

Production and perception of English stops by native Spanish speakers

James Emil Flege

Department of Biocommunication, University of Alabama at Birmingham, Birmingham, AL, U.S.A.

and

Wieke Eefting

Institute of Phonetics, University of Utrecht, Utrecht, The Netherlands

Received 2nd April 1986, and in revised form 2nd September 1986

This study examined production of /b, d, g/ and /p, t, k/ in the initial position of English and Spanish words by two groups of native Spanish adults and native Spanish 9–10 year-olds who began learning English as a second language (L2) by the age of 5–6 years. It also examined identification by the same subjects of a VOT continuum ranging from /da/ to /ta/. The subjects in all three groups produced /p, t, k/ with significantly longer voice onset time (VOT) values in English than Spanish words, but with significantly shorter VOT values in English words than age-matched English monolinguals. The subjects also realized Spanish /p, t, k/ with significantly shorter VOT values than age-matched Spanish monolinguals. The mean /da/–/ta/ category boundaries of the two groups of adult subjects, but not the child subjects, occurred at significantly shorter mean values than those of age-matched monolingual speakers of English. The results are interpreted to mean that although the native Spanish speakers of English developed phonetic categories for English [p^h], [t^h], and [k^h], they did not realize /p, t, k/ authentically because much of their L2 input was likely to have been Spanish-accented English rather than the English spoken by English native speakers.

1. Introduction

The question addressed here is whether individuals who learn a second language (henceforth L2) establish phonetic categories for phones found in L2 but not the native language (L1). Weinreich (1953) illustrated this question for native Russian speakers of English exposed to voiceless unaspirated tokens of [t] in Russian words, and voiceless aspirated [t^h] tokens in English words. Weinreich (1953, p. 8) suggested that Russians are likely to realize /t/ as [t] rather than [t^h] in the initial position of English words because of the “natural identification” of these two phones. He hypothesized that Russian speakers of English are more likely to develop a “coexistent” phonetic system, in which [t] and [t^h] are produced in Russian and English words as realizations of a single

category, than to develop a "merged" system, in which /t/ and /t^h/ exist as separate categories in an enriched system¹.

Flege & Hillenbrand (1984) hypothesized that L2 phones which differ acoustically from easily identifiable counterparts in L1, but are nonetheless judged to be realizations of the same category as the L1 phones (called "similar phones"), cannot be produced authentically² by L2 learners because of equivalence classification. Flege (1986b) tested this hypothesis in a study which rested on the assumption that (for example) French speakers of English judge tokens of [t] in French words and [t^h] in English words to be realizations of a single category (i.e. {t}). Acoustic analyses showed that experienced native French speakers of English (and also English speakers of French) realized /t/ with significantly different VOT values L1 and L2. This suggested that they had noted at least some of the acoustic differences distinguishing the short-lag dental [t] used to realize /t/ in French from the long-lag alveolar [t^h] realizations of /t/ in English. However, the French subjects realized English /t/ with significantly shorter (French-like) VOT values than English monolinguals; and the native English subjects realized /t/ in French words with longer (English-like) VOT values than French monolinguals. Thus, this study replicated earlier studies showing that adults who learn English as an L2 realize English /p, t, k/ with significantly shorter VOT values than native English speakers if /p, t, k/ are implemented as short-lag stops in their L1 (Caramazza, Yeni-Komshian, Zurif & Carbone, 1973; see also Flege & Port, 1981; Port & Mitleb, 1983), and also showed that individuals who learn an L2 (e.g. French) in which /p, t, k/ are implemented as short-lag stops will realize those stops with VOT values that are too long if /p, t, k/ are implemented as long-lag stops in their L1.

The aim of this study was to further test the hypothesis (Flege, 1981, 1986b; Flege & Hillenbrand, 1984) that L2 learners cannot realize /p, t, k/ authentically in L2 if they are implemented with a different phonetic category in L1 and L2. The study focused on the production of English stops by native speakers of Spanish. Like French /p, t, k/, Spanish /p, t, k/ are implemented as short-lag stops in utterance-initial position (Lisker & Abramson, 1964; Williams, 1977a)³. To realize /p, t, k/

¹The terminology and symbols used here have been taken from Keating (1984). Phonemic categories (e.g. /t/) are said to be implemented by one of several possible universal phonetic categories (e.g. {t} or {t^h}). The phonetic category used to implement a more abstract phonemic category establishes the basic acoustic properties (such as modal VOT value) of a phone. In addition, language-specific detail (such as the absolute value, in ms, of the VOT interval) is provided by realization rules, which are shaped by learning and by universal phonetic principals. Keating's (1984) model seems to us to imply that substantial VOT differences in stop production can be accomplished by selecting different phonetic categories of stops (e.g. {t} versus {t^h}), but not by selecting between alternative phonetic realization rules. This interference will need to be tested empirically.

²Our use of the term "authentic" refers to statistically significant acoustic differences between subject groups. For example, a group of L2 learners is said to produce L2 phone "non-authentically" if they differ significantly from a group of native speakers in terms of measured VOT values. Such a difference does not, of course, in itself guarantee that the divergence of the L2 subjects from the phonetic norms of the L2 being learned is linguistically, socially, or psychologically relevant. These issues fall outside the scope of the present study. However, recent studies (Flege & Hammond, 1982; Flege, 1984) have suggested that small VOT differences, such as those that may distinguish native from non-native speakers of English, are auditorily detectable.

³Developmental studies indicate that English-learning children realize /b, d, g/ and /p, t, k/ with significantly different VOT values by about the age of three years, suggesting the establishment of separate phonetic categories. By the age of about 10 years monolingual English children realize English stops with values that closely resemble those of English adults (Kewley-Port & Preston, 1984; Zlatin & Koenigsknecht, 1976; Bailey & Haggard, 1980; Flege & Eefting, 1986). This suggests that they have established English-specific realization rules. It may take children learning Spanish longer to establish realization rules for /p, t, k/. Macken (1980) found that Spanish-learning children as old as four years did not produce a significant VOT difference between /p, t, k/ and /b, d, g/, perhaps because the VOT distinction between Spanish /b, d, g/ and /p, t, k/ is neutralized in many phonetic contexts (Harris, 1969).

authentically in English, a native speaker of Spanish must either: establish new long-lag phonetic categories with which to implement English /p, t, k/ (i.e. {p^h}, {t^h}, {k^h}); modify the realization rules used for outputting existing phonetic categories (i.e. {p}, {t}, {k}); or establish new realization rules to be used when speaking English.

