

# The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis

Robert Allen Fox

Department of Speech and Hearing Science, The Ohio State University, 110 Pressey Hall, 1070 Carmack Road, Columbus, Ohio 43210-1002

James Emil Flege and Murray J. Munro

Department of Biocommunication, University of Alabama at Birmingham, VH503, Birmingham, Alabama 35294

(Received 9 January 1992; revised 11 July 1994; accepted 23 November 1994)

A group of Spanish- and English-speaking listeners participated in a multidimensional scaling (MDS) study examining perceptual responses to three Spanish and seven English vowels. The vowel stimuli represented tokens of Spanish /i/, /e/, and /a/ and English /i/, /ɪ/, /e/, /ɛ/, /æ/, /ʌ/, and /ɑ/. Each vowel had been spoken by three monolingual talkers of Spanish or English and all possible vowel pairs (405 pairs) were presented to listeners (excluding pairs representing the same vowel category). Thirty monolingual English listeners and thirty native Spanish listeners who had learned English as a second language rated these vowel pairs on a nine-point dissimilarity scale. These perceptual distances were then analyzed using the individual-differences version of ALSICAL. Results demonstrated that the English monolinguals used three underlying dimensions in rating vowels while the Spanish-English bilinguals used just two. The most salient perceptual dimension for both groups distinguished vowel height. However, for the English listeners, this dimension was most significantly correlated with duration and indicated a language-dependent sensitivity to this phonetic feature. The second dimension for the English listeners represented a front-back distinction, while the third reflected a central/noncentral distinction. For the Spanish listeners, the second dimension was less easily interpreted. However, the perceptual data for the Spanish listeners was more interpretable in terms of the distribution of the vowels in the two-dimensional perceptual plane. The vowels were distributed in terms of three separate vowel clusters, each cluster near the location of a Spanish vowel. Separate MDS analyses were carried out for subgroups of Spanish listeners who were relatively proficient or nonproficient in English. The vowel space of the proficient Spanish listeners was more Englishlike than that of the nonproficient Spanish listeners, suggesting that the perceptual dimensions used by listeners in identifying vowels may be gradually modified as proficiency in the second language improves.

PACS numbers: 43.71.An, 43.71.Hw, 43.71.Es

## INTRODUCTION

Languages differ not only according to the number of vowels used to contrast meaning, but also in terms of the phonetic properties that are used to distinguish the vowels they possess. Such differences should have implications in terms of how listeners perceive vowels, especially in the case of identifying phonetic qualities that are not represented in a listener's native language (L1).

A number of different studies have examined the perception of vowels by speakers of different languages. For example, Stevens *et al.* (1969) asked American and Swedish listeners to identify the tokens of two synthetic vowel continua: One composed of unrounded vowels ranging from /i/ to /ɛ/ and a second composed of vowels ranging from unrounded /i/ to rounded /ʌ/. As could be expected, the Swedish listeners identified fewer vowels as /i/ than did the English listeners (who have no phonemically distinct /ʌ/), demonstrating that differences in the identification of vowels can be significantly related to the vowel inventory of the listener's L1. However, ABX discrimination of these vowel tokens was very similar for the two groups, suggesting that

discrimination may have been used on more universal auditory factors. Indeed, Terbeek (1977) suggested that listening to vowels in a *psychophysical mode* might reduce or eliminate such cross-language differences (although this mode of processing is likely to be quite different from that employed in natural speech perception).

Scholes (1967, 1968) required listeners from many different language groups to identify, using keywords, the tokens from a matrix of synthetic vowels formed by factorially combining  $F1$  and  $F2$  values. He found different patterns of identification depending upon the language background of the listener. Again, these results indicated that the identification of vowels depends, at least in part, on the number and nature of the vowel categories in the listener's native language.

Terbeek (1977) obtained dissimilarity judgments among a set of 12 monophthongs from speakers of English, German, Thai, Turkish, and Swedish (languages that have significantly different vowel inventories). Using multidimensional analysis, he found significant perception differences as a function of language group—differences that depended upon

both vowel inventory and patterns of phonological opposition.

More recently, Bradlow (1993) examined the relationship between vowel inventory and vowel production and perception, specifically comparing English and Spanish listeners on three different vowel contrasts: [i-e], [u-o], and [e-o]. In her perceptual studies (using Kuhl's, 1991, "radial-rating" methodology), she found that these two groups of listeners exhibited different response patterns (identifications and goodness ratings) to the same stimuli and related these differences to the structural differences between the English and Spanish vowel space.

These studies (and others) have provided insight into cross-language differences in vowel perception, revealing likely differences in perception that depend on the nature of the listener's L1 vowel inventory. For example, Butcher (1976) suggests that a pair of vowels drawn from a portion of the acoustic vowel space that is crowded in the listener's L1 will be perceived as more dissimilar than an equidistant pair of vowels drawn from an uncrowded portion of this acoustic space. This, in turn, implies that the psychological vowel space of individuals who speak different languages may differ, perhaps through the use of different perceptual dimensions or the differential weighting of these dimensions.

The present study—an extension of an earlier study by Flege *et al.* (1994)—compares vowel perception in monolingual English listeners and bilingual Spanish-English listeners. Of particular interest is how these adult bilinguals perceive English vowels that do not have a counterpart in Spanish (i.e., new vowels) as opposed to English vowels that do have a counterpart in Spanish (i.e., similar vowels).

A total of 60 listeners (30 monolingual English and 30 Spanish-English bilinguals) rated pairs of vowels for degree of perceived dissimilarity. The vowel stimuli used were natural tokens of Spanish /i/, /e/, and /a/ and English /i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ʌ/, and /ɑ/. Three monolingual Spanish talkers each contributed one token of the Spanish vowels, and three English monolinguals did so for the English vowels. Listeners used a nine-point scale to rate vowel pairs for degree of perceived dissimilarity. The mean ratings for each possible pairing of the ten vowel categories (3 Spanish, 7 English) were submitted to multidimensional scaling (MDS) analysis. Earlier MDS analyses have examined the mean ratings of vowels by listeners differing in L1 background (Pols *et al.*, 1969; Terbeek, 1977; Terbeek and Harshman, 1971, 1972; Butcher, 1976), but this technique has apparently not been applied to ratings of multiple natural tokens of vowels drawn from two different languages.

In our earlier study (Flege *et al.*, 1994), we analyzed the obtained raw dissimilarity scores looking for significant differences in the perceptual distances between individual pairs of vowels but did not use MDS analysis. This approach had the disadvantage that no direct representation of the underlying perceptual processes (e.g., the perceptual "features" or "dimensions" used in the identification or comparison of vowel quality) of the two language groups were compared. The use of MDS techniques will allow us to develop and compare models of these underlying perceptual structures. The primary aim of the present study was thus to determine

the extent to which the Spanish-English bilinguals and English monolinguals would demonstrate different underlying perceptual dimensions in making their vowel dissimilarity judgments (and, by extension, in the perception of vowels in general). The vowel systems of Spanish and English differ considerably. English has far more vowels than Spanish (15 vs 5), and makes use of differences in at least two phonetic features to make vowel contrasts that are not found in Spanish (e.g., duration and diphthongization; Stockwell and Bowen, 1965; Delattre, 1964, 1966, 1969; Skelton, 1969; Guirao and de Marique, 1972, 1975; Godinez, 1978; MacDonald, 1989; Flege, 1989, 1991). This difference raises the possibility that the native English listeners would use more underlying dimensions in perceiving vowels than the native Spanish listeners.

A secondary aim of this study was to follow up on one incidental finding of this earlier study. The native Spanish listeners had been assigned to two subgroups based on differences in English-language experience but in no instance did the experienced and inexperienced Spanish listeners differ significantly for the 45 pairings of Spanish and/or English vowels. These results support the claim that vowel perception may change relatively little as bilinguals gain experience with the vowel system of the L2. However, "experience" in a second language (however quantified) may not always be directly linked to "proficiency" in speaking or understanding in that language. In fact, there was some evidence that listeners who differed in their ability to pronounce and comprehend spoken English (based on self-ratings as well as experimenter's ratings) did differ in their perception of vowel quality. Using correlational analysis, the relationship between the 30 Spanish listeners' mean dissimilarity ratings and both their amount of L2 experience and rated proficiency in English was investigated. Of the 45 correlations between dissimilarity ratings and L2 proficiency, five were significant. However, none of the correlations between mean ratings and amount of L2 experienced reached significance at the 0.05 level.

