Korean vowel identification by English and Mandarin listeners: Effects of L1-L2 vowel inventory size and acoustic relationship*

Na-Young Ryu
University of Toronto

This paper examined how English and Mandarin listeners identify Korean vowels, focusing on native (L1) vowel inventory size and acoustic similarity between native and second language (L2) vowels. It was found that English listeners (with a larger vowel inventory) have higher identification accuracy scores for Korean vowels than Mandarin listeners (with a small vowel inventory), supporting the predictive role of vowel inventory size in non-native perception. To predict perceptual difficulties for L2 listeners, linear discriminant analysis (LDA) was undertaken. The LDA model predicted that English listeners will have the most perceptual difficulties with Korean vowels [o, u], while Mandarin listeners will have the most perceptual difficulties with [i, ʌ, o, u]. The results of a Korean vowel identification task revealed that both groups showed the same perceptual difficulty order: ɨ > ʌ > o > u, indicating that acoustic similarity between L1 and L2 vowels partially predicts vowel perceptual difficulty for L2 listeners.

1 Introduction

Adult L2 learners often struggle with the perception and production of non-native sounds. Models of L2 sound acquisition, such as Flege’s (1995) Speech Learning Model (SLM), Best’s (1994) Perceptual Assimilation Model (PAM) and Best and Tyler’s (2007) extension of the Perceptual Assimilation Model (PAM-L2) claim that cross-language perceptual and phonetic similarity can account for the relative ease or difficulty of the perception and production of various non-native sounds. According to the SLM, focusing primarily on the L2 speech production of experienced learners and having a strong perceptual component, the success of L2 learning depends on the perceived cross-language similarity between L1 and L2 sounds, as well as the age at which learning starts and the amount of L1 use. With respect to the cross-language perception for L2 listeners, it proposes that learners’ perceptual difficulties can be predicted by whether the relevant sound in their target language is identical, new or similar to their L1 phonetic categories. If L2 phones are identical to L1 phones, they are not expected to have difficulties as a result of positive transfer. In addition, new L2 phones, which do not exist in L1, are perceived with little interference from L1 and eventually establish a new L2 category, given enough exposure to the L2. In contrast, similar L2 sounds, which are perceived as phonetically close to an L1 category, are expected to be more difficult to acquire.

* The research reported in this paper was supported by an Academy of Korean Studies research grant (ASK-2017-R23) awarded to the author. I would like to thank Dr. Yoonjung Kang for her indispensable support of this research, as well as Dr. Kyoungrok Ko and Dr. Yujeong Choi for help with recruiting participants. I am also grateful to audiences at the 2017 CRC Summer Phonology/Phonetics Workshop for their valuable comments.
because learners are less likely to create separate L2 categories for these acoustically similar sounds. While the SLM predicts perceptual difficulties of nonnative sounds based on a comparison of individual sounds in L1 and L2, the PAM and its extension to L2 acquisition, PAM-L2, compare L2 sound contrasts to L1 categories. The PAM proposes that accuracy in the discrimination of nonnative sound contrasts are on the basis of the way the members of each contrast are assimilated to L2 sound categories. For example, the greatest L2 listeners’ discrimination difficulty is expected when two nonnative sounds are assimilated to a single L1 category (single-category assimilation). On the other hand, for two nonnative sounds which are mapped onto two different L1 categories, L2 categories will be accurately discriminated, regardless of L2 experience (two-category assimilation). More recently, Escudero’s (2005, 2009) Second Language Linguistic Perception model (L2LP) proposes that phonetic-acoustic similarity between L1 and L2 sounds influences L2 perception and allows for predictions of listeners’ initial state of L2 acquisition and the degree of difficulty. The L2LP claims that learners are able to exhibit native-like perception on both L1 and L2 languages, because both languages are handled by two separate systems, contrary to the claims made by the SLM. The L2LP makes predictions about the relative ease of non-native sound perception on the basis of a phonetic-acoustic comparison of L2 contrasts to L1 categories. According to their scenarios, if learners map an L2 contrast to a single L1 category, known as single-category assimilation in PAM’s terms, or as a new scenario in the terms of L2LP, they will have to create a new L2 category or split an existing L1 category to acquire the new contrast. In contrast, if learners map a new contrast to two different L1 categories, via two-category assimilation in the PAM and similar scenario in the L2LP, it is expected to be easier to acquire than a new scenario.