Previous results suggest that Spanish learners of English will realize /p, t, k/ in English with VOT values that are significantly shorter than those of native English speakers if they establish "coexistent" phonetic systems, as hypothesized by Weinreich (1953). This hypothesis assumes, of course, that changes in VOT values effected through use of alternate realization rules will necessarily be smaller than the substantial VOT differences that distinguish realizations of /p, t, k/ in Spanish and English (see Keating, 1984, and footnote 1). The changes in VOT that can be effected through the selection of different phonetic categories are likely to be much greater. L2 learners might, therefore, realize English /p, t, k/ authentically if they establish new phonetic categories for English /p, t, k/ (and thereby develop a "merged" phonetic system).

The finding of previous studies that L2 learners were unable to realize English /p, t, k/ authentically may be due to the age at which the subjects examined began learning English. Fokes, Bond & Steinberg (1985) found that native Arabic children aged 2–11 years realized /p, t/ in English words with English-like VOT values (82 ms). Although the Arabic children's speech production was not compared to that of age-matched native English children, this suggests that they may have approximated the VOT norm of English more closely than adult L2 learners in previous studies. Williams (1979), on the other hand, found that 8–10 and 14–16 year-old native Spanish (Puerto Rican) children produced English /p/ with Spanish-like mean VOT values of about 40 and 20 ms, respectively. These values were probably shorter than would be observed for age-matched native English children, suggesting that even child L2 learners may not realize L2 stops authentically.

The findings cited above lead to the hypothesis that adults who learn English L2 as young children, but not those who learn English as an L2 later in life, will succeed in realizing English /p, t, k/ authentically. "Early L2 learners" would have the potential advantage of being exposed to tokens of English [p^h], [t^h], and [k^h] at a time when it may be relatively easy to establish new phonetic categories, and when they have many years to refine the phonetic realization rules needed to output those phonetic categories. Two recent studies support the "early L2 learning" hypothesis. Mack (1983) did not observe a significant VOT difference for /t/ between English monolinguals and adults who learned French and English as young children; and Williams (1977b) found no VOT difference for /p/ between English monolinguals and adult subjects who learned Spanish and English as young children⁴. Caramazza *et al.* (1973), on the other hand, found that French Canadians who began learning English in school "no later" than the age of seven years produced

⁴Mack's (1983) subjects began learning their L2 (English or French) at an average age of 4.5 years, and Williams' (1977b) subjects began learning their L2 (English or Spanish) "upon entering school . . . or earlier". The lack of a statistically significant difference between English monolinguals and Williams' (1977b) subjects may have been due to the subjects whose L1 was English. Williams' report did not include measures of individual or group means, nor measures of central tendency for the monolingual and bilingual groups. Moreover, VOT was measured in imitated rather than spontaneous speech. Monolingual children produce the same VOT values in both context (Bond & Korte, 1983), but imitating might affect the speech of adults who speak two languages differently than it affects monolingual children's.

/p, t, k/ in English with VOT values that were significantly shorter (51 ms) than English monolinguals (74 ms)⁵.

This study directly compared the production of English stops by native Spanish adults who first began learning English L2 as young children to that of adults monolingual native speakers of English. It also compared the production of English stops by native Spanish 8–9 year-olds who began learning English upon entering school (at the age of 5–6 years) to that of monolingual English children. Subjects in the three native Spanish groups should not differ significantly from age-matched native English subjects if children who learn an L2 establish phonetic categories for the long-lag stops of English, and if doing so results in authentic production of English /p, t, k/. However, if several years are needed to establish realization rules for new phonetic categories, the native Spanish adults might resemble age-matched native English subjects more closely than the native Spanish children.

It is generally believed that perception in some sense “leads” production in L1 speech acquisition. The extent to which L2 learners approximate the phonetic norms of L2 for stop consonant production might be limited by the perceptual representations established for L2 stops. A recent study (Flege & Eefting, 1986) employing the VOT continuum to be used here showed that the mean /d/–/t/ boundary of native English adults occurred at significantly longer values (43 ms) than that of native Spanish adults (23 ms); and the boundaries of native English children occurred at significantly longer values (29 ms) than native Spanish children’s (17 ms). Williams (1977*b*) found that adults who learned Spanish and English as young children manifested category boundaries at VOT values closely resembling those of either Spanish or English monolinguals. This supports the existence of a merged phonetic system. Williams (1979), on the other hand, found that the category boundaries of Puerto Rican children who began learning English in the U.S.A. between the ages of 8–16 years occurred at values that were generally intermediate to those observed for Spanish and English monolinguals, supporting the existence of coexistent phonetic systems.

The second aim of this study was, therefore, to determine whether the native Spanish subjects would have category boundaries that closely resembled native Spanish or English subjects, or whether their boundaries would occur at intermediate values. The category boundaries obtained for the Spanish speakers of English examined here were directly compared to those obtained earlier by Flege & Eefting (1986) for native speakers of Spanish and English. If early exposure to English permits native Spanish speakers to establish phonetic categories for long-lag English stops, we might expect native Spanish speakers of English to closely resemble age-matched monolingual English subjects in identifying the members of /da/–/ta/ continuum differing in VOT. However, if early L2 learning does not facilitate the formation of long-lag stop categories, the native Spanish speakers of English—adults as well as children—might be expected to show category boundaries at values that are intermediate to those obtained for native Spanish and English subjects who did not speak an L2.

⁵The native French subjects examined by Caramazza *et al.* (1973) may have differed from Mack’s (1983) and Williams’ (1977*b*) subjects in beginning to learn English L2 somewhat later in life (see McLaughlin, 1978, for a discussion of early childhood bilingualism). It is also possible that, for psychosocial reasons, they desired to speak English with a “French accent”. If so, the results obtained may not have reflected their optimal pronunciation of English stops.

2. Methods

2.1. Subjects

Seven groups of subjects differing in age and/or linguistic experience were formed, each consisting of five male and five female subjects with normal hearing and speech (according to self-report or that of their teacher).

The native Spanish children in the group designated BC ("bilingual children") were 8–9 year-olds born in Puerto Rico of native Spanish parents who had attended a private English-speaking school in Mayaguez for an average of 3.6 years. The language background questionnaire indicated that these children had received little if any exposure to English prior to the age of 5–6 years. They spoke English without an obvious Spanish accent in the authors' opinion.

The adult subjects in the group designated "LCB" ("later childhood bilinguals") were born and raised in Puerto Rico of native Spanish parents and had never lived in an English-speaking environment. They began learning English at the age of 5–6 years upon entering a private elementary school like the one attended by subjects in group BC, where they were enrolled for an average of 7.1 years. All but one of them later attended a Spanish-speaking public high school, and all were enrolled in a Spanish-speaking university at the time of the study.

