

L2 Spanish Vowel Perception;
Empirical Findings and Implications for Pedagogical Practice

Leslie Gordon

Georgetown University

It is commonly accepted that adult L2 perception is heavily influenced by L1 phonology. Whether or not perception must precede production is a matter of some debate. While most evidence suggests that perception precedes production (Barry, 1989; Rochet, 1995), other studies have demonstrated that correct production may occur in the absence of correct perception (Goto, 1971; Sheldon & Strange, 1982). Even if both scenarios are true, research suggests that adult L2 learners have “perceptual foreign accents” which stem from the difficulty learners face in perceiving contrasts that are not functional in their native language (Strange, 1995; Beddor & Strange, 1982; Best, McRoberts & Sithole, 1988). Not all classes of L2 sounds present the same degree of difficulty. There is strong evidence to indicate that the more “similar” an L2 sound is to an L1 sound, the more difficult it is for the L2 learner to create a distinct L2 category for that sound (Wode, 1978; Flege, 1987a,b; 1988). This effect, known as “equivalence classification” (Flege, 1987a) is a topic which has not yet been explored in one of the largest groups of L2 learners: adult English-speaking learners of Spanish. In fact, L2 perception studies which examine L2 English perception by Spanish speakers far outnumber studies of English-speaking learners of Spanish (Flege et al., 1994; Fox et al., 1995; Flege, Munro and Fox, 1994; Flege and Bohn, 1989; Flege, 1991). Of the existing perception studies focused on English-speaking learners of Spanish, studies of L2 consonant perception and production historically outnumber studies of vowel perception (Flege & Hammond, 1982; Hammond & Flege, 1988).

The scarcity of data on English speakers’ perception of Spanish vowels may stem from the assumption that the five Spanish vowels are nearly identical to select categories in the English inventory and should therefore present no perceptual difficulty. A similar

claim has been made by Escudero and Boersma (2002) in a study of the perceptual challenges faced by Dutch speaking learners of Spanish. However, despite the apparent equivalence, English and Spanish vowels do not occupy exactly the same position in the vowel space and are acoustically different. Native English speakers are often observed to have difficulty suppressing some of the acoustic characteristics of their native vowels when producing Spanish vowels. What is not known, however, is whether this difficulty is correlated with or caused by flawed perception of Spanish vowels. This paper presents a preliminary study which preceded current, more in-depth research on the topic of the perceptual abilities of native English speaking adults who acquire Spanish as an L2.

The researcher's interest in this population is driven by over a decade of observation of language learners and the pedagogical practices which shape their learning. To date, data gathered from studies of L2 perception in general has yet to be fully implemented in pedagogical practice. In the increasingly rare instances in which L2 phonology is targeted in the classroom the focus is overwhelmingly on production, and in the case of L2 Spanish, much more time is devoted to problematic consonants than to vowels. Little if any regard is given to perception and its potential effects on production. Therefore, this paper addresses two core issues. First, do English speakers correctly distinguish between English and Spanish vowels? Second, how can an understanding of L2 perception inform pedagogical practice? This paper will report on results of a pilot study which examined L2 Spanish vowel perception of English speakers at different levels of instruction. Results of the pilot study will be reviewed and followed by a discussion of the pedagogical implications of this and future studies on L2 perception.

English and Spanish vowel systems.

There are thirteen vowels in the American English inventory, (Figure 1). Of these, the categories /i^y/, /e^y/, /o^w/, /u^w/ and /ʌ/ are of central interest to the present study, as these vowels most closely approximate the five Spanish vowels /i/, /e/, /o/, /u/ and /a/ (Figure 2). The English vowels /i^y/, /e^y/, /o^w/ and /u^w/ are diphthongized vowels and are the most tense of the American English vowels, however all are less tense than their counterparts in the Spanish inventory (Hammond 2001). These four English vowels are diphthongized in stressed syllables but are reduced in other contexts (to /ɪ/ or /ʊ/). Because the English stimuli in this study are monosyllabic they are stressed by default and therefore feature the diphthongized vowels /i^y/, /e^y/, /o^w/ and /u^w/

Figure 1. The American English Vowel Inventory (Hammond, 2001)

