

Differential use of temporal cues to the /s/-/z/ contrast by native and non-native speakers of English

James Emil Flege

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

James Hillenbrand

RIT Research Corporation, 75 Highpower Road, Rochester, New York 14623-3435

(Received 21 September 1984; accepted for publication 3 October 1985)

This study examined the effect of linguistic experience on perception of the English /s/-/z/ contrast in word-final position. The durations of the periodic ("vowel") and aperiodic ("fricative") portions of stimuli, ranging from *peas* to *peace*, were varied in a 5×5 factorial design. Forced-choice identification judgments were elicited from two groups of native speakers of American English differing in dialect, and from two groups each of native speakers of French, Swedish, and Finnish differing in English-language experience. The results suggested that the non-native subjects used cues established for the perception of phonetic contrasts in their native language to identify fricatives as /s/ or /z/. Lengthening vowel duration increased /z/ judgments in all eight subject groups, although the effect was smaller for native speakers of French than for native speakers of the other languages. Shortening fricative duration, on the other hand, significantly decreased /z/ judgments only by the English and French subjects. It did not influence voicing judgments by the Swedish and Finnish subjects, even those who had lived for a year or more in an English-speaking environment. These findings raise the question of whether adults who learn a foreign language can acquire the ability to integrate multiple acoustic cues to a phonetic contrast which does not exist in their native language.

PACS numbers: 43.71.Es, 43.70.Hs

INTRODUCTION

An adult's perception of speech is shaped to some extent by the system of phonemic oppositions found in his or her native language (L1), and by the way phonemes are phonetically realized in L1. One consequence of perceptual "tuning" to the acoustic properties of L1 phones is that subjects differing in L1 background may label the same phone differently (e.g., Williams, 1977). Cross-language differences in speech perception which arise as the result of early exposure to the phones found in a specific variety of speech imply that individuals who learn a foreign language (L2) may perceive the phones of that language differently than L2 native speakers.

Several studies have shown that L2 learners differ from L2 native speakers in the perception of L2 phonetic contrasts which differ acoustically from a similar contrast in L1. For example, several studies have shown that the phoneme boundaries of individuals who learned English as an L2 occur at different values along a voice onset time (VOT) continuum than those of English native speakers (Caramazza *et al.*, 1973; Williams, 1980). Other studies have shown that both children and adults may be unable to identify and/or discriminate L2 phonetic contrasts which do not exist in L1, or fail to do so on the basis of acoustic parameters that are sufficient perceptual cues to the L2 contrasts for L2 native speakers (Bond and Adamescu, 1979; Oller and Eilers, 1983; Werker and Tees, 1984a; Goto, 1971; Miyawaki *et al.*, 1975; Werker *et al.*, 1981; Elsendoorn, 1984).

The L2 learners are probably auditorily capable of perceiving phonetic contrasts in L2, but fail to exploit the full range of their sensory capabilities when identifying and discriminating L2 phones. For example, Werker and Tees

(1984b; see also Flege, 1984b) found that adults could discriminate certain L2 phones in isolation but not in a CV context. This suggested that processing speech in a "phonetic" mode (as opposed to an "auditory" mode; see, e.g., Repp, 1981) may invoke patterns of perceptual processing that have been shaped by linguistic experience. Processing L2 speech in a phonetic mode may induce the listener to use only cues that have proven efficacious in L1. This may make it difficult for L2 learners to discriminate or identify L2 phones not distinguished by acoustic dimensions that serve as perceptual cues to a phonetic contrast in L1 (Weinreich, 1953). Thus the goal of this study was to further probe the extent to which adults who have learned an L2 become sensitive to acoustic dimensions that do not distinguish phonetic contrasts in their native language.

Several studies have shown that adults can learn to identify and discriminate L2 phones on the basis of acoustic dimensions that do not distinguish L1 phones, either as the result of short-term training or naturalistic L2 learning (Strange and Dittman, 1981; Gillette, 1980; Pisoni *et al.*, 1982; McClasky *et al.*, 1980; Mochizuki, 1981; Shimizu and Dantsuji, 1983; MacKain *et al.*, 1981; Elsendoorn, 1984). Moreover, there is evidence that the processing of an acoustic dimension (such as VOT) may change as the result of exposure to L2 phones if the L2 phones are realized differently than their L1 counterparts (e.g., Caramazza *et al.*, 1973; Williams, 1980).

Despite such evidence, it is not certain that adults can learn to perceive all of the phonetic contrasts found in an L2, or that L2 learners who appear able to perceive an L2 phonetic contrast use the same perceptual cues or the same processing mode as L2 native speakers (see Redmond, 1977).

The L2 learners may develop sensitivity to some acoustic parameters before others and, as a result, learn to perceive certain L2 phonetic contrasts more rapidly than others. For example, Werker and Tees (1984a,b) found that laboratory training enabled English-speaking subjects to discriminate one Hindi phonetic contrast (the distinction between voiceless aspirated and breathy voiced stops) but not another (the distinction between dental and retroflex stops). It is unclear whether this difference had an auditory basis or resulted from a difference in the relative acoustic similarity of the Hindi phones being tested and phones in English.

Another recent study provided preliminary evidence that non-native speakers differ from native speakers of English in responding to temporal cues to the English /s/-/z/ contrast. Arabic /s/-/z/ are distinguished by voicing (i.e., glottal pulsing) in word-final position, but not by the duration of fricative noise or by the duration of the preceding vowel (Mittleb, 1982). In English, on the other hand, the periodic portion ("vowel," for short) of a word like *peas* (/piz/) is longer than that of *peace* (/pis/), whereas the aperiodic noise ("fricative" for short) associated with /s/ is longer than that associated with /z/ (Denes, 1955; Peterson and Lehiste, 1960; Raphael, 1971; Haggard, 1978). Lengthening the vowel and shortening the fricative noise in CVC stimuli increases /z/ judgments by native English speaker (Denes, 1955; Raphael, 1971; Gruenenfelder, 1979; Derr and Massaro, 1980; Soli, 1982). Flege (1984a) constructed three five-member continua ranging from *peace* to *peas* in which vowel duration was lengthened, fricative duration was shortened, or vowel and fricative duration varied together. Experienced and relatively inexperienced native Arabic speakers of English resembled native English subjects in identifying members of the continuum in which both vowel and fricative durations were manipulated. The experienced native Arabic speakers of English closely resembled native English speakers in showing an increase in /z/ judgments when either vowel duration was lengthened or fricative duration was shortened. However, the inexperienced native Arabic subjects did *not* show an increase in /z/ judgments when fricative duration alone was shortened.

