CROSS-LANGUAGE PHONETIC INTERFERENCE: ARABIC TO ENGLISH*

JAMES EMIL FLEGE
Northwestern University
and
ROBERT PORT
Indiana University

This study compares phonetic implementation of the stop voicing contrast produced in Arabic by Saudi Arabians and by both Americans and Saudis in English. The English stops produced by Saudis manifested temporal acoustic correlates of stop voicing (VOT, stop closure duration, and vowel duration) similar to those found in Arabic stops. Despite such phonetic interference from Arabic to English, however, American listeners generally had little difficulty identifying the English stops produced by the Saudis, with the exception of /p/. This phoneme, which is absent in Arabic, was frequently produced with glottal pulsing during the stop closure interval. The timing of /p/, however, suggests that the Saudis did grasp the phonological nature of /p/ (i.e., that the contrast between /p–b/ is analogous to that between /t–d/ and /k–g/) but were unable to control all the articulatory dimensions by which this sound is produced.

Perhaps the most important and obvious aspect of foreign-accented speech is sound substitutions, such as [s] for /θ/ in French-accented "I sink so." But a large part of what leads to the perception of accentedness probably cannot be adequately represented by a segmental phonetic transcription. We began this study with the hypothesis that both the phonological structure and phonetic characteristics of a speaker's native language will influence his pronunciation of sounds in a foreign language learned in adulthood. Cross-language interference may occur at several levels of organization. First, a speaker might mispronounce a sound in a foreign language because no comparable sound exists in the **phonemic** inventory of his native language (Lado, 1957). But if such a novel sound is composed of features that specify sounds which *do* exist in the speaker's native language, however, a contrastive analysis based on phonemic principles (see Flege, 1979) predicts that it will be learned with relatively little difficulty (Weinreich, 1953, p. 22). If distinctive features are indeed "commutable" (Jakobson, 1962, p. 420) and can thus be transferred from sound to sound, then a foreign language speech sound that represents a "hole in the pattern" of the native language phonemic inventory should be easy to learn. Second, interference might occur at the level of **segmental phonetic** features even if the

* *This research is based on an Indiana University Ph.D. thesis by the first author which was supervised by the second author. It was funded in part by NICHD grant HD12511 to Indiana University, and by a Post-doctoral Fellowship (NIH grant NS 07107) to the first author through the Institute for the Advanced Study of the Communication Processes, University of Florida.*

more abstract phonological features that specify a sound have been correctly combined. Support for the existence of this kind of interference would exist if language learners were to mispronounce only certain allophones of a novel foreign language phoneme. And, third, interference might result from cross-language differences in the **phonetic implementation** of a feature.

It has been claimed that a segmental phonetic transcription can, in principle, describe all the linguistically controllable aspects of speech (Chomsky and Halle, 1968) but even the best phonetic transcription can probably not capture perfectly an idiolect or accent.[1] Research on speech timing, for example, suggests that similar sounds found in different languages may have quite different patterns of temporal implementation (Lehiste, 1970; Kohler, 1979; Port, Al-ani and Maeda, 1980). Such cross-language timing differences may not be directly perceptible at a segmental level to most listeners, but they may well contribute to the perception of accentedness and even, in some cases, result in diminished intelligibility (Jonasson and McAllister, 1972; Huggins, 1976). Although there is relatively little cross-language research on coarticulation, it seems likely that this aspect of sub-segmental phonetic implementation might sometimes also prove incommensurable across languages. For example, the degree to which vowels preceding nasal consonants are nasalized seems to vary from language to language (Clumeck, 1976).

In the present study we examined several acoustic dimensions that are phonetic correlates of the phonological contrast between voiced and voiceless stops. Voice-onset time (VOT) is a measure of the time between release of stop closure and the onset of glottal pulsing (voicing). This acoustic dimension often distinguishes classes of stops like /ptk/ and /bdg/, and may be sufficient to cue the perceptual distinction between such stop categories (Lisker and Abramson, 1964, 1967, 1971). Duration of the closure interval of a stop as well as the duration of vowels preceding a stop are two other important temporal acoustic correlates of the voicing contrast in many languages (Lehiste, 1970; Klatt, 1976). And, finally, the presence or absence of glottal pulsing (voicing) during the closure interval of a stop is very often an important spectral dimension by which voiced and voiceless stops are distinguished (Lisker, 1978).

We recorded and measured phonetically similar material representing colloquial Saudi Arabian Arabic, American English, and the accented English produced by Saudi Arabians using the same instrumental techniques. Arabic was chosen as the counterpoint to English in this study because the phonetic contrast between voiced and voiceless stops in Arabic appears to differ from that of English (Yeni-Komshian, Caramazza and Preston, 1977; Port, Al-ani and Maeda, 1980) and because Arabic lacks one of the stops found in English, the voiceless labial stop /p/ (Al-ani, 1970). These cross-language differences offered the opportunity to assess how a difference in phonological inventory as well as more subtle differences in the phonetic implementation of a phonological contrast would affect production of foreign language speech sounds by adult language learners.

[1] *See Monsen (1976) for an interