Is more always better? The perception of Dutch vowels by English versus Spanish listeners

Samra Alispahic, Paola Escudero & Karen E. Mulak

The MARCS Institute, University of Western Sydney, Australia

s.alispahis@uws.edu.au,paola.escudero@uws.edu.au,k.mulak@uws.edu.au

Abstract

The present study investigates Australian English (AusE) monolingual listeners' perception of non-native vowels in Dutch, a language with fewer vowels compared to AusE. AusE listeners' performance was compared to native monolingual Peruvian Spanish (PS) listeners whose L1 contains fewer vowels than Dutch. Results show that compared to PS listeners, AusE listeners were better at discriminating only one contrast, /ɪ -ʏ /. AusE listeners took full advantage of their vowel inventory size by mapping Dutch vowels across multiple English categories. Surprisingly, they also appear to duplicate the PS single category assimilation by mapping Dutch /ɪ / and /i/ to both English /i/ and /ɪ /.

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

1. Introduction

Second-language (L2) learners frequently struggle with the perception and production of novel L2 sounds. Vowels are particularly difficult to master, due in part to native accent variation found in vowel inventories across different languages. The initial state of L2 learning is comparable to that of naïve listeners [1], as exemplified in the Second Language Linguistic Perception model (L2LP) [2] as well as the extension of the Perceptual Assimilation Model to L2 acquisition (PAM-L2) [3]. Difficulties in L2 acquisition may occur as a result of discrimination of acoustic and/or phonological distinctions not present in a listener's native language (L1). L2LP and PAM-L2 propose that naïve listeners will perceive L2 contrasts in line with the features of their native phonemes, which in turn predict attainment and differentiation of contrasts during L2 acquisition. For example, if two non-native contrasts are perceived as belonging to two separate native categories, they will be easier to distinguish and acquire than if they are perceived within a single native category [4].

Furthermore, the L2LP model stipulates that a learner's L2 speech perception (and production) is influenced by the sounds of their L1 vowel inventory [2]. Learning new vowel categories would be required by learners whose L1 has fewer categories than the target language [5], whereas unlearning certain categorieswould be required if the native language has a greater number of categories [6]. The model states that unlearning or adjusting existing categories is easier than learning new ones.

Recent studies suggest 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 inventories are smaller than that of English. 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 /a/, /aɪ /, and /aʊ / were all assimilated to the Spanish and French /a/ equivalent). Likewise, PS listeners, whose vowel inventory is smaller than that of Dutch, also have substantial difficulty discriminating the Dutch contrasts /a–ɑ /, /i–ɪ / and /y–ʏ /, as these contrasts are not present in PS [8]. Consequently, in order to avoid misidentifying non-native vowel categories, these listener groups would need to learn new categories. However, the discrimination of Dutch vowel contrasts by PS listeners has not been compared to those of a listener group whose L1 vowel inventory is larger in size.

The present study contributes to this line of research by comparing Peruvian Spanish (PS) and Australian English (AusE) listeners' discrimination and categorization of Dutch vowels. Standard Northern Dutch (SND) has 8 monophthongal vowels (/ɑ , ɛ , ɪ , i, ɔ , u, ʏ , y/), 4long mid vowels (/a, e, o, ø/) and 3 diphthongal vowels (/ɛ ɪ , ɔ u, œy/) [9]. Compared to Dutch, AusE has an additional 3monophthongs and 2 diphthongs, for a total of 12 monophthongs (/i□ , ɪ , e, e□ , æ, ɐ □ , ɐ , ɔ , o□ , ʊ , ʉ □, ɜ □ /) and 6 diphthongs (/æɪ , ɑ e, æɔ , əʉ , oɪ , ɪ ə/) [10]. Comparatively,PS has a smaller inventory, comprising5 vowels (/a, e, i, o, u/) [8]. Figure 1 shows the F1 and F2 acoustic values for the monophthongs of these three languages.

