*Linguistic Research* 40(Special Edition): 127-149
DOI: 10.17250/khisli.40..202309.005

# The role of L2 experience on the perceived similarity and identification of British English vowels by Vietnamese speakers*

**Thi Lan Anh Doan** · Eunhae Oh*****
**(Konkuk University)**

**Doan, Thi Lan Anh and Eunhae 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. This study investigated the effects of L2 experience on the cross-language influences on the perception of English vowels by Vietnamese learners of English. Two groups of Vietnamese learners of English, inexperienced and experienced, completed L2-to-L1 vowel mapping and L2 vowel identification tasks. The mapping task results showed that the inexperienced learners mapped English vowels more closely to their corresponding Vietnamese vowels, while experienced learners showed more dispersed mapping responses. In the L2 vowel identification task, we found that the experienced learners showed a significantly higher accuracy in identifying English vowels, especially those that lacked clear counterparts in Vietnamese. As participants gained more experience with the L2, their perception was found to be more sensitive and gradient. However, the accuracy of identifying some similar vowels appeared to vary with temporal and articulatory features, indicating the importance of considering other phonetic features and their interactions with spectral features for a more comprehensive understanding of the difficulty associated with "new" or "similar" L2 vowel identification. The findings shed light on the intricate and dynamic interplay of L1 and L2 experience in shaping the perception of L2 vowels. **(Konkuk University)**

**Keywords** L2 experience, L1 and L2 perceptual similarity, L2-to-L1 vowel mapping, vowel identification, Vietnamese learners of English

---

* This paper was modified and developed from the Master's thesis of the first author. Many thanks to the anonymous reviewers for their valuable comments and feedback.
** First author
*** Corresponding author

## 1. Introduction

### 1.1 Previous research

Studies on cross-language influences have shown that learners of a second language (L2) encounter challenges in the perception of non-native vowels and consonants that are not perceptually distinctive in their native language (L1) (Jamieson and Morosan 1986; Polka 1991, 1995; Best and Strange 1992; Aoyama et al. 2004; Levy and Strange 2008). The Speech Learning Model (SLM) by Flege (1995, 2002) predicts the extent of perceptual difficulties of L2 phonetic categories based on the existing knowledge of the L1. The framework acknowledges that certain sounds may be easier or more challenging to acquire based on the similarity or dissimilarity between the L1 and L2 sound systems. It categorizes the L2-to-L1 phonetic mappings into three types: identical, similar, and new. The SLM predicts that L2 learners would be more successful in learning 'new' rather than 'similar' L2 phones as learners can readily establish distinct and separate categories for the new L2 sounds that are dissimilar from any L1 sounds. In Flege (1987), English learners of French who had lived in France for 10 years showed no significant difference from French native speakers in terms of F2 frequency for the "new" French vowel /y/, but a significant difference for the "similar" vowel /u/. Bohn and Flege (1992) investigated the "new" and "similar" English vowels produced by native Dutch adults. English vowels /i/, /u/, /a/, /ʌ/ were categorized as "similar" to Dutch vowels and /æ/ as "new". Contrary to predictions, however, some Dutch speakers were able to produce the similar vowels as accurately as the new vowel, leading to the exclusion of the "new-similar-identical" trichotomy from the SLM.

The revised version of SLM (SLM-r) (2021) further proposes that L2 adult learners' ability to discern LI- L2 phonetic differences improves across the life-span when they receive consistent and meaningful L2 input. It proposes that "the mechanisms and processes used to establish the elements making up the Ll phonetic system, including the ability to form phonetic categories, remain intact and available for L2 learning" (2021: 18). Munro (1993) provided supporting evidence by examining the production of ten English vowels by Arabic-speaking learners of English who acquired the language later in life. The 23 native Arabic speakers between the age of 19 and 57 years who resided in the U.S. for an average of 5.7 years were asked to produce were asked to read a list of /bVt/ words in a sentence frame. The result showed that some Arabic speakers'

production of English vowels was perceived as native-like by five native English speakers, which suggest that, with sufficient L2 experience, L2 learners of all ages retain the capacity to establish L2 phonetic categories in a native-like manner. In Kondo (2010), the production of English multisyllabic words by eight Japanese adult speakers with extended experience in English-speaking environments was examined. Comparisons were made between the Japanese speakers and native English speakers in terms of vowel duration, fundamental frequency (F0), intensity, and unstressed vowel spectral quality. The results revealed that, among these acoustic features, only the quality of unstressed vowels significantly differed between the two groups. This suggests that the experienced adult Japanese learners were able to acquire and produce L2 prosodic features in a native-like manner, even though the prosodic features were not present in their L1.

L2 experience not only influences the acquisition of the L2 sounds but also has an impact on the L1 segmental features. For instance, Oh et al. (2011) conducted a longitudinal study on English vowels produced by Japanese children residing in the United States. They found that the children were able to accurately produced English "new" vowels /ɪ/, /ɛ/, /ɑ/, /ʌ/ and /ʊ/ in a year's time. They also reported that their Japanese vowels /i/ and /a/ shifted due to the L2 exposure. The changes observed in the L1-L2 vowel inventory were interpreted to suggest significant effects of L2 experience on the dynamic bi-directional interaction between the L1 and L2 vowel systems. Additionally, de Leeuw, Menne, and Scobbie (2013) focused on the lateral phoneme /l/ in the L1 and L2 of ten late German-English bilinguals who all learned the L2 in late adolescence to adulthood. The acoustic analyses of German and English words containing the lateral revealed clear evidence of L1 attrition across groups and individuals, particularly in terms of F1 rather than F2 values. The results showed that, with extensive L2 experience, a significant amount of adaptability was observed in the L1 phonetic system during post-adolescence and throughout the lifespan. In Cabrelli et al. (2019), they investigated the impact of an L2 phonological structure that isn't allowed in the participants' L1 on their perception. A total of 15 Brazilian Portuguese (BP) native speakers who were immersed in an English-speaking environment were given an explicit metalinguistic vowel identification task and two ABX tasks. The aim was to determine if the participants' perception of coda stops in English influenced their perception of similar tasks in their native language, BP. The results revealed a notable distinction between the participants' BP sensitivity data and that of BP monolingual individuals. The participants perceived coda obstruents differently from what would be expected in

monolingual BP speakers, providing evidence of L2 phonotactic influence in L1 perception.

