

# Can Late EFL Learners Attain Nativelike Pronunciation? Evidence from Catalan Speakers' Production of English Low Vowels

Lucrecia Rallo Fabra

*Universitat de les Illes Balears*

## Abstract

Catalan learners of English have many difficulties to produce the three English low vowels /æ ʌ ɑ/ accurately. This is because, like Spanish, Catalan has only one low vowel /a/ so learners tend to perceive the three target vowels as instances of the same Catalan category (Rallo Fabra, 2005). Flege (1995) predicts that nonnative speakers can produce L2 sounds authentically if they perceive the differences between the native and the target sounds. This paper investigates production accuracy of the three English low vowels by three groups of Catalan learners of English differing in foreign accent (FA) and a group of native English speakers. The data were obtained in an elicitation task in which participants were asked to pronounce a series of monosyllabic words containing one of the three target vowels. Production accuracy was measured quantitatively in terms of spectral data of first and second formant vowel frequencies. Results indicate that learners can produce nativelike instances of vowels /æ/ and /ʌ/ but not of /ɑ/. These findings do not really follow Flege's predictions, since learners were able to produce nativelike instances of two English vowels although these vowels were heard as "similar" to Catalan /a/.

## 1. Introduction

Accented speech is a common trait among learners who acquired the target language after what is known as the critical period for language acquisition. A great body of research has shown that foreign accent and pronunciation errors persist even after lengthy periods of residence in the host country or daily use of the target language (Flege, MacKay and Meador, 1999, Piske, Flege, MacKay and Meador, 2002). Various factors influence successful L2 speech learning.

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

Flege Schirru and MacKay (2003) classify these factors in three types: maturational factors, amount and nature of L2 input and interaction of L1-L2 sound systems.

*Maturational factors: “early is better”.* Learners who were first exposed to the target language in childhood tend to be more successful than learners whose exposure started at adolescence. A growing number of speech perception studies have proved that the mechanisms that are operative for L1 acquisition are less effective as age of exposure to the L2 increases. For instance, Strange (1995) refers to pre-linguistic infants as “language universal perceivers” because they are able to distinguish any sound contrast regardless of their linguistic backgrounds. Unlike children, adults are highly resistant to modify the perceptual patterns established in childhood and, consequently, these perceptual foreign accents interfere with learning new L2 sounds.

*Amount and nature of L2 input.* Early and late learners represent two separate populations in terms of L2 speech learning not just because of the age factor. Often the age factor interacts with the amount and type of L2 input learners are exposed to. Flege *et al.* (2003) argue that early bilinguals are bound to receive a higher amount of L2 input because they attend school so they are forced to interact with speakers of the target language on a daily basis. But late bilinguals receive less input because at their workplace they speak and listen both their first and second language.

Failure to produce the target sounds inaccurately can be also attributed to the quality of the input learners have received. The instructional setting not always provides the best conditions for speech learning because many foreign language teachers themselves do not succeed in pronouncing the sounds of the target language accurately. It follows that nativelike pronunciation will be out of reach to students who are permanently exposed to foreign-accented speech.

*Interaction of the L1 and L2 sound systems.* Speaking two languages does not imply that the sound systems of each language operate separately. One of the tenets of the Speech Learning Model (Flege, 1995: 242) is that “the L1 and L2 sound systems exist in a *common phonological space*” and hence influence each other. This interaction involves two mechanisms: category assimilation and category dissimilation. An L2 sound assimilates to an L1 sound when it is perceived as an instance of the L1 sound, despite audible differences

between both sounds. Category dissimilation is the result of the learners' effort to maintain contrast between the target sound and the closest native sound. And it occurs when learners perceive the target sound as different from the closest L1 sound.

Strange (2007) describes two methods to establish "cross-language similarity relationships". These two methods are acoustic/articulatory descriptions of sound inventories and direct assessment of similarity by means of perceptual tests. Table I shows spectral values of the three target English vowels (Hillendbrand and cols, 1995) and the closest Catalan vowels (Recasens, 1986). According to these values English /æ/ is acoustically close to Catalan /ɛ/. The F1 values of English /ʌ/ and /ɑ/ are close to Catalan /a/ but the F2 values of these two English vowels are lower than Catalan /a/, indicating the central nature of the Catalan vowel as compared to the English targets.

