Spanish listeners’ perception of American and Southern British English vowels

Paola Escuderoa and Kateřina Chládková
Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012 VT Amsterdam, The Netherlands
escudero@uva.nl, k.chladkova@uva.nl

Abstract: L2 studies demonstrate that learners differ in their speech perception patterns. Recent explanations attribute this variation to the different initial stages with which learners start their L2 development. Spanish listeners’ categorization of Standard Southern British English and American English vowels is compared. The results show that, on the basis of steady-state F1 and F2 values, listeners classify the vowels of these two English varieties differently. This finding suggests that the dialect to which learners are exposed determines their initial stage for L2 perception and the tasks they need to perform to successfully acquire a new sound system.

© 2010 Acoustical Society of America
PACS numbers: 43.71.Hw, 43.71.Es, 43.71.An [JH]
Date Received: June 26, 2010 Date Accepted: August 10, 2010

1. Introduction

It is common knowledge that second-language (L2) speakers differ from one another with respect to how native-like they perceive and produce L2 sounds. Researchers have shown that a multiplicity of non-linguistic factors can explain the individual variation in the level of L2 speech attainment. Among these factors the most notable are the age of onset of L2 learning and the length of residence in the L2-speaking country. Other factors include amount of L2 use and motivation to learn and use the L2 (Piske et al., 2001).

Other studies have explained L2 individual variation through the comparison of the L2 variety to which learners are exposed and their native language. For instance, Escudero and Boersma (2004) proposed that Spanish learners of Standard Southern British and Scottish English go through different developmental paths depending on the variety of English they are learning. Interestingly, however, two recent studies have shown that these developmental paths are also learner-dependent (Morrison, 2009; Mayr and Escudero, 2010). In these studies, there was systematic variation between learners, despite the fact that they had all been exposed to a single variety of their target language. As the authors propose, a plausible explanation for this individual variation is that each learner starts their process of acquisition with a different initial stage. In that respect, Escudero and Boersma (2004) and Morrison (2009) showed that learners’ initial attention to the specific dimensions that are involved in the production of an English vowel contrast, specifically their relative perception of duration versus the vowels’ F1 and F2 values, affected their subsequent L2 development. Similarly, Mayr and Escudero (2010) showed that L2 contrasts that were assimilated to a single L1 category in a non-native identification task where listeners classify the L2 sounds in terms of native sounds turned out to be the most difficult ones in an L2 categorization task.

The Perceptual Assimilation Model (PAM) (Best, 1995) accounts for cross-language perception by defining a variety of assimilation or classification patterns that occur when non-native listeners encounter the sounds of a new language. Listeners who classify non-native sounds as equally good examples of a single L1 category are expected to have the most difficulties with discriminating these sounds in the L2. The approach toward perceptual assimila-
tions described in the PAM model was incorporated within the Second Language Linguistic Perception model (L2LP) (Escudero, 2005). The L2LP model posits that learners’ initial L2 perception is equal to how monolingual listeners perceive the target language sounds through their native language sound system (see also Best and Tyler, 2007). The L2LP also presupposes that non-native perception is a reliable predictor of later L2 development.

The present study aims at assessing the cross-language perception of two English varieties by Peruvian Spanish listeners in order to infer their initial stage and predict their further L2 development. Listeners were presented with Standard Southern British English (SSBE) and American English (AE) vowels, which are varieties that have been shown to differ in their phonetic realizations, as seen in the comparison of the acoustic properties of SSBE vowels reported in Deterding (1997) and AE vowels in Hillenbrand et al. (1995). Variety-dependent differences in perceptual assimilations are found and specific learning scenarios and developmental paths are inferred for Peruvian learners of each of the two English varieties.

Unlike the previous studies that mostly tested a limited number of L2 vowel sounds or contrasts (e.g., Escudero and Boersma, 2004; Morrison, 2009), the present study focuses on all of the vowels that are present in two English varieties. That is, it is assessed how Spanish listeners perceive the F1 and F2 values of the nine phonemes that have been traditionally described as part of the vowel inventories of SSBE and AE, i.e., /i, ɛ, æ, ʌ, ɔ, ʊ, u/.

2. Method

2.1 Participants

The 40 participants (20 men and 20 women) were young monolinguals from Lima, with an age range of 18 to 30 years, who were students at the Pontificia Universidad Católica del Perú. They were raised by monolingual Peruvian parents and they estimated their knowledge of any language other than Spanish as no more than 2 on a scale from 0 to 7 (0=no knowledge of the language and 7=native-like knowledge). All participants reported to have normal hearing. They received a fixed hourly rate for their participation.