In line with the new-similar L2 sound perception within the framework of the SLM, Park and de Jong (2008) developed the L1 mapping model which aims to predict the L2 consonant identification accuracy using L1 categories. Korean listeners participated in the mapping task, where they labeled English consonants with the similar Korean phonetic segments. They reported that the accuracy in identifying L2 consonants can be better predicted when mapping to L1 categories for similar consonants. Building upon this research, Lee and Cho (2018) conducted a subsequent study to investigate whether the predictive power of the model remains consistent when applied to the identification of L2 English vowels by Korean listeners with varying amount of L2 experience. The findings revealed that the model accurately predicted the identification of English vowels that closely resembled Korean vowels, especially when the listeners' goodness ratings were taken into consideration. However, the effectiveness of the L1 mapping model, based on the new-similar distinction, was lower compared to consonants, due to the less categorical nature of vowels. In support of this perceptual difference, Gerrits and Schouten (2004) conducted research that demonstrated the absence of categorical perception for vowels in both isolated and text conditions. They argued that there was a significant difference in the degree of categorical perception between stop consonants and vowels and concluded that categorical perception might be influenced by the type of task used, particularly forced-choice discrimination tasks, which may not fully capture the listeners' ability to access detailed perception of phonological categories. Considering that the perceptual challenges posed by L2 sounds vary from categorical stop consonants to more continuous vowel categories, the gradual and continuous nature of learners' perceptual improvement for L2 vowels may be better achieved through identification tasks (Fry et al. 1962; Repp 1981; Gerrits and Schouten 2004; Mirman et al. 2004).

Several studies have examined the relationship between learners' L2 vowel identification accuracy and the level of exposure to the L2. Kitikanan (2022) explored the effect of L2 exposure on the identification of British English vowels by 52 Thai speakers with or without an extensive phonetic training. The identification of British English “new” vowels /e/, /ɔː/, /ɪ/, and /ɜː/ showed significantly higher correct identification scores in the trained group compared to the non-trained group. Lee and Cho (2020) compared the vowel identification accuracy of long-LOR (mean 11 years of length of residence in the United States) and short-LOR (4 years) Lorean listeners in a relatively

less familiar Standard Southern British English. The results showed that the short-LOR listeners' identification accuracy of the vowels was overall lower than that of the long-LOR listeners. Although both studies align with previous studies on the meaningful influence of L2 experience on the improvement of L2 vowel identification accuracy, L1-L2 mapping did not accurately predict the results. Lee and Cho (2020) proposed the need for a more comprehensive investigation into the distinction between similar and new L1-L2 vowel mapping, They argued that the acquisition of L2 vowels occurs on a gradient scale rather than being strictly categorical as factors, including L2 proficiency, learners' familiarity with the L2 dialect, and the extent of dissimilarities between the L1 and L2 vowels, may influence learners' L2 category development.

### 1.2 Present study

While many researchers have explored the role of L2 experience on accurate L2 vowel identification, there has been relatively limited investigation into the changes in perceived similarity between L1 and L2 vowels as a function of L2 experience, and how these changes contribute to improved L2 vowel identification accuracy. In addition, the existing literature on Vietnamese speakers of English in L2 speech perception is more scarce. Some studies have examined the perception of English prosody, stress, and intonation among adult Vietnamese learners (Nguyen, Ingram, and Pensalfini 2008; Cunningham 2013; Nguyen 2018), while others have centered on the difficulties encountered by Vietnamese speakers in producing English consonants (Nguyen 2015; Bui 2016). Most relevant to the current research, Winn et al. (2008) conducted a study examining the production of Vietnamese vowels by native English speakers. The English-speaking adult learners encountered difficulties in distinguising between Vietnamese central and back vowels. Furthermore, adult learners showed difficulties in accurately producing the long (i.e., ơ and a) and short (i.e., â and ă) vowels due to the absence of phonemic quantity distinctions in English. More recently, Đào and Nguyễn (2018) conducted research with the goal of examining the production of Vietnamese 9 vowels by adult L2 Korean learners. The production of 9 Vietnamese vowels produced by Korean learners of Vietnamese showed a merging of the three Vietnamese vowels /ɔ/, /o/, and /ɤ/ as a result of phonetically similar Korean vowels /o/ and /ʌ/. The study provided insights into the interaction between Korean and Vietnamese vowel systems and the impact of

cross-language similarity. However, there remains a significant gap in the existing literature regarding the perceptual challenges in discriminating L2 vowels by Vietnamese adult learners. The main objectives of the present study are 1) to explore the perceived phonetic relationship between Vietnamese vowels (L1) and English vowels (L2) based on the learners' L2 experience and 2) to investigate whether this perceived relationship leads to higher accuracy in identifying L2 vowels.

Vietnamese has been characterized by three main dialects: northern, central, and southern. While these dialects do not differ significantly in terms of grammar, they vary in certain aspects of pronunciation, particularly consonants. The dialects correspond to the speech patterns of specific regions such as Hanoi (northern), Hue (central), and Ho Chi Minh City (southern). However, when it comes to vowels, there is a general similarity observed across these dialects (Alves 2006). Regarding the Northern Vietnamese vowel inventory, researchers do not unanimously agree on the number of monophthongs. Crothers (1978) described Vietnamese as having ten monophthong vowels, including /i, e, ɛ, a, ɔ, o, u, ʌ, ɤ, ɯ/. Other researchers do not classify central [ɐ] as a separate phoneme and categorize it with /a/ (Smalley 1954; Lê Văn Lý 1960; Crothers 1978). According to Haudricourt (1952), Đoàn, (1977) and Kirby (2011), the Vietnamese vowel system is composed of nine monophthongs /i, e, ɛ, a, ɔ, o, u, ɤ, ɯ/ and the vowels /ɐ/ and /ʌ/ as allophones of phonemes /a/ and /ɤ/[1]. As shown in Table 1, the 9-vowel system was adopted for the current study.