In a previous study (Rallo Fabra, 2005), the perceptual relatedness of English vowels to Catalan vowels was examined using the same direct estimation method as in previous work by Flege and colleagues (Flege, McKay and Meador, 1999; Guion, Flege, Akahane-Yamada and Pruitt, 2000; Tsukada et al., 2005). The same group of learners who participated in the present study were asked to identify English vowels as instances of the seven Catalan vowel categories. What is interesting here is that three English low vowels /æ ʌ ɑ/ were mapped onto a single Catalan vowel category /a/, of these three, /ɑ/ was the best fit followed by /ʌ/ and /æ/. Indicating that cross-language perceptual similarity may not parallel similarities based on acoustic descriptions.

The aim of the present study is to investigate pronunciation of the three English low vowels by three groups of adult Catalan learners. We can predict that it is unlikely that these learners pronounce the three low vowels /æ ʌ ɑ/ accurately because these three vowels were assimilated to the same Catalan vowel /a/. It is uncertain to which extent the type of L2 input learners received can influence pronunciation of the target vowels. As undergraduate students, participants were exposed to both native and non-native English speakers and they had finished a course on English Phonetics by the time they were tested. This profile of learners favours a good disposition toward the learning task and as Moyer (1999: 85) suggests

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

“places these subjects in a unique position to overcome potential maturational constraints”.

Table 1. Average formant frequencies in Hz of the three target American English vowels (Hillenbrand, Getty, Clark and Wheeler, 1995) and the closest Catalan vowels (Recasens, 1986).

	AmEng /æ/	AmEng /ʌ/	AmEng/a/
<b>F1</b>	588	626	768
<b>F2</b>	1952	1200	1333
<b>F3</b>	2601	2550	2522

	Cat/e/	Cat/ɛ/	Cat/a/
<b>F1</b>	367	570	713
<b>F2</b>	1996	1852	1498
<b>F3</b>	2543	2542	2445

## 2. Method

### 2. 1. Participants

Two groups of speakers participated in the study: a group of experienced learners of English from the URV in Tarragona and a control group of native AmEng speakers. They were all female because in most English Philology courses male students are outnumbered by female students. Vowel production accuracy was evaluated by means of spectral values. It is well known that the differences in vocal tract size and vocal fold length account for variation in fundamental frequency and consequently, on formant frequencies between male and female voices. At the time the speech samples were obtained they were all finishing an American English Phonetics at their and Phonology course which included in-depth study of the English sound system as well as intensive stimulability.

The Catalan learners had similar amounts of exposure to English in terms of formal instruction and stay in an English-speaking country but they exhibited different levels of proficiency in English.

For this reason, the group was divided into three subgroups, according to their foreign accent, as shown in Table 2: low FA (n=10), mid FA (n=7), high FA (n=7). The FA estimations were based in the score of a series of recordings that students had to provide as course requirement. The speech samples were tape recorded for later evaluation by the English Phonetics instructor, an experienced Catalan phonetician who had lived in the U.S. for ten years and spoke English on a regular basis. The native AmEng speakers were visiting students from an exchange program with the University of Barcelona. They all came from different states in the US (Illinois, Connecticut, Florida, California and Washington).

Table 2. Pronunciation scores obtained by the 24 Catalan learners of English tested in the experiment

	Mean	Minimum	Maximum
'Low FA'	8.72	7.9	9.8
'Mid FA'	6.66	6	7.5
'High FA'	4.25	2.5	5.5

## 2. 2. *Speech elicitation*

As shown in Table 3, each speaker produced a total of 13 monosyllabic words containing four or five tokens of each of the three target vowels. The words were selected from *The Oxford Acoustic Database*, which provides word lists for various languages classified by sounds.

Participants were tested individually in a sound-proof booth while the experimenter showed them a card with a word printed on it through the booth window. This way talkers could not see which word was coming next. The recordings were made with a *Marantz* tape recorder, Model CP-300 and a *Shure* (model *SM58*) microphone.