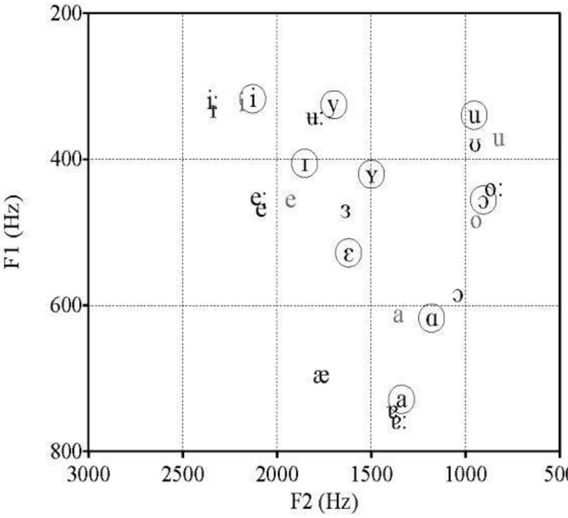

Figure 1. *Average F1 and F2 acoustic values for vowels produced by a male native speaker of Australian English (in black: [10]), Northern Standard Dutch (circled: [11]), and Peruvian Spanish (grey: [12]).*

According to the L2LP model, comparison between the acoustic properties of listeners' L1 vowels and the target vowels can predict L2 acquisition. If our findings are in line with the vowel inventory predictions and past research [7, 13], AusE listeners, having a larger vowel inventory than that of Dutch, should outperform PS listeners, whose vowel inventory is smaller than that of Dutch, in their discrimination of Dutch vowel contrasts. Similarly, having a larger vowel inventory should aid AusE listeners' categorization of Dutch vowels. Additionally, as proposed by the L2LP model, the acoustic properties of listeners' L1 vowels inventories will influence their categorization and discrimination patterns.

As shown in Figure 1, AusE and PS have very different F1 and F2 values. In their perception of Dutch vowels, AusE listeners are expected to easily discriminate the Dutch contrast /a-ɑ /, due to the Dutch vowel /a/ being acoustically closer to AusE /ɐ / and /ɐ □ / thantoDutch /ɑ /, whose F1 and F2 values lie closer to AusE /ɔ /. Conversely, PS listeners may have difficulty with this contrast as Dutch /ɑ / is not present in the PS vowel inventory. However, the F1 and F2 values for the PS /a/ are more similar to Dutch /ɑ / than to Dutch /a/.

The AusE vowels/ɪ / and /i/ and the PS vowel /i/ have higher F1 and F2 values compared to Dutch /ɪ /, which is instead more comparable to AusE and PS /e/. As PS has only one /i/ vowel and AusE /ɪ / and /i/ have similar acoustic values, both listener groups should have comparable difficulty discriminating the Dutch /ɪ -i/ contrast.

We also predict that PS listeners will have difficulty discriminatingthe Dutch vowels /ɪ /, /y/, and /ʏ /,as these are not present in the PS vowel inventory [8]. However, even though /ɪ / and /ʏ / have a high F1 and low F2 value, they are acoustically most similar to PS /e/. On the other hand, AusE /ʉ / is acoustically more similar to Dutch /y/ than AusE /e/ is, with AusE /e/ more comparable to Dutch /ɪ /, as mentioned above. Even though Dutch /ʏ / has higher F1 values compared to AusE /ɜ /, its F2 value is lower than that of Dutch /ɪ / or /y/, making it more comparable to AusE /ɜ / and allowing for easier discrimination based on the acoustic properties.

These differences in acoustic variation can affect how Dutch vowels are perceived by AusE and PS listeners. Based on the stipulations put forth by the L2LP model [2], we further predict that the acoustic properties of both AusE and PS listeners' native vowel inventories will affect discrimination of the Dutch vowel contrasts. We thus examined to what extent the L1 vowel inventory sizes and vowel acoustic properties affect listeners' discrimination and perceptual assimilation patterns by directly comparing AusE and PS listeners' XAB discrimination of Dutch contrasts, and comparing AusE listeners' performance in a Dutch vowel categorization task to reported results from PS participants who completed the same task.

