

# CHALLENGES IN THE PERCEPTION AND PRODUCTION OF ENGLISH FRONT VOWELS BY NATIVE SPEAKERS OF EUROPEAN PORTUGUESE

## DESAFIOS NA PERCEÇÃO E PRODUÇÃO DAS VOGAIS ANTERIORES DO INGLÊS POR FALANTES DE PORTUGUÊS EUROPEU

Anabela Rato | Andréia Schurt Rauber | Letícia Piske Soares | Liane Regio Lucas  
UNIVERSIDADE DO MINHO, PORTUGAL | APPEN, ESTADOS UNIDOS | UNIVERSIDADE CATÓLICA DE PELOTAS, BRASIL | UNIVERSIDADE CATÓLICA DE PELOTAS, BRASIL  
asrato@gmail.com | asrauber@gmail.com | soareslepiske@gmail.com | liane19@hotmail.com

Este artigo apresenta os resultados de um estudo que investigou a percepção e a produção das vogais anteriores do inglês (/i/, /ɪ/, /ɛ/, /æ/) por um grupo de 18 falantes nativos de português europeu. A percepção das vogais-alvo foi testada com uma tarefa de identificação e os dados de produção de fala foram analisados acusticamente. Os resultados mostram que os ouvintes portugueses tendem a discriminar as duas vogais-alvo do par /i/-/ɪ/ com acuidade, mas produzem-nas com alguma sobreposição. Relativamente às vogais baixas, /æ/ foi produzida mais alta e mais posterior do que /ɛ/, com considerável sobreposição, e os resultados do teste de percepção relativos a este par indicam que /æ/ foi maioritariamente identificada como /ɛ/. Considerando que o /ɛ/ inglês é mais baixo do que o /ɛ/ português (Escudero *et al.*, 2009), ambas as vogais inglesas /ɛ/ e /æ/ foram percebidas como baixas o suficiente para serem identificadas como /æ/. Os resultados também sugerem que a percepção precede a produção das vogais altas, uma vez que as taxas de correção elevadas relativas à identificação de /i/ e /ɪ/ não corresponderam a uma produção com distância euclidiana suficiente entre essas categorias no espaço acústico. Relativamente às vogais baixas, os participantes tenderam a perceber ambas as vogais como baixas o suficiente para serem consideradas /æ/, mas produziram-nas altas o suficiente para serem consideradas /ɛ/.

**Palavras-chave:** aprendizagem fonológica não-nativa; vogais do inglês; percepção; produção

This paper presents the results of a study that investigated the perception and production of the English front vowels (/i/, /ɪ/, /ɛ/, /æ/) by a group of 18 native speakers of European Portuguese. Perception of the target vowels was tested with an identification test, while production data were analyzed acoustically. The results show that the Portuguese listeners tended to discriminate the two target vowels of the /i/-/ɪ/ pair accurately, but produced them with some overlap. As regards the low vowels, /æ/ was produced slightly higher and further back than /ɛ/, with considerable overlap, and the results of the perception test for this pair indicated that /æ/ was mostly identified as /ɛ/. Considering that English /ɛ/ is lower than Portuguese /ɛ/ (Escudero *et al.*, 2009), both English /ɛ/ and /æ/ were perceived as low enough to be identified as /æ/. The results also suggest that perception precedes production of the high vowels, since the high accurate scores in the perception of /i/ and /ɪ/ did not mean that the two categories were produced with enough euclidean distance in the acoustic space. As for the low vowels, the participants tended to perceive both vowels as low enough to be considered /æ/, but produced both high enough to be considered /ɛ/.

**Keywords:** non-native phonological learning; English vowels; perception; production.

## 1. Cross-language speech perception and production

It is widely acknowledged that second/foreign language (L2) speech learning is a challenge to adult learners in terms of articulation and perception of non-native phonetic contrasts, that is, L2 speech sounds that do not exist in the native language (L1) or are not phonologically distinctive in the L1. Adult L2 learners are, thus, frequently characterized as having not only a foreign pronunciation but also “accented” perception (Strange, 1995: 2). Difficulties to recognize and produce some non-native segments and suprasegments result from the interaction of different factors. On the one hand, there are the aspects related to the phonological similarities and differences between L1 and L2 phonetic inventories. On the other hand, there are various interrelated learner variables that include age of L2 learning, quantity and quality of L2 input, amount of L1 and L2 use over time, L2 formal instruction, and other individual differences such as motivation and language learning aptitude (Munro & Bohn, 2007; Piske, 2007; Sebastian-Gallés, 2005).

Studies have revealed that “difficulties in perception of non-native vowel contrasts are a significant part of the problems many L2 learners have in mastering the L2 phonology” (Strange, 2007: 36). Therefore, although this fact has been attested by a large number of studies that have investigated the cross-language perception and production of non-native vowels by learners with various L1 backgrounds, to our knowledge, the learning of L2 English vowel contrasts by European Portuguese learners has not yet been examined. Therefore, the purpose of this paper is to investigate the identification and production of English front vowels by adult native speakers of European Portuguese and examine whether learners’ perceptual and productive competences are interrelated.