1 In the original mapping task, both allophones of the phonemes /a/ and /ɤ/ were included. However, the goodness ratings for the allophones were consistently below 3, and the responses for these allophones did not contribute significantly to the results, leading to the decision to combine the results of the phonemes with those of the allophones for the current study.

Table 1. Vietnamese vowels

| | Front | Back [-round] | Back [+round] |
|---|---|---|---|
| High | /i/ | /ɯ/ | /u/ |
| Mid | /e/ | /ɤ/ | /o/ |
| Lower-mid | /ɛ/ | | /ɔ/ |
| Low | /a/ | | |

The Standard Southern British English (SSBE) vowels were chosen for the mapping and identification tasks based on a study conducted by Lee and Cho (2020). They demonstrated that an L1-based model accurately predicted relatively less experienced Korean listeners' identification of SSBE vowels more effectively than their identification of North American English (NAE) vowels, indicating that the learners' familiarity with the specific L2 dialect can influence the effects of L2-learning experience on L2 vowel identification. Our research goal was to achieve better comparability with the findings from Lee and Cho (2018, 2000) by exploring whether L2 experience still exhibits similar effects when the dialect is less familiar to native Vietnamese speakers, an aspect that has not been extensively researched. The vowel system of SSBE comprises eleven monophthongs /i/, /ɪ/, /e/, /æ/, /ʌ/, /ɑ/, /ɒ/, /ɜ/, /ɔ/, /ʊ/, and /u/. Among the eleven vowels, four tense vowels (/i/, /ɑ/, /ɔ/, /u/) and six lax vowels (/ɪ/, /e/, /æ/, /ʌ/, /ɒ/, /ʊ/) were chosen for the current study (see Table 2). The open-mid central unrounded vowel (represented by the IPA symbol /ɜ/), which is typically spelled with the letters 'e-r', as in words like "her" was not included due to the use of nonse words.

Table 2. Standard British English vowels

| | Front | Central | Back |
|---|---|---|---|
| High | /i/, /ɪ/ | | /ʊ/, /u/ |
| Mid | /e/ | /ə/ | /ɒ/, /ɔ/ |
| Low | /æ/ | /ʌ/ | /ɑ/ |

Figure 1 illustrates vowel diagrams of Vietnamese (left) and SSBE (right) vowels. As L1 vowel categories play a critical role in predicting the accuracy of L2 vowel identification, the dissimilarities between Vietnamese and English are expected to contribute to the challenges of perceiving the lax and tense vowel distinction in English. Figure 1 shows

that Vietnamese and English vowels /i/ overlap in the vowel space, leading to an expectation that these L2 vowels would be perceived as equivalent to the acoustically "identical" vowels in the L1. Note, however, that a problem arises when determining the "identical" vowels based on the International Phonetic Alphabet (IPA), as some IPA symbols (e.g., /e/, /ɔ/, /u/) are shared by both languages but do not correspond to the same acoustic information (F1, F2) in the vowel space. Figure 1 illustrates that the F1 values for /e/ are higher and for /ɔ/ are lower in Vietnamese compared to the English counterparts based on the IPA. English /u/ is located in between Vietnamese /ɯ/ and /u/. Conversely, different IPA symbols between the two languages may represent similar vowel qualities, such as Vietnamese /o/, rather than /ɔ/, being acoustically closer to English /ɔ/. While the IPA provides a useful framework for categorizing vowel sounds based on articulatory features, it may not accurately correspond to the acoustic vowel space when plotting the first formant frequencies (see Vaissière 2011).

The identification of "new" and "similar" vowels between the two languages presents similar challenges. Vietnamese lacks direct counterparts for lax vowels /ʊ/, /ɪ/, as well as /æ/, /ʌ/, and /ɒ/. These would be considered as "new" L2 vowels and are expected to exhibit sufficient differences for accurate perceptual identification. Mapping English /ɪ/ and /ɒ/ onto the Vietnamese vowel space reveals that no acoustically similar vowels are located around them, making these "new" vowels distinctive enough to be identified. However, English /æ/, /ʌ/, and /ʊ/ share similarities with Vietnamese /ɛ/, /a/, and /ɯ/, respectively, in the Vietnamese vowel space. This leads to varying perceptual distances among the "new" L2 vowels. These discrepancies between IPA symbols and vowel qualities in Vietnamese and English need to be acknowledged and accounted for in predicting perceptual identification of L2 vowels based on the L1. That is, the absence of corresponding IPA symbols in the L2 may not necessarily predict their perceptual novelty.

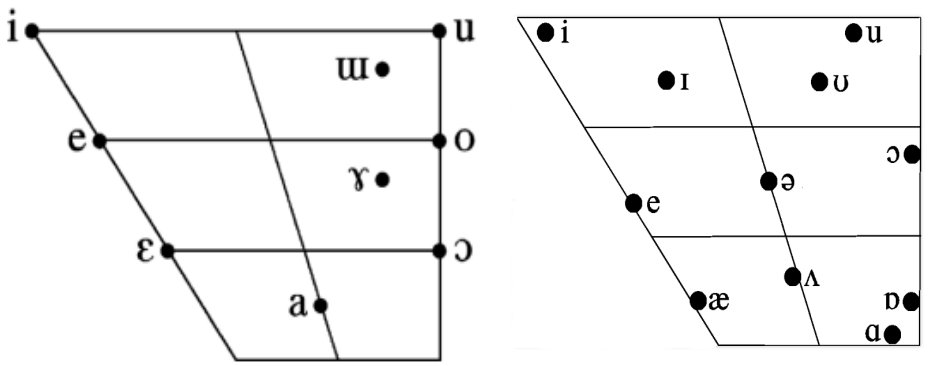


Figure 1. Vietnamese (left, Kirby, 2011) and Standard Southern British English (right, Roach, 2006) monophthong centroids in a schematic F1-F2 space