The adult subjects in the group designated "ECB" ("earlier childhood bilinguals") had the same mean age (19 years) as those in group LCB and, like those in LCB, were attending a Spanish-speaking university at the time of the study. They differed from the subjects in LCB in several respects. They were born on the mainland U.S.A. (where English is the predominant language) or had been taken there shortly after birth. The subjects in ECB had lived for an average of 9.7 years in the U.S.A. and were enrolled for 6.4 years in an English-speaking elementary school there. All but one later attended a Spanish-speaking public high school in Puerto Rico. Although the subjects in ECB indicated that Spanish was the language usually spoken in their home, one parent of two subjects was a native speaker of English. Perhaps because of their somewhat earlier, and probably more massive, exposure to English, three subjects in ECB regarded English as their first language, four considered Spanish to be their first language, and three were unsure.

The Spanish speakers of English were compared to age-matched monolingual English subjects. The subjects in the group designated "EA" ("English adults") were monolingual English speakers with a mean age of 26 years who were affiliated with the University of Alabama at Birmingham. Subjects in group "EC" ("English children") were 9–10 year-olds enrolled in a fourth-grade class at a parochial school in Birmingham, Alabama. None of the subjects in either native English group had been exposed to any language other than English.

The Spanish speakers of English were also compared to age-matched native Spanish speakers who had not learned English as an L2. The subjects in these two groups were born in Puerto Rico and had never lived in an English-speaking environment. Subjects in group "SA" ("Spanish adults") were engineering students at the University of Mayaguez with a mean age of 20 years; those in "SC" ("Spanish children") group were fourth graders enrolled in a public school in Mayaguez, Puerto Rico with a median age of nine years.

The children in group SC had received only one year of English-language instruction (nominally one hour per day) from a non-native speaker of English at the time of the

study. It appears they had received little if any exposure to English or any other foreign language. (TV and radio programs are broadcast exclusively in Spanish in Mayaguez). It would probably be unreasonable to refer to the adult subjects in group SA as "monolingual" from a linguistic, although not phonetic, perspective. They had all studied English as a foreign language in school for 12 years as required by law, but their exposure to English stops seems to have been limited. None of them had pursued their study of English beyond the official language requirement; their formal instruction in English focused on grammar and reading, and came from native Spanish teachers likely to have produced English words with Spanish-like stops. Most importantly, one of us who does not speak Spanish (JEF) was unable to engage these subjects in a simple conversation in English prior to the experiment.

2.2. Stimuli

The production of isolated English and Spanish words was elicited using written lists consisting of eight randomizations of the following stop-initial words with /a/:

English: parking/barking target/darling carver/garner

Spanish: pato/bato tato/dato cato/gato

The VOT continuum used in the identification experiment consisted of consonant-vowel syllables with VOT values ranging from -60 to $+90$ ms in 10-ms steps. The continuum thus provided exemplars of the three modal VOT categories used to implement stops in Spanish and English: lead, short-lag, and long-lag. The stimuli were synthesized using a modified version of Klatt's (1980) parallel/cascade software synthesizer. The parameters selected yielded endpoints closely resembling careful isolated productions of "da" and "ta" by the first author (a native speaker of American English). The spectrum and intensity of the 10-ms release burst occurring in all 16 stimuli represented an average of values observed in multiple tokens of /da/ and /ta/. The F_1 – F_3 formant transitions were linearly interpolated from values appropriate for an alveolar stop ($F_1 = 206$, $F_2 = 1545$, $F_3 = 2600$) to values appropriate for /a/ ($F_1 = 700$, $F_2 = 1220$, $F_3 = 2600$) over 100 ms⁶. F_4 and F_5 frequencies were held constant at 3300 and 3750 Hz. Aspiration between the release burst and voicing onset was simulated by setting the aspiration source (AH) to 66 dB and widening the F_1 – F_3 bandwidths. Amplitude of the voicing source (AV) was interpolated from 0–60 dB over the first 15 ms of the "vowel", then held constant at 60 dB near the end of the stimuli. F_0 was set at 105 Hz up to the release burst, at 133 Hz for the next 85 ms, then interpolated down to 115 Hz over 320 ms. Creaky voice was simulated in the final 75 ms of the stimuli by decreasing F_0 from 60–50 Hz, and the amplitude of the voicing source from 60–50 dB.

Prevoicing was simulated by setting the amplitude of the sinusoidal voicing source (AVS) to 56 dB, F_1 frequency to 180 Hz, F_1 amplitude to 30 dB, and the F_1 bandwidth to 180 Hz. VOT values in the lead range varied according to when, relative to the release burst, the voicing source was first increased above 0 dB. VOT in the short- and long-lag ranges varied according to when amplitude of the aspiration source was decreased to 0 dB, and the voicing source amplitude increased above

⁶Lisker, Liberman, Erickson, Dechovitz & Mandler (1977) reported that 115 ms represents the upper limit on the duration of formant transition in acceptable tokens of synthetic /da/ and /ta/.

0 dB⁷. The F_1 cutback seen in English /t/ was simulated by setting the amplitude of F_1 to 0 dB during the aspiration interval.

2.3. Procedures

Information pertaining to language background was obtained prior to data collection. The subjects' speech was recorded on a high-quality cassette tape recorder (Technics Model RS-M235) with an omnidirectional dynamic microphone (Shure Model 578). The experimenter (WE, a native speaker of Dutch) cued production by moving an index card down a written list at a rate of about 24 words per minute. The Spanish speakers of English produced both the Spanish and English speech material. Half of them answered the language-background questions and produced the first word list in Spanish, and the other half in English, to neutralize the possible effect of language "set" (Elman, Diehl & Buchwald, 1977).

The speech perception stimuli were recorded on a professional-quality portable cassette tape recorder (Sony Model TCD5M) for binaural presentation over headphones (TDH-49) at a comfortable level. The 16 stimuli were presented in ascending order of VOT in a practice block for familiarization prior to data collection. The inter-stimulus interval of 2.0 s provided ample time for all subjects to circle "ta" or "da" on the answer sheet. A 10-s pause was placed between the 10 separate randomizations of the 16 stimuli. The subjects were told to guess in case of uncertainty. The instructions were given in English if the language used for collecting language background information was Spanish, and vice-versa.

