Auditory and categorical effects on cross-language vowel perception

James Emil Flege and Murray J. Munro
Department of Biocommunication, University of Alabama at Birmingham, VH503, Birmingham, Alabama 35294

Robert Allen Fox
Department of Speech and Hearing Science, The Ohio State University, Columbus, Ohio 43210

(Received 9 January 1992; revised 19 November 1993; accepted 7 January 1994)

English monolinguals and native Spanish speakers of English rated the dissimilarity of tokens of two Spanish vowel categories, two English vowel categories, or one Spanish and one English vowel category. The dissimilarity ratings of experienced and inexperienced Spanish subjects did not differ significantly. For both the native Spanish and English subjects, perceived dissimilarity increased as the distance between vowels in an $F1-F2$ acoustic space increased. This supported the existence of a universal, sensory-based component in cross-language vowel perception. The native English and Spanish subjects' ratings were comparable for pairs made up of vowels that were distant in an $F1-F2$ space, but not for pairs made up of vowels from categories that were adjacent in an $F1-F2$ space. The inference that the differential classification of a pair of vowels augments perceived dissimilarity was supported by the results of experiment 2, where subjects rated pairs of vowels and participated in an oddity discrimination task. Triads in the oddity task were made up of tokens of vowel categories that were either adjacent (e.g., /a/–/æ/–/a/) or nonadjacent (e.g., /a/–/i/–/i/) in an $F1-F2$ space. The native English subjects' discrimination was better than the native Spanish subjects' for adjacent but not nonadjacent triads. The better the Spanish subjects performed on adjacent triads—and thus the more likely they were to have differentially classified the two phonetically distinct vowels in the triad—the more dissimilar they had earlier judged realizations of those two categories to be when presented in pairs. Results are discussed in terms of their implications for second language acquisition.

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

INTRODUCTION

Although the notion of "phonetic similarity" has played a central role in recent discussions of second-language (L2) acquisition and cross-language speech perception, the factors that may influence phonetic dissimilarity judgments are poorly understood. One factor likely to be important is the extent of the match between the articulatory movement and contact patterns used in forming sounds found in two languages (alternatively, the extent to which their underlying "gestures" are perceived to match; e.g., Best and Stange, 1992). One might suppose that the greater is the articulatory/gestural match between sounds in two languages, the greater will be their degree of perceived phonetic similarity. The present study did not examine the role of articulatory/gestural similarity, however, but instead focused on the role of three other factors: the auditory-acoustic difference between a pair of sounds, their categorical status (i.e., whether the sounds were identified in terms of one or two phonetic categories), and typicality (i.e., the extent to which the sounds matched representations for native language phonetic categories). In experiment 1, native speakers of Spanish and English rated pairs of vowels for dissimilarity. The pairs were made up of tokens of two Spanish vowel categories, two English vowel categories, or one Spanish and one English vowel category. Experiment 1 also assessed the effect of English-language experience on the Spanish subjects' ratings. Experiment 2 was carried out to test a conclusion drawn from the experiment 1 results, viz., that the differential classification of a pair of vowels increases their degree of perceived dissimilarity.

A number of studies have shown the importance of an auditory factor in the perception of foreign vowels. For example, results obtained by Stevens *et al.* (1969) suggested the existence of "natural" auditory-based boundaries between vowels.[1] Native speakers of English and Swedish identified the members of synthetic vowel continua. One ranged from /i/ to /ɛ/. The other, which ranged from /i/ to /ʉ/, contained front rounded vowels that do not occur systematically in English. As expected from differences in how the vowel /i/ is produced in English and Swedish, the Swedish subjects identified fewer vowels as /i/ than did the native English subjects (see also Kuhl *et al.*, 1992). However, the ABX discrimination of both the native English and Swedish subjects was similar, even for the /i/–/ʉ/ continuum.

The role of an auditory factor was tested in the present experiment by comparing native Spanish and English subjects' ratings of vowel pairs. Degree of *auditory* difference was not assessed directly. Instead, we made the simplifying assumption that the degree of auditory difference between certain pairs of vowels would be related to their distance in an $F1-F2$ acoustic space.[2] Tests of the auditory factor were restricted to pairs of vowels that were relatively distant in an $F1-F2$ acoustic space (e.g., /i/–/æ/, /i/–/a/) to avoid the potentially confounding effect of categorical status (see below). Our approach assumed that vowels which are sufficiently distant from one another in an $F1-F2$ space would be identified in terms of two phonetically distinct categories by all subjects, regardless of L1 background. Two findings were deemed necessary for the existence of an auditory factor to receive support: (1) the ratings of perceived dissimilarity for pairs of vowels should vary according to the vowels' distance in an $F1-F2$ space, and (2) much the same pattern of ratings should be obtained for the native English and Spanish subjects. That is, for a factor to be called "auditory," it must be universal in the sense of not showing an effect of previous linguistic experience.

Terbeek (1977) suggested that access to a "psychophysical" mode of perception might reduce or eliminate cross-language differences in vowel perception, and that auditory representations may provide a "universal" framework for vowel perception (pp. 68, 221). However, a great deal of research has shown that adults are *language-specific* perceivers of speech. For example, speakers of different native languages are apt to label vowels in conformity to the number and nature of phonemic categories in their L1 (Stevens *et al.*, 1969; Scholes, 1967a,b, 1968; Butcher, 1976; Terbeek, 1977; Terbeek and Harshman, 1972). Previous linguistic experience may also alter listeners' weighting of acoustic dimensions relevant to perceived vowel quality, such as duration (e.g., Bennett, 1968; Gottfried and Beddor, 1988; Munro, 1992). Given this, we thought it likely that previous linguistic experience would influence cross-language judgments of vowel dissimilarity independently of (or at least, in addition to) the effects of the auditory factor mentioned earlier.

Kuhl and her colleagues (e.g., Kuhl *et al.*, 1992) have shown that vowels in the vicinity of an L1 vowel category "prototype" are discriminated less readily than are pairs of vowels of equivalent auditory difference that are not located near an L1 vowel prototype. This "perceptual magnet" effect has an important possible implication for L2 acquisition. It may mean that an L2 vowel which is phonetically similar, but not identical to an L1 vowel will be judged to be more similar phonetically to the L1 vowel than it would be judged otherwise based solely on its *auditory* difference from the L1 vowel. Best and her colleagues (e.g., Best *et al.*, 1988; Best and Strange, 1992) have shown that the discriminability of consonants in an L2 or foreign language is influenced importantly by how they are categorized. For example, a pair of unfamiliar L2 consonants that are "assimilated" by (i.e., identified in terms of) two distinct L1 consonant categories will be discriminated better than will a pair of unfamiliar L2 consonants that are assimilated by a *single* L1 category. This seems to hold true even if the pair of readily discriminable L2 consonants differ phonetically from the L1 categories which have assimilated them.

It seems reasonable to think that the *discriminability* of vowels will be related to their perceived phonetic *dissimilarity*. If so, then one might suppose that the categorical status of a pair of vowels will also influence their perceived dissimilarity. All else being equal, a pair of vowels that has been identified in terms of two (L1 or L2) categories may be judged to be more dissimilar than will a pair of vowels identified in terms of a *single* category (see also Repp and Crowder, 1990, and references therein). Support for this "differential classification" hypothesis was provided by Kewley-Port and Atal (1989), who examined synthetic vowels which occupied a small portion of the acoustic vowel space. However, all subjects in the Kewley-Port and Atal study were monolingual speakers of English. At present, then, there is no direct evidence showing that the categorical status of *natural* vowels drawn from *two different* languages influences degree of perceived vowel dissimilarity.

The differential classification hypothesis raises an important question related to L2 acquisition by adults: Does the perceived phonetic dissimilarity of pairs of vowels in the L2 (and/or L1) remain constant over the course of L2 acquisition, or does it change as function of experience in the L2? It may be that certain L2 vowels which differ sufficiently from any L1 vowel are treated as falling outside the L1 inventory (Delattre, 1964, 1969; Wode, 1981). Such vowels, if they exist, might evade the "perceptual magnet" effect described by Kuhl (see also Schouten, 1975, p. 14). Flege hypothesized that phonetic category formation for L2 sounds becomes less likely with increasing age, but that even adult learners may establish categories for L2 sounds that differ "sufficiently" from the nearest L1 sound (e.g., Flege, 1981, 1987, 1988). The results of several studies suggested that vowels which are acoustically distinct from the nearest L1 vowel may ultimately be produced more accurately than are L2 vowels which are more similar acoustically, but not identical to an L1 vowel (Major, 1987; Bohn and Flege, 1990, 1992; Flege and Bohn, 1989; Flege, 1991, 1992a,b). Differences in vowel learnability, in turn, might reflect differences in the likelihood of category formation for vowels encountered in an L2.

Cross-language vowel perception research (Butcher, 1976; Terbeek, 1977) has suggested that vowels in a "crowded" perceptual space tend to be judged as more dissimilar than vowels in a relatively uncrowded space (see Jusczyk, 1991). Thus, if bilinguals establish additional phonetic categories for vowels in an L2, and if their psychological vowel space becomes more crowded as a result, the addition of vowel categories should augment perceived vowel dissimilarity (but cf. Kuhl *et al.*, 1992). For example, if a Spanish learner of English were to establish a phonetic category for the English vowel /ɪ/, it might augment the perceived dissimilarity of vowels in the high, front portion of the vowel space.

The results of several recent vowel studies, when taken together, suggest that at least some adult Spanish L2 learners treat English /ɪ/ as distinct from /i/-quality vowels (Flege, 1991; Flege and Bohn, 1989; Blankenship, 1991).[3] On the assumption that learning to recognize phonetic differences between L1 and L2 vowels requires a certain amount of time, we hypothesized that experienced Spanish speakers of English would be more likely than inexperienced Spanish speakers to identify vowels in /i/–/ɪ/ pairs in terms of two distinct vowel categories, and would therefore rate the vowels in such pairs as more dissimilar than would inexperienced Spanish speakers of English. Best *et al.* (1988), on the other hand, suggested that foreign consonants either map onto an L1 category or, if very dissimilar from any L1 consonant, might be treated as nonspeech sounds (i.e., not as "new" consonants). If this were also true for vowels it would mean that adult L2 learners perceive no vowels in an L2 as new. This implies that the categorical status of L2 vowels with respect to vowels in an L1 will not change. If so, then increasing L2 experience would not be expected to lead to increases in the perceived dissimilarity of pairs of L2 vowels (or pairs of L1 and L2 vowels), at least not due to changes in perceived categorical identity.

Identification data were not obtained in the present study because of the difficulty inherent in having phonetically untrained individuals label unfamiliar vowels.[4] The differential classification hypothesis was tested by comparing the ratings given to two kinds of vowel pairs. We assumed that the vowels in pairs designated "nonadjacent" would be identified in terms of two phonetically distinct categories by both the native Spanish and English subjects because of their relatively great distance in an $F1-F2$ acoustic space. On the other hand, we assumed that vowels in "adjacent" pairs would be identified differentially by the native English subjects but not by the native Spanish subjects. So, for example, we assumed that all subjects would identify tokens of English /i/ and /ɑ/ in terms of two distinct categories (the English subjects in terms of English /i/ and /a/, the Spanish subjects in terms of Spanish /i/ and /a/). For pairs containing /i/ and /ɪ/, on the other hand, we assumed identification in terms of two English categories by the native English subjects (viz., /i/, /ɪ/) but identification in terms of a single, Spanish vowel category (viz., /i/) by at least some native Spanish subjects. The prediction being tested was that the English and the Spanish subjects' ratings of vowels in nonadjacent pairs would not differ, whereas the English subjects would rate vowels in *adjacent* pairs as being more dissimilar than would the Spanish subjects. A second experiment was carried out to validate the assumptions made concerning the categorical status of vowels in adjacent and nonadjacent pairs.

The last factor evaluated in the present study was typicality. As used here, the term "typicality" refers to the extent to which phones match listeners' phonetic representations in long-term memory. Terbeek (1977, p. 234) suggested that dissimilarity judgments for vowels are influenced by how closely they conform to L1 categories. Such differences might derive, at least in part, from learned differences in feature weighting (Terbeek, 1977, p. 238) or from the "stretching" of perceptual dimensions arising from differences in allocation of attention (e.g., Jusczyk, 1991). Also, a vowel token that closely matches a "best exemplar" or "prototype" vowel in the L1 inventory might be processed more thoroughly, or stored and retrieved more easily from memory, than a token that does not (see Hebb, 1949). Our *a priori* expectation was that L1 vowels would match listeners' phonetic representations better than would vowels drawn from a second or foreign language. Given the assumptions of a multiple-trace model (e.g., Hintzman, 1986), one might hypothesize that pairs of typical (i.e., familiar) L1 vowels will be rated as more dissimilar than will pairs of L2 (or foreign) vowels.