Several theoretical models explain how non-native sounds are perceived by L2 learners, but we will only focus on the *Perceptual Assimilation Model* (PAM and PAM-L2) (Best, 1995; Best & Tyler, 2007) and the *Speech Learning Model* (SLM) (Flege, 1995). Both models depart from cross-language (L1 and L2) phonetic similarity to predict learners’ success or failure in the acquisition of non-native segmental contrasts.

The PAM predicts the occurrence of three main patterns in the perceptual assimilation of L2 contrasts. If non-native sounds are similar to native segments, (i) they can be perceived as good or bad exemplars of an L1 phonetic category (the higher the degree of similarity between L2 and L1 phonetic categories, the poorer the perceptual discrimination). If an L2 segment is very dissimilar from any L1 category, that is, if it is a sound category non-existing in the L1, (ii) it can be uncategorizable in the L1 system (uncategorized speech sounds) or (iii) perceived as non-speech sounds (non-assimilable).

The SLM also hypothesizes that the acquisition of L2 phonetic categories depends on the perceived similarity of L1 and L2 sounds. Flege (1987, 1995) claims that if an L2 sound is perceived as new, i.e., sufficiently dissimilar from L1 categories, the learner’s tendency is to establish a new category for that sound. However, if it is perceived as identical or similar, i.e., as an allophone of an L1 sound perceptually equivalent to an L1 phoneme, the learner fails to create a new category for that segment. Thus, the smaller the perceived cross-linguistic distance between L1 and L2 sounds, the higher the chances of perceiving an L2 sound as an allophone of an L1 sound; conversely, the greater the perceived phonetic dissimilarity between L1 and L2 sounds, the more likely it is that phonetic differences between the sounds will be perceived and a new category will be created. According to

the author, this process of perceptual equivalence classification helps determine L2 learners' production accuracy.

The relation between perception and production of non-native sounds has been extensively discussed in the field of L2 speech learning, and several studies have provided consistent empirical evidence for a link between learners' perceptual and productive abilities. We will briefly review some of the literature on the relation between perception and production of non-native vowel contrasts.

Rochet (1995) analyzed Portuguese and English speakers' perception and production of the French vowels /u/, /i/ and /y/ and found a correlation between inaccurate perception and production of the vowel /y/. When /y/ was produced incorrectly, Portuguese speakers tended to produce it more like /i/, whereas English speakers produced it more like /u/. In perceptual identification tests, Portuguese and English speakers also performed differently from native French speakers. Portuguese speakers identified /y/ as /i/ and English speakers as /u/, which followed the same pattern as that observed in production.

Flege, Bohn and Jang (1997) investigated the effect of L2 experience in the perception and production of the English front vowels /i/, /ɪ/, /e/, and /æ/ by 80 native speakers of German, Spanish, Mandarin and Korean. Significant differences between experienced and inexperienced non-native speakers of English were found in terms of production and perception accuracy, with both dimensions being related. Differences in degrees of accuracy in producing and perceiving the English vowels depended on L1 background, possibly because of the differences in the perceived relation between the L1 and L2 vowel inventories.

Another study by Flege, Mackay and Meador (1999) assessed the relation between perception and production of a larger set of English vowels by native Italian speakers, whose age of arrival (AOA) in Canada differed. As in Flege *et al.* (1997), the results showed that L2 experience influenced both production and perception, since speakers' accuracy in producing and perceiving English vowels diminished as AOA increased. Moreover, a correlation between L2 vowel production and perception was found, and the authors concluded that L2 vowel production accuracy was limited by how accurately they were perceived.

Rauber, Rato and Silva (2010) carried out two experiments to test perception and production of English front vowels by native speakers of Mandarin. The findings showed that perception accuracy outperformed production in the case of vowels /i/ and /æ/, which were better identified than produced.

They found that perception and production are related: Higher identification accuracy rates were related to better production results, whereas lower identification rates were related to poorer production.

Considering the cross-language similarity between European and Brazilian Portuguese, the results of three studies on English vowel perception and production by Brazilian learners of English as a foreign language (EFL) will be reported. Rauber, Escudero, Bion and Baptista (2005) investigated the identification and production of English vowels, and their findings confirmed that inaccurate production was related to poor categorical discrimination. Bion, Escudero, Rauber and Baptista (2006) further investigated this relation by examining the perception and production of a smaller subset of English front vowels (/i/, /ɪ/, /ɛ/, /æ/) and provided further evidence for a strong link between perception and production, since greater discrimination in the perception test was related to better production results. Rauber (2010) also examined whether there was a correlation between the perception and production of (/i/, /ɪ/, /ɛ/, /æ/, /u/, /ʊ/) by proficient English learners. Her findings showed that the rate of accurate perception of L2 vowels was higher than that of production, indicating that vowels which were better perceived were also the ones produced more accurately by L2 learners, thus confirming the interrelation between perceptual and productive abilities.

