# THE PERCEPTION OF ENGLISH VOWELS BY NATIVE KOREAN AND MANDARIN SPEAKERS

by

**Alexis Zhou**

**A Thesis**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Master of Arts**

School of Interdisciplinary Studies

West Lafayette, Indiana

May 2022

# THE PURDUE UNIVERSITY GRADUATE SCHOOL
# STATEMENT OF COMMITTEE APPROVAL

**Dr. Olga Dmitrieva, Chair**

School of Languages and Cultures

**Dr. Daniel J. Olson**

School of Languages and Cultures

**Dr. Mariko Moroishi Wei**

School of Languages and Cultures

**Approved by:**

Dr.  Olga Dmitrieva

*Dedicated to my family*

# ACKNOWLEDGMENTS

I first and foremost would like to express my gratitude to my advisor, Dr. Olga Dmitrieva, for her support and guidance on this project, and continued encouragement on all my academic endeavors. I would also like to thank my committee members, Dr. Daniel J. Olson and Dr. Mariko Moroishi Wei, for their continued mentorship and advice.

I am also grateful for my family, who have supported me throughout this journey. I am also thankful to my husband, who pushes me to be my best daily.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AOA: | Age of acquisition |
| L1: | First language |
| L2: | Second language |
| L2LP: | Second language linguistic perception |
| PAM: | Perceptual assimilation model |
| SLM: | Speech learning model |

# ABSTRACT

Second language (L2) learners experience difficulties discriminating L2 contrasts that are absent in their L1 (Aoyama et al., 2004; Bohn, 1995; Flege et al., 1997), often due to interreference from the learner's L1 inventory (Flege, 1987, 1995). Popular theoretical frameworks, such as SLM (Flege, 1995), PAM (Best & Tyler, 2007), and L2LP (Escudero, 2005; van Leussen & Escudero, 2015), all make predictions about the acquisition of L2 phonemes. However, these frameworks do not consider the extent of L1 allophonic variability as a factor in the acquisition of L2 contrasts.

In order to begin addressing this, the present study compared the perceptual discrimination of English vowel contrasts /ɪ-i/ and /ɛ-æ/ by two groups, L1 Mandarin and Korean learners. Both languages have only one phonemic category in the vowel areas where English makes the two relevant contrasts. Both members of each contrast are likely to be assimilated to the only existing L1 category. Furthermore, Mandarin /ɛ/ is but one allophone of a great number of possible contextual realizations of the mid vowel phonemic /E/, while Korean lacks this allophonic variability. While the two groups of leaners are expected to be comparable in their perception of the /ɪ-i/ contrast, Mandarin speakers could have more difficulty in their perception of the /ɛ-æ/ distinction. This prediction is based on the assumption that Mandarin listeners are more likely to accept a non-native realization as a deviant but possible version of the L1 mid-vowel phoneme due to its high variability.

To test this, 16 L1 Standard Mandarin and 14 L1 Korean speakers residing in the United States performed an online AX perceptual discrimination task. The results showed that the two groups discriminated /ɪ-i/ vowels with similar accuracy. The Mandarin group was less accurate than the Korean group on the discrimination of the /ɛ-æ/ pairs. This finding is compatible with the hypothesis that greater allophonic variability in the phonemic category of the L1 can make the acquisition of the relevant L2 categories more challenging. These findings suggest that allophonic variability in the L1 needs to be taken into account when making theory-driven predictions concerning the acquisition of specific phonological categories of the L2.

# CHAPTER 1. INTRODUCTION

## 1.1 Introduction & Goals of the Present Study

It has been proven that L2 learners have difficulties discriminating L2 contrasts that are not found in their L1 phonemic inventories, often due to interference from the learner's L1 inventory (Flege, 1987; 1995). The purpose of this study is to compare the perceptual discrimination of two English vowel contrasts, /ɪ-i/ and /ɛ-æ/, by L1 Mandarin and L1 Korean speaker groups, to see which group would be more accurate in their discriminations. Both speaker groups have difficulties distinguishing /ɪ/ and /æ/ from vowels present in their L1 phonemic inventories, however, it is unclear which speaker group is more successful in their perceptions of these vowels when compared with vowels close by in the vowel space, specifically /i/ and /ɛ/. Unlike Korean, the Mandarin vowel system has a great amount of allophonic variability in the mid-region of the vowel space, which may play a role in the acquisition of the L2 contrasts explored in this study. The results of this study will determine whether allophonic variation does indeed play a role in the perceptual discrimination of these contrasts, and which speaker group will be more successful in their discriminations.

## 1.2 Outline of the Thesis

Chapter 2 provides a review of literature on relevant theoretical frameworks, previous perception studies, and will finish with an in-depth exploration of the vowel systems of Korean and Mandarin, specifically focusing on the allophonic variation of the mid-region of the Mandarin vowel space. Chapter 3 discusses the methodology of the study, providing information on how data was collected, the stimuli selection, and a detailed explanation of the perception task. Chapter 4 presents the results of this study, discussing the models chosen to analyze the perception data. Chapter 5 presents the discussion and conclusions drawn from the data collected along with possible limitations to this study.

# CHAPTER 2. LITERATURE REVIEW

## 2.1 On the Perception of L2 Contrasts

It is well known that speaker's first language (L1) impacts the acquisition of second language (L2) phonemes in a number of ways (Flege, 2002; Flege et al., 1995; Munro et al., 1996, among others). This is true for the perception of phonemes present in an L2 (Flege et al., 1997; Jenkins et al., 1995). Language learners often have difficulties discriminating L2 contrasts that are absent in their L1. Some examples are: the English /l- ɹ/ contrast for L1 Japanese learners of English (Aoyama et al., 2004; Iverson et al., 2003) and the English /ɪ-i/ contrast for L1 Spanish learners of English (Escudero, 2000, 2005; Morrison, 2008), among other L2 contrasts examined in previous perception studies (Bohn, 1995; Bohn & Flege, 1992; Chen et al., 2001; Flege et al., 1997; Flege & MacKay, 2004; Guion et al., 2000). This difficulty in differentiating sounds can be due to the interference from the speaker's L1 inventory (Flege, 1987; 1995). Novel contrasts are thought to be more difficult for L2 learners to perceive precisely because they have not encountered by them in their L1s.

## 2.2 Theoretical Frameworks to Explain Perceptual Difficulties

The Speech Learning Model (SLM) (Flege, 1995), the Perceptual Assimilation model (PAM) & PAM-L2 (Best, 1995; Best & Tyler, 2007), and the Second Language Linguistic Perception model (L2LP) (Escudero, 2005; 2009) are the most common theoretical frameworks used to explain the perceptual difficulties L2 learners have in acquiring L2 phonemes.

### 2.2.1 Speech Learning Model (SLM)

The Speech Learning Model (SLM) makes predictions about the perception of single phonemes. SLM states that the phonic elements making up the L1 and L2 phonetic systems exist in a common phonological space, and therefore will influence each other (Flege, 1995; 1999; 2003; Flege et al., 2002). According to Flege (1988, 1992), L2 vowels can fall into three categories; "identical", where the L1 and the L2 phones are acoustically indistinguishable from each other, "similar", where the new L2 phoneme is acoustically similar but not identical to an L1 category, or "new", where the L2 phoneme does not correspond to any existing L1 category. SLM states existence of the L1 categories that are similar to the L2 ones make it more difficult to perceive these L2 sounds as distinct from similar L1 phones, due to the *equivalence classification* (Flege, 1995). Depending on the category that L2 sound falls into, learners will either create a separate category for the “new” L2 phoneme, or assimilate the “similar” L2 phoneme to an existing L2 category (Flege, 2003) via equivalence classification. Indirectly, SLM makes predictions regarding L2 contrasts discrimination: if two L2 phones are both perceived as ‘equivalent’ to the same L1 category (e.g. one ‘identical’ and one ‘similar’), their discrimination is likely to suffer, since they are both categorized as the same L1 phoneme.

### 2.2.2 Perceptual Assimilation Model (PAM) & PAM-L2

PAM-L2 also makes predictions about how L2 categories will be formed by learners of a second language. One of the main focuses of PAM-L2 is predicting if an L2 learner has “perceived equivalence between and L2 and L1 phonological category”, or in other words, if the learner has assimilated the L2 target phoneme to the phonological category of the L1 (Best & Tyler, 2007: 24). PAM-L2 predicts several perception scenarios for pairs of L2 contrastive sounds. One scenario occurs when separate L2 phonological categories are assimilated to separate categories in the L1, known as a *two-category assimilation* (Best & Tyler, 2007). A second scenario can happen when the two phonemes in the L2 pair are perceived as equivalent to the same L1 phonological category, but “one is perceived as being more deviant than the other”, in other words, one of the two is perceived to be a better representation of the L1 category (2007: 26). For this scenario, known as a *category-goodness assimilation* in PAM, L2 learners are more likely to acquire a new category but with a perceived difference in how well the two phonemes fit phonetically. The third scenario occurs when both L2 phonemes are perceived as equivalent to the same L1 category, but as equally good or poor examples of that category, known as *single-category L2 contrast assimilation* (Best & Tyler, 2007). The last scenario described by PAM-L2 occurs when there is no L1-L2 assimilation, and the phonemes have similarities to several L1 phonological categories, but no specific ones. These phonemes are considered *uncategorized.*

### 2.2.3 Second Language Linguistic Perception Model (L2LP)

The L2LP model differs from SLM (Flege, 1995) and PAM & PAM-L2 (Best, 1995; Best & Tyler, 2007) in that it aims to represent the whole developmental process of L2 acquisition, from naïve, non-native learners to advanced, native-like learners through precise learning tasks and developmental trajectories for learners (van Leussen & Escudero, 2015). In terms of perception of L2 phonemes, Escudero's (2005, 2009) *optimal perception hypothesis* posits that L2 learners initially perceive L2 phonemes in a manner similar to the production of these sounds in their L1 environment, predicting that acoustical differences and similarities between the phonemes of the L1 and L2 will shape learning development (van Leussen & Escudero, 2015). L2LP predicts three scenarios, which can be compared to the scenarios of PAM (Best, 1995). One possible scenario occurs when a pair of contrastive L2 sounds are both acoustically similar to a single L1 sound. Learners then needs to either create a new L2 category or split their existing single L1 category, known as a NEW scenario in L2LP, or *single category assimilation* in PAM (Best, 1995).