2.4. Analyses

Spectrograms were made (Kay Model 7800) of 30 Spanish or English words spoken by the native Spanish subjects in groups SA and SC, and by the native English subjects in groups EC and EA. Twice that number (two languages \times six stops \times five tokens) were made for the 30 native Spanish speakers of English. One of us (WE) segmented all VOT intervals to the nearest 0.5 mm (2.0 ms) from the beginning of the release burst to the onset of periodicity in F_2 and higher formants. When /b, d, g/ were implemented as short-lag stops, VOT was measured the same way. When these stops were realized with lead VOT values, the VOT interval was measured from the beginning of low-frequency periodicity to the beginning of the release burst. (According to convention, prevoiced stops were assigned negative VOT values). Reliability of the VOT measurements was estimated by remeasuring three stops produced by five randomly selected adult and child subjects from duplicate spectrograms. The mean difference between the two sets of measurements averaged 0.27 mm (1.1 ms), with a range of 0.0–0.5 mm (0.0–2.0 ms).

Mean VOT values were calculated from five tokens of /p/, /t/, and /k/ spoken by each subject. These values were submitted to mixed-design ANOVAs to determine the effect of place of articulation and subject group. The mean VOT values for prevoiced realizations of Spanish /b/, /d/, and /g/ were also analyzed in a Group \times Place ANOVA. The

⁷The duration of VOT in the 16 stimuli was measured from spectrograms (Kay Model 7800). The difference between nominal and actual VOT values averaged 1.4 ms ($SD = 1.1$), which is less than the temporal resolution (2.0 ms) of the measurement technique employed. Acoustic measurements revealed that the prevoicing was 21.7 dB less intense than peak intensity of the "vowel", the release burst was 32.4 dB less intense, and the aspiration interval was 22.6 dB less intense.

percentage of times /b, d, g/ were realized with short-lag VOT values in Spanish and English was computed for each subject and submitted to χ -square analyses.

The functions obtained in the identification experiment showed rapid cross-over from consistent ($\geq 90\%$) /ta/ responses to consistent /da/ responses at some point along the VOT continuum. A category boundary was computed for each subject by linear interpolation of the 50% cross-over point in each subject's function. These values were submitted to a one-way randomized block ANOVA examining the effect of group.

3. Results

3.1. Speech production

3.1.1. Spanish stops

Table I presents the mean VOT in stops initiating Spanish words as a function of place of articulation for the 20 subjects who did not speak English and the 30 who did. Spanish /b/ was realized with a mean value of -81 ms, /d/ with a mean value of -82 ms, and /g/ with a mean of -75 ms. The ANOVA showed that the main effect of Place was significant [$F(2,90) = 3.18$; $p < 0.05$]. Post-hoc tests showed that /b/ and /d/ differed significantly from /g/ ($p < 0.05$). This finding agrees with results obtained for native English adults by Smith (1978), who interpreted it to be the result of greater difficulty maintaining the transglottal pressure difference needed for voicing in /g/ than in /b/ or /d/ owing to the relatively small cavity volume behind /g/. The effect of Group was

TABLE I. The mean VOT in Spanish word-initial stops produced by native Spanish children (SC) and adults (SA) who did not speak English, native Spanish children who did speak English (BC), and two groups of Spanish adults who spoke English (ECB, LCB), in ms. SD is the standard deviation; n is the number of tokens analyzed. Separate mean values were calculated for lead and short-lag realizations of /b, d, g/

Group		Lead VOT			Short-lag VOT					
		/b/	/d/	/g/	/b/	/d/	/g/	/p/	/t/	/k/
SC	\bar{x}	-81	-82	-89	7	14	26	16	17	27
	SD	(35)	(41)	(20)	(12)	(6)	(10)	(5)	(4)	(5)
	n	46	44	42	4	6	8	50	50	50
SA	\bar{x}	-82	-80	-68	-	16	34	18	22	39
	SD	(23)	(29)	(22)	-	(0)	(10)	(5)	(6)	(7)
	n	50	48	46	0	2	4	50	50	50
BC	\bar{x}	-91	-87	-76	16	12	19	7	10	18
	SD	(34)	(24)	(23)	(12)	(3)	(3)	(4)	(5)	(4)
	n	47	46	45	3	4	5	50	50	50
ECB	\bar{x}	-74	-75	-64	8	14	25	17	19	31
	SD	(9)	(17)	(18)	(0)	(3)	(3)	(7)	(5)	(6)
	n	49	45	44	1	5	6	50	50	50
LCB	\bar{x}	-80	-90	-82	16	16	24	11	15	27
	SD	(25)	(25)	(26)	(0)	(0)	(11)	(4)	(4)	(4)
	n	49	49	46	1	1	4	50	50	50

non-significant, which suggests that learning English did not alter the Spanish speakers' realizations of Spanish /b, d, g/.

The native Spanish subjects did not prevoice Spanish /b, d, g/ in every instance. Of the 750 tokens examined, 7.2% were short-lag stops (or at least did not have prevoicing of sufficient amplitude to be visible on spectrograms). This was probably not the result of exposure to English. Keating's (1984) data showed that five Polish monolinguals recorded in Poland sometimes realized Polish /b, d, g/ with short-lag VOT values. (As in Spanish, Polish /p, t, k/ are implemented as short-lag stops, while /b, d, g/ are implemented with lead VOT values.) It may be physiologically easier to realize stops with short-lag than long-lag or lead VOT values (Kewley-Port & Preston, 1974). Moreover, the native Spanish children who spoke English (BC) realized Spanish /b, d, g/ with short-lag VOT values somewhat less, rather than more, often than the native Spanish children (SC) who had not learned English. The 20 child subjects realized /b, d, g/ with short-lag VOT values significantly more often (10% of instances) than the 30 adults (6%; $X = 5.43$; $p < 0.05$). This probably occurred because the children's control of the gestures needed to ensure voicing in the closure interval of stops was not fully developed (Flege, McCutcheon & Smith, 1986).

The Spanish subjects who spoke English realized Spanish /p, t, k/ with shorter VOT values than the native Spanish speakers who did not. The children in group BC realized /p, t, k/ with a shorter mean VOT value (12 ms) than the Spanish monolingual children in SC (20 ms). The Spanish adults in groups ECB and LCB realized Spanish /p, t, k/ with somewhat shorter mean VOT values (23 ms and 18 ms, respectively) than the Spanish adults in group SA (26 ms). The interaction of Group \times Place was significant [$F(8,90) = 3.89$; $p < 0.01$]. The simple main effect of Group was significant for /p/ [$F(4,45) = 7.19$], /t/ [$F(4,45) = 9.15$], and /k/ [$F(4,45) = 18.0$; $p < 0.01$]. Post-hoc comparisons using the Newman-Keuls procedure showed that: the native Spanish children who spoke English (BC) realized /p/, /t/ and /k/ with significantly shorter VOT values than the monolingual Spanish children in group SC; the adult Spanish speakers of English (groups LCB and ECB) produced /k/ with significantly shorter mean VOT values than the adults in group SA; and the subjects in LCB produced /t/ with shorter VOT values than the subjects in SA ($p < 0.05$).