The current study aims to investigate the perceived phonetic relationship between Vietnamese vowels and English vowels based on Vietnamese learners' experience with English. We also seek to determine whether the experience leads to higher accuracy in identifying L2 vowels, especially for the new and similar vowels. Two experiments were conducted. In Experiment 1, two groups of participants, inexperienced and experienced Vietnamese listeners, were given an L2-to-L1 mapping task to examine how English vowel categories were mapped onto Vietnamese vowel categories. In Experiment 2, an English vowel identification task was conducted to explore whether experienced Vietnamese listeners were more accurate in identifying the English vowels that do not exist in the learners' native language. Goodness ratings were also included to determine the degree of similarity between the L1-L2 vowels and identification accuracy for each experiment.

## 2. Experiment 1

### 2.1 Method

#### 2.1.1 Participants

Ten native Vietnamese listeners participated in both English-to-Vietnamese vowel mapping and English vowel identification experiments. All of the listeners were students

at a university located in the Northern Vietnam (mean age = 23 years). The listeners were divided into two groups based on the extent of English-learning experience. The inexperienced group consisted of five listeners who had never resided in an English-speaking country and did not have English as their major. Nonetheless, the listeners have received English education as a regular course at school starting from the age of 12 years old. The experienced group comprised five participants who had the experience of living in an English-speaking country, with a mean length of residence of 2.1 years. The experienced listeners also possessed an IELTS score of over 7, indicating an advanced level of English proficiency. None of the participants reported any speech or hearing disorders.

### 2.1.2 Speech stimuli

Speech stimuli were recorded by a female native speaker of SSBE from London (age = 24) for both experiments. The speaker was recorded in a soundproof booth at a university, using a Shure KSM 44 microphone and a Tascam (HD-P2) solid-state recorder. To minimize the possible effect of word familiarity on inexperienced learners' perception of English words as well as the effect of consonantal contexts on vowel perception, nonce words were used. Target English vowels were embedded in the context of /h V d/ for 10 vowels (/i/, /ɪ/, /e/, /æ/, /ʌ/, /ɑ/ /ɒ/, /ɔ/, /ʊ/, /u/) and produced in a framed sentence of "I say _____ again." two times. Each target vowel was extracted for the perception tasks, using Praat (Boersma and Weenink 2016). The vowel was determined by referencing the onset and offset of clear energy in the second formant frequency on the sound spectrogram and the waveform. The same speech stimuli were used for both experiments.

### 2.1.3 Procedure

Both experiments were conducted in the same soundproof laboratory at a university. The L2-to-L1 mapping vowel task was completed before the English vowel identification task and prior to each task, the participants received instructions in their native language by a native Vietnamese speaking experimenter. For the mapping task, the participants listened to the English stimuli in a random order and were tasked with matching the

given English vowel to the most similar Vietnamese vowel from a set of alternatives displayed in Vietnamese orthography on the computer monitor (i,y-/i/, e-/ɛ/, ê-/e/, a-/a/, o-/ɔ/, ô-/o/, ơ-/ɤ/, u-/ɯ/, u-/u/). Afterwards, the participants rated the goodness of the vowels on a scale 5-point scale, where 1 indicated “low similarity” and 5 “high similarity”.

### 2.2 Results and discussion

Experiment 1 examined the perceived similarity between Vietnamese and British English vowels by native Vietnamese listeners with varying levels of English experience. The results were deemed to predict the accuracy of English vowel identification in Experiment 2. Tables 3 and 4, respectively, show the inexperienced and experienced listeners’ response percentages of Vietnamese labeling chosen for each English stimulus, along with their corresponding mean goodness rating.

Table 3. Vietnamese inexperienced listeners’ mean percentage of labeling English vowels with Vietnamese vowels with mean goodness ratings in parentheses. Percentages above 50 are marked in bold.

| English \ Vietnamese | /i/ | /ɛ/ | /e/ | /a/ | /ɔ/ | /o/ | /ɤ/ | /ɯ/ | /u/ |
|---|---|---|---|---|---|---|---|---|---|
| /i/ | **100** (4.2) | | | | | | | | |
| /ɪ/ | **60** (4.0) | | 40 (4.3) | | | | | | |
| /e/ | | 20 (3.0) | **60** (3.8) | 20 (3.5) | | | | | |
| /æ/ | | | | **50** (3.3) | | | **50** (3.0) | | |
| /ʌ/ | | | | **60** (3.8) | 20 (3.0) | | 20 (2.5) | | |
| /ɑ/ | | | | **60** (3.5) | 40 (3.3) | | | | |
| /ɒ/ | | | | | **60** (4.0) | 40 (3.5) | | | |
| /ɔ/ | | | | | | **100** (4.1) | | | |
| /ʊ/ | | | | | | | 30 (3.0) | **50** (4.4) | 20 (3.0) |
| /u/ | | | | | | | | **80** (3.8) | 20 (4.0) |

Table 4. Vietnamese experienced listeners' mean percentage of labeling English vowels with Vietnamese vowels with mean goodness ratings in parentheses. Percentages above 50 are marked in bold.

| English \ Vietnamese | /i/ | /ɛ/ | /e/ | /a/ | /ɔ/ | /o/ | /ɤ/ | /ɯ/ | /u/ |
|---|---|---|---|---|---|---|---|---|---|
| /i/ | **100** (4.0) | | | | | | | | |
| /ɪ/ | 30 (3.0) | | **60** (2.8) | | | | | | 10 (1.0) |
| /e/ | | **60** (3.3) | 10 (3.0) | 10 (3.0) | | | 20 (1.0) | | |
| /æ/ | | | | **70** (3.3) | | | 30 (2.0) | | |
| /ʌ/ | | | | 40 (3.75) | 30 (2.0) | | 30 (2.7) | | |
| /ɑ/ | | | | 30 (3.7) | **70** (3.6) | | | | |
| /ɒ/ | | | | | **80** (3.8) | 20 (3.0) | | | |
| /ɔ/ | | | | | | **100** (4.0) | | | |
| /ʊ/ | | | | | | | 30 (2.7) | 40 (3.8) | 30 (3.3) |
| /u/ | | | | | | | | 40 (3.8) | **60** **(4.5)** |