The findings of Flege *et al.* (1994) are thus consistent with a claim that the perceptual vowel space of Spanish-English bilinguals who are relatively proficient in English differs from that of relatively nonproficient bilinguals. Therefore, supplementary MDS analyses were carried out here to compare the vowel spaces of subgroups of relatively proficient and nonproficient Spanish-English bilinguals.

## I. METHOD

### A. Stimuli

The methods used to select English and Spanish CVs used as stimuli in the present study are described in detail in Flege *et al.* (1994). Briefly, five Spanish monolinguals produced tokens containing the Spanish /i/, /e/, and /a/, and five English monolinguals produced tokens with the English vowels /i/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ʌ/, and /ɑ/. For the Spanish monolinguals these tokens included the Spanish words *pito*, *peto*, and *pató*, spoken in the carrier phrase *Digo ahora* \_\_. The English monolinguals read randomized lists of /bVC\$/to/ and /bV\$/to/ nonwords (where \$ indicates a syllable division)

in the carrier phrases *Now I say* \_\_. The first syllables of these tokens, which were stressed, were then edited out and normalized for peak intensity. All of the talkers were male (which removes gender differences as a distracting factor).

A pilot experiment requiring native listeners of each language to identify and rate for goodness CVs from their native language, was used to select good exemplars of each vowel category and to reduce the CVs to a manageable number. On the basis of these data, the CVs spoken by two of the English talkers were eliminated. The Spanish corpus was reduced proportionally by eliminating the CVs spoken by the Spanish talkers whose tokens received the lowest average goodness ratings. The goodness ratings were then used to select one of the three available tokens of each vowel category from each talker.

Acoustic measurements were made of the nine Spanish (3 Spanish vowels  $\times$  3 talkers) and 21 English CVs (7 vowels  $\times$  3 talkers) that were selected. These measurements included voice onset time (VOT) and fundamental frequency ( $F_0$ ). Formant frequencies were estimated at three locations in each vowel: At a point 20 ms from the onset of periodicity (designated the vowel "onset"), at the acoustic midpoint, and at a point 20 ms from the end of the periodic portion (the vowel "offset"). These values are shown in Table I in Flege *et al.* (1994, p. 3627).<sup>1</sup>

As expected, the Spanish /p/s and English /b/s had similar mean VOT values (12 and 9 ms, respectively). The English tense and lax vowels differed in duration and low vowels were longer than high vowels. Also as expected, the English vowels tended to be longer than Spanish vowels of the same height, perhaps due to differences in the effect of stress on vowel duration in English and Spanish (Delattre, 1964, 1966). The format frequency values were also consistent with previous studies (Skelton, 1969; Guirao and de Manrique, 1972, 1975; Godinez, 1978).

These naturally produced vowels were not monophthongs. In particular, vowel onsets and offsets were affected by the adjacent consonants, and several English vowels show intrinsic dynamic movement (e.g., /eɪ/). The movement of the first two formants in the Spanish and English vowels are illustrated in Fig. 1. In this figure, the mean onset and offset formant values (converted to the bark scale; Zwicker and Terhardt, 1980) of  $F_1$  and  $F_2$  are shown for the ten vowel categories. This bark-scale transformation was done because this critical band scale is likely to be a more appropriate representation of the auditory properties of the stimuli than is the Hz scale (see discussion in Syrdal and Gopal, 1986; Kewley-Port and Atal, 1989). Unlike Syrdal and Gopal, we did not use Traunmüller's (1981) proposed low-frequency end correction because that would have eliminated the differences associated with  $F_0$  values in this study.

The monophthongal Spanish /e/ showed much less formant movement than the diphthong English /eɪ/, although the direction of movement was the same for both vowels. Both the onset and offset values for Spanish /a/ were intermediate to those for English /ʌ/ and /ɑ/ and all three vowels demonstrated formant movements of roughly the same magnitude and direction. The English vowels /ɪ/, /ɛ/, and /æ/ and the Spanish vowel /i/ showed little net formant movement.

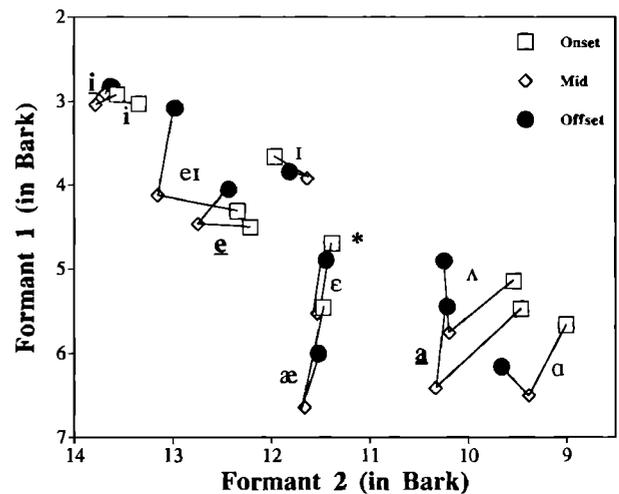


FIG. 1. The mean frequencies (in Bark) of the first two vowel formants of three Spanish /i e a/ and seven English vowels /i ɪ ɛ æ ʌ ɑ / at three temporal locations: 20 ms into the periodic portion of CVs (vowel "onset"), at midpoint, and 20 ms from the end (vowel "offset"). Each mean is based on three tokens spoken by three different talkers. The Spanish vowels are indicated by boldfaced, underlined symbols. The asterisk indicates the location of the "neutral vowel" ( $F_1=500$ ,  $F_2=1500$  Hz).

## B. Listeners

Sixty subjects were recruited in Birmingham, Alabama to serve as paid listeners. Half (10 males, 20 females) were monolingual speakers of American English and half (7 males, 23 females) were native listeners of Spanish who had learned English as an L2 (more detail about the background of these listeners can be found in Flege *et al.*, 1994). One aim of the present study was to determine if the proficient Spanish listeners would resemble native English listeners in perceiving vowels more closely than native Spanish listeners who were relatively nonproficient in English. Using an estimate of each Spanish listener's proficiency (based on self-ratings and experimenter's ratings of their overall ability to pronounce English), the Spanish listeners were assigned to a "proficient" or "nonproficient" subgroup. Each subgroup contained 13 listeners (four listeners with proficiency rankings near the middle of the group were not assigned to either group; this was intended to help avoid overlap in actual English-language proficiency between the two subgroups).

These two subgroups differed in several respects. The proficient listener's self-ratings, and those obtained from the experimenter, were higher than those for the nonproficient listeners (5.4 vs 3.; 5.1 vs 3.7). The proficient listeners were somewhat younger (30 vs 38 years) and had lived slightly longer in the US (4.1 vs 3.7 years) than the nonproficient listeners. They had also arrived at a somewhat earlier age in the US (23 vs 31 years), had studied English longer in school than the nonproficient listeners (9.3 vs 6.6 years), and reported using English more on a daily basis than did the nonproficient listeners (69% vs 43%). The proficiency rankings were not significantly correlated with age, length of residence in the US, or number of years studying English. There was a significant correlation between the proficiency rankings and percentage daily use of English ( $r=0.537$ ,  $p<0.01$ ),

and a significant negative correlation between these rankings and age of arrival in the US ( $r=0.435$ ,  $p<0.02$ ).

### C. Procedure

Each listener rated 405 CV pairs in a single session. There were nine exemplars each of 45 different kinds of pair types. For example, each Spanish vowel was paired with every other Spanish vowel, resulting in three basic Spanish-Spanish pair types ( $/i/-/e/$ ,  $/e/-/a/$ , and  $/i/-/a/$ ). There were nine instances of each of these pair types (3 talkers $\times$ 3 talkers) which produced 27 distinct vowel pairs (e.g., each of the three tokens of Spanish  $/i/$  were each paired with each of the three tokens of Spanish  $/e/$ , and so on). The same procedure was used to form pairs consisting of two different English vowels, yielding 189 English-English pairs, and to form pairs with one Spanish and one English vowel, yielding 189 Spanish-English pairs. Each vowel token in the corpus was paired just once with every other vowel token (except its two cogeners in the same category).<sup>2</sup> In the Spanish-English pairs, the Spanish and English vowel tokens occurred in the first and second position an equal number of times. The realizations of the three Spanish and seven English vowel categories occurred an equal number of times in the first and second positions of these pairs.