Effects comparable to the one being proposed here for vowel perception have been observed in other domains, such as voice recognition (Thompson, 1987; Goggin *et al.*, 1991) and facial recognition (Chance *et al.*, 1982; Brigham *et al.*, 1982; Shapiro and Penrod, 1986; Bothwell *et al.*, 1989).[5] To test the typicality hypothesis, we compared Spanish and English subjects' ratings of pairs made up of two phonetically distinct Spanish vowels to their ratings of pairs made up of two phonetically distinct English vowels. For the typicality hypothesis to receive support, the Spanish subjects would need to rate Spanish–Spanish pairs as more dissimilar than the English subjects, and the native English subjects would need to rate English–English vowel pairs as more dissimilar than the native Spanish subjects.

I. EXPERIMENT 1

The aims of this experiment were to (1) determine if the perceived dissimilarity of English and Spanish vowels changes as native speakers of Spanish gain experience in English, and (2) test the hypothesized role of three factors (viz., auditory difference, categorical status, typicality) on vowel dissimilarity ratings.

A. Method

1. Stimuli

The Spanish consonant–vowel (CV) syllables examined here contained the vowels /i/, /e/, and /a/; the English CVs examined contained, /i/, /ɪ/, /e$^{\text{I}}$/, /ɛ/, /æ/, /ʌ/, and /ɑ/. Stimulus variability has been found to promote generalization in many cognitive domains. Therefore, multiple tokens of each vowel category were used as stimuli to encourage listeners to process vowels in a general rather token-specific mode (see, e.g., Cohen and Strauss, 1979; Fox, 1985; Uchanski *et al.*, 1991; Van Tassell *et al.*, 1992). Five Spanish monolinguals, all newly arrived in Austin, Texas, read a single randomized list that included five tokens each of the Spanish words *pito, peto, pato* at the end of the carrier phrase *Digo ahora* ___. Only the first CV in the Spanish words was examined in the present study.

A difficulty we faced in eliciting the production of English vowels in the same phonetic context as the Spanish vowels was that while most Spanish words are formed from open CV or CCV syllables, English does not permit lax vowels to occur in open syllables. However, we managed to elicit the production of lax English vowels in open syllables by devising a list of nonwords. Five monolingual English speakers living in Birmingham, Alabama read a randomized list of disyllables at the end of the carrier phrase *Now I say* ___. The English disyllables (like the Spanish words) had stress on the first syllable. The native

English talkers read seven short lists, each with just one vowel in the first syllable. For example, the /ɪ/ list included the nonce forms *bick+toe*, *bin+toe* and *bit+toe* (all of the form /bVC+to/) as well as *bi+toe* (/bV+to/), in which /ɪ/ occurred in an open syllable.[6] The native English talkers produced the lax vowels in the open syllables accurately and without hesitation, apparently by analogy to lax vowels in closed syllables on the same list.

The middle three of the five available Spanish and English /CV+to/ tokens were low-pass filtered (4.8 kHz) and digitized at 10.0 kHz. The initial CVs were edited out of their original disyllables at the point of complete constriction of the intervocalic /t/, then normalized for peak intensity. A preliminary experiment was carried out to reduce the corpus of 45 Spanish CVs (5 talkers×3 vowels×3 tokens) and 175 English CVs (5 talkers×7 vowels×3 tokens). The three English monolinguals and three Spanish/English bilinguals who participated in the pilot experiment identified, then rated for goodness, CVs from their native language. There is a wide separation between the five vowels of Spanish in an $F1-F2$ acoustic space (Stockwell and Bowen, 1965; Delattre, 1964, 1966, 1969; Skelton, 1969; Guirao and de Manrique, 1972, 1975; Godinez, 1978). Not surprisingly, then, vowels in the Spanish CVs were never misidentified. However, the English subjects did misidentify a few vowels produced by two native English talkers, whose CVs were eliminated. To reduce the Spanish corpus proportionally, CVs spoken by the two native Spanish talkers whose CVs had received the lowest average goodness ratings were eliminated. The three Spanish monolinguals whose CVs were retained were from northern Mexico. Two of the three retained native English talkers were originally from Ohio, and one was from California. The goodness ratings were used to select one of the three available tokens of each vowel category produced by each talker.

Acoustic measurements were made of the nine Spanish CVs and the 21 English CVs that were retained. Voice onset time (VOT) was measured from the release burst of the initial consonant to the first upward-going zero crossing of the periodic portion of the CVs. Fundamental frequency ($F0$) was the reciprocal of the average duration (in ms) of three glottal pulses at the vowel midpoint. Formant frequencies were estimated using LPC analysis at three locations in each vowel. A 21.2-ms Hamming window was centered 20 ms from the onset of periodicity (designated the "vowel onset"), at the acoustic midpoint, and 20 ms from the end of the periodic portion (the "vowel offset"). Fourteen LPC coefficients were calculated at each location. The frequencies of $F1$, $F2$, and $F3$ were estimated by applying a peak picking algorithm to smoothed LPC spectra.

The acoustic measurements are presented in Table I. The Spanish /p/'s and English /b/'s had similar average VOT values (12 vs 9 ms). The formant frequency values were much as expected from previous research (Peterson and Barney, 1952; Skelton, 1969; Guirao and de Manrique, 1972, 1975; Godinez, 1978). The English /i/'s were longer than the English /ɪ/'s (130 vs 81 ms); the English /e$^{\text{I}}$/'s were longer than the /ɛ/'s (165 vs 115 ms); and both the /æ/'s and /ɑ/'s were longer than the /ʌ/'s (165 and 168 ms vs 97 ms). As expected (Delattre, 1964, 1966), English vowels were longer than the corresponding Spanish vowels (/i/-78, /e/-90, /a/-109 ms). The overall difference in the duration of English and Spanish vowels might be attributed to language-specific differences in the effect of stress on vowel duration (Delattre, 1966).

Kewley-Port and Atal (1989) noted a relation between judgments of dissimilarity and the distance in an $F1-F2$ acoustic space between pairs of steady-state synthetic vowels. The relationship was significantly stronger for $F1-F2$ values that had been converted to a bark scale, which approximates a frequency-to-place transformation along the basilar membrane, than for $F1-F2$ values expressed in Hz. The bark distance metric advocated by Kewley-Port and Atal (1989) was used here. One difficulty in applying it, however, was that $F1$ and $F2$ values in our natural vowels changed over time, especially those in the English /e$^{\text{I}}$/'s tokens. This raised the issue of which spectral "slice" to consider. We might have taken $F1$ and $F2$ measures from the acoustic midpoint, which is sometimes called the vowel "target." However, doing so would have belied the presence of formant movement in our natural stimuli. We considered using the endpoint of the deflections of the $F1$ and $F2$ trajectories but, when present, they often occurred at different locations in the vowel.

These considerations led us to use *average* $F1$ and $F2$ frequencies. A series of LPC analyses were carried out for each CV by moving a 12.8-ms Hamming window in 5-ms steps through the periodic portion of each CV. An editing program was used to remove obviously spurious frequency

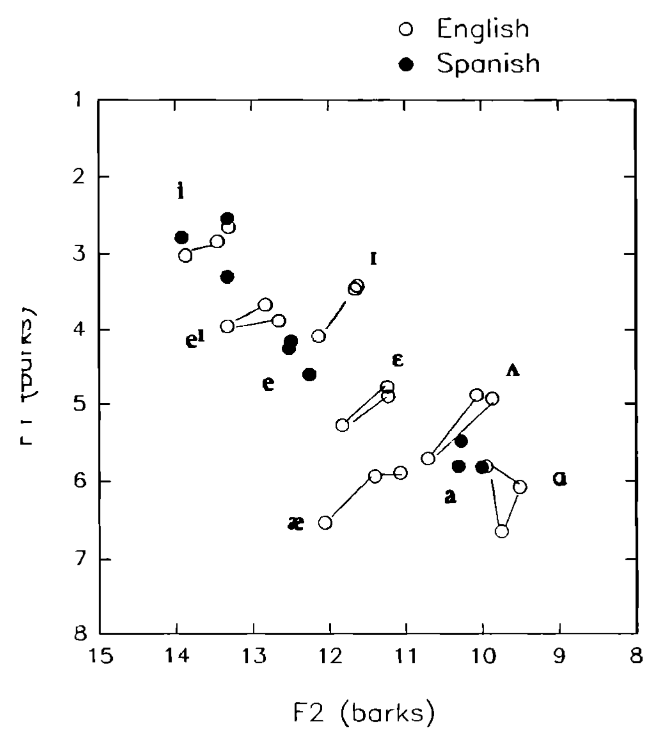

FIG. 1. The average frequency of the first two vowel formants ($F1$, $F2$) of three tokens each of Spanish vowels (/i e a/) and English vowels (/i ɪ e$^{\text{I}}$ ɛ æ ʌ ɑ/), in barks. See text for how average frequencies were computed, and Table I for frequencies at three temporal locations.

TABLE I. Mean acoustic values for Spanish and English vowels used in experiment 1. Each mean is based on tokens produced by three monolingual talkers (Ss).

					Onset[a]			Midpoint[b]			Offset[c]		
	Ss	VOT[d]	$F0$[e]	DUR[f]	$F1$	$F2$	$F3$	$F1$	$F2$	$F3$	$F1$	$F2$	$F3$
						English vowels							
/i/	1	10	119	131	312	2205	2726	317	2299	2897	320	2257	2829
	2	13	93	130	304	1995	2383	281	2079	2516	256	2064	2462
	3	7	109	129	318	2024	2478	300	2180	2628	293	2185	2673
/ɪ/	1	11	123	73	415	1806	2430	476	1616	2571	469	1718	2498
	2	5	94	85	357	1615	2322	373	1618	2393	369	1607	2384
	3	8	120	86	366	1615	2437	374	1561	2472	361	1599	2489
/e$^{\text{ɪ}}$/	1	12	119	140	457	1926	2488	443	2117	2531	358	2072	2489
	2	4	88	161	468	1666	2318	438	1903	2474	274	1868	2445
	3	7	114	195	428	1746	2549	411	2020	2508	318	1935	2500
/ɛ/	1	10	110	122	520	1673	2534	629	1653	2628	580	1648	2578
	2	5	89	104	513	1482	2488	600	1495	2607	496	1507	2535
	3	2	114	118	452	1466	2395	552	1575	2533	478	1509	2460
/æ/	1	7	116	165	653	1668	2604	777	1788	2562	685	1702	2583
	2	7	84	170	555	1538	2362	690	1563	2423	615	1546	2447
	3	12	107	160	548	1474	2294	732	1468	2493	654	1472	2399
/ɑ/	1	10	115	155	699	1050	2190	769	1129	2161	743	1208	2186
	2	10	88	178	581	1052	2217	694	1108	2493	652	1161	2473
	3	4	110	172	551	1136	1857	681	1190	1968	619	1209	1955
/ʌ/	1	11	116	100	623	1269	2355	703	1425	2191	554	1394	2365
	2	19	87	95	519	1048	2364	599	1200	2367	485	1236	2356
	3	12	119	96	503	1191	2200	562	1249	2348	519	1268	2316
						Spanish vowels							
/i/	1	16	191	78	350	2140	2571	354	2238	2576	355	2144	2582
	2	12	109	80	264	2074	2705	293	2126	2933	229	2072	2761
	3	10	112	77	286	2222	2719	292	2295	2773	286	2273	2745
/e/	1	9	168	99	468	1629	2449	513	1864	2634	485	1758	2573
	2	15	104	86	492	1838	2332	436	1900	2587	393	1822	2475
	3	11	110	85	458	1761	2508	457	1903	2751	388	1831	2683
/a/	1	17	166	110	517	1195	2591	646	1331	2610	567	1309	2612
	2	6	103	116	652	1148	1933	757	1313	2135	584	1257	1992
	3	13	111	100	593	1125	1876	705	1308	1982	599	1316	2043

[a–c]Vowel frequencies measured 20 ms into the vowel interval, at the acoustic midpoint of the vowel, and 20 ms from the end of the vowel, in Hz.
[d]VOT, Voice onset time (in ms)
[e]$F0$, fundamental frequency at the vowel midpoint (Hz).
[f]DUR, Vowel duration (ms).

values (e.g., the values reported occasionally between the $F1$ and $F2$ tracks). The average $F1$ and $F2$ frequency for each of the CVs was then calculated based on values in the initial consonant (/p/ or /b/) transitions, the "vowel," and the transitions leading into the /t/ of the forthcoming CV (which had been edited out of the stimuli).

The average values in Hz were then converted to a bark scale (Zwicker and Terhardt, 1980), and are shown in Fig. 1. Our approach implicitly assumed that information in consonant transitions into and out of a "vowel" contributes to the vowel's perceived quality. This assumption is consistent with results obtained by Fowler (1984), who found that listeners tended to integrate information in the initial release burst and the steady state portion of a following "vowel" portion of CVs when judging the similarity of pairs of CVs. Also, the "silent center" research of Strange and her colleagues (e.g., Strange, 1987) has shown that vowels in CVCs can be identified fairly well on the basis of consonant transitions alone. The Euclidean bark distances between vowels were not expected to account fully for the listeners' dissimilarity ratings. This is because the natural vowel stimuli examined here differed along many dimensions in addition to average $F1$ and $F2$ frequencies (see Table I and also Fox, 1989). Nevertheless, based on the findings of Kewley-Port and Atal (1989), we expected to observe a systematic relationship between degree of perceived dissimilarity and $F1-F2$ bark distances.