Both groups showed a complete one-to-one mapping of the English phonemes /i/ and /ɔ/ to the Vietnamese vowels /i/ and /o/, respectively, with mean goodness ratings exceeding 4. As predicted by Figure 1, inexperienced listeners mapped English /e/, /ʌ/, /u/ to the acoustically similar Vietnamese /e/, /a/, /ɯ/, respectively. However, the responses changed with L2 experience. First, experienced listeners not only demonstrated more one-to-many mapping patterns compared to inexperienced listeners, but also had lower mean percentages of labeling English vowels with Vietnamese vowels. Specifically, inexperienced listeners showed one-to-many mapping for three English vowels /e, ʌ, ʊ/ with the mean percentages ranging from 50 to 60, while experienced listeners mapped four vowels /ɪ, e, ʌ, ʊ/ with the mean percentages ranging from 40 to 60. As inexperienced listeners gained more experience, their initial associations with English /ʌ/ and /ʊ/ changed from being mapped to /a/ (60%) and /ɯ/ (50%) in Vietnamese to having no apparent mapping. That is, unlike inexperienced listeners, experienced listeners did not show any strong preference to a single category in Vietnamese for /ʌ, ʊ/ in English

(marked in shade). Notably, as indicated with the red box marked in Tables 3 and 4, inexperienced listeners were more likely to label different English vowels with the same Vietnamese vowels compared to experienced listeners: English /i, ɪ/ as Vietnamese /i/, English /æ, ʌ, ɑ/ as Vietnamese /a/, English /ʊ, u/ as Vietnamese /ɯ/. These instances exemplify "Single-category assimilation" according to the Perceptual Assimilation Model (PAM, Best 1995), where L2 phonemes are assimilated into the same phonological category as the corresponding L1 phonemes.

Taken together, L2 experience and acoustic similarity played significant roles in the mapping Vietnamese and SSBE vowels. Inexperienced listeners appeared to heavily rely on the degree of similarity between the target vowel and the corresponding vowel in their L1, leading to a more limited range of mapping options. On the other hand, greater experience led to increased dispersion in mapping responses among experienced listeners. This may indicate that experienced listeners exhibit a greater ability to differentiate L2 vowels from exisiting L1 vowel categories compared to inexperiened listeners. Also, experienced listeners showed a decrease in ratings of goodness for one-to-many mapping, implying a more “refined perception of higher-order invariants in the L2”, providing further support for the PAM-L2 (Perceptual Assimilation Model for Second Language Acquisition, Best and Tyler 2007).

By analyzing the changes in the L2-to-L1 mapping patterns, we may be able to predict the SSBE vowel identification accuracy based on the L2 experience. Inexperienced listeners predominantly classified English /ɪ/ as Vietnamese /i/ (60%), while experienced listeners were more likely to identify it as /e/ (60%). With an increase in experience, English /e/ also shifted from being identified as /e/ (60%) to /ɛ/ (60%) in Vietnamese. The evident downward shift in labeling from /i/ to /e/ and /e/ to /ɛ/ reflects experienced listeners’ perceptual sensitivity to the subtle acoustic differences among English front vowels along the F1 continuum. Similarly, the initial mapping of English /u/ as Vietnamese /ɯ/ (80%) shifted to back vowel /u/ (60%), indicating that experienced listeners may have become more attuned to the changes in vowel backness (i.e., F2 values) as well. Consequently, the acquired sensitivity to the acoustic differences between L1 and L2 vowel categories was expected to lead to higher identification accuracy for /ɪ/, /e/ and /u/. Also, assuming that the English vowels without clear counterparts in Vietnamese suggest that experienced listeners were able to form distinct L2 vowel categories separate from similar L1 vowels, we can expect higher identification accuracy for /ʌ/ and /ʊ/.

## 3. Experiment 2

### 3.1 Introduction

In Experiment 2, SSVE vowel identification task was carried out with two objectives: firstly, to assess the identification accuracy of new and similar vowels for inexperienced Vietnamese listeners based on the English (L2)-to-Vietnamese (L1) vowel category mapping observed in Experiment 1, and secondly, to investigate the effects of L2 experience on increasing L2 vowel identification accuracy.

### 3.2 Method

#### 3.2.1 Procedure

The Vietnamese listeners who participated in the vowel mapping task were also involved in the English vowel identification experiment in a soundproof booth. Prior to Experiment 2, a training session was conducted primarily for the inexperienced listeners to familiarize them with English phonetic symbols. During the training session, all participants were given a script containing the International Phonetic Alphabet (IPA) representation of 10 SSBE vowels. They were then asked to listen to audio recordings and mouse-click on the IPA symbol they believed represented the sound they heard. After each response, all participants received feedback indicating whether their response was correct or incorrect before proceeding to the next stimulus. Each vowel was provided three times in a random order regardless of the participants' response accuracy.

After the training session, the participants were presented with the same set of SSBE stimuli used in Experiment 1. Their task was to choose the English vowel they heard by clicking the corresponding option among ten alternatives provided in IPA symbols on the computer screen. Only IPA symbols were given to minimize the confusion arising from the same orthography used in Vietnamese.