/i ^y / /ɪ/	[ɔ]	/u ^w / / /
/e ^y / /ɛ/	[χ] / /	/o ^w / /ʊ/
/æ/	/ʌ/	

Figure 2. The Spanish Vowel Inventory (Hammond, 2001)

/i/		/u/
/e/ [ɛ]		/o/
	/a/	

English speakers learning Spanish frequently encounter difficulties producing the vowels /i/, /e/, /o/, and /u/ without carrying over the diphthongized quality of their L1 tense vowels. Does the presence of this acoustic feature in English vowels also affect English speakers' correct perception of Spanish vowels? This pilot study explores the perceptual abilities of L2 learners at different levels of proficiency in order to characterize their ability – or inability – to distinguish English and Spanish vowel categories which are similar, but not identical, in quality. The vowels and test words used in this study are naturally produced by a balanced English/Spanish bilingual and have not been modified. A separate study by this researcher delves further into the role of vowel duration and diphthongization in vowel perception and utilizes synthetically modified vowels to further probe listeners' degree of sensitivity to acoustic cues which differ from English to Spanish. Plans for future research will be reviewed in later sections of this paper. The present study addresses the following questions:

1. Does native vowel acoustic quality have an effect on correct perception of stimuli by language (English or Spanish)?
2. Do 1) initial consonants and 2) syllabic structure (CV, CVC) have an effect on correct perception of native and L2 vowels?
3. Does increased exposure to Spanish have an effect on correct perception of native and L2 vowels?

Method

Participants

Three groups of university students participated in this experiment. Participants were native English-speaking students in intact classes at three levels; Introductory Spanish, Intermediate Spanish and Advanced Spanish. At each of these levels, classes are split over two semesters. The participants were drawn from the second class of each level (Intro 2 [n=14], Intermediate 2 [n=13], Advanced 2 [n=8]). Participants completed profile questionnaires regarding the amount of their previous exposure to Spanish and their proficiency in other languages.

Stimuli

The stimuli utilized in this study were CV and CVC words containing the five English and five Spanish vowels under study. The CV stimuli consisted of initial /b/ or /m/ followed by the test vowels. For the CVC stimuli, each vowel was inserted in all of the following contexts: /b_s/, /b_m/, /m_s/, and /m_m/. One bilingual English/Spanish speaker recorded the test stimuli. To avoid confusion of English and Spanish vowels, the speaker was instructed to first rehearse and then record only the English set of stimuli. This procedure was then repeated for the Spanish stimuli. The “words” were then edited

for noise and compiled into a playlist utilizing Mac i-Tunes software. The final playlist also included three repetitions of each word. Thus the final playlist included a total of 180 tokens (30 words X 2 languages X 3 repetitions). The order of tokens was then randomized before being split into two playlists consisting of 90 tokens each. In each playlist, the words were separated by 2 second intervals.

Procedure

The intact classes were brought to a computer lab to be administered the playlists. Due to technical difficulties in loading the playlists on two separate networks operating in the room, the playlists were played for all subjects via one central computer and were heard through centralized speakers. Participants were told they would hear a list of randomized English and Spanish nonsense words and were instructed to listen to each and indicate whether each word sounded to them like an English or a Spanish word. Participants recorded their answers on a score sheet on which they circled “E” or “S” for each item. The first list of 90 words was played at the beginning of the class hour (50 minutes) and the second list was played at the end of the class hour. The total time for each list was three minutes and five seconds. The Intermediate group was tested in their classroom due to the unavailability of the computer lab at their class hour.

Each student’s response sheet was hand-corrected according to a master score sheet made after building the playlists. After all score sheets were corrected, a final count of correctly identified words was calculated for each participant. In other words, “Participant X correctly identified the Spanish /bi/ token 2 out of 3 times”, and so on through the list of all test words. In the last phase of scoring, a final count of correct identification was made for each level to arrive at a group percentage of correct

identifications. Due to technical difficulties which led to the questionable sound quality of some items, one or two test words were thrown out for each group. Therefore calculations for these five items were based on 2 repetitions rather than 3.

Results

Question 1: Perception of English and Spanish stimuli

Table 1 compares the total percentage of accurately perceived English and Spanish stimuli, by learner level.

Table 1.