All listeners were tested in a sound booth, where the CVs were presented binaurally over headphones at a comfortable listening level after having been converted from digital to analog form (at a 12-bit sampling rate) and low-pass filtered at 4.8 kHz. The listeners used a nine-point scale to rate the vowels in CV pairs for degree of dissimilarity. The scale end points were defined by the labels "very similar" (1) and "very dissimilar" (9). The listeners were told to use the whole scale, and to guess if uncertain. Listeners were not trained on the rating task, but they were given practice with very dissimilar, moderately dissimilar, and very similar CV pairs (pseudorandomly sampled from the test set) before the experiment began. Each pair was presented 1.0 s after the last response. A relatively long ISI of 1.2 s was used to encourage the use of phonetic codes in long-term memory (e.g., Pisoni, 1973, 1975; Fox, 1983, 1985, 1989). A total of 45 mean values were computed for each listener, one for each of the 45 CV pairs (3 Spanish-Spanish, 21 English-English, and 21 Spanish-English).

### D. Analyses

The mean ratings obtained from each of the 60 listeners were then examined using MDS analyses. This technique is used to account for the perceived difference between pairs of stimuli by locating the stimuli within an  $n$ -dimensional perceptual space. Each listener's mean ratings ( $n=45$ ) were entered into symmetrical vowel-by-vowel matrices and then analyzed using ALSCAL, a program which assumes that dissimilarity judgments for any pair of stimuli reflects the underlying perceptual "distance" between them (Takane *et al.*, 1976).

If adults' perception of vowels changes during the process of L2 learning, one might expect more heterogeneity among the bilingual Spanish listeners than the monolingual

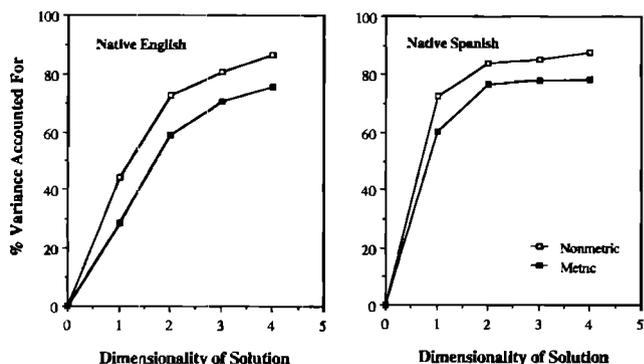


FIG. 2. Fit curves representing the cumulative variance accounted for by metric and nonmetric multidimensional scaling analyses as a function of the number of dimensions. The solutions depicted are for 30 monolingual English listeners (left) and for 30 native speakers of Spanish who had learned English as a second language (right).

English listeners. The version of ALSCAL used here was selected with this consideration in mind; namely, a weighted individual differences scaling procedure that maintains the distinction between the matrices derived for each listener. Like INDSCAL (Carroll and Chang, 1970), it develops a group space (with coordinate values for the stimuli on the extracted dimensions) and a set of subject weights indicating how salient each dimension is for each individual listener.

We made no *a priori* assumptions concerning the number or the nature of the dimensions underlying the listeners' vowel perception except to assume that the perceptual space was Euclidean. The similarity matrices from the English and Spanish listeners were analyzed separately. For each group, both ordinal (nonmetric) and interval (metric) solutions were obtained in two to four dimensions. The version of ALSCAL used here does not permit the determination of a one-dimensional solution using the weighted individual-differences Euclidean distance model. Therefore, one-dimensional solutions for the English and Spanish groups (both metric and nonmetric) were determined using ALSCAL's default Euclidean distance model. All solutions are presented unrotated.<sup>3</sup>

## II. RESULTS AND DISCUSSION

### A. Selection of dimensionality

The choice of the optimal MDS solution was based on a consideration of the amount of variance accounted for by various solutions, and the results of split-half similarity analyses. A two-dimensional, nonmetric solution was eventually chosen for the Spanish listeners whereas a three-dimensional, nonmetric solution was chosen for the English listeners. Figure 2 presents fit curves showing the cumulative variance accounted for by both metric and nonmetric solutions as a function of the number of perceptual dimensions. One reason for rejecting the metric solutions was that the nonmetric solutions accounted for at least 7% more variance than the metric solutions at every dimensionality for both listener groups. Also, it seems reasonable to view the mean ratings examined here as ordinal-level data.

For the Spanish listeners, increases in the cumulative variance accounted for by both the nonmetric and metric solutions became very gradual after two dimensions. This resulted in an “elbow” in their fit curve. Most of the variance (72.4%) was accounted for by a single dimension in the nonmetric solution. Adding a second dimension accounted for an additional 12.0% of the variance in the Spanish listeners’ data. Smaller proportions were accounted for by dimensions 3 and 4.

The fit curves for the English listeners, on the other hand, were more gradual than those for the Spanish listeners, and did not flatten out after two dimensions. The nonmetric one-, two-, three-, and four-dimensional solutions accounted for 44.2%, 72.7%, 80.6%, and 86.5% of the variance, respectively. Since the addition of dimension 3 and 4 increased the amount of variance accounted for by about the same amount, it was not possible to determine the optimal number of dimensions for the English listeners solely on the basis of the fit curves. Split-half analyses (see, e.g., Gandour and Harshman, 1978) were therefore carried out to select the dimensionality of the optimal solutions.

The Spanish listeners were evenly (and randomly) subdivided into two subgroups of 15, designated “Sp-A” and “Sp-B”. The same was done for the English listeners, with the resulting subgroups designated “En-A” and “En-B”. Identical MDS analyses (as described above) were performed on the distance data from these subgroups. The spatial solutions obtained for the two same-language subgroups were compared by computing the simple correlations between stimulus coordinates on the corresponding dimensions. A perceptual dimension was considered reliable, and thus likely to represent a psychologically real perceptual dimension for a group as a whole, if a significant correlation was obtained between the corresponding dimensions of two split-halves.

Analysis of the Spanish data confirmed the conclusion drawn from an inspection of the fit curve. The two-dimensional solutions for the split-halves (i.e., groups Sp-A and Sp-B) were nearly identical. The correlations between Sp-A and Sp-B were 0.976 for dimension 1 and 0.914 for dimension 2. Comparison of the two Spanish subgroups for a three-dimensional solution showed that while two of the dimensions were similar (and were highly correlated with dimension 1 and dimension 2 of the two-dimensional solution), the correlation between groups Sp-A and Sp-B for dimension 3 was only 0.760. This was considered to be low, given that only ten pairs of coordinates were being compared.

The split-half analysis therefore supported the two-dimensional solution for the Spanish listeners shown in Fig. 3. Dimension 1 in the Spanish listeners’ space seems to reflect a vowel height distinction. Dimension 2 is more difficult to interpret. We have suggested a possible interpretation of this vowel space by drawing ellipses around clusters of vowels in the perceptual plane where the vowels are distributed in three separate clusters, each cluster centered near the location of a Spanish vowel. This may be a demonstration of a “perceptual magnet” effect for these Spanish vowels. Specifically, while looking at L1 identification and discrimination, Kuhl and her colleagues (Kuhl *et al.*, 1992) have dem-

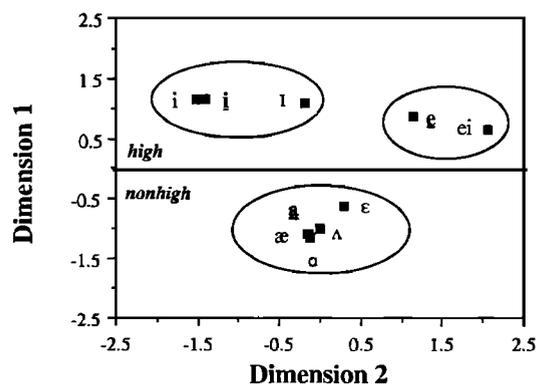


FIG. 3. The 2-dimensional ALSCAL solution obtained for 30 Spanish/English bilingual listeners. The Spanish vowels are indicated by boldfaced, underlined symbols.

onstrated that vowels in the vicinity of a vowel category “prototype” are not discriminated as easily as are vowels of equivalent auditory difference that are not located near a vowel prototype. As Flege *et al.* (1994, p. 3624) speculated, if an L2 vowel is phonetically similar to, but not identical with an L1 vowel, it may be judged as relatively more similar phonetically to the L1 vowel than it would be judged otherwise based solely on its auditory characteristics. In terms of the vowel space, we thus might expect phonetically similar but not identical L2 vowels to cluster more closely to the most similar L1 vowel if the listener has not established a new vowel category (an L2 prototype) in this space.