### 3.3 Results and discussion

Tables 5 and 6 report the results of inexperienced and experienced listeners' SSBE vowel identification task with mean accuracy, respectively. Using the 'lme4' package (Bates et al. 2015) in R (R Core Team 2022), a logistic regression model was conducted to examine the impact of L2 experience (experienced, inexperienced) and vowel types (10 vowels) on the accuracy (0 or 1) of participants. In order to assess the effects of L2 experience on accuracy, the coefficient and the corresponding significance levels were examined. The model incorporated Participant as a random effect to account for individual differences. Experience showed a significant positive effect on Accuracy (Estimate = 2.634, $p < 0.001$), suggesting that as experience increases, there is an increased likelihood of achieving higher identification accuracy. The main effect of vowel type was also significant, with certain vowels being more accurately identified than others. Experienced listeners showed substantially higher accuracy for four SSBE vowels /i/ (Estimate = 3.573, $p < 0.001$), /ɪ/ (Estimate = 1.832, $p = 0.02$), /ʊ/ (Estimate = 2.139, $p = 0.009$) and /u/ (Estimate = 2.789, $p = 0.001$) compared to inexperienced listeners. While no significant interactions were found for the remaining vowels at the conventional $p < 0.05$ level, marginally significant interactions were observed for /ɔ/ ($p = 0.05$), /e/ ($p = 0.05$) and /ɒ/ ($p = 0.05$).

Table 5. Vietnamese inexperienced listeners' mean accuracy in SSBE vowels identification (in percentage). Vowels that were inaccurately identified are shaded and responses above 50% are marked in bold.

| | Inexperienced | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Target vowels | /i/ | /ɪ/ | /e/ | /æ/ | /ʌ/ | /ɑ/ | /ɒ/ | /ɔ/ | /ʊ/ | /u/ |
| /i/ | **70** | 30 | | | | | | | | |
| /ɪ/ | **50** | 30 | | 20 | | | | | | |
| /e/ | | | 40 | **50** | 10 | | | | | |
| /æ/ | | | | | 40 | **60** | | | | |
| /ʌ/ | | | 10 | | 40 | 20 | 30 | | | |
| /ɑ/ | | | | | 20 | 10 | **50** | 20 | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| /ɒ/ | | | 20 | | 20 | | 20 | 30 | 10 | |
| /ɔ/ | | | | 20 | 40 | | 10 | 20 | | 10 |
| /ʊ/ | | | | | | | | | **50** | **50** |
| /u/ | | | | | | | | | **50** | **50** |

Table 6. Vietnamese experienced listeners' mean accuracy in SSBE vowels identification (in percentage). The vowel that was inaccurately identified is shaded and responses above 50% are marked in bold.

| | Experienced | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Target vowels | /i/ | /ɪ/ | /e/ | /æ/ | /ʌ/ | /ɑ/ | /ɒ/ | /ɔ/ | /ʊ/ | /u/ |
| /i/ | **100** | | | | | | | | | |
| /ɪ/ | | **90** | 10 | | | | | | | |
| /e/ | | | **70** | 30 | | | | | | |
| /æ/ | | | | 20 | **50** | 10 | 10 | | | |
| /ʌ/ | | | | 10 | **60** | | 10 | 10 | | |
| /ɑ/ | | | | | 10 | **50** | 10 | 30 | | |
| /ɒ/ | | | | | | | **90** | 10 | | |
| /ɔ/ | | | | | | 10 | | **90** | | |
| /ʊ/ | | | | | | | | | **80** | **20** |
| /u/ | | | | | | | | | | **100** |

Tables 5 and 6 display the comparison between experienced and inexperienced listeners' accuracy in identifying the 10 SSBE vowels. The findings demonstrate that experienced listeners outperformed inexperienced ones in all cases. Notably, the accuracy for /i/ (70 to 100), /ɪ/ (30 to 90), /ʊ/ (50 to 80), and /u/ (50 to 100) significantly improved with L2 experience. On the other hand, inexperienced listeners exhibited better accuracy for /i/ (70), /ʊ/ (50), and /u/ (60) compared to other vowels, while facing considerable challenges with new SSBE vowel categories, such as /ɪ/ (30), /æ/ (0), /ɒ/ (20), and /ɔ/ (20). The discrepancies in identification accuracy between SSBE vowels having acoustic similarities to corresponding Vietnamese vowels (/i, ʊ, u/) and those without (/ɪ, æ, ɒ, ɔ/) indicate that inexperienced listeners relied more on their L1 vowel categories to

identify L2 vowels.

Previously, it was hypothesized that the "new" SSBE vowels /ɪ/ and /ɒ/, lacking acoustic similarities with Vietnamese vowels, would be relatively easier to learn compared to the "similar" vowels /æ/, /ʌ/, /ʊ/, which had corresponding Vietnamese vowels (/ɛ/, /a/, /ɯ/). The results from experienced listeners supported this hypothesis, as they demonstrated higher accuracy for /ɪ/ (90) and /ɒ/ (90), while encountering more difficulty with /æ/ (20) and /ʌ/ (60). Even though /ʊ/ was expected to pose a challenge due to its spectral similarity to Vietnamese /ɯ/, experienced listeners were able to distinguish them better, likely due to the lip rounding feature present in English /ʊ/. Similarly, the SSBE low back vowel /ɒ/, being rounded and short, was differentiated from the Vietnamese vowel /a/, which is unrounded and inherently long. Experienced listeners appeared to use these additional features to distinguish similar sounds effectively.

## 4. General discussion

This study investigates the relationship between perceived phonetic similarity of vowels in the L1 and L2 and its impact on L2 vowel identification accuracy. Building on Flege's Speech Learning Model, the current study aimed to understand the challenges L2 learners face in perceiving non-native vowels and how L2 experience influences the accurate perception of L2 phonetic categories. Two experiments were conducted by inexperienced and experienced Vietnamese learners of English. In Experiment 1, participants performed an L2-to-L1 mapping task, where they were asked to match the SSBE vowels to the most similar Vietnamese vowels. The aim was to understand how SSBE vowels were categorized or perceived by the Vietnamese listeners in terms of their native Vietnamese vowel system and how having experience in English would influence the mapping patterns.