The words were digitized using *Kay Elemetrics Computerised Speech Lab CSL 5500* at 10000 Hz sampling rate. The audio files were then synthesized using the *Analysis and Synthesis (ASL)* utility. Only the vowel portion of each file was synthesized. Each synthesized file was finally converted to text format.

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

Vowel production accuracy was evaluated by means of acoustical parameters, more specifically first (F1) and second (F2) formant frequencies of the steady-state portion of the vowel segment. F1 and F2 are known to distinguish vowel sounds across languages (Ladefoged, 2001). F1 indicates vowel height: high vowels such as /i/ have low formant frequencies. In contrast, low vowels such as /æ/ have higher formants. Second formant indicates the front/back position of the tongue, so the second formants of front vowels such as /i/ or /e/ are high in frequency as compared to back vowels such as /u/ or /ɔ/.

Table 3. Speech materials used in the experiment.

/æ/	/ɑ/	/ʌ/
cap	rob	rub
dad	sob	shut
had	knob	nun
man	cop	cup
lass		

### 3. Results

A first inspection of Table 4 with the mean values obtained across vowels and groups reveals similar performance in all groups for the low front vowel /æ/. Small differences between the Mid FA group and the other two groups of learners, in terms of F1. And finally the low back vowel /ɑ/ reveals different F1 values of the three groups of learners with respect to the native English group. To test the significance of these between-group differences, the values were submitted to three separate one-way ANOVAs for each target vowel. The *F* values for /æ/ show that there was no significant difference between the vowel productions of the non-native and English native groups ( $F(3,155)=2.87, p>0.01$  and  $F(3,155)=2.11, p>0.01$ ), indicating that the three groups of learners produced /æ/ authentically.

All three groups of learners produced nativelike tokens of the English central vowel / $\Lambda$ /. The ANOVA only yielded a marginally significant effect of the F1 ( $F(3,122)=4.26, p<0.01$ ) but no significant effect of F2 ( $F(3,122)=2.24, p>0.01$ ). Tukey's post hoc analysis indicated that the Mid FA group's productions of the central mid-low vowel / $\Lambda$ / were significantly higher than the productions by the low and high FA groups.

Table 4. Mean frequency values obtained from the vowel elicitation of three groups of Catalan learners of English varying in experience with the target language and a group of native AmEng controls.

		Low	Mid	High	NAmE
/æ/	F1	782	831	794	862
	F2	2083	2003	2003	1921
/ɑ/	F1	788	828	751	905
	F2	1386	1419	1478	1346
/Λ/	F1	728	804	711	785
	F2	1605	1548	1620	1684

For the low back vowel /ɑ/, the ANOVA revealed a significant effect on first formant frequencies ( $F(3,124)=8.36, p<0.001$ ) but not of second formant frequencies ( $F(3,124)=2.24, p>0.01$ ). Pairwise comparison with Tukey's post hoc test showed that the F1 frequencies of the tokens produced by the low and high FA groups were significantly lower than the control group'. The F1 of the mid FA learners' productions did not differ significantly from the native English group.

The F1 and F2 formant values for the three vowels tested have been plotted in four separate charts for each group. The values of one NE speaker were not included in this chart because she exhibited backing of the mid-low central vowel / $\Lambda$ / as it is common among speakers of the Northern US cities like Chicago (Labov, 1998). This phenomenon caused expansion of the ellipsis for this vowel. Two aspects should be noted in the charts that correspond to the Catalan learners' vowel productions. First, the size of the ellipses increases in

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

the High FA group indicating that many of these learners produce vowel instances that deflect away from the target. Second, the ellipses for the three vowel categories present an increasing degree of overlapping, especially in the ellipses of vowels / $\Lambda$ / and / $\alpha$ / produced by the high FA group. The acoustic distances between these two vowels narrow indicating that learners strive to maintain the contrast between both sounds in speech production.

Figure 11: Vowel chart representing the vowel productions of five native American English speakers (“e” = / $\text{æ}$ /, “v” = / $\Lambda$ / and “a” = / $\alpha$ /).