The split-half analysis led to a different conclusion regarding optimal dimensionality for the English listeners. The three-dimensional solutions obtained for the listeners in En-A and En-B were similar. The correlations between these groups were 0.963 for dimension 1, 0.957 for dimension 2, and 0.972 for dimension 3 (all significant at the 0.001 level). The between-group correlations for the four-dimensional solutions were lower than those obtained thus far, however. In particular, the correlations between En-A and En-B dropped to 0.671 ( $p < 0.05$ ) and 0.583 (n.s.) for two of the four dimensions. Results of the split-half procedure therefore supported selection of a three-dimensional solution for the English listeners, which is shown in Fig. 4.

As with the Spanish listeners, the English listeners’ dimension 1 seem to represent a high/nonhigh distinction. Dimension 2 corresponds to a front/back distinction. Note that the upper right (+high, +back) quadrant would normally contain /u/ and /ʊ/, vowels not included in the present study. Given this limitation, the perceptual space represented here for the English listeners is remarkably consistent with previous MDS studies examining English vowels. The absence of high back vowels might be expected to warp the plane defined by the high/low and front/back dimensions that has been found in previous research (e.g., Singh and Woods, 1971; Shepard, 1972; Fox, 1983).

The English listeners’ dimension 3 is less readily interpretable. However, it is the case that the vowels near the center of the acoustic vowel space cluster near one end of this dimension while vowels further from the center are distributed toward the other end of the dimension. This dimen-

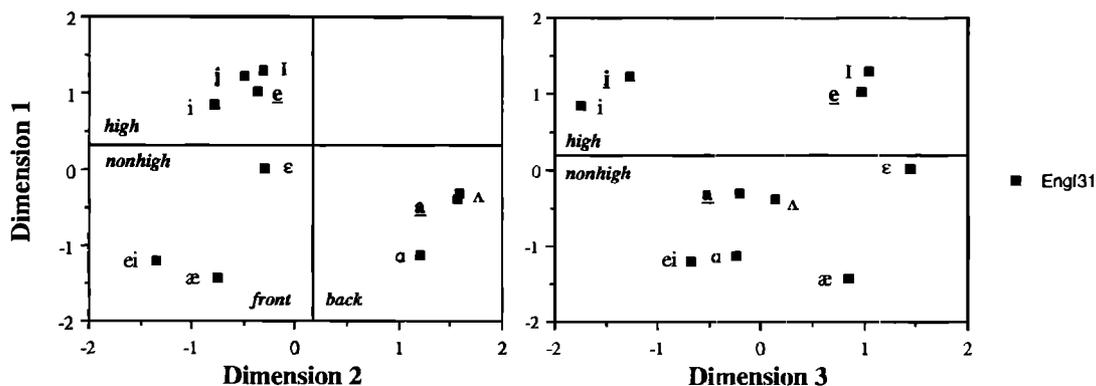


FIG. 4. The 3-dimensional ALSCAL solution obtained for 30 English monolingual listeners. The left panel is a plot of dimension 1 by dimension 2; the right panel is a plot of dimension 1 by dimension 3. The Spanish vowels are indicated by boldfaced, underlined symbols.

sion can tentatively be interpreted as a central/noncentral dimension.

### B. Interpretation of perceptual dimensions

Describing perceptual dimensions as we have just done provides only a qualitative view of their nature. A technique which allows a more quantitative evaluation of these dimensions in terms of the physical characteristics of the stimuli involves comparing the coordinates of each vowel with its auditory/acoustic properties using correlational analyses (as in Terbeek, 1977; Fox, 1983; Fox and Trudeau, 1988; Kempster *et al.*, 1991). This is a standard practice in MDS studies, but we feel it necessary to exercise more caution than is customary in evaluating the results. In most MDS studies, each vowel category is represented by a single token. Listeners then typically rate (usually more than once) all possible pairings of a single exemplars of each vowel category being examined. In the present study, on the other hand, the three Spanish and seven English vowels were represented by three tokens that had been produced by three different talkers. Accordingly, the acoustic values examined in the correlation tests had to be the means of the three tokens of each category. Although this ignores potentially important token-to-token (i.e., speaker-to-speaker) acoustic variation, it must be considered that all tokens were well-recognized exemplars of their vowel category and it is the general trend in the perception of these categories with which we are most concerned.

A total of 33 acoustic measures were examined in the correlation analyses. Most of these acoustic measures were described previously in Sec. I. In addition,  $F1-F0$ ,  $F2-F1$ , and  $F3-F2$  bark-difference values were computed as suggested by Syrdal and Gopal (1986). Bark differences were computed for frequencies observed at the onset, midpoint, and offset of vowels. As a measure of formant movement, the slopes of the  $F1$ ,  $F2$ , and  $F3$  trajectories from onset-to-midpoint, midpoint-to-offset, and onset-to-offset were calculated. In addition, given the suggested interpretation that dimension 3 of the English vowel space reflects a central-noncentral distinction, the Euclidean distances between the vowel onsets, midpoints, and offsets in the formant 1 by formant 2 plane (see Fig. 1) and the location of the neutral vowel ( $F1=500$ ,  $F2=1500$ ) in this plane were also

calculated. Correlations between these acoustic parameters and the vowel coordinates for the perceptual spaces of the English and Spanish listeners are presented in the Appendix (also contained in the Appendix are correlations between all perceptual dimensions obtained and discussed in this paper).

Dimension 1 in both the Spanish and English groups was significantly correlated with the  $F1$  and  $F1-F0$  measures (at the 0.05 level for the English and 0.001 level for the Spanish vowel spaces). This is consistent with the interpretation of these two dimension as reflecting a high-low distinction. However, dimension 1 in the English space was most highly correlated with duration (at the 0.002 level). Duration was not significantly correlated with any of the Spanish dimensions. Since duration differences are associated with vowel quality distinctions not found in Spanish, this may demonstrate a language-dependent sensitivity to this phonetic feature by English listeners, but not Spanish listeners. The  $F2$ ,  $F3$ ,  $F2-F1$ , and  $F3-F2$  measures were also significantly correlated with dimension 1 for the Spanish listeners. At first glance, this might seem to undermine the interpretation of dimension 1 as a high-low dimension. It does not, however, because vowel height was correlated with front/back distinctions in the small set of vowel stimuli included in this study. The slope of the  $F1$  onset-to-midpoint trajectory was also significant, but it was much lower than many other correlations.

For the English listeners, dimension 2 was most highly correlated with the  $F2$ ,  $F2-F1$  and  $F3-F2$  measures (at the 0.01 level) and can therefore be termed a "front/back" dimension (but note the caveat above). These results are basically consistent with scaling studies that have examined the perception of English vowels by native English-speaking listeners (e.g., Fox, 1983, 1988; Fox and Trudeau, 1988; Rakerd and Verbrugge, 1985), although the front/back dimension is usually more salient than the vowel height dimension. The different outcome obtained here is likely due to the absence of high, back vowels (e.g., /u/, /ʊ/). Dimension 2 of the Spanish vowel space was significantly correlated only to the Euclidean distance between the vowel onsets and the neutral vowel position. This correlation is significant only at the 0.05 level and may be spurious, given the large number of correlations computed. This dimension shows some general similarity to English dimension 3 although these two

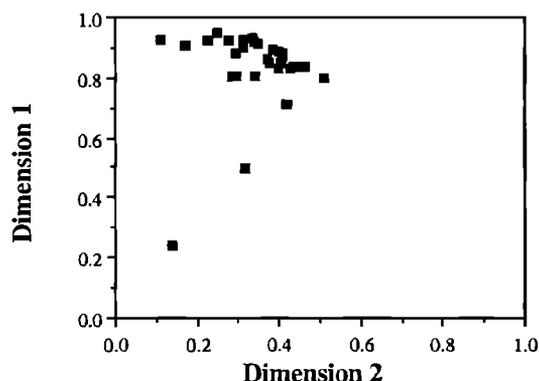


FIG. 5. Subject weights obtained for individual Spanish listeners on dimensions 1 and 2.

dimensions are not significantly correlated (see the Appendix).