A number of studies have also demonstrated that L2 learners who are able to recognize and produce non-native vowel contrasts may accomplish the task in a different way than native speakers do, that is, their selective perceptual strategies may differ. We will now refer to two studies that investigated the weighting of acoustic cues used to differentiate non-native vowel contrasts, whether or not they are used in L2 learners' native languages. Both studies assessed the weighting of temporal (i.e., vowel duration) and spectral (i.e., degree of openness and frontness/backness of the vowel) cues in English lax-tense vowel contrasts.

For instance, Bohn and Flege (1990) investigated the perception of the English front vowels /i/, /ɪ/, /ɛ/, and /æ/ by adult native speakers of German, and observed that the German speakers used duration as the acoustic cue to distinguish the English /ɛ/-/æ/ contrast, instead of the spectral cue native speakers use. In the synthetic continua (*beat-bit*, *bet-bat*), in which vowel formants and duration varied, German speakers of English perceived the similar English contrast /ɪ/ and /i/ as native speakers did, and both groups of speakers used spectral cues to do so. Thus, it seems that German speakers use the same perceptual cues used by English speakers to distinguish

both vowels. However, in the continuum involving the “new” vowel /æ/, German speakers used durational cues, whereas native English speakers used spectral cues.

Cebrian (2006) assessed the relative weighting of quality and duration cues in the identification of the English lax-tense vowel contrast by adult Catalan speakers, whose L1 has no durational contrasts. This was evaluated by means of an identification test with a synthetic vowel continuum from /i/ to /ɪ/ to /ɛ/. Results showed that L2 learners made a greater use of temporal cues than did native English speakers.

This perceptual strategy used by L2 learners is explained by Bohn (1995) as the *desensitization hypothesis*. This principle states that “whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast” (Bohn, 1995: 294-5).

Finally, numerous studies investigating the effects of perceptual training on both L2 speech perception and production have also provided robust evidence on the interrelation between perceptual and productive abilities by showing that learning focused exclusively on perceptual identification and discrimination tasks can transfer to improvement in production (e.g., Aliaga-Garcia, 2013; Iverson, Pinet & Evans, 2012; Nobre-Oliveira, 2007; Pereira & Hazan, 2013).

## 2. The present study

In this study, we investigated how Portuguese EFL learners identify and produce two non-native vowel contrasts evenly distributed in the front area of the acoustic vowel space.

The European Portuguese phoneme inventory comprises three front vowels (/i, e, ɛ/), which differ only in spectral quality, whereas the American English vowel system includes four monophthongs (/i, ɪ, e, æ/) and one homogenous diphthong or semi-diphthong (/e/), and these vowels differ both in quality and length.

The high-front English vowel pair (/i-ɪ/) and the low-front vowel pair (/ɛ-æ/) differ both in terms of spectral quality and duration. The frequency of the first formant, which is the acoustic parameter of vowel height, is lower in /i/ than in /ɪ/, and also lower in /ɛ/ than in /æ/. In terms of duration, /i/ is longer than /ɪ/, and /æ/ is longer than /ɛ/.

Thus, based on the results of the studies reported in the previous section, this study aims at answering the following questions:

1. Are perceptual and productive abilities interrelated, i.e., is accurate perception related to accurate production of the target segments in L2 learners (Bion *et al.*, 2006; Flege, Bohn, & Jang, 1997; Flege, Mackay, & Meador, 1999; Rato, Rauber, & Silva, 2010, Rauber *et al.*, 2005; Rauber, 2010)?

2. Do native speakers of European Portuguese make use of both temporal and spectral cues to produce the English front vowels /i/, /i/, /ɛ/, /æ/ (Bohn & Flege, 1990; Cebrian, 2006)?

In order to answer these research questions, two experiments were elaborated: a perception test and a production test. Information about the participants, tests and analyses will be provided in the next subsections.

### 2.1. Participants

Eighteen students enrolled in the second semester of the European Languages and Literatures undergraduate course of a Portuguese university participated in this study, eight were women (ages ranging from 19 to 31 years, mean = 23.0, standard deviation (SD) = 4.9) and 10 were men (ages ranging from 18 to 29 years, mean = 20.7, SD = 3.3). Participants started learning English at school at the age of 10 years (age of learning (AOL) ranging from 6 to 10 years, mean = 9.6, SD = 1.2), and had been exposed to English mainly through formal classroom instruction in a Foreign Language (FL) context for seven years (length of formal instruction (LFI) ranging from 6 to 9 years, mean = 7.3, SD = 1.1). The number of weekly hours of formal English instruction in a monolingual Portuguese classroom ranged from two hours between the first and sixth grades to four hours between the seventh and twelfth grades. According to English proficiency tests taken at regular English classes at the university, all had an intermediate level of English proficiency, or B1 level (Common European Framework of Reference for Languages). The participants were born in Braga or in neighboring cities in the Minho region, Portugal, and had never lived in an English-speaking country. All reported to have no hearing impairment and participated as volunteers.

## 2.2. Production and perception tests

Two experiments were designed to test the participants' production and perception of the target vowels. The production test consisted of the reading of the carrier phrase "CVC. Say CVC now" (C = consonant, V = vowel). The CVC words used in this experiment were the same as those in Hagiwara (1997) and Rauber *et al.* (2010) and are presented in Table 1.