The greater sensitivity of the experienced than inexperienced Arab subjects to changes in fricative duration probably resulted from their exposure to English rather than auditory processing differences, for listeners are sensitive to variations in fricative duration that do not signal categorical distinctions (Karno and Porter, 1980; cf. Lehiste and Shockey, 1980). However, it was not certain that even the experienced Arabs made the same *phonetic* use of temporal cues as the native English subjects. They might have simply labeled stimuli on the basis of perceived auditory changes in either vowel or fricative duration since these dimensions were varied separately in two of the continua. If non-native subjects process temporal cues to the English /s/-/z/ contrast in a truly *English-like* manner, both vowel and fricative duration should affect their voicing judgments simultaneously, as they do for native speakers of English.¹

Since vowel and fricative duration were not varied factorially by Flege (1984a), the aim of the present study was to determine whether adults who learn English as an L2 estab-

lish the ability to integrate (i.e., use simultaneously) both temporal cues to the English /s/-/z/ contrast. To do so, we constructed a two-dimensional array of stimuli and compared the identification responses of native speakers of French, Swedish, and Finnish to those of native speakers of English.²

Native speakers of Swedish and Finnish do not learn a contrast between /s/-/z/ in their native language³ for neither language possesses a /z/ phoneme. However, both Swedish and Finnish (like Arabic) possess a distinction between phonemically long and short phonemes. For example, the Swedish words *viss* and *vis* differ according to the duration of the vowel and final fricative noise. Phonologically, *viss* can be analyzed as a /CVCC/ word containing a short /i/ and a long /s:/; *vis* can be analyzed as a /CVVC/ word with a long /i:/ and short /s/ (Eliasson, 1982).⁴ As in English word pairs like *peace* versus *peas*, the durations of the vowel and final fricative portions of a monosyllabic Swedish word are not independent of one another. A long fricative is preceded by a short vowel, and vice versa. In Finnish, on the other hand, vowel and consonant quantity patterns are generally independent of one another (Lehiste, 1984). For example, a short /s/ can be preceded by a long or short vowel in the intervocalic position of two-syllable words (compare *pasi* and *paasi*), and both /s/ and /s:/ may occur after a short vowel (compare *passi* and *pasi*). Only short fricatives can occur in word-final position, where they can be preceded by either a long or short vowel (compare *taas* and *Tass*).

Native speakers of French were included as a control group. The French contrast between /s/-/z/ in word-final position appears to resemble the corresponding phonetic contrast in English. The effect of changing vowel and fricative duration on perception of the French /s/-/z/ contrast has not been tested, but existing speech production data suggest it will be roughly the same as for English listeners. Vowels are considerably longer before /z/ than /s/ in French words and, conversely, a longer fricative noise is associated with /s/ than /z/ (Kohler, 1979, 1984; Kohler *et al.*, 1981; Chen, 1970; Mack, 1982; O'Shaughnessy, 1984).

I. METHODS

A. Subjects

A total of eight groups were established. The subjects, who participated without being paid for their services, were chosen on the basis of language background, lack of evident speech or hearing disorder, and availability. A profile of subjects who had learned English as an L2 is presented in Table I. No attempt was made to ensure dialectal homogeneity among the non-native subjects, nor to evaluate their competence in English. It should be noted, however, that all of them were sufficiently competent in English to understand the experimental instructions, which were given in English.

The members of the three "inexperienced" non-native groups consisted of individuals who had never lived outside their native land. Their English-language experience was essentially restricted to English classes, movies, books, and music. Subjects in the three non-native groups designated "experienced" consisted of individuals who had lived in an

TABLE I. Group profile of the non-English subject groups. "Ed." refers to number of years of formal academic instruction in the English language; "mos." indicates the mean number of months spent in an English-speaking environment; "exp." designates subjects who were relatively experienced in English, "inexp." subjects who were less experienced. Standard deviations are in parentheses.

Subject group	N	Gender	Age	Ed.	Mos.
French (inexp.)	15	9M/6F	32.7 (9.3)	6.9 (2.9)	1.4 (2.2)
French (exp.)	10	5M/5F	38.2 (9.3)	7.0 (0.0)	36.5 (53.5)
Swedes (inexp.)	15	7M/8F	36.1 (6.9)	8.6 (1.5)	1.4 (1.2)
Swedes (exp.)	15	10M/5F	43.0 (7.7)	9.4 (2.1)	35.3 (21.2)
Finns (inexp.)	11	7M/4F	16.5 (0.5)	8.2 (0.8)	0.2 (0.4)
Finns (exp.)	10	6M/4F	27.0 (7.3)	10.2 (2.4)	15.1 (5.0)

English-speaking country (the United States or Great Britain) for at least 10 months. They undoubtedly had a great deal more exposure to the English spoken by native speakers than subjects in the corresponding "inexperienced" groups.

All but three of the non-native subjects were recruited in their native country. Most of the French subjects were affiliated with the Centre Nationale de la Recherche Scientifique in Paris or Strasbourg. The experienced Swedish subjects were affiliated with the Department of Linguistics at Stockholm or Uppsala University, or with the the Royal Institute of Technology (KTH) in Stockholm. The inexperienced Swedish subjects were affiliated with the Linguistics Department at Uppsala or Stockholm University. Most of the experienced Finnish subjects were members of the Faculty of Humanities at the University of Joensuu; the remaining three were Helsinki high school students who had just returned from a year in the United States. All of the inexperienced Finnish subjects were high school students from Joensuu.

To balance the design, we also examined two groups of monolingual native speakers of English. The data now available show no difference in the production and perception of temporal correlates of the /s/-/z/ distinction between native speakers of American and British English (the variety of English learned by most French, Swedish, and Finnish students). However, the first author has observed that native English speakers from Alabama seem to devoice final /z/ more frequently than speakers from Chicago, suggesting the possibility that they may produce (and therefore perceive) temporal differences between /s/-/z/ differently. A secondary aim of this study was therefore to determine whether speakers of these two dialects of American English make the same use of vowel and fricative duration as perceptual cues to the /s/-/z/ distinction. The members of one native English group consisted of Speech-Language Pathology students at Northwestern University in Chicago. Members of the other group consisted of Alabama residents who were

Speech-Language Pathology students at Montevallo College near Birmingham. The members of both groups were predominantly females in their early twenties who indicated that they were unable to speak any language other than English.

B. Stimuli

A 25-member stimulus set ranging from *peace* to *peas* was synthesized using the Klatt synthesizer (Klatt, 1980). The syllable-initial /p/ was synthesized with a 5-ms release burst and 50-ms formant transitions which terminated in values appropriate for English /i/ ($F1 = 280$ Hz, $F2 = 2350$ Hz, $F3 = 3000$ Hz, $F4 = 3700$ Hz, $F5 = 4000$ Hz). $F0$ varied from 128 to 110 Hz in three linear piecewise segments.

Soli (1982) found that the amplitude rise/fall times of postvocalic fricative noise in synthetic speech affects voicing judgments, so the rise/fall times used here were modeled on the first author's production of *peace* and *peas* (as determined by the average amplitude contours on spectrograms). The amplitude of the final fricative noise increased in a linear piece-wise fashion for 85 ms, and decreased in two linear segments over the final 60 ms. Amplitude was decreased in two linear segments over the final 100 ms of the periodic portion. Although periodic and aperiodic excitation overlapped by about 50 ms (see Flege, 1984a, for details), no periodicity was evident on spectrograms during the final fricative portion of the stimuli.