The next possible scenario occurs when the productions of the L2 contrasting sounds are acoustically closest to the productions of two different L1 sounds. This is known as a SIMILAR scenario in L2LP, or a *two-category assimilation* in PAM. When learners encounter a SIMILAR scenario, L2LP states the existing L1 categories are replicated and adjusted so the boundaries will line up with those of the L2 contrast. It is at this point where the predictions of L2LP and PAM contrast with the predictions of SLM. L2LP and PAM predict shifting boundaries is easier for learners than creating new categories (Elvin et al., 2014). SLM on the other hand, predicts that similar sounds are more difficult to learn than new sounds, however, it is important to note again that SLM makes explicit predictions about single sounds and not sound contrasts (Flege, 1995).

The third type of scenario can arise when a non-native sound is perceived as more than one L1 category, known as *multiple category assimilation* in the L2LP framework (Escudero & Boersma, 2002; Escudero, 2005). This scenario is called the SUBSET scenario in L2LP, or can be compared to what's known as *uncategorized* or *categorized-uncategorized assimilation* in PAM. Both L2LP and PAM predict this kind of scenario to be less challenging for learners than the NEW or *single category assimilation* scenario (van Leussen & Escudero, 2015).

### 2.2.4 Predictions on Phonemes and Allophones

Although each model has its own unique predictions, all three of the perceptual models discussed above address issues facing L2 learners when acquiring new sounds. However, these models fail to take into account possible influences of allophonic variation on the perception of new L2 contrasts. Since the inventory of an L1 can strongly affect the acquisition of new phonemes in an L2, a learner may face difficulties when trying to differentiate an L2 contrast in a space where there is high allophonic variability in their L1.

## 2.3 The Vowel Inventories of Mandarin & Korean and their Relationships to this Study

### 2.3.1 To be Explored in this Study

In this study, the perception of two Midwestern American English vowel pairs by two different L1 listener groups, L1 Standard Mandarin and L1 Korean speakers of English, is analyzed to test which group is more successful in their discrimination of the L2 contrasts. The two vowels of each L2 English contrast, /ɛ-æ/ and /ɪ-i/, are close to each other in the acoustic-perceptual space. Moreover, only one member of both contrasts is argued to be represented in either Mandarin or Korean inventory.

### 2.3.2 Korean Vowel Inventory

Although there is disagreement on the exact number of vowels in Korean, with researchers proposing 10, 9, 8, and 7 vowel systems (10 vowel system: Kim, 1992; Lee, 1996; Kang 2003; Yang, 1996, 9 vowel system: Oh, 1997, 8 vowel system: Lee, 1999; Bae, 1996; Kim, 1998, 7 vowel system: Ahn & Iverson, 2007; Kim-Renaud, 2009; Lee & Iverson, 2012; Shin & Cha, 2003; Shin et al., 2012), all agree that /ɪ/ is not phonemic in Korean, while vowels /i/ and /ɛ/ are included in the inventory. For example, the 10 vowel system includes the following vowels: /i, ɨ, e, ɛ, a, o, ø, u, ʌ, y/ (Yang, 1996). The status of /æ/ in Korean is debated, some researchers argue /ɛ/ can be produced with as [æ]-type variants, especially for Koreans who still have the phonemic vowel /e/ (Hong, 1988). The status of non-high-front vowels has changed in Korean recently. Lee & Iverson (2012) argue /e/ still prevails in the formal standard, but that the contrast between non-high-front vowels has been merged, leading to a 7 vowel system with one low vowel, three mid, and three high vowels (Ahn & Iverson, 2007; Kim-Renaud, 2009; Lee & Iverson, 2012). More specifically, Lee & Iverson (2012) argues the Korean vowel inventory is made up of the following vowels, /i, ɛ, ɨ, a, o, u, ʌ/. Due to this merge, the presence of /e/ has been eliminated, and all instances of /ɛ/ are produced in the same way, shifting support away from the idea that /ɛ/ could be produced in a more [æ]-like manner in Korean in order to avoid confusability with [e]. Due to the recent support of this theory of merging in Korean, this study will rely on the 7 vowel system in analysis. Relying on a vowel system with a different number of vowels (e.g. 10, 8) would not significantly change the predictions made in this study, as the vowels that would be added would not affect the mid region significantly, which is the focus of the current experiment. A schematic representation of the Korean vowel system was created to provide a visual interpretation of the vowel system proposed by Lee & Iverson (2012). This visual representation is shown below in Figure 1.

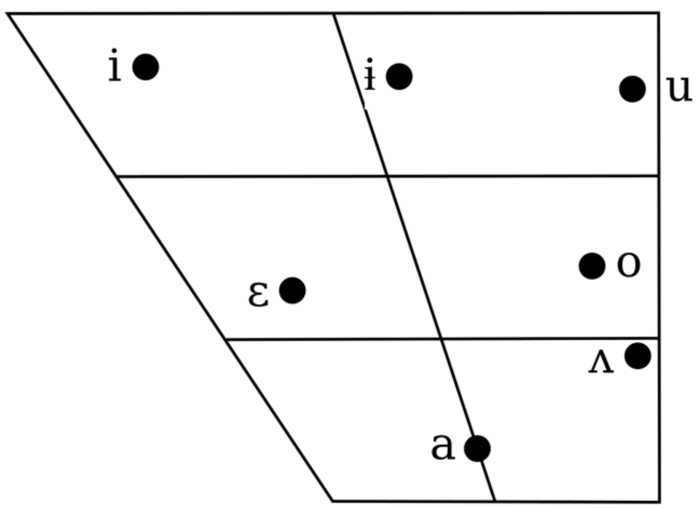


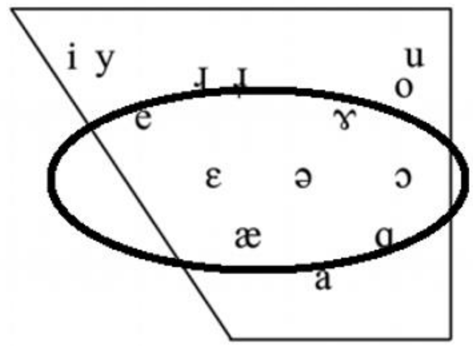

Figure 1. Visualization of the Korean vowel space

### 2.3.3 Mandarin Vowel Inventory

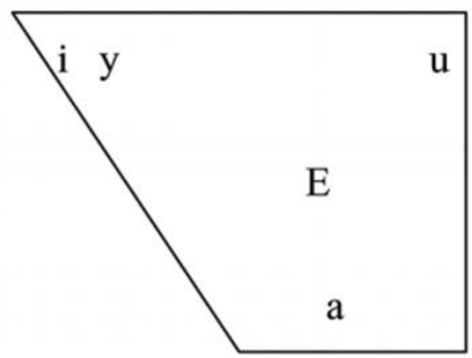
There is also disagreement on the specific vowels present in Standard Mandarin's phonemic inventory. Most scholars agree that there are three high vowels, /i/, /y/, and /u/ (Duanmu, 2000; Liu & Ng, 2009; Lin, 2007; Norman, 1988; Wan & Jaeger, 2003; Zee & Lee, 2001), and one low vowel /a/ (Duanmu, 2000, 2007; Liu & Ng, 2009; Mok, 2012; Zee & Lee, 2001). The situation for the mid vowel(s) is more complicated. There is disagreement on whether the "apical" vowels, which occur after dental and retroflex fricatives and affricates, should be grouped as allophones of /i/, or should be considered separate phonemes. The issue of the two apical vowels, represented as [ɿ] and [ʅ] by (Chen et al., 2019, Wen et al., 2017), and with other symbols by (Chen et al., 2019; Zee & Lee, 2001), will not be discussed in further detail, as the current study will be focusing on the vowel qualities in the mid-range.