2.2 Stimuli

The stimuli were isolated vowels synthesized using the Praat program (Boersma and Weenink, 1992). The synthesis included the nine monophthongs common to SSBE and AE, namely, /i, ɛ, æ, ʌ, ɔ, ʊ, u/, and the five monophthongs of Spanish, namely, /i, ɛ, a, o, u/. The synthetic vowels differed mainly in their first formant (F1) and second formant (F2) values, which were kept relatively constant throughout the duration of the vowel. Steady-state formants were considered because the main objective of the present study was to compare the Spanish perception of the two varieties based on the same grounds. That is, SSBE and AE have the use of static F1 and F2 values for signaling vowel identity in common but not formant movement, because while AE monophthongs exhibit large formant movement (Hillenbrand and Nearey, 1999), SSBE vowels are true monophthongs (Hawkins and Midgley, 2005; Ferragne and Pellegrino, 2010). In addition, the Spanish vowels included in the stimuli set as well as the Spanish response categories available to the listeners are all produced as monophthongs, and naive Spanish listeners seem to use static F1 and F2 values and ignore formant movement to classify, for instance, English /i/ and /ɛ/ as Spanish /i/ or /ɛ/ (Morrison, 2009).

The F1 and F2 values for the synthetic SSBE vowels were the mean F1 and F2 values of the vowels produced by 5 male professional speakers reported in Deterding (1997), the F1 and F2 of the AE vowels were the mean values of the vowels produced by 45 male speakers reported in Hillenbrand et al. (1995), and the F1 and F2 of the Spanish vowels were the mean values of the vowels produced by 10 (normal-voiced) male speakers reported in Cervera et al. (2001). In total, 23 different vowels were synthesized. Spanish vowel tokens were included for the stimulus set to sound more native-like to the Spanish listeners, and to assess, on the basis of the results, whether participants paid attention to the task and used the five response categories correctly.

The vowels had the same fundamental frequency (F0), duration, and higher formants
They had a falling F0 contour which started at 150 Hz at the vowel onset and fell down to 100 Hz at the vowel offset. They had the same duration, namely 128 ms. This was because monolingual Spanish listeners do not use vowel duration as a cue to vowel identity when classifying native and non-native vowels (Morrison, 2008). Although early L2 studies show that Spanish learners of English use vowel duration to discriminate English vowels (Bohn, 1995; Flege et al., 1997), more recent studies demonstrate that this durational use is developmental. That is, Spanish listeners learn to use vowel duration through exposure with the English language (Escudero and Boersma, 2004; Kondaurova and Francis, 2008; Morrison, 2009). For vowel tokens with F2 lower than 1500 Hz, the third formant (F3) was 2500 Hz; for tokens with F2 higher than 1500 Hz, F3 was computed as the vowel’s F2+1000 Hz. Higher formants were added in a similar fashion to create a flatter spectrum.

Figure 1 shows the F1 and F2 values of the English and Spanish stimuli together with values of male Peruvian vowel productions from Chládková et al. (submitted). Based on how the SSBE and the AE vowels compare acoustically to Spanish vowels, one can expect differences between how Spanish listeners will classify the SSBE and AE vowel tokens. Specifically, they will classify /æ/ differently when produced in the two English varieties because SSBE /æ/ is closest to Spanish /a/ while AE /æ/ is closest to the Spanish /e/. Similarly, the closest Spanish vowels to SSBE /æ/ seem to be different from the closest Spanish vowels to AE /i, e, a, ë, u/, which will lead Spanish listeners to choose a different set of Spanish vowels for these five vowels in the two dialects.

2.3 Procedure

The testing took place in a quiet room at the Pontificia Universidad Católica del Perú in Lima. The experiment was a multiple-forced choice task set up with the program Praat (Boersma and Weenink, 1992). The stimuli were presented via headphones at a comfortable hearing level. The experimenter was a female native speaker of Peruvian Spanish who addressed the participants in Spanish throughout the testing session.

The task was to identify the incoming auditory stimuli with one of the Spanish response categories presented on a computer screen. The labels of the five Spanish vowel categories were presented orthographically. Listeners were told that they would hear Spanish vowels which were extracted from recordings of running speech and might thus sometimes sound like poor examples of the vowels. They were instructed to choose one of the five categories even when unsure. Prior to the experiment, listeners were given practice trials with the five Spanish vowel tokens, played once each, to ensure that they understood the task and became familiar with the nature of the stimuli. Their responses to these practice trials were not counted for the analysis of the results. Participants were able to easily classify these tokens as the intended Spanish vowel during the five-trial practice session.