The duration of the periodic and final aperiodic portions of the CVC stimuli were each varied in five 50-ms steps. This was done by iterating a 5-ms frame from the steady-state portion of the vowel and fricative intervals. Vowel durations ranged from 150 to 350 ms; fricative duration ranged from 100 to 300 ms. These temporal values were selected on the basis of values reported in previous production studies (Peterson and Lehiste, 1960; Denes, 1955; Raphael, 1971) and on values measured in the first author's production of *peace* and *peas* in isolation.

The 25 stimuli were normalized for overall rms amplitude, converted to analog form with 12-bit amplitude resolution at a 10-kHz sampling rate, and low-pass filtered at 4.9 kHz. The fricative noise was not highly natural sounding because of the limited bandwidth, although the continuum endpoints were readily identifiable as *peace* or *peas* in the authors' estimation. The stimuli were recorded on a professional-quality cassette tape recorder (Sony, model TCD5M) for later presentation. The interstimulus interval was set at 2 s, with a 5-s pause between different random orderings of the 25 stimuli.

C. Procedures

Each subject was tested individually in a quiet room. The stimuli were presented binaurally over headphones (TDH-49) at a comfortable listening level of 80 dB SPL(A). The subjects' task was to identify the voicing feature of the final fricatives by circling "peas" or "peace" on an answer sheet. They were told to guess in case of uncertainty. All of the subjects were given a pretest consisting of ten randomized presentations each of a single token of *peace* and

peas produced by a native speaker of American English. No subject required more than two attempts to correctly identify the voicing feature of the final fricative 19 out of 20 times. The experimenter closely monitored subjects during the experiment to ensure that they responded to each stimulus as it was presented.

D. Analysis

An ANOVA in which vowel and fricative duration served as repeated measures (five levels each) was performed to test the effects of temporal changes on voicing judgments.⁵ A total of seven planned orthogonal comparisons between subject groups were specified to determine the effect of linguistic experience on subjects' use of vowel and fricative duration. The orthogonal contrasts compared the native French and English subjects; the native Swedish and Finnish subjects; native speakers of quantity and nonquantity languages (i.e., the Swedish-Finnish versus English-French subjects); the native English speakers from Chicago and Alabama; and the experienced versus inexperienced speakers of French, Swedish, and Finnish.

E. Speech production data

Speech perception differences between subjects differing in native language background are often interpretable in terms of differences in speech production. It was impossible to compare the English /s/-/z/ contrast to a corresponding contrast in Swedish and Finnish since those languages lack a /z/ phoneme. However, since French possesses an /s/-/z/ contrast which has not been directly compared to the corresponding English contrast, we gathered a small amount of production data for those two languages. Four monolingual speakers of French (in Paris) and four monolingual speakers of American English (two from the South, two from the Midwest) were asked to read the following five French or English word pairs:

	French				
/z/	vase,	base,	chose,	mise,	biz
/s/	face,	basse,	chausse,	fil,	bis
	English				
/z/	saws,	laws,	(to) close,	seize,	peas
/s/	sauce,	loss,	close,	cease,	peace.

Written lists containing the test words (along with foils) were recorded twice (Sony model TCD5M) in sound-treated rooms. The duration of vowel and fricative intervals was measured spectrographically (Kay model 7800) to the nearest 0.5 mm (2.02 ms) in two replicate productions of each word, along with the duration of voicing in the final fricative. Final "fricative" duration was measured from the beginning to the end of aperiodic energy in the frequency range 4-8 kHz. Fricative voicing was measured from the beginning of the fricative interval to the end of a low-frequency periodic energy component (or to the end of the fricative interval, in those instances where the entire fricative interval showed periodicity). "Vowel" duration was measured from the beginning of periodic energy in the second

formant to the beginning of the final fricative interval. (For vowels preceded by /l/ and /m/, vowel onset was further defined by a sudden increase in amplitude.)

The mean values of these intervals for words ending in /s/ and /z/ were calculated for each subject and submitted to an ANOVA in which language was the grouping variable and voicing was a repeated measure. A mean value for each of these intervals was also calculated for each word (four talkers \times two replicate productions). These means were submitted to a language \times voicing ANOVA.

II. RESULTS

A. Identification data

The mean percentages of /z/ (i.e., "peas") judgments are summarized in Table II for subjects in the eight groups as a function of vowel and fricative durations taken separately. The mean identification data plotted in Fig. 1, which have been collapsed across experienced and inexperienced subjects for each of the four native language backgrounds (English, French, Swedish, Finnish), show the results as a joint function of vowel and fricative duration. The number of observations per data point in Fig. 1 varies from a minimum of 210 for the native Finnish subjects (21 subjects \times 10 judgments) to a maximum of 300 for the native Swedish subjects (30 subjects \times 10 judgments).⁶

TABLE II. Mean percent /z/ judgments by subjects in eight subject groups as a function of vowel duration (upper panel) or fricative duration (lower panel). Each mean value represents responses to five stimuli with equal fricative duration (upper panel) or vowel duration (bottom panel).

Subject group	Vowel duration				
	150	200	250	300	350
English	1.9	11.0	39.6	71.0	81.6
Birmingham	1.4	11.4	39.7	73.5	84.6
Chicago	2.3	10.7	39.5	68.8	78.9
French	6.9	17.0	39.8	66.6	75.4
inexperienced	7.3	16.5	40.9	65.7	76.1
experienced	6.2	17.6	38.0	67.8	74.7
Swedish	5.7	17.3	46.7	78.3	87.0
inexperienced	8.1	17.5	51.7	80.0	89.7
experienced	3.3	17.2	41.7	76.7	84.3
Finnish	9.3	19.5	55.0	80.0	86.2
inexperienced	4.9	20.0	62.7	86.4	90.7
experienced	14.2	19.0	46.4	73.0	81.2
	Fricative duration				
	150	200	250	300	350
English	57.9	48.3	39.5	31.4	27.9
Birmingham	53.9	48.6	41.7	35.5	30.9
Chicago	61.5	48.1	37.6	27.7	25.2
French	54.6	46.9	37.8	34.8	31.5
inexperienced	56.3	46.0	38.7	33.9	31.9
experienced	52.2	48.2	36.4	36.2	31.0
Swedish	47.5	50.5	47.1	45.6	44.5
inexperienced	49.3	52.3	48.9	48.0	48.5
experienced	45.6	48.7	45.2	43.2	40.1
Finnish	49.9	51.9	49.0	51.4	47.8
inexperienced	53.8	55.8	51.1	53.6	50.4
experienced	45.6	47.6	46.6	49.0	45.0

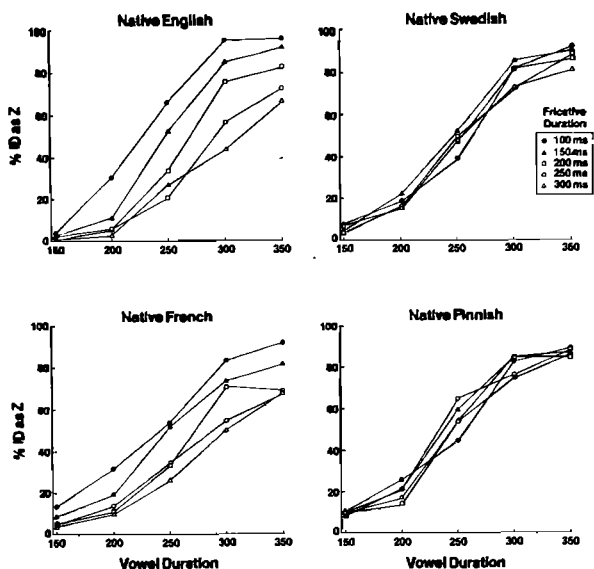


FIG. 1. Mean percent /z/ identification by subjects differing in native-language background. Changes in vowel duration (150, 200, 250, 300, 350 ms) are represented along the abscissa; changes in fricative duration (100, 150, 200, 250, 300 ms) are represented by grouping responses to all stimuli with the same fricative duration in a separate function. The 28 English, 25 French, 30 Swedish, and 21 Finnish subjects judged all 25 members of the *peas-peace* continuum ten times each.