#### *2.3.3.1 Allophonic Variation in the Mid-Region*

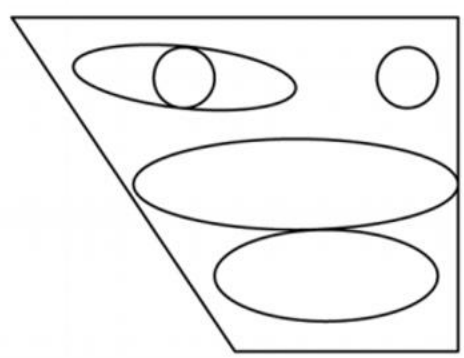
There are many different vowel qualities in the mid-range, including [e, ɛ, ə, ɤ, ɔ, o]. Since the distribution of these vowels is complementary, most sources agree that these are allophones of the same phoneme, but the specific underlying phoneme for these allophones is debated due to the wide variation of the allophones (Mok, 2012). Some researchers have proposed /ɤ/ as the underlying vowel (Cheng, 1973), others /ə/ (Duanmu, 2000; Howie, 1976; Wen et al., 2017; Zee & Lee, 2001). Some have proposed two underlying phonemic vowels, /ɛ/ and /ɔ/ (Howie, 1976; Liu & Ng, 2009). Others have proposed four mid vowel phonemes, specifically /e/, /o/, /ə/, /ɤ/ (Lin, 1991). The idea of an underspecified mid vowel, typically represented by /E/, that changes in frontness/backness and rounding, but not in height, has also been proposed (Wiese, 1997), and supported by others (Mok, 2012). This study will rely on the analysis of Wiese (1997), and consider /E/ to be the underlying phoneme in the mid-range. Choosing a different type of analysis of the mid vowel phonemes present in Mandarin would not change the predictions made in this study, as all allophones would still be present in the mid region, leading to a wide range of vowel qualities regardless of which phonemes are present. A visualization of the Mandarin vowel space and the wide range of the underspecified mid vowel /E/, with all of its allophones circled below, from Mok's (2012) study can be seen in Figure 2 (194). One thing is consistent across different approaches: both /ɪ/ and /æ/ are not phonemic in Mandarin, as in Korean. Similar to Korean, Mandarin also features a high front vowel /i/ and mid front /ɛ/, the latter one at least as an allophone of a highly variable mid vowel phoneme.




Mandarin allophonic vowels




Mandarin phonemic vowels



Mandarin phonemic vowel size

Figure 2. Visualization of the Mandarin vowel space

### 2.3.4 Single-Category Assimilation Prediction

Looking at the surface level, both Korean and Mandarin learners come to the task of acquiring the two English contrasts with similar L1 hindrances. The existence of one, but not another phoneme from each contrast in the L1, and the acoustic proximity of the phonemes within each contrast, suggests that the L2 phoneme without the exact L1 equivalent will be perceived as 'similar' to the closest existing L1 phone (i.e. English /ɪ/ will be perceived as similar to Mandarin and Korean /i/ and English /æ/ will be perceived as similar to Korean /ɛ/ and the underspecified mid vowel in Mandarin), and categorized as such. As a result, the scenario of the *single-category assimilation*, in PAM-L2 terminology, is likely to arise (i.e. both English /ɪ/ and /i/ will be assimilated to Korean/Mandarin /i/), which is predicted to present the greatest challenge for the contrast discrimination by the L2 learners in all three theoretical models. In addition to this challenge, Mandarin has a wide range of allophones in the mid-region of its vowel space, possibly leading to L1 Mandarin speakers facing a greater challenge in acquiring the two English contrasts discussed.

## 2.4 Previous Literature on Difficulties Distinguishing English Vowel Contrasts

It has been shown that L1 Standard Mandarin learners of English have difficulties distinguishing [ɪ] and [æ] from vowels present in their L1 phonemic inventory. Since the Mandarin learners of English were shown to perceive the two L2 phonemes [ɪ] and [æ] as the L1 phonemes [i] and [E] respectively, this indeed would fall under Best & Tyler's (2007) *single-category L2 contrast assimilation.* Flege et al. (1997) examined the production and perception of these two English vowel pairs by Spanish, Mandarin, German and Korean L1 participants with differing levels of L2 experience, broken down into experienced and inexperienced speakers based on length of residence in the United States. The relevant results of this study show that L1 Korean and L1 Mandarin speakers had difficulties in perceiving the differences between the two vowels in each pair **(**/ɛ-æ/**, /ɪ-i/)**. Flege et al. (1997) also demonstrated that Korean and Mandarin participants had difficulty producing [æ] in a target-like manner and speculated that L1 allophony may have hindered the English /æ/ productions by L1 Mandarin and L1 Korean speakers, as '[æ]-quality vowels' occur in some contexts in Mandarin, and can occur as a realization of /ɛ/ in Korean, but does not explore this topic any further. As discussed above, with the more recent merger in Korean, it is unlikely that /ɛ/ is still being realized in an [æ]-like manner, suggesting that allophony likely no longer plays a role in Korean speakers' production and perception of English [æ]. The possibility of allophony effects for the Mandarin speakers in Flege et al.'s (1997) study can be illustrated with the accuracy of the identification of the Mandarin-accented vowels by native speakers of English, for /ɛ/, experienced: 63%, inexperienced: 60%, for /æ/, experienced: 77%, inexperienced: 58%. The results for the /ɛ-æ/ pair were significantly less accurate than those for the /ɪ-i/ pair, for /i/, experienced: 84%, inexperienced: 80%, for /ɪ/, experienced: 90%, inexperienced 83%, which may show that the allophony in the mid-vowels may play a role in the successful perception and production of the /ɛ-æ/ vowel contrast. The difficulty in distinguishing the vowels /ɪ-i/, due to the fact that /ɪ/ is often mis-identified as /i/, was also shown in Tsukada et al. (2005), Hwang & Lee (2015), and Kim et al. (2017) for L1 Korean learners and Jia et al. (2006) and Chen et al. (2001) for L1 Mandarin learners.

Many previous studies (Flege et al., 1997) and theories such as SLM (Flege, 1995) and PAM (Best & Tyler, 2007) have proposed that there will be issues arising from category similarity in L2 phonological acquisition. However, these theories and studies have largely not accounted for the role of allophonic variability in the discriminatory abilities of L2 learners. This study seeks to determine if allophonic variability does indeed play a role, or if category acquisition is shaped by existing phonemic categories alone. To do so, the perceptual discrimination of L2 contrasts by two learner groups with L1s that differ in the presence or absence of allophonic variability in the mid-region of the vowel space, L1 Mandarin and L1 Korean, will be explored. Looking at previous literature, it is clear that both learner groups struggle with the distinguishing English /ɪ/ and /æ/ from vowels present in their L1 phonemic inventories, likely stemming from the single category assimilations as described in PAM and L2LP. However, the results from studies such as Flege et al. (1997) raise the question as to which group is more successful in their perceptions of the two vowels when compared with vowels that are close by in the vowel space, specifically /i/ and /ɛ/ respectively.

## 2.5 Acoustic Profiles of the Vowels in this Study

To begin answering the question of the role of allophonic variability in the discrimination of L2 contrasts, the acoustic profiles of the vowels looked at in this study will be analyzed. The average F1-F2 values, reported by Peterson & Barney (1952) for the vowels tested in this study,

/ɛ-æ/ and /ɪ-i/, in General American English, can be seen in Table 1 below. The F1-F2 values for Korean vowels [i] and [ɛ] are shown in Table 2 (Igeta & Arai, 2011). Standard Mandarin phoneme's [i] average F1-F2 values shown in Table 3 (Zee & Lee, 2001), but as discussed previously the situation for [ɛ] is more complicated. A visualization of these vowels for the three languages for male and female speakers is given in Figure 3 (male) and 4 (female). These numbers demonstrate that the vowels within each pair are acoustically proximal to each other across the three languages, supporting the possibility of perceptual confusion by L2 learners given the lack of the relevant phonemic distinction in their L1s.

Table 1. Average American English F1-F2 Values

| Vowel | F1 | | F2 | |
|---|---|---|---|---|
| | Male | Female | Male | Female |
| /i/ | 270 | 310 | 2290 | 2790 |
| /ɪ/ | 390 | 430 | 1990 | 2480 |
| /ɛ/ | 530 | 610 | 1840 | 2330 |
| /æ/ | 660 | 860 | 1720 | 2050 |

Table 2. Average Korean F1-F2 Values

| Vowel | F1 | | F2 | |
|---|---|---|---|---|
| | Male | Female | Male | Female |
| /i/ | 217 | 238 | 2318 | 2980 |
| /ɛ/ | 484 | 663 | 2067 | 2517 |

Table 3. Average Standard Mandarin F1-F2 Values

| Vowel | F1 | | F2 | |
|---|---|---|---|---|
| | Male | Female | Male | Female |
| /i/ | 300 | 401 | 2443 | 3036 |