**Table 1. Words read by the participants to elicit the production of the target vowels**

	/i/	/ɪ/	/ɛ/	/æ/
b_t	beat	bit	bet	bat
h_d	heed	hit	head	had
t_k	teak	tick	tech	tack

The participants read the carrier sentence twice, totaling 432 productions (4 vowels x 3 phonological contexts x 2 repetitions x 18 participants). Each carrier sentence containing a target word was presented in isolation on a computer screen. The sentences were randomly organized to avoid ordering effects. To minimize the influence of orthography, before each carrier sentence, a picture representing a CVC word containing one of the four target vowels was presented (pictures of a bee (/i/), a pig (/ɪ/), an egg (/ɛ/), and a cat (/æ/)), and the participants were asked to read this word aloud and keep in mind that the sound of the vowel of the word in the picture should rhyme with the vowel of the target word in the sentence. Thus, for example, the participants would (1) see the picture of a bee; (2) say the word "bee"; (3) the experimenter would click on the screen and the carrier sentence "Say beat now" would be displayed so that (4) the participants could read it aloud.

The participants' productions were recorded in a silent room of the Portuguese university with a digital Roland Edirol R-09HR recorder and an Edirol CS-15 unidirectional condenser microphone, at 44,100 Hz, 16 bit, mono. However, to save space in computer memory, the recordings were later downsampled to 22,050 Hz. Each recording session lasted approximately 15 minutes.

The same perception test used by Rauber *et al.* (2010) was used in the present study and consisted of an identification test with natural stimuli designed in the software *Praat* (Boersma & Weenink, 2012). The stimuli were recorded by four native speakers of American English, two men and two women, all from the state of California. The stimuli were normalized according to peak amplitude to keep the volume constant and included CVC words containing one of the vowels /i, ɪ, ε, æ, ʌ, u, ʊ/ within two consonants that could be voiceless stops or fricatives or the voiced stop /b/ (e.g., *beat, book, sit, Pete, cat, but, boot*, see Appendix for the complete list). The participants heard each word in isolation four times, twice recorded by a woman, and twice recorded by a man. The exposure to different voices aimed at encouraging participants to ignore audible (voice quality) differences and focus on phonetically relevant differences in terms of category variation. The total number of stimuli was 168 (7 vowels x 6 contexts x 2 genders (every word was recorded once by a man and once by a woman) x 2 repetitions of the whole set). Figure 1 shows the identification test screen. After adjusting the volume to a comfortable level, the participants were instructed to choose one of the seven options displayed on the screen. Although only the results of the front vowels will be discussed in this paper, the experiment tested the perception of four front vowels (/i, ɪ, ε, æ/), one central vowel (/ʌ/), and two back vowels (/u, ʊ/). Since the students had not had any phonetics class by the date of the data collection, besides the phonetic symbol, a word whose stressed vowel contained one of the target segments was also included in the response buttons. The participants were instructed to hear a word and click on the button that contained the option that most resembled the stressed vowel of the stimulus. After choosing an option, the participant had to click on the OK button to hear the next stimulus. In case of doubt, it was possible to click on the Replay button to hear the stimulus again, but repetition was restricted to three times, i.e., after repeating the stimulus for the third time, the Replay button was automatically disabled. If after clicking OK and hearing the next stimulus participants noticed that they had made a mistake in the response to the previous stimulus, they could click on the “oops” button, which would then eliminate the previous answer and replay the previous stimulus, thus allowing participants to proceed in the test after having corrected a mistake. The “oops” button was rarely used. The perception test was administered in a computer lab of the Portuguese university, and the data collection session lasted approximately 30 minutes.

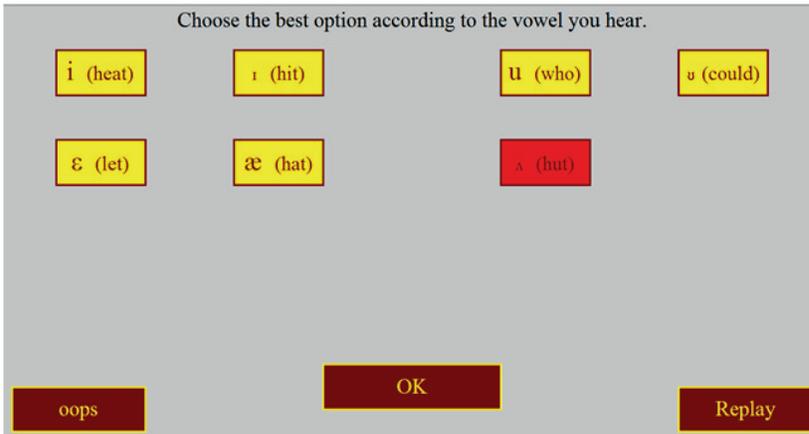


Figure 1. Identification test screen in *Praat*

### 2.3. Data analysis

As regards the production data, the target vowels were analyzed in terms of spectral quality (first two formants: F1 and F2) and duration in *Praat*. The vowels were manually segmented at zero crossings, i.e., when the wave crosses the zero amplitude line. The beginning of the vowel was considered the first periodic pulse on the waveform that had considerable amplitude, while the end point of the vowel was the last periodic pulse with some amplitude and that still resembled the vowel period. The darker areas showing the formants in the spectrogram were also important to help determine the vowel start and end points. After segmentation and labeling of the vowels, a *Praat* script (same used in Rauber, 2010) was run to measure (1) total duration of the segmentation, and (2) the mean F1 and F2 values at the 40% central part of the segmentation, which is expected to be the most stable portion of the vowel, less affected by the articulation of the previous and following segments.