2. Method

2.1 Listeners

AusE listeners were seven female and four male students aged 18-45 years (*M* = 24.19) attending the University of Western Sydney. Participants were born and raised in Greater Western Sydney, and reported little to no experience with languages other than English, and no prior experience with Dutch.

PS listeners for the XAB task were a subset of 11 participants randomly selected from a prior study [8]. Listeners were monolingual PS 18- to 28-year-olds (*M* = 20.95) who had lived in Lima, Peru their entire life.They reported no prior knowledge of Dutch or English.

2.2 Stimuli

Stimuli for the XAB task were 20 tokens of each of the 5 Dutch vowels /a/,/ɑ /,/ɪ /,/i/,/y/,and/ʏ /. For the categorization task, stimuli were 20 tokens of each of the 12 Dutch vowels /ɑ , a, e, ɛ , ɪ , i, ɔ , o, ø, u, ʏ , y/. These stimuli were used in previous studies examining the identification and discrimination of Dutch vowels [1, 15], and were extracted from corpus recordings [9] of 10 male and 10 female native speakers of Northern Standard Dutch who produced the vowels in a non-word /sVs/ context embedded in a carrier sentence.

.

2.3 Procedure

AusE listeners were tested at the MARCS Institute at the University of Western Sydney. Following the same method used for testing PS listeners [1, 8, 15], participants completed an XAB task and a categorization task, and were advised that they would be listening to AusE vowel tokens in order to avoid any influence upon their native perception. Data collection for both tasks took approximately 1 hour to complete and short breaks were provided throughout testing.

As reported in [8], PS listeners took part in two forced-choice discrimination XAB tasks in Lima, Peru. We only report on the data for their first task, which was the same auditory discrimination XAB task presented to AusE listeners. In this task, PS and AusE listeners were presented with the five naturally produced Dutch vowel contrasts, /a-ɑ /, /ɪ -i/, /y-ʏ /, /i-ʏ / and /ɪ -ʏ /.

For the XAB task, participants heard three sounds, which were presented at a comfortable volume level through headphones, and were asked to decide if the first sound (X) was more like the second (A) or more like the third (B). 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 s and an inter-trial interval of 0.5s between the participant's click and presentation of the next trial.

To determine how the acoustic properties of both listener groups impacted discrimination of the Dutch vowels, the same AusE listeners were presented with a second task in which they categorized the Dutch vowels. We compared their results to those of another set of 20 male and 20 female PS listeners reported in [1] who completed the same task as the AusE listeners. AusE and PS listeners completed a forced choice identification task in which participants were presented with one of the 12 Dutch vowels mentioned above, and asked to match the sound they heard either with one of the 12 English words, *heed, hid, hood, who'd, hair, head, heard, hall, had, hut, hot, hard* (AusE listeners) or with vowels (PS listeners) presented on the screen.

3. Results

Figure 2 presents participants'difference in accurate discrimination of the Dutch vowel contrastsfrom chance. 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 PS listeners who have a smaller vowel inventory compared to Dutch.Further,we were interested inhow the acoustic properties of both listener groups impacted their perception of Dutch vowels.

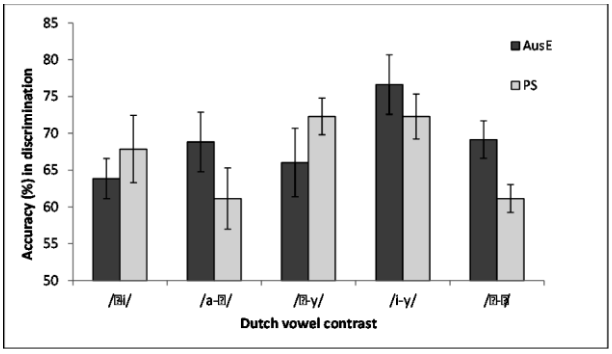

Figure 2.*Discrimination accuracy for five Dutch vowel contrasts by native AusE and native PS listeners. Error bars represent one standard error of the mean.*

We used a mixed-effects logistic model[17, 18, see also 19] to compare accuracy across contrasts. Participant and XAB trial were included as random effects, and contrast (/ɪ -i/, /a-ɑ /, /ʏ -y/, /i-y/, /ʏ -ɪ /) was included as a fixed effect.