2. Subjects

A total of 60 subjects, all of whom passed a pure-tone hearing screening, were recruited as paid listeners in Birmingham, Alabama. Ten males and 20 female subjects were monolingual speakers of American English. The re-

TABLE II. Characteristics of the four groups of subjects ($N=15$/group) in experiment 1. Standard deviations are in parentheses.

	Subject group			
	SA[a]	SB[b]	EnA[c]	EnB[c]
Chronological age, in years	33.1(7.9)	33.6(9.2)	30.2(7.3)	29.1(5.0)
Age of L2 learning, in years	27.0(8.2)	24.9(7.6)	···	···
Length of residence in U.S., in years	1.8(1.2)	7.0(4.8)	···	···
Self-estimated % daily use of English	49.6(28.7)	60.8(14.3)	···	···
L2 input (LOR×% use)	0.9(0.6)	4.1(2.7)	···	···
Experimenter ratings[d]	4.3(0.8)	4.4(0.9)	···	
Self-ratings[e]	4.3(0.8)	4.9(1.0)	···	···
L2 proficiency[f]	8.6(1.4)	9.3(1.8)	···	···

[a]SA, relatively inexperienced Spanish speakers of English.
[b]SB, relatively experienced Spanish speakers of English.
[c]EnA and EnB, randomly selected subgroups of native English speakers.
[d]Experimenter ratings of ability to pronounce/comprehend English.
[e]Self-ratings by subjects of ability to pronounce/comprehend English.
[f]L2 proficiency, the sum of the experimenter and self-ratings.

maining subjects (7 males, 23 females) were native speakers of Spanish who had learned English as an L2. The native English subjects were randomly assigned to two subgroups of 15 subjects each (EnA, EnB) according to their order of enrollment in the study. As summarized in Table II, the native Spanish subjects were assigned to subgroups according to English-language experience. The relatively inexperienced Spanish subjects (SA) and the relatively experienced Spanish subjects (SB) were of approximately the same chronological age, had studied English in school for about the same number of years, and had arrived in the U.S. at about the same age. The SB subjects had lived longer in the U.S. than the SA subjects (7.0 vs 1.8 years) and reported using English more on a daily basis (61% vs 50%). The SB subjects were estimated to have had roughly four times more English-language input than the SA subjects (4.1 vs 0.9 years).[7]

3. Procedure

The 60 subjects (listeners) each rated 405 vowel pairs in a single session. The 30 CV stimuli were paired once with every other CV except the other two CVs containing a realization of the same vowel category. There were nine exemplars of each of 45 vowel pair types. Each Spanish vowel was paired with every other Spanish vowel, resulting in three Spanish–Spanish pair types (viz., /i/–/e/, /i/–/a/, /e/–/a/). The three tokens of Spanish /i/ were each paired with all 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. 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 also occurred an equal number of times in the first and second positions.

Subjects were tested individually in a sound booth, where stimuli were presented binaurally at a comfortable level after being low-pass filtered (4.8 kHz). The subjects used a scale ranging from "very similar" (1) to "very dissimilar" (9) to rate each pair of vowels. They were instructed to use the whole scale, and to guess if uncertain. No training was administered, but the subjects were given practice before the experiment on pairs thought likely to span the dissimilarity scale. The delay between the subjects' responses and the next trial was 1.0 s. A relatively long ISI of 1.2 was used between the CVs in each pair to encourage the use of phonetic codes stored in long-term memory (Pisoni, 1973; Fox, 1983, 1985, 1989; Shigeno, 1991). For each subject, the mean of nine dissimilarity ratings was calculated for each of three Spanish–Spanish, 21 English–English, and 21 Spanish–English vowel pair types. A series of ANOVAs indicated that the four groups of subjects used the dissimilarity scale in much the same manner.[8]

B. Results

1. L2 experience

The mean ratings obtained for the 45 vowel pair types are presented in Table III along with the results of one-way ANOVAs which tested the effect of group on the ratings for each vowel pair type. When the group effect was significant, pairwise differences between the two native Spanish groups differing in English-language experience (SA, SB) and the two randomly selected native English groups (EnA, EnB) were examined using Tukey's HSD test. The effect of group on the dissimilarity ratings was significant for 15 of the 45 vowel pair types. The ratings were sometimes higher (indicating greater perceived dissimilarity) for the native English than for the Spanish subjects, but they were also sometimes lower. The bases for these between-group differences will be discussed below.

As expected, groups EnA and EnB never differed significantly. Somewhat surprisingly, neither did the experienced and inexperienced Spanish subjects. In only 20 of 45

TABLE III. Mean ratings of vowel dissimilarity obtained in experiment 1 for the subjects in four groups. "HSD" indicates the results of a Tukey's HSD test (alpha=0.05) comparing the four groups when the group factor reached significance.

	(1) SA[a]		(2) SB[b]		(3) EnA[c]		(4) EnB[c]			
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	$F(3.59)$	HSD test
					Spanish–Spanish pairs					
/i/–/e/[d]	5.4	1.9	5.1	2.3	4.8	1.2	4.0	1.1	1.8	
/i/–/a/[d]	7.6	1.3	7.6	1.2	6.8	1.1	6.0	1.2	5.8*	1,2 > 4
/e/–/a/[d]	6.0	1.1	6.6	1.3	6.0	1.4	5.5	1.2	1.9	
					English–English pairs					
/i/–/ɪ/	3.6	1.2	4.0	1.5	6.0	1.1	5.6	1.3	12.2**	3,4 > 1,2
/i/–/e^{I}/[d]	7.2	1.2	7.0	0.6	7.1	0.9	6.2	1.1	2.9	
/i/–/ɛ/[d]	7.3	1.3	6.9	1.2	6.0	1.2	5.5	1.3	5.8*	1,2 > 4
/i/–/æ/[d]	7.8	1.7	7.7	1.0	7.5	0.8	6.9	1.3	1.6	
/i/–/ɑ/[d]	8.0	1.3	8.3	0.6	7.9	0.6	7.0	1.2	5.0*	1,2 > 4
/i/–/ʌ/[d]	7.6	1.2	7.8	0.9	7.4	0.9	6.9	1.2	1.7	
/I/–/e^{I}/	6.0	1.5	5.2	1.2	7.1	1.1	6.9	1.0	7.5*	3,4 > 2
/I/–/ɛ/	5.7	1.3	4.4	1.4	5.0	1.3	4.3	1.3	3.3	
/I/–/æ/[d]	7.3	1.1	6.8	1.2	7.2	1.3	7.0	1.2	0.4	
/I/–/ɑ/[d]	7.5	1.2	7.0	1.1	7.1	1.1	6.7	1.0	1.1	
/I/–/ʌ/[d]	6.5	1.2	6.8	1.4	6.6	1.3	6.3	1.1	0.4	
/e^{I}/–/ɛ/	6.0	1.3	6.0	1.3	6.1	1.3	5.9	1.4	0.1	
/e^{I}/–/æ/[d]	6.4	1.4	6.8	1.3	5.8	1.5	5.3	1.7	3.0	
/e^{I}/–/ɑ/[d]	6.7	1.5	6.9	1.3	6.8	1.5	6.2	1.6	0.7	
/e^{I}/–/ʌ/[d]	6.6	1.1	6.5	1.2	7.0	1.4	6.5	1.1	0.6	
/ɛ/–/æ/[d]	3.0	1.3	3.0	1.1	4.1	1.1	4.3	1.3	5.5*	4 > 2,1
/ɛ/–/ɑ/[d]	4.2	1.8	4.4	1.5	6.3	1.3	5.8	1.4	7.2**	3,4 > 1 3 > 2
/ɛ/–/ʌ/	3.0	1.7	2.9	1.4	5.1	1.4	4.8	1.4	8.5**	3,4 > 1,2
/æ/–/ɑ/	2.5	1.4	2.4	1.6	5.1	1.8	4.4	1.3	12.1**	3,4 > 1,2
/æ/–/ʌ/	3.2	1.6	2.7	1.7	6.5	1.7	6.6	1.1	27.5**	3,4 > 1,2
/ɑ/–/ʌ/	2.5	0.9	2.3	1.4	4.4	1.1	4.3	1.2	13.8**	3,4 > 1,2
					Spanish–English pairs					
/i/–/i/	2.6	1.5	1.9	0.8	1.8	0.5	2.0	0.9	2.0	
/e/–/i/[d]	5.7	1.6	5.9	1.6	5.8	1.1	5.4	1.3	0.4	
/a/–/i/[d]	7.8	1.4	7.7	1.1	7.8	0.7	6.8	1.2	2.9	
/i/–/ɪ/	3.6	1.0	3.3	1.2	4.6	1.3	4.2	1.2	3.9	
/e/–/I/[d]	3.1	1.2	2.7	1.1	2.0	0.9	1.9	1.0	3.9	
/a/–/ɪ/[d]	6.9	1.6	6.7	1.4	6.9	1.1	6.2	1.1	0.9	
/i/–/e^{I}/	7.0	1.5	6.5	1.0	7.2	0.9	6.7	0.9	1.2	
/e/–/e^{I}/	5.3	1.2	4.5	1.0	5.8	1.2	5.9	0.9	5.0*	3,4 > 2
/a/–/e^{I}/[d]	6.8	1.4	7.1	1.3	7.2	1.4	6.9	1.2	0.3	
/i/–/ɛ/[d]	7.0	1.3	6.8	1.1	6.2	0.9	5.5	1.0	5.5*	1,2 > 4
/e/–/ɛ/	4.9	0.9	4.7	1.6	4.2	1.4	3.7	0.9	2.9	
/a/–/ɛ/[d]	2.9	1.2	3.3	1.7	5.8	1.6	4.8	1.3	12.6**	3,4 > 1,2
/i/–/æ/[d]	7.6	1.5	7.8	1.3	7.8	0.4	7.2	1.1	0.9	
/e/–/æ/	6.7	1.3	6.8	1.1	6.8	0.9	6.2	0.7	1.2	
/a/–/æ/	3.1	1.3	2.9	2.0	6.5	1.3	6.1	0.8	28.2**	3,4 > 1,2
/i/–/ɑ/[d]	8.0	1.3	8.1	0.9	7.9	0.6	7.2	1.0	2.4	
/e/–/ɑ/[d]	6.8	1.1	6.7	1.2	6.8	1.3	6.4	1.2	0.4	
/a/–/ɑ/	2.1	0.9	2.0	0.8	4.0	1.0	3.6	1.1	17.6**	3,4 > 1,2
/i/–/ʌ/[d]	7.6	1.5	7.6	1.3	7.2	0.8	6.6	1.2	2.1	
/e/–/ʌ/	5.8	1.5	6.2	1.4	6.4	1.3	5.8	1.2	0.7	
/a/–/ʌ/	1.9	0.9	1.7	0.6	1.7	0.6	1.7	0.9	0.4	

[a]SA, relatively inexperienced Spanish speakers of English.
[b]SB, relatively experienced Spanish speakers of English.
[c]EnA and EnB, randomly selected subgroups of native English speakers.
[d]Examined in correlation analyses.

instances were the mean ratings obtained for group SB more like those of the native English subjects than were the means obtained for group SA. Also, there was not a systematic difference between groups SA and SB for the English–English pairs considered separately. These results suggest that amount of L2 experience may, in itself, have little effect on the perceived dissimilarity of pairs of English and/or Spanish vowels.

TABLE IV. The Spanish–Spanish, English–English, and Spanish–English pairs used in experiment 1 to test the hypothesis that $F1$–$F2$ distances influence judgements of perceived vowel dissimilarity. Frequency values for the first two vowel formants ($F1$, $F2$) are expressed in barks. "EBD" is the Euclidean bark distance between the two vowels in each pair (see text).

	Vowel 1		Vowel 2		
Vowel 1–Vowel 2	$F1$	$F2$	$F1$	$F2$	EBD
Spanish /i/–Spanish /a/	2.88	13.6	5.70	10.21	4.41
Spanish /i/–Spanish /e/	2.88	13.6	4.34	12.43	1.87
English /i/–English /ɑ/	2.84	13.55	6.18	9.74	5.07
English /i/–English /ɛ/	2.84	13.55	4.99	11.43	3.02
English /i/–Spanish /a/	2.84	13.55	5.70	10.21	4.40
Spanish /i/–English /ɛ/	2.88	13.6	4.99	11.43	3.03

2. Auditory difference

The mean ratings obtained for the six vowel types listed in Table IV were examined to test the hypothesis that degree of auditory difference influences perceived vowel dissimilarity. Three criteria governed the selection of these six pairs. First, the pairs had to differ substantially in $F1$-$F2$ distances because the metric used here to estimate auditory differences did not take into account many other auditorily relevant properties (e.g., duration, $F0$, $F1$, and $F2$ movement, $F3$ frequency). Second, the pairs had to be made up of vowels that were likely to be identified differently by *all* subjects. This is because, as discussed later, differential classification was found to increase degree of perceived similarity. Finally, we tried to hold vowel typicality (familiarity) as constant as possible by including an equal number of pairs containing two Spanish vowels (/i/–/a/, /i/–/e/), two English vowels (/i/–/ɑ/, /i/–/ɛ/), and one Spanish and one English vowel (English /i/–Spanish /a/, Spanish /i/–English /ɛ/). Most importantly, of course, the vowel pairs selected had to differ according to *acoustic distances.* Of the six pairs selected, three pairs were made up of a high vowel and a low vowel (Spanish /i/–Spanish /a/, English /i/–English /ɑ/, English /i/–Spanish /a/) and three were made up of a high vowel and a midvowel (Spanish /i/–Spanish /e/, English /i/–English /ɛ/, and Spanish /i/–English /ɛ/). The distances in an $F1-F2$ space between the high and low vowels were, of course, greater than the distances between the high and midvowels.

