The relationship between Australian English speakers' non-native perception and production of Brazilian Portuguese vowels

Jaydene Elvin[1,2], Paola Escudero[1,2], Daniel Williams[1,2,4] and Catherine T. Best[1,2,3]

[1]The MARCS Institute , Western Sydney University, Australia
[2]ARC Centre of Excellence for the Dynamics of Language, Australia
[3]School of Humanities and Communication Arts, Western Sydney University, Australia
[4]Linguistics Department, Area of Excellence – Cognitive Sciences, University of Potsdam, Germany

{j.elvin, paola.escudero, c.best}@westernsydney.edu.au, daniel.williams@uni-potsdam.de

Abstract

This study investigates the relationship between non-native perception and production of Brazilian Portuguese (BP) vowels by six Australian English (AusE) monolinguals. Participants' non-native categorization and discrimination patterns were used to predict performance in non-native production. We further investigated the acoustic similarity between participants' non-native and native vowel productions. The findings indicate a perception-production link. In particular, non-native vowel production was acoustically similar to native vowel production. We also found that perceptually difficult non-native vowel contrasts were predominantly produced as one single non-native category.

Index Terms: non-native perception, non-native production, interrelation

1. Introduction

One of the main goals of a second language (L2) learner is to be able to speak the L2 in a native-like manner. Unfortunately, for many, this goal is not achieved [1]. This is because factors such as the age of L2 acquisition, length of residence in an L2 environment, language experience and L1 vs. L2 use [2] all contribute to the degree of foreign-accented speech. Common to all of the factors is the influence of an individual's native language (L1) on their L2 pronunciation.

Problems in L2 pronunciation are thought to be caused by their L2 perceptual difficulties. Models of L2 speech perception such as the Speech Learning Model [SLM; 3], the Perceptual Assimilation Model [PAM; 4] and the Second Language Linguistic Perception model [L2LP; 5,6] all claim that L2 speech sounds are filtered by listeners' L1. For example, two non-native sounds that are perceived as one single category are known as single-category assimilation in PAM and as the NEW scenario in L2LP. A scenario of this type has been shown to result in perceptual discrimination difficulties, whereas no discrimination difficulties are found for scenarios where two non-native sounds are mapped on to two separate native categories (two-category assimilation in PAM and the SIMILAR scenario in L2LP). When two sounds in a non-native contrast are mapped on to two or more native categories (uncategorised assimilation in PAM and the SUBSET scenario in L2LP) discrimination difficulty may occur if these non-native vowels are mapped to the same overlapping native categories.

While the above scenarios predict difficulty in non-native or L2 perception, the L2LP model posits that these learning scenarios will also influence an individual's production of non-native or L2 vowels. Although the L2LP model is not the only model to suggest this (e.g. SLM investigates the limitations of a learner's ability to perceive and produce native-like sounds due to experience and age-related limitations), it is the only model to account for a relationship between perception, spoken word recognition and production at all stages of development. According to the L2LP theoretical framework, at the onset of learning, L2 production will closely match the acoustic properties of sounds produced in the speaker's L1. L2 pronunciation develops as learners adjust their perceptual mappings to match those of the L2 with the help of their lexical (word) representations.

Studies investigating the perception-production relationship have produced mixed results. A number of studies [1, 6] have indeed identified a link between perception and production and suggest that native language perceptual patterns do seem to influence L2 production patterns. However, there is a long-standing debate regarding the perception-production link which is unresolved due to mixed empirical results and the problematic nature of testing this link [7]. In particular, there are a number of methodological reasons that make the investigation of the interrelation difficult. For example, there are different task demands with different techniques of analysis. Additionally, previous interrelation studies have tended to focus on groups rather than individuals [6]. In order to appropriately investigate the relationship between perception and production, [7] suggest that data should ideally be collected from the same participants performing both tasks.

The aim of the present study is to investigate the relationship between the perception and production of non-native Brazilian Portuguese (BP) vowels by Australian English (AusE) monolinguals. In particular, we will investigate whether our participants' non-native categorization patterns and discrimination difficulties influence the way in which these same vowels are produced in the non-native production task. We control for the methodological issues suggested by [7] by using the same participants and the same stimuli across all tasks, as well as collecting native vowel production data to compare with their non-native vowel productions.