In the current study, I investigated how L1 vowel inventory size and acoustic similarity between L1 and L2 vowels affect English and Mandarin listeners’ identification difficulties of Korean vowels. Many cross-language studies find that L1 vowel inventory size influences L2 vowel perception (Escudero et al., 2014; Fox et al., 1995; Iverson & Evans, 2007, 2009; Bundgaard-Nielsen et al., 2011). Having a smaller L1 vowel inventory than L2 often results in more perceptual difficulties, as two vowels in L2 are more likely to correspond to a single L1 category perceptually. In contrast, having a larger L1 vowel inventory than L2 often facilitates L2 perception, because there are enough different L1 categories for all L2 sounds to map distinctly. For example, Iverson and Evans (2007) report that native speakers of German and Norwegian, having a larger and more complex vowel system than English, identify English vowels more accurately than French and Spanish native speakers, whose L1 vowel inventories are smaller. Other studies, however, find that the L1 inventory size is not always a good predictor of L2 perceptual difficulty. Elvin et al. (2014) demonstrate that Iberian Spanish listeners, who have a smaller vowel inventory in comparison to speakers of Australian English, outperform Australian English listeners in their discrimination of six Brazilian Portuguese vowel contrasts. In short, there are conflicting findings on the effect of L1 vowel size on L2 perception and it is important to consider how and to what extent L1 and L2 vowel inventories interact in L2 perception.

In order to examine the effects of L1 vowel inventory size on L2 vowel perception, the present study chose three languages with different vowel inventory sizes: Mandarin, English and Korean. Mandarin has the smallest number of vowels with five monophthongs [i, y, a, u, ə], with several allophonic variations (Duanmu, 2007; Lin, 2007). Canadian English has the largest vowel inventory with 10 monophthongs [i, ɪ, ɛ, ɛ, æ, a, ð, o, ʊ, ʊ] (Boberg, 2008), while Korean falls in the middle with seven monophthongs [a, e, i, o, u, i, ɨ] (Sohn, 2001).

In addition to L1 vowel inventory size, acoustic-phonetic properties of L1 and L2 also affect L2 vowel perceptual accuracy. Numerous L2 perception studies examine the role of L1-L2 cross-language acoustic similarities on the perception of L2 vowels (e.g., Feige et al., 1997; Best, 1995, Best & Tyler 2007; Escudero, 2005; Escudero, 2009). Strange et al. (2005) compare the acoustic properties of American English and Northern German vowels using Klecka’s (1980) linear discriminant analysis (LDA) – that is, how an LDA model constructed based on L1 (Northern German) vowels classifies L2 (American English) vowels – and find that the model’s classifications do not accurately predict Northern German listeners’ perceptual assimilations of American English vowels. However, some recent studies show that cross-
linguistic acoustic similarity can successfully predict difficulty in non-native vowel perception (Escudero & Chládková, 2010; Escudero & Vasiliev, 2011; Alispahic et al., 2017).

The present study contributes to this debate by examining to what extent acoustic similarity between Korean and English as well as Korean and Mandarin can predict L2 learners’ Korean vowel difficulties through LDA models. Specific questions and predictions addressed in the present study are as follows:

(1) Question 1: Are English listeners (with a large and complex vowel system) more accurate at identifying Korean vowels than Mandarin listeners (with a small vowel inventory)?
   a. Prediction: If vowel inventory size is a reliable predictor of non-native vowel perception, English listeners, whose vowel inventory is larger than that of Korean, are predicted to outperform Mandarin listeners, whose vowel inventory is smaller compared to Korean in their identification of Korean vowels.

(2) Question 2: How accurately does the acoustic similarity between English and Korean vowels and between Mandarin and Korean vowels predict identification difficulty of Korean vowels for English and Mandarin listeners?

2 Experiment 1: A comparative acoustic study of Korean, English and Mandarin vowels

The aim of Experiment 1 was to examine the vowel characteristics of three languages (Korean, Mandarin, English) to predict which Korean vowels are easier or more difficult for English and Mandarin listeners to perceive.