Percentage of Accurately Perceived Stimuli, Across Levels

	English	Spanish
Introductory learners	59%	64%
Intermediate learners	51%	68%
Advanced learners	63%	57%

Two observations can be made from the results of Table 1. First, the percentages for English tokens are barely above chance for Introductory Spanish students (59%) and Intermediate Spanish students (51%). The percentage of correctly perceived stimuli was only slightly higher for Advanced Spanish students (63%). The second observation to be made is that Spanish stimuli were more accurately perceived than English stimuli by most learners. Curiously, Advanced learners correctly perceived a higher percentage of English tokens than Spanish tokens. Possible explanations for this will follow in the discussion section.

Question 2a: Effect of initial and final consonants

Regarding the effect of initial and final consonants, there is no specific perceptual pattern which can be deduced from the present pilot study. However, there are a few observations to be made. Figures 3-14 (Appendix) give the percentage of correctly identified tokens by language (English, Figures 3-8; Spanish, Figures 9-14) and by syllable type. For English words, the overall percentage of correct identification is higher for words with an initial nasal (/m/). Here I use the term “overall percentage” to describe the range of percentages between levels for each of the words in the figures. For example, the range for English /bi/ is 50-75%; these tokens were correctly identified as “English” 50-75% of the time. Figures 2-8 show higher percentages for the words with initial /m/ than words with initial /b/. This is true for all words except /mis/, /mes/ and /mas/. I will offer my hypothesis for these lower percentages in the discussion to follow. For the Spanish test words, the reverse appears to be true. Figures 9-14 show that for most of the test words with initial /b/, the overall percentages are higher than for words with initial /m/. Notable exceptions to this pattern are Spanish /mi/, /mam/ and /mom/ (Fig. 14), all of which reflect lower overall percentages than /bi/, /bam/ and /bom/ (Fig. 11). At this point I am not certain why these words would pattern differently than what is generally observed in Figures 3-8.

Question 2b: Effect of syllable type

Total percentages for correct identification of CV and CVC tokens are given in Table 2 below.

Table 2.

Percentages of Correctly Perceived CV and CVC tokens.

Syllable type	Test tokens	Percentage correct
CV	/m__/	67%
	/b__/	68%
CVC	/m__m/	53%
	/m__s/	72%
	/b__m/	68%
	/b__s/	85%

Generally, CVC tokens were more accurately perceived than CV tokens. The exception to this pattern is syllable type /m_m/, which corresponds to the lowest overall percentage of accuracy. Looking at the CVC tokens alone, the percentages of accurately perceived tokens are highest for words with final /s/. Perhaps the explanation for this finding lies in the acoustic qualities of the fricative /s/ and the nasal /m/. The fricative /s/ is characterized by high frequency noise and does not sustain the formants of the vowel which precedes it. By contrast, nasals – /m/ in the present case – are voiced and the transition from the preceding vowel into the nasal is more continuous than the transition from vowel to fricative. Perhaps in the case of /m_s/ and /b_s/ words the abruptness of the transition is a more salient cue to vowel quality and identification. It appears then, that the listeners in this study were better aided in their identification of vowels by the presence of final consonants. Listeners were not better at distinguishing the vowels in open syllables.

Question 3: Perception across learner levels

From Table 1 we observe that performance between groups was highly variable. English words were more accurately perceived by Advanced learners but the second highest percentage is for Introductory learners rather than Intermediate learners. For Spanish words, the Advanced group performed at the lowest level of accuracy, while Intermediate students were better than Introductory students, as would be predicted. There is indeed great variability between groups. From these graphs alone it is impossible to make the claim that more accurate vowel perception correlates with more experience with the target language. Despite the fact that the different levels performed differently than expected, there were a few items that were shown to be significantly different between groups. Tables 3 and 4 list the CV and CVC stimuli that were shown to be significantly different between groups.

Table 3.

Results of a One-Way ANOVA (Group x CV)

Source	df	Sums of squares	Mean Squares	F	p
ENG /bo/					
Between groups	2	11.009	5.504	5.768	.007
Within groups	32	30.534	.954		
Total	34	41.543			
ENG /ma/					
Between groups	2	14.261	7.131	8.635	.001
Within groups	32	26.424	.826		
Total	34	40.686			
SPAN /be/					
Between groups	2	7.709	3.854	13.074	.000
Within groups	32	9.434	.295		
Total	34	17.143			
SPAN /mo/					
Between groups	2	8.236	4.131	3.441	.044
Within groups	32	38.423	1.201		
Total	34	46.686			

Table 4.