Dimension 3 in the English space was significantly correlated with the Euclidean distance measures (at the 0.001 level). This dimension is similar to the peripheral/central dimension obtained by Terbeek (1977) when analyzing the perceptual data from all language groups and to the mid/nonmid dimension obtained for the English data only. However, it is likely that this dimension is not very salient in the normal process of vowel identification. For example, dimension 3 accounts for only 7.9% of the variance in this study and the analogous dimensions in Terbeek's study were the last to be extracted (and accounted for the least amount of variance).

### C. Analysis of subject weights

The subject weights provide a measure of the importance of the perceptual dimensions for each listener. The weights for the Spanish listeners are shown in Fig. 5. The mean of the subject weights for dimension 1 was 0.834, with two of the 30 listeners (one proficient, one nonproficient) showing outlying values. The mean of the weights for dimension 2 was only 0.340, suggesting that the Spanish listeners placed greater emphasis on the high-low dimension (dimension 1) than on dimension 2. It should be noted that the three Spanish vowels /i e a/ can be distinguished completely on the basis of vowel height (since neither /u/ nor /o/

were included in the stimulus set). Thus the Spanish listeners may have naturally allocated more attention to the vowel height differences among these vowels.

Correlational analyses were carried out to determine if the individual Spanish subjects' weights for either dimension 1 or 2 were related to questionnaire variables. The dimension 1 subject weights were not correlated significantly with self-estimated percent daily use of English ( $r=0.260$ ), self-estimated English-language proficiency ( $r=0.040$ ), or the experimenter's estimates of the listeners' L2 proficiency ( $r=0.060$ ;  $p>0.10$  in each instance). There was, however, a small but significant negative correlation between the dimension 2 subject weights and both the self-estimates of L2 proficiency ( $r=-0.414$ ,  $p<0.05$ ) and the experimenter's ratings ( $r=-0.404$ ,  $p<0.05$ ). Although these negative correlations might indicate that the more proficient Spanish listeners depend less on a second perceptual dimension than do less-proficient Spanish listeners, it more likely demonstrates that dimension 2 of the group space for all Spanish listeners did not accurately represent the perceptual space of the more proficient listeners alone (an interpretation that is supported by the separate analyses of the two groups presented in the next section).

Figure 6 shows the subject weights for dimension 1 (high-low), dimension 2 (front-back), and dimension 3 (central/noncentral) for the English listeners. The mean weights were higher for dimension 1 and dimension 2 than for dimension 3 (0.598, 0.542, and 0.370, respectively). The English listeners appeared to rely equally on the high-low and front-back dimensions in rating vowel dissimilarity. The central/noncentral dimension was considerably less salient to these listeners.

### D. Proficient versus nonproficient bilinguals

Additional nonmetric MDS analyses were carried out to determine if the perceptual dimensions of the Spanish listeners who were relatively proficient in English would more closely resemble those of English listeners than nonproficient Spanish listeners. Data for the both groups were compared using the nonmetric individual-differences model of ALSCAL. As shown in Fig. 7, the fit curves obtained in the analyses of data for both groups resemble closely the ALSCAL analysis of data for all 30 Spanish listeners. In-

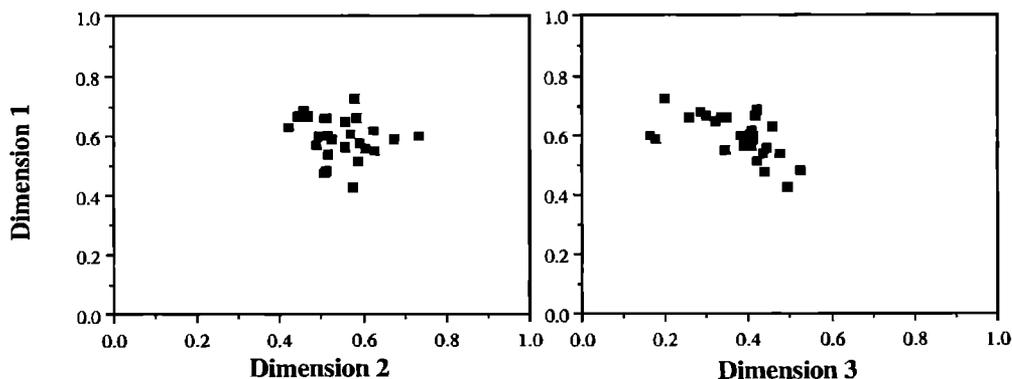


FIG. 6. Subject weights obtained for individual English listeners for dimensions 1 and 2 (left) and dimensions 1 and 3 (right).

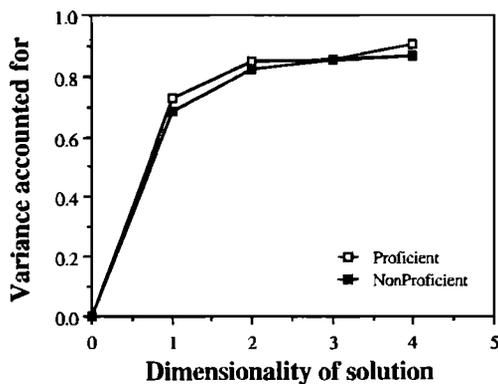


FIG. 7. Fit curves obtained in nonmetric MDS analyses for 13 relatively “proficient” Spanish listeners (unfilled boxes) and 13 relatively “nonproficient” Spanish listeners (filled boxes).

creases in the cumulative variance accounted for in both the proficient and nonproficient listeners’ data became much more gradual after two dimensions which produced an elbow in the fit curves at that dimensionality. The majority of the variance was accounted for by dimension 1 (proficient: 73.0%, nonproficient: 68.6%). Adding a second dimension accounted for an additional 11%–14% variance (proficient: 84.7%, nonproficient: 82.6%).

It thus appears that a two-dimensional solution was optimal both for proficient and nonproficient Spanish listeners. These two vowel spaces are shown in Fig. 8. In general, both of these vowel spaces are similar to that shown in Fig. 3 for all 30 Spanish listeners. Again, dimension 1 was significantly correlated with the  $F1$  and  $F1 - F0$  measures (at the 0.001 level) and dimension 2 was significantly correlated with the Euclidean distance between vowel onsets and the neutral vowel location (at the 0.05 level). However, several differences between the proficient and nonproficient spaces are evident. The location of vowels seems to be more closely clustered in the nonproficient Spanish space than the proficient Spanish space. The vowels /e/ and especially /ɛ/ have become more separated from the other phonetically similar vowels in the perceptual plane for the proficient listeners, perhaps indicating evolution towards a perceptual space that is more like the vowel space of the English listeners. As a result, English /e/ and /ɛ/ stand apart from other Spanish and

English vowel categories for the proficient but not the nonproficient Spanish listeners.

A relatively straightforward method (suggested by an anonymous reviewer) to determine whether the proficient Spanish listeners’ vowel space is more Englishlike than the nonproficient Spanish listeners’ space is to compare the intervowel distances based on the coordinates provided by the MDS solutions. Specifically, Euclidean perceptual distances were calculated between all pairs of vowels ( $n=45$ ) for each of these three groups. Correlational analysis showed that the distances in the three-dimensional English space were more highly correlated with the proficient Spanish distances ( $r=0.47, p<0.001$ ) than nonproficient Spanish distances ( $r=0.33, p<0.05$ ).<sup>4</sup> These results would support the claim that the vowel space of the proficient Spanish listeners is more similar to the English space than is the nonproficient Spanish vowel space. We readily agree, however, that these differences are small and need to be corroborated in subsequent studies.

Correlations between the acoustic measures and perceptual dimensions for the proficient and nonproficient Spanish listeners are presented in the Appendix. As in the analysis of all 30 Spanish listeners, dimension 1 obtained for both the proficient and nonproficient Spanish subgroups was correlated most significantly with  $F1$  and  $F1 - F0$  acoustic measures, and less significantly with the  $F2, F3, F2 - F1,$  and  $F3 - F2$  measures. Again, dimension 2 for both groups was correlated significantly only with Euclidean distance from vowel onset to the neutral vowel position. The correlations between dimension 1 and the acoustic measures were generally higher for the proficient than the nonproficient subgroup. No other systematic difference between the two subgroups was evident. Figure 9 shows the subject weights obtained for the proficient and nonproficient Spanish listeners. As before, the dimension 1 subject weights for two individual listeners (one proficient, one nonproficient) represented nonsystematic outliers.