For AusE listeners, the model revealed a main effect of contrast (χ^2(4, *N* = 4620) = 14.08, *p* = .007). Fisher's Least Significant Difference (LSD)-corrected post-hoc pairwise comparisons revealed that AusE listeners had more correct responses for contrast /i-y/ than for /ɪ -i/ (*p*< .001, 95% CI [0.07, 0.19]), /ʏ -y/ (*p*= .007, [0.03, 0.18]), and /ɪ -ʏ / (*p* = .001, [0.03, 0.12]. Thus, AusE listeners' ranking from most difficult contrast to the least difficult is: (1) /ɪ -i/, /ʏ -y/, /a-ɑ /, /ʏ -ɪ / and /i-y/.

A second mixed effects logistics model also revealed a main effect of contrast (χ^2(4, *N* = 4620) = 38.92, *p*< .001) forPS listeners. Fisher's LSD-corrected post-hoc pairwise comparisons revealed that PS participants had more correct responses to the contrast /ʏ -y/ than /ɪ -ʏ / (*p*< .001, [0.07, 0.15]) and /a-ɑ / (*p* = .013, [0.02, 0.20]), and also had more correct responses to /i-y/than /ɪ -ʏ / (*p*< .001, 95% CI [0.06, 0.16]). Compared to/a-ɑ /, /i-y/ showed a trend towards more correct responses (*p* = .056, [0.00, 0.23]).A trend towards more correct responses for /ɪ -i/compared to/ɪ -ʏ / was also observed (*p*= .099,[-0.01, 0.15]). Although our participants comprised a subset of those tested in [8], and data in [8] were analyzed 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) /i-y/~/y-ʏ /.

To address our second research question of whether discrimination of the Dutch vowel contrasts differed across participants whose native language had more (AusE) or fewer (PS) vowels compared to Dutch, we compared AusE listeners' correct and incorrect responses of the XAB task to the subset of PS listeners from [8] described in section 2.2.

A mixed-effects logistic model with participant and XAB trial as random effects and vowel contrast and language background as fixed effects revealed a main effect of contrast (χ2(4,*N* = 9240) = 28.9, *p* = < .001). There was no main effect of language background (χ2(1, *N* = 9240) = .326, *p* = .568). However, there was an interaction of vowel contrast and language background (χ2(4, *N* = 9240) = 13.2, *p* = .010). Thus, participants' discrimination of the vowel contrasts differed based on their language background and the particular Dutch contrast. Fisher's LSD-corrected post-hoc pairwise comparisons revealed that AusE listeners had more correct responses than PS listenersfor the contrast /ɪ -ʏ / (*p* = .008, [.02, .14]).Table 3 shows the percentage of times a Dutch vowel was classified as an AusE vowel. The L2LP model states that initially, a learner will perceive and categorize sounds of their target L2 in line with the sound inventory of their first language. Taking into consideration that the PS vowel inventory is smaller than both the AusE and Dutch vowel inventories, we are able to observe similar yet also different assimilation patterns between the two groups.

