The Interrelation between the Perception and Production of English Vowels by Native Speakers of Brazilian Portuguese

Andréia S. Rauber, Paola Escudero, Ricardo A. H. Bion, Barbara O. Baptista

Graduate Program in Applied Linguistics
Federal University of Santa Catarina, Florianópolis, Brazil &
University of Amsterdam, Amsterdam, The Netherlands
asrauber@hotmail.com, Paola.Escudero@uva.nl, bion@cce.ufsc.br, b_baptista@terra.com.br

Abstract

This study investigated the relationship between the perception and production of English vowels by highly proficient Brazilian EFL speakers. Two experiments were carried out: A production test to measure the first two formants of the learners' English and Brazilian Portuguese (BP) vowels, and an oddity discrimination test to investigate the L2 (second language) categorial perception of English vowels. The production results reveal that the learners' F1 and F2 values for English resembled the F1 and F2 values for BP. The results of the L2 perception test indicate that the learners discriminated different English vowels with different degrees of accuracy. Importantly, some relationship between perception and production was found because the two English (i.e., the learners' L2) vowel sounds that were poorly discriminated were produced with F1 and F2 values similar to those of a single vowel in the learners' L1 (first language).

1. Introduction

Currently, some of the most influential models that try to explain nonnative sound perception are Flege's Speech Learning Model (SLM) [1], and Best's Perceptual Assimilation Model (PAM) [2]. The SLM claims that L1 phonetic categories will limit the possibility of L2 category formation because it is assumed that L1 and L2 sounds co-exist in a single phonological space and that L2 sounds are filtered through the learner's L1 sounds. According to the SLM, a mechanism called equivalence classification establishes additional categories for "new" sounds, but not for "similar" sounds. A "new" L2 sound would be considered acoustically different from any L1 sound, whereas a "similar" sound would be perceived as some L1 category.

However, while Flege only distinguishes between "new" and "similar" sounds, Best's model proposes several assimilation processes, with their own relative difficulties. For instance, the PAM says that if two L2 sounds are perceived as two distinct L1 sounds, i.e., a two-category assimilation, non-native listeners will successfully discriminate such L2 sounds. However, if two L2 sounds are perceived as only one L1 sound, i.e., single category assimilation, and both L2 sounds are equally good examples of the L1 sound, their nonnative discrimination will be poor. In addition, if two L2 sounds cannot be classified as L1 sounds, discrimination is also expected to be poor. Another possible process of assimilation is proposed by Escudero and Boersma: When the L1 has a greater number of categories than the L2, one binary contrast, for instance, can be perceived as more than two categories in the L1, which is called multiple category assimilation [3].

Concerning the relationship between L2 perception and production, Rochet found evidence that the former precedes the latter [4]. The results of his study indicate that the inaccurate L2 productions of the French vowel /y/ by Canadian English and Brazilian Portuguese (BP) learners of French were perceptually motivated. Likewise, Flege et al. found that the higher the discrimination score, the more accurate the vowels were produced [5].

With respect to vowel acquisition, the comparison between two vowel systems can predict and explain the difficulties L2 learners may have in perceiving and producing vowels. That is, similarity between L1-L2 vowels [6], the different cues that signal vowel contrasts (e.g., spectral quality, duration) [3], [6], the differences in size between vowel systems [3], and the presence or absence of vowels in different inventories [3] are specific sources or difficulty that lead us to predict and explain L2 learners' non-target-like or inaccurate production. In the specific case of the acquisition of English vowels by BP learners of English, the particular ways in which the two vowel systems compare, as shown in Figure 1, leads to the types of difficulties outlined in the previous paragraph.