2.1 Methods

2.1.1 Participants

A total of 68 female speakers participated in Experiment 1. To avoid potential variation due to gender difference, only female speakers were recorded. For the Mandarin group, 37 female native Mandarin speakers who were born and raised in China (mean age = 21 years) were recruited; for the English group, 23 native English speakers, who were born and raised in Canada (mean age = 21 years) were recruited. Moreover, eight native Korean speakers who were born and educated in the Seoul/Kyenggi region (mean age = 24 years) acted as a comparison group. The task was conducted individually in the Phonetics Laboratory at the University of Toronto. All participants reported no hearing and speaking disorders and they were paid $10 for their participation.

2.1.2 Materials

For the acoustic study of vowels, seven Korean vowels [a, e, i, o, u, ɨ, ʌ], 10 Canadian English vowels, [ɪ, i, æ, ɛ, ə, ʌ, ɑ, o, u] and five Mandarin vowels [i, y, ɤ, u, ɑ] were recorded by the participants in this experiment. Each vowel appeared in similar phonetic environments to avoid coarticulation effects (Hillenbrand et al, 2001). Following Yang (1996), an [hVda] context for the Korean speakers and an [hVd] context for the English speakers were used. Following Liao (2006), Mandarin words contained vowels in an [hVda] context with Tone1 on the first syllable followed by the syllable [da], ɤ, with a neutral tone. When there was no corresponding Mandarin character, Pinyin was used. See Appendix A for Korean words, B for English words and C for Mandarin words used in this experiment.
2.1.3 Procedure

Participants were instructed to read a word list presented in their native language on a computer screen, using PsychoPy (Peirce, 2007). Each stimulus appeared three times in isolation. They were asked to read each word as naturally as possible at a normal speaking rate, and practiced before the actual recording. The recordings were made with a Sound Device 722 digital audio recorder and a DPA 4011 cardioid microphone. The stimuli were recorded at a sampling rate of 44.1 kHz.

2.1.4 Acoustic measurements

1,341 tokens (English tokens: 723, Mandarin tokens: 472, Korean tokens: 146) were acoustically analyzed. For each token, the first and second formant frequencies (F1 and F2) were measured at the 50% temporal points of the vocalic duration in Praat (Boersma & Weenink, 2018).

2.1.5 Data analysis

I used linear discriminant analysis (LDA) models to test whether acoustic similarities are predictive of vowel classification pattern for L2 listeners. This was accomplished by using the F1/F2 bark values as well as the seven target Korean vowels. Note that the LDA model works on the assumption that similarity of auditory parameters (e.g., F1, F2 or duration) can predict how L2 vowels are perceived in L1. LDA models were built for English and Mandarin vowels according to their F1 and F2 bark values and these models were used to predict how the Korean vowels are perceptually classified into Mandarin and English vowels.

2.2 Results

2.2.1 Cross-language acoustic comparison

This section shows the results of within-language and cross-language acoustic comparisons. Figure 1 shows results of the acoustic vowel analysis of the three languages produced by female native speakers. The ellipses represent one standard deviation from the mean.

Figure 1: First formant (F1) and second formant (F2) (bark) plots of Korean, English1 and Mandarin vowels produced by female native speakers of each language.

1 English vowels [e] and [o] are somewhat diphthongized.
As shown in Figure 1, all vowels are clearly separated from each other in English and Mandarin, while Korean has [u] and [o], which overlap substantially. The finding confirms that Korean vowels [o] and [u] are in the process of merging in Seoul Korean (Seong, 2004; Han & Kang, 2013). Given the extensive overlap of [o] and [u] in Korean production, I expect that L2 learners will have difficulty distinguishing the two vowels.

The acoustic similarities are presented in Figure 2 and Figure 3, respectively. Figure 2 compares vowel qualities of English produced by female native Canadian English speakers and Korean vowels produced by female native Korean speakers.

*Figure 2: Average F1 and F2 values (Hz) of Canadian English (red) and Korean vowels (black) produced by adult female native English and Korean speakers*

As seen in Figure 2, cross-language acoustic comparison of Korean and English vowels revealed that Korean [i] was acoustically very similar to English [i], leading to the prediction that that would not cause difficulties in acquisition. However, there are greater cross-language differences for the vowels [ʌ, o, u, i]. Korean [ʌ] is much closer to English [a] than to English [ʌ]. Korean [u] is higher in F1 values and lower in F2 values than English [u], while Korean [o] is lower both in F1 and F2 values which could lead to confusion and discrimination difficulty for this contrast. Moreover, English [u] appears to be acoustically intermediate between Korean [i] and [u]. It can be predicted that English learners of Korean will have difficulty distinguishing the Korean [u]-[i] contrast.