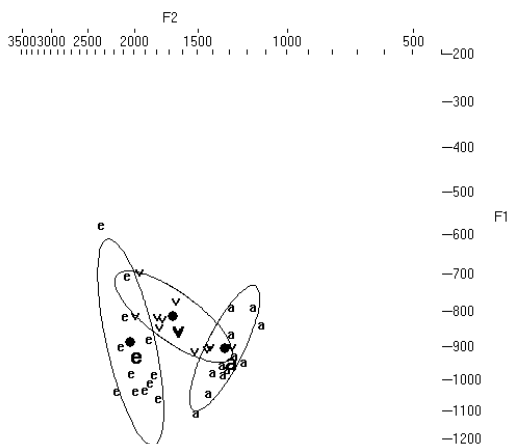




Figure 2: Vowel chart representing the vowel productions of five ten Catalan learners of English with low FA. (“e” = /æ/, “v” = /ʌ/ and “a” = /ɑ/).

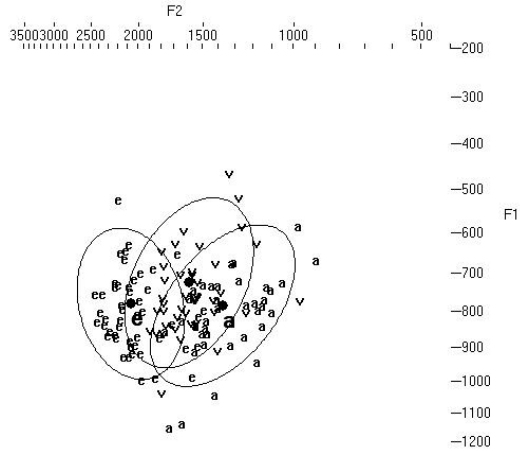
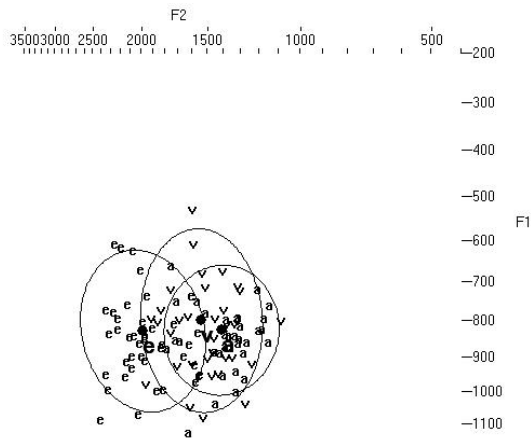
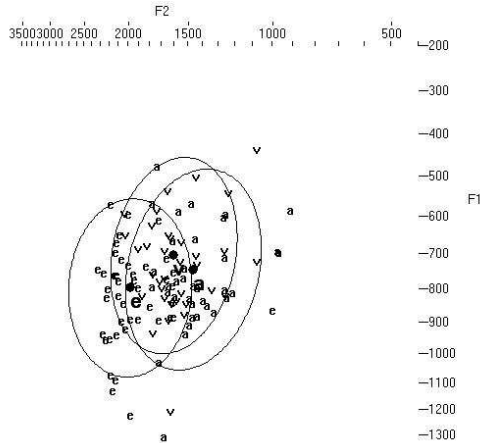


Figure 3: Vowel chart representing the vowel productions of seven Catalan learners of English with mid FA. (“e” = /æ/, “v” = /ʌ/ and “a” = /ɑ/).



## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

Figure 4: Vowel chart representing the vowel productions of seven Catalan learners of English with high FA. (“e” = /æ/, “v”= /ʌ/ and “a”= /ɑ/).



### 4. Discussion

The results just reported show that some late learners of English with limited exposure to the target language can produce some nonnative sounds authentically. There was no significant difference in the spectral data of vowels /æ/ /ʌ/ between the three groups of learners and the English native group. However, two groups, the Low and High FA, failed to produce good instances of vowel /ɑ/, more specifically, they produced /ɑ/ tokens with significantly lower F1 than the native English group. Remember that in a prior work in which I tested on perceptual similarity between Catalan and English vowels I found that English vowels /æ/ /ʌ/ and /ɑ/ were all perceived as similar to Catalan /a/. The Speech Learning Model predicts poor pronunciation if a certain L2 sound is perceived as similar to an L1 sound.