In order to have comparable vowel spaces, the data of the Portuguese participants were normalized according to the minimum and maximum formant values of the vowels produced by the American participants (see Rauber, 2010: 90-92, for a detailed explanation about the normalization procedure). After normalizing the formant values, we calculated the Euclidean distance between the two vowels of each pair. The Euclidean distance is the space in Hertz (Hz) between /i/-/ɪ/, and /ɛ/-/æ/ in terms of F1 and F2

values (for details on Euclidean distance, see Bion *et al.*, 2006: 1364). This procedure allowed the comparison between the Euclidean distance of the vowels produced by the Portuguese and the American participants, i.e., we were not interested in the comparison of absolute values between the two groups, but of the space between the vowels of each pair. All the measures were then exported to IBM SPSS v. 17 for statistical analysis.

The perception data were analyzed after extracting the results from *Praat's* experiment files. The percentage of correct and incorrect identifications was organized in confusion matrices for comparisons and the data was also statistically analyzed in IBM SPSS.

The following sections will report the results obtained in the production and perception tests, respectively.

#### 2.4. Production results

The results of the production test are organized according to gender. Table 2 presents duration, F1 and F2 values of the vowels produced by the Portuguese L2 participants, as well as the relevant acoustic measures of the vowels produced by American monolinguals. To facilitate comparisons, the data about the male American English monolinguals are those published by Rauber (2010), and the female American monolingual data are those published by Rauber *et al.* (2010). These two studies will be our main references for English vowel duration and formant values because their data collection procedures were exactly the same as those in the present study. Figure 2 shows the plot with the formant measures of the L2 participants.

Table 2. Mean duration and formant values measured in the present study (standard deviation in parentheses), American monolinguals' values from Rauber (2010, male participants), and Rauber *et al.* (2010, female participants)

	Portuguese L2 speakers					
	Duration		F1		F2	
	F	M	F	M	F	M
i	114 (39)	100 (25.3)	396 (34.8)	289 (31.1)	2795 (144.2)	2381 (132.0)
ɪ	97 (28)	103 (33.1)	432 (52.0)	334 (54.5)	2718 (177.2)	2309 (180.6)
ɛ	130 (34)	103 (28.0)	787 (54.7)	649 (48.7)	2297 (108.7)	1955 (131.8)
æ	143 (39)	100 (23.1)	799 (55.7)	680 (45.4)	2270 (125.7)	1946 (92.9)

American monolinguals						
i	134 (14.4)	140 (24.4)	393 (29.4)	276 (21.2)	2744 (146.5)	2332 (171.7)
ɪ	82 (16.4)	118 (20.3)	565 (32.1)	403 (48.2)	2228 (154.5)	1860 (181.2)
ɛ	104 (18.0)	134 (25.6)	713 (15.7)	538 (71.2)	1968 (79.1)	1732 (143.9)
æ	154 (23.3)	179 (29.3)	816 (27.8)	642 (86.4)	1998 (97.1)	1664 (130.2)

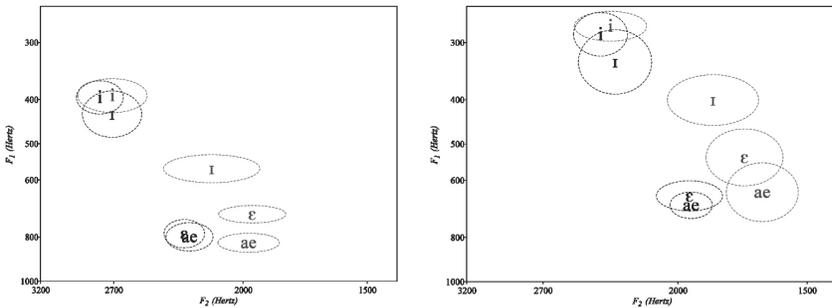


Figure 2. F1 and F2 values of the Portuguese learners (black) and the American monolinguals (grey), the ellipses represent standard deviation; women's results on the left, and men's results on the right.

As can be observed in Table 2 and Figure 2, there is partial overlap in the production of the high front vowel pair and almost complete overlap of the low front vowel pair by both female and male Portuguese L2 speakers. Table 3 shows the Euclidean distance means between the vowels produced by both the Portuguese learners and the American monolinguals. Since the data was not normally distributed, the non-parametric Mann-Whitney test was run, and the results revealed that both female and male L2 participants produced vowels with a Euclidean distance significantly smaller than that of American monolinguals (*women* /i-ɪ/:  $Z = -3.06, p = .002, /ɛ-æ/ : Z = -1.96, p = .05; men /i-ɪ/ : Z = -3.24, p = .001, /ɛ-æ/ : Z = -2.78, p = .005$ ).