*Figure 3: Average F1 and F2 values (Hz) of Mandarin (blue) and Korean (black) vowels produced by adult female Mandarin and Korean speakers*
The general result of this comparison of the Korean and Mandarin vowels in Figure 3 is that there are acoustically similar and different vowels between the two languages. Korean [i] and [a] are acoustically most similar to Mandarin [i] and [ɑ], respectively, which can be expected to be easy to acquire for Mandarin L2 learners. However, two distinct Korean vowels [o] and [u] are acoustically very close to a single Mandarin [ʊ] category. More specifically, Mandarin [u] and Korean [o] have similar F2 values, but Mandarin [u] has lower F1 values than Korean [ʊ]. Also, Mandarin [u] has a lower F1 and F2 value than Korean [u]. Due to the acoustically short distance between Korean [o] and [u], Mandarin [u] can correspond to both of them, making it difficult for Mandarin L2 learners to accurately acquire Korean [u] and [o].

2.2.2 Discriminant analysis and predictions for perceptual assimilation

I used LDA models to predict the patterns of perceptual assimilation from the acoustic properties of Korean vowels and those of the closest vowels in Mandarin and English. That is, LDA allows us to predict how Mandarin and English speakers perceive Korean vowels in relation to their L1 vowels, by using F1 and F2 measurements converted to the bark scale, as the bark scale is a more accurate representation of how the human auditory system processes acoustic stimuli (Escudero et al., 2012). The results of the LDA model predict different perceptual assimilation of Korean vowels by English listeners and Mandarin listeners. The results are presented in Table 1 and 2 respectively.

Table 1: Korean vowel tokens classified across English vowels using cross-language LDA

<table>
<thead>
<tr>
<th>Korean vowel</th>
<th>æ</th>
<th>a</th>
<th>e</th>
<th>e</th>
<th>i</th>
<th>i</th>
<th>o</th>
<th>o</th>
<th>u</th>
<th>o</th>
<th>ʌ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>14</td>
<td>2</td>
<td>9</td>
<td>20</td>
<td>1</td>
<td>14</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>12</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>2</td>
<td>20</td>
<td>6</td>
<td>1</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>6</td>
<td>1</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o</td>
<td></td>
<td>19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td>1</td>
<td>19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʌ</td>
<td>16</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the overall LDA classifications as seen in Table 1, predicted perceptual patterns for Korean vowels by English listeners are as follows: Korean [i] is mostly classified to English [i], indicating English speakers are more likely to perceive the vowel correctly. Similarly, Korean vowels [a], [e], [i] and [ʌ] are mostly assimilated to English vowels [æ], [ɛ], [ʊ] and [ɑ] respectively. Therefore, it is predicted that these Korean vowels are easy for English listeners to identify. However, English listeners would have greater difficulty discriminating and identifying the two different Korean vowels [o] and [u], since they are mapped to a single English vowel [o].

Table 2: Korean vowel tokens classified across Mandarin vowels using cross-language LDA

<table>
<thead>
<tr>
<th>Korean vowel</th>
<th>a</th>
<th>ɤ</th>
<th>i</th>
<th>u</th>
<th>ʊ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>19</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>1</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>1</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>10</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>1</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u</td>
<td>2</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ʌ</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6
Table 2 presents the tokens of Korean vowels being classified as any of the seven Mandarin vowels. The results of cross-language LDA showed that for Korean [a] and [i], the model yields 100% to Mandarin [a] and [i] classification, predicting little perceptual difficulty for Mandarin listeners. Korean vowels [i] and [ʌ] are mostly categorized as single Mandarin [y]; Mandarin listeners are predicted to find the two Korean vowels challenging to discern. In addition, two distinct Korean vowels, [o] and [u], are mapped to a single Mandarin vowel category [u], predicting perceptual difficulties for Mandarin listeners for these vowels.

In summary, the overall results predicted by the LDA model showed that English and Mandarin listeners have different perceptual assimilation for Korean vowels. English listeners will have difficulties perceiving Korean vowels as follows: (easiest) [a, e, i, ɨ, ο, u] >> [o, u] (most difficult). Mandarin listeners, on the other hand, will have difficulties perceiving Korean vowels as follows: (easiest) [a, e, i] >> [i, ο, u] (most difficult). Experiment 2 examines Korean vowel identification by English and Mandarin listeners and tests how well perceptual assimilation patterns can be predicted on the basis of acoustic similarity between L1 and L2 vowels, specifically as measured by the LDA models.

3 Experiment 2: Perception of Korean vowels by Mandarin and English listeners

The primary goal of Experiment 2 was to assess the accuracy of English and Mandarin listeners’ Korean vowel identification and to determine whether the acoustic similarity between L1 and L2 predicted by the LDA model affects perception of Korean vowels. The secondary goal was to investigate whether listeners’ L1 vowel inventory size plays a role in L2 perception performance.

3.1 Methods

3.1.1 Participants

In total, 76 participants took part in Experiment 2: 38 Mandarin (35 female, three male, mean age = 21 years) and 28 English participants (22 female, six male, mean age = 22 years) who were undergraduate students who have taken Korean courses at the University of Toronto. They were native speakers of Mandarin and Canadian English living in Toronto at the time of testing. Ten native Korean speakers participated as controls. All participants were the same as those in Experiment 1, and male participants were included in Experiment 2.

3.1.2 Materials

The materials consisted of 92 Korean monosyllabic words, including the target four vowels [ɨ, ο, u, Λ] which are relatively difficult to acquire for L2 learners (E. Kim, 2010, 2016; Han & Kim, 2011). The full set of stimuli is provided in Appendix D. Auditory stimuli were natural recordings by a male native speaker of Korean (46 years old). Each item was recorded five times with a DPA 4011 cardioid microphone at a sampling rate of 44.1 kHz. From these recordings, the most natural and clear tokens were selected by the author for the stimuli.

3.1.3 Procedure

Experiment 2 was an identification task. The participants were instructed to listen to a Korean stimulus over headphones and to identify the stimulus out of four choices presented on a computer screen by pressing corresponding numbers 1, 2, 3, or 4 on the keyboard. For example, on a given trial, a participant might hear ‘ko <고>’. The task is then to choose from four visually presented stimuli, e.g., ki <고>, ko <고>, ku <구>, kΛ <거>, by pressing the corresponding number on the keyboard. The inter-trial interval was 1000 ms.
The participants were asked to respond as accurately and quickly as possible. All the instructions were provided in Mandarin for the Mandarin speakers, in English for the English speakers and in Korean for the Korean speakers. The experiment was deployed using the PsychoPy experiment builder software and conducted in a sound attenuated booth located in the University of Toronto Phonetics Laboratory.

Before the actual experiment, each participant completed a practice session to ensure familiarity with the identification task. There were 92 target trials and all trials were randomly presented for each participant. There was no time-out and participants had to choose a stimulus in order to listen to the next stimulus.

3.1.4 Statistical analysis

To determine whether identification accuracy differs between participants whose native language has more or fewer vowels compared to Korean, I compared performance between English and Mandarin listeners. The statistical analysis was performed with a mixed-effects logistic regression model (Baayen et al., 2008) using the lme4 package (Bates et al., 2011) in R (R Development Core Team, 2014). In the model, response was the dependent variable (correct response 1 vs. incorrect response 0) and the fixed effects predictor was the L1 language (Korean, Mandarin, and English). The L1 language was Helmert coded to compare non-Korean listeners against Korean listeners and then compare English and Mandarin listeners. This model also included subjects and items as random effects.

3.2 Results

3.2.1 Effects of L1 vowel inventory size

This section presents the results of the identification task by comparing response accuracy of Korean vowels by adult Mandarin and English listeners who have different vowel inventory systems.

Table 3: The output of a mixed-effects logistic regression model in the identification task

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>Z-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(intercept)</td>
<td>2.3744</td>
<td>0.1385</td>
<td>17.147</td>
<td>&lt; 2e-16 ***</td>
</tr>
<tr>
<td>Korean vs. non-Korean</td>
<td>-1.8547</td>
<td>0.3596</td>
<td>-5.158</td>
<td>2.5e-07 ***</td>
</tr>
<tr>
<td>Mandarin vs. English</td>
<td>0.4923</td>
<td>0.2314</td>
<td>2.128</td>
<td>0.0334 *</td>
</tr>
</tbody>
</table>

Significance codes: <0.001 ‘***’, <0.01 ‘**’, <0.05 ‘*’, <0.1 ‘.’