Results of a One-Way ANOVA (Group x CVC)

Source	df	Sums of squares	Mean Squares	F	p
ENG /bos/					
Between groups	2	8.591	4.295	5.210	.011
Within groups	32	26.380	.824		
Total	34				
ENG /bam/					
Between groups	2	7.989	3.994	4.007	.028
Within groups	32	31.897	.997		
Total	34	39.886			
ENG /mis/					
Between groups	2	24.466	12.233	21.153	.000
Within groups	32	18.505	.578		
Total	34	42.971			
SPAN /bis/					
Between groups	2	11.246	5.623	5.641	.008
Within groups	32	31.897	.997		
Total	34	43.143			
SPAN /bes/					
Between groups	2	2.632	1.316	3.649	.037
Within groups	32	11.540	.361		
Total	34	14.171			
SPAN /mis/					
Between groups	2	20.796	10.398	17.431	.000
Within groups	32	19.089	.597		
Total	34	39.886			

While only a small number of syllable types are shown to be significantly different between groups, Table 4 shows that the majority of significant CVC items are tokens ending in final /s/. This fact mirrors the data given in Table 2, perhaps suggesting that a final /s/ appears to make vowel quality more salient for the listeners in this study.

In the interest of highlighting any tendencies within or between groups, I can refer to the descriptive statistics of each group for each set of tokens. Tables 5 and 6 report the mean, standard deviation and mode for each group for both English and Spanish words. Shaded cells contain results of particular interest. I will begin with Table 5, which reports statistics for English words. First, I would like to repeat that the numbers reported in these tables are calculated over 3 trials. The cells highlighted in Table 5 report means of 2.5 or higher, indicating high accuracy on those particular tokens. Also, the mode reported for those cells is consistently 3, a further indication that these items were easily perceived by most listeners.

Table 5.

Descriptive Statistics for Perception of English Tokens by Learners of All Levels

	Intro			Intermediate			Advanced		
	Mean	Standard. Dev.	Mode	Mean	Standard Dev.	Mode	Mean	Standard Dev.	Mode
/bi/	1.5	1.22	0	2	1	3	2.25	0.88	3
/be/	1.28	1.32	0	1.46	0.96	1	0.75	1.03	0
/ba/	0.78	0.97	0	0.69	0.85	1	0.87	1.12	0
/bo/	2.57	0.85	3	1.46	0.87	2	1.3	1.3	0
/bu/	2.71	0.46	3	2.07	1.32	3	2.5	0.75	3
/bis/	1.35	1.08	1	1.23	1.09	0	1.25	1.28	0
/bes/	0.57	0.75	0	0.69	0.94	0	1.25	1.28	0
/bas/	0.57	0.93	0	0.38	0.65	0	1	1.19	0
/bos/	2.42	0.64	3	1.38	1.19	0	2.3	0.74	3
/bus/	1.28	1.13	1	0.61	0.86	0	1.12	1.12	0
/bim/	1.57	1.22	0	1.84	1.06	2	2.37	0.91	3
/bem/	1.14	1.09	0	1.23	1.09	1	1.12	0.83	2
/bam/	1.14	1.16	0	1.76	0.92	2	2.37	0.74	3
/bom/	2.42	0.64	3	2.61	0.65	3	2.75	0.46	3
/bum/	2.71	0.61	3	2.69	0.75	3	2.5	0.75	3
/mi/	2.07	0.82	3	1.61	1.19	3	2	0.92	1
/me/	1.07	0.91	1	1.15	0.89	1	2	1.19	3
/ma/	2.28	0.99	3	0.84	0.55	1	1.37	1.18	1
/mo/	2.78	0.42	3	2.76	0.43	3	2.5	0.75	3
/mu/	2.57	0.64	3	2.46	0.96	3	2.37	0.74	3
/mis/	0.42	0.75	0	0.61	0.86	0	2.5	0.53	2
/mes/	0.57	1.01	0	0.46	0.51	0	1.12	1.35	0
/mas/	0.28	0.61	0	0.30	0.63	0	0.62	0.51	1
/mos/	2.71	0.46	3	2.69	0.85	3	2.37	0.74	3
/mus/	2.64	0.63	3	2.07	1.18	3	2.37	0.91	3
/mim/	2.42	0.75	3	1.69	1.10	3	2.5	0.75	3
/mem/	1.07	0.73	1	1.61	1.26	0	1.25	0.88	2
/mam/	2.78	0.57	3	2.3	1.03	3	2	0.92	2
/mom/	2.78	0.42	3	2.15	1.14	3	2.25	0.88	3
/mum/	2.5	0.65	3	1.76	1.16	3	2.37	0.74	3

Table 6.