3.1.2. English stops

Table II presents the mean VOT in stops initiating English words as a function of place of articulation for the 30 native Spanish and 20 native English subjects. The native Spanish subjects (BC, LCB, ECB) realized English /b, d, g/ with short-lag VOT values significantly less often (26%) than the native English subjects (64%) in groups EC and EA ($X^2(1) = 60.3$; $p < 0.01$). The native English children realized /b, d, g/ with short-lag VOT values significantly more often (91%) than the native English adults (37%; $X^2(1) = 34.3$, $p < 0.01$). The native Spanish children, on the other hand, showed a non-significant tendency to realize /b, d, g/ with short-lag VOT less frequently (20%) than the Spanish adults in groups LCB and ECB (29%; $X^2(1) = 2.88$, ns).

The VOT values measured in voiceless English stops varied as a function of place of articulation. For the five groups taken as a whole, /k/ had a mean VOT of 76 ms, /t/ a mean of 71 ms, and /p/ a mean VOT of 58 ms. The native speakers of Spanish realized /p, t, k/ with shorter VOT values than age-matched native English subjects. The native Spanish adults in ECB and LCB realized English /p, t, k/ with shorter (Spanish-like) mean VOT values (75 ms and 57 ms, respectively) than the English adults in EA (94 ms).

TABLE II. The mean VOT produced in English words by monlingual English children (EC) and adults (EA), native Spanish children who spoke English (BC), and two groups of native Spanish adults who spoke English (ECB, LCB), in ms. SD is the standard deviation; *n* is the number of tokens. Separate mean values were calculated for lead and short-lag realizations of /b, d, g/

Group		Lead VOT			Short-lag VOT					
		/b/	/d/	/g/	/b/	/d/	/g/	/p/	/t/	/k/
EC	<i>x</i>	-65	-50	-81	15	22	32	70	82	84
	SD	(55)	(6)	(48)	(7)	(9)	(8)	(15)	(16)	(14)
	<i>n</i>	5	4	5	45	46	45	50	50	50
EA	<i>x</i>	-98	-85	-81	19	18	30	78	89	94
	SD	(16)	(16)	(17)	(14)	(8)	(5)	(16)	(15)	(15)
	<i>n</i>	32	35	28	18	15	22	50	50	50
BC	<i>x</i>	-70	-68	-69	5	9	20	40	54	60
	SD	(26)	(19)	(25)	(4)	(4)	(11)	(10)	(10)	(14)
	<i>n</i>	42	43	34	8	6	16	50	50	50
ECB	<i>x</i>	-66	-62	-56	18	23	35	57	74	75
	SD	(18)	(16)	(19)	(7)	(14)	(15)	(14)	(16)	(17)
	<i>n</i>	34	32	37	14	18	13	50	50	50
LCB	<i>x</i>	-66	-62	-70	25	21	38	48	56	67
	SD	(23)	(20)	(17)	(14)	(5)	(20)	(20)	(16)	(20)
	<i>n</i>	41	35	32	9	15	17	50	50	50

The native Spanish children in group BC also realized English stops with a shorter mean VOT value (51 ms) than the English children in EC (79 ms).

The ANOVA revealed that the effects of Place [$F(2,90) = 57.1$; $p < 0.01$] and Group were significant [$F(4,45) = 11.6$; $p < 0.01$]. Post-hoc tests revealed that the native Spanish adults in Groups ECB and LCB produced English /p, t, k/ with significantly shorter VOT values than the native English adults in EA; and that the native Spanish children in BC realized stops with shorter VOT values than the English children in group EC ($p < 0.05$). The difference between ECB and LCB was non-significant.

3.2. Cross-language switching

The native Spanish speakers of English tended to realize /b, d, g/ with short-lag VOT values more often in English than Spanish. The difference averaged 6% for the children in group BC, 6% for the adults in ECB, and 8% for the adults in LCB. Chi-square tests indicated that none of these differences was significant ($p < 0.05$).

There were significant differences in the way /p, t, k/ were realized in Spanish and English. The subjects in all three bilingual groups (LCB, ECB, EC) realized /p, t, k/ with longer VOT values in English than Spanish. The VOT difference averaged 46 ms for group ECB, 41 ms for group LCB, and 41 ms for group BC. The lack of a Language \times Group interaction [$F(2,27) = 0.08$, ns] indicated that differences in age and linguistic experience did not enable the subjects in any one group to produce a more substantial between-language difference than the subjects in the other groups. The significant Language \times Place interaction [$F(2,54) = 8.45$; $p < 0.01$] derived from the fact that

/p/ was realized with significantly shorter VOT values than /t/ and /k/ in Spanish but not English.

3.3. Stop identification

As reported earlier (Flege & Eefting, 1986), the subjects in groups EA, EC, SA and SC consistently identified the endpoint stimuli in the VOT continuum as /da/ and /ta/, and showed a rapid cross-over from /da/ to /ta/ at a point along the continuum which varied according to both age and native-language. The same was true for the native Spanish subjects in groups LCB, ECB and EC. The identification functions of all but one of the subjects in these groups was perfectly monotonic. All 30 of them identified the -60 ms stimulus as /da/ in 100% of instances. Only one of these subjects identified the +90 ms stimulus as /ta/ in less than 100% of instances. The identification functions of the Spanish speakers of English were quite steep: the number of "ambiguous" stimuli (i.e. stimuli judged with less than 90% consistency) averaged 0.9 for group ECB, 1.0 for LCB, and 1.9 for BC.

The mean category boundaries of the adult Spanish speakers of English in LCB (29 ms) and ECB (27 ms) occurred at shorter VOT values than that of the monolingual English adults in group EA (43 ms). However, the mean boundary of the native Spanish children who spoke English occurred at about the same value (33 ms) as that of the native English children in group EC (30 ms). The main effect of Group was highly significant [$F(6,63) = 15.1$; $p < 0.01$]. The post-hoc tests showed that: the boundaries of the monolingual English adults (EA) occurred at significantly longer VOT values than those of subjects in the six other groups; the boundaries of the Spanish children who spoke English (BC) occurred at longer values than the Spanish adults and children who did not speak English (SA, SC); and the boundaries of the Spanish monolingual children (SC) occurred at significantly shorter VOT values than those of the monolingual English children (EC) and the native Spanish adults in groups SA, ECB and LCB ($p < 0.05$).

The /d-/t/ boundaries obtained for the 10 subjects in groups LCB, ECB and BC are presented in Table III, along with the mean boundaries obtained for English monolinguals (EA, EC) and the native Spanish subjects who did not speak English (SA, SC).