The L2-to-L1 mapping patterns observed in the study showed significant differences based on the listeners' experience in English. Vietnamese listeners with limited L2 experience not only demonstrated stronger and more consistent one-to-one mapping (/i/ as /i/, /e/ as /e/, /ɑ/ as /a/, /u/ as /ɯ/) but also more instances of single-category mapping (/i, ɪ/ as /i/, /æ, ʌ, ɑ/ as /a/, /ʊ, u/ as /ɯ/) between SSBE vowels and Vietnamese vowels. In contrast, experienced listeners exhibited more varied and dispersed mapping responses, indicating in a less reliance on one-to-one correspondences. The goodness ratings

provided by experienced listeners were also overall lower. The results were interpreted to suggest that experienced listeners were able to attune to the subtle acoustic differences between L2 and similar L1 vowel categories. That is, as individuals gain more experience and exposure to the L2, their perception of L2 vowel sounds is likely to become more refined and distinct from the similar L1 vowel sounds. Consequently, this sensitivity allows them to better discriminate and categorize L2 vowels separately from their L1 vowel system, contributing to higher accuracy in SSBE vowel identification in Experiment 2. This is in line with the findings reported in previous studies (Flege 1991; Cebrian 2006; Lee and Cho 2018). Specifically, Flege (1991) conducted a study involving Spanish participants who were tasked with recognizing English vowels using Spanish orthographic symbols. When a perfect match wasn't found, participants were rpompted to employ "none" label. The findings showed that experienced listeners used the "none" label more frequently (42%) compared to inexperienced listeners (18%), indicating that the level of L2 experience yielded a more gradient perception for L2 vowels.

Note that the acoustic similarity of the two languages should not be assumed from the same phonetic symbols they share. Among the four IPA symbols (/i, e, ɔ, and u/) shared by Vietnamese and SSBE, only /i/ appears to be identical, as it matches 100% of the time in both the experienced and inexperienced groups. However, for /ɔ/, both groups, regardless of English experience, showed a 100% mapping responses with Vietnamese /o/, indicating that /ɔ/ in the two languages do not represent same vowel category. Moreover, different results are shown for /e/ and /u/ as a function of L2 experience. SSBE and Vietnamese /e/ returned a 60% mapping responses for inexperienced listeners, but this dropped significantly to 10% with L2 experience. In contrast, for SSBE /u/, only 20% of inexerpienced listeners, akin to naïve listners, mapped it to Vietnamese /u/, while experienced listeners showed 60% mapping responses, suggesting that the acoustic properties of /u/ also differ between the two languages. The findings indicate two important points: First, relying on the same phonetic symbols to predict the degree of learning difficulties or the perceptual assimilation of the L1-L2 phonemic categoires can be misleading. Second, the L2 experience of participants in L2-to-L1 mapping perceptual experiments can significantly influence how similar L2 listeners perceive the two sounds to be. The different results between experienced and inexperienced listeners calls for cautious interpretation of the perceptual mapping between identical or simiilar vowels in different languages.

Experiment 2 focused on English vowel identification accuracy for inexperienced and

experienced learners. The results showed that experienced listeners outperformed inexperienced learners in identifying SSBE vowels. Experienced learners exhibited significantly higher accuracy especially for the L2 vowels that lacked direct counterparts such as /ɪ/, /ʊ/, indicating their ability to establish separate L2 vowel categories. Despite displaying higher accuracy in their overall performance, experienced listeners still encountered difficulties for some L2 sounds. For example, SSBE vowel /æ/ showed only 20% accuracy and /ɑ/ 50% accuracy. Both vowels demonstrated strong L2-to-L1 mapping responses (70%), with Vietnamese /a/ and /ɔ/ serving as corresponding matches, respectively. Thus, not all "new" or “similar” vowels can be expected to demonstrate comparable accuracy based solely on their spectral features. In our study, SSBE vowels such as /ɪ/ and /ɒ/ exemplify how additional reinforcing properties, such as vowel length or lip rounding, play an important role in helping listeners differentiate between the two sounds.

The significance of the present study resides in its exploration of languages that have received limited attention within the overarching research domain of L2 perception. We believe that the current research will serve to enhance the reliability and understanding of the documented features pertaining to the effects of L2 experience on dynamic interaction between L1 and L2 perception. However, we also acknowledge a limitation related to our small sample size. A larger and more diverse sample is needed to provide a comprehensive understanding of the cross-language perception of L2 vowels and offer stronger support for the observed patterns. Overall, the study highlights the crucial role of L2 experience in shaping learners' perception of L2 vowels. Experienced listeners demonstrated increased sensitivity to acoustic differences between L1 and L2 vowels, leading to different L2-to-L1 mapping patterns as well as improved vowel identification accuracy. The findings have significant implications for language teaching and learning development, emphasizing the importance of providing consistent and meaningful L2 input to enhance learners' ability to establish native-like L2 phonetic categories.

## References

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada, and Tsuneo 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.

Alves, Mark. 2006. Linguistic research on the origins of the Vietnamese language: An overview. *Journal of Vietnamese Studies* 1(1-2): 104-130.

Bates, Douglas, Martin Mächler, Benjamin Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1): 1-48. https://doi.org/10.18637/jss.v067.i01

Best, Catherine and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20(3): 305-330.

Best, Catherine and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Ocke-Schwen Bohn and Murray J. Munro (eds.), *Language experience in second language learning: In honor of James Emil Flege,* 13-34. Amsterdam: John Benjamins Publishing Company.

Boersma, Paul and David Weenink. 2016. *Praat: doing Phonetics by computer*. Retrieved from http://www.praat.org/

Bohn, Ocke-Schwen and James E. Flege. 1992. The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14: 131-158.

Bui, Sao. 2016. Pronunciation of consonants /ð/ and /θ/ by adult Vietnamese EFL learners. *Indonesian Journal of Applied Linguistics* 6: 125-134.

Cabrelli, Jennifer, Alicia Luque, and Irene Finestrat-Martínez. 2019. Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception. *Journal of Phonetics* 73: 55-69.