The mean ratings were submitted to a (3) group$\times$(3) pair composition type$\times$(2) acoustic distance ANOVA, with repeated measures on the last two factors. For the auditory factor to receive support, it would be necessary to show that the subjects in *all four* groups judged the high-low pairs to be more dissimilar than the high-mid pairs. As expected, the difference between high–low and high–mid vowels (7.4 vs 5.9) was significant [$F(1,56)=150.3$, $p<0.01$]. As shown in Fig. 2, much the same effect of acoustic distance was evident for the subjects in all four groups, so the group$\times$distance interaction was nonsignificant [$F(3,56)=0.83$, $p>0.10$]. The lack of a significant three-way interaction [$F(6,112)=1.52$, $p>0.10$] indicated that much the same effect of $F1-F2$ distances was evident for all four groups, irrespective of possible differences in typicality (see below).

Pearson product-moment correlation analyses were carried out to quantify the influence of auditory differences (as indexed by distances in an $F1-F2$ bark space) on the dissimilarity ratings. The correlations examined the mean dissimilarity ratings obtained for the 27 pair types whose members were deemed likely, on the basis of $F1-F2$ distances, to be identified in terms of two distinct vowel categories by all 60 subjects. (The 27 pairs selected for this analysis are indicated by a superscript d symbol in Table III.) The coefficients of correlation, which averaged 0.647, were significant at the 0.05 level for all but one of the 60 subjects tested (a native Spanish speaker). Much less variance was accounted for here by the bark distance metric (viz., 42%) than in the Kewley-Port and Atal (1989) study. This was probably because our natural stimuli differed along dimensions other than $F1$ and $F2$ (see above).

Of the vowels examined here, the diphthongal English /e[I]/ tokens were probably the ones misrepresented to the greatest extent by estimating audtitory difference through an average of $F1$ and $F2$ frequencies over the entire "vowel" (see above). The role of formant movement was

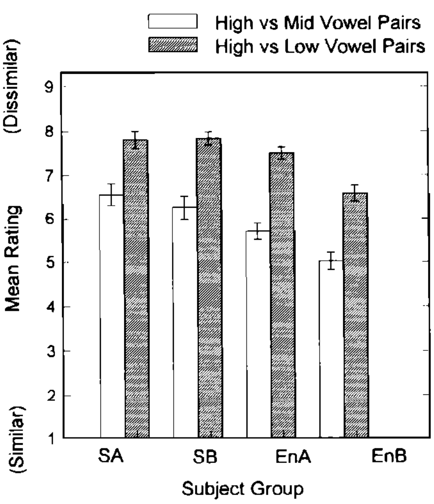

FIG. 2. Mean dissimilarity ratings of vowel pairs by relatively inexperienced and experienced Spanish speakers of English (SA,SB) and two randomly selected groups of native English speakers (EnA,EnB). The vowels in pairs made up of a phonologically high and mid vowel were more distant from one another in an $F1-F2$ acoustic space than were pairs made up of phonologically high and low vowels; see text. The brackets enclose ± 1 standard error.

assessed indirectly by eliminating the five (of 27) pairs containing /eᴵ/ tokens, then recomputing the correlations. If the listeners made use of formant movement information, one would expect the strength of correlations to increase despite the reduction in degrees of freedom. In fact, the correlation coefficients increased from the earlier average of $r=0.647$ to $r=0.732$, accounting for 54% of variance in the mean dissimilarity ratings. The two sets of correlation coefficients were examined in a (4) group×(2) diphthongal status ANOVA. The coefficients computed for pairs including /eᴵ/ were significantly smaller than those computed for the 22 pairs remaining after the /eᴵ/ pairs were excluded [$F(1,56)=150.8$, $p<0.01$]. The lack of a significant group main effect [$F(3,56)=1.05$, $p>0.10$], or a significant group×status interaction [$F(3,56)=0.32$, $p>0.10$], suggested that the subjects in all four groups used formant movement information in the /eᴵ/ tokens.

3. Typicality

According to the typicality hypothesis, a pair of familiar L1 vowels should be rated as being more dissimilar than a pair of vowels of the same degree of auditory dissimilarity if the vowels in the second pair represent poor matches to L1 vowels. The typicality hypothesis was tested here by examining listeners' ratings of three Spanish–Spanish pairs (/i/–/e/, /e/–/a/, /i/–/a/), three English–English pairs (/i/–/eᴵ/, /eᴵ/–/ɑ/, /i/–/ɑ/), and three Spanish–English pairs (/e/–/i/, /a/–/eᴵ/, /i/–/ɑ/). Each subset, referred to as a "language composition type," contained a phonologically mid–high vowel pair, a mid–low pair, and a high–low pair. The Spanish vowel system has three degrees of phonological vowel height, so we assumed that even the Spanish subjects would identify the vowels in each of the pairs in terms of two phonetically distinct vowel categories (albeit not the same ones as the native English subjects). The phonological distances, of course, corresponded to acoustic distances (see Fig. 1).

Our primary aim was to select vowel pairs which contained vowels likely to differ in terms of how well they matched the subjects' long-term memory representations. We assumed that vowels from an L2 or foreign language would represent poorer matches to listeners' long-term memory representations for vowels than would realizations of L1 categories. If so, then Spanish vowels should match the Spanish subjects' vowel representations better than the English subjects' vowel representations, and vice versa.[9] For example, we expected many Spanish subjects to identify the English /eᴵ/ tokens as "poor" exemplars of Spanish /e/ because of the presence of diphthongization (see Scholes, 1967a; Blankenship, 1991; see also Best *et al.*, 1988) and the English subjects to judge the Spanish /e/'s as poor exemplars of the English /eᴵ/ category because of the *absence* of diphthongization. So, too, we expected the English /ɑ/'s to match the native English subjects' /ɑ/ category better than the Spanish subjects' /a/ category, and vice versa (see Table I, and Flege, 1989, 1991).

The mean ratings obtained for the three language composition types are shown in Fig. 3 (top). The means shown

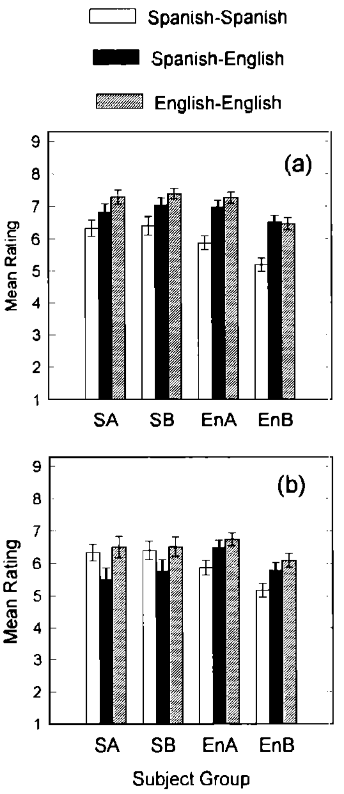

FIG. 3. Mean dissimilarity ratings of pairs made up of two Spanish vowels, two English vowels, or one English and one Spanish vowel by relatively inexperienced and experienced Spanish speakers of English (SA,SB) and two native English groups (EnA,EnB). The ratings in (a) included pairs containing the English vowel /eᴵ/ whereas the ratings in (b) did not; see text. The error bars enclose ±1 standard error.

here have been averaged over the pairs differing in $F1-F2$ distances (i.e., the mid–high, mid–low, and high–low pairs) because, as expected from the analyses in Sec. I B 2, a similar pattern of between-pair differences was observed for all four listener groups. Perceived dissimilarity was greater for English–English than Spanish–English than Spanish–Spanish pairs for the subjects in groups SA, SB, and EnA. However, for group EnB, much the same ratings were obtained for English–English and Spanish–English pairs.

The ratings were submitted to a (4) group×(3) language composition type×(3) phonological distance ANOVA, which yielded a significant group×language composition interaction [$F(6,112)=2.96$, $p<0.05$]. The simple effect of group was significant for the Spanish–Spanish pairs [$F(3,56)=4.33$, $p<0.05$] and for the English–English pairs [$F(3,56)=3.27$, $p<0.05$], but not for the Spanish–English pairs [$F(3,56)=0.88$, $p>0.10$]. The lack of a significant difference between groups for the Spanish–English pairs is consistent with the typicality hypothesis. An HSD test revealed that the subjects in groups SA and SB judged the vowels in Spanish–Spanish pairs to be more dissimilar than did the subjects in group EnB ($p<0.05$). This, too, is consistent with the typicality hypothesis.

However, contrary to hypothesis, an HSD test revealed that the subjects in group SB judged English–

English pairs to be *more* rather than less dissimilar than the subjects in group EnB ($p<0.05$). One might attribute this to the inclusion of English /e^{I}/'s. Spanish allows /e/ +/i/ sequences but does not have a diphthongal /e^{I}/ vowel phoneme (Stockwell and Bowen, 1965). Perhaps the SB subjects were sufficiently familiar with the English sound system to recognize that the English /e^{I}/ tokens were not two-vowel sequences (i.e., /i/+/e/) but had not established a new category for English /e^{I}/. If so, then the offglide in the English /e^{I}/'s might have sounded anomalous to them, heightening perceived dissimilarity.

A second analysis was therefore carried out in which the pairs with /e^{I}/ (viz., /i/–/e^{I}/, /e^{I}/–/ɑ/, /a/–/e^{I}/) were replaced with /i/–/ɛ/, /ɛ/–/ɑ/, and /a/–/ɛ/. As shown in Fig. 3 (bottom), perceived dissimilarity was greater for English–English than Spanish–English than Spanish–Spanish pairs for the two native English groups (EnA, EnB). For the subjects in SA and SB, on the other hand, the English–English and Spanish–Spanish pairs were more dissimilar than the Spanish–English pairs. Most importantly, the English–English pairs were not rated as more dissimilar by the native English than Spanish subjects. A significant group×language composition interaction was obtained in the second test of the typicality hypothesis [$F(6,112)=9.62$, $p<0.01$]. The effect of language composition was significant for all four groups [$F(2,28)=6.78$ to 21.8, $p<0.01$]. More importantly, the simple effect of group was significant for the Spanish–Spanish pairs [$F(3,56)=4.36$, $p<0.01$] and for the Spanish–English pairs [$F(3,56)=3.89$, $p<0.01$] but not for the English–English pairs [$F(3,56)=1.45$, $p>0.10$].

The lack of a significant difference between groups for the English–English pairs in the second test failed to support the typicality hypothesis, but neither did it contradict the hypothesis, as in the first analysis. An HSD test revealed, however, that the Spanish–English pairs were more dissimilar for the subjects in EnA than for the Spanish subjects in SA. Such a difference would not be expected if perceived dissimilarity varied according to typicality. Thus, taken together, the results failed to support the typicality hypothesis.

4. Categorical status

Several analyses were undertaken to test the hypothesis that differential classification increases perceived dissimilarity. The pairs examined in the first two analyses are listed in Table V. Half of these pairs were classified as "adjacent" pairs, the other half as "nonadjacent." As shown in Fig. 1, vowels in the adjacent pairs were less distant in an $F1-F2$ space than were the vowels in the nonadjacent pairs. Also, with the possible exception of one English–English pair (viz., /ɑ/–/ʌ/), vowels in the adjacent pairs were not separated by another Spanish or English vowel category in the same way as were vowels in most nonadjacent pairs.

Four mean values were calculated for each subject: one each for the adjacent and nonadjacent English–English pairs, and one each for the adjacent and nonadjacent Spanish–English pairs. Each mean was based on the means obtained for the pairs shown in Table V. The means for English–English pairs are shown in Fig. 4 (top). As expected, vowels in the nonadjacent pairs were rated as more dissimilar than vowels in the adjacent pairs by both the Spanish and English subjects. The difference was much

TABLE V. Vowel pairs used to test the hypothesis that the categorical status of the two vowels in a pair influence their perceived dissimilarity (see text). All English–English pairs contained realizations of two different English vowel categories. The Spanish–English pairs had one Spanish and one English vowel (order counterbalanced).