It is likely that phonetic training enhanced correct pronunciation of two of the target sounds. García Lecumberri *et alii*. (2003) claim that transcription raises learners' awareness of

pronunciation errors. This would explain the fact that the Catalan learners produced nativelike instances of two of the target vowels, despite these vowels being assimilated to the same Catalan vowel.

The vowels elicited by the three groups of learners showed a tendency to centralization relative to the vowels elicited by the AmEng group. Overall, this trend increased as a function of foreign accent, which means that the acoustic distances between the three target vowels were larger for the low FA group compared to the high FA group. The less expanded vowel spaces of the three groups of learners have immediate consequences on vowel intelligibility, because they maximize potential confusions between neighbouring vowel categories.

### References

- Flege, J.E. 1995. "Second-language Speech Learning: Theory, Findings and Problems". *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. W. Strange. Ed. 229–273. Timonium, MD: York Press.
- Flege, J. E., I. MacKay and D. Meador. 1999. "Native Italian Speakers' Production and Perception of English Vowels". *J. Acoust. Soc. Am.* 106. 2973–2987.
- Flege, J. E., C. Schirru and I. R. A. MacKay. 2003. "Interaction between the Native and Second Language Phonetic Subsystems". *Speech Communication* 40. 467–491.
- García Lecumberri, M. L., J. Maidment, M. Cooke, A. Ericsson and M. Giurgiu. 2003. "A Web-based Transcription Tool". *Proceedings of the 15th International Congress of Phonetic Sciences*. M. J. Solé, D. Recasens and J. Romero. Eds. 981–984. Barcelona: Causal Productions.
- Guion, S. G., J. E. Flege, R. Akahane-Yamada, R. and J. C. Pruitt. 2000. "An Investigation of Current Models of Second Language Speech Perception: The Case of Japanese Adults' Perception of English Consonants". *Journal of the Acoustic Society of America* 10:5. 2711–2724.
- Hillenbrand, J., L. A. Getty, M. J. Clark and K. Wheeler. 1995. "Acoustic Characteristics of American English Vowels". *Journal of the Acoustical Society of America* 97: 5. 3099–3101.
- Ladefoged, P. 2001. *Vowels and Consonants*. London: Blackwell.
- Labov, W. 1998. "The Three Dialects of English". *Handbook of Dialects and Language Variation*. M. D. Linn. Ed. 39–81. San Diego: Academic Press.
- Moyer, A. 1999. "Ultimate Attainment in L2 Phonology. The Critical Factors of Age, Motivation and Instruction". *Studies in Second Language Acquisition* 21. 81–108.

## PROCEEDINGS 31<sup>ST</sup> AEDEAN CONFERENCE

- Piske, T., J. E. Flege, I. MacKay and D. Meador. 2002. "Early and Late Bilinguals' Production of English Vowels in Words, Non-words and Conversational Speech". *Phonetica* 59. 49–71.
- Rallo Fabra, L. 2005. "Predicting Ease of Acquisition of L2 Speech Sounds. A Perceived Dissimilarity Test". *Vigo International Journal of Applied Linguistics* 2. 75–92.
- Recasens, D. 1986. *Estudis de fonètica experimental del català oriental central*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Strange, W. 1995. "Cross-language Studies of Speech Perception. A Historical Review". *Speech Perception and Linguistic Experience*. W. Strange. Ed. 3–46. Baltimore: York Press
- . 2007. "Cross-language Phonetic Similarity of Vowels. Theoretical and Methodological Issues". *Language Experience in Second Language Speech Learning. In honor of James E. Flege*. O. S. Bohn and M. J. Munro. Eds. 35–55. Amsterdam: John Benjamins.
- Tsukada, K., D. Birdsong, E. Bialystok, M. Mack, H. Sung and J. E. Flege. 2005. "A Developmental Study of English Vowel Production and Perception by Native Korean Adults and Children". *Journal of Phonetics* 33. 263–290.