$T$ -tests revealed that the subject weights obtained for the proficient and nonproficient subgroups ( $M=0.784$  vs  $0.854$ , respectively) did not differ significantly ( $t=-1.23, p>0.10$ ) on dimension 1, but did differ significantly ( $t=2.28, p<0.05$ ) on dimension 2 ( $M=0.411$  and  $0.315$ , respectively). This significant difference may also reflect a subtle

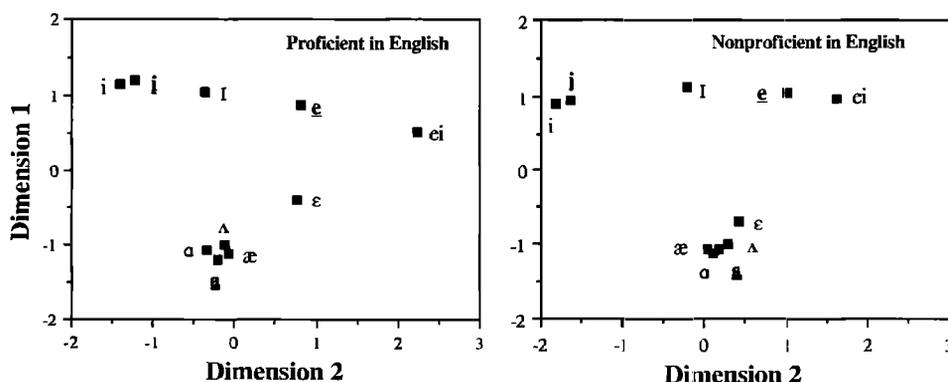


FIG. 8. The 2-dimensional ALSCAL solutions obtained for relatively “proficient” (left) and “nonproficient” (right) Spanish listeners. The Spanish vowels are indicated by boldfaced, underlined symbols.

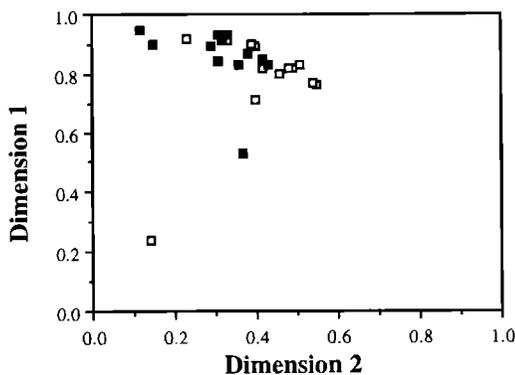


FIG. 9. Subject weights for 13 proficient (unfilled boxes) and 13 nonproficient Spanish listeners (filled boxes) for the two obtained perceptual dimensions.

change in the underlying perceptual system of adult Spanish learners of English as a function of developing greater proficiency in producing spoken English.

### III. SUMMARY

English has a larger number of contrastive vowel categories than Spanish (15 vs 5, by most counts). Not surprisingly, English uses more phonetic features than Spanish to distinguish vowels, such as the combination of spectral and temporal differences that distinguish tense from lax vowels. From this, one might expect individuals who have learned English and Spanish as native languages to perceive vowels differently. Specifically, one might expect native listeners of English to use more perceptual dimensions than native listeners of Spanish. At the very least, one might expect English listeners' reliance upon phonetic features such as duration to be greater than that of Spanish listeners. A multidimensional scaling (MDS) analysis was carried out to test this. Consistent with the hypothesis, the optimal MDS solution obtained for the native English listeners involved three dimensions (high/low or duration, front/back, and central/noncentral) whereas the optimal solution for the Spanish listeners involved just two dimensions, one reflecting a high/low distinction and a second that did not lend itself to an easy interpretation. However, the most straightforward interpretation of the Spanish vowel space was in terms of the perceptual plane in which the vowels were distributed into three clusters, each cluster corresponding to one of the Spanish vowels. In addition, one perceptual dimension of the English listeners (dimension 1) demonstrated a language-dependent sensitivity to duration. These data thus support the claim that the structure of a listener's vowel space is significantly affected by the vowel inventory of the listener's L1.

Several findings of the study bear on the issue of whether vowel perception changes as the result of L2 learning. The mean similarity ratings from proficient and nonproficient Spanish listeners were submitted to separate MDS analyses. The optimal solutions for both subgroups involved just two dimensions. Learning English apparently did not increase the dimensionality of the Spanish listeners' psychological vowel space in a manner analogous to the increase in

dimensionality of the space defining voice quality shown by individuals who undergo training on the labeling of voices (Kreiman *et al.*, 1990).

The MDS analyses did reveal two differences between the proficient and nonproficient listeners, however. First, in each of the Spanish vowel spaces, dimension 2 was identified as representing a possible central/noncentral dimension (bearing some broad similarity to dimension 3 in the English space). The greater the Spanish listener's proficiency in English, the more the reliance on dimension 2. Second, the perceptual distances between vowels in the vowel space of the proficient listeners were more highly correlated with the intervowel distances in the English vowel space than were the perceptual distances in the nonproficient Spanish space. These differences would support a claim that listeners' perceptual processes change gradually as a function of their developing production capabilities.

Clearly further research is needed to understand how vowel perception and production may change during adult second language acquisition. For example, in the present study, there were at least two limitations. First, the lack of high, back vowels in the stimulus set may have warped the perceptual space of both sets of listeners to some degree and it may have contributed to the difficulty in interpreting dimension 2 in the Spanish vowel space. Second, the use of tokens from more than a single talker was problematic because it introduced intracategory variations that could not be handled well from the MDS approach utilized. It would be useful to conduct a similar MDS study in a future experiment utilizing tokens from a single talker of a representative sample of the vowel distinctions found in English and Spanish.

### ACKNOWLEDGMENTS

This research was supported in part by NIH grant DC00257-08 to the University of Alabama at Birmingham. The authors thank Randy Diehl for help obtaining recordings from Spanish monolinguals in Austin, TX. The authors would also like to thank Diane Kewley-Port and an anonymous reviewer for suggesting many changes that significantly improved this paper.

### APPENDIX

As described in the main body of this paper, in MDS studies it is common to relate the perceptual dimensions extracted with the acoustic/auditory properties of the stimuli themselves. A technique commonly used to examine these relationships is correlational analysis (e.g., Terbeek, 1977; Fox, 1983). As noted in the body of the paper, the three Spanish and seven English vowels were represented by three tokens that had been produced by three different talkers. Accordingly, the acoustic values examined in the correlation tests had to be the means of the three tokens of each category. A total of 33 acoustic measures were examined in the correlation analyses including the following: VOT,  $F_0$ , vowel duration, frequencies of  $F_1$ ,  $F_2$ , and  $F_3$ , and the  $F_1 - F_0$ ,  $F_2 - F_1$ , and  $F_3 - F_2$  bark-difference values (Syrdal and Gopal, 1986) at vowel onset, midpoint, and offset.

TABLE AI. Pearson product-moment correlations between selected acoustic parameters of the stimulus vowels and the vowel coordinates of the perceptual dimensions for the English-only, Spanish-only, proficient Spanish, and nonproficient Spanish vowel spaces. All frequency measurements used in the calculation of correlation statistics were converted to the Bark scale (see text for details).<sup>a</sup>