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

Crothers, James. 1978. Typology and universals of vowel systems in phonology. *Universals of human language* Volume 2, 93-152. Stanford, CA: Stanford University Press.

Cunningham, Ulla. 2013. Teachability and learnability of English pronunciation features for Vietnamese-speaking learners. In Ewa Waniek-Klimczak and Linda R. Shockey (eds.), *Teaching and researching English accents in native and non-native speakers*, 3-14. Heidelberg: Springer.

de Leeuw, Esther, Ineke Mennen, and James M. Scobbie. 2013. Dynamic systems, maturational constraints and L1 phonetic attrition. *International Journal of Bilingualism* 17(6): 683-700. https://doi.org/10.1177/1367006912454620

Đoàn, Thiện Thuật. 1977. *Ngữ âm tiếng Việt (Vietnamese phonology).* Hanoi: University and Vocational College Publisher.

Flege, James. 1991. The interlingual identification of Spanish and English vowels: Orthographic evidence. *The Quarterly Journal of Experimental Psychology* 43: 701-731.

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

Flege, James. 2002. Interactions between the native and second-language phonetic systems. In Peter Burmeister, Thorsten Piske, and Andreas Rohde (eds.), *An integrated view of language development: Papers in honor of Henning Wode,* 217-244. Trier: Wissenschaftlicher Verlag.

Flege, James and Ocke Bohn. 2021. The revised speech learning model (SLM-r). In Ratree Wayland

(ed.), *Second language speech learning: Theoretical and empirical progress*, 3-83. Cambridge: Cambridge University Press.

Fry, Dennis B., Arthur S. Abramson, Peter D. Eimas, and Alvin M. Liberman. 1962. The identification and discrimination of synthetic vowels. *Language and Speech* 5: 171-189.

Gerrits, Ellen and Marten Egbertus, Hendri Schouten. 2004. Categorical perception depends on the discrimination task. *Perception and Psychophysics* 66(3): 363-376.

Haudricourt, André-Georges. 2017. *The short vowels of Vietnamese*. halshs-01631474f.

Jamieson, Donald G. and David E. Morosan. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception and Psychophysics* 40(4): 205-215.

Kirby, James P. 2011. Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association* 41(3): 381-392.

Kitikanan, Patchanok. 2020. The effect of L2 experience on the perceptual assimilation of British English monophthongs to Thai monophthongs by L2 Thai learners. *English Language Teaching* 13(5): 1-10.

Kondo, Yuko. 2000. Production of schwa by Japanese speakers of English: An acoustic study of shifts in coarticulatory strategies from L1 to L2. In Michael Broe and Janet Pierrehumbert (eds.) *Papers in laboratory phonology 5: Acquisition and the lexicon*, 29-39. Cambridge: Cambridge University Press.

Lý, Lê Văn. 1960. *Le parler vietnamien, sa structure phonologique et morphologique fonctionnelle*. (2nd edition). Viên Kháo company.

Lee, Shinsook and Mi-Hui Cho. 2018. Predicting L2 vowel identification accuracy from cross-language mappings between L2 English and L1 Korean. *Language Science*s 66: 183-198.

Lee, Shinsook and Mi-Hui Cho. 2020. The impact of L2-learning experience and target dialect on predicting English vowel identification using Korean vowel categories. *Journal of Phonetics* 82: 100983.

Levy, Erika S. and Winifred Strange. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36(1): 141-157.

Mirman, Daniel, Lori L. Holt, and James L. McClelland. 2004. Categorization and discrimination of nonspeech sounds: Differences between steady-state and rapidly-changing acoustic cues. *Journal of the Acoustical Society of America* 116(2): 1198-1207.

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

Nguyễn, Anh-Thư. 2018. L2 English rhythm by Vietnamese speakers: a rhythm metric study. *Linguistic Journal* 121: 22-44.

Nguyễn, Anh-Thư, John C. Ingram, and Rob J. Pensalfini. 2008. Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics* 36(1): 158-190.

Nguyễn, Anh-Thư. 2015. Vowel reduction in conversational speech in Vietnamese. *International Journal of Asian Language Processing* 23(2): 91-109.

Oh, Grace E., Susan Guion-Anderson, Katsura Aoyama, James E. Flege, Reiko Akahane-Yamada, and Tsuneo Yamada. 2011. A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking country. *Journal of Phonetics* 39: 156-167.

Park, Hanyong and Kenneth 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(4): 704-723.

Polka, Linda. 1991. Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *The Journal of the Acoustical Society of America* 89(6): 2961-2977.

Polka, Linda. 1995. Linguistic influences in adult perception of non-native vowel contrasts. *The Journal of the Acoustical Society of America* 97(2): 1286-1296.

Repp, Bruno H. 1981. Two strategies in fricative discrimination. *Perception and Psychophysics* 30(3): 217-227.

Roach, Peter. 2004. *English Phonetics and Phonology* (4th edition). Cambridge: Cambridge University Press.

Smalley, William A. and Nguyên Văn Van. 1954. *Vietnamese for missionaries: a course in the spoken and written language of central Vietnamese.* Dalat: Imprimerie Evangelique.

Vaissière, Jacqueline. 2011. On the acoustic and perceptual characterization of reference vowels in a cross-language perspective. *The 17th International Congress of Phonetic Sciences (ICPhS XVII)*, 52-59.

**Thi Lan Anh Doan**
Graduate student
Department of English Language and Literature
Konkuk University
120 Neungdong-ro, Gwangjin-gu
Seoul 05029, Korea
E-mail: lananhdt1910@gmail.com

**Eunhae Oh**
Professor
Department of English Language and Literature
Konkuk University
120 Neungdong-ro, Gwangjin-gu
Seoul 05029, Korea
E-mail: grace1111@konkuk.ac.kr

Received: 2023. 08. 11.
Revised: 2023. 09. 14.
Accepted: 2023. 09. 14.