Adjacent pairs	
English–English	Spanish–English
/i/,/ɪ/	/i/,/ɪ/
/ɛ/,/æ/	/e/,/e^{I}/
/ɛ/,/ʌ/	/e/,/ɛ/
/æ/,/ɑ/	/a/,/ɑ/
/æ/,/ʌ/	/a/,/ʌ/
/ʌ/,/ɑ/	/a/,/æ/
Nonadjacent pairs	
English–English	Spanish–English
/i/,/ɑ/	/i/,/ɛ/
/ɪ/,/ɑ/	/e/,/æ/
/e^{I}/,/ɑ/	/e/,/ɑ/
/i/,/æ/	/a/,/i/
/ɪ/,/æ/	/a/,/ɪ/
/e^{I}/,/æ/	/a/, /e^{I}/

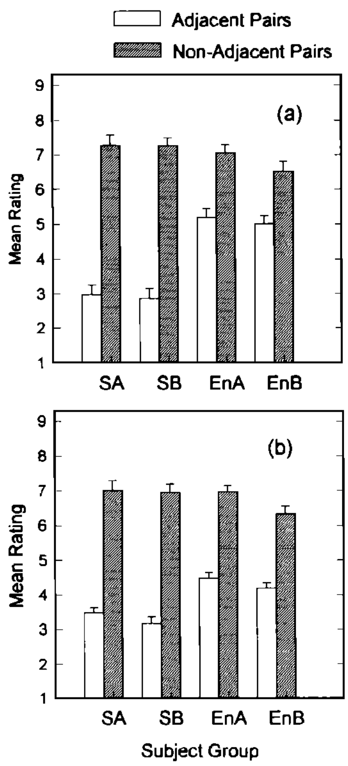

FIG. 4. Mean dissimilarity ratings for pairs of vowels that were relatively close ("adjacent") or distant ("nonadjacent") in an $F1-F2$ acoustic space by relatively inexperienced and experienced Spanish speakers of English (SA,SB) and two native English groups (EnA,EnB). The pairs in (a) were made up of two English vowels whereas those in (b) were made up of one English and one Spanish vowel (see text). The error bars enclose ± 1 standard error.

greater for the native Spanish than for the native English subjects, however.

The means were examined in (4) group$\times$(2) pair type ANOVAs, one for the adacent and nonadjacent English–English pairs, and one for the adjacent and nonadjacent Spanish–English pairs. The prediction being tested was that the Spanish and English subjects would differ more for the adjacent pairs than for the nonadjacent pairs. In fact, the ANOVA examining English–English pairs yielded a significant two-way interaction [$F(3,56)=25.5$, $p<0.01$]. The simple main effect of the group factor was significant for adjacent [$F(3,56)=22.4$, $p<0.01$] but not nonadjacent pairs [$F(3,56)=1.79$, $p>0.10$]. An HSD test revealed that, for the adjacent pairs, mean dissimilarity was significantly greater for the two English groups (EnA, EnB) than for the two Spanish groups (SA, SB) ($p<0.05$). Groups SA and SB did not differ, nor did groups EnA and EnB ($p>0.10$).

As shown in Fig. 4 (bottom), much the same results were obtained for the Spanish–English pairs. Once again, the difference in perceived dissimilarity between the adjacent and nonadjacent pairs was greater for the native Spanish than native English subjects, leading to a significant group$\times$pair type interaction [$F(3,56)=10.1$, $p<0.01$]. The simple effect of group was significant for adjacent [$F(3,56)=13.9$, $p<0.01$] but not non-adjacent pairs [$F(3,56)=1.78$, $p>0.10$]. An HSD test revealed that, for adjacent pairs, mean dissimilarity was significantly greater for the two English groups than for the two Spanish groups ($p<0.05$). Groups SA and SB did not differ, nor did groups EnA and EnB ($p>0.10$). This was consistent with the hypothesis that differential classification increases perceived dissimilarity.

Results of a third analysis also supported the differential classification hypothesis. It examined pairs containing Spanish /a/ and one of five English vowels that varied in $F1-F2$ distance from Spanish /a/ (viz., /i ɪ æ ɛ ʌ/). As shown in Fig. 5, the subjects in all four groups rated /a/–/i/ pairs as more dissimilar than /a/–/ɪ/ pairs, as expected from the $F1-F2$ distances between vowels in these two pairs. All four groups of subjects rated the /a/–/ʌ/ pairs as being very similar, also in accord with $F1-F2$ distances. However, the native English and Spanish subjects differed considerably for the /a/–/æ/ and /a/–/ɛ/ pairs. The subjects in both English groups (EnA, EnB) rated these pairs as more dissimilar than the subjects in both Spanish groups (SA, SB), which yielded a significant group$\times$pair interaction [$F(12,224)=12.9$, $p<0.01$]. The simple main effect of group was nonsignificant for /a/–/i/, /a/–/ɪ/, and /a/–/ʌ/, but it was significant for the /a/–/æ/ [$F(3,56)=28.2$, $p<0.01$] and /a/–/ɛ/ pairs [$F(3,56)=12.6$, $p<0.01$]. An HSD test revealed that, for these last two vowel pairs, perceived dissimilarity was significantly greater for the subjects in EnA and EnB than for the subjects in SA and SB ($p<0.05$).

C. Discussion

The experiment 1 results suggested that the perception of vowel dissimilarity changes little overall as adult native Spanish adults gain experience with English vowels. Relatively experienced and inexperienced Spanish subjects did not differ significantly for any of the 45 vowel pair types examined here. A difference might have been obtained had we compared native Spanish groups differing to a greater extent in L2 experience (see Flege and Fletcher, 1992) or, more likely perhaps, native Spanish groups differing in L2 *proficiency*. In a *post-hoc* analysis, we examined the correlation between the 30 Spanish subjects' mean dissimilarity ratings and (a) their overall experience in English and (b) their overall ability to pronounce English ("proficiency"). Five correlations with L2 proficiency reached significance ($p<0.05$), whereas none involving L2 experience did so.[10] Additional research will be needed to determine if L2 proficiency is indeed linked more closely to adult L2 learners' perception of dissimilarity in L1 and L2 vowels than is amount of L2 experience.

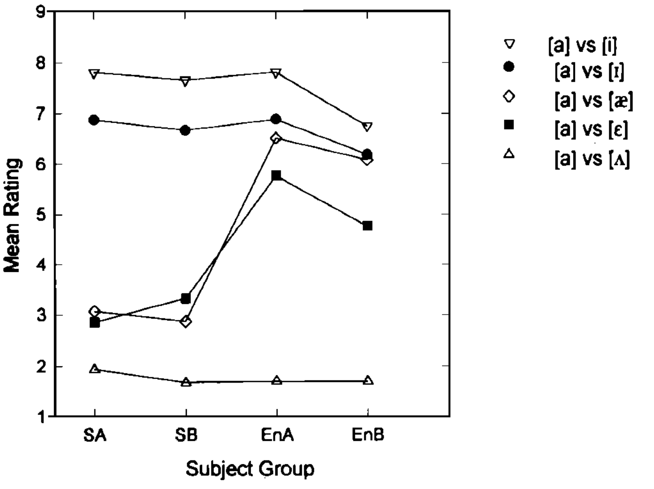

FIG. 5. Mean dissimlarity ratings for pairs consisting of a token of Spanish /a/ and a token of one of five English vowel categories differing in distance from Spanish /a/ in an $F1-F2$ acoustic space by relatively inexperienced and experienced Spanish speakers of English (SA,SB) and two native English groups (EnA,EnB).

Support was provided for two of the three factors hypothesized to influence ratings of vowel dissimilarity. One factor was auditory difference, which was roughly estimated here as the distance between vowels in an $F1-F2$ bark space. The subjects in the two English groups and the two Spanish groups all judged the vowels in high–low pairs to be significantly more dissimilar than the vowels in high–mid pairs. This is a striking finding when one considers that English has far more vowels than Spanish (15 vs 5), that the vowels said to be "shared" by English and Spanish differ phonetically (Flege, 1989), and that native English speakers use more perceptual dimensions in classifying vowels than do native speakers of Spanish (Fox *et al.*, under review).

Kewley-Port and Atal (1989) observed that the perceived dissimilarity of synthetic vowels varied as a function of Euclidean bark $F1-F2$ distances. The results obtained here extend this finding to naturally produced vowels that were more distant in an $F1-F2$ space, and support the observation by Fox (1985, p. 215) that vowel dissimilarity judgments are not based solely on "phonetic labels or long-term memory prototypes." Given the lack of significant differences between the native Spanish and English subjects, one might conclude that some part of vowel perception is "universal" (e.g., Stevens *et al.*, 1969).

Other results, to be discussed later, confirmed the inference by Fox (1985, p. 216) that pairs of vowels being judged for dissimilarity are categorized implicitly. If one assumes that identification occurs after an auditory (sensory) stage of analysis (e.g., Pisoni, 1973; Pisoni and Luce, 1987; Bladon and Lindblom, 1981; Bladon, 1986), then one might infer from experiment 1 that listeners maintain auditory (sensory) information after vowels are categorized. This is consistent with evidence that "redundant" information, including information relevant to talker identity, may be maintained after categorization (Goldinger, 1992).

Other analyses suggested differences in perceived vowel dissimilarity as a function of how the Spanish and English subjects categorized vowels. There was little difference between the English and Spanish subjects for vowels in "nonadjacent" pairs. The vowels in these pairs were relatively distant in an $F1-F2$ acoustic space. However, the English subjects judged the less distant vowels in "adjacent" pairs to be more dissimilar than the Spanish subjects. We assumed that both the Spanish and English subjects would identify the vowels in nonadjacent pairs in terms of two phonetically distinct vowel categories. However, based on their distance in an $F1-F2$ space, we assumed that at least some Spanish subjects would identify vowels in adjacent pairs in terms of just *one* category.

The difference between adjacent and nonadjacent pairs for the native Spanish subjects suggested that the differential classification of a pair of vowels increases their degree of perceived dissimilarity (see also Kewley-Port and Atal, 1989; Bladon and Lindblom, 1981). Perhaps the availability of two distinct phonetic codes increases perceived dissimilarity (Fujisaki and Kawashima, 1970; Pisoni, 1973; Repp and Crowder, 1990). This inference assumes, of course, that the Spanish and English subjects in experiment 1 were equally likely to identify vowels in nonadjacent pairs in terms of distinct vowel categories, whereas the English subjects were more likely than the Spanish subjects to identify the vowels in *adjacent pairs* in terms of distinct categories.

This assumption seems plausible in the light of previous research.[11] Also, given the nature of the hypothesis being tested, it was only necessary that there be a *greater likelihood* of differential classification of adjacent pairs by the English than Spanish subjects. Thus, if our assumption concerning any of the so-called adjacent pairs was incorrect, it would work *against* the hypothesis, which was nevertheless supported. Still, given the ambiguities just discussed, we decided to carry out a second experiment (see below) to further test the conclusion that differential classification increases perceived dissimilarity.

Experiment 1 also tested a typicality hypothesis, viz., that pairs of familiar L1 vowels would be rated as more dissimilar than a pair of vowels drawn from a second or foreign language. ANOVAs examined three Spanish–Spanish pairs, three English–English pairs, and three Spanish–English pairs. Given the distances from one another in an $F1-F2$ acoustic space, we expected the vowels in all nine pairs to be identified in terms of two different vowel categories by both the native Spanish and English subjects. We assumed, further, that realizations of vowel categories from a subject's L1 would match the subject's long-term memory representations for vowels to a greater extent than would vowels drawn from a second or foreign language. This led to the prediction that the English subjects would judge English–English pairs to be more dissimilar than the Spanish subjects, whereas the Spanish subjects would rate Spanish–Spanish pairs to be more dissimilar than the native English subjects. As predicted, the Spanish subjects rated the Spanish–Spanish pairs to be more dissimilar than the English subjects. However, the Spanish subjects also rated certain English–English pairs to be more dissimilar than the English subjects. Although experiment 1 did not support the typicality hypothesis, neither did it disconfirm the hypothesis. It will therefore be necessary to further test the typicality hypothesis with Spanish monolinguals.

II. EXPERIMENT 2

The results of experiment 1 suggested that, all else being equal, a pair of vowels identified in terms of two categories will be perceived as more dissimilar than a pair of vowels identified in terms of a *single* category. This conclusion was based on the assumption that both the native English and Spanish subjects identified vowels in pairs designated "nonadjacent" in terms of two phonetically distinct categories, whereas the Spanish subjects were less likely than the English subjects to do so for vowels in "adjacent" pairs. The categorical status of vowels in the two kinds of pairs was not tested in experiment 1, however. The purpose of experiment 2 was therefore to further test the conclusion that the differential classification of a pair of vowels augments their degree of perceived dissimilarity. This was done by comparing the results of a rating task to the results obtained in a oddity discrimination task administered following the rating task. We reasoned that if the native English subjects but not the native Spanish subjects were able to correctly choose the odd item out in triads containing two phonetically distinct vowels drawn from the same general portion of the vowel space (e.g., Spanish /a/–English/æ/–Spanish /a/), it would suggest that the native English subjects, but not the native Spanish subjects, treated the two phonetically distinct vowels contained in each triad as phonetically distinct. The conclusion concerning the effect of categorical status would be supported if pairs of vowels that were successfully discriminated in the oddity task were judged to be more dissimilar than pairs of vowels that were not successfully discriminated.