AusE listeners made full use of their vowel inventory when classifying the Dutch vowels, often mapping a single Dutch vowel to two or three English vowel counterparts, e.g. /ʏ / → /ɪ /, /e/ and /ʊ /. Likewise, participantsmappedthe Dutch /y/ vowel to the AusE /ʊ / and /ʉ / vowels, the Dutch/o/ vowel to AusE /o□/ and /ɔ /, and the Dutch /e/ to the AusE/e□/ and /e/ vowels. PS listeners also categorized tokens of Dutch, /a, e, ɛ , i, ɔ , o, u/, to two or more PS vowel categories, namely, /i-e, i-u, e-u, e-ei-eu/ [1].
However, AusE participants appear to duplicate PS listeners' single category assimilation [1], for instance, by mapping Dutch /ɪ / and /i/ to both English /i/ and /ɪ /. This pattern can

Dutch Stimuli	English Vowels											
	iː	ɪ	æ	ʊ	ʉ	ɐː	eː	e	ɜ	oː	ɔ	ɐ
y	5.9	17.3	0.0	23.6	34.1	0.9	2.3	5.0	0.9	0.5	5.0	4.5
ɛ	5.5	9.1	4.5	2.3	0.5	3.6	16.4	41.8	9.1	1.4	1.4	4.5
o	3.2	0.5	0.9	16.8	16.4	2.3	0.0	2.7	5.9	29.5	21.8	0.0
ɔ	0.0	1.4	0.5	14.5	5.9	2.7	0.9	3.2	0.5	12.3	52.3	5.9
e	15.0	3.6	11.8	2.3	3.2	5.5	21.8	23.2	6.4	6.8	0.5	0.0
ɪ	7.7	46.8	1.8	5.0	6.4	1.4	4.1	16.8	6.4	0.5	1.4	1.8
ɑ	0.9	4.1	7.3	2.7	1.8	8.6	3.6	2.3	2.7	6.8	41.4	17.7
ʏ	3.2	16.8	2.3	17.7	15.0	2.3	6.8	20.5	6.8	1.4	3.2	4.1
i	24.1	53.2	0.5	2.3	2.7	0.0	0.9	8.2	3.6	0.9	1.8	1.8
u	0.0	2.7	0.0	50.0	29.5	2.3	0.5	0.5	2.7	0.5	7.7	3.6
ø	3.6	0.0	3.6	10.5	19.5	3.6	5.9	10.5	31.4	2.3	6.4	2.7
a	0.9	0.5	15.0	2.3	1.4	45.5	10.5	4.5	5.9	5.5	1.4	6.8

Table 3. Percentage categorization of Dutch vowels as English vowels by 11 AusE participants.

be observed inthe mapping of Dutch vowels /ɔ / and /ɑ / to AusE /ɔ /, which lies acoustically between the two Dutch contrasts. Much like PS listeners [1], AusE listeners also assimilated some Dutch vowels to the AusE counterparts with the closest F1/F2 category values, specifically /ɛ , u, ø, a/ to the AusE/e, ʊ , ɜ , ɐ /.

4. Discussion

The present study directly compared the non-native vowel perception of two listener groups, whose vowel inventories differ in size. Specifically, AusE participants, whose L1 has 18 vowels, were tested to see whether their Dutch vowel perception would be better than that of PS listeners, who haveonly 5vowels in their inventory.

While language background did not affect discrimination of Dutch vowel contrasts, there was an interaction between language background and contrast. Specifically, AusE listeners were better at discriminating the /ɪ -ʏ / contrast compared to PS listeners.While this advantage was not seen for all Dutch contrasts, which are also not present in PS, it does suggest that the vowel inventory of the learners' L1 in relation to their L2 may affect the ability to learn some non-native vowel contrasts.

However, while PS listeners' discrimination of the /a-ɑ / contrast may not be very surprising, what is surprising is that the AusE participants found the /i-ɪ / contrast to be most difficult. Both groups assimilated these contrasts to a single category: /i/ for AusE participants and /a/ for PS [1] participants. These results are in line with L2LP which states that the more a contrast assimilates to a single L1 category, the more problematic it will be for a listener to discriminate and acquire.

The present findings suggest that there may be only a small advantage in the non-native vowel perception for listeners whose L1 has a larger vowel inventory than that of the target language. However, further research with larger listener groupsneeds to be conducted to discern whether variation across listener groups for the present study remains the same.

5. References

[1] Escudero, P., and Williams, D., "Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners", J. Acoust. Soc. Am., 129(1): EL1-EL7, 2011.

