Difficulty in discriminating non-native vowels: Are Dutch vowels easier for Australian English than Spanish listeners?

Samra Alispahic, Paola Escudero, and Karen E. Mulak

MARCS Institute, University of Western Sydney, Penrith, Australia
s.alispahic@uws.edu.au, paola.escudero@uws.edu.au, k.mulak@uws.edu.au

Abstract

Previous studies have shown that the number of vowels present in one’s L1 inventory may affect the ability to learn and discriminate non-native vowel contrasts. Specifically, learners whose L1 contains fewer vowels compared to the target language may find many non-native vowel contrasts novel and have discrimination performance lower than learners whose L1 contains more vowels than the target language for whom most of the non-native vowel contrasts will be familiar. The present study tested monolingual Australian English (AusE) listeners’ discrimination of non-native vowels in Dutch, which has fewer vowels compared to AusE. We further compared AusE listeners’ performance to that of native monolingual Spanish listeners whose L1 contains fewer vowels than Dutch. AusE listeners were able to discriminate all Dutch vowel contrasts above chance. While there was no main effect of language background, an interaction language background x contrast revealed that AusE listeners more accurately discriminated the /ɪ-ʏ/ contrast compared to Spanish listeners, suggesting some advantage for AusE listeners. The findings are discussed in relation to models of non-native and L2 speech perception together with a comparison of vowel acoustic properties across AusE, Spanish and Dutch.

Index Terms: non-native speech perception, vowel discrimination, acoustic phonetics, phonetics and phonology.

1. Introduction

Second-language (L2) acquisition is a difficult task, due in part to difficulties in the discrimination of acoustic and/or phonological distinctions not present in the learner’s native language (L1). The initial stage of L2 learning is comparable to that of naïve listeners [1], as described in the Second Language Linguistic Perception model (L2LP) [2] and in the extension of the Perceptual Assimilation Model to L2 acquisition (PAM-L2) [3]. These two models suggest that naïve listeners perceive L2 contrasts in accordance to the features of their native phonemes, and that this predicts whether and to what extent contrasts will be discriminated and learned during L2 acquisition. If two non-native contrasts are perceived within a single native category, they will be challenging to distinguish and acquire. Alternatively, they will be more easily discriminated if they are perceived as belonging to two separate native categories [4].

Additionally, the L2LP model stipulates that the sounds in a learner’s L1 vowel inventory will determine a learners’ perception (and production) of L2 sounds [2]. Learners whose L1 has fewer categories than the target language are required to learn new categories [5], while learners whose native language has a greater number of categories are required to unlearn certain categories [6]. However, the model predicts that unlearning or adjusting existing categories will be easier for a learner than learning new ones.

Recent research has also demonstrated that having a larger L1 vowel inventory may aid the acquisition of L2 vowels [7]. German and Norwegian listeners more accurately identified English vowels in two English vowel identification tasks than French and Spanish listeners whose L1 vowel inventory is smaller than that of English. The French and Spanish listeners’ difficulty in identifying English vowels was apparent in their assimilation of three or more English vowels to a single L1 vowel category (e.g., English /æ/, /ˈæ/, and /aʊ/ were all assimilated to the Spanish and French /a/ equivalent). Thus, these listener groups would need to learn new categories in order to avoid misidentifying non-native vowel categories.

In line with these results, Spanish listeners, whose vowel inventory is smaller than that of Dutch, had substantial difficulty discriminating the Dutch contrasts /a-ɑ/, /i-ɪ/ and /y-ʏ/ because these contrasts do not exist in Spanish [8]. However, Spanish listeners’ discrimination of Dutch vowel contrasts has not been compared to that of a listener group with a larger L1 vowel inventory.

The present study contributes to this line of research by directly comparing Peruvian Spanish and Australian English listeners’ discrimination of five Dutch vowel contrasts: /a-ɑ/, /i-ɪ/, /y-ʏ/, /ɛ-ɛ/ and /y-ʏ/. While Spanish has a small vowel inventory consisting of five vowels (/a, ɛ, i, o, u/) [8], Standard Northern Dutch has eight short or monophthongal vowels (/ɑ, r, i, ɪ, o, u, v, y/) and four long mid vowels (/a, ɛ, o, u/) and three diphthongal vowels (/ɛi, ɔu, ey/) [9]. Compared to Dutch, Australian English (AusE) has an additional three monophthongs and two diphthongs, for a total of 12 monophthongs (/i, ɪ, c, e, ɛ, v, ɔ, o, ʌ, ʊ, u, y/) and 6 diphthongs (/æɪ, ɑe, æə, ɔɪ, ɪə, ɪɔ/) [10]. Figure 1 shows the F1 and F2 acoustic values for the monophthongs of these three languages.