Descriptive Statistics for Perception of Spanish Tokens by Learners of All Levels

	Intro			Intermediate			Advanced		
	Mean	Standard Dev.	Mode	Mean	Standard Dev.	Mode	Mean	Standard Dev.	Mode
/bi/	1.07	0.82	1	1.92	1.18	3	1	1.3	0
/be/	1.71	0.46	2	2.61	0.65	3	2.75	0.46	3
/ba/	2.42	0.64	3	2.46	0.96	3	2.12	0.83	3
/bo/	1.42	1.34	3	1.07	1.11	0	1.87	0.99	2
/bu/	2.07	0.91	2	2.15	1.06	3	1.75	1.03	2
/bis/	2.85	0.36	3	2.23	1.16	3	1.37	1.40	3
/bes/	2.78	0.42	3	2.76	0.43	3	2.12	0.99	3
/bas/	2.78	0.42	3	2.92	0.27	3	2.25	1.16	3
/bos/	2.71	0.61	3	2.53	0.96	3	2	1.41	3
/bus/	2.85	0.36	3	2.76	0.59	3	2.37	1.06	3
/bim/	2.14	1.09	3	2	1.15	3	1	1.06	1
/bem/	2	1.3	3	2.15	1.14	3	1.5	0.92	2
/bam/	2.57	0.75	3	2.23	1.01	3	2	1.06	2
/bom/	2.3	1.00	3	2	1	3	2.12	0.83	2
/bum/	2	1.17	3	2.3	0.94	3	1.62	1.06	1
/mi/	0.78	0.97	0	1.09	1.09	0	0.87	1.24	0
/me/	2.64	0.63	3	0.63	0.63	3	2.37	0.91	3
/ma/	1.21	1.31	0	1.11	1.11	3	2.12	0.83	3
/mo/	2	1.03	3	1.18	1.18	0	1.75	1.03	2
/mu/	1.78	1.12	3	1.01	1.01	2	1.5	1.19	1
/mis/	2.64	0.74	3	0.57	0.57	2	0.62	1.06	0
/mes/	1.92	1.26	3	0.85	0.85	3	1.62	1.18	1
/mas/	2.78	0.42	3	0.43	0.43	3	2	1.41	3
/mos/	2.57	0.75	3	0.92	0.92	3	2.37	0.74	3
/mus/	2	0.96	2	1.09	1.09	3	1.62	1.18	2
/mim/	1.42	1.01	2	0.55	0.55	2	0.75	0.75	1
/mem/	1.57	1.15	3	1.03	1.03	3	0.83	0.83	2
/mam/	1.57	1.22	3	1.19	1.19	0	0.74	0.74	1
/mom/	1.28	0.99	1	1.05	1.05	2	0.70	0.70	1
/mum/	1.85	1.09	3	1.06	1.06	2	1.19	1.19	3

What is most interesting about the results in Table 5 is that accuracy was consistently high for English words containing /o/ or /u/. These are two of the diphthongized English vowels discussed in the introduction to this paper. It is curious that the results were not similar for the other two diphthongized vowels /i/ and /e/. Is there a special quality inherent in the back vowels /o/ and /u/ that is not found in the front vowels /i/ and /e/? This study does not address front/back vowel differences, but such differences could be explored in future research.

Table 6 reports the statistical results for the Spanish words. One can see that while perception of tokens containing /o/ and /u/ are sometimes high, they do not dominate the results as we saw for the English items. Instead, we see that accuracy is also high for other vowels, such as /i/, /e/ and /a/. Once again, most participants correctly identified these words for all three repetitions. However, as in Table 5, the standard deviation is consistently less than 1.00. Additionally, it appears that the participants with less target language exposure continue to outperform higher level learners on tests of vowel perception. While this finding is certainly contrary to what one would expect, results of previous studies have also shown that L2 perception can lag behind production even in late learners (Sheldon & Strange, 1982; Flege, 1993).