TABLE III. The /d-/t/ boundary of the subjects in groups LCB, ECB and BC, (in ms). S and E indicate whether a subject began the experiment in Spanish or English (see Methods)

Mean of monolingual English adults (Group EA) = 42.6 ms									
Mean of monolingual Spanish adults (Group SA) = 23.1 ms									
Phoneme boundaries of the early childhood bilinguals (ECB)									
S	E	S	E	E	S	E	S	E	S
30.0	22.9	25.6	26.7	24.4	27.1	30.0	26.3	28.3	32.9
Phoneme boundaries of the late childhood bilinguals (LCB)									
E	S	E	S	E	S	E	S	E	S
35.0	34.0	34.4	22.0	33.3	23.3	32.0	25.0	18.3	32.9
Mean of monolingual English children (Group EC) = 29.9 ms									
Mean of monolingual Spanish children (Group SC) = 16.6 ms									
Phoneme boundaries of the child bilinguals (BS)									
E	S	E	S	E	S	E	S	E	S
38.0	32.0	24.3	25.0	35.0	30.0	30.0	26.0	41.7	44.0

The half-way point between the mean values observed for group SA (23 ms) and EA (43 ms) is 33 ms. The adult Spanish speakers of English either had boundaries closely resembling the mean boundary of the native Spanish adults in group SA, or a boundary close to that half-way point. Three subjects in ECB had a category boundary near the half-way point, while seven had boundaries resembling the mean for group SA. Of the subjects in LCB, viz. six, had boundaries close to the half-way point, and four had boundaries close to the mean for group SA.

The category boundaries of the Spanish children who spoke English (BC) either closely approximated or exceeded the mean value observed for the native English children (EC), or came close to the half-way point between the means observed for the monolingual Spanish children (17 ms) and the monolingual English children in EC (30 ms). Of the children in group BC, three had boundaries near the half-way point (i.e. 23 ms), and seven had boundaries resembling those of the native English children in group EC⁸.

4. Discussion

Based on previous research (e.g. Caramazza *et al.*, 1973; Flege & Port, 1981; Port & Mitleb, 1983; Flege & Hillenbrand, 1984; Flege & Eefting, 1986; Flege, 1986b), it was hypothesized that native Spanish speakers of English would be unable to produce English /p, t, k/ authentically. More specifically, it was predicted they would produce /p, t, k/ with significantly shorter (i.e. Spanish-like) VOT values than native English subjects as the result of judging the [p, t, k] and [p^h, t^h, k^h] phones in Spanish and English words, respectively, to be realizations of the same phonetic category.

The results confirmed this prediction for two groups of adults and one group of children who began learning English as an L2 by the age of 5–6 years. Even though subjects in the three native Spanish groups realized /p, t, k/ with significantly longer VOT values in English than Spanish words, they realized /p, t, k/ with significantly shorter VOT values in English words than age-matched native speakers of English. In agreement with results obtained by Caramazza *et al.* (1973) for French speakers of English, the mean VOT for /p, t, k/ observed for the native Spanish subjects in group LCB was almost perfectly intermediate (57 ms) to the means observed for the Spanish adults who did not speak English (26 ms) and the monolingual English adults (87 ms). The native Spanish children who learned English as an L2 also realized English /p, t, k/ with a mean VOT (51 ms) that was almost perfectly intermediate to the means observed for Spanish and English monolingual children (20 ms and 79 ms), respectively⁹.

⁸To minimize the likelihood that the bilingual subjects would show an effect of "language set" (see Elman *et al.*, 1977), half of the subjects were using English, and half Spanish, at the time of the identification experiment (see Methods). Neither the effect of Language Order, nor the interaction of Group \times Order, significantly affected the boundary values. All of the subjects heard the 16 stimuli in ascending order of VOT prior to the experiment. It is likely that since the continuum contained long-lag stops with VOT values of up to +90 ms (which could obviously not occur in Spanish), the Spanish speakers of English adopted an "English" set for identifying the stimuli insofar as they were able.

⁹The means obtained for individual subjects were generally consistent with the grouped means. The intermediate VOT values obtained for the native Spanish speakers of English were not simply an artifact of averaging Spanish-like VOT values for some subjects with English-like values for others. All 10 native Spanish children in BC realized /p/, /t/ and /k/ with shorter mean values than the mean value observed for the native English children in EC. The Spanish adults in ECB realized English voiceless stops with shorter mean values than those observed for native English adults in all but one instance (one subject realized /k/ with a mean (95 ms) that slightly exceeded the native English mean for /k/). In only three instances did the Spanish adults in LCB realize English voiceless stops with mean values exceeding the mean values observed for the native English adults.

The adult subjects in ECB realized English /p, t, k/ with a mean VOT (69 ms) that resembled somewhat more closely the phonetic norm of English than Spanish. Perhaps the somewhat earlier exposure to English of the subjects in ECB than LCB, or a greater amount of L2 input from native speakers of English, enabled them to more nearly approximate the phonetic norm of English (as estimated by the native English adults).

These findings strongly suggest that the age at which L2 learning commences does not in itself prevent L2 learners from realizing L2 stops authentically. We can be fairly confident that the observed differences between native and non-native speakers of English did not derive from an inability to detect acoustic differences distinguishing between Spanish and English stops, for the native Spanish subjects realized /p, t, k/ with significantly longer VOT values in Spanish than English words. The switch in production indicated they had at least partial tacit awareness of acoustic differences distinguishing the /p, t, k/ of Spanish and English which must have come through the auditory modality.

Flege (1986b) suggested that L2 learners will be unable to realize English /p, t, k/ authentically if they are implemented as short-lag stops in their L1 because equivalence classification will prevent them from establishing phonetic categories for English stops. As applied to this study, the hypothesis rests on the assumption that, for example, Spanish learners of English judge the [t] and [t^h] phones in Spanish and English to be realizations of the same phonetic category (i.e. Spanish {t}). It also assumes that talkers cannot differentiate stops to the extent required by the cross-language VOT difference distinguishing /t/ in Spanish and English by applying realization rules to the parameters specified for the short-lag stop category {t} (see footnote 1 and Keating, 1984). Spanish learners of English might note acoustic differences between Spanish and English /t/, and might produce /t/ differently in Spanish and English. However, the “equivalence classification” hypothesis leads to the prediction that Spanish learners of English will never produce English /t/ authentically because they are prevented from establishing a long-lag stop category, {t^h}.