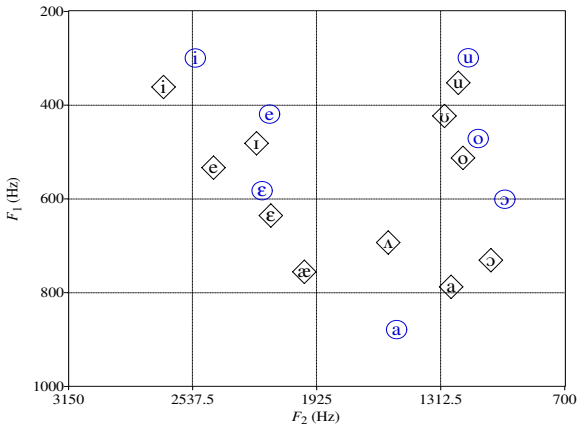

Figure 1: English vowels (in diamonds) and BP vowels (in circles) taken from [7] and [8], respectively.

In Figure 1 we see that, in stressed position, there are eleven vowels in English (/i, ɪ, eɪ, ɛ, æ, ʌ, ɑ, ɔ, oʊ, ʊ, u/), and seven oral vowels in Portuguese (/i, e, ɛ, a, ɔ, o, u/), which means that the English vowels /ɪ/, /eɪ/, /æ/, /ʌ/, /ɑ/, /oʊ/, and /ʊ/ do not exist in BP. We can observe that the BP vowels /ɪ/, /e/ and /ɛ/ are nearly at the same height as the English vowels /i/, /ɪ/ and /ɛ/, respectively. The BP /a/ is lower than, but as central as the English /ʌ/. As for the back vowels, the BP /ɔ/ is in an intermediary position between the English /ɔ/ and /oʊ/. The BP /o/ and /u/ are nearly at the same height as the English /oʊ/ and /u/, respectively.

Taking into account vowel inventory size differences together with the PAM's claim that two nonnative sounds can be assimilated into a single category, we formulated three research questions, based on the acoustic analysis of the F1 and F2 values of English and BP vowels: (1) how the formant values of the participants' L2 vowels will resemble those of their L1 vowels in production; (2) what specific English vowels will be poorly discriminated by BP learners; and (3) whether the L2 (i.e., English) vowels that are poorly discriminated are also inaccurately produced, i.e., produced with F1 and F2 values similar to those of the learners' L1 (BP). Three hypotheses to answer our questions were constructed based on previous studies: (1) the English vowels /ɪ, æ, ɑ, ʊ/ will be produced with formant values similar to those of BP /i, ɛ, ɔ, u/ [9]; (2) the English vowels /ɪ, æ, ɑ, ʊ/, which do not exist in the BP inventory, will have low discrimination rates [8]; and (3) participants who perceive vowel contrasts will also produce them differently [5]. These three hypotheses were tested with the experimental design described in the next section.

2. Method

In order to investigate the relationship between the perception and production of English vowels by BP learners, two experiments were designed. The first experiment consisted of a production test that provided information about the first two formant values of the English and BP vowels. Immediately after, the learners performed a perceptual experiment in which they discriminated L2 vowel constrasts in a categorial test, as will be described below.

2.1. Participants

Sixteen Brazilian Portuguese speakers of American English participated in our study. Due to the availability of participants, all of them were women. They were masters and doctoral students of English Language and Literature at the Federal University of Santa Catarina (UFSC). Their ages ranged from 26 to 30 years, the mean being 27.3 years. All the participants had already taught English for over 5 years. Thus, it was expected that their high degree of English proficiency and their constant exposure to English would have already contributed for them to have formed phonetic categories for the L2 vowels.

2.2. Experiment 1: L1 and L2 production

2.2.1. Material

The participants' production was elicited by their reading of 66 short sentences containing the English vowels (11 vowels x 6 sentences), and 42 sentences containing the BP vowels (7 vowels x 6 sentences). The English vowels were inserted in consonant-vowel-consonant (CVC) real words, and were both preceded and followed by a voiceless consonant. The BP vowels were inserted in disyllabic real words (CVCV) with penultimate stress. The BP words had to be disyllabic, because this language does not have monosyllabic words ending in voiceless stops. Only the first CV was examined. All the target words were in sentence-final position.