The L2LP theoretical framework is the most applicable to the present study as it specifically accounts for learners, such as our AusE participants who have no prior experience with the target language. If the findings are consistent with the L2LP theoretical framework, we would expect that our participants should produce non-native BP vowels similarly to the productions of their own closest L1 vowel categories and that participants will have difficulties producing vowels that are similar to two separate BP categories for those contrasts that are difficult to discriminate.

2. Method

2.1. Participants

This paper reports a subset of six AusE monolinguals (3 male) from [8] which is a study that investigated the non-native perception of BP vowels. They were all Australian English speakers born and raised in Western Sydney and aged between 18 and 30. The AusE participants reported little to no knowledge of any foreign language and provided informed consent.

2.2. Stimuli and procedure

We first recorded participants' productions of the 13 AusE vowels, namely, /iː, ɪ, ɪə, e, eː, ɜː, ɐ, ɐː, æ, oː, ɔ, ʊ and ʉː/ produced in the fVf context. The auditory stimuli were the same for the non-native categorization, discrimination and repetition tasks and consisted of naturally produced BP pseudo-words in the fVfe context. Target words were produced by five male and five female speakers from São Paulo, selected from the [9] corpus. The vowel in the first syllable was always stressed and corresponded to one of the seven Portuguese target vowels /i, e, ɛ, a, o, ɔ, u/. The perception and production tasks were counterbalanced. In perception, participants first completed a 2 alternate forced choice task in an XAB format, followed by a non-native categorization task. A non-native repetition task was used to elicit production data as it was the most appropriate task for monolingual participants with no experience with Brazilian Portuguese.

2.3. Data analysis for native and non-native vowel production

WebMaus [10], an online tool for automatically segmenting and labelling speech sounds was used to segment both the native and non-native vowel productions. The automatically generated start and end boundaries were checked and manually adjusted to ensure accuracy. Vowel duration was measured as the time (ms) between these start and end boundaries. Formant measurements for each vowel token were extracted at three time points (25%, 50%, 75%) following the optimal ceiling method reported in [9]. In the optimal ceiling method, for every vowel, per speaker, the "optimal ceiling" is chosen as the one that yields the least amount of variation for both the first and second formant values within the set number of annotated tokens for the vowel. Formant ceilings ranged between 4500 and 6500 Hz for females and 4000 and 6000 for males.

3. Results

3.1. Non-native perception

3.1.1. Non-native categorization

Table 1 shows the results from the non-native categorization task. Both BP /i/ and /e/ were perceived as AusE /iː/ and /ɪ/, with BP /e/ also being categorized as AusE /ɪə/ and /eː/. BP /ɛ/ was mostly classified as AusE /eː/, with some tokens also being categorized as AusE /æ/. Categorization for BP /a/ was split between AusE /æ/ as well as /ɐː/. Most tokens of BP /o/ were perceived as AusE /oː/, with a smaller amount also being classified as AusE /ʊ/. Likewise, BP /ɔ/ was generally perceived as AusE /oː/ and finally, BP /u/ was predominately categorized as AusE /ʊ/. In regards to predicting difficulty in discrimination and non-native production based on these categorization patterns, we would expect that BP /a-ɔ/ and /a-ɛ/ should be easier to discriminate and produce than the remaining contrasts as the target vowels in these contrasts were mapped to different native categories resulting in little to no perceptual overlap. The non-native categorization patterns further suggest that participants may have difficulties discriminating and producing BP /i/-/e/ and /o/-/u/ as both vowels are mapped to the same native categories causing a large amount of perceptual overlap. Additionally, the vowels in the BP contrasts, /e/- /ɛ/ and /o/- /ɔ/ are mapped to some of the same native categories and this may make these contrasts difficult to discriminate and produce.

Table 1: *Classification percentages for the six AusE listeners. The native vowel category with the highest classification percentage appears in bold and those percentages below chance (i.e. 8%) appear in grey.*

	AuE vowels												
BP	**iː**	**ɪ**	**ɪə**	**e**	**eː**	**ɜː**	**æ**	**ɐː**	**ɐ**	**ɔ**	**oː**	**ʊ**	**ʉː**
i	**48**	38	7	5				2					
e	**27**	13	22	5	**27**		3	3					
ɛ	7	2	5	3	**53**	2	20	8					
a			2		3	3	**43**	**43**			5		
o						7			8	7	**62**	15	2
ɔ					3	10		2		3	**75**	2	5
u			2			8			15	2	7	**45**	22

3.1.2. Non-native discrimination

Figure 1 shows the average accuracy scores for each of the six BP contrasts.