Discussion

I have reported higher percentages of accuracy for Spanish words (Table 3) and high means for Spanish words containing /i/, /e/, /o/ and /u/, which contrast with high means for only /o/ and /u/ of English words. If Spanish words containing /i/ and /e/ are accurately perceived more often than English words containing the same vowels, does this indicate any listener sensitivity to diphthongized front vowels? More specifically,

are listeners sensitive to the *absence* of formant transitions associated with diphthongized front vowels? While it is tempting to consider this telling evidence, I can make no such claims without further testing. Ongoing research includes development of methods designed to more carefully examine the effect of varying degrees of formant transitions on the perception of English and Spanish vowels.

One of the more perplexing findings of this study is that correct perception of English items was at or near the level of chance. Why would the participants of this study, native or near-native English speakers, incorrectly perceive English vowels with such frequency? The answer may lie in the nature of the task itself. Participants were instructed to identify each item as “English” or “Spanish”. It is possible that some participants assumed that because this procedure was carried out in their language classroom, the playlists were loaded in favor of Spanish tokens. In such a situation, it is conceivable that in cases of doubt, participants would reply with an answer of “Spanish”. Feedback received from some participants after the experiment were revealing as to the difficulty of the task. Comments such as “I was never sure”, “They all sounded the same”, and “Are you sure they were different?” indicate that for these listeners, accurately distinguishing L2 from L1 vowels was not easy.

There are other factors which potentially contribute to the results I have obtained, and I would like to briefly discuss those factors. Earlier I reported that for English /mi/, /mis/ and /mes/, percentages of accuracy were lower than for other English words with initial /m/. I propose a potential interference of vocabulary and semantic knowledge in the perception of these and other items which are meaningful words in Spanish or English. Words such as *mes* (‘month’), *mi* (‘my’, sing.) *mis* (‘my’, pl.) may have lead

participants to assume that these were Spanish words even when responding to these tokens as realized with English vowels. It is also possible that /mi/, realized with a Spanish /i/, could be interpreted as the English word ‘me’. The same explanation could be applied to /bi/. Other words in the list are also meaningful in English: /be/ ‘bay’; /bes/ ‘base’; /bum/ ‘boom’; /bu/ ‘boo’; /mo/ ‘mow’; /bo/ ‘bow’, etc. Any future study should either control for the inclusion of words that are meaningful in one language but not in the other, so that participants are not unnecessarily influenced by semantics, or ensure by other means that participants are not ascribing semantic meaning to the tokens. For the languages in question, this is a challenge to accomplish while also keeping a balance between initial and final tokens.

Pedagogical implications

As previously stated, the goal of this pilot study was to make initial observations regarding the perceptual skills of English-speaking learners of Spanish at different levels of instruction for the purpose of informing a larger-scale study currently being undertaken. While the data gathered here show high variability both between and within groups and fail to reach significance on most measures, there are nevertheless important observations to be made regarding the performance of these participants and the relevance of that performance to current pedagogical practice.

While it is logical to expect that students with more semesters of Spanish instruction would outperform beginning learners on measures of perception, data from this study shows that this was not the case. The commonly held belief that speaking follows naturally from exposure to the target language is a philosophy that has been in place since Krashen (1985) proposed that speaking skills emerge on their own after a

sufficient exposure to comprehensible input. Most L2 classrooms today follow the Communicative Approach and as a result, the emphasis is on effective delivery of spoken communication and comprehension of spoken input (i.e., understanding the whole message). However, because comprehension can encompass other factors, such as visual input and context clues, the ability of the learner to comprehend spoken input may not always coincide with correct perception of segments smaller than the sentence or word.