2.2.2. Procedure

The participants were asked to read the sentences containing the target vowels at natural speed in a language lab, and their productions were recorded by using Sony educational cassette recorders EF 5030, and Sony headsets H5-95.

In order to analyze the data, the sentences were digitized at 10 kHz, with 16-bit accuracy. The first two formants were measured by selecting a steady-state 25-millisecond portion of each vowel. In the specific case of the semi-diphthongs only the first element of each vowel was measured.

With the aim of comparing the participants' L2 production with that of monolingual English speakers, the data from Ohnishi were used. In his study, the vowels were produced with natural speed by 21 Californians within a bVd frame, except for /ɔ/, which was elicited in the word *orange*, since speakers of Californian dialect do not use this vowel consistently in monosyllabic words [7].

2.3. Experiment 2: L1 and L2 perception

A categorial discrimination test (CDT), based on Flege et al., was designed to investigate the discrimination rate of the English vowel pairs /i/-/ɪ/, /ɪ/-/eɪ/, /ɛ/-/æ/, /u/-/ʊ/, /ɔ/-/ɑ/, /ʊ/-/ʌ/, /ʊ/-/oʊ/, and /ʌ/-/ɑ/ [10]. The CDT we used was an oddity discrimination test in which every trial contained an odd item, or all the three items had the same target vowel. In the former case, the set is called *change trial*, since there is one vowel that differs from the other two, and in the latter the set is called *catch trial*, since all the vowels are the same. In the change trials, the odd item varied in position: In some trials it was the first item; in others, the second, and it also appeared in the third position. This format was used to avoid bias in the answers due to the position of the odd item.

2.3.1. Stimuli

The CDT contained 108 trials of three items (8 vowel contrasts x 8 trials + 11 non-contrasted vowels x 4 trials). The sequences were recorded by five native speakers (2 men and 3 women) from different U.S. states: Kentucky, Massachusetts, Michigan, New York and Pennsylvania. Although the native speakers were from different US regions, their vowels did not

differ considerably from each other, as perceptually judged by an American phonetician. The sample was recorded in the CSL program, at 10 kHz, with 16-bit accuracy. Following Flege et al., the five native speakers produced words that were formed by the insertion of one of the English vowels into the /bVt/ frame [10]. Each word was said at the end of the carrier sentence "This is a __". The target sentences were segmented and randomly organized in the Praat 4.2 program. The interval between the three sentences in each trial was 1.3 seconds and the interval between trials was 2.8 seconds.

2.3.2. Procedure

The participants were given a sheet of paper with four alternatives for each set and were asked to check alternative (a), (b) or (c) to indicate the odd item, or alternative (d) to indicate that all of the items were the same. Feedback was provided in a 5-set practice session before the experiment began. All the 16 participants were simultaneously tested in a language lab.

3. Results and discussion

3.1. Experiment 1: L1 and L2 production

Tables 1 and 2 show the mean and standard deviation (SD) of F1-F2 values of the BP, IL, and English front and back vowels. To facilitate comparison, the formant values of the BP vowels /e/, /a/, and /o/, are in the same columns as those of the IL and English vowels /eɪ/, /ʌ/, and /oʊ/, respectively.

Table 1: Mean and SD (in parentheses) of F1 and F2 values of Brazilian Portuguese (BP), interlanguage (IL), and English (E) front vowels.

Front Vowels		**/i/**	**/ɪ/**	**/e/ - /eɪ/**	**/ɛ/**	**/æ/**	**/a/ - /ʌ/**
B	F1	302		420	713		996
P	F2	(53)		(61)	(92)		(78)
		2574		2474	2190		1669
		(341)		(135)	(82)		(88)
I	F1	327	423	477	848	832	696
L	F2	(61)	(110)	(97)	(86)	(46)	(74)
		2607	2384	2569	2074	2153	1671
		(249)	(229)	(134)	(137)	(154)	(60)
E	F1	360	480	531	634	754	691
	F2	2674	2214	2423	2143	1977	1560

In Tables 1 and 2, it is possible to observe that although the distinction was small, 50% of the participants produced a distinction between the vowels of the /i/-/ɪ/ contrast, and 56.25% distinguished the /ɪ/-/eɪ/ contrast. In the /ɪ/-/eɪ/ pair, there was a tendency to invert the positions of /ɪ/ and /eɪ/, contrary to Baptista's findings, which showed that her Brazilian participants' /eɪ/ was produced lower than /ɪ/ [9]. None of the participants made a distinction between the /ɛ/-/æ/ contrast.