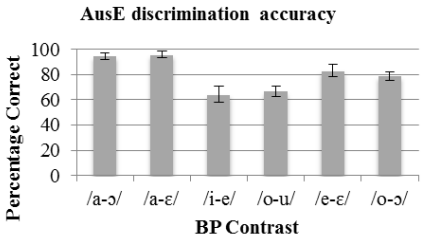

Figure 1: *Discrimination accuracy across the six BP contrasts*

The discrimination results indicate that AusE listeners had higher performance on BP /a-ɔ/ and /a-ɛ/, which is unsurprising given that there was little to no perceptual overlap for these contrasts. We would therefore predict that participants should be able to produce two distinct vowels that are similar to two different BP vowels. Furthermore, in line with the non-native categorization predictions, participants did indeed have overall lower discrimination accuracy scores for BP /i/-/e/ and /o/-/u/ and we would therefore predict that participants may produce both vowels in each contrast as the same non-native vowel category. Additionally, BP /e/-/ɛ/ and /o/-/ɔ/ were not the easiest to discriminate, however they were also not the hardest to discriminate. It is possible that participants may still have difficulty producing these vowels given perceptual overlap identified in non-native categorization, but perhaps they may not be as unstable as those contrasts with the lowest discrimination accuracy.

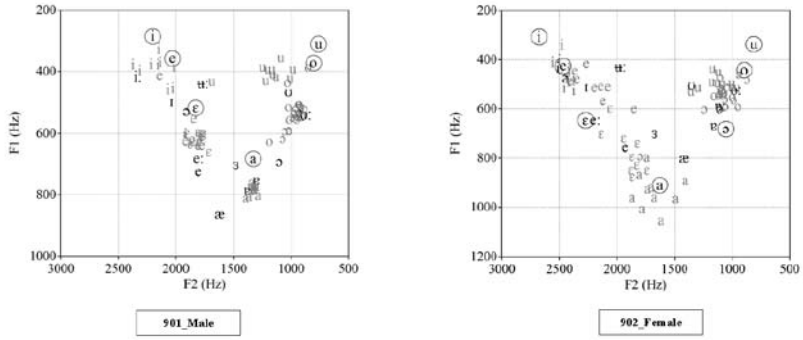

Figure 2: *A sample of one male and one female participants' F1 and F2 values for all tokens of their non-native productions of the BP vowels (grey) are displayed against the mean values for their native AusE vowel productions (black), and for the target BP vowels (black, with circles).*

3.2. Non-native production

3.2.1. Acoustic similarity between non-native vowel productions and native vowel production

Figure 2 shows an example of the F1 and F2 values for one male and one female participant's non-native production of the 7 BP vowels, together with the averaged F1 and F2 values of their own native vowel productions and the target BP vowels. We show these example vowel plots to demonstrate how non-native production of the BP vowels varies across participants and that they have not yet formed stable vowel categories for each vowel. In fact, it appears that at the initial stage of learning, non-native vowel productions are indeed acoustically more similar to their own L1 vowel categories.

Table 2: *Predicted group membership for participants' non-native vowel productions tested on AusE vowel productions. The native vowel category with the highest predicted probability appears in bold and those probabilities below chance (i.e. 8%) appear in grey.*

BP	AusE vowels												
	iː	ɪ	ɪə	e	eː	ɜː	æ	ɐː	ɐ	ɔ	oː	ʊ	ʉː
i	**47**	26	10	1							2		14
e	**24**	18	11	19	6	2							20
ɛ	8	8	7	**24**	23	12	6	1					15
a				2		4	**40**	17	27	8	1	2	
o						2		1	4	23	29	**40**	
ɔ						6	2	4	8	**29**	**29**	23	
u				1		1			1	11	10	**74**	2

