Proceedings of ISCA Workshop on Plasticity in Speech Perception (PSP2005); London, UK; 15-17 June 2005

Vowel discrimination in early bilinguals: How plastic?

Anders Højen

Speech and Hearing Sciences Division
University of Alabama at Birmingham, U. S.
AndersHojen@yahoo.dk

Abstract

Previous research has suggested a lack of plasticity for second-language (L2) perceptual learning even in individuals who learned their L2 in childhood. This study examined the discrimination of English vowels by native (L1) Spanish speakers who learned English in childhood. Results showed that early bilinguals' performance was much more similar to that of English monolinguals (though differences were found) than to that of Spanish monolinguals, suggesting extensive plasticity for L2 perceptual learning in early childhood.

1. Introduction

Recent neuroscience research has revealed far more plasticity in many brain functions than previously thought possible, [1], [2], [3]. This has spurred renewed interest in whether, or to what extent, the systems underlying human speech perception remain plastic.

Recent L2 speech perception research has shown that even early bilinguals may differ from native speakers in their perception of L2 vowels or consonants, see [4] for a review. These native-nonnative differences were interpreted as a lack of plasticity for L2 perceptual learning in early childhood. Opposite results were later reported which suggested that early bilinguals may obtain native-like L2 vowel perception, but only if they used their L1 much less than L2 [5].

These conflicting reports raise the question of "to what extent" L2 perceptual learning is constrained following acquisition of the L1 perceptual system in early childhood. To evaluate this, two reference points are useful. Bilinguals' perception of the L2 should be compared to that of monolingual speakers of the target L2, to reveal how much the bilinguals have yet to learn, and to monolingual speakers of the L1, to determine how much they have already learned.

This study used the categorial AXB discrimination test to compare the discrimination of English vowels by English monolinguals, Spanish monolinguals, and early Spanish-English bilinguals. However, a problem had to be solved. Previous research using the categorial AXB format for cross-language vowel discrimination has often yielded near-ceiling scores for non-native vowel contrasts, presumably because the format allows listeners to discriminate foreign vowels based on auditory codes, e.g., [6]. To evaluate the extent of phonetic learning, a "phonetically sensitive" test is needed, which yields high scores only for participants who generate two different phonetic codes for contrasting sounds.

A preliminary cross-language experiment was carried out which devised a variant of the AXB format that yielded low discrimination scores for foreign vowel contrasts when both vowels were heard as instances of a single L1 vowel category. At the same time, the test yielded high scores when members of a foreign vowel contrast were heard as two different L1 vowels. Thus, the test could be said to be "phonetically sensitive". This pattern of results was obtained via two manipulations. Trials testing four vowel contrasts were presented randomly in a single block (with provision for breaks) rather than blocked by vowel contrast as in previous research. Moreover, the fundamental frequency (F0) of tokens varied independently of category identity within the AXB trials testing each contrast. These procedures presumably impeded auditory but not phonetic discrimination of vowels, and were therefore used in the experiment reported below.

2. Methods

2.1. Participants

Three groups of listeners (n=20 each) were recruited in Birmingham, Alabama. The "English monolinguals" (11 f, 9 m, mean age=31) spoke English only, had native English parents, and had been born and raised in the U.S. The "Spanish monolinguals" (8 f, 12 m, mean age=29) spoke Spanish only, had native Spanish speaking parents, and had been born and raised in a Spanish speaking country in South or Central America. "Early bilinguals" (13 f, 7 m, mean age=31) had been born in a Spanish speaking country in South or Central America, had native Spanish speaking parents, and had started learning English in the U.S. between the ages of 2-10 years (mean=6). Participants passed a pure tone hearing screening (250-4000 Hz, 25 dB) and were paid.

2.2. Stimuli

A native English male produced 64 non-word stimuli in a declarative or an interrogative carrier sentence, "I say ___ now again" or "Can you say ___ now again?" The non-words were formed by inserting each of eight English vowels [i u ɪ eɪ ɑ ʌ ʊ oʊ] in a CCVC frame. Each vowel was produced four times in the declarative sentence with a relatively low F0 ("L" tokens), and four times in the interrogative sentence with a relatively high F0 ("H" tokens). Recordings were digitized and the non-words were edited out. This yielded two subsets of tokens (H vs. L) for each non-word with non-overlapping F0 at vowel midpoint. F0 ranges for H and L subsets are shown in Table 1. The eight tokens of each of the eight English vowels made up four vowel contrasts each with a different consonant frame: [gl_s] for [i]-[u], [kl_s] for [ɪ]-[eɪ], [tʃ_s] for [ɑ]-[ʌ], and [fl_s] for [ʊ]-[oʊ].