Inspection of Fig. 1 reveals that changes in vowel duration had a substantially greater effect on voicing judgments than comparable changes in fricative duration. Averaging across all 104 subjects, the mean percentage of /z/ responses increased about 77% (from 6% to 83%) as vowel duration was lengthened by 200 ms; /z/ judgments increased by only 15% (from 38% to 53%) as fricative duration was shortened by the same amount. The main effects of both vowel duration [$F(4,384) = 837.17, p < 0.001$] and fricative duration [$F(4,384) = 49.63, p < 0.001$] were significant.

Lengthening vowel duration had a somewhat different effect on voicing judgments by subjects differing in linguistic experience. Vowel duration exerted a significantly larger effect on the voicing judgments of Finns who had never lived outside Finland than Finns who had lived in an English-speaking country. This led to a significant interaction between the factor vowel duration and the orthogonal contrast between experienced versus inexperienced Finns [$F(4,384) = 3.93, p < 0.005$]. The interaction between vowel duration and the orthogonal contrast between English-French and Swedish-Finnish subjects nearly reached significance [$F(4,384) = 2.04, p < 0.10$]. This appeared to be due to the relatively small effect of vowel duration on the French subjects' identification responses. Lengthening vowel duration increased the French subjects' /z/ responses less (69%) than it increased percent /z/ responses by the English, Swedish, and Finnish subjects (77%–81%). The interaction of vowel duration and the orthogonal contrast between French-English subjects reached significance [$F(4,384) = 2.71, p < 0.05$] but not the interaction between vowel duration and the orthogonal contrast between Swedish-Finnish subjects [$F(4,384) = 0.96$]. Neither of the orthogonal comparisons involving experienced versus inexper-

enced subjects reached significance ($p < 0.10$), indicating that learning English did not influence these non-native subjects' use of vowel duration.

Changing final fricative duration exerted a smaller and less consistent effect than changing vowel duration. The planned comparisons revealed two significant differences between subject groups. First, shortening fricative duration increased /z/ judgments by native English speakers from Chicago to a greater extent than it did for native English speakers from Birmingham. This cross-dialect difference led to a significant interaction between fricative duration and the orthogonal contrast between the Alabama and Chicago subjects [$F(4,384) = 3.16, p < 0.025$]. Second, shortening fricative duration increased /z/ judgments by the English and French subjects to a considerably greater extent than by Swedish and Finnish subjects. In fact, fricative duration exerted no systematic effect on the judgments of the Swedes and Finns, leading to a significant interaction between fricative duration and the orthogonal contrast between Swedish-Finnish and English-French subjects [$F(4,384) = 31.40, p < 0.001$]. None of the orthogonal comparisons involving experienced versus inexperienced subjects reached significance ($p < 0.10$), indicating that exposure to English did not influence the non-native subjects' use of fricative duration.

The significant vowel duration \times fricative duration interaction [$F(16,1536) = 13.29, p < 0.001$] seems to be due to the fact that, when the preceding vowel was very short (i.e., 150 ms in duration), variations in fricative duration did not influence voicing judgments. No three-way interaction involving an orthogonal contrast between experienced-inexperienced subjects and vowel duration \times fricative duration reached significance ($p < 0.10$).

B. Production data

The duration of vowel and fricative intervals, as well as the duration of voicing (i.e., glottal pulsing) during final fricatives produced by monolingual native speakers of French and English, is summarized in Table III. Talkers in both the French and English groups made vowels longer before /z/ than /s/; produced /s/ with a longer period of frication than /z/; and produced the /z/ closure interval with a longer period of voicing than the /s/ closure interval. Both the ANOVA based on talker means and word means indicated that these effects of voicing were significant ($p < 0.01$).

The native English speakers made vowels considerably longer (by 50% or 117 ms) than the French talkers. Both the ANOVA based on talker means and the ANOVA based on word means revealed this difference to be significant [$F(1,6) = 9.4, F(1,8) = 33.6, p < 0.05$]. The native English talkers produced a contrast that was larger on average (50% or 139 ms) than that of the native French talkers (43% or 82 ms). Figure 2, which plots the magnitude of the voicing effect on vowel duration in words produced by the native French and English talkers, shows that this was mainly due to two of the four English talkers. As a result, the ANOVA based on word means [$F(1,8) = 13.1, p < 0.01$] but not talker means [$F(1,8) = 3.03$] yielded a significant interaction

TABLE III. The mean duration in ms of intervals measured in French or English words ending in /s/ or /z/. Each mean value is based on 40 observations (five words \times two replicate productions \times four talkers). Standard deviations are in parentheses.

	English talkers				
	1	2	3	4	\bar{x}
Vowel /s/	263(34)	215(56)	373(33)	266(62)	279
Duration /z/	364(41)	407(72)	451(40)	452(16)	418
Fricative /s/	212(22)	307(25)	209(28)	288(39)	254
Duration /z/	148(30)	208(28)	166(12)	173(24)	174
Voicing /s/	2 (7)	10(12)	17(10)	1 (3)	7
Duration /z/	30(15)	71(29)	36(16)	10 (6)	36

	French talkers				
	1	2	3	4	\bar{x}
Vowel /s/	163(58)	227(50)	119(50)	250(39)	190
Duration /z/	233(20)	273(40)	219(20)	363(37)	272
Fricative /s/	304(16)	271(29)	445(23)	175(28)	299
Duration /z/	167(21)	110(13)	186(36)	88(11)	138
Voicing /s/	14(15)	14(14)	21(18)	15 (6)	16
Duration /z/	167(21)	85(22)	186(36)	29(11)	117

of language \times voice.

There was little difference between the French and English talkers in the average duration of word-final fricatives (219 vs 214 ms). However, whereas the French talkers made /s/ 45 ms longer than the English talkers, they made /z/ 36 ms shorter than the English talkers. This meant that the French talkers produced a greater temporal difference between /s/-/z/ (117% or 160 ms) than the English talkers (46% or 80 ms). The interaction of language \times voicing reached significance in the ANOVA based on word means [$F(1,8) = 44.5, p < 0.001$], but missed reaching significance in the ANOVA based on talker means [$F(1,6) = 4.19, p < 0.10$].