Table 3. Euclidean distance values (in Hertz) between the vowels of the two pairs by the Portuguese L2 learners and the American monolinguals (US mono), standard deviation within parentheses

Vowel pair	Group	ED (SD)	
		Women	Men
/i-ɪ/	L2 learners	114 (63)	104 (66)
	US mono	580 (81)	490 (55)
/ɛ-æ/	L2 learners	62 (35)	65 (49)
	US mono	129 (42)	126 (39)

As regards duration, Table 1 and Figure 3 show the participants' mean values. In the group of Portuguese women, Mann-Whitney tests revealed that the duration values of the vowels produced by the L2 learners were significantly greater than those produced by the American monolinguals for the vowels /i/ ( $Z = -2.12$ ,  $p = .033$ ) and /ɛ/ ( $Z = -3.88$ ,  $p < .001$ ). No significant difference in duration was found for /i/ ( $Z = -1.90$ ,  $p = .057$ ) and /æ/ ( $Z = -1.24$ ,  $p = .22$ ). Differently, the male L2 speakers produced all vowels with duration values significantly smaller than their English counterparts (/i/:  $Z = -7.54$ ,  $p < .001$ ; /ɪ/:  $Z = -2.83$ ,  $p = .005$ ; /ɛ/:  $Z = -5.82$ ,  $p < .001$ ; /æ/:  $Z = -9.50$ ,  $p < .001$ ). In order to analyze the duration differences within groups, Wilcoxon tests were run for the L2 participants. Significant differences were only found in the group of women, showing that they produced the vowels /ae/ and /i/ longer than /ɛ/ ( $Z = -2,708$ ,  $p = .007$ ) and /ɪ/ ( $Z = -3,629$ ,  $p < .001$ ), respectively. This indicates that only the female participants used the duration cue to distinguish the vowels of each pair.

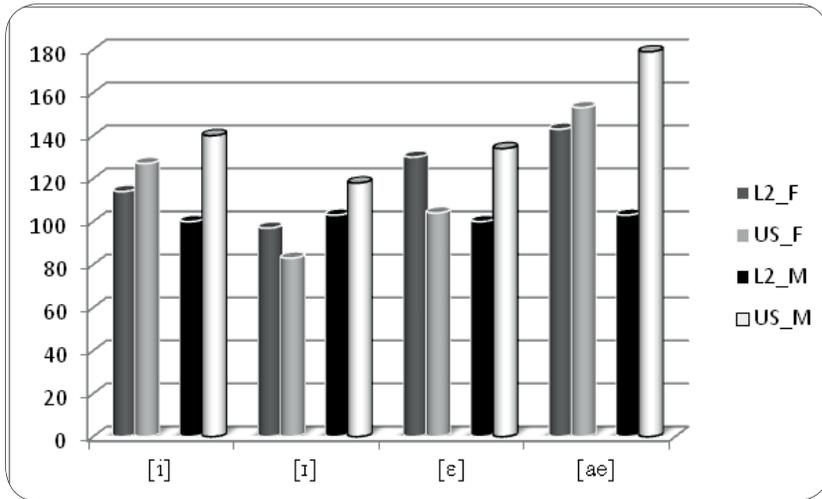


Figure 3. Mean duration values (in milliseconds) of the vowels produced by the female and male L2 learners and American monolinguals (US)

### 2.5. Perception results

The vowel identification rates of the group of L2 female participants are reported in Table 4. The first column shows the vowels heard and the first line shows the vowels selected in the response buttons, i.e., the vowel categories they could map the sound heard to. Since the data was not normally distributed, all statistical tests run to analyze perception data were non-parametric. We can observe that the highest rate of correct recognition was for /æ/ (86.81%) and that, according to the result of a Wilcoxon test, the vowel /ɛ/ was identified significantly less accurately (36.11%) than its counterpart ( $Z = -2.21$ ,  $p = .027$ ). The low rates for the correct identification of /ɛ/ might be explained by the fact that, according to Escudero *et al.* (2009), the F1 value of this vowel produced by Portuguese speakers from Lisbon is 511 Hz, while the F1 of American English (Rauber *et al.*, 2010) is 713 Hz. Thus, the English /ɛ/ is so much lower than the Portuguese /ɛ/ that the participants might have categorized it as /æ/. As for the vowels /i/ and /ɪ/, both had around 70% of correct responses and no significant difference ( $Z = -1.09$ ,  $p = .28$ ) was found in the accurate perception of this pair. We can observe that the main confusion was between themselves, that is, /i/ was consistently misidentified as /ɪ/ and vice versa.