The complete stimulus set to which listeners were exposed during the testing session included 10 repetitions of each of the 18 English vowels, and 5 repetitions of each of the 5
Spanish vowels, i.e., 205 stimuli in total. The between-trial interval was one second. There was a possibility to take a short pause every 25 trials. The participants took on average 20 minutes to complete the experiment.

3. Results

As expected from the comparison of the acoustic properties of Cervera et al. and Chládková et al., listeners identified the Spanish tokens as the intended Spanish vowel 95.6 percent of the time (99% confidence interval of the mean percentage=93–98). They thus seem to have used the Spanish response categories correctly and according to the tokens’ acoustic properties. This suggests that it is likely that their English responses were also based on the tokens’ acoustic properties.

3.1 Classification percentages

Table 1 presents the results for the nine SSBE and AE vowels. The table shows the mean percentage of times, averaged across 40 listeners, that each vowel was classified as each of the five Spanish vowels. Only responses that occur in at least 5% of the time are reported.

It can be seen that only SSBE and AE /i/ and /u/ show comparable assimilations. A repeated-measures analysis of variance was conducted to test the significance of the observed dialectal differences. The dependent measure was the classification percentage and the within-subject factors were English variety (2 levels=SSBE and AE), English vowel (9 levels=9 English vowels), and Spanish response category (5 levels=5 Spanish responses). There was a significant triple interaction English variety × English vowel × Spanish response category (F[32,1248,ε=0.107]=118.46, p=6.1·10⁻⁴⁰). This suggests that the classification of some English vowels is influenced by whether they are produced with SSBE or AE values.

It was then tested which vowels yielded differential response patterns across dialects. Thus, for every possible combination of Spanish response and English vowel, the difference score was computed, i.e., a _difference score_. For each English vowel, the difference score of all the possible pairs of the five Spanish responses were computed, i.e., 10 comparisons per vowel. As is seen in Table 1, there were mostly two responses per stimulus, but sometimes also three. For instance, Spanish /a/ and /e/ were chosen for SSBE /æ/, while /e/, /i/ and /u/ were chosen for SSBE /u/. Given that there were on average two Spanish responses for each English vowel, the significance level was corrected accordingly for each of the nine English vowels, i.e., α=0.05/(9×2)=0.003. The analysis revealed that 7 out of the 9 English vowels, i.e., /i, e, æ, a, ɔ, ʌ, u/, yielded different patterns for AE than for SSBE [for all of them t(39)=3.215, p(two-tailed)<0.0013].

The results show that the acoustic predictions were borne out because the vowels /i, e, æ, a, ɔ, ʌ, u/, which compared differently to Spanish vowels when they had SSBE or AE values, indeed yielded a significant between-variety difference in their classification. These distinct response patterns for SSBE and AE vowels imply that the initial stage of Spanish learners acquiring SSBE is different from that of Spanish learners acquiring AE.

3.2 Perceptual assimilation patterns

The L2LP model (Escudero, 2005) predicts that, initially, Spanish learners of English will exhibit a number of single category assimilations. That is, given that there are (almost) twice as
many categories in English as in Spanish, beginning learners will likely assimilate some of the English contrasts to a single Spanish category, and therefore will not be able to hear the contrast. Figure 2 illustrates the perceptual assimilations found in the Spanish listeners of the present study. It can be readily seen that SSBE and AE vowel contrasts undergo different types of assimilation in Spanish listeners. For instance, Spanish listeners use a single Spanish label, namely /e/, for three different AE vowels, namely /i/, /ε/, /æ/, which suggests that they may perceive the AE three-way contrast /i/-/ε/-/æ/ as Spanish /e/. On the other hand, it seems that they hear at least two different vowels for the same three-way contrast in SSBE because /ε/ and (partially) /æ/ are perceived as Spanish /e/, while /æ/ is perceived as Spanish /a/.

It can also be observed that, although SSBE /u/ is partially assimilated to Spanish /o/, the SSBE contrast /æ/-/u/ is mostly perceived as Spanish /a/, which suggests that Spanish listeners have trouble discriminating it. In addition, the two AE vowels /o/ and /u/ are perceived as Spanish /o/, while SSBE /u/-/æ/ seems to be a difficult contrast because both vowels are perceived as Spanish /u/. Interestingly, some English vowels were heard as more than one Spanish category, which exemplifies Escudero and Boersma’s (2002) multiple category assimilation. This occurs differently for the two English varieties: while SSBE /æ/ assimilates to Spanish /æ/ and /a/, AE /ε/ is heard as Spanish /a/ and /o/.