Finally, the French talkers produced a larger distinction in voicing between /s/-/z/ than the English talkers. The talkers in both groups produced /s/ with very little voicing. However, whereas the French talkers produced /z/ with a voicing duration of 117 ms, the native English talkers produced /z/ with voicing that averaged only 36 ms. The interaction of language \times voicing reached significance in the ANOVA based on word means [$F(1,8) = 40.6, p < 0.001$], but missed reaching significance in the ANOVA based on talker means [$F(1,6) = 3.65, p < 0.10$]. This difference in voicing duration may be attributed, at least in part, to a phonological difference between French and English (see Casagrande, 1984, and footnote 2). The French, but not the English, talkers inserted a schwa-like vowel after /z/ (but not after /s/) in every instance. This could be expected to preserve voicing in /z/, as well as shorten its period of frication.

III. DISCUSSION

This study examined the extent to which native speakers of French, Swedish, and Finnish used vowel and fricative duration as perceptual cues to the English /s/-/z/ contrast when those two parameters were varied factorially in synthetic CVC stimuli. Two groups of native English subjects differing in dialect showed a significant increase in the percentage of /z/ judgments as vowel duration was lengthened;

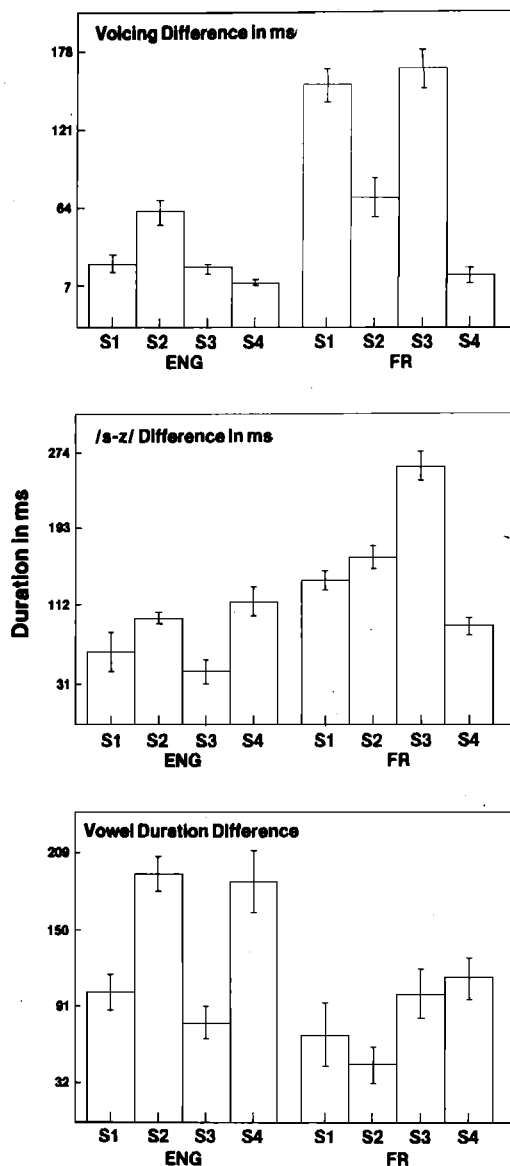


FIG. 2. Mean temporal contrasts in ms produced by four monolingual native speakers of French and English. The upper panel plots the duration by which voicing in /z/ exceeded the voicing in /s/; the middle panel plots the duration by which /s/ frication exceeded the fricative noise associated with /z/; the bottom panel plots the duration by which vowels preceding /z/ exceeded vowels preceding /s/ in minimally (or near minimally) paired French or English words. The brackets enclose one standard error.

they also showed a significant increase in /z/ judgments as the duration of postvocalic fricative noise was shortened by the same amount.

The identification of fricatives as /s/ or /z/ by non-native English subjects who had lived in an English-speaking environment for a year or more was compared to that of subjects who had not lived outside their homeland. The most important finding of the study was that exposure to English seemed to have little or no effect on the voicing judgments of the non-English subjects, at least as assessed by means of a two-alternative forced-choice test. The subjects in all six non-native groups differed in some respect from the native English subjects.

The Swedish and Finnish subjects, whose native language does not possess an /s/-/z/ contrast, differed from

the English (and French) subjects in showing no significant effect of fricative duration. Massive exposure to the English spoken by English native speakers apparently did not induce phonetic sensitivity to fricative duration as a cue to the English /s-/z/ contrast on the part of the Swedish and Finnish subjects. Only five of the 51 native Swedish and Finnish subjects showed as large a fricative duration effect as the average English and French subject (i.e., a 26% increase in /z/ judgments as fricative duration was shortened). This finding differs from the results of several other recent studies showing an effect of L2 experience on the perception of phonetic contrasts present in L2 but not L1 (Caramazza *et al.*, 1973; Williams, 1980; MacKain *et al.*, 1981; Werker and Tees, 1984a,b; Elsendoorn, 1984). In particular, the present results differ from the finding by Flege (1984a) that shortening fricative duration increased /z/ judgments by Saudi Arabians who had lived for 6 years in the U.S. but not by Saudis who had just arrived in the U.S.

We can think of several possible explanations for this apparent discrepancy. It is possible that since Swedish and Finnish do not possess a phonemic contrast between /s-/z/, Swedes and Finns do not immediately note the existence of the phonetic contrast between /s-/z/ and this, in turn, slows the rate at which they develop sensitivity to acoustic correlates of the /s-/z/ contrast (see footnote 3, and Flege and Hillenbrand, 1984). Fricative duration was varied separately in one continuum in the Flege (1984a) study, but factorially along with vowel duration in the present study. This may have made it easier for the experienced native Arabic subjects to focus on changes in fricative duration. It is possible that an effect of fricative duration would have been noted for at least the experienced Swedish and Finnish subjects in this study had only fricative duration been varied.

A final possible explanation is that the Swedish and Finnish subjects had not had sufficient exposure to English to induce sensitivity to fricative duration as an acoustic cue to the /s-/z/ contrast in English words. Werker and Tees (1984a,b) found that the overall amount of naturalistic L2 experience may crucially affect the ability of L2 learners to perceive non-native phonetic contrasts (see also MacKain *et al.*, 1981). In their study, native English speakers with 5 years' Hindi experience, but not those with only a year's experience, were able to discriminate a Hindi phonetic contrast not exploited in English. Even the experienced non-native subjects in the present study had lived a much shorter time in an English-speaking environment (about 2 years, on average) than the experienced Arab subjects in Flege's (1984a) study (about 6 years). Further, the Swedish and Finnish subjects were tested in their native country, and most had not lived in an English-speaking country for at least several years prior to the study. Perhaps the experienced Swedish and Finnish subjects examined here *did* develop sensitivity to fricative duration, but later lost the ability to use fricative duration as a voicing cue once they were no longer regularly exposed to the English /s-/z/ contrast.