If our findings are in line with the vowel inventory predictions and the results of [7, 13] AusE listeners, having a larger vowel inventory than that of Dutch, should outperform Spanish listeners, whose vowel inventory is smaller than that of Dutch, in their discrimination of Dutch vowel contrasts.
Figure 1. Average F1 and F2 acoustic values for vowels produced by a male native speaker of Australian English (in black: [10]), Northern Standard Dutch (circled: [11]), and Peruvian Spanish (grey: [12]).

We hypothesize that AusE listeners will easily discriminate the Dutch vowel contrast /a-/a/. Alternatively, they may perceive the Dutch vowel /a/ more accurately as its values are more similar to the AusE /e/ and /ɛ/ than the Dutch vowel /a/, which is instead closer to the AusE /ɛ/ vowel. Listeners should be able to differentiate the /i-i/ contrast as a result of the AusE front vowels /i/ (as in HEED) and /u/ (as in HID) being present in their L1. Listeners may, however, discriminate the Dutch vowel /i/ easier as its acoustic values are more similar to the AusE /i/ and /ɪ/ than the Dutch /ɪ/, which is more acoustically similar to the AusE vowel /ɛ/.

AusE listeners should not find it difficult to discriminate between the short Dutch monophthongal vowels presented in the contrast /ɪ-/y/, as the acoustic values discriminating these vowels align with the differences between the AusE vowels /ɑ/ (as in WHO’D) and /ɔ/ (as in HOOD). The same pattern should be observed for the Dutch contrast /i-y/, as AusE listeners should find it easy to differentiate these two vowels due to the durational contrast present between the AusE front vowel /i/ (HEED) and the high vowel /ʊ/ (WHO’D). Finally, it is predicted that AusE listeners should be able to discriminate the Dutch contrast /ɪ-/y/, as they are differentiated by durational cues similar to those differentiating the AusE vowels /i/ (HID) and /ʊ/ (HOOD).

Conversely, we predict that Spanish listeners should find it relatively difficult to discriminate the five Dutch vowel contrasts. This is because four of the Dutch vowels presented, /ɑ/, /ɛ/, /ʏ/, and /ʊ/, are not part of the Spanish vowel inventory [8]. While it could be expected that Spanish listeners might use the durational difference present in the /a/a/ contrast as a discriminatory feature, this possibility was removed by equalizing the duration of vowel stimuli in order to ensure that discrimination was on the basis of spectral features. Furthermore, Spanish listeners are predicted to have difficulty discriminating the Dutch contrast /i-/ɪ/ as Spanish has only the category /ɨ/, and no corresponding category for /ɪ/, and thus both Dutch tokens are expected to be perceived as /ɨ/. However, this could assist the Spanish listeners when discriminating between the/i-y/ and /ɪ-/y/ contrasts as /i/ and /ɪ/ may be perceived as Spanish /u/.

Based on the stipulations put forth by the L2LP model [2], we predict that the size of listeners’ native vowel inventories will affect discrimination of the Dutch vowel contrasts. We further predict that the acoustic properties of listeners’ L1 vowels will influence their discrimination patterns.

## 2. Method

### 2.1 Listeners

AusE listeners were 11 students aged 18 to 45 years ($M = 24.19$) attending the University of Western Sydney, who participated for course credit. Participants were born in Australia, and grew up in the Greater Western Sydney area. Participants reported no to hardly any experience with a language other than English ($<1$ on a [hardly at all] – 7 [highly fluent] scale), and reported no prior experience with Dutch.

Spanish listeners were 11 participants randomly selected from a data set of 22 listeners from [8]. Listeners were Peruvian Spanish monolingual males and females who had lived in Lima, Peru their entire life and reported no prior knowledge of Dutch or English. They were between the ages of 18 to 28 years ($M = 20.95$). They were paid 10 dollars an hour for their participation.

### 2.2 Stimuli

Listeners heard 20 isolated tokens of each of the Dutch vowels /a/-a/ /i/-i/ /y/-y/ and /u/-u/, produced by 10 male and 10 female native speakers of Northern Standard Dutch, which were selected from the corpus reported in [9]. Vowels were extracted from words produced in an /s/-V-/s/ context, where V stands for the vowel. Each word was produced in a Dutch carrier sentence. These stimuli were used in previous studies examining the identification and discrimination of Dutch vowels [1, 15]. The Dutch sentences, their IPA transcriptions, and their English translations are included in Table 1 below.

### Table 1. Carrier sentences for words containing Dutch vowels [9].

<table>
<thead>
<tr>
<th>Carrier sentences for short vowels and monophthongs</th>
<th>Dutch sentence</th>
<th>IPA transcription</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>In sVs en in sVse zit de V</td>
<td>/in sVs en in sVse zit de V</td>
<td>/in sVs en in sVse zit de V</td>
<td>In sVs en in sVse zit de V</td>
</tr>
<tr>
<td>Carrier sentences for long vowels and diphthongs</td>
<td>Dutch sentence</td>
<td>IPA transcription</td>
<td>English translation</td>
</tr>
<tr>
<td>In sVs en in sVze zit de V</td>
<td>/in sVs en in sVze zit de V</td>
<td>/in sVs en in sVze zit de V</td>
<td>In sVs en in sVze zit de V</td>
</tr>
</tbody>
</table>

### 2.3 Procedure

Peruvian Spanish listeners took part in two forced-choice discrimination XAB tasks in Lima, Peru. We only report on the data for the first task, which was an auditory XAB (AUDI-XAB) discrimination task. In this task, listeners were presented with the five naturally produced Dutch vowel
contrasts, /a-ɑ/, /ɪ-i/, /y-ʏ/, /i-ʏ/ and /ɪ-ʏ/, described above. Participants heard three sounds, which were presented at a comfortable volume through headphones, and were asked to decide if the first sound (X) was more like the second (A) or third (B) sound.

Using the computer program Praat [16], listeners were presented with two yellow squares on a computer screen, viz. “2” and “3,” and were asked to click on a square that corresponded to the sound that best matched the first sound. Each XAB task contained one of the five contrasts and included 80 trials. Participants were offered four practice trials, which could be repeated once if needed. The X stimuli and the A and B response options had an inter-stimulus interval of 1.2 seconds and there was an inter-trial interval of 500 ms between the participant’s click and the presentation of the next trial. On average, listeners took 32.5 minutes to perform the five AUDI_XAB tasks.

Likewise, AusE listeners were tested at the MARCS Institute at the University of Western Sydney. As for the Spanish listeners, stimuli were presented at a comfortable volume level through headphones. After completing four practice trials, participants partook in the same vowel discrimination experiment as was presented to the Peruvian Spanish listeners. That is, AusE participants preformed the same five naturally produced Dutch vowel contrasts as the Peruvian Spanish listeners. Following the same method used in previous studies [1, 8, 15], participants were advised that they would be listening to AusE vowel tokens in order to avoid any influence upon their natural perception. Data collection for AusE listeners took approximately seven minutes per contrast to complete, and short breaks were provided throughout testing.

3. Results

Participants’ difference in accurate discrimination of vowel contrasts from chance are presented in Figure 2. We were interested in whether native AusE listeners, having a larger vowel inventory than Dutch, could discriminate non-native Dutch vowel contrasts, and how their performance compared to native Spanish listeners, who have a smaller vowel inventory compared to Dutch.

To first test AusE listeners’ discrimination of Dutch vowel contrasts, we tested participants’ correct and incorrect responses in the XAB task in a mixed-effects logistic model, which is appropriate for analyzing data comprised of categorical responses [17, 18, see also 19]. Participant number and XAB trial were included as random effects, and contrast (/ɪ-i/, /a-ɑ/, /y-ʏ/, /i-ʏ/, /ɪ-ʏ/) was included as a fixed effect. This revealed a main effect of contrast ($\chi^2(4, N = 4620) = 14.08, p = .007$). Fisher’s Least Significant Difference (LSD)-corrected post-hoc pairwise comparisons revealed that AusE participants had more correct responses for contrast /ɪ-y/ than for contrast /ɪ-i/ ($p < .001, 95\% CI [0.07, 0.19]$), for contrast /y-ʏ/ ($p = .007, [0.03, 0.18]$), and for contrast /ɪ-ʏ/ ($p = .001, [0.03, 0.12]$). There were also more correct responses for contrast /ɪ-y/ than /ɪ-i/ ($p = .011, [0.01, 0.09]$). There was a trend towards more correct responses for /ɪ-y/ compared to /a-ɑ/ ($p = .074, [0.01, 0.16]$).

To examine whether Spanish listeners discriminated the Dutch contrasts, we similarly analyzed Spanish listeners’ correct and incorrect responses to an XAB task in a mixed-effects logistic model. Participant number and XAB trial were included as random effects, and contrast as a fixed effect. This revealed a main effect of contrast ($\chi^2(4, N = 4620) = 38.92, p < .001$). Fisher’s LSD-corrected post-hoc pairwise comparisons revealed that Spanish participants had more correct responses to the contrast /ɪ-y/ than contrast /ɪ-ʏ/ ($p < .001, [0.07, 0.15]$) and contrast /a-ɑ/ ($p = .013, [0.02, 0.20]$), and also had more correct responses to the contrast /ɪ-ʏ/ than the contrast /ɪ-i/ ($p < .001, 95\% CI [0.06, 0.16]$). There was a trend towards more correct responses for contrast /ɪ-y/ than contrast /ɑ-ɑ/ ($p = .056, [0.00, 0.23]$), and towards more correct responses for contrast /ɪ-i/ than contrast /ɪ-ʏ/ ($p = .099, [-0.01, 0.15]$). Although our participants comprised a sub-sample of those tested in [8], and data in [8] were analysed using a Repeated-Measures ANOVA, these results were nonetheless comparable, as the resulting difficulty ranking, from more difficult to less difficult, was the same across the two studies: (1) /a-ɑ/~/ɪ-ʏ/ >> (2) /ɪ-i/ >> (3) /ɪ-y~/~y-ʏ/.

![Figure 2. Difference in accurate discrimination of non-native Dutch vowel contrasts from chance (50%), by native AusE and native Spanish participants. Error bars represent one standard error.](image)

To address our second research question of whether discrimination of the Dutch vowel contrasts would differ across participants whose native language had more (AusE) or fewer (Spanish) vowels compared to Dutch, we compared AusE listeners’ correct and incorrect responses to the XAB task to the subset of Peruvian Spanish listeners from [8] described in section 2.2. This analysis comprised a second mixed-effects logistic model with participant and XAB trial as random effects and vowel contrast and language background as fixed effects. This revealed a main effect of contrast ($\chi^2(4, N = 9240) = 28.9, p < .001$). There was no main effect of language background ($\chi^2(1, N = 9240) = 0.33, p = .568$). However, there was an interaction of vowel contrast and language background ($\chi^2(4, N = 9240) = 13.2, p = .010$). Thus, participants’ discrimination of the vowel contrasts differed based on their language background and the particular contrast. Fisher’s LSD-corrected post-hoc pairwise comparisons revealed that AusE listeners had more correct responses for the contrast /ɪ-ʏ/ than Spanish listeners ($p = .008, [0.02, 0.14]$).

4. Discussion

The aim of the present study was to directly compare the non-native vowel discrimination of two groups of listeners, one with a larger and one with smaller vowel inventory than that of the target language. Specifically, we tested whether AusE participants, whose L1 has 18 vowels, would be better at
discriminating Dutch vowel contrasts than Peruvian Spanish listeners, who have only five vowels in their inventory.

The results showed that AusE listeners’ discrimination of the Dutch vowel contrast /i/-/y/ was better than for the /i/-/ɪ/ contrast. AusE listeners also had more correct responses to the /ɪ/-/y/ contrast compared to /i/-, and /i/-/ɪ/ compared to /i/-/y/ respectively. A trend towards more correct responses for /i/-/ɪ/ compared to /a/-/a/ was also observed and trended towards better performance compared to the /y/-/ɪ/ contrast. Spanish listeners had more correct responses in the XAB task for the contrast /ɪ/-/y/ in comparison to /i/-/ɪ/ and /a/-/a/ and also showed a larger number of accurate responses to the contrast /i/-/y/ compared to /i/-/ɪ/. Trends for better discrimination were observed for the /i/-/y/ and /i/-/ɪ/ contrasts in comparison to /a/-/a/ and /i/-/ɪ/, respectively. Furthermore, we found that while language background did not affect discrimination of Dutch vowel contrasts overall, there was an interaction between language background and contrast. Specifically, AusE listeners were better at discriminating the /i/-/ɪ/ contrast compared to Spanish listeners.

This last finding suggests that perhaps the vowel inventory of the learners’ L1 in relation to their L2 may affect the ability to learn some non-native vowel contrasts, as AusE listeners, whose vowel systems contain more vowels than Dutch, performed better on the contrast compared to Spanish listeners, whose vowel inventory contains fewer vowels compared to Dutch. However, this advantage was not seen for all of the Dutch contrasts that are also not present in Spanish, which indicates that having more L1 vowels does not always yield better discrimination performance.

The present findings suggest that there may be a language background advantage in perception of non-native vowel contrasts. Further research conducted in our lab will help determine whether vowel inventory size indeed affects non-native vowel perception in a larger population sample.

5. References