The results of a cross-language classification experiment by the Spanish monolinguals showed that English [i] and [u] were most often classified as Spanish /i/ and /u/. Therefore, [i]-[u] was expected to be relatively easy for Spanish speakers to discriminate and was thus designated the "easy" contrast.

Table 1: Ranges of F0 values for two subsets of each vowel (n=4 each) having either a relatively high (H) or low (L) F0 value at vowel midpoint.

Vowel	H F0 (Hz)	L F0 (Hz)
i	164-178	110-120
u	170-182	107-111
ɪ	160-177	113-123
eɪ	163-174	113-114
ɑ	153-170	119-122
ʌ	160-170	113-132
ʊ	166-176	118-148
oʊ	158-172	124-139

The vowels in the other three contrasts were most often classified as the same Spanish vowel ([ɪ]-[eɪ] 92% of the time, mainly as Spanish /e/; [ɑ]-[ʌ] 78%, mainly as /a/; and [ʊ]-[oʊ] 88%, mainly as /o/). These contrasts could thus reasonably be expected to be difficult for Spanish monolinguals to discriminate in a phonetically sensitive test and were designated the three "difficult" contrasts (D1, D2, and D3).

2.3. Procedure

Participants were tested individually in a sound booth and heard the stimuli over headphones. The 64 trials testing each of the four contrasts were presented for AXB discrimination in a single randomized block. The 256 test trials were preceded by 32 extra trials, which were not analyzed.

The block was presented two times with a 1000 ms or a 0 ms inter-stimulus interval (ISI) in counterbalanced order. Previous research suggested that the 0-ms condition would be more difficult than the 1000-ms condition, [7], [8]. The interval between response and the next trial was 1200 ms.

Each trial consisted of three physically different stimuli. The first stimulus (A) and the third stimulus (B) in each trial belonged to a different vowel category; the second stimulus (X) belonged to the same category as either the first or the third stimulus. The participants' task was to push a button labeled "1-2" if they heard a match between the first and the second stimuli, or push a button labeled "2-3" if they heard a match between the second and the third stimuli.

Each contrast was tested with an equal representation of AAB, ABB, BAA, and BBA trial types. In each trial type, two F0-types occurred equally often. In the "incongruent" F0-type, the F0 level (H or L) of the X stimulus was different from that of the stimulus belonging to the same vowel category (for example an AAB trial with a LHH pattern). If participants were to base their response on F0 similarities rather than vowel category match, they would always get the incorrect answer in the incongruent F0-type. In the "neutral" F0-type, participants could not base their response on F0 similarities because the X stimulus always had a different F0 level than the A and B stimuli (i.e., HLH or LHL).

3. Results

A preliminary 3 (Group) × 4 (Contrast) × 2 (ISI) × 2 (F0-type) ANOVA showed a uniform effect of F0-type across the levels of the other factors. Therefore, it was decided to average over the results for the two F0-types.

As expected, the English monolinguals obtained very high scores for all four contrasts in both ISI conditions. The Spanish monolinguals obtained high scores for the easy contrast, but near-chance scores for the difficult contrasts, see Fig. 1. The early bilinguals obtained high scores for all contrasts, though slightly lower than the English monolinguals' for the three difficult contrasts.

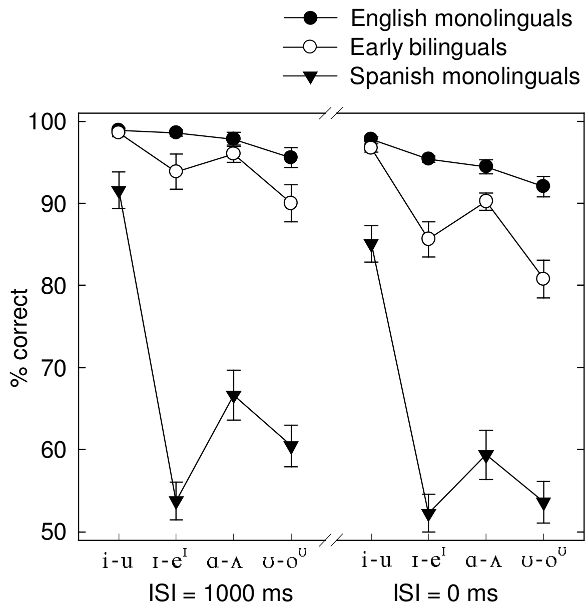

Figure 1: Mean percent correct discrimination of four vowel contrasts by three groups of listeners at an ISI of 1000 or 0 ms. Error bars bracket +/– 1 standard error.