	English			Spanish (all Ss)		Spanish proficient		Spanish nonproficient	
	D1	D2	D3	D1	D2	D1	D2	D1	D2
VOT	0.27	0.43	-0.41	0.04	-0.29	0.02	-0.43	0.02	-0.26
F0	0.59	0.07	-0.21	0.40	-0.19	0.40	-0.31	0.38	-0.24
Duration	-0.85	-0.15	-0.12	-0.42	0.29	-0.45	0.37	-0.37	0.29
Onset F1	-0.76	0.60	0.45	-0.92	0.40	-0.93	0.35	-0.85	0.60
Onset F2	0.56	-0.87	-0.27	0.87	-0.18	0.88	-0.10	0.83	-0.39
Onset F3	0.50	-0.79	-0.16	0.71	-0.18	0.73	-0.08	0.67	-0.36
Onset F1-F0	-0.79	0.56	0.46	-0.92	0.41	-0.93	0.37	-0.85	0.60
Onset F2-F1	0.66	-0.79	-0.35	0.91	-0.27	0.93	-0.20	0.87	-0.49
Onset F3-F2	-0.57	0.87	0.30	-0.90	0.17	-0.90	0.10	-0.87	0.39
Mid F1	-0.73	0.61	0.48	-0.96	0.24	-0.97	0.20	-0.92	0.46
Mid F2	0.49	-0.84	-0.37	0.84	-0.09	0.85	-0.02	0.82	-0.32
Mid F3	0.60	-0.81	-0.11	0.77	-0.18	0.80	-0.09	0.72	-0.36
Mid F1-F0	-0.75	0.58	0.48	-0.96	0.25	-0.97	0.22	-0.92	0.47
Mid F2-F1	0.63	-0.76	-0.44	0.94	-0.17	0.94	-0.11	0.90	-0.40
Mid F3-F2	-0.42	0.82	0.46	0.84	0.05	-0.83	-0.02	-0.82	0.29
Offset F1	-0.63	0.61	0.50	-0.92	0.10	-0.92	0.05	-0.90	0.32
Offset F2	0.51	-0.85	-0.40	0.87	-0.13	0.88	-0.05	0.84	-0.36
Offset F3	0.58	-0.80	-0.17	0.77	-0.22	0.79	-0.12	0.72	-0.40
Offset F1-F0	-0.66	0.58	0.50	-0.92	0.11	-0.92	0.07	-0.90	0.33
Offset F2-F1	0.58	-0.76	-0.46	0.92	-0.12	0.92	-0.05	0.90	-0.35
Offset F3-F2	-0.46	0.83	0.47	-0.87	0.10	-0.87	0.02	-0.86	0.33
Onset-mid F1 slope	0.31	-0.67	-0.44	0.80	0.15	0.77	0.17	0.82	-0.08
Onset-mid F2 slope	0.18	-0.42	0.29	0.27	-0.21	0.28	-0.14	0.25	-0.24
Onset-mid F3 slope	-0.69	0.16	-0.22	-0.51	0.02	-0.52	0.09	-0.50	0.06
Mid-offset F1 slope	-0.24	0.53	0.14	-0.47	0.37	-0.47	0.31	-0.42	0.47
Mid-offset F2 slope	0.19	-0.27	-0.11	0.24	0.14	0.25	0.13	0.24	0.06
Mid-offset F3 slope	0.61	-0.43	0.13	0.56	-0.04	0.58	-0.06	0.55	-0.15
Onset-offset F1 slope	0.09	-0.27	-0.35	-0.45	0.53	0.42	0.51	0.52	0.37
Onset-offset F2 slope	0.41	-0.69	0.32	0.50	-0.21	0.52	-0.14	0.47	-0.29
Onset-offset F3 slope	-0.46	-0.19	-0.17	-0.21	-0.01	-0.21	0.12	-0.21	-0.02
Onset dist. from neutral vowel	0.27	0.22	-0.87	0.27	-0.65	0.25	0.69	0.19	-0.70
Mid dist. neutral vowel	-0.02	-0.13	-0.87	0.23	-0.50	0.21	-0.48	0.17	-0.60
Offset dist. from neutral vowel	0.15	-0.36	-0.88	0.54	-0.35	0.51	-0.33	0.50	-0.52

<sup>a</sup>Note:  $0.64 \leq r \leq 0.76$ ,  $p < 0.05$ ;  $0.76 < r \leq 0.87$ ,  $p < 0.01$ ;  $r > 0.87$ ,  $p < 0.001$ .

The slopes of the F1, F2, and F3 trajectories from the onset to midpoint, midpoint to offset, and onset to offset were also calculated. In addition, given the suggested interpretation of dimension 3 of the English vowel space as reflecting a central-noncentral distinction, the Euclidean distances be-

tween the vowel onsets, midpoints, and offsets in the formant 1 by formant 2 plane (see Fig. 1) and the location of the neutral vowel (F1=500, F2=1500) were also calculated. These correlations are shown in Table AI.

It is also useful to have a quantitative measure of the

TABLE AII. Intercorrelations between all perceptual dimensions extracted.<sup>a</sup>

	English			Spanish (all Ss)		Spanish proficient		Spanish nonproficient	
	D1	D2	D3	D1	D2	D1	D2	D1	D2
English D1	...	-0.18	-0.06	0.75	-0.41	0.78	-0.45	0.68	-0.50
English D2	-0.18	...	0.03	-0.67	-0.15	-0.67	-0.27	-0.68	0.05
English D3	-0.06	0.03	...	-0.30	0.43	-0.28	0.40	-0.25	0.55
Spanish D1	0.75	-0.67	-0.30	...	-0.10	0.99	-0.09	0.99	-0.32
Spanish D2	-0.41	-0.15	0.43	-0.10	...	-0.13	0.97	0.05	0.97
Spanish proficient D1	0.78	-0.67	-0.28	0.99	-0.13	...	-0.11	0.98	-0.34
Spanish proficient D2	-0.45	-0.27	0.40	-0.09	0.97	-0.11	...	0.04	0.93
Spanish nonproficient D1	0.68	-0.68	-0.25	0.99	0.05	0.98	0.04	...	-0.17
Spanish nonproficient D2	-0.50	0.05	0.55	-0.32	0.97	-0.34	0.93	-0.17	...

<sup>a</sup>Note:  $0.64 \leq r \leq 0.76$ ,  $p < 0.05$ ;  $0.76 < r \leq 0.87$ ,  $p < 0.01$ ;  $r > 0.87$ ,  $p < 0.001$ .

similarity between one perceptual dimension and another (e.g., the similarity between dimension 3 in the English vowel space and dimension 2 in the Spanish vowel space). To do this, we have calculated the correlation between all possible pairs of perceptual dimensions described in this paper. These correlations are shown in Table AII. Please note that the signs of the correlations in this table are determined by the orientation of the dimensions that were selected in our figures (e.g., one commonly selects the orientation of the perceptual space so that the “high” vowels appear at the top of the figure and the “low” vowels at the bottom of the figure). In actuality, the signs of these correlations are relatively meaningless since for any given vowel space, any of the perceptual axes could be “flipped” (i.e., the sign of each coordinate value switched) without changing the fit of the solution to the data.

<sup>1</sup>Flege *et al.* (1994) also calculated what they termed “average”  $F_1$  and  $F_2$  frequencies for each vowel. These average formant frequencies were based on the formant values obtained in a separate series of LPC analyses done on each token. Specifically, a 12.8-ms Hamming window was shifted in 5-ms steps through the entire periodic portion of each CV (spurious frequency values were later eliminated using an editing program) and the means of these frequencies were computed. However, these average values fail to capture the dynamic nature of the tokens. In addition, the correlations between these values and the perceptual dimensions (presented in the Appendix) never represent the most significant correlations. Consideration of these average values has therefore been eliminated in the present study. For technical details concerning the acoustic measurements made on the stimulus vowels (e.g.,  $F_0$  duration), the reader is referred to Flege *et al.* (1994).

<sup>2</sup>It is a commonplace in MDS studies to exclude identical vowel pairs from the stimulus set for the dissimilarity rating task. Dissimilarity ratings represent distances between two objects in a perceptual space and it is assumed that the distance between an object and itself is zero. The experimental procedure utilized here is somewhat different than is normally encountered since we are excluding vowel pairs that are not identical but which represent the same vowel category (e.g., English /i/, Spanish /e/, etc.). However, the principle is the same. Namely, we are assuming that the distance between one instance of a vowel category and another instance of the same category is zero. Clearly it would be of great interest to examine intracategory differences perhaps using an approach like that taken by Kewley-Port and Atal (1989), but that will be left to future studies.

<sup>3</sup>As was noted by an anonymous reviewer, another way of comparing the vowel spaces of these two groups of listeners would be to analyze the entire set of data together and then determine whether the two groups placed differential reliance on the obtained dimensions by analyzing the subject weights. The disadvantage of this approach is that if the vowel spaces of the two groups are very different, then the vowel space obtained in such an analysis might be unrepresentative of either group’s underlying perceptual space. For example, note the differences between the overall group vowel space and the vowel spaces from the individual language groups in Terbeek (1977). However, in order to be thorough, this type of analysis was done using data from the 30 English listeners and 26 Spanish listeners (the 13 proficient and 13 nonproficient listeners used in the supplementary MDS analyses). A two-dimension vowel space was obtained. Dimension 1 was interpreted as a high/low dimension and was similar to that obtained for the English listeners only. The second dimension was much less interpretable (it could be considered a weak diphthongal/monophthongal dimension since it separated /e/ and to a lesser extent, /æ/ from the other eight vowels). The subject weights were examined using a two-way ANOVA with the between-subject factor language group and the within-subject factor dimension. There was a significant main effect of dimension [ $F(1,54)=388.2$ ,  $p<0.001$ ]. The subject weights were considerably higher for dimension 1 ( $M=0.762$ ) than dimension 2 ( $M=0.342$ ). There was no significant main effect of language group [ $F(1,54)=0.34$ ], but there was a significant dimension by language group interaction [ $F(1,54)=26.2$ ,  $p<0.01$ ]. *Post-hoc* Scheffé tests showed that the subject weights were significantly higher (at the 0.02 level) for the Spanish listeners than the English listeners on dimension 1 ( $M=0.806$  and  $0.724$ , respectively), but were significantly smaller (at the 0.001 level) for the Spanish listeners on dimension 2 ( $M$

$=0.267$  and  $0.407$ , respectively). These results are completely compatible with (and to large extent, predictable from) the analyses presented in the text.