Table 4. Confusion matrix with the mean percentage rates of category recognition by the L2 female participants

Vowels heard	Identified vowels						
	/æ/	/ɛ/	/i/	/ɪ/	/u/	/o/	/ʌ/
/æ/	<b>86,81</b>	6,94				0,69	5,56
/ɛ/	59,73	<b>36,11</b>		0,69	0,69	0,69	2,09
/i/	0,69	0,69	<b>70,14</b>	25,70	0,69		2,09
/ɪ/	1,39	2,09	22,22	<b>73,61</b>			0,69

The results of the L2 male participants follow the same trend as that of the females. Table 5 shows that /æ/ was accurately identified 93.75% of the times, i.e., significantly more accurately than /ɛ/ (53.34%), as revealed by the Wilcoxon test ( $Z = -2.81$ ,  $p = .005$ ). Although the Portuguese men confused /ɛ/ with other vowels more often than did the women, when misidentified, this vowel was frequently confused with /æ/. According to Escudero *et al.* (2009), the F1 value of /ɛ/ produced by Portuguese monolinguals from Lisbon is 455 Hz, considerably lower than the English /ɛ/ (538 Hz). Thus, similarly to the female participants, the male L2 speakers might have considered the English /ɛ/ so low that they often identified it as /æ/. No significant difference was found between the correct identification rates for /i/ (66.24%) and /ɪ/ (70.83%), as shown by the results of the Wilcoxon test ( $Z = -.53$ ,  $p = .59$ ). The perceptual confusion in the identification of these two vowel categories shows evidence that they have not been fully established in the participants' phonological system.

Table 5. Confusion matrix with the mean percentage rates of category recognition by the L2 male participants

Vowels heard	Identified vowels						
	/æ/	/ɛ/	/i/	/ɪ/	/u/	/o/	/ʌ/
/æ/	<b>93.75</b>	3.75	1.67				0.83
/ɛ/	39.17	<b>53.34</b>	2.08	3.33	0.83	0.83	0.42
/i/		0.42	<b>66.24</b>	32.5	0.42	0.42	
/ɪ/	1.67	6.25	21.25	<b>70.83</b>			

### 3. Discussion and conclusions

In this paper, we aimed at investigating how native speakers of European Portuguese perceive and produce English front vowels, and whether their production was based on both temporal and spectral acoustic cues. The production and perception data provided evidence that the vowel contrasts /i/-/ɪ/ and /ɛ/-/æ/ are not well established in the L2 learners' phonological system: The degree of overlap in production is greater for the pair of low vowels, but also occurs in the pair of high vowels. Interestingly, the English /ɛ/ produced by the Portuguese participants was much lower than their L1 /ɛ/. Thus, instead of producing /æ/ higher than the American monolinguals, the Portuguese participants lowered the /ɛ/ so much that almost no distinction can be made between the vowels of this pair in production. The Euclidean distance between the high vowels is somewhat more visible in the L2 data, but the partial overlap in production can be easily understood by the perception results, namely the considerable confusion in the identification of /i/ and /ɪ/.

The results of vowel duration show that only the female L2 speakers used this acoustic cue to make distinctions between the vowels of the two contrasts, but still the values for the lax vowels /ɪ/ and /ɛ/ were significantly greater than those for the Americans. Thus, the L2 female participants made use of both temporal and spectral cues to produce the English vowels, but this use was still quite different from the American monolinguals' cue weighting, corroborating previous studies (e.g., Bohn & Flege, 1990; Cebrian, 2006).

The findings provide further evidence for the interrelation between perception and production. The pair of high vowels was produced with some Euclidean distance, and the confusion in perception was smaller than that of the low vowels, which were produced with almost complete overlap. Furthermore, these results support the predictions of both L2 speech learning models (PAM-L2 and SLM) and previous studies that investigated the interrelation between perception and production of L2 vowels (e.g., Bion *et al.*, 2006; Flege *et al.*, 1997, 1999; Rauber *et al.*, 2005, 2010; Rochet, 1995). Participants were not able to establish phonetic categories for non-native sounds that, although differing acoustically from corresponding vowels in the L1, are perceptually similar to native sounds. The high degree of perceived similarity between L2 and L1 front vowels led to the merging of two distinct L2 vowel categories into one L1 category. This could be observed both in the production and perception of the low-front contrast and, to

a lesser extent, of the high-front contrast. The study of other non-native vowel contrasts is needed to further understand this interrelation between perceptual and productive abilities in the learning of L2 phonetic segments. Moreover, follow-up studies are necessary to investigate whether perceptual categorization and production accuracy can be attained by means of phonetic training, focused on raising L2 learners' awareness of perceptual differences between L1 and L2 phonological inventories.