Note: Underlines indicate the reference categories in each comparison.

Results of the mixed-effects logistic regression model revealed a main effect of language background. Korean listeners were extremely accurate at identifying Korean vowels, while none of L2 listener groups achieved native Korean speakers’ levels of perceptual vowel accuracy. Between English and Mandarin listener groups, it was predicted that English L2 learners, having a larger vowel inventory, will outperform Mandarin L2 learners, having a smaller vowel inventory. The results showed that English listeners had more correct responses for Korean vowels than Mandarin listeners, consistent with the hypothesis that there is an effect of L1 vowel inventory size on L2 vowel perception.
As shown in Figure 4, native speakers of Korean as a control group reached 99% identification accuracy for all Korean vowels except for [u] (89%), which was misidentified to [o] (11%). This might be due to [u] becoming increasingly approximated with [o] by young Korean speakers. This shows that the contrast between [o] and [u] poses some difficulty even for native Korean listeners, as we may expect from the significant overlap in their distribution.

The perceptual difficulty of L2 listeners from least to most difficult was independent of their L1 background: i > ʌ > u > o. Interestingly, although [i] is a new L2 phone to both English and Mandarin listeners, L2 listeners were most likely to accurately identify Korean [i], suggesting that, with L2 experience, listeners can establish a new L2 phonological or phonetic category and perceive it correctly, as proposed by the PAM and SLM models.

English listeners were more successful at identifying [ʌ, u, o] than Mandarin listeners, while they were less accurate at identifying [i] than Mandarin listeners. This result was partly confirmed by the LDA model which predicted that Mandarin listeners have perceptual difficulties with four vowels [i, ʌ, u, o], whereas English listeners have perceptual difficulties with only two vowels [o, u]. Mandarin listeners showed little perceptual difficulty with [i] in their identification task, which was not a part of the prediction made by the model. Nonetheless, the model made correct predictions that both listener groups have low identification scores for Korean [o, u], since the two distinct Korean vowels are classified to a single English and Mandarin vowel category [u].

Figure 5 illustrates how English and Mandarin listeners identified Korean vowels and the degrees of perceptual difficulty for the vowels. Interestingly, Mandarin listeners (with a smaller vowel system) are more successful in perceiving [i] than English listeners (with a larger vowel system). This might be due to
the fact that L2 listeners with a smaller L1 vowel inventory have unoccupied vowel space to learn new vowels. As Flege (1995, 2003) argued, changing existing categories is more difficult for L2 learners than learning entirely new categories.

For the target vowels [ʌ] and [ʊ], both groups misidentified to [o]. For the target vowel [o], Mandarin listeners misidentified to either [u] (10%) or [ɨ] (19%), while English listeners misidentified to mostly [ʌ] (18%), showing that misidentification patterns for [o] differs by L1 background.

4 Discussion and conclusion

This study investigated whether L1 vowel inventory size affects L2 vowel perception and to what extent cross-language acoustic similarity (using the LDA model) plays a role in predicting L2 perceptual difficulty by comparing English and Mandarin listeners’ identification patterns of Korean vowels. With respect to the influence of L1 vowel inventory, English listeners were generally better at identifying Korean vowels than Mandarin listeners, suggesting that individuals with a large L1 vowel inventory have an advantage in L2 vowel perception. However, Mandarin listeners had higher accuracy than English listeners in Korean [i], which is a new L2 phone to both Mandarin and English listeners. This may be due to the fact that learners with a larger L1 vowel system have a relatively crowded vowel space that interferes with formation of new categories, whereas those with smaller L1 vowel system have more room to adapt new categories. Thus, new L2 phones can be relatively easy to establish for learners with small L1 vowel inventory and are more accurately perceived than by those with a complex and large L1 vowel inventory, since it is harder to create a separate category for new L2 vowels when they are close to an existing category. That is, individuals with a larger L1 vowel system may rely more on L1 assimilation and less on new learning than the L2 listeners with a smaller L1 vowel system.