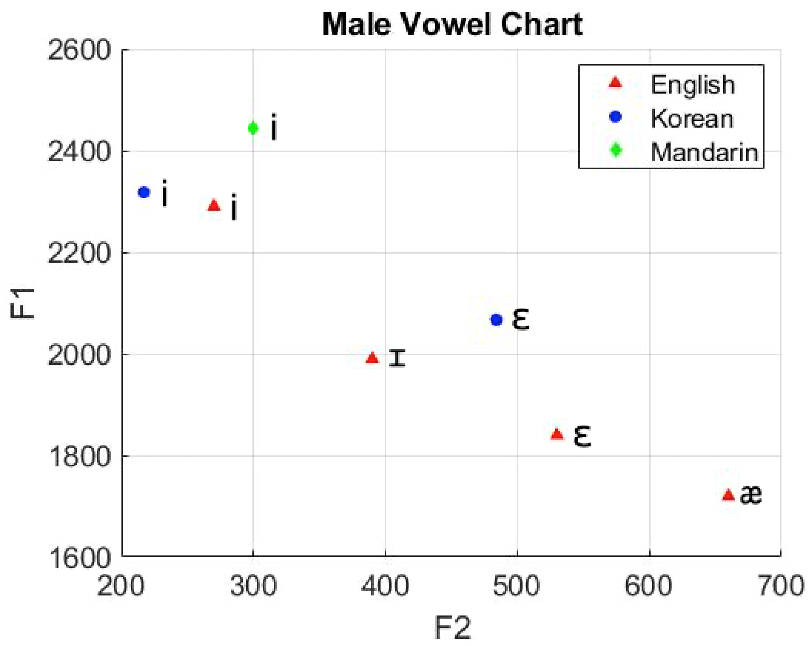



Figure 3. Vowel plot for male speakers

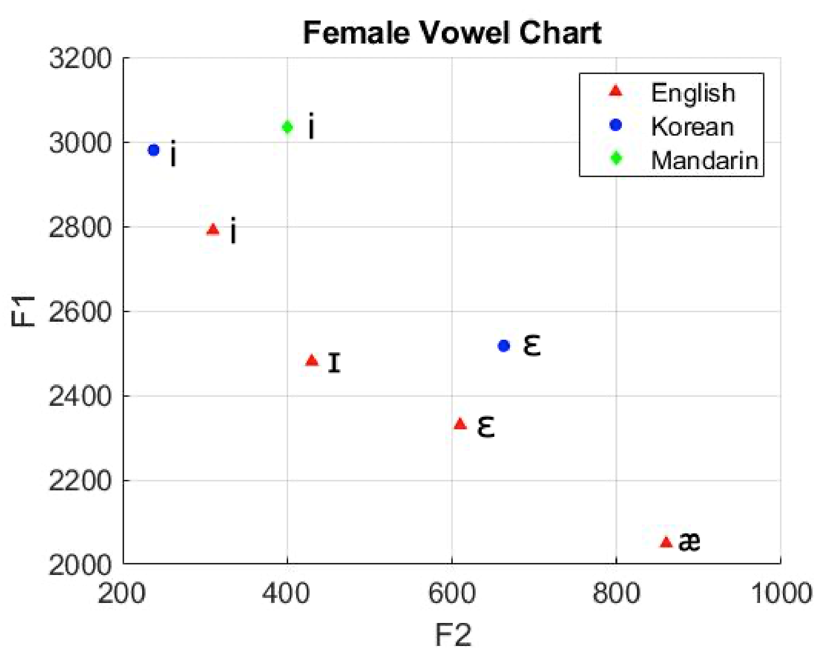



Figure 4. Vowel plot for female speakers

## 2.6 The Current Study

Unlike Korean, Mandarin presents a great allophonic variability in the mid-region of the vowel space. The models of L2 acquisition discussed above do not address the issue of allophonic variability as a possible factor in the acquisition of L2 contrasts. Nevertheless, it is possible that allophonic variability may play a role. This exploratory study compares the discrimination accuracy of the two English vowel contrasts by Mandarin and Korean learners in order to begin addressing the question. Based on the previous literature, the following hypothesis was proposed: the two groups of leaners will be comparable in their discrimination of the /ɪ-i/ contrast but will diverge in their discrimination of the /ɛ-æ/ contrast. Specifically, we predict that Mandarin learners will be less successful than their Korean counterparts in the discrimination of the /ɛ-æ/ contrast due to the high variability of the Mandarin mid vowel. The basis for this prediction is the assumption that such high allophonic variability will condition Mandarin listeners to be even more likely to accept a non-native, acoustically proximal, realization as a deviant but possible version of the L1 mid-vowel phoneme.  Looking at this issue under the lens of popular theoretical models, this conditioning will solidify the tendency for the equivalence classification between English /æ/ and Mandarin mid phoneme, and result in the robust single category assimilation, described in both PAM and L2LP, of English /ɛ-æ/ to Mandarin mid vowel. Korean speakers, on the other hand, are used to operate in a more well-defined, narrower space for their mid phoneme /ɛ/, and therefore may be less likely to accept English /æ/ as a version of L1 /ɛ/. One caveat to this prediction is that Mandarin mid-vowel variability appears to be largely limited to the horizontal dimension, which may mean that despite all this variability /æ/ would not be readily acceptable since it deviates in terms of height, not only backness.

# CHAPTER 3. METHODS

## 3.1 Participants

In this experiment, 16 L1 Standard Mandarin speakers from Mainland China (10 female, 6 male), and 14 L1 Korean speakers from South Korea (4 female, 10 male) participated in the discrimination task. All participants were born in their respective home countries but were living in the United States at the time of data collection. The mean length of time spent living in the United States was 14.57 years for the L1 Korean speakers and 12.38 years for the L1 Mandarin speakers. The Korean participants had an average age of 25.36 years (*MIN* = 18, *MAX* = 37, *SD* = 5.47), and Mandarin participants had an average age of 25.31 years (*MIN* = 18, *MAX* = 37, *SD* = 5.97). A two-tailed t-test was performed to verify that age was not statistically significantly different between the two groups. There was no significant difference, $t(28) = 0.021$, $p = .98$, between the ages of the L1 Mandarin group and the ages of the L1 Korean group. The L1 Korean group had an average of 6.23 out of 7 for their self-rated English proficiency levels, while the L1 Mandarin group had an average of 5.81 out of 7 for their self-rated proficiency levels. The age when the participants started learning English in their home countries (age of acquisition or AOA) was also collected. The L1 Mandarin group had an average AOA of 8.56 years (*MIN* = 1, *MAX* = 15, *SD* = 3.95), while the L1 Korean group had an average AOA of 7.71 years (*MIN* = 5, *MAX* = 10, *SD* = 1.39). The age when the participants arrived in the United States was also collected. The L1 Mandarin group had an average age of arrival of 13.69 years (*MIN* = 4, *MAX* = 30, *SD* = 8.31), while the L1 Korean group had an average age of arrival of 10.43 years (*MIN* = 6, *MAX* = 25, *SD* = 4.89).

Out of the 30 participants, 15 learned a foreign language other than English, 12 out of 15 participants learned the foreign language in middle or high school, none of these 12 participants spent more than 5 years learning the additional language in school. Among the remaining 3 participants, who were part of the L1 Mandarin group, 2 spoke Cantonese and 1 spoke Hokkien (Taiwanese) in addition to standard Mandarin. These 3 participants spoke their respective additional languages at home, but indicated standard Mandarin as their first language. The average proficiency scores were also collected from participants for these additional languages on a scale from 1 to 7, with an average score of 3.60 for L1 Korean participants and an average score of 2.78

for L1 Mandarin participants. The specific languages and how they were learned are detailed below in Table 4 for L1 Mandarin participants, and Table 5 for L1 Korean participants. One participant learned 2 languages in school, leading to 16 total participants in the tables instead of 15.

Table 4. Foreign Languages Learned by L1 Mandarin Participants

| Number of participants | Language learned | How language was learned |
|---|---|---|
| 3 | Spanish | 2: school, 1: school + work |
| 2 | French | 2: school |
| 3 | Japanese | 1: school, 1: television, 1: unknown[a] |
| 2 | Cantonese | 2: home |
| 1 | Hokkien | 1: unknown[a] |

[a]2 participants incorrectly answered the question of how they learned their foreign language.

Table 5. Foreign Languages Learned by L1 Korean Participants

| Number of participants | Language learned | How language was learned |
|---|---|---|
| 4 | Spanish | school |
| 1 | Mandarin | school |

## 3.2 Stimuli

The stimuli consisted of the following vowel pairs, /ɛ-æ/ and /ɪ-i/, inserted in monosyllabic minimal pairs with a CVC structure, for example *bed.* There were 10 minimal pairs total, made up of 5 pairs for the /ɛ-æ/ distinction and 5 pairs for the /ɪ-i/ distinction, shown below in Table 6. A female native speaker of Midwestern American English recorded the English stimuli for the perceptual experiment (age 23). For the recording session, the words were read in a list, with 3 repetitions for each word, from which the best recordings were chosen for the experiment.

Table 6. Stimuli

| Pair | Frequency |
|---|---|
| /bɛd-bæd/ | 114848, 286811 |
| /dɛd-dæd/ | 178699, 144072 |
| /pɛt-pæt/ | 19434, 28491 |
| /gɛs-gæs/ | 157624, 86977 |
| /dɛb-dæb/ | 3035, 1487 |
| /sɪt-sit/ | 108786, 67899 |
| /sɪk-sik/ | 66276, 44440 |
| /pɪtʃ-pitʃ/ | 23255, 5148 |
| /pɪp-pip/ | 1614, 1740 |
| /tʃɪp-tʃip/ | 18796, 33782 |

For each of the ’different’ pairs, e.g., /bɛd- bæd/ and /bæd-bɛd/, the phonetic context surrounding the vowel was acoustically identical, with different vowels spliced in depending on the word. The vowels were extracted from the original recording starting from the end of the release portion of the initial consonant and ending at the offset of periodicity and formant structure of the vowel. For the ’same’ pairs, e.g., /bɛd-bɛd/ and /bæd-bæd/, the phonetic context surrounding the vowels was also acoustically identical but the two vowels in each pair came from two different recordings of the same word. All the stimuli were real English words, with the exception of Deb, which is a popular female name. The words had an average frequency of 69660.7 (min: 1487, max: 178699), with frequencies taken from the Corpus of Contemporary American English (COCA), which contains 950 million words taken from various spoken and written texts (Davies, 2008-). Two-tailed t-tests were performed for each vowel contrast to verify that the word frequency between the word pairs was not statistically significant. There was no significant difference, $t(8) = 0.24$, $p = .82$, between the frequency for the words containing /ɛ/ ($M = 94728$, $SD = 79825.81$) and the words containing /æ/ ($M = 109567.60$, $SD = 113325.79$). There was also no significant difference, $t(8) = 0.57$, $p = .58$, between the frequency for the words containing /ɪ/ ($M = 43745.40$, $SD = 43460.59$) and the words containing /i/ ($M = 30601.80$, $SD = 27719.98$).