As predicted, Spanish listeners perceived English vowels as the Spanish vowel that has the closest F1 and F2 values. More specifically and perhaps not surprisingly, AE and SSBE vowels with relatively high F2 values such as /i/ or /ε/ are mapped to Spanish vowels with high F2 values such as /i/ or /æ/, while AE and SSBE vowels with relatively low F2 values such as /o/ or /u/ are mapped to Spanish vowels with low F2 such as /o/ or /u/.

SSBE vowels are also mapped to the Spanish category that is closest in terms of F1 but not in terms of F2. That is, SSBE /u/ which has a high F2 value is perceived 41 percent of the time as Spanish /u/ which has a low F2 value. Similarly, SSBE /ε/ is assimilated to either /u/ or /o/ 20 percent of the time. Figure 1 explains these surprising patterns: the SSBE vowels /i/ and /ε/ have markedly lower F2 values than those of Spanish /i/ and /ε/. That is, the F2 values of SSBE /i/ and /ε/ lie outside the 2σ ellipse (which is shown in the figure) of Spanish /i/ and /ε/. In contrast, neither AE /i/ nor /ε/ exhibits this pattern because their F2 values are within those of Spanish /i/ and /ε/. Assuming that F2 is a cue to vowel backness, the fact that AE /i/ and /ε/ were always assimilated to a Spanish vowel with a high F2 but SSBE /i/ and /ε/ were sometimes assimilated to a Spanish vowel with a low F2 value suggests that the Spanish perceptual boundary between front and back vowels is different from that of SSBE. That is, the Spanish boundary seems to lie within the F2 space of SSBE front vowels, which results in many SSBE front vowels being classified as Spanish back vowels.

4. Discussion

In this study, Spanish responses to the F1 and F2 values of the SSBE and AE vowels were analyzed and the findings show between-variety differences in the classification of seven Eng-
lish vowels. On the basis of the classification percentages, we inferred the patterns of perceptual neutralization for both English varieties. The differences across the two English varieties suggest that the initial stage of Spanish learners is specific to their target English variety.

With respect to the L2LP model, several predictions can be made about Spanish listeners’ further L2 development. The model states that learners who hear most tokens of the two categories in an L2 contrast as the same L1 category will face the task of either splitting their existing L1 category or creating a new one. On the other hand, learners who are able to classify some tokens of one of the L2 categories as a different L1 vowel, i.e., exhibit two-category assimilation of an L2 contrast, will just need to shift their perceptual boundary for this contrast. That is, the degree of single category assimilation defines the learning path and the degree of difficulty. According to the L2LP model, a learner who has assimilated an L2 contrast to two L1 categories, despite a lack of perfect match between the L1 and L2 distributions of the categories, will have a less demanding L2 task than a learner who has assimilated the same two categories to a single L1 category (Escudero, 2005).

These differences in the perception of AE and SSBE vowels imply that the perceptual development of learners of these varieties will be different. For instance, a learner of SSBE will need to split her existing L1 category /u/ to successfully acquire the English /u/-/ʊ/ contrast, while a learner of AE will only need to slightly shift her boundary for this contrast, if at all. Further, a learner of AE will face the tricky task of splitting her /ɛ/ category in three to successfully acquire English /ɛ/-/ɛ/-/æ/, while a SSBE learner will have more difficulty with /æ/-/æ/-/ɛ/. Moreover, the present findings suggest that the SSBE learner, but not the AE learner, will have to learn to redefine her cue specification for the backness feature in order to shift her front-back boundary toward lower F2 values.

It is worth mentioning that Spanish listeners might have assimilated AE vowels differently if the synthetic stimuli considered in this study had the formant movement observed in natural productions of many of the AE monophthongs. Further research would need to show if formant movement plays the same important role in non-native perception of AE vowels as it does for native listeners (Hillenbrand and Nearey, 1999).

The present study has clearly shown what the cross-language perceptual assimilation patterns are for the vowel system and the spectral properties that are common to SSBE and AE. There are thus important differences between learners of SSBE and AE at the initial stage of their L2 development. Predictions were made about L2-variety-specific developmental scenarios for various English vowel contrasts. Results of future experiments with beginning and advanced learners will show whether these predictions are borne out.

Acknowledgments

This research was supported by Grant No. 275.75.005 from The Netherlands Organization for Scientific Research awarded to P.E. We are grateful to Paul Boersma and Evan Cohen for their valuable comments.

References and links