Flege (1981, 1986a) further proposed that equivalence classification will cause stops in L1 to resemble corresponding stops in L2. This was confirmed in a recent study (Flege, 1986b) showing that highly experienced native English speakers of French produced English /t/ with significantly shorter (French-like) VOT values than English monolinguals; and, conversely, that highly experienced native French speakers of English produced /t/ in French words with significantly longer means than French monolinguals. However, contrary to prediction, the L2 effect on L1 production observed here represented a differentiation rather than assimilation of VOT values for L1 and L2 voiceless stops. The native Spanish speakers of English in all three groups produced Spanish /p, t, k/ with significantly shorter VOT values than age-matched native Spanish subjects who had not learned English as an L2. That is, their Spanish stops were somewhat less English-like (at least in terms of VOT) than those of Spanish subjects who had not learned English as an L2.

The data presented here are insufficient to provide a satisfactory explanation for the divergent L2 effect on L1 production noted in the two studies. One possibility is that the direction of the L2 effect on L1 depends on the extent to which the L1 phonetic system has been developed when L2 learning commences. Flege (1986b) examined subjects who began learning their L2 (English or French) in early adulthood. Subjects in the present study, on the other hand, began learning their L2 (English) by the age of 5–6 years.

Age of learning might affect whether L2 learners establish phonetic categories for stops found in L2 but not L1. An individual exposed to [t] and [t^h] tokens in Spanish and English words, respectively, might regard the short- and long-lag stops as being the realizations of separate categories because they occur in some of the same phonetic contexts. Alternatively, they might regard them as being realizations of the same category because these phones tend to be heard in different languages and social contexts, and thus might be regarded as being in complementary distribution (Weinreich, 1953). It is conceivable that young children are more likely than adults to arrive at the "new category" solution when exposed to L2 phones not found in L1 because they are still in the process of establishing phonetic categories based on the phonetic input they receive in L1.

Degree of literacy might also be an important factor. For example, adults who learn an L2 through formal instruction may be more inclined to judge realizations of /p, t, k/ in L1 and L2 as belonging to the same category because they are spelled the same than young children or illiterate adults.

We speculate that the native Spanish speakers of English examined here established separate phonetic categories for English [p^h], [t^h], and [k^h] phones because they began learning English as young children. If so, these "early L2 learners", but not the "late L2 learners" examined by Flege (1986*b*), may have differentiated L1 stops from corresponding L2 stops as the result of "polarization". According to Keating (1984), three universal phonetic categories of stops might exist at a single place of articulation (i.e. voiced, voiceless unaspirated, and voiceless aspirated). Keating proposed that the universal phonetic principle of polarization leads to a small VOT difference between the short-lag stops of Spanish and English. The short-lag realizations of Spanish /p, t, k/ have somewhat longer VOT values than the short-lag realizations of English /b, d, g/. This is because the short-lag realizations of Spanish /p, t, k/ "polarize" away from Spanish /b, d, g/ (which are realized with lead VOT values), while the short-lag realizations of English /b, d, g/ polarize away from English /p, t, k/ (which are realized with long-lag VOT values). Note that, in both instances, polarization would tend to enhance the phonetic contrast between phonologically voiced and voiceless stops.

If universal principles affect the speech of individuals, and if the early L2 learners examined here established phonetic categories for English [p^h], [t^h] and [k^h], the shortening of VOT in Spanish /p, t, k/ observed for the Spanish speakers of English could be the result of polarization. They may have realized /p, t, k/ in Spanish in such a way as to enhance the acoustic difference between the short-lag categories used to implement /p, t, k/ in Spanish (i.e. {p, t, k}) and the long-lag categories used to implement /p, t, k/ in English (i.e. {p^h, t^h, k^h}). A similar effect would not be expected for late L2 learners if they do not establish phonetic categories for English /p, t, k/.

The hypothesis just offered might explain why early and late L2 learners differ, but it leads to another problem. If the early L2 learners examined here established phonetic categories for long-lag English stops, they should have been able to realize /p, t, k/ authentically in English. That is, if the native Spanish speakers of English did not judge tokens of /p, t, k/ in Spanish and English to be realizations of the same categories, there is no reason to expect an upper limit on the authenticity with which they produced English /p, t, k/. Even assuming a "critical period" exists for human speech learning (see Flege, 1986*c*, for arguments to the contrary), age cannot be invoked to explain why the subjects of this study differed significantly from native English speakers for the reasons noted above.

It is possible that the incomplete approximation to L2 phonetic norms observed here may have been the result of non-authentic L2 input. Martinet (1968, p. vii) observed that a linguistic community is "never homogeneous", and that each individual is a "battle-field for conflicting linguistic types and habits". English is viewed by many Puerto Ricans as a vehicle for economic advancement. For this and other reasons, the native Spanish-speaking parents of many subjects in this study were reported to speak English in the home occasionally to encourage their children to learn English. It is likely that our native Spanish subjects were exposed to English spoken by native speakers of Spanish in which tokens of /p, t, k/ were realized with VOT values intermediate to the short-lag and long-lag values typifying Spanish and English. We can only speculate concerning what part of their English L2 input came from native speakers of English, and what part of it came from non-native speakers. However, the subjects examined here may have produced English /p, t, k/ with about the same intermediate VOT values they heard.

This interpretation is consistent with the results of the identification experiment. About half (9/20) of the adult Spanish speakers of English had category boundaries at VOT values that were intermediate to the mean values obtained for Spanish adults who did not speak English (SA) and the monolingual English adults. The remaining subjects' boundaries occurred at values similar to those obtained for group SA. This suggests that the adult Spanish speakers of English judged the VOT continuum either in terms of the /t/-/d/ contrast of Spanish, or in terms of the English /d/-/t/ contrast they heard spoken by other native speakers of Spanish. Further research will be needed to directly determine the relationship between L2 phonetic input and category boundaries.

The category boundary locations of the Spanish children who spoke English are more difficult to understand. Of these children, three had boundaries intermediate to those of monolingual Spanish and English children, and seven had boundaries close to the mean obtained for the English children. It is unclear to us why these children produced English /p, t, k/ with significantly shorter VOT values than English monolingual children, but did not differ significantly from them in labelling stops. This finding is, of course, consistent with the general view that the development of speech perception "leads" the development of speech production. It suggests that the Spanish children would eventually produce English /p, t, k/ authentically.

However, the results obtained for the Spanish adults in group LCB indicate this will probably not happen. Like the native Spanish children in BC, the adults in LCB began learning English in an English-speaking school in Puerto Rico at the age of 5-6 years. We do not know if they closely resembled native English children in producing English /p, t, k/ when they themselves were children. What we do know is that, as adults, they produced English /p, t, k/ with significantly shorter VOT values than native English adults. It is therefore possible that as the children in group BC get older, their /t/-/d/ category boundaries will not move to progressively longer values as it does for monolingual native speakers of English (Flege & Eefting, 1986). If so, in the future they will differ from age-matched English monolinguals in both the perception and production of English stops.