### 3.3 Recruitment & Procedures

Participants were recruited using Prolific, an online recruitment platform. In Prolific, the following constraints were used to recruit participants, location: United States, native language (labeled first language in the Prolific platform): Korean or Mandarin (depending on group), age: 18-40 years. After meeting the specific requirements, participants were redirected to a Purdue Qualtrics survey. Participants first had to answer background information questions (a full list of these questions from the online survey can be found in the Appendix), then a separate block of perception questions. Since the recruitment process was done online, in the first block of the survey, there were 5 questions written in the participant's respective first language to confirm they were indeed native speakers of that language, since an in-person assessment could not be done. These questions were simple fill in the blank questions that were created by native speakers of each respective language. The rest of the questions in the background information section were chosen to capture the specific language backgrounds of each participant, and included questions about English proficiency, more specifically a self-rating scale, other languages spoken and proficiency in those languages, and other useful information such as the languages the participants' parents spoke. As stated above, one question asked about the age participants started learning English, or the age of acquisition (AOA) of English. In previous research it has been shown that the later an L2 learner is exposed to the L2, the stronger the foreign accent tends to be (Abbrahamsson & Hyltenstam, 2009; Flege et al., 2006), so it was important to capture the age of acquisition of English for participants as a possible factor that could influence the perception task results. Age of arrival to the United States was also collected for the participants, as it has also been found to be a factor that can impact second language acquisition (Oyama, 1976), and therefore impact the perception task results. A self-rated proficiency question was chosen to avoid the need for a lengthy proficiency test. Although there is some controversy on the validity of a self-rating, it has been shown to be a fairly accurate measure of the actual proficiency in ESL speakers (Wilson, 1999).

### 3.4 Perception Task

All participants performed an AX discrimination task directly following the language background question portion of the survey. The stimuli in the perception task were randomized for presentation for every participant. For each trial, participants heard an audio clip containing two words consecutively, with a 200 ms gap between the words in a pair, and were then asked to indicate whether ‘the two vowels in the words’ they heard were different or the same. The 200 ms gap was chosen based on similar AX discrimination task studies (Seo et al., 2022). The response was not timed, but participants were unable to replay the audio. Both ‘same’ (two words with the same vowel) and ‘different’ pairs (two words with different vowels) were included as stimuli, in equal numbers. Different vowels were also presented in both orders. For example, for the vowel pair /ɛ-æ/ the stimuli included /ɛ-æ/, /æ-ɛ/, /ɛ-ɛ/, and /æ-æ/. Including ‘same’ pairs of this type allowed to test if participants were truly discriminating the vowel categories, or if they were marking them as different simply because they were different sound files. Each stimulus pair was presented to each participant four times. Therefore, each participant completed 40 discrimination trials. A total of 1200 tokens were included in the final analysis (30 participants x 10 word pairs x 4 vowel orientations (e.g. /æ-ɛ/, /ɛ-ɛ/)). Participants were compensated 5 USD per 30 minutes through the Prolific platform.

# CHAPTER 4. RESULTS

## 4.1 Data Analysis

The participants' answers to the perception questions were marked as either correct or incorrect and inputted in MATLAB R2021b for statistical analysis using the 'statistics and machine learning toolbox' (The MathWorks, 2019).

## 4.2 Main Mixed Effects Model

A logistic regression mixed-effects model was chosen to analyze the data based on previous studies with similar perceptual discrimination tasks (Ryu, 2018; Seo et al., 2022). In this model perceptual accuracy, coded as '1' (correct response) or '0' (incorrect response), as a binary categorical dependent variable. This model included group (L1 Mandarin or L1 Korean), trial type (same or different), and contrast type (/i-ɪ/ or /ɛ-æ/), the interaction between group and contrast type, and the interaction between group and trial type as fixed effects. It also included item and subject as random intercepts. The results of this mixed effects model are shown below in Table 7.

Table 7. Results of Mixed-Effects Model 1

| | Estimate | SE | *t*-value | *p*-value | Odd ratio |
|---|---|---|---|---|---|
| Intercept | 2.37 | 0.34 | 6.91 | 7.65e-12 | 10.71 |
| Group-Mandarin | -1.04 | 0.40 | -2.58 | 0.0099 | 0.35 |
| Contrast type-/i-ɪ/ | -0.29 | 0.37 | -0.78 | 0.44 | 0.75 |
| Trial type-different | 1.31 | 0.42 | 3.16 | 0.0016 | 3.71 |
| Group-Mandarin: Contrast type-/i-ɪ/ | 0.97 | 0.44 | 2.20 | 0.028 | 2.63 |
| Group-Mandarin: Trial type- different | 0.24 | 0.51 | 0.48 | 0.63 | 1.27 |

### 4.2.1 Group Effects

The results of this model showed that the L1 Mandarin group was significantly different from the baseline group (L1 Korean) ($\beta = -1.04$, $SE = 0.40$, $p < 0.001$) in terms of the accuracy of their perceptual discrimination answers. This demonstrates that, although both groups were largely accurate in their discrimination responses, the L1 Mandarin group was significantly less accurate than the L1 Korean group, as illustrated in Figure 5.

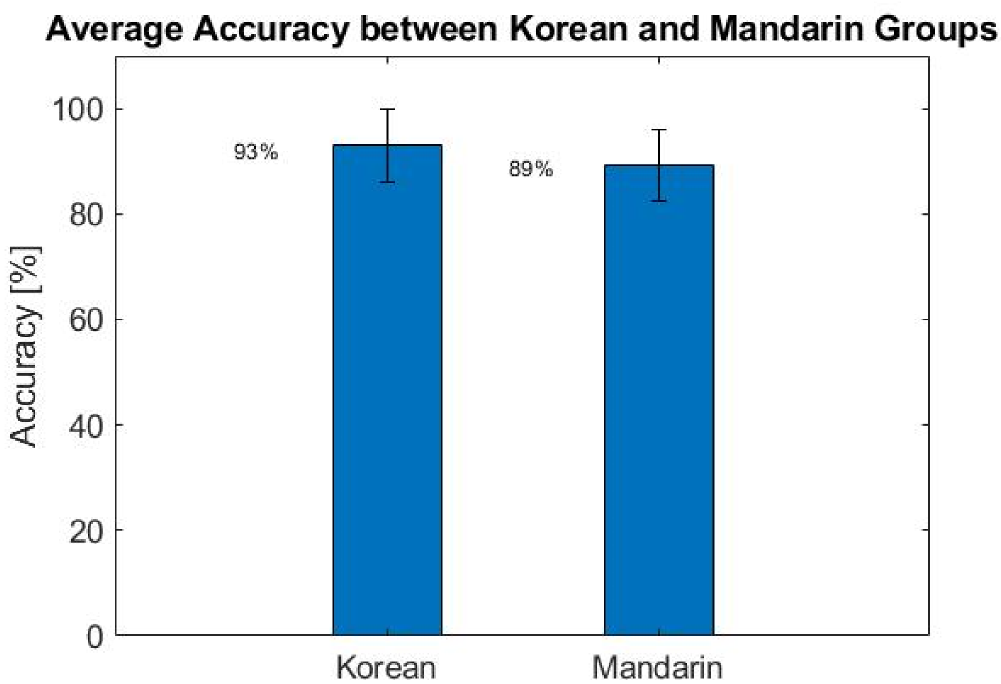



Figure 5. Average perceptual accuracy for speaker groups

*Note.* Error bars represent +/- 1 SD

### 4.2.2 Contrast Type Effects

The effect of contrast type (/i-ɪ/ or /ɛ-æ/) was not found to significantly affect the discriminatory accuracy of the participants ($p = 0.44$), when adjusted for group and speaker. However, the effect of the interaction between group (L1 Mandarin or L1 Korean) and contrast type (/i-ɪ/ or /ɛ-æ/) was found to significantly affect the discriminatory accuracy in the questions involving the /i-ɪ/ pair for the L1 Mandarin group by increasing the odds of an accurate response by a factor of 2.63 ($\beta = 0.97$, $SE = 0.44$, $p = 0.028$) compared to the baseline group (L1 Korean).

The significance of the interaction between group and contrast type is that, while the contrast type did not affect the L1 Korean participants in their judgement accuracies, it increased the judgement accuracies of the L1 Mandarin participants when the contrast type was /i-ɪ/ when compared to the baseline contrast /ɛ-æ/. This also indicates that the main effect of Group is driven primarily by the /ɛ-æ/) contrast, where Korean and Mandarin participants performed differently. Figure 6 below shows the average scores of the L1 Mandarin and L1 Korean groups for the two different contrast types.