Table 2: Mean and SD (in parentheses) of F1 and F2 values of Brazilian Portuguese (BP), interlanguage (IL), and English (E) back vowels.

Back Vowels		**/ɑ/**	**/ɔ/**	**/o/ - /oʊ/**	**/ʊ/**	**/u/**
BP	F1		752	467		328
	F2		(69)	(56)		(48)
			1194	1099		994
			(96)	(87)		(184)
IL	F1	857	811	501	391	355
	F2	(97)	(82)	(113)	(119)	(52)
		1271	1116	1037	1212	1327
		(276)	(99)	(149)	(325)	(219)
E	F1	806	629	629	471	400
	F2	1251	1054	1389	1283	1114

Concerning the back vowels, /ɑ/-/ɔ/ were produced as only one vowel (/ɔ/) by 56.25% of the participants, and a small distinction was made by the remaining participants. The semi-diphthong /oʊ/ was produced close to the Portuguese /o/, in other words, too high by all the participants. Finally, some distinction was made between the vowels in the /ʊ/-/u/ contrast by 18.75% of the participants; however, 25 % of them did not make a distinction between the two back vowels, and 56.25% of the participants produced /ʊ/ and /u/ with nearly the same F1 values, but the former was produced with lower F2 values, which means that the distinction was inappropriate.

3.2 Experiment 2: L1 and L2 perception

As can be seen in Table 3, the discrimination of the contrasts /ɛ/-/æ/, /ʊ/-/u/, and /ɔ/-/ɑ/ was poor, their discrimination rate being less than 55%. The /ʌ/-/ɑ/ contrast was the least accurately discriminated (20.83%). This low discrimination rate might be explained by the small F1-F2 difference between the vowels in these pairs. The /i/-/ɪ/ contrast, which has some considerable distance between F1 and F2 values, was the most accurately discriminated contrast (93.83%). Some modest difficulty was found in the discrimination of the /ʊ/-/ʌ/ contrast (71%). The contrasts involving the semi-dipthongs were discriminated over 85% of the times. The slight diphthongization was probably a facilitating factor in perception.

Table 3: Percentages and SD (in parentheses) of accurate perception of change trials

	/i/-/ɪ/	**/ɪ/-/eɪ/**	**/ɛ/-/æ/**	**/u/-/ʊ/**
Mean	93.83 (2)	87.83 (2)	44.0 (3)	54.33 (3)
	/ɔ/-/ɑ/	**/ʊ/-/ʌ/**	**/ʊ/-/oʊ/**	**/ʌ/-/ɑ/**
Mean	29.5 (2)	71.0 (4)	85.67 (2)	20.83 (4)

Regarding the catch trials, /æ/ was the only vowel to present a rather high error rate (54.17%), which shows that the participants had not formed a category for this vowel and could not discriminate it from /ɛ/. Another vowel that caused

moderate difficulty in the catch trials was the back vowel /ɑ/ (70.83%). This difficulty is consistent with the results obtained in the change trials, since the /ʌ/-/ɑ/ and /ɔ/-/ɑ/ contrasts had the lowest discrimination rates. All the other vowels in the catch trials were accurately discriminated over 83.33% of the times.

3.3 Discussion

In this sub-section, we combine the results of the two experiments in order to verify whether the hypotheses are borne out. Recall that Hypothesis 1 states that the English vowels /ɪ, æ, ɑ, ʊ/ would be produced with formant values similar to those of BP /i, ɛ, ɔ, u/. The results of Experiment 1 show that, although proficient in English, the participants tended to make use of their L1 vowel system to produce L2 vowels. They were able to provide some slight contrast between similar sounds, but these contrasts still differed from those produced by native English speakers.

The results of Experiment 2 partially corroborate Hypothesis 2, which states that the English vowels /ɪ, æ, ɑ, ʊ/ would have low discrimination rates. The findings show that the participants had no difficulty discriminating the vowel /ɪ/, probably because it differs considerably from both /i/ and /eɪ/ in F1 and F2. However, the similar vowels /æ/, /ɑ/, and /ʊ/ were poorly discriminated in most of the contrasts.

Finally, the results of the two experiments corroborate Hypothesis 3, which says that participants who perceived vowel contrasts would also produce them differently. The comparison between the results of the perception and production tests gives more evidence to the fact that perception precedes production of sounds and that it is a prerequisite for accurate L2 production. The results of both experiments confirmed that the vowel contrasts with high discrimination rates were the ones produced with different F1 and F2 values.

Some special attention must be given to the results of the /ɛ/-/æ/ contrast in both experiments. In the production test, the formant values of these two vowels tended to be inverted. In the perception test, the participants had low discrimination rates in both catch and change trials. Thus, considering that this contrast is produced mainly through F1, the difference in F2 being quite subtle, participants with extensive exposure to English may have learned to differentiate the vowels by relying on duration. This thought is supported by Escudero and Boersma, who found that duration was the cue used by Spanish learners of English to discriminate this vowel contrast [3]. Even though the participants were not using the appropriate F1 and F2 values to produce the /ɛ/-/æ/ contrast, the reliance on duration might explain why some of the participants were able to discriminate the vowel in this contrast in some occasions. Taking into account Flege's concepts of what is "similar" and "new", the English vowel /ɛ/ should be considered a "similar" vowel by BP speakers, and /æ/, a "new" vowel. However, in terms of F1 and F2 values, the English /æ/ is "similar" and the English /ɛ/ is almost identical to BP /ɛ/. Thus, it is not possible to say which forms a new category, because the two are in the same category, which is what may cause the participants' lack of consistency when producing and perceiving them.

4. Conclusion

The results of this study show that every English vowel is acoustically similar to a BP vowel, and within the group of similar vowels, some are acoustically identical. This means that similarity, as proposed by the SLM, does not explain the lack of accurate production and perception of L2 vowels. Importantly, the assimilation of two English vowels into a single L1 category seems to be the source of difficulty the participants had to accurately produce and discriminate L2 vowels. In addition, the findings reveal that inaccurate production is related to inaccurate perception, as has been found in several L2 sound acquisition studies [3], [4], [5], [6].

5. References

[1] Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

[2] Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-206). Timonium: York Press.

[3] Escudero, P., & Boersma, P. (2002). The subset problem in L2 perceptual development: Multiple-category assimilation by Dutch learners of Spanish. In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 208-219). Cascadilla.

[4] Rochet, B. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410).Timonium, MD: York Press.

[5] Flege, J. E., Bohn, O. –S, & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

[6] Escudero, P. (2002). The perception of English vowel contrasts: Acoustic cue reliance in the development of new contrasts. In A. James & J. Leather (Eds.), *New Sounds 2000: Proceedings of the fourth international symposium on the acquisition of second language speech* (pp. 122-131). University of Amsterdam. Klagenfurt: Univiversity of Klagenfurt.

[7] Ohnishi, M. (1991). *A spectrographic investigation of the vowels of Californian English (Southwest General American).* Paper presented at the 1991 Convention of the Phonetic Society of Japan. Osaka, Japan.

[8] Rauber, A. S. (2002). The perception and production of English vowels by Brazilian Portuguese EFL learners. Unpublished paper. UFSC.

[9] Baptista, B. O. (2000). *The acquisition of English vowels by Brazilian Portuguese speakers*. ARES.Florianópolis:UFSC.

[10] Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America, 95* (6), 3623-3641.