[2] Escudero, P., "Linguistic perception and second-language acquisition: explaining the attainment of optimal phonological categorization", PhD thesis, Utrecht University, The Netherlands, 2005.

[3] Best, C. T., and Tyler, M. D., "Non-native and second-language speech perception: Commonalities and complementarities", in O.-S. Bohn and M. Munro [Eds], Language experience in Second Language speech learning: in Honor of James Emil Flege, 13-34, John Benjamins, Amsterdam, 2007.

[4] Best, C. T., "The Emergence of Native-Language Phonological Influences in Infants : A Perceptual Assimilation Model", In J. C. Goodman and H.C. Nusbaum, Eds., The Development of Speech Pereption: The transition from Speech Sounds to Spoken Words, Cambridge, MA: MIT Press, pp.167-224, 1994.

[5] Escudero, P., and Chládková, K., "Spanish listeners' perception of American and Southern British English vowels", J. Acoust. Soc. Am., 128(5): EL254-EL260, 2010.

[6] Escudero, P., and Boersma, P., "The subset problem in L2 perceptual development: Multiplecategory assimilation by Dutch learners of Spanish," in Proceedings of the 26th Annual Boston University Conference on Language Development, edited by B. Skarabela, S. Fish, and A. Do (Cascadilla, Somerville), pp. 208–219, 2002.

[7] Iverson, P., & Evans, B. G., "Learning English vowels with different first-language vowel systems: perception of formant targets, formant movement, and duration." The Journal of the Acoustical Society of America, 122(5), 2842–54.doi:10.1121/1.2783198, 2007.

[8] Escudero, P., &Wanrooij, K., "The effect of L1 orthography on non-native vowel perception," Lang. Speech 53, pp. 343–365, 2010.

[9] Adank, P., Van Hout, R., & Smits, R., "An acoustic description of the vowels of northern and southern standard Dutch," J. Acoust.Soc. Am. 116, 1729–1738, 2004.

[10] Cox, F., and Palethorpe, S., "Illustrations of the IPA," Journal of the International Phonetic Association, 37/3, pp.341-350. doi:10.1017/S0025100307003192, 2007.

[11] vanLeussen, J. W., Escudero, P., & Williams, D., "Acoustic properties of Dutch steady-state vowels: Contextual effects and a comparison with previous studies." Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, pp. 1194-1197, 2011.

[12] Chládková, K., Escudero, P., and Boersma, P., "Context-specific acoustic differences between Peruvian and Iberian Spanish vowels", J. Acoust. Soc. Am., 130(1): 416-428, 2011.

[13] Iverson, P., & Evans, B. G., "Learning English vowels with different first-language vowel systems II : Auditory training for native Spanish and German speakers a)". The Journal of the Acoustical Society of America, 126, 865. doi:10.1121/1.3148196, 2009.

[14] Adank, P., Van Hout, R., & Smits, R., "An acoustic description of the vowels of northern and southern standard Dutch," J. Acoust.Soc. Am. 116, 1729–1738, 2004.

[15] Escudero, P., & Williams, D., "Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch", Journal of the Acoustical Society of America, 131 (5), 406-412.doi:10.1121/1.3701708, 2012.

[16] Boersma, P., and Weenink, D., "PRAAT: Doing phonetics by computer (Version 5.2.26)", [Computer program], retrieved May 2011. Online: http://www.praat-org/ (University of Amsterdam, Amsterdam) 1992-2011.

[17] Baayen, R. H., Davidson, D. J., & Bates, D. M., "Mixed-effects modeling with crossed random effects for subject and items," J. Mem. Lang. 59, 390-412. doi: 1.1016/j.jml.2007.12.005, 2007.

[18] Jaeger, T. F. "Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models," J. Mem. Lang. 59, 434-446. doi: 10.1016/j.jml.2007.11.007, 2007.

[19] Arnon, I. "Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew," J. Child Lang. 37, 27-57.doi: 1.1017/S030500090900943X