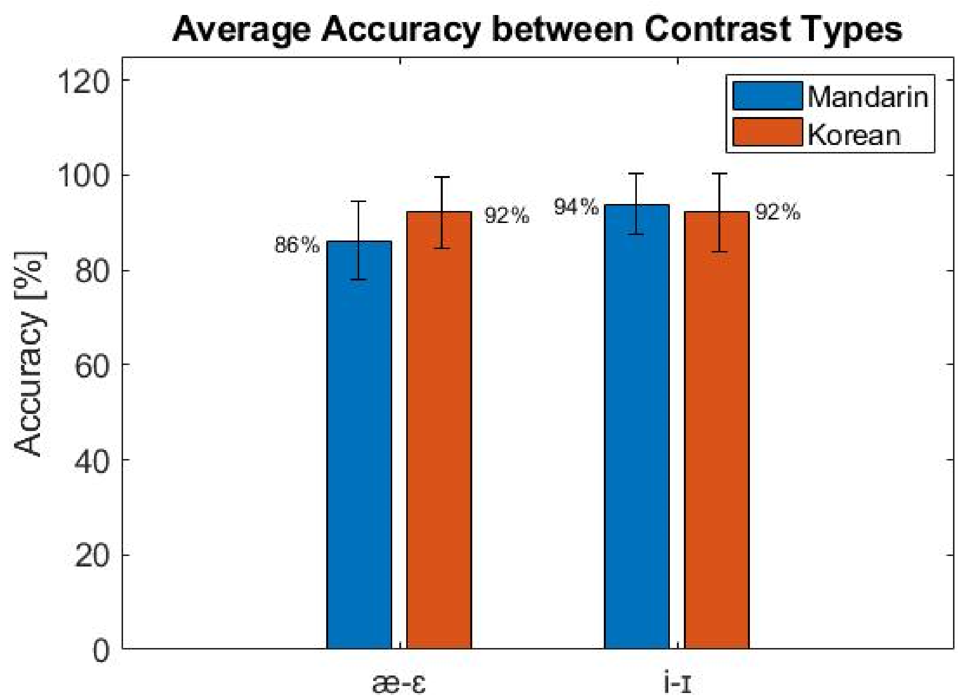



Figure 6. Group accuracies for contrast types

*Note.* Error bars represent +/- 1 SD

### 4.2.3 Follow-Up Models

Two separate follow-up logistic regression mixed-effects models were implemented, one within each vowel contrast type. Both used perceptual accuracy, coded as ‘1’ (correct response) or ‘0’ (incorrect response), as a binary categorical dependent variable and Group (L1 Mandarin or L1 Korean) as the only fixed effect. They also included item and subject as random intercepts. The results, reported in Table 8 and 9, indicate that the effect of Group was significant in the /ɛ-æ/ model only, confirming that Mandarin participants were significantly outperformed by Korean participants in the discrimination of the /ɛ-æ/ contrast, but not in the discrimination of the /i-ɪ/ contrast.

Table 8. Effect of Group in the /i-ɪ/ Model

| | Estimate | SE | *t*-value | *p*-value | Odd ratio |
|---|---|---|---|---|---|
| Intercept | 2.69 | 0.45 | 5.93 | 5.26e-09 | 14.73 |
| Group-Mandarin | -0.0038 | 0.40 | -0.0095 | 0.99 | 0.997 |

Table 9. Effect of Group in the /ɛ-æ/ Model

| | Estimate | SE | *t*-value | *p*-value | Odd ratio |
|---|---|---|---|---|---|
| Intercept | 3.00 | 0.53 | 5.65 | 2.52e-08 | 20.09 |
| Group-Mandarin | -0.93 | 0.34 | -2.78 | 0.0056 | 0.40 |

### 4.2.4 Trial Type Effects

Looking at main statistical model, the effect of trial type (same or different) was found to significantly affect the discriminatory accuracy of the participants by increasing the odds of an accurate response for the 'different' word questions by a factor of 3.71 ($\beta = 1.31$, $SE = 0.42$, $p = 0.0016$), when adjusted for group and speaker. The fixed effect of the interaction between group and type was not found to significantly affect the discriminatory accuracy in the 'different' word questions for the L1 Mandarin group when compared to the baseline group (L1 Korean) ($p = 0.63$). These results indicate that the two participant groups showed higher accuracy when the perception question contained 'different' words, possibly due to a 'different' bias in responses, meaning that participants tended to say 'different' more often, even for 'same' trials. Figure 7 below shows the average scores of the two participant groups for the 'same' and 'different' trials.

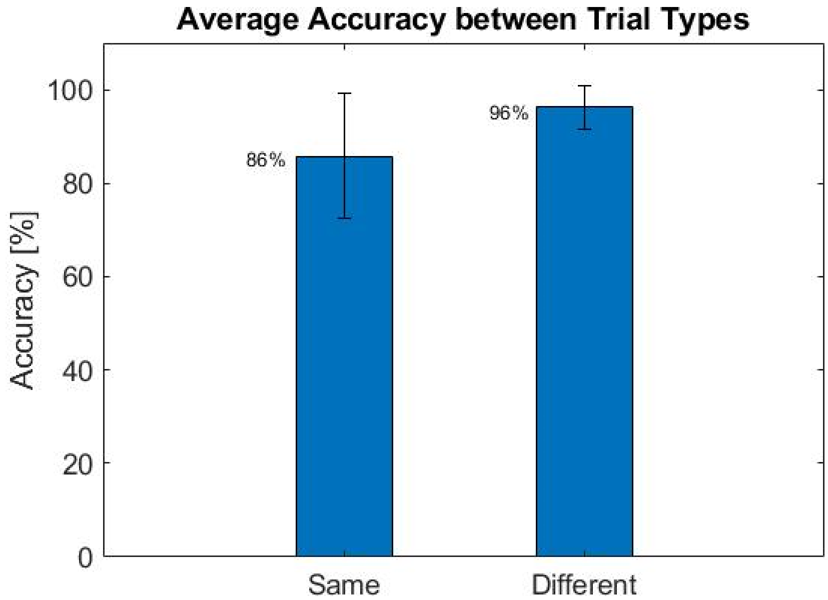



Figure 7. Average accuracy for groups by trial

*Note.* Error bars represent +/- 1 SD

### 4.3 Mixed Effects Model to Test Attributes

A fourth logistic regression mixed-effects model was implemented to see if certain attributes of participants' language backgrounds were predictive of their discrimination accuracy. Perceptual accuracy, coded as '1' (correct) or '0' (incorrect), was used again as a binary categorical dependent variable. The fixed effects in this model were, group (L1 Mandarin or L1 Korean), age of acquisition (AOA) of English (in years), age of arrival to the United States (in years), and self-rated proficiency (on a scale from 1-7), with item and subject as random intercepts. The results of this model are shown below in Table 10.

Table 10. Results of Mixed-Effects Model 4

| | Estimate | SE | *t*-value | *p*-value | Odd ratio |
|---|---|---|---|---|---|
| Intercept | -1.65 | 1.16 | -1.42 | 0.155 | 0.52 |
| Group-Mandarin | 0.23 | 0.33 | 0.69 | 0.49 | 1.26 |
| AOA | -0.044 | -0.056 | -0.79 | 0.429 | 0.96 |
| Age of arrival | 0.044 | 0.023 | 1.89 | 0.059 | 1.05 |
| proficiency | -0.21 | 0.15 | -1.33 | 0.182 | 0.82 |

The results indicate that AOA, age of arrival, and self-rated proficiency did not significantly affect the discriminatory accuracy of the two speaker groups.

# CHAPTER 5. DISCUSSION & CONCLUSION

## 5.1 Revisiting the Hypothesis

The goal of this study was to test if allophonic variability could be a factor in the acquisition of L2 phonemes through an AX perceptual discrimination task looking at two vowel contrasts, /ɪ-i/ and /ɛ-æ/ by two L1 speaker groups, L1 Mandarin and L1 Korean. Both pairs have been shown to be difficult for L1 Mandarin and L1 Korean speakers in previous research (Chen et al., 2001; Flege et al., 1997; Hwang & Lee, 2015; Jia et al., 2006; Kim et al., 2017; Tsukada et al., 2005). Since the vowel inventories of both Korean and Mandarin contain the /i/ phoneme but lack the /ɪ/ phoneme, it was predicted that the two groups would perform in a similar manner in their discriminations of the /ɪ-i/ pair. However, due to the significant allophonic variability in the mid-region of the Standard Mandarin vowel space, it was predicted that the Mandarin learners would be less accurate in the discrimination of the /ɛ-æ/ contrast than Korean learners.

## 5.2 Discussion of the Perception Results

The results of this study supported the predictions made about the discriminatory accuracies of the two vowel contrasts. Overall, the L1 Mandarin group participants had lower discriminatory accuracy when compared to the L1 Korean group participants, and the significant interaction between Group and Vowel contrast revealed that it is the differential performance of the two groups on the /ɛ-æ/ contrast, where Mandarin participants were less accurate than Korean ones, that resulted in the significant main effect. It was expected that the two groups would perform similarly for the /ɪ-i/ contrast, and they did. In fact, Mandarin learners were slightly, but not significantly, more accurate for this contrast on average. Thus, the prediction that Mandarin learners would have more difficulty discriminating /ɛ/ and /æ/ than Korean participants was supported. The fact that the two groups of participants were very similar to each other in terms of the average self-reported English proficiency minimizes the possibility that the difference in performance on the /ɛ-æ/ contrast could be due to the uneven English proficiency levels across the two groups. Moreover, a very comparable performance on the /ɪ-i/ contrast further supports the assumption that the two groups were similar in terms of English proficiency. Furthermore, neither AOA nor English proficiency were predictive of participants' performance on the discrimination task, across groups. Therefore, the difference between Mandarin and Korean participants in discriminating /ɛ-æ/ contrast should be attributed to a factor other than English proficiency.