Eight percent correct scores were obtained for each participant (4 contrasts × 2 ISI conditions). These scores were submitted to a 2 (Group) × 2 (ISI) × 4 (Contrast) ANOVA with ISI and Contrast as repeated measures. The ANOVA yielded significant main effects of Group [$F(2,57)=175.3$, $p<0.001$], ISI [$F(1,57)=53.6$, $p<0.001$], and Contrast [$F(3,171)=95.8$, $p<0.001$] and a significant three-way interaction [$F(3,171)=2.4$, $p=0.033$].

The three-way interaction was explored through simple effects tests. The simple effect of Group was significant for all four contrasts at the 1000-ms ISI [$F(2,57)=9.7$ to 180.3, $p<0.001$]. Tukey post-hoc tests with a 0.05 alpha level for each test revealed that in the 1000-ms condition, the Spanish monolinguals obtained significantly lower scores for all four contrasts than the English-monolinguals and early bilinguals, who did not differ significantly from one another.

The simple effect of Group was also significant for all four contrasts at the 0-ms ISI [$F(2,57)=15.4$ to 158.6, $p<0.001$]. However, Tukey tests revealed a different pattern of significant between-group differences than were observed at the 1000-ms ISI. The Spanish monolinguals again obtained lower scores for all four contrasts than the English monolinguals and early bilinguals did. However, the early bilinguals obtained significantly lower scores than the

English monolinguals for two of the difficult contrasts, [ɪ]-[eᶦ] and [ʊ]-[oᵘ]. This suggested that the early bilinguals' perception of English vowels was not "functionally equivalent" [5] to that of the English monolinguals.

The simple effect of Contrast was significant for all six Group × ISI combinations [$F(3,57)=4.3$ to 65.5, $p<0.01$]. Tukey tests revealed different patterns of between-contrast differences as a function of Group and ISI. For English monolinguals, [ʊ]-[oᵘ] scores were lower than [i]-[u] and [ɪ]-[eᶦ] scores at ISI=1000, but [ʊ]-[oᵘ] and [ɪ]-[eᶦ] did not differ at ISI=0. For Spanish monolinguals, higher scores were obtained for [i]-[u] than D1-D3 at ISI=0 and ISI=1000, and for [ɑ]-[ʌ] than [ɪ]-[eᶦ] at ISI=1000 ms. For early bilinguals, higher scores were obtained for [i]-[u] than D1-D3 at ISI=0, for [i]-[u] than [ɪ]-[eᶦ] and [ʊ]-[oᵘ] at ISI=1000 ms, and for [ɑ]-[ʌ] than [ʊ]-[oᵘ] at both ISIs.

Finally, significantly lower scores were obtained at ISI=0 than ISI=1000 in ten of the 12 Contrast × Group combinations [$t(19)=2.1$ to 4.8, $p<0.05$]. The two exceptions were the easy contrast for English monolinguals, and the [ɪ]-[eᶦ] contrast for Spanish monolinguals. These two pairs of scores were the highest (98% vs. 99%) and the lowest (52% vs. 54%) of the twelve pairs of scores, suggesting that the absence of a significant ISI effect could be due to ceiling and floor effects.

Individual analyses. The early bilinguals but not the Spanish monolinguals obtained native or near-native discrimination scores. However, the early bilinguals differed significantly from the English monolinguals for two contrasts in the 0-ms ISI condition. Examination of individual mean scores across the three difficult contrasts in the 0-ms condition revealed that four out of 20 early bilinguals obtained percent correct scores (64-76%) that fell outside the range of scores obtained by the English monolinguals (81-99%). The responses by these four early bilinguals to a background questionnaire were compared to the responses by the four highest-scoring early bilinguals in an attempt to identify factors might have influenced the extent of phonetic learning in the two subgroups. This comparison revealed several potentially important differences, see Table 2.

Table 2: Self-reported mean characteristics of the four highest- vs. four lowest-scoring early bilinguals. S.d. are in parentheses. AOE, age of exposure to English; English ability, self-rated ability to speak English (1="poor", 7="good").