Table 2 shows the results of a cross-language discriminant analysis conducted to determine acoustic similarity between participants' non-native productions of BP vowels and their own native vowels. We trained the model on the participants' own native vowel productions and tested it against their non-native productions of BP using F1, F2, F3 and duration as input parameters. Instead of reporting the percentage of times a BP vowel was categorized as a native, we report the probabilities of group membership averaged across the BP vowel tokens. The benefit of reporting probabilities in the present study is that it takes into account that some BP tokens may be acoustically close to more than one vowel, which can be masked by categorization percentages. The results indicate that participants' productions of BP vowels at the initial state are indeed acoustically similar to their own native vowel categories. In particular, BP /i/ and /e/ are acoustically similar to AusE /iː/, /ɪ/, /ɪə/ and /ʉː/. Participants produced BP /ɛ/ as acoustically similar to AusE /e/, /eː/ and /ʉː/. The non-native productions of BP /a/ are acoustically similar to /æ/, /ɐː/ and /ɐ/. We also found that participants produced BP /ɔ/ with equal chance that the vowel would be classified as AusE /ɔ/ and /oː/ with 23% chance the tokens would be classified as AusE /ʊ/. Furthermore, there was a 40% chance that the non-native productions of BP /o/ would be classified as AusE /ʊ/ (with 29% chance for AusE /oː/ and 23% chance for AusE/ɔ/). Additionally, there was a 75% chance that the productions for BP /u/ would be classified as AusE /ʊ/ with an 11% chance for produced similarly to AusE /ɔ/ and 10% for AusE /oː/.

3.2.2. Acoustic similarity between non-native BP vowel production and target BP stimuli

Table 3: *The percentage of non-native vowel tokens classified as the intended BP vowel category. The highest classification percentage appears in bold.*

Non-native ↓	BP vowel category						
	a	e	ɛ	i	o	ɔ	u
a	**80**					20	
e		**49**	46	5			
ɛ	11	14	**70**	2		4	
i		**63**	3	32		2	
o			2		**66**	28	5
ɔ	5		3		32	**60**	
u			8		**67**	2	23

To determine acoustic similarity between the target BP tokens and non-native productions we ran an additional cross-language discriminant analysis with the model trained on target BP vowel tokens and tested on our participants' non-native vowel productions, using the same input parameters from the previous discriminant analysis. The model yielded 85.7% correct classification for the trained BP vowels and 54.1% correct classification for the non-native vowel productions. This suggests that only half of the non-native BP vowel productions were produced close to the target BP vowels. Table 3 shows the percentage of times that each non-native vowel production was correctly classified as the target BP token.

As predicted it seems that participants were able to produce vowels that were acoustically similar to two different vowels for the BP contrasts /a/-/ɛ/ and /a/-/ɔ/ which were perceptually easy to discriminate. There was very little acoustic overlap between the non-native productions of BP /a/ and /ɛ/, with BP /a/ being correctly classified for 80% of the non-native vowel productions and 70% correct classification for BP /ɛ/. Likewise, there is no overlap between BP /a/ and /ɔ/ with 60% correct classification for BP /ɔ/.

We further find that the non-native productions of the BP contrasts with lower discrimination accuracy scores, namely BP /i/-/e/ and /o/-/u/ were indeed less stable and more varied as participants seem to be unsuccessful in producing non-native vowels that were acoustically similar to two separate BP vowel categories. For example, in the BP /i/-/e/ contrast only 32% of the BP /i/ productions were correctly classified as the majority of tokens were classified as BP/e/. Likewise, only 23% of the productions of the BP /u/ vowel were correctly classified, with 67% of the non-native productions of BP /u/ classified as BP /o/. The AusE participants' productions of the vowels in the BP /e/-/ɛ/ and BP /o/-/ɔ/ contrasts was also less stable and more varied. While 49% of the non-native vowel productions of BP /e/ were correctly classified, 46% were incorrectly classified as BP /ɛ/. For BP /o/- /ɔ/, 66% of the non-native vowel productions of BP /o/ were correctly classified, with 28% incorrectly classified as /ɔ/. Furthermore, 60% of the non-native productions of BP /ɔ/ were correctly classified, with 32% being incorrectly classified as BP /o/.

4. Discussion

The aim of the present study was to investigate the relationship between the perception and production of non-native Brazilian Portuguese (BP) vowels by Australian English (AusE) monolinguals. The L2LP theoretical framework posits that learners will initially perceive and produce vowels of the target language in the same manner in which they perceive and produce vowels in their own native language. Our findings support this claim as we found that our participants' perception of BP was influenced by their native language and produced BP vowels similar to their own native categories.