There were two significant between-group differences pertaining to the effect of vowel duration. The French subjects, whose native language possesses a phonemic contrast between /s/ and /z/, showed a significantly smaller increase

in /z/ judgments as vowel duration was lengthened than the English subjects (as well as the Swedish and Finnish subjects). This may have occurred because the French /s-/z/ contrast differs from the analogous contrast in English: Previous reports indicated that /s/ and /z/ are distinguished in the final position of French words by means of vowel and fricative duration (Kohler, 1979, 1984; Kohler *et al.*, 1981; O'Shaughnessy, 1981). This was confirmed by the production data reported here for French and English monolinguals. Our production data also suggested that French talkers may produce a larger temporal contrast in fricative noise and closure voicing between /s-/z/ than native English speakers, but a *smaller* contrast in preceding vowel duration (cf. Chen, 1970; Mack, 1982). These preliminary production results, which need to be replicated in a larger production study with more monolingual talkers, suggest the possibility that French listeners focus attention more on acoustic cues found near the end of syllables ending in /s/ and /z/ (i.e., the duration of voicing and fricative noise) than English listeners, whereas English listeners focus more on vowel duration than French listeners (see Flege and Hillenbrand, 1985).

Even though they do not have an /s-/z/ contrast in their native language, the Swedish and Finnish subjects showed as large an effect of vowel duration as the English subjects. It is possible that they "reinterpreted" (Weinreich, 1953; Flege, 1985) the role of vowel duration, that is, processed vowel duration as if it were a cue to a phonemic length distinction in their native language. The experienced Finnish subjects showed a significantly smaller effect of vowel duration than the inexperienced Finns. We speculate that this occurred as a result of their exposure to phonetic contrasts found in English. The results of L2 speech production studies (see Flege, 1985, for a review) suggest that native speakers of quantity languages reinterpret the distinction between tense and lax English vowels (e.g., /i-/I/) as a phonemic length contrast. Garnes (1977) found that Icelandic subjects who had lived in the U.S. made less consistent use of vowel duration as a cue to the contrast between phonemically short and long Icelandic vowels than those who had not lived outside Iceland (see also Gottfried and Beddor, 1984). If the experienced Finns reinterpreted the vowel duration correlate of the English /s-/z/ contrast as a *phonemic length* contrast, their exposure to the frequent neutralization of the phonetic contrast between /s/ and /z/ in running English speech (see, e.g., Umeda, 1975; Crystal and House, 1982) may have reduced their sensitivity to vowel duration as a perceptual cue (both to phonemic length contrasts, in L1, and to the voiced-voiceless distinction in English).

The experienced Swedish subjects may not have shown a similar reduction in the effect of vowel duration vis-à-vis the inexperienced Swedish subjects because their experience with English was less recent. Unlike any of the Swedish subjects, three of the experienced Finnish subjects had just returned from a year in the U.S. They showed a much smaller increase in /z/ judgments (about 50%) as vowel duration was lengthened than the remaining experienced Finnish subjects whose English L2 experience was less recent (an average of about 75%).

One question raised by this study is why the native Swedish and Finnish subjects used vowel duration rather than fricative duration in identifying stimuli, despite the fact that both stimulus dimensions were varied in five 50-ms steps, and despite the fact that phonemic categories are distinguished on the basis of both vowel and fricative duration in their native languages. Debriefing after the experiment suggested that the Swedish and Finnish subjects were more aware of changes in vowel than fricative duration. Most of these subjects reported hearing changes in vowel duration, but few of them reported hearing the changes in fricative duration.

The greater effect of vowel than fricative duration is consistent with the results of several previous studies which have suggested that listeners are more sensitive to changes in vowel than fricative duration (Soli, 1982; Carlson and Granström, 1975; Johansson, 1975; Garnes, 1977). It is also possible that the subjects focused more attention on vowel than fricative duration changes because the vowel portion preceded the fricative portion in CVC stimuli used in the present study. It may be difficult, in general, for L2 learners to focus attention on multiple correlates of an unfamiliar L2 phonetic contrast, especially if one is more salient than the other(s). Focusing attention on one stimulus dimension probably simplifies an otherwise difficult classification task.

We are left with the intriguing general question of whether L2 learners can learn to integrate multiple acoustic cues to a phonetic contrast which exists in L2 but not L1. The present study does not, of course, demonstrate that L2 learners are *incapable* of doing so. For example, our subjects might have more closely resembled the native English subjects had natural speech or more natural-sounding synthetic stimuli been used (see Miyawaki *et al.*, 1975; Mochizuki, 1981; Shimizu and Dantsuji, 1983), or had a more sensitive testing paradigm been employed (see, e.g., Werker and Tees, 1984a,b).

Kohler and Kunzel (1978) hypothesized that perception of inversely related temporal cues to the voicing feature in postvocalic stops derives from a single, centrally programmed pattern of articulation. Previous studies (Suomi, 1976; Niemi *et al.*, 1981; Flege and Port, 1981; Lauttamus, 1981) have shown that native speakers of quantity languages typically do not produce the same voicing-related differences in vowel and/or closure duration as English native speakers when producing word-final obstruents. This suggests that listeners might be unable to perceptually integrate multiple acoustic cues to a phonetic contrast if they themselves cannot control the articulatory gestures which yield those cues.

Others have hypothesized that listeners possess an *innate* ability to relate the multiple acoustic cues associated with a phonetic distinction. Mann (1984) examined the labeling of a /da-/ga/ continuum by two groups of native Japanese subjects who were either able or unable to correctly identify English /r/ and /l/. Both Japanese groups resembled native English speakers in giving more /ga/ responses when stimuli were preceded by /l/ than /r/. Since previous research had shown that this context effect occurs in a phonetic but not auditory mode of perception, Mann concluded

that the Japanese subjects had tacit knowledge of the acoustic consequences of producing /r/ and /l/, even though some of them could not identify /r/ and /l/.

The present results suggest either that listeners do not integrate multiple acoustic cues to phonetic category distinctions through an innate knowledge of vocal tract properties, or that such an ability does not extend to the temporal (as opposed to the spectral) domain. It seems to us that temporal correlates of the English /s/-/z/ contrast must be largely, if not entirely, learned. Although both vowel duration and fricative duration fluctuate over wide ranges in running speech, native speakers are likely to manifest perceptual knowledge of optimal (or "prototypical") temporal values (see, e.g., Oden and Massaro, 1978; Elsendoorn, 1984). It is possible that the ability to perceptually integrate *multiple* temporal cues (see Repp, 1982) depends on exposure at an early age to phonetic categories distinguished by those multiple cues (see Flege, 1985, for a review).

In summary, the present results in combination with previous related studies suggest that the perception of phonetic contrasts which exist in a foreign language but not the native language, may develop slowly, at least for adult learners. The extent to which the perception of speech evolves is likely to be influenced by many factors including age, the amount and kind of L2 experience, the nature and extent of phonological differences between L1 and L2, and especially the acoustic nature of the phonetic contrast being learned. Further research is clearly needed to address the specific and general questions raised by this study regarding the development of simultaneous use of multiple acoustic cues in the perception of L2 phonetic contrasts. The perception of native Swedish and Finnish subjects with somewhat more prolonged—and recent—exposure to English than subjects tested in the present study, as well as individuals chosen on the basis of English pronunciation proficiency, should be examined. Data from these populations may help to resolve the question of whether adults can establish temporal trading relations in the perception of L2 phonetic contrasts.