Challenges posed by the Communicative Approach

Most L2 classrooms today follow the tenets of the Communicative Approach. The Communicative Approach to second language teaching rose to popularity following the decline of the Audio Lingual Method (ALM) of the 1960's and 1970's. The Communicative Approach offers advantages that the ALM did not, such as an emphasis on communication and the elimination of drill-based classroom environments that raised students' affective filters. However, the Communicative Approach makes no provision for the teaching of the target sound system in the classroom (Omaggio, 1986). Terrell (1989) notes that "communicative approaches [likewise] have not known what to do with pronunciation." Because the communicative classroom strives to keep students' affective filters low, overt correction of errors is generally frowned upon. Therefore instructors are faced with somewhat of an "either/or" scenario; either emphasize oral communication at the expense of improving perception and pronunciation, or take the time to explicitly target and perhaps correct speech at the risk of raising students' affective filters. If communicative classrooms make no provision for explicit instruction of pronunciation, it stands to reason that perceptual training is virtually unheard of. Both philosophically

and in practice, perceptual acquisition of the L2 sound system may be hindered by the goals of the communicative classroom.

In the instances where instructors do in fact choose to correct their students' mistakes, the correction is made to production. This reflects another important gap in communicative pedagogy: problems in production are overwhelmingly approached from an articulatory standpoint. Depending upon the degree of concern for fostering a communicative environment and maintaining low affective filters, different forms of correction may take place. Recasting is a method frequently used in the communicative classroom. In a recast, the instructor responds to student error (of any type) by forming his response in such a way that the correct form is repeated to the student. Overt correction would involve pointing out the student's error and perhaps asking for repetition until the target is reached. These two very different correction methods have one thing in common; both are based on the assumption that repetition of the correct form will be sufficient for raising the student's awareness of her error and that she will adjust her production accordingly. However, problems in production may stem from problems in perception and students might not be responsive to recasting or multiple overt repetitions. Both perceptual and articulatory habits of the L1 take years to shape, and we must be aware that the adult learner recurs to this system when learning an L2.

Many materials used in L2 classrooms today also reflect the lack of perception research informing classroom practice. Materials that do mention pronunciation do so without reference to perception. Those that do make reference to perception do so by equating L2 and L1 sounds where possible. An example of such an approach comes from an introductory Spanish workbook currently used at a major research institution, which

equates Spanish vowels to their equivalent English vowels, giving examples such as “/a/ is like the sound in ‘father’”. A more phonetically-informed approach, used by the occasional textbook, is to describe Spanish vowels as “short and tense; tenser than English vowels”. The problem with both of these approaches is that they assume that the L2 learner 1) has a meta awareness of even his L1 that allows him to make the comparison or 2) is sensitive to the acoustic differences between the L1 and L2.

Lastly, classroom materials which include phonological instruction almost always relegate it to the final pages of the chapter, sending the message that phonological learning is secondary to lexical and grammatical acquisition, or perhaps only present it in workbooks, making phonological learning “student work”. In the past some language programs have required students to complete pronunciation exercises in the language laboratory. Today almost all L2 textbooks include media which allow students to complete the exercises at home, making language lab experience increasingly a thing of the past. This places the burden once again on the instructor to make sure that listening and pronunciation practice is completed and to perhaps make time for class review of those activities. If phonological learning and practice is left to independent study time, the same questions arise as previously discussed; how does one know that the student correctly perceives the target differences presented in his workbook? Therefore it seems that relegating listening and pronunciation practice to workbooks or final pages of text chapters cannot successfully address the issue.

Looking forward

Despite the present custom of separating perceptual and pronunciation training from other areas of L2 instruction, it remains possible and even beneficial to train

perception and production. Previous research has shown that the perceptual skills of L2 students can be improved and that such improvement can lead to an improvement in production (Rochet, 1993; Rochet & Chen, 1992; Pisoni et al., 1982; Tees & Werker, 1984; Jamieson & Morsosan, 1986, 1989). Furthermore, it appears that such training need not take place in isolation of other L2 instruction, but can in fact be accomplished in a communicative setting. Papers by Llisteri (2003, 2006) discuss the different facets of teaching L2 sounds and include suggestions for not only teaching and correction strategies but also appropriate timelines by which certain classes of sounds should be targeted in the classroom. Research currently being conducted by this author includes a training component to test the effects of explicit instruction on L2 perceptual performance.