<sup>4</sup>If the English intervowel distances are computed on the basis of the dimension 1 and 2 coordinates only (more closely matching the two-dimensional space of the two Spanish groups), the proficient Spanish group still shows a higher correlation ( $r=0.58$ ,  $p<0.01$ ) with the English distances than does the nonproficient Spanish group ( $r=0.54$ ,  $p<0.02$ ), although this difference is smaller.

- Bradlow, A. R. (1993). *Language-Specific and Universal Aspects of Vowel Production and Perception: A Cross-Linguistic Study of Vowel Inventories* (DMLL, Cornell University, Ithaca, NY).
- Butcher, A. (1976). *The Influence of the Native Language on the Perception of Vowel Quality* (Arbeitsberichte Institut für Phonetik, Univ. Kiel, 6).
- Carroll, J., and Chang, J. (1970). “Analysis of individual differences in multidimensional scaling via a  $n$ -way generalization of ‘Eckart-Young’ decomposition,” *Psychometrika* **35**, 283–319.
- Delattre, P. (1964). “Comparing the vocalic features of English, German, Spanish, and French,” *Inter. Rev. Appl. Ling.* **7**, 71–97.
- Delattre, P. (1966). “A comparison of syllable length conditioning among languages,” *Inter. Rev. Appl. Ling.* **4**, 183–198.
- Delattre, P. (1969). “An acoustic and articulatory study of vowel reduction in four languages,” *Inter. Rev. Appl. Ling.* **12**, 295–325.
- Flege, J. (1989). “Differences in inventory size affect the location but not the precision of tongue positioning in vowel production,” *Lang. Speech* **32**, 123–147.
- Flege, J. (1991). “The interlingual identification of Spanish and English vowels: Orthographic evidence,” *Q. J. Exp. Psychol.* **43A**, 701–731.
- Flege, J. E., Munro, M. J., and Fox, R. A. (1994). “Auditory and categorical effects for cross-language vowel perception,” *J. Acoust. Soc. Am.* **95**, 3623–3641.
- Fox, R. (1983). “Perceptual structure of monophthongs and diphthongs in English,” *Lang. Speech* **26**, 21–60.
- Fox, R. (1985). “Multidimensional scaling and perceptual features: Evidence of stimulus processing or memory prototypes,” *J. Phon.* **13**, 205–217.
- Fox, R. (1989). “Dynamic information in the identification and discrimination of vowels,” *Phonetica* **46**, 97–116.
- Fox, R., and Trudeau, M. (1988). “A multidimensional scaling study of esophageal vowels,” *Phonetica* **45**, 30–42.
- Gandour, J., and Harshman, R. (1978). “Cross-language differences in tone perception: A multidimensional scaling investigation,” *Lang. Speech* **21**, 1–33.
- Godinez, Jr., M. (1978). “A comparative study of some Romance vowels,” in *UCLA Working Papers in Phonetics* (Department of Linguistics, UCLA, Los Angeles, CA), Vol. 41, pp. 3–19.
- Guirao, M., and de Manrique, B. (1972). “Identification of Argentine Spanish vowels,” in *Proceedings of the Seventh International Congress of Phonetic Sciences*, edited by A. Rigault and R. Charbonneau (The Hague, Mouton).
- Guirao, M., and de Manrique, B. (1975). “Identification of Argentine Spanish vowels,” *J. Psycholinguist. Res.* **4**, 17–25.
- Kempster, G., Kistler, D., and Hillenbrand, J. (1991). “Multidimensional scaling analyses of dysphonia in two speaker groups,” *J. Speech Hear. Res.* **34**, 534–543.
- Kewley-Port, D., and Atal, B. (1989). “Perceptual distances between vowels located in a limited phonetic space,” *J. Acoust. Soc. Am.* **85**, 1726–1740.
- Kreiman, J., Gerratt, B., and Precoda, K. (1990). “Listener experience and perception of voice quality,” *J. Speech Hear. Res.* **33**, 103–115.
- Kuhl, P. (1991). “Human adults and human infants exhibit a ‘perceptual magnet’ effect for the prototypes of speech sounds, monkeys do not,” *Percept. Psychophys.* **50**, 93–107.
- Kuhl, P., Williams, K., Lacerda, F., Stevens, K., and Lindblom, B. (1992). “Linguistic experience alters phonetic perception in infants by 6 months of age,” *Science* **255**, 606–608.
- MacDonald, M. (1989). “The influence of Spanish phonology on the English spoken by United States Hispanics,” in *American Spanish Pronunciation*, edited by P. Bjarkman and R. Hammond (Georgetown U.P., Washington, DC).
- Pisoni, D. (1973). “Auditory and phonetic memory codes in the discrimination of consonants and vowels,” *Percept. Psychophys.* **13**, 253–260.
- Pisoni, D. (1975). “Auditory short-term memory and vowel perception,” *Mem. Cognit.* **3**, 7–18.

- Pols, L., van der Kamp, L., and Plomp, R. (1969). "Perceptual and physical space of vowel sounds," *J. Acoust. Soc. Am.* **46**, 458–467.
- Rakerd, B., and Verbrugge, R. R. (1985). "Linguistic and acoustic correlates of the perceptual structure found in an individual differences scaling study of vowels," *J. Acoust. Soc. Am.* **77**, 296–301.
- Scholes, R. (1967). "Phoneme categorization of synthetic vocalic stimuli by speakers of Japanese, Spanish, Persian, and American English," *Lang. Speech* **10**, 46–68.
- Scholes, R. (1968). "Phonemic interference as a perceptual phenomenon," *Lang. Speech* **11**, 86–103.
- Shepard, R. (1972). "Psychological representation of speech sounds," in *Human Communication: A Unified View*, edited by E. David and P. Denes (McGraw-Hill, New York), pp. 67–113.
- Singh, S., and Woods, G. (1971). "Perceptual structure of 12 American English vowels," *J. Acoust. Soc. Am.* **49**, 1861–1866.
- Skelton, R. (1969). "The pattern of Spanish vowel sounds," *Int. Rev. Appl. Linguist.* **7**, 231–237.
- Stevens, K., Liberman, A., Studdert-Kennedy, M., and Ohman, S. (1969). "Cross Language study of vowel perception," *Lang. Speech* **12**, 1–23.
- Stockwell, D., and Bowen, R. (1965). *The Sounds of English and Spanish* (Univ. of Chicago, Chicago).
- Syrdal, A., and Gopal, H. (1986). "A perceptual model of vowel recognition based on the auditory representation of American English vowels," *J. Acoust. Soc. Am.* **79**, 1086–1100.
- Takane, Y., Young, F., and DeLeeuw, J. (1976). "Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features." *Psychometrika* **42**, 7–67.
- Terbeek, D. (1977). "A Cross-language Multidimensional Scaling Study of Vowel Perception," in *Working Papers in Phonetics* (Dept. of Linguistics, UCLA, Los Angeles, CA), Vol. 37, pp. 1–27.
- Terbeek, D., and Harshman, R. (1971). "Cross-language differences in the perception of natural vowel sounds," in *Working Papers in Phonetics* (Dept. of Linguistics, UCLA, Los Angeles, CA), Vol. 19, pp. 26–28.
- Terbeek, D., and Harshman, R. (1972). "Is vowel perception non-Euclidean?" in *Working Papers in Phonetics* (Dept. of Linguistics, UCLA, Los Angeles, CA), Vol. 22, pp. 13–29.
- Trautmüller, H. (1981). "Perceptual dimension of openness in vowels," *J. Acoust. Soc. Am.* **69**, 198–212.
- Zwicker, E., and Terhardt, E. (1980). "Analytical expressions for critical band–band rate and critical bandwidth as a function of frequency," *J. Acoust. Soc. Am.* **68**, 1523–1525.