A. Method

1. Subjects and stimuli

Three groups of subjects participated: monolingual speakers of English (1 male, 12 females), relatively inex-

TABLE VI. Characteristics of the three groups of subjects in experiment 2. Standard deviations are in parentheses.

	CA[a]	AOA[b]	UND[c]	SPK[d]	LOR[e]	USE[f]	INPUT[g]
Native English ($n=13$)	24.6 (4.9)	...	...	...	...	...	...
Experienced Spanish ($n=12$)	35.9 (8.9)	18.3 (12.9)	22.8 (11.7)	24.1 (11.4)	10.5 (7.1)	72.8 (18.6)	7.6 (5.3)
Inexperienced Spanish ($n=12$)	30.3 (8.6)	23.9 (8.6)	26.7 (7.2)	27.5 (6.9)	1.4 (1.1)	45.7 (24.0)	0.6 (0.5)

[a]CA, chronological age at testing, in years.
[b]AOA, age of arrival in the U.S., in years.
[c]UND, age at which English could first be understood "comfortably," in years.
[d]SPK, age at which English could first be spoken "comfortably," in years.
[e]LOR, length of residence in the U.S., in years.
[f]USE, self-estimated percentage daily usage of English.
[g]INPUT, LOR × USE, expressed in years.

perienced native Spanish speakers of English (7 males, 5 females), and relatively experienced native Spanish speakers of English (2 males, 10 females). As summarized in Table VI, the relatively experienced native Spanish subjects (SB) had considerably more English-language experience than the relatively inexperienced (SA) subjects. Only one subject had participated previously in experiment 1.

Twenty-one CVs from experiment 1 were used as stimuli: three CV stimuli containing Spanish /a/ and three CVs each containing English /i/, /ɪ/, /e$^{\mathrm{I}}$/, /ɛ/, /ɑ/, /ʌ/. (One digitally corrupted Spanish /a/ token from experiment 1 had to be replaced.) As can be seen in Table I, the three "nonadjacent" pairs formed from these tokens (Spanish /a/–English /i/, Spanish /a/–English /ɪ/, Spanish /a/–English /e$^{\mathrm{I}}$/) were more distant from one another in an $F1-F2$ space than were the three "adjacent" pairs (Spanish /a/–English /ɛ/, Spanish /a/–English /ɑ/, Spanish /a/–English /æ/). For the rating task, the three Spanish /a/ tokens were paired twice with all three tokens of each English vowel, once in the first position and once in the second position. This yielded a total of 108 pairs.

The triads used in the oddity discrimination task were formed using vowels that had been paired in the rating task. Most triads contained a single realization of one vowel category and two realizations of another category. In roughly half of the triads, a single Spanish /a/ token occurred with all possible combinations of two different realizations of an English vowel, once in the first position ($n=18$), once in the second position ($n=18$), and once in the third position of the triad ($n=18$). The same procedure was used in forming 54 triads in which a single English vowel token occurred along with two Spanish /a/ realizations. In addition to the 108 triads just described, which each contained realizations of two phonetically distinct vowel categories (one Spanish, one English), there were 14 "catch" triads. These triads contained three different realizations (tokens) of a single category (e.g., the three tokens of Spanish /a/ in the corpus, which were spoken by three male talkers). Six practice trials presented at the beginning of the rating and oddity discrimination tasks were not analyzed.

2. Procedure

Subjects were tested individually in a sound booth, where they heard stimuli presented over headphones at a comfortable level using a Sony (Model TC-D5 Pro II) tape recorder. The subjects all participated in the oddity task after the rating task to ensure that the oddity task would not influence the subjects' dissimilarity ratings. The interstimulus interval between the pairs of CVs in the rating task was 1.2 s. The subjects rated each pair using a scale ranging from "very similar" (1) to "very dissimilar" (9). The subjects were strongly encouraged to use the whole scale, and were told to guess if unsure. The intertrial interval between pairs was sufficiently long, at 2.8 s, to permit the subjects to circle their response on an answer sheet. Following a short break, the subjects identified the odd item out in the oddity discrimination task. They were told that if one vowel was different from two other vowels in a triad, they were to indicate its serial position by circling "1", "2", or "3" on an answer sheet. They were to circle "N" (no odd item out) if there was no odd item out. The subjects were required to respond to each triad, and told to guess if unsure.

The oddity discrimination test did not assess vowel identification directly. There were nevertheless reasons to think that it provided a useful index of whether the Spanish subjects did or did not identify a pair of vowels in terms of two categories. The three tokens making up each triad were spoken by different individuals, and so differed according to dimensions that would not be expected to influence a vowel's identity. The subjects were told that there would not always be an odd item out. Indeed, catch triads made up of three realizations of a single vowel category were included to encourage subjects to disregard variations along dimensions such as talker identity that were not relevant to vowel identity. Finally, the interstimulus interval between the CVs in each triad was relatively long (viz., 1.2 s) to encourage the use of phonetic codes (Pisoni, 1973; Repp and Crowder, 1990).

TABLE VII. Mean ratings (top) and percent correct oddity discrimination scores (bottom) obtained in experiment 2. Vowels in the "nonadjacent" pairs were more distant from one another in an $F1-F2$ space than were the vowels in "adjacent" pairs (see text).

Pair type	Pair	NE[a]	SB[b]	SA[c]	$F(2,34)$	HSD test
			Rating task			
Nonadjacent	/a/–/i/	6.7	7.9	8.3	11.03[e]	SA,SB > NE
		(1.3)	(0.8)	(0.5)		
Nonadjacent	/a/–/ɪ/	6.0	6.9	7.2	4.23[d]	SA > NE
		(1.5)	(0.9)	(0.9)		
Nonadjacent	/a/–/eᴵ/	6.8	7.0	7.6	2.47	···
		(1.0)	(1.0)	(0.9)		
Adjacent	/a/–/ɛ/	4.4	4.5	4.4	0.01	···
		(1.3)	(1.5)	(1.4)		
Adjacent	/a/–/æ/	5.3	4.0	3.8	4.16[d]	SA < NE
		(1.4)	(1.1)	(1.6)		
Adjacent	/a/–/ɑ/	3.8	2.7	2.6	8.18[e]	SA,SB < NE
		(1.0)	(0.5)	(1.0)		
			Oddity task (% correct)			
Nonadjacent	/a/–/i/	98	95	95	1.09	···
		(4)	(6)	(5)		
Nonadjacent	/a/–/ɪ/	91	96	89	0.69	···
		(15)	(7)	(17)		
Nonadjacent	/a/–/eᴵ/	98	95	97	0.78	···
		(5)	(5)	(6)		
Adjacent	/a/–/ɛ/	80	66	51	3.62[d]	NE > SA
		(29)	(26)	(27)		
Adjacent	/a/–/æ/	77	59	38	6.60[e]	NE > SA
		(31)	(24)	(25)		
Adjacent	/a/–/ɑ/	57	26	18	8.14[e]	NE > SA,SB
		(36)	(18)	(17)		

[a]NE, native speakers of English.
[b]SB, relatively experienced Spanish speakers of English.
[c]SA, relatively inexperienced Spanish speakers of English.
[d,e]Probability at the 0.05 and 0.01 levels.

B. Results and discussion

1. Dissimilarity ratings

A mean dissimilarity rating was calculated for each subject for each of the three adjacent and the three nonadjacent pair types. These ratings are presented in Table VII, along with the percent correct scores obtained in the later, oddity discrimination task for the six pairs. Each was based on 18 judgments. An average rating for the three adjacent and three nonadjacent pairs was then calculated for each subject.

The two overall mean ratings obtained for each subject were examined in a (3) group×(2) pair type ANOVA, with repeated measures on the pair type factor. The mean ratings for adjacent and nonadjacent pairs are shown in Fig. 6 (top). As expected, vowels in the nonadjacent pairs were rated as more dissimilar than vowels in the adjacent pairs by the subjects in all three groups. The difference in ratings between adjacent and nonadjacent difference was larger, however, for the inexperienced than experienced Spanish subjects. The difference for the experienced Spanish subjects, in turn, was greater than for the native English subjects.

Not surprisingly, the ANOVA yielded a significant two-way interaction $[F(2,34)=11.2,\ p<0.01]$. The group effect was significant for both the nonadjacent pairs $[F(2,34)=6.57,\ p<0.05]$ and the adjacent pairs

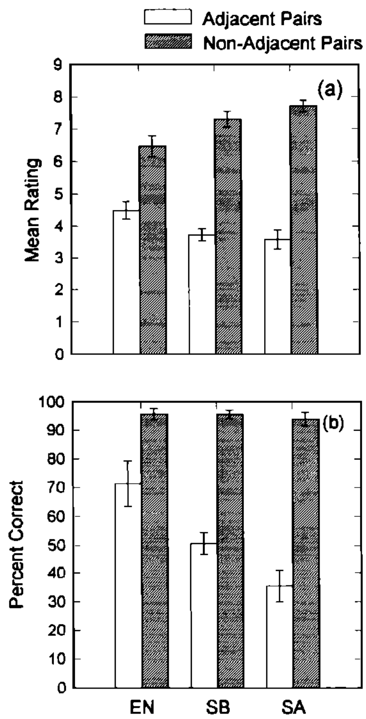

FIG. 6. (a) Mean dissimilarity ratings obtained in experiment 2 for vowels in three adjacent and three nonadjacent pairs. Ratings were obtained from native speakers of English (NE), experienced Spanish speakers of English (SB), and relatively inexperienced Spanish speakers of English (SA). (b) Percent correct scores from an oddity discrimination test.

$[F(2,34)=3.87, \quad p<0.05]$. An HSD test revealed that the English subjects rated adjacent vowels as significantly more dissimilar than the inexperienced Spanish subjects (<0.05), but they did not differ significantly from the experienced Spanish subjects ($p>0.05$), nor did the two Spanish groups differ significantly from one another ($p>0.05$). This finding can be taken as support for the differential classification hypothesis provided, of course, that the Spanish subjects were less likely than the native English subjects to treat vowels in the adjacent pairs as phonetically distinct. For nonadjacent pairs, the native English subjects rated vowels as being significantly *less dissimilar* than the inexperienced Spanish subjects, but they did not differ significantly from the experienced native Spanish subjects ($p<0.05$). This may mean that Spanish–English bilinguals are more sensitive to differences between Spanish and English vowels than are English monolinguals (see Sec. I. C).

2. Oddity discrimination

The oddity test results support an important assumption made in experiment 1, viz., that the native Spanish subjects would be less likely than the native English subjects to identify vowels in adjacent pairs (which were drawn from categories that are close in an $F1-F2$ acoustic space) in terms of two categories. Six percent correct discrimination scores were obtained for each subject in the oddity task, each based on 18 responses. Three scores were for adjacent pairs, and three were for nonadjacent pairs. Two average percent correct scores were then calculated for each subject, one for the adjacent pairs and one for the nonadjacent pairs.

The mean percentage of correct responses on the oddity task are shown in Fig. 6 (bottom). As expected, the subjects in all three groups responded correctly more often to triads made up of vowels drawn from categories that were relatively distant in an $F1-F2$ space (i.e., the nonadjacent triads) than to triads made up of vowels that were less distant (i.e., the adjacent triads). The three groups differed little for the nonadjacent triads, for which the correct response rate always exceeded 90%. Much larger between-group differences were evident for the adjacent triads, however. For these triads, the rate of correct responses averaged 71% for native English subjects and 51% for experienced Spanish subjects. The rate obtained for the inexperienced Spanish subjects, 35% correct, was only slightly higher than guessing.

The percent correct discrimination scores were not examined in group$\times$pair ANOVA because there was so little variance in the scores for nonadjacent triads. Instead, separate one-way ANOVAs examining the effect of the group factor were carried out for the adjacent and nonadjacent triads. As expected, the effect of group was nonsignificant for the nonadjacent triads $[F(2,34)=0.27, \quad p>0.10]$. The effect of group was significant, however, for adjacent triads $[F(2,34)=9.64, \quad p<0.05]$. An HSD test revealed that the native English subjects' scores for the adjacent pairs were significantly higher than those of subjects in groups SA and SB ($p<0.05$), but that the two Spanish groups did not differ significantly from one another ($p>0.10$).

3. Correlations between the rating and oddity tests

We wanted to learn if the more dissimilar the Spanish subjects judged a pair of vowels to be, the more often they would correctly choose the odd item out in triads made up of those vowels. To determine this, the 24 native Spanish subjects' oddity discrimination scores and their mean ratings were submitted to Pearson product-moment correlation analyses. Correlations for all three nonadjacent pairs were nonsignificant owing to ceiling effects for the oddity discrimination task $[r(22)=-0.008$ to $0.285, \quad p>0.10]$. Significant correlations were obtained, however, for the adjacent pairs /a/–/ɛ/ $[r(22)=0.679, \quad p<0.01]$ and /a/–/æ/ $[r(22)=0.489, \quad p<0.01]$.