References

- Barry, W. (1989). Perception and production of English vowels by German learners: Instrumental-phonetic support in language teaching. *Phonetica*, 46 (155-168).
- Beddor, P. & Strange, W. (1982). Cross-language study of perception of the oral-nasal distinction. *Journal of the Acoustical Society of America*, 71, 1551-1561.
- Best, C.T., McRoberts, G.W., & Sithole, N.M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14 (3), 345-360.
- Flege, J.E. (1987a). Effects of equivalence classification on the production of foreign language speech sounds. In Leather, J. & James, A. (Eds.), *Sound patterns in second language acquisition* (pp. 9-40). Providence: Foris Publications.
- (1987b). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.
- (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders: A review 1988* (pp. 244-401). Norwood, NJ: Ablex.
- (1991). Orthographic evidence for the perceptual identification of vowels in Spanish and English. *Quarterly Journal of Experimental Psychology*, 43, 701-731.
- & Bohn, O.S. (1989). The perception of English vowels by native Spanish speakers. *Journal of the Acoustical Society of America*, 1, 85, S85(A).

----- & Hammond, R. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, 5(1), 1-17.

-----, Munro, M.J., & Fox, I. R.A. (1994). Auditory and categorical effects for cross-language vowel perception. *Journal of the Acoustical Society of America*, 95, 3623-3641.

Fox, R.A., Flege, J.E. & Munro, M.J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America*, 97, 4, 2540-2551.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds “l” and “r.” *Neuropsychologia*, 9, 317-323.

Hammond, R. (1995). Foreign accent and phonetic interference: The application of linguistic research to the teaching of second language pronunciation. In F. Eckman, D. Highland, P. Lee, J. Mileham and Weber, R. (Eds.), *Second language acquisition theory and pedagogy*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

----- (2001). *The sounds of Spanish: Analysis and application*. Somerville, MA: Cascadilla Press.

----- & Flege, J.E. (1988). Attitudes, Experience, and the Mimicry of Sounds – Implications for Second Language Acquisition. Paper presented at The Seventh International Symposium on International Perspectives on Language, Literature, and Culture, George Mason University, Fairfax, Virginia.

Jamieson, D. & Morsosan, D. (1986). Training non-native speech contrasts in adults:

- Acquisition of the English / / - / / contrast by francophones. *Perception & Psychophysics*, 40, 201-15.
- & Morsosan, D. (1989). Training new, nonnative speech contrasts: A comparison of the prototype and perceptual fading techniques. *Canadian Journal of Psychology*, 43, 88-96.
- Krashen, S. (1985). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Llisteri, J. (2003). La enseñanza de la pronunciación. *Cervantes, Revista del Instituto Cervantes en Italia*, 4(1), 91-114.
- (2006). Relationships between speech production and speech perception in a second language. In K. Elenius and P. Branderud (Eds.), *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Vol. 4 (pp. 92-99). Stockholm: University of Stockholm.
- Omaggio, A. (1986). *Teaching language in context: Proficiency oriented instruction*. Boston: Heinle & Heinle.
- Pisoni, D., Aslin, R., Perey, A. and Hennessy, B. (1982). Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 297-314
- Rochet, B (1993). The role of auditory training in teaching non-native speech contrasts. Paper presented at the Third Conference on Second Language Acquisition and Foreign Language Learning, Purdue University, West Lafayette, Indiana.
- (1995). Perception and production of second-language speech sounds by

- adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore: York Press.
- & Chen, F-X. (1992). Acquisition of French VOT contrastw by adult speakers of Mandarin Chinese. In J.J. Ohala, T.M. Nearey, B.L. Derwing, M.M. Hodge, and G.E. Wiebe (Eds.), *ICSLP 92 Proceedings, International Conference on Spoken Language Processing*, Vol. 1. Alberta: University of Alberta.
- Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, 243-261.
- Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-48). Baltimore: York Press.
- Tees, R. & Werker, J. (1984). Perceptual flexibility: Maintenance or recovery of ability to discriminate non-native speech sounds. *Canadian Journal of Psychology*, 38, 579-90.
- Terrell, T. (1989). Teaching Spanish pronunciation in a communicative approach. In P. Bjarkman and R. Hammond (Eds.), *American Spanish pronunciation – Theoretical and applied perspectives*. Washington, D.C.: Georgetown University Press.
- Wode, H. (1978). The beginnings of non-school room L2 phonological acquisition. *International review of applied linguistics*, 16, 109-125.

Appendix