Our findings also support the L2LP claim that perception is linked to and perhaps precedes production. In particular, we found that the non-native contrasts which are easy to perceive, the two vowels in that contrast were produced as similar to two separate BP categories. We also found that non-native vowel production was less stable and more varied for those vowel contrasts which had lower overall discrimination accuracy. This finding suggests that not only do learners struggle to perceive a difference between the two vowels in the BP contrast, but they also find it difficult to produce these vowels as separate categories. Given the particularly low accuracy scores in the discriminant analysis for BP /i/ and /u/ we would expect that speakers will likely have a heavy accent when producing these vowels. We further found that non-native vowels produced in the BP /e/-/ɛ/ and /o/-/ɔ/ contrasts were less stable and more varied, even though these contrasts were not the most difficult to discriminate. It is highly likely that is because their non-native categorization patterns show a considerable amount of acoustic overlap.

While our findings are in-line with the L2LP model, our findings are also consistent with predictions found in the PAM framework as listeners' non-native categorization patterns predicted difficulty in non-native discrimination. Although PAM does not explicitly account for non-native production, the model could be extended to non-native production as there were cases where categorization patterns with perceptual overlap resulted in both discrimination difficulty and less stable non-native vowel production.

In sum, the results from this small-scale study indicate that there is indeed a relationship between perception and production at the initial state of learning. The findings of the present study are in line with the L2LP model which posits that L2 production at the initial state is largely similar to the learner's own native vowel productions. Furthermore, AusE participants also had difficulty producing separate categories for BP contrasts that were perceptually difficult to discriminate. However, our production findings are only based on measurements of acoustic similarity and it would also be beneficial to have native BP speakers rate these tokens to determine the degree of foreign accent in these vowel productions. Finally, the L2LP suggests that L2 production will improve as listeners' perception improves and future studies should test the L2LP claims for L2 development as the participants in the present study were all naïve to BP and therefore representative of the initial state of learning.

5. Acknowledgements

This work was supported by an Australian Postgraduate Award, the MARCS Research Training Scheme and the Australian Research Council Centre of Excellence for the Dynamics of Language (CE140100041).

6. References

[1] Rallo Fabra, L. and Romero, J. "Native Catalan learners' perception and production of English vowels," J. Phon., 40(3), 491–508, 2012.

[2] Piske, T., MacKay, I.R.A. and Flege, J.E. "Factors affecting degree of foreign accent in an L2 : a review," J. Phon., 29(2), 191–215, 2001.

[3] Flege, J.E. "Second language speech learning: Theory, findings, and problems," in Speech perception and linguistic experience: Issues in cross-language research, W. Strange, Ed. Timonium, MD: York Press, 233-276, 1995.

[4] C. T. Best, "A direct realist perspective on cross-language speech perception.," in Speech perception and linguistic experience: Issues in cross-language research, W. Strange, Ed. Timonium, MD: York Press,171-204, 1995.

[5] Escudero, P. "Linguistic Perception and Second Language Acquisition," PhD Dissertation, Utrecht University, 2005.

[6] van Leussen, J.-W. and Escudero, "Learning to perceive and recognize a second language: the L2LP model revised," Front. Psychol., 6, 1–12, 2015.

[7] Levy E.S. and Law F.F. "Production of French vowels by American-English learners of French: language experience, consonantal context, and the perception-production relationship.," J. Acoust. Soc. Am., 128(3), 1290–305, 2010.

[8] Elvin, J., Escudero, P., Williams, D., Shaw, J.A. and Best, C.T. "The role of acoustic similarity, non-native categorisation and individual differences in predicting non-native perception: Brazilian Portuguese vowels by English vs. Spanish listeners," under review

[9] Escudero, P., Boersma, P., Rauber, A. and Bion, R. "A cross-dialect acoustic description of vowels: Brazilian and European Portuguese.," J. Acoust. Soc. Am., 126(3), 1379–93, 2009.

[10] Kisler, T., Schiel, F. and Sloetjes, H. "Signal processing via webservices: the use case WebMAUS," in Proc. of Digital Humanities, 30-34, 2012.