The correlation for the third adjacent pair, /a/–/ɑ/, was nonsignificant $[r(22)=0.312, p>0.10]$. However, the results shown in Table VII for the three groups are consistent with the correlation results obtained for the two other adjacent pairs. The native English subjects' oddity discrimination scores for /a/–/ɑ/ were significantly higher than those of the SA and SB subjects and, at the same time, the native English subjects judged /a/ and /ɑ/ to be significantly more dissimilar than the native Spanish subjects ($p<0.05$). When the data for the 13 native English subjects were added to that obtained for the native Spanish subjects, a significant correlation between oddity discrimination and dissimilarity ratings for /a/–/ɑ/ was obtained $[r(35)=0.609, \quad p<0.01]$.

The findings for /a/–/ɛ/ and /a/–/æ/ are consistent with the hypothesis that perceived dissimilarity is increased by the differential classification of a pair of vowels. Flege (1991) found that the frequency with which Spanish speakers of English identified English /ɛ/ tokens in terms of Spanish /a/ (or /e/) declined with increasing English-language experience. It may be that adult Spanish learners of English discover only gradually the phonetic dissimilarity between English /æ/ and Spanish /a/ and that those who do so may, as a result, be better at discriminating Spanish /a/ and English /ɛ/ as the result of generating two distinct phonetic codes. The generation of two distinct phonetic codes may, in turn, increase the degree of perceived dissimilarity. The significant correlation for /a/–/æ/ might be interpreted in the same way, and might be taken as support for the hypothesis that some native Spanish subjects establish a category for English /æ/. The lack of a correlation for /a/–/ɑ/, on the other hand, suggests that few if any Spanish subjects noted the phonetic distinction between those vowels, perhaps due to the proximity of Spanish /a/ and English /ɑ/ in an $F1-F2$ space.

III. GENERAL DISCUSSION

One aim of this study was to assess three factors that might influence listeners' judgments of the perceived dissimilarity of vowels drawn from Spanish and/or English. Two of the three factors were demonstrated to influence perceived dissimilarity. The effect on dissimilarity ratings of vowels' distance in an $F1-F2$ space was much the same for native Spanish and English subjects. The greater the distance between vowels—and thus probably the greater the magnitude of their auditory differences—the more dissimilar the vowels were judged to be. This finding agrees with previous research (e.g., Pols *et al.*, 1969) in showing the importance of $F1$ and $F2$ frequency on the perception of vowel quality and vowel dissimilarity. It suggests that listeners can gain access to an auditory (sensory) code that shows little if any effect of language-specific experience. Use of this code may be responsible for what has been referred to as the "universal" basis of vowel perception (Terbeek, 1977).

Much previous research has suggested, nevertheless, that linguistic experience may "distort" or "warp" auditory-based judgments of similarity (Terbeek and Harshman, 1972). Such effects were evident here in analyses that focused on pairs of vowels that were likely to be identified in terms of two phonetically distinct categories by the native English subjects but not by the native Spanish subjects. The hypothesis that differential classification of a pair of vowels increases their perceived dissimilarity was tested by comparing the mean ratings obtained for two types of vowel pairs. "Nonadjacent" pairs (e.g., /i/–/ɑ/) contained vowels that were sufficiently distant in an $F1-F2$ space to virtually ensure identification in terms of two distinct vowel categories by all subjects, including the native Spanish subjects. Vowels in "adjacent" pairs such as /i/–/ɪ/, on the other hand, contained vowels that were closer in an $F1-F2$ space. We supposed that the English subjects would identify the vowels in adjacent pairs such as /i/–/ɪ/ in terms of the two categories, but that at least some Spanish subjects would not do so. This led to the prediction that native English and Spanish subjects would not differ in rating the vowels in nonadjacent pairs, whereas the English subjects would judge the vowels in adjacent pairs to be more dissimilar than the Spanish subjects.

The predicted effect of differential classification was observed in experiment 1, but this experiment provided no independent assessment of whether vowels were or were not identified in terms of two distinct categories. In experiment 2, subjects rated pairs of vowels for dissimilarity, then participated in an oddity discrimination task. Triads in the oddity task were formed from the pairs of vowels rated earlier for dissimilarity. Several features of the oddity task (use of multiple talkers, a relatively long ISI, inclusion of catch trials) were deemed to make it a good indicator of whether or not the vowels in a pair were classified in terms of two phonetically distinct categories.

The experiment 2 rating results replicated findings obtained in experiment 1. The English and Spanish subjects' correct response rates in the oddity test did not differ significantly for nonadjacent triads, but the English subjects' rates were significantly higher than the Spanish subjects' rates for the *adjacent* triads. Correlations between the Spanish subjects' dissimilarity ratings and their oddity test scores were nonsignificant for nonadjacent pairs. This was expected, for we supposed that the Spanish subjects would identify vowels in nonadjacent pairs in terms of two distinct categories, leading to ceiling effects in the oddity discrimination test results. The Spanish subjects' ratings and oddity test scores for adjacent pairs, on the other hand, correlated significantly. The more often they correctly chose the odd item out in triads containing realizations of Spanish /a/ and English /æ/ (or realizations of Spanish /a/ and English /ɛ/), the more dissimilar they judged realizations of these vowel categories to be. No such correlations were observed for the native English subjects, even though their performance on the oddity test was not at ceiling.

Taken together, the results of experiments 1 and 2 indicated that differential classification of a pair of vowels increases their degree of perceived dissimilarity. It may be that the presence of two different phonetic codes increases the vowels' perceived dissimilarity (e.g., Fujisaki and Kawashiima, 1970; Pisoni, 1973). Alternatively, the effect of perceiving two categorically distinct vowels may be to bias dissimilarity judgments post-perceptually. The exact locus of the effect of phonetic classification on dissimilarity judgments remains to be determined in future research.

We also tested the effect of typicality (familiarity) on vowel dissimilarity ratings. It was hypothesized that a pair of familiar L1 vowels would closely match vowels in long-term memory than would a a pair of vowels drawn from a second or foreign language, and that this would increase the perceived dissimilarity of the familiar (L1) vowels. The specific prediction being tested was that the Spanish subjects would rate Spanish–Spanish pairs as being more dissimilar than the English monolingual subjects and, conversely, that the English subjects would rate English–English pairs as being more dissimilar than the Spanish subjects. The predicted effect was observed for Spanish–Spanish pairs. However, the native Spanish subjects did not rate the English–English pairs as less dissimilar than did the native English subjects. In fact, in some instances the Spanish subjects rated English–English vowel pairs as *more* dissimilar than the English subjects.

Perhaps the Spanish subjects' familiarity with the vowels in two languages generally heightened their awareness of between-vowel differences. Experience may influence which features in complex stimuli are attended to, and the relative weights assigned to each during feature integration (see Hintzman, 1986; O'Toole *et al.*, 1991). The psychological vowel space of Spanish/English bilinguals might be regarded as more "crowded" than that of the English monolinguals. One effect of crowding seems to be an increase in perceived dissimilarity (see Butcher, 1976; Jusczyk, 1991). It would therefore be useful to test the typicality hypothesis with Spanish monolinguals. They may, as predicted, judge Spanish–Spanish pairs to be more dissimilar than English monolinguals and also judge English–English pairs to be *less dissimilar* than English monolinguals.

The other major aim of this study was to determine if the perception of vowel dissimilarity changes as Spanish speakers gain experience in English. The dissimilarity ratings obtained in experiment 1 for relatively experienced and inexperienced Spanish learners of English, all of whom learned English as adults, did not differ significantly. This suggested that simple exposure to an L2 is unlikely, in itself, to result in changes in the perception of vowel dissimilarity. One incidental finding of the study pointed to the importance of individual differences in L2 acquisition. No correlations were found to exist between the amount of English-language experience of the Spanish subjects in experiment 1 and their ratings of dissimilarity. However, a few significant correlations were noted between the subjects' English-language proficiency (i.e., estimated ability to pronounce and comprehend spoken English) and their vowel dissimilarity ratings. In the oddity discrimination task of experiment 2, somewhat better performance was noted for experienced than inexperienced Spanish subjects for triads in which Spanish /a/ tokens were juxtaposed to tokens of neighboring English vowels (/æ/, /ɛ/, or /ɑ/). The difference was nonsignficant, however, owing to the high degree of intersubject variability. Of the 24 Spanish subjects tested, six experienced and four inexperienced subjects' oddity tests scores for triads with /a/ and English /æ/ fell within the native English subjects' range. One tentative conclusion that might be drawn from these findings is that success in detecting cross-language vowel differences depends on individual patterns of auditory processing and/or phonetic organization. It would be useful to determine if those Spanish subjects who performed well on triads with English /æ/, which suggests the possibility of category formation for English /æ/ (e.g., Flege, 1988) pronounce this vowel more accurately than those who were unable to discriminate English /æ/ from Spanish /a/.

In summary, the results of this study suggested that Spanish/English bilinguals' perception of dissimilarity in English vowels, and in pairs of Spanish and English vowels, changes relatively little overall as a function of English-language experience. Three factors were hypothesized to influence cross-language vowel similarity ratings. Support was not provided for the hypothesis that vowel typicality influences judgments of vowel dissimilarity, although this hypothesis merits further attention. It was shown, however, that the perception of vowel dissimilarity varies as a function of the acoustic (hence auditory) distance between vowels in an $F1-F2$ acoustic space, and that the categorical status of the vowels in a pair being rated exerts an important effect on perceived dissimilarity.

ACKNOWLEDGMENTS

This research was supported by NIH Grant No. DC00257. The authors are very grateful to M. Carroll, A. Rodriguez, W. Strange, and A. Valentin for their help in acquiring data from native Spanish subjects in Tampa, Florida. Thanks are extended to C. Best, O.-S. Bohn, J. Ohala, and W. Strange for comments on earlier versions of this article.

[1] Insofar as natural boundaries exist, they may derive from nonlinearities in the auditory representations of acoustic properties of vowels (e.g., Bladon and Lindblom, 1981; Bladon, 1986).

[2] This assumption seemed to be warranted based on the close relationship that has been shown to exist between vowel identification and the frequencies of the first two or three formants (e.g., Fairbanks and Grubb, 1961; Pols *et al.*, 1969; Singh and Woods, 1971); the fact that the vowels most likely to be confused in an identification test are often adjacent in an $F1-F2$ space (e.g., Peterson and Barney, 1952); and the fact that the vowels most likely to be confused in short-term memory have similar $F1$ and $F2$ frequencies (e.g., Wicklegren, 1965).

[3] Flege (1991) had native Spanish subjects use five letters (i, e, a, o, u) to identify English /i/ and /ɪ/ tokens as one of the five vowels of Spanish. In another study, subjects identified the members of a synthetic continuum ranging from *beat-bit* as /i/ or /ɪ/ (Flege and Bohn, 1989). Blankenship (1991) had native Spanish subjects identify steady-state synthetic vowels in an $F1 \times F2$ matrix.

[4] Labeling of unfamiliar vowels might also be difficult for trained listeners. Despite the suggestion by Stevens *et al.* (1969) concerning natural boundaries between rounded and unrounded vowels, Ladefoged (1967) found that trained phoneticians sometimes confused unfamiliar rounded and unrounded vowels because of difficulty in separating the acoustic effects of lip rounding and tongue positioning.

[5] A simulation study by O'Toole *et al.* (1991) suggested that the well-known "other race" effect in facial recognition studies may arise from the perception of greater dissimilarity for pairs of faces belonging to a familiar (majority) culture than to an unfamiliar (minority) culture. The associative network that was developed appeared to "tune" its features and feature weightings so as to better accommodate the faces that had been presented most frequently.

[6] The "+" symbol, which indicated a syllable boundary, occurred on lists but did not seem to distract the talkers who read the lists.

[7] The subjects' experience in English was estimated by multiplying self-reported percentage daily usage of English (at school or work, at home, in social settings) by length of residence in an English-speaking environment. Surprisingly, the subjects in groups SA and SB did not seem to differ greatly in ability to pronounce and comprehend spoken English. One author (MJM) used a 7-point scale to rate each subject's ability to pronounce and comprehend English after he had administered a language background questionnaire. The subjects also used 7-point scales to rate their own pronunciation and comprehension abilities. As expected (Flege and Eefting, 1987), the average of the experimenter's ratings and the average of the listeners' self-ratings correlated significantly [$r=0.636$, $p<0.01$]. When the two sets of average ratings were summed, groups SB and SA differed little (9.3 vs 8.6).

[8] The overall mean ratings obtained for subjects in the four groups (SA-5.6, SB-5.5, EnA-6.0, EnB-5.6) did not differ significantly [$F(3,56)=1.71$, $p>0.10$], nor did the *range* of ratings used by the four groups (SA-6.7, SB-7.2, EnA-7.0, EnB-6.4) differ significantly [$F(3,56)=1.77$, $p>0.10$].

[9] One might argue that this would not hold true for /i/. However, although the Spanish and English /i/ stimuli differed little in terms of $F1-F2$ frequency (see Fig. 1), they did differ in duration.