ACKNOWLEDGMENTS

This study was supported, in part, by NIH grants NS20572 and NS20963, and by a Travel Grant from Dean James Pittman of the UAB School of Medicine. The authors would like to thank M. Eskenazi (France), S. Elliason, S. Hunnicutt, T. Janson, J. Lubker (Sweden); J. Niemi, S. Niemi, S.-L. Savy, and M. O'Dell (Finland), and W. Smith (The United States) for their invaluable help recruiting subjects. The authors would also like to thank S. Fletcher, I. Lehiste, V. van Heuven, J. Niemi, S. Niemi, and T. Janson for commenting on a previous draft of this report, and K. Kirk for statistical consultation.

¹ Port (1982) hypothesized that native English-speaking listeners use the ratio of consonantal constriction to preceding vowel duration as a temporal cue to the voicing feature in postvocalic obstruents. Such a unitary cue could conceivably be linked to the articulatory difference distinguishing voiced and voiceless consonants (see Kohler and Kunzel, 1978), but Massaro (Derr and Massaro, 1980; Massaro and Cohen, 1983) has presented

evidence suggesting that native English-speaking listeners assess consonant closure and preceding vowel duration independently. In this view, the perception of voicing is based on the perceived duration of the two intervals (along with other cues) after they have been weighted against centrally stored "prototypical" values established on the basis of linguistic experience.

² French, Swedish, and Finnish differ from one another (and English) in a number of important ways. For example, the kinds of syllables permitted in these languages differ. About 80% of the syllables which occur in spoken French are open syllables, compared to only about 40% in English (see Gottfried, 1984). Finnish has few one-syllable (CVC or CVVC) words, for the canonical word form in Finnish is disyllabic. There are only about 20 monosyllabic Finnish words which end in a consonant, most of which are function words (e.g., *siis*, "therefore"). The canonical word form in Swedish, on the other hand, is monosyllabic; many words end in consonants, including fricatives.

³ An important difference among the languages contrasted in this study pertains to the occurrence of phonological oppositions based on the voicing feature. Contrasts between voiced and voiceless stops and fricatives exist in English and French. Finnish can be analyzed as having no opposition between voiced and voiceless consonants, although a case can be made for the existence of a /d/. Finnish /d/ is restricted to intervocalic position, where it occurs only singly (unlike voiceless stops, which may be either long or short). The stop consonants /p,t,k/ (or their long cogeners /p:/, /t:/, and /k:/) generally manifest little voicing in any position within the word, whereas /d/ is ordinarily produced with voicing throughout the entire closure interval according to Suomi (1976, 1980). Swedish possesses a contrast between voiced and voiceless stops, but no /s/-/z/ contrast owing to the absence of a /z/ phoneme. As a result, the English /s/-/z/ distinction is especially difficult for Swedish learners of English, who are taught at school that English /z/ is a kind of /s/ produced with voicing.

⁴ There is often a vowel quality difference distinguishing pairs of phonemically long and short vowels in Swedish and Finnish. For example, the short vowel /i/ in Swedish *viss* resembles IPA [ɪ] or even [e], whereas the long vowel /i:/ in *vis* closely resembles IPA [i] in quality (T. Janson, 1984).

⁵ The same statistical analysis was performed after the data had been transformed using an arcsine transformation. Only the analysis based on untransformed data is presented here since both analyses yielded the same results.

⁶ The results presented in Fig. 1 and Table II do not include the responses of 12 non-native subjects who reversed the category labels used by 76 other non-native subjects and by all 28 native English subjects. Five native speakers of French (one experienced, four inexperienced), six native speakers of Finnish (three experienced, three inexperienced), and one native Swede (an experienced subject) were excluded. The fact that 12 of the non-native subjects identified stimuli with long vowels as "peace" (/s/) and stimuli with relatively short vowels as "peas" (/z/) is surprising since they correctly identified naturally produced tokens of *peas* and *peace* in the pre-test. The identification responses of the excluded Finnish and French subjects were analyzed in separate ANOVAs. Except for the category reversal, their responses closely resembled those of the Finnish and Swedish subjects who were included.

Bond, Z., and Adamescu, L. (1979). "Identification of novel phonetic segments by children, adolescents, and adults," *Phonetica* 36, 182-186.

Caramazza, A., Yeni-Komshian, G., Zurif, E., and Carbone, E. (1973). "The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals," *J. Phonet.* 54, 421-428.

Carlson, R., and Granström, B. (1975). "Perception of segmental duration," in *STL-QPSR* 1, 1-16.

Casagrande, J. (1984). *The Sound System of French* (Georgetown U. P., Washington, DC).

Chen, M. (1970). "Vowel length variation as a function of the voicing of the following consonant environment," *Phonetica* 22, 129-159.

Crystal, T., and House, A. (1982). "Segmental duration in connected speech: A preliminary analysis," *J. Acoust. Soc. Am.* 72, 705-716.

Denes, P. (1955). "Effect of duration on the perception of voicing," *J. Acoust. Soc. Am.* 27, 761-764.

Derr, M., and Massaro, D. (1980). "The contribution of vowel duration, F₀ contour and fricative duration to the /jus-juz/ distinction," *Percept. Psychophys.* 27, 51-59.

Eliasson, S. (1982). "Transfer as evidence for phonological solutions," *Stud. Anglica Posaniens.* 14, 185-196.

Elsendoorn, B. A. G. (1984). "Tolerances of durational properties in British English vowels," unpublished Ph.D. thesis (Department of Phonetics, University of Utrecht, Holland).

Flege, J. (1984a). "The effect of experience on Arabs' perception of the English /s/ vs. /z/ contrast," *Folia Linguist.* 18, 117-138.

Flege, J. (1984b). "The detection of French accent by American listeners," *J. Acoust. Soc. Am.* 76, 692-707.

Flege, J. (1985). "The production and perception of foreign language speech sounds," in *Human Communication and Its Disorders, Vol. 1*, edited by H. Winitz (Ablex, Norwood, NJ, in press).

Flege, J., and Hillenbrand, J. (1984). "Limits on pronunciation accuracy in adult L2 speech production," *J. Acoust. Soc. Am.* 76, 708-721.

Flege, J., and Hillenbrand, J. (1985). "Integration of cues to the /g-/k/ contrast by native speakers of English and French," to be published in *J. Acoust. Soc. Am.*

Flege, J., and Port, R. (1981). "Cross-language phonetic interference: Arabic to English," *Lang. Speech* 24, 125-146.

Garnes, S. (1977). "Effects of bilingualism on perception," in *Papers in Psycholinguistics and Sociolinguistics*, edited by O. Garnica (Dept. of Linguistics, The Ohio State University, Columbus, OH), pp. 1-10.

Gillette, S. (1980). "Contextual variation in the perception of R and L by Japanese and Korean speakers," *Minnesota Pap. Ling. Philos. Lang.* 6, 59-72.

Gottfried, T. (1984). "Effects of consonant context on the perception of French vowels," *J. Phonet.* 12, 91-114.

Gottfried, T., and Beddor, P. S. (1984). "Perception of temporal versus spectral information in French vowels," *J. Acoust. Soc. Am. Suppl.* 1 76, S80.