The predictions of perceptual difficulty for L2 listeners using LDA models were borne out. The model predicted that Mandarin listeners had the most difficulty perceiving Korean [i, ʌ, u, o] and English listeners had the most difficulty perceiving Korean [o, u]. The findings of the Korean vowel identification task showed that perceptual difficulties for both English and Mandarin listeners were as follows: i > ʌ > u > o. Both listeners were most likely to accurately perceive Korean [i], while they were least likely to accurately perceive [o]. Findings from the identification task were partially in line with acoustic predictions.

This work had several limitations and raised several questions for future research. There are possible reasons why perceptual vowel difficulty is not fully predicted by LDA models. First, contextual allophones and Mandarin diphthongs were not taken into account for acoustic stimuli and only F1 and F2 were used for input parameters. If F1, F2, F3 values and vowel duration were implemented for the input parameters, the LDA model might yield a more correct classification of perceptual difficulty. In addition, the audio stimuli were recorded by a male Korean speaker who is in his 40s, but all participants for the perception experiments were in their 20s, and the L2 learners are likely more experienced with younger Koreans’ speech through Korean pop culture. Moreover, acoustic properties of vowels used in LDA models are extracted from vowels in the [hVd(a/a)] context, while Korean vowels in the identification task appeared in all contexts. Thus, future study would benefit by considering these limitations to generate more accurate perception predictions.

References


**Appendix A: Korean words in the [hVda] context**

<table>
<thead>
<tr>
<th>Korean vowel</th>
<th>a</th>
<th>e</th>
<th>i</th>
<th>o</th>
<th>u</th>
<th>ɨ</th>
<th>ʌ</th>
</tr>
</thead>
<tbody>
<tr>
<td>word</td>
<td>하다</td>
<td>헤다</td>
<td>히다</td>
<td>호다</td>
<td>후다</td>
<td>호다</td>
<td>허다</td>
</tr>
</tbody>
</table>

**Appendix B: English words in the [hVd] context**

<table>
<thead>
<tr>
<th>English vowel</th>
<th>i</th>
<th>ɨ</th>
<th>æ</th>
<th>ɛ</th>
<th>a</th>
<th>ʌ</th>
<th>ʊ</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>word</td>
<td>heed</td>
<td>hid</td>
<td>had</td>
<td>head</td>
<td>hod</td>
<td>Hudd</td>
<td>hood</td>
<td>who’d</td>
</tr>
</tbody>
</table>

**Appendix C: Mandarin words in the [hVda] context**

<table>
<thead>
<tr>
<th>Mandarin vowel</th>
<th>ɑ</th>
<th>ə</th>
<th>i</th>
<th>u</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>word</td>
<td>哈的 (hade)</td>
<td>喝的 (hede)</td>
<td>hi 的 (hide)</td>
<td>乎的 (hude)</td>
<td>hü 的 (hüde)</td>
</tr>
</tbody>
</table>
Appendix D: Korean monosyllabic nonsense words including vowels [i, ʌ, u, o] for the perception task

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Coda</th>
<th>onset</th>
<th>ㄱ</th>
<th>ㄴ</th>
<th>ㄷ</th>
<th>ㅂ</th>
<th>ㅅ</th>
<th>ㅁ</th>
<th>ㅇ</th>
<th>ㄹ</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>코다</td>
<td>ㄱ</td>
<td>ㄴ</td>
<td>ㄷ</td>
<td>ㅂ</td>
<td>ㅅ</td>
<td>ㅁ</td>
<td>ㅇ</td>
<td>ㄹ</td>
<td></td>
</tr>
<tr>
<td>그</td>
<td>국</td>
<td>길</td>
<td>금</td>
<td>길</td>
<td>슨</td>
<td>슨</td>
<td>슨</td>
<td>슨</td>
<td>슨</td>
<td></td>
</tr>
<tr>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td>음</td>
<td></td>
</tr>
<tr>
<td>ㅏ</td>
<td>거</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td>걷</td>
<td></td>
</tr>
<tr>
<td>ㅗ</td>
<td>고</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td>골</td>
<td></td>
</tr>
<tr>
<td>ㅜ</td>
<td>구</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td>국</td>
<td></td>
</tr>
</tbody>
</table>

KOREAN VOWEL IDENTIFICATION BY ENGLISH AND MANDARIN LISTENERS