[10] Amount of L2 experience was estimated by multiplying the number of years each subject had lived in the United States by the subjects' estimates of their daily usage of English. L2 proficiency was estimated by adding the subjects' average self-rating of ability to comprehend and pronounce English (obtained using 7-point scales) to the ratings accorded each subject by one of the authors (MJM). Groups SA and SB differed little according to the L2 proficiency measure (see Table II). Four of the five pairs involved in significant correlations with L2 proficiency were classified as adjacent (viz., Spanish /i/–English /ɪ/, English /ɛ/–English /æ/, English /ɛ/–English /ʌ/, English /ɛ/–Spanish /a/). For these pairs, we would expect dissimilarity ratings to vary as a function of whether the vowels in them were, or were not, identified as realizations of two categories (see below).

[11] The assumption that the Spanish subjects would be as likely as the English subjects to differentially identify the vowels in the nonadjacent pairs seems uncontroversial, for they all differed according to phonological vowel height, and many also differed in duration and formant movement patterns. The assumption that vowels in adjacent pairs would *not* be differentially identified by some Spanish subjects was less certain, however. The six English–English pairs classified as adjacent were /i/–/ɪ/, /ʌ/–/ɑ/, /æ/–/ɑ/, /ɛ/–/æ/, /æ/–/ʌ/, and /ɛ/–/ʌ/.

Two studies supported our assumption regarding /i/–/ɪ/. Flege and Bohn (1989) found that most, but not all, Spanish subjects had difficulty identifying /i/ and /ɪ/ based on spectral differences. Flege (1991) found that many, but not all, Spanish speakers identified English /i/'s and /ɪ/'s in terms of the Spanish /i/ category. We assumed that many Spanish subjects would identify both of the vowels in /ʌ/–/ɑ/ pairs as realizations of the Spanish /a/ category because the English /ʌ/ and /ɑ/ tokens used in experiment 1 clustered around the realizations of Spanish /a/, and realizations of all three vowel categories had similar patterns of formant movement. Support for the assumption regarding /æ/–/ɑ/ pairs came from the finding (Flege and Bohn, 1989) that Spanish subjects had no difficulty differentially identifying vowels in a *bet-bat* (/ɛ/–/æ/) continuum based on spectral differences. Given their difficulty identifying vowels in a *beat-bit* (/i/–/ɪ/) continuum, the authors inferred that the Spanish subjects must have identified the /æ/ end point in terms of the Spanish /a/ category. Also, Spanish subjects in the Flege (1991) study usually identified English /æ/'s in terms of the Spanish /a/ category. Finally, data reported by Scholes (1967a,b) suggested indirectly that the Spanish subjects would identify English /ɑ/'s in terms of Spanish /a/. The rationale for including /ɛ/–/æ/ among the adjacent pairs was that Spanish subjects in the Flege (1991) study used both "e" and "a" responses for English /ɛ/ tokens. Indirect support for the assumption that some Spanish subjects would not differentially identify the vowels in /æ/–/ʌ/ pairs was provided by Scholes (1967a). Data from that study suggested that synthetic vowels labeled /a/ by Spanish subjects may be labeled /ʌ/ by native speakers of English. Data from the Flege (1991) study suggest that English /æ/'s are often identified by Spanish subjects in terms of the Spanish /a/ category. No previous study provided evidence concerning the remaining adjacent pair (viz., /ɛ/–/ʌ/).

Bennett, D. C. (**1968**). "Spectral form and duration as cues to the recognition of German and English vowels," Lang. Speech **11**, 65–85.

Best, C., and Stange, W. (**1992**). "Effects of phonological and phonetic factors on cross-language perception of approximants," J. Phon. **20**, 305–331.

Best, C., McRoberts, G., and Sithole, N. (**1988**). "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants," J. Exp. Psychol.: Hum. Percept. Perform. **14**, 345–360.

Bladon, A. (**1986**). "Phonetics for hearers," in *Language for Hearers*, edited by G. McGregor (Pergamon, Oxford), pp. 1–24.

Bladon, A., and Lindblom, B. (**1981**). "Modeling the judgment of vowel quality differences," J. Acoust. Soc. Am. **69**, 1414–1422.

Blankenship, B. (**1991**). "Vowel perception in a second language," M.A. thesis (unpublished), Univ. California, Los Angeles.

Bohn, O.-S., and Flege, J. (**1990**). "Interlingual identification and the role of foreign language experience in L2 vowel perception," Appl. Psychol. **11**, 303–328.

Bohn, O.-S., and Flege, J. E. (**1992**). "The production of new and similar vowels by adult German learners of English," Stud. Sec. Lang. Acquis. **14**, 131–158.

Bothwell, R., Brigham, J., and Malpass, R. (**1989**). "Cross-racial identification," Person. Soc. Psychol. Bull. **15**, 19–25.

Brigham, J., Maass, A., Snyder, L., and Spaulding, K. (**1982**). "Accuracy of eyewitness identifications in a field setting," J. Person. Soc. Psychol. **42**, 673–681.

Butcher, A. (**1976**). "The Influence of the Native Language on the Perception of Vowel Quality," Arbeitsberichte Institut für Phonetik 6, Univ. Kiel.

Chance, J., Turner, A., and Goldstein, A. (**1982**). "Development of differential recognition of own and other-race faces," J. Psychol. **112**, 29–37.

Cohen, L. B., and Strauss, M. (**1979**). "Concept acquisition in the human infant," Child Devel. **50**, 419–424.

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," Int. Rev. Appl. Ling. **4**, 183–198.

Delattre, P. (**1969**). "An acoustic and articulatory study of vowel reduction in four languages," Int. Rev. Appl. Ling. **12**, 295–325.

Fairbanks, G., and Grubb, P. (**1961**). "A psychological investigation of vowel formants," J. Speech Hear. Res. **4**, 203–219.

Flege, J. (**1981**). "The phonological basis of foreign accent," TESOL Quart. **15**, 443–455.

Flege, J. (**1987**). "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification," J. Phon. **15**, 47–65.

Flege, J. (**1988**). "The production and perception of speech sounds in foreign languages," in *Human Communication and Its Disorders, A Review 1988*, edited by H. Winitz (Ablex, Norwood, NJ) pp. 224–401.

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," Quart. J. Exp. Psychol. **43A**, 701–731.

Flege, J. (**1992a**). "The Intelligibility of English vowels spoken by British and Dutch talkers," in *Intelligibility in Speech Disorders: Theory, Measurement, and Management*, edited by R. Kent (John Benjamins, Amsterdam), pp. 157–232.

Flege, J. (**1992b**). "Speech learning in a second language," in *Phonological Development: Models, Research, and Application*, edited by C. Ferguson, L. Menn, and C. Stoel-Gammon (York, Parkton, MD), pp. 565–604.

Flege, J., and Bohn, O.-S. (**1989**). "The perception of English vowels by native Spanish speakers," J. Acoust. Soc. Am. Suppl. 1, **85**, S85(A).

Flege, J., and Eefting, W. (**1987**). "Cross-language switching in stop consonant production and perception by Dutch speakers of English," Speech Commun. **6**, 185–202.

Flege, J., and Fletcher, K. (**1992**). "Talker and listener effects on the perception of degree of foreign accent," J. Acoust. Soc. Am. **91**, 370–389.

Fowler, C. (**1984**). "Segmentation of coarticulated speech in perception," Percept. Psychophys. **36**, 359–368.

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., Flege, J. E., and Munro, M. J. (**1994**). "A multidimensional scaling study of Spanish and English vowels," J. Acoust. Soc. Am. (submitted).

Fujisaki, H., and Kawashima, T. (**1970**). "Some experiments on speech perception and a model for perceptual mechanisms," Ann. Rep. Eng. Res. Inst., Univ. Tokyo **29**, 207–214.

Godinez, M. Jr., (**1978**). "A comparative study of some Romance vowels," UCLA Work. Papers Phon. **41**, 3–19 (Department of Linguistics, UCLA, Los Angeles, CA).

Goldinger, S. (**1992**). "Words and voices: Implicit and explicit memory for spoken words," Res. Speech Percept. Tech. Report 7, Speech Res. Lab., Indiana Univ.

Gottfried, T., and Beddor, P. (**1988**). "Perception of temporal and spectral information in French vowels," Lang. Speech **31**, 57–75.

Goggin, J., Thompson, C., Strube, G., Simental, L. (**1991**). "The role of language familiarity in voice identification," Mem. Cognit. **19**, 448–458.

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. Psychol. Res. **4**, 17–25.

Hintzman, D. (**1986**). "Schema abstraction' in a multiple-trace memory model," Psychol. Rev. **93**, 411–427.

Hebb, D. (**1949**). *The Organization of Behavior, A Neuropsychological Theory* (Wiley, New York).

Jusczyk, P. (**1991**). "How word recognition may evolve from infant speech perception capabilities," in *Cognitive Models of Speech Perception*, edited by G. Altmann and R. Shillcock (MIT, Cambridge, MA).

Kewley-Port, D., and Atal, V. (**1989**). "Perceptual distances between vowels located in a limited phonetic space," J. Acoust. Soc. Am. **85**, 1726–1740.

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.

Ladefoged, P. (**1967**). *Three Areas of Experimental Phonetics* (Oxford U.P., London).

Major, R. (**1987**). "Phonological similarity, markedness, and rate of L2 acquisition," Stud. Sec. Lang. Acquis. **9**, 63–82.

Munro, M. (**1992**). "Perception and production of English vowels by native speakers of Arabic," Ph.D. dissertation (unpublished), Univ. of Alberta.

Nosofsky, R. (**1986**). "Attention, similarity, and the identification-categorization relationship," J. Exp. Psychol. Gen. **115**, 39–57.

Nosofsky, R., (**1987**).

Nosofsky, R., Clark, S., and Shin, H. (**1989**). "Rules and exemplars in categorization, identification and recognition," J. Exp. Psychol.: Learn. Mem. Cognit. **15**, 282–304.

O'Toole, A., Deffenbacher, K., Abdi, H., and Bartlett, J. (**1991**). "Simulating the 'other-race effect' as a problem in perceptual learning," Connect. Sci. **3**, 163–178.

Peterson, G., and Barney, H. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 175–184.

Pisoni, D. (**1973**). "Auditory and phonetic memory codes in the discrimination of consonants and vowels," Percept. Psychophys. **13**, 253–260.

Pisoni, D., and Luce, P. (**1987**). "Acoustic-phonetic representations in word-recognition," Cognition **25**, 21–52.

Pols, L., van der Kamp, L., and Plomp, R. (**1969**). "Perceptual and physical space of vowel sounds," J. Acoust. Soc. Am. **46**, 458–467.

Repp, B., and Crowder, R. (**1990**). "Stimulus order effects in vowel discrimination," J. Acoust. Soc. Am. **88**, 2080–2090.

Scholes, R. (**1967a**). "Phoneme categorization of synthetic vocalic stimuli by speakers of Japanese, Spanish, Persian, and American English," Lang. Speech **10**, 46–68.

Scholes, R. (**1967b**). "Categorical responses to synthetic stimuli by speakers of various languages," Lang. Speech **10**, 252–282.

Scholes, R. (**1968**). "Phonemic interference as a perceptual phenomenon," Lang. Speech **11**, 86–103.

Schouten, M. E. H. (**1975**). "Native-language interference in the perception of second-language vowels," Ph.D. dissertation, University of Utrecht.

Shapiro, P., and Penrod, S. (**1986**). "Meta-analysis of facial identification studies," Psychol. Bull. **100**, 139–156.

Shigeno, S. (**1991**). "Assimilation and contrast in the phonetic perception of vowels," J. Acoust. Soc. Am. **90**, 103–111.

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. Ling. **7**, 231–237.

Stevens, K., Liberman, A., Studdert-Kennedy, M., and Öhman, 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).

Strange, W. (**1987**). "Information for vowels in formant transitions," J. Mem. Lang. **26**, 550–557.

Terbeek, D. (**1977**). A Cross-language Multidimensional Scaling Study of Vowel Perception. Working Papers in Phonetics, **37** (Dept. of Linguistics, UCLA, Los Angeles, CA).

Terbeek, D. and Harshman, R. (**1972**). "Is vowel perception non-Euclidian?" Work. Papers Phon. **22**, 13–29 (Dept. of Linguistics, UCLA, Los Angeles, CA).

Thompson, C. (**1987**). "A language effect in voice identification," Appl. Cognit. Psychol. **1**, 121–131.

Uchanski, M., Millier, D., Ronan, C., Reed, M., and Braida, L. (**1991**). "Effects of token variability on resolution for vowel sounds," J. Acoust. Soc. Am. Suppl. 1, **90**, 2254(A).

Van Tassell, D., Greenfield, D., Logemann, J., and Nelson, D. (**1992**). "Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants," J. Acoust. Soc. Am. **92**, 1247–1257.

Weinreich, U. (**1953**). *Languages in Contact, Findings and Problems* (Mouton, The Hague).

Wicklegren, W. (**1965**). "Distinctive features and errors in short memory for English vowels," J. Acoust. Soc. Am. **38**, 583–588.

Wode, H. (**1981**). *Learning a Second Language, I. An Integrated View of Language Acquisition* (Gunter Narr, Tubingen).

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.