In summary, this study showed that individuals who begin learning English L2 as young children do not realize English /p, t, k/ authentically if these stops are implemented with different phonetic categories in English and their L1. This suggests that the divergences noted between native and non-native adults in previous production studies did not result from passing a "critical period" for speech learning. The results indirectly suggested that the native Spanish speakers of English may have established phonetic

categories for the long-lag realizations of /p, t, k/ found on the phonetic surface of English. However, the effect of learning English on the production of Spanish /p, t, k/ showed that these categories were probably not fully independent of the categories established previously for short-lag realizations of /p, t, k/ in Spanish. They seem to have mutually influenced one another through the "inter-lingual" identification described by Weinreich (1953), probably at an abstract phonological level of organization. The production differences observed between the Spanish subjects and the native speakers of English further suggested that L2 production depends importantly on the nature of the phonetic input received during L2 learning.

This research was sponsored by NINCDS grant NS20963. The authors would like to thank E. Blau, C. Casablanca, and E. Flores for help locating subjects in Puerto Rico; J. Serota for providing native English child subjects in Birmingham; and S. Fletcher and B. Repp for their comments on a previous draft of this article.

References

- Abramson, A. & Lisker, L. (1970) Discriminability along the voicing continuum: Cross-language tests, *Proceedings of the 6th International Congress of Phonetic Sciences*, pp. 569–573. Prague: Academia.
- Abramson, A. & Lisker, L. (1973) Voice-timing perception in Spanish word-initial stops, *Journal of Phonetics* 1, 1–8.
- Bailey, P. & Haggard, M. (1980) Perception-production relations in the voicing contrast for initial stops in 3-year-olds, *Phonetica*, 37, 377–396.
- Bond, Z. & Korte, S. (1983) Children's spontaneous and imitative speech: An acoustic-phonetic analysis, *Journal of Speech and Hearing Research*, 26, 464–466.
- Caramazza, A., Yeni-Komshian, G., Zurif, E. & Carbone, E. (1973) The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals, *Journal of the Acoustical Society of America*, 54, 421–428.
- Elman, J., Diehl, R. & Buchwald, S. (1977) Perceptual switching in bilinguals, *Journal of the Acoustical Society of America*, 62, 971–974.
- Flege, J. (1980) Phonetic approximation in second language acquisition, *Language Learning*, 30, 117–134.
- Flege, J. (1981) The phonological basis of foreign accent: A hypothesis, *TESOL Quarterly*, 15, 443–445.
- Flege, J. (1984) The detection of French accent by American listeners, *Journal of the Acoustical Society of America*, 76, 692–707.
- Flege, J. (1986a) The production and perception of foreign language speech sounds, In *Human Communication and Its Disorders, Vol. II* (H. Winitz, editor). Norwood, N.J.: Ablex Publishing, in press.
- Flege, J. (1986b) The production of "new" and "similar" phones in a foreign language, *Journal of Phonetics*, under review.
- Flege, J. (1986c) A critical period for learning to pronounce foreign languages? *Applied Linguistics*, in press.
- Flege, J. & Eefting, W. (1986) Developmental and linguistic effects on the production and perception of stop consonants, *Phonetica*, under review.
- Flege, J. & Hammond, R. (1982) Mimicry of non-distinctive phonetic differences between language varieties, *Studies in Second Language Acquisition*, 5, 1–17.
- Flege, J. & Hillenbrand, J. (1984) Limits on pronunciation accuracy in adult foreign language speech production, *Journal of the Acoustical Society of America*, 76, 708–721.
- Flege, J., McCutcheon, M. & Smith, S. (1986) The development of skill in producing word-final English stops, *Journal of the Acoustical Society of America*, to appear.
- Flege, J. & Port, R. (1981) Cross-language phonetic interference: Arabic to English, *Language and Speech*, 24, 125–46.
- Fokes, J., Bond, Z. S. & Steinberg, M. (1985) Acquisition of the English voicing contrast by Arab children, *Language and Speech*, 28, 81–91.
- Harris, J. (1969) *Spanish Phonology*. Cambridge, MA: MIT Press.
- Keating, P. (1984) Phonetic and phonological representation of stop consonant voicing, *Language*, 60, 286–319.
- Kewly-Port, D. & Preston, M. (1974) Early apical stop production, *Journal of Phonetics*, 2, 195–210.
- Klatt, D. (1980) Software for a cascade/parallel formant synthesizer, *Journal of the Acoustical Society of America*, 67, 971–995.
- Lisker, L. & Abramson, A. (1964) Cross-language study of voicing in initial stops: Acoustical measurements, *Word*, 20, 384–422.

- Lisker, L., Liberman, A., Erickson, D., Dechovitz, D. & Mandler, R. (1977) On pushing the voice-onset-time (VOT) boundary about, *Language and Speech*, **20**, 209–216.
- Mack, M. (1983) *Psycholinguistic Consequences of Early Bilingualism: A Comparative Study of Performance of English Monolinguals and French-English Bilinguals in Phonetic, Syntactic, and Semantic Experiments*. Unpublished Ph.D. thesis, Brown University.
- Macken, M. (1980) The acquisition of stop systems: A cross-linguistic perspective, In *Child Phonology, Vol. 1, Production* (G. Yeni-Komshian, J. Kavanagh & C. Ferguson, editors), pp. 143–168. New York: Academic Press.
- Martinet, A. (1968) Preface to U. Weinreich, *Languages in Contact, Findings and Problems*, The Hague: Mouton.
- McLaughlin, B. (1978) *Second Language acquisition in childhood*. Hillsdale, N.J: Lawrence Erlbaum.
- Port, R. & Mitleb, F. (1983) Segmental features and implementation in acquisition of English by Arabic speakers, *Journal of Phonetics*, **11**, 219–231.
- Smith, B. (1978) Effects of place of articulation and vowel environment on “voiced” stop consonant production, *Glossa*, **12**, 163–175.
- Weinreich, U. (1953) *Languages in contact, Findings and problems*, The Hague: Mouton.
- Williams, L. (1977a) The voicing contrast in Spanish, *Journal of Phonetics*, **5**, 169–184.
- Williams, L. (1977b) The perception of stop consonant voicing by Spanish-English bilinguals, *Perception and Psychophysics*, **21**, 289–297.
- Williams, L. (1979) The modification of speech perception and production in second-language learning, *Perception and Psychophysics*, **26**, 95–104.
- Zlatin, M. & Koenigsknecht, R. (1976) Development of the voicing contrast: A comparison of voice onset time in stop perception and production, *Journal of Speech and Hearing Research*, **19**, 93–111.