However, it is still important to discuss the fact that the participants overall rated themselves as having high proficiency in English. Most of the participants in this study rated themselves a 6 or 7 on a 1-7 self-rating scale for English proficiency, with 1 being 'very poor' and 7 being 'native-like'. Only six participants rated themselves less than a score of 6, and of those six only two participants rated themselves less than 5, 'good' proficiency. The high self-reported proficiency of the participants, as well as their relatively long duration of residency in the English-speaking country, may explain on the overall high accuracy in discriminating between the two vowels in the /ɪ-i/ pair and the /ɛ-æ/ pair. Age of arrival was another factor that could have potentially impacted the perception results, however, age of arrival was also not predictive of participants' performance on the discrimination task. It is possible that age of arrival did not play a role in the perception task results because, the tuning of phonetic categories occurs in infancy (Kuhl, 2005), which may suggest that these participants, even those who arrived in the United States at a young age, were already fine-tuned to the phonetic categories of their L1s, that the effects of age of arrival in the United States were negligible.

The outcome of this study is compatible with the hypothesis that Mandarin speakers of English experience more difficulty with the /ɛ-æ/ contrast because of the high degree of allophonic variation present in Mandarin in the mid-vowel region. As a result, many different phonetic realizations in the mid region of the vowel space could be potentially acceptable for Mandarin listeners as variants of the same, highly variable, native vowel phoneme, including the non-native English /ɛ/ and /æ/. Korean, on the other hand, although also without /æ/ as a separate vowel phoneme, has a more restricted space of variability for its native /ɛ/. Therefore, a non-native /æ/ is less likely to be accepted by Korean learners as a possible realization of the native /ɛ/. In other words, using Flege's 'similar' and 'different' dichotomy, English /æ/ is less 'similar' to native vowels for Korean learners than for Mandarin learners. As a consequence, equivalence classification between a similar native vowel and English /æ/ is more likely for Mandarin learners, while separate category formation (and therefore, more successful discrimination) for English /æ/ is more likely for Korean learners.

### 5.2.1 On Allophonic Variation

These results demonstrate that allophonic variation in native vowels can play a role in determining the acquisition trajectory for the non-native vowel categories. Therefore, the study emphasizes the importance of considering the full complexity of the native vowel inventory, including the type and number of allophonic realizations for each phonemic category, in making predictions about the relative difficulty of acquiring non-native vowels. Flege's original formulation of the SLM specified that the acoustic-perceptual similarities between the sounds of L1 and L2 need to be established on the basis of specific positional allophones. Nevertheless, much of the subsequent work on the acquisition of L2 vowels overlooked the detailed allophonic variation and instead compared languages based on the presence of certain vowel qualities as independent phonemes in the L1 and the L2 inventories (Bohn & Flege, 1992; Chen et al., 2001; Morrison, 2008). This, admittedly more straightforward approach, was probably sufficient to generate relatively accurate predictions for many L1-L2 language pairings. However, the present study suggests that in cases of especially variable realization of native vowels, such as the case of mid vowels in Mandarin, taking allophonic variability into account can help fine-tune our expectation concerning the acquisition of certain L2 vowels.

## 5.3 Conclusion

The discriminatory accuracy of two English vowel pairs, /ɛ-æ/ and /ɪ-i/, by two different L1 participant groups, L1 Mandarin and L1 Korean groups, were tested to see which group would be more successful in their discrimination of the L2 contrasts that have been proven to be difficult for these particular groups (Chen et al., 2001; Flege et al., 1997; Hwang & Lee, 2015; Jia et al., 2006; Kim et al., 2017; Tsukada et al., 2005). It was predicted that the two participant groups would perform similarly for the /ɪ-i/ contrast but would differ in the performances for the /ɛ-æ/ contrast due to the large allophonic variability in the mid-region of the vowel space in Mandarin. This hypothesis was supported, with the two participant groups performing similarly for the questions containing the /ɪ-i/ pair. As expected, the Mandarin group was less accurate in the /ɛ-æ/ pair questions than the Korean group, confirming the possibility that L1 allophonic variation could have affected the acquisition of L2 vowel contrasts.

### 5.3.1 Limitations to this Study and Areas for Further Research

The results of this study indicate that the two participant groups showed higher accuracy on ‘different’ trials. This suggests a possible bias of participants to respond ‘different’, independently of the trial type. It is unclear what is causing this bias, however, it may be due to the way the instructions for the perception task were worded. Participants were asked to indicate whether the two vowels they heard in the words were the same or different for each trial. It is possible that participants misinterpreted the question and responded ‘different’ even on the ‘same’ trials, because on those trials two different recordings of the same vowel were used, which meant that the two vowels were acoustically non-identical, although they did belong to the same vowel category. In future work, more careful attention needs to be paid to wording the instructions and ensuring that participants understood the intended meaning.

Some participant recruitment requirements had to be relaxed in Prolific in order to reach the goal for the number of participants in this study, for example, the constraint on foreign languages learned (excluding English) was relaxed, creating the issue of having a few of the L1 Mandarin participants have other languages spoken at home. Also, the constraint on age had to be relaxed, leading to a wide age range of participants. Another issue in this study was the lack of detailed information about some of the other languages the participants spoke (outside of their first language and English). In particular, 3 of the L1 Mandarin participants indicated that they spoke another Chinese language at home (Cantonese or Hokkien). Although two out of three of these participants indicated a low proficiency of these languages, it would be useful to know what percentage of time this language was spoken at home compared to standard Mandarin. In future research, recruitment and data collection should be done in-person, to both widen the participant pool so the necessary participant requirements can be met, and to control for the time spent on each test item. In addition, other task types, such as an AXB task, should be used to corroborate the findings of the current study. Finally, studies looking at other languages that have high allophonic variability in a specific region, and how the speakers of these languages acquire new L2 phonemes, would be an important contribution to this underexplored area.

# APPENDIX: LANGUAGE BACKGROUND QUESTIONS

Q1. What is your biological sex (select one)?

A. Man

B. Woman

Q2. How old are you in years?

Q3. Have you ever had a vision problem, hearing impairment, language disability, or learning disability?

A. Yes (please explain)

B. No

Q4. Where were you born?

-Please, provide more information and list the countries and US states lived in, and the duration of residence

| | | |
|---|---|---|
| Country, City and/or US State: | From (age): | To (age): |
| Country, City and/or US State: | From (age): | To (age): |
| Country, City and/or US State: | From (age): | To (age): |
| Country, City and/or US State: | From (age): | To (age): |
| Country, City and/or US State: | From (age): | To (age): |
| Country, City and/or US State: | From (age): | To (age): |

Q5. Are you a native speaker of Korean or Standard Mandarin?

A. Korean

B. Standard Mandarin

Please fill in the blank for the following questions (depending on the answer to the previous question)

B. Standard Mandarin

Q6. “野火烧不尽，___吹又生”

Q7. _____，曲项向天歌。白毛浮绿水，红掌拨清波。

Q8. 床前明__光，疑是地上霜。

Q9. 谁知盘中餐，粒粒皆_____。

A. Korean

Q6. 개울가에 올챙이 _____ 꼬물꼬물 헤엄치다.

Q7. 동해물과 _____ 마르고 닳도록.

Q8. 텔레비전에 _____ 나왔으면 정말 좋겠네.

Q9. 아빠 힘내세요, _____ 있잖아요.

Q10. What age were you when you started learning English?

Q11. How many years have you been learning English?

Q12. Rate your estimated English proficiency using the scale below.

1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Q13. Think of the adults who raised you. What was the first language for all of them?

Relationship to you: Language:

Relationship to you: Language:

Relationship to you: Language:

Relationship to you: Language:

Q14. Do you speak any other languages?

A. Yes

B. No

If yes:

Q15. How did you learn this/these languages (this is excluding English and your native language)?

Q16. For the other language(s) you speak (this is excluding English and your native language), please indicate each language's name and how many years you have learned this language.

Language: # years learned:

Language: # years learned:

Language: # years learned:

Language: # years learned:

Language: # years learned:

Q17. For the other language(s) you speak this is excluding English and your native language), rate your estimated proficiency using the scale below.

Language: 1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Language: 1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Language: 1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Language: 1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Language: 1. Very poor, 2. Poor, 3. Fair, 4. Functional, 5. Good, 6. Very good, 7. Native-like

Q18. Do you have any musical or vocal training?

A. Yes

B. No

If yes:

Q19. Please provide more information.

## REFERENCES

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*(2), 249–306. https://doi.org/10.1111/j.1467-9922.2009.00507.x

Ahn, S., & Iverson, G. (2007). Structured imbalances in the emergence of the Korean vowel system. In J. C. Salmons & S. Dubenion-Smith (Eds.), *Historical linguistics 2005*, (pp. 275–293). Philadelphia, PA: John Benjamins.

Aoyama, K., Flege, J. E., Guion-Anderson, S., Akahane-Yamada, R.R., & Yamada, T. (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. https://doi.org/10.1016/S0095-4470(03)00036-6

Bae, J. (1996). *Kwuke umwunlon kaysel* [Introduction of Korean Phonology]. Seoul: Sinkwumunhwasa

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, (pp. 171–204). Timonium, MD: York Press, Inc.

Best, C.T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.17.07bes

Bohn, O.-S. (1995). What determines the perceptual difficulty encountered in the acquisition of nonnative contrasts? In *13th International Conference of Phonetic Sciences*, Stockholm.

Bohn, O.-S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition, 14*(2), 131–158. https://doi.org/10.1017/S0272263100010792

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*(6), 427–440. https://doi.org/10.1080/02699200110044804

Chen, Y., Zhang, J., Sieg, J., & Chen, Y. (2019). Is [ɤ] in Mandarin a transitional vowel? - Evidence from tongue movement by ultrasound imagining. *Journal of Chinese Linguistics, 47*(2), 371–405. https://doi.org/10.1353/jcl.2019.0016

Cheng, C. C. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.