## References

- ALIAGA-GARCIA, Cristina. 2013. "Effects of identification and articulatory training methods on L2 vowel production." Paper presented at *New Sounds 2013*, Concordia University, Montreal, Canada, May 17-19.
- BEST, Catherine. 1995. "A direct realist view of cross-language speech perception." In *Speech Perception and Linguistic Experience: Issues in Cross Language Research*, edited by Winifred Strange, 171-204. Maryland: York Press.
- BEST, Catherine T., and TYLER, Michael D. 2007. "Nonnative and second-language speech perception: Commonalities and complementarities." In *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, edited by Ocke-Schwen Bohn and Murray J. Munro, 13-34. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- BION, Ricardo A. H., ESCUDERO, Paola, RAUBER, Andreia S., and BAPTISTA, Barbara O. 2006. Category formation and the role of spectral quality in the perception and production of English front vowels. *Proceedings of Interspeech 2006*. 1363-66
- BOHN, Ocke-Schwen. 1995. "Cross-language speech perception in adults: First language transfer doesn't tell it all." In *Speech Perception and Linguistic Experience: Issues in Cross Language Research*, edited by Winifred Strange, 279-304. Maryland: York Press.
- BOHN, Ocke-Schwen, and Flege, James E. 1990. "Interlingual identification and the role of foreign language experience in L2 vowel perception." *Applied Psycholinguistics* 11: 303-328.
- BOERSMA, Paul, and WEENINK David. "PRAAT: Doing phonetics by computer" (Version 5.3.23) (Computer program). Accessed August 8, 2012. <http://www.praat.org>.
- CEBRIAN, Juli. 2006. "Experience and the use of non-native duration in L2 vowel categorization." *Journal of Phonetics* 34: 372-87.
- FLEGE, James E. 1987. "The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification." *Journal of Phonetics* 15: 56-65.
- FLEGE, James E. 1995. "Second language speech learning: Theory, findings, and problems." In *Speech Perception and Linguistic Experience: Issues in Cross Language Research*, edited by Winifred Strange, 233-278. Maryland: York Press.

- FLEGE, James E., BOHN, Ocke-Schwen, and JANG, Sunyoung. 1997. "Effects of experience on non-native speakers' production and perception of English vowels." *Journal of Phonetics* 25: 437-470.
- FLEGE, James E., MACKEY, Ian R. A., MEADOR, Diane. 1999. "Italian speakers' perception and production of English vowels." *Journal of the Acoustical Society of America* 106/5: 2973-2987.
- HAGIWARA, Robert. 1997. "Dialect variation and formant frequency: The American English vowels revisited." *Journal of the Acoustical Society of America*, 102/1: 655-658.
- IVERSON, Paul, PINET, Melanie, and EVANS, Bronwen G. 2012. "Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels." *Applied Psycholinguistics* 33: 145-60. Accessed June 4, 2013. doi: 10.1017/S0142716411000300.
- MUNRO, Murray J., and BOHN, Ocke-Schwen. 2007. "The study of second language speech learning: A brief overview." In *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, edited by Ocke-Schwen Bohn and Murray J. Munro, 3-12. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- NOBRE-OLIVEIRA, Denise. 2007. "The Effect of Perceptual Training on the Learning of English Vowels by Brazilian Portuguese Speakers." PhD dissertation. University of Santa Catarina, Brazil.
- PEREIRA, Yasna and HAZAN, Valerie. 2013. "Impact of different training modes on the perception and production of English vowels by L2 learners." Paper presented at *New Sounds 2013*, Concordia University, Montreal, Canada, May 17-19.
- PISKE, Thorsten. 2007. "Implications of James Flege's research for the foreign language classroom." In *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, edited by Ocke-Schwen Bohn and Murray J. Munro, 315-30. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- RAUBER, Andreia S. 2010. *Acoustic Characteristics of Brazilian English vowels: Perception and Production results*. Saarbrücken: Lambert Academic Publishing.
- RAUBER, Andreia S., ESCUDERO, Paola, BION, Ricardo A. H., BAPTISTA, Barbara O. 2005. "The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese." *Proceedings of Interspeech 2005*. 2913-16.
- RAUBER, Andreia S., RATO, Anabela, and SILVA, Lúcia. 2010. "Percepção e produção de vogais anteriores do inglês por falantes nativos de mandarim." *Diacrítica* 24/1: 5-23.
- ROCHET, Bernard. 1995. "Perception and production of second-language speech sounds by adults." In *Speech Perception and Linguistic Experience: Issues in Cross Language Research*, edited by Winifred Strange, 379-410. Maryland: York Press.
- SEBASTIÁN-GALLÉS, Núria. 2005. "Cross-language speech perception." In *The Handbook of Speech Perception*, edited by David B. Pisoni and Robert E. Remez, 546-566. Oxford: Blackwell Publishing.

- STRANGE, Winifred. 1995. "Cross-language studies of speech perception: a historical review." In *Speech Perception and Linguistic Experience: Issues in Cross Language Research*, edited by Winifred Strange, 3-45. Maryland: York Press.
- STRANGE, Winifred. 2007. "Cross-language phonetic similarity of vowels: Theoretical and methodological issues". In *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, edited by Ocke-Schwen Bohn and Murray J. Munro, 35-56. Amsterdam/Philadelphia: John Benjamins Publishing Company.

## Appendix

### Stimuli used in the perception test

/i/	/ɪ/	/ɛ/	/æ/	/ʌ/	/ʊ/	/u/
beat	bit	bet	bat	but	book	boot
keep	kit	kept	cat	cut	cook	coot
Pete	pit	pet	pat	putt	put	poop
seat	sit	set	sat	shut	soot	suit
teak	tick	tech	tack	tuck	took	toot
teat	tit	tet	tat	tut	took	tuke