Variable	Highest-scoring	Lowest-scoring
AOE, years	3 (2)	8 (2)
Years in Spanish schools	0 (0)	3 (2)
Years in U.S. schools	18 (3)	13 (3)
% English use last 5 years	79 (18)	63 (18)
% English use first 5 years	70 (20)	25 (17)
English ability	7 (0)	5 (1)

Compared to the highest-scoring bilinguals, the lowest-scoring bilinguals were exposed to English somewhat later in childhood, had fewer years of education in U.S. schools, and used less English (particularly during their first five years of exposure to English). In addition, the four lowest-scoring bilinguals were the only ones who did not assign themselves the maximum rating of English speaking proficiency.

To further evaluate the effect of age of exposure (AOE) on the discrimination of the three difficult vowel contrasts in the 0-ms condition, the correlation between AOE and mean score for all 20 early bilinguals was calculated. The correlation was significant [$r=0.67(18)$, $p<0.001$].

4. Discussion

This study compared the discrimination of one easy and three difficult English vowel contrasts by three groups of listeners. The discrimination scores of the English monolinguals were very high, but slightly off ceiling, attesting to the general difficulty of the test. The Spanish monolinguals' scores were high for the easy contrast but near chance for the three difficult contrasts. This pattern of results suggests that the test was phonetically sensitive.

The scores of the early bilinguals closely resembled those of the English monolinguals. The early bilinguals differed significantly from the English monolinguals in two instances, but the bilinguals' scores were always much closer to the English monolinguals' than to the Spanish monolinguals' scores. It appears that the bilinguals' perceptual systems had undergone a substantial reorganization, which can probably be ascribed to their being exposed to English from childhood. This result suggests that L2 perceptual learning is not severely constrained in early childhood, cf. [4].

However, the results of the individual analyses showed that AOE and language use patterns may influence the extent to which bilinguals achieve native-like L2 vowel perception. These results are important for the discussion of competing views on the basis for L2 speech learning difficulty.

Given that all bilinguals began learning their L2 before age 10, the AOE effect would not be predicted by a critical period account of L2 learning difficulty. But the results support two related models of L2 speech learning difficulty.

The Speech Learning Model (SLM) [9] posits that native-nonnative differences in L2 speech perception are caused by interference from the L1 phonetic system. Because the L1 is strengthened by use through childhood and adolescence, L1 interference will be stronger the later the L2 is learned.

Similar predictions have been derived from the domain-general Hebbian hypothesis of synaptic modification [10]. When one neuron fires another, the connection between the two is strengthened such that subsequent activation of one neuron is more likely to fire the other. For the perception of L1 sounds, this means that when a given pattern of activity in sensory input neurons elicits a particular percept (e.g., /i/) in neurons further downstream, the connection from the sensory neurons to the downstream neurons is strengthened. Subsequently, a similar pattern of activation in sensory neurons will activate neurons corresponding to an /i/ percept with greater likelihood and efficiency; the L1 perceptual system is strengthened. Similarly, if the patterns of activity in sensory input neurons for two distinct L2 sounds elicit the same percept in neurons further downstream, the tendency for this error to happen in the future is strengthened.

It is hypothesized by the SLM and follows from the Hebbian learning mechanism that the likelihood that new phonetic distinctions will be discerned and learned decreases with age already in childhood (because of the strengthening of the L1 perceptual system). In early childhood when the L1 phonetic system has not yet been strongly reinforced, small phonetic differences between L2 and L1 sounds might be enough for the sounds to activate different neural circuits, which might develop into separate phonetic categories.

However, the later an L2 is learned, the more likely learners are to get stuck in a “perceptual rut” because subtle phonetic differences are filtered out by the increasingly efficient L1 perceptual system. When this happens, added L2 experience will not help learners, but only serve to maintain the tendency to make certain perceptual errors.

Only when phonetic differences are sufficiently large to (sometimes) elicit different cortical responses might adult L2 learners gain from added L2 experience.

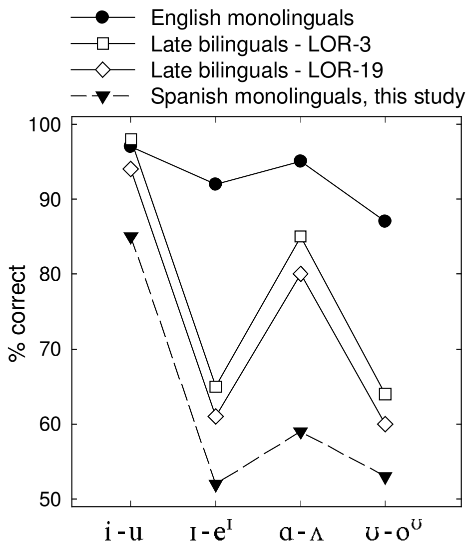

Figure 2: Mean percent correct discrimination of four vowel contrasts by three groups of listeners from [11] and by Spanish monolinguals from this study at ISI 0 ms.

Support for these predictions were obtained in a study of two groups of late Spanish-English bilinguals differing in length of residence in the U.S. (LOR=3 vs. 19 years) [11]. We examined their discrimination of the vowel stimuli used in the present study with an ISI of 0 ms. The results of the two late bilingual groups and a native English control group (n=20 each) are shown in Fig. 2; the results of the Spanish monolinguals from the present study are juxtaposed and connected with a dashed line.

Two important points may be noted. First, assuming that the discrimination ability of the two groups of late bilinguals before they started learning English was comparable to that of the Spanish monolinguals, the late bilinguals showed the most learning for the easiest of the three difficult contrasts, [ɑ]-[ʌ] (which also showed the lowest percentage of cross-language classification overlap in Spanish monolinguals, see section 2.2). This suggests that adult L1 Spanish learners of English have a greater chance to get a perceptual “foothold” to initiate perceptual learning of this contrast than the other two difficult contrasts.

Second, the LOR-19 group showed no higher (actually non-significantly lower) discrimination ability than the LOR-3 group. This result supports the Hebbian prediction that adult L2 learners can end in a perceptual rut such that added L2 experience will not result in perceptual learning.

5. Conclusion

Early Spanish-English bilinguals showed native or near-native discrimination of difficult L2 English vowel contrasts that Spanish monolinguals discriminated near chance. This is evidence for extensive plasticity for perceptual learning in early childhood. Age of exposure effects and language use effects were observed. These effects are consistent with the Speech Learning Model [9] and the Hebbian account of perceptual learning [10].

6. Acknowledgements

This research was supported by NIH Grant DC000257, the Danish Research Council for the Humanities, and the Oticon Foundation.

7. References

[1] Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B. and Taub, E. "Increased cortical representation of the fingers of the left hand in string players", *Science, Vol. 270, 1995, p 305-307.*

[2] Allard, T., Clark, S. A., Jenkins, W. M. and Merzenich, M. M. "Reorganization of somatosensory area 3b representations in adult owl monkeys after digital syndactyly", *J. Neurophysiology, Vol. 66, 1991, p 1048-1058.*

[3] Recanzone, G. H., Schreiner, C. E. and Merzenich, M. M. "Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys", *J. Neuroscience, Vol. 13, 1993, p 87-103.*

[4] Sebastián-Gallés, N. and Kroll, J. F. in *Phonetics and Phonology in Language Comprehension and Production* (eds. Schiller, N. O. and Meyer, A. S.), Mouton de Gruyter, Berlin, 2003, p 279-317.

[5] Flege, J. and MacKay, I. R. A. "Perceiving vowels in a second language", *SSLA, Vol. 26, 2004, p 1-34.*

[6] Polka, L. "Linguistic influences in adult perception of non-native vowel contrasts". *J. Acoust. Soc. Amer., Vol. 97, 1995, p 1286-1296.*

[7] Cowan, N. and Morse, P. A. "The use of auditory and phonetic memory in vowel discrimination", *J. Acoust. Soc. Amer., Vol. 79, 1986, 500-507.*

[8] Pisoni, D. B. "Auditory and phonetic memory codes in the discrimination of consonants and vowels" *Percept. Psychophys., Vol. 13, 1973, p 253-260.*

[9] Flege, J. in *Speech perception and linguistic experience: Issues in cross-language research* (ed. Strange, W.), York Press, Timonium, MD, 1995, p 233-277.

[10] McClelland, J. L., Fiez, J. A. and McCandliss, B. D. "Teaching the /r/-/l/ discrimination to Japanese adults: behavioral and neural aspects", *Physiology & Behavior, Vol. 77, 2002, p 657-662.*

[11] Højen, A. and Flege, J. "Second-language vowel discrimination and short-term memory in late bilinguals", *in preparation.*