9(2), 69-76.
- Lee, S-S., J-K. Kang and H-S. Nam. 2022. Identification of English vowels by non-native listeners: Effects of listeners' experience of the target dialect and talkers' language background. *Second Language Research* 38(3), 449-475.
- Lee, S-S., J-K. Kang, D-J. Shin and H-S. Nam. 2019. The effect of dialect experience and talker variation on Korean EFL listeners' English vowel identification. *Studies in Foreign Language Education* 33(4), 1-27.
- Lee, S-S. and M-H. Cho. 2020. The impact of L2-learning experience and target dialect on predicting English vowel identification using Korean vowel categories. *Journal of Phonetics* 82, 1-20.
- Lee, S-Y., J-Y. Kim, K-Ch. Nam, H-A. Jeon and Y-J. Kim. 2018. Neurocognitive study on the ongoing merge of Korean vowel /e/ and /E/: Comparing MMNs of younger and older generations. *Studies in Phonetics, Phonology and Morphology* 24(2), 209-224.
- Lee, Y-C. 2023. Revisiting the merger of mid front vowels in Seoul Korean. *Studies in Phonetics, Phonology and Morphology* 29(2), 193-209.

- Levy, E. S. and W. Strange. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36, 141-157.
- Llompарт, G. M. and E. Reinisch. 2019. Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech* 62(3), 594-622.
- 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.
- Moyer, A. 1999. Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition* 21, 81-108.
- Park, H-Y. and K. J. de Jong. 2008. Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics* 36, 706-723.
- Park, H-Y. and K. J. de Jong. 2017. Perceptual category mapping between English and Korean obstruents in non-CV positions: Prosodic location effects in second language identification skills. *Journal of Phonetics* 62, 12-33.
- R Core Team. 2019, 2023. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Available online at <http://www.R-project.org/>
- Saito, K. and A. Tierney. 2025. Roles of domain-general auditory processing in second language speech learning revisited: What degree of precision makes a difference? *Language Learning*, 1-42.
- Strange, W. 2007. Cross-language phonetic similarity of vowels: Theoretical and methodological issues. In O. S. Bohn and M. J. Munro, eds., *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, 35-55. John Benjamins.
- Strange, W., O. S. Bohn, S. A. Trent and K. Nishi. 2004. Acoustic and perceptual similarity of North German and American English vowels. *Journal of Acoustical Society of America* 115, 1791-1807.
- Strange, W., R. Akahane-Yamada, R. Kudo, S. A. Trent, K. Nishi and J. J. Jenkins. 1998. Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics* 26, 311-344.
- Trofimovich, P., S. Kennedy and J. Foote. 2015. Variables affecting L2 pronunciation development. In M. Reed and J. Levis, eds., *The Handbook of English Pronunciation*, 353-373. John Wiley & Sons.
- Tyler, M. D. 2021. Phonetic and phonological influences on the discrimination of non-native phones. In R. Wayland, ed., *Second Language Speech Learning: Theoretical and Empirical Progress*, 157-174. Cambridge University Press.
- Yang, B-G. 1996. A comparative study of English and Korean monophthongs produced by male and female speakers. *Journal of Phonetics* 24, 245-261.
- Yang, B-G. 2014. A link between perceived and produced vowel spaces of Korean learners of English. *Speech Sciences* 6(3), 81-89.

Examples in: English

Applicable Languages: English

Applicable Level: Secondary, Tertiary

**Appendix 1. Average Formant Frequencies (F1, F2, and F3 in Hz) and Vowel Durations (in ms) of
Stimulus Words: Native English Speakers**

	beat	bit	bait	bet	bat	but	Burt	Bart	bot	bought	boat	boot
F1	333	499	426	703	913	721	513	677	780	715	479	375
F2	2644	2017	2068	1882	1709	1451	1438	1307	1175	1115	1260	1384
F3	3078	2720	2686	2789	2641	2783	1724	1970	2594	2549	2678	2546
Duration	192	157	202	169	215	178	190	214	209	215	218	187

Appendix 2. Average Formant Frequencies (F1, F2 in Hz) of Stimulus Words: EFL Learners

	beat	bit	bait	bet	bat	but	Burt	Bart	bot	bought	boat	boot
F1	418	428	483	748	788	723	586	749	805	628	464	427
F2	2490	2399	2347	1690	1704	1240	1414	1358	1226	1078	957	1305

Appendix 3. EFL Learners' English-to-Korean Vowel Mapping Probability

	beat /i/	bit /ɪ/	bait /eɪ/	bet /ɛ/	bat /æ/	but /ʌ/	Burt /ɜ:/	Bart /ɑ:/	bot /ɑ/	bought /ɔ/	boat /oʊ/	boot /u/
/i/ < ㅇㅣ >	1.00 (5.6)	.94 (5.4)										
/e/ < ㅓㅣ >			.36 (2.6)	.30 (5.7)	.30 (5.8)							
/ɛ/ < ㅓㅣ >			.64 (3.2)	.67 (5.6)	.67 (5.4)							
/ʌ/ < ㅓㅣ >						.97 (5.6)	.92 (4.6)	.18 (4.8)	.59 (4.7)	.68 (4.2)		
/u/ < ㅜㅣ >												
/a/ < ㅓㅣ >								.82 (4.4)	.35 (4.5)			
/o/ < ㅜㅣ >										.23 (3.1)	.92 (3.1)	
/u/ < ㅜㅣ >												.94 (5.6)
/jɛ/ < ㅓㅣ >												
/wa/ < ㅓㅣ >												
/jʌ/ < ㅓㅣ >												
/jo/ < ㅜㅣ >												
/ju/ < ㅜㅣ >												
/wɛ/ < ㅓㅣ >												

Note: Only mapping probability above chance level (i.e., more than 8.3) is provided. Goodness-of-fit ratings are in parentheses.