Davies, M. (2008–) *The corpus of contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Duanmu, S. (2000). *The phonology of standard Chinese*. Oxford: Oxford University Press.

Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford: Oxford University Press.

Elvin, J., Escudero, P., & Vasiliev, P. (2014). Spanish is better than English for discriminating Portuguese vowels: acoustic similarity versus vowel inventory size. *Frontiers in Psychology*, *5*, 1188. https://doi.org/10.3389/fpsyg.2014.01188

Escudero, P. (2000). *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers*. [M.Sc. dissertation, University of Edinburgh]. Retrieved from https://www.fon.hum.uva.nl/paul/p2/papers/escuderoMScThesis.pdf

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. [Doctoral dissertation, Utrecht University]. Retrieved from https://www.lotpublications.nl/Documents/113_fulltext.pdf

Escudero, P. (2009). Linguistic perception of "similar" L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 151–190). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/978311021934.151

Escudero, P., & Boersma, P. (2002). The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 208–219). Somerville, MA: Cascadilla.

Flege, J.E. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–65. https://doi.org/10.1016/S0095-4470(19)30537-6

Flege, J. E. (1988). The production and perception of speech sounds in a foreign language. In H. Winitz (Ed.), *Human communication and its disorders, a review* (pp. 224–240). Norwood, NJ: Ablex.

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn & C. Stoel-Gammon (Eds.), *Phonological development: models, research, and application* (pp. 565-604). Timonium, MD: York Press.

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* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E. (1999). The relation between L2 production and perception. In J. Ohala, Y. Hasegawa, D. Granveille & A. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetics Sciences* (pp. 1273–1276). San Francisco, California.

Flege, J. E., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics, 34*(2), 153–175. https://doi.org/10.1016/j.wocn.2005.05.001

Flege, J.E., Bohn, O.S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*(4), 437–470. https://doi.org/10.1006/jpho.1997.0052

Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics, 23*(4), 567–598. https://doi.org/10.1017/S0142716402004046

Flege, J. E., & MacKay, I. R. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*(1), 1–34. https://doi.org/10.1017/S0272263104026117

Flege, J.E., Munro, M.J., MacKay, I.R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America, 97*(5), 3125–3134. https://doi.org/10.1121/1.413041

Flege, J.E., Schirru, C., & MacKay, I.R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*(4), 467–491. https://doi.org/10.1016/S0167-6393(02)00128-0

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America, 107*, 2711–2724. https://doi.org/10.1121/1.428657

Hong, Y. (1988). *A sociolinguistic study of Seoul Korean*. (Publication No. 8824748) [Doctoral dissertation, University of Pennsylvania]. ProQuest Dissertations Publishing.

Howie, J. (1976). *Acoustical studies of Mandarin vowels and tones* (Vol. 18). Cambridge: Cambridge
University Press.

Hwang, H., & Lee, H. Y. (2015). The effect of high variability phonetic training on the production of English vowels and consonants. In B. Beattie, J. Livingstone, M. MacMahon, J. M. Scobbie, R. Smith, J. Stuart-Smith, M. Wolters (Eds.), *18th International Congress of Phonetic Sciences, ICPhS 2015*, Glasgow, UK: University of Glasgow.

Igeta, T., & Arai, T. (2011). A case study on comparison of male and female vowel formants by native speakers of Korean. In W. S. Lee & E. Zee (Eds.), *17th International Congress of Phonetic Sciences, ICPhS 2011* (pp. 934–937). Hong Kong: City University of Hong Kong.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*(1), 47–57. https://doi.org/10.1016/S0010-0277(02)00198-1

Jenkins, J. J., Strange, W., & Polka, L. (1995). Not everyone can tell a "rock" from a "lock": Assessing individual differences in speech perception. In D. J. Lubinski & R. V. Dawis (Eds.), *Assessing individual differences in human behavior: New concepts, methods, and findings* (pp. 297–325). Palo Alto, CA: Davies-Black Publishing.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, *119*(2), 1118–1130. https://doi.org/10.1121/1.2151806

Kang, O. (2003). *Hankwuke umwunlon* [Phonology of Korean]. Seoul: Thayhaksa.

Kim, Z. S. (1998). Seoul thopakimaluy holsoli cocikey tayhan cosa yenkwu [A survey on the vowel system of Seoul Korean]. *Journal of the Institute of Seoul Studies, 10*, 271–286.

Kim-Renaud, Y. (2009). *Korean: An essential grammar*. New York, NY: Routledge.

Kim, D., Clayards, M., & Goad, H. (2017). Individual differences in second language speech perception across tasks and contrasts: The case of English vowel contrasts by Korean learners. *Linguistics Vanguard, 3*(1), 1–11. https://doi.org/10.1515/lingvan-2016-0025

Kuhl, P. K., Conboy, B., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period”. *Language Learning and Development, 1*, 237–264. https://doi.org/10.1080/15475441.2005.9671948

Lee, H. Y. (1996). *Kwuke umsenghak* [Korean Phonetics]. Seoul: Thayhaksa.

Lee, H. B. (1999). "Korean". In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet* (pp. 120–123). Cambridge: Cambridge University Press.

Lee, S. A. S., & Iverson, G. K. (2012). Vowel category formation in Korean–English bilingual children. *Journal of Speech, Language, and Hearing Research, 55*, 1449–1462. https://doi.org/10.1044/1092-4388(2012/11-0150)

Lee, W. S., & Zee, E. (2003). Illustrations of the IPA: Standard Chinese (Beijing). *Journal of the International Phonetic Association, 33*(1), 109–112. https://doi.org/10.1017/S0025100303001208

Lee-Ellis, S. (2012). *Looking into bilingualism through the heritage speaker’s mind*. (Publication No. 3518528) [Doctoral dissertation, Univrsity of Maryland, College Park]. ProQuest Dissertations Publishing.

Lin, Y. H. (2007). *The sounds of Chinese with audio CD* (Vol. 1). Cambridge: Cambridge University Press.

Liu, H. & Ng, M. (2009). Formant characteristics of vowels produced by Mandarin esophageal speakers. *Journal of Voice, 23*(2), 255–260. https://doi.org/10.1016/j.jvoice.2007.09.002

MacMillian, N. A. & Creelman, C. D. (1991). *Detection theory: A users guide*. Cambridge: Cambridge University Press.

The MathWorks, I. (2019). *Statistics and machine learning toolbox*. Natick, Massachusetts, United States. Retrieved from https://www.mathworks.com/help/stats/

Mok, P. P. K. (2012). Does vowel inventory density affect vowel-to-vowel coarticulation? *Language and Speech, 56*(2), 191–209. https://doi.org/10.1177/0023830912443948

Morrison, G. S. (2008). L1-Spanish speakers’ acquisition of the English /i/-/ɪ/ contrast: Duration-based perception is not the initial developmental stage. *Language and Speech, 51*(4), 285–315. https://doi.org/10.1177/0023830908099067

Munro, M., Flege, J. E., & MacKay, I. R. (1996). The effect of age of second language learning on the production of English vowels. *Applied Psycholinguistics*, *17*(3), 313–334. https://doi.org/10.1017/S0142716400007967

Norman, J. (1988). *Chinese*. Cambridge University Press.

Oh, J.R. (1997). *Hyentaykwuke umwunlon* [Phonology of Modern Korean]. Seoul: Hyengsel chwulphansa.

Peterson G. & Barney L. (1952). Control methods in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

Ryu, N. Y. (2018). Korean vowel identification by English and Mandarin listeners: Effects of L1-L2 vowel inventory size and acoustic relationship. *Toronto Working Papers in Linguistics*, *40*, 1–13. Retrieved from https://twpl.library.utoronto.ca/index.php/twpl/article/view/29209

Seo, Y., Dmitrieva, O., & Cuza, A. (2022). Crosslinguistic Influence in the Discrimination of Korean Stop Contrast by Heritage Speakers and Second Language Learners. *Languages*, *7*(1), 6. https://doi.org/10.3390/languages7010006

Shin, J. & Cha, J. (2003). *Wulimalsoliuy chyeykyey* [The Sound Pattern of Korean]. Seoul: Hankwukmunhwasa.

Shin, J., Kiaer, J., & Cha, J. (2012). *The Sounds of Korean*. Cambridge: Cambridge University Press.

Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, *33*(3), 263–290. https://doi.org/10.1016/j.wocn.2004.10.002

Van Leussen, J.W. & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, *6*, 1–12. https://doi.org/10.3389/fpsyg.2015.01000

Wan, I. P. & Jaeger, J. J. (2003). The phonological representation of Taiwan Mandarin vowels: A psycholinguistic study. *Journal of East Asian Linguistics*, *12*(3), 205–257. https://doi.org/10.1023/A:1023666819363

Wen, B., Wang, P., & Shi, F. (2017). On the mid vowel phoneme in Mandarin phonology. *Lingua*, *196*, 1–9. https://doi.org/10.1016/j.lingua.2017.06.003

Wiese, R. (1997). Underspecification and the description of Chinese vowels. In J. L. Wang & N. Smith (Eds.), *Studies in Chinese phonology* (pp. 219–249). Berlin: Mouton de Gruyter.

Wilson, K. (1999). Validity of global self-ratings of ESL speaking proficiency based on an FSI/ILR-referenced scale. *ETS Research Report Series, 2,* i-64. https://doi.org/10.1002/j.2333-8504.1999.tb01811.x

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics, 24*(2), 245–261. https://doi.org/10.1006/jpho.1996.0013