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

Gruenenfelder, T. (1979). "Fundamental frequency as a cue to postvocalic consonantal voicing: Some data from perception and production," in *Speech Communication Papers*, edited by J. Wolf and D. Klatt (Acoustical Soc. Am., New York), pp. 57-61.

Haggard, M. (1978). "The devoicing of voiced fricatives," *J. Phonet.* 6, 95-102.

Janson, T. (1984). Personal communication.

Johansson, S. (1975). *Papers in Contrastive Linguistics and Language. Lund Studies in English*, 50 (CWK Gleerup, Lund, Sweden).

Karno, E., and Porter, R. J. (1980). "Imitation of /s/ duration in VCVs," *J. Acoust. Soc. Am. Suppl.* 1 67, S38.

Klatt, D. (1980). "Software for a cascade/parallel formant synthesizer," *J. Acoust. Soc. Am.* 67, 971-995.

Kohler, K. J. (1979). "Parameters in the production and the perception of plosives in German and French," *Inst. of Phonetics, Univ. of Kiel, Arbeitsberichte* 12, 261-292.

Kohler, K. J. (1984). Personal communication.

Kohler, K. J., and Kunzel, H. J. (1978). "The temporal organization of closing-opening movements for sequences of vowels and plosives in German," *Inst. of Phonetics, Univ. of Kiel, Arbeitsberichte* 10, 117-167.

Kohler, K. J., van Dommelen, W., and Timmerman, G. (1981). "Die Merkmalpaare stimmhaft/stimmlos und fortis/lenis in der Konsonantenproduktion und Perception des heutigen Standard-Französisch," *Inst. of Phonetics, Univ. of Kiel, Arbeitsberichte*, 14.

Lauttamus, T. (1981). "Havaintoja suomen ja englannin sibilanttien kontekstuaalisesta variaatisosta," *Virittaja* 4, 340-354.

Lehiste, I. (1984). Personal communication.

Lehiste, I., and Shockey, L. (1980). "Labelling, discrimination and repetition of stimuli with level and changing fundamental frequency," *J. Phonet.* 8, 469-474.

Mack, M. (1982). "Voicing-dependent vowel duration in English and French: Monolingual and bilingual production," *J. Acoust. Soc. Am.* 71, 173-178.

MacKain, K., Best, C., and Strange, W. (1981). "Categorical perception of English /r/ and /l/ by Japanese bilinguals," *Appl. Psycholing.* 2, 369-390.

Mann, V. (1984). "Perception of [l] and [r] by native speakers of Japanese: A distinction between articulatory tracking and phonetic categorization," *Ann. Bull. Inst. Logoped. Phoniat.* 18, 127-134.

Massaro, D., and Cohen, M. (1983). "Consonant/vowel ratio: An improbable cue in speech," *Percept. Psychophys.* 33, 501-505.

McClasky, C., Pisoni, D., and Carrell, T. (1980). "Effects of transfer of training on identification of a new linguistic contrast in voicing," *Res. Speech Percept.* 6, 205-234 (Dept. of Psychology, Indiana Univ., Bloomington).

- Mitleb, F. (1982). "Voicing effect on vowel duration is not an absolute universal," *J. Acoust. Soc. Am. Suppl.* 1 71, S23.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., and Fujimura, O. (1975). "An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English," *Percept. Psychophys.* 18, 331-340.
- Mochizuki, M. (1981). "The identification of /r/ and /l/ in natural and synthesized speech," *J. Phonet.* 9, 283-303.
- Niemi, J., Hartikainen, P., Ikonen, U., and Inkinen, R. (1981). "Contextual constraints on timing in English: A study of sound change in progress," *Occas. Pap. Dept. Lang., Joensuu Univ.* 1, 52-64.
- Oden, G., and Massaro, D. (1978). "Integration of featural information in speech perception," *Percept. Psychophys.* 85, 172-191.
- Oller, D. K., and Eilers, R. (1983). "Speech identification in Spanish- and English-learning 2-year-olds," *J. Speech Hear. Res.* 26, 50-53.
- O'Shaughnessy, D. (1981). "A study of French vowel and consonant durations," *J. Phonet.* 9, 385-406.
- O'Shaughnessy, D. (1984). "A multispeaker analysis of durations in read French paragraphs," *J. Acoust. Soc. Am.* 76, 1664-1672.
- Peterson, G., and Lehiste, I. (1960). "Duration of syllable nuclei in English," *J. Acoust. Soc. Am.* 32, 693-703.
- Pisoni, D., Aslin, R., Perey, A., and Hennessy, B. (1982). "Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants," *J. Exp. Psychol.: Hum. Percept. Perform.* 8, 297-314.
- Port, R. (1982). "Consonant/vowel ratio as a cue for voicing in English," *Percept. Psychophys.* 32, 141-157.
- Raphael, L. (1971). "Vowel duration as a cue to the perceptual separation of cognate sounds in American English," *Haskins Lab. Stat. Rep. on Speech Res.*, August 1971.
- Redmond, L. D. (1977). "Learning to recognize foreign speech sounds: Strategies of auditory processing," unpublished Ph.D. thesis (University of California at San Diego).
- Repp, B. (1981). "Two strategies in fricative discrimination," *Percept. Psychophys.* 30, 217-227.
- Repp, B. (1982). "Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception," *Psychol. Bull.* 92, 81-110.
- Shimizu, K., and Dantsuji, M. (1983). "A study of the perception of /r/ and /l/ in natural and synthetic speech sounds," *Stud. Phonolog.* XVII, 1-14.
- Soli, S. (1982). "Structure and duration of vowels together specify fricative voicing," *J. Acoust. Soc. Am.* 72, 366-378.
- Strange, W., and Dittman, S. (1981). "Effects of discrimination training on the perception of /r-l/ by Japanese adults," paper presented at the 22nd meeting of the Psychonomic Society.
- Suomi, K. (1976). *English Voiceless and Voiced Stops as Produced by Native and Finnish Speakers, Jyväskylä Contrastive Studies*, 2 (Dept. of English, Univ. of Jyväskylä).
- Suomi, K. (1980). *Voicing in English and Finnish Stops* (Dept. Finnish and General Linguistics, Univ. of Turku, Finland), Publ. 10.
- Umeda, N. (1975). "Vowel duration in American English," *J. Acoust. Soc. Am.* 58, 434-455.
- Weinreich, U. (1953). *Languages in Contact, Findings and Problems* (Mouton, The Hague).
- Werker, J., Gilbert, J. H., Humphrey, I., and Tees, R. C. (1981). "Developmental aspects of cross-language speech perception," *Child Dev.* 52, 349-355.
- Werker, J. F., and Tees, R. C. (1984a). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," *Inf. Behav. Dev.* 7, 49-63.
- Werker, J. F., and Tees, R. C. (1984b). "Phonemic and phonetic factors in adult cross-language speech perception," *J. Acoust. Soc. Am.* 75, 1866-1878.
- Williams, L. (1977). "The voicing contrast in Spanish," *J. Phonet.* 5, 169-184.
- Williams, L. (1980). "Phonetic variation as a function of second-language learning," in *Child Phonology, Vol. 2, Perception*, edited by G. Yeni-Komshian, J. Kavanagh, and C. Ferguson (Academic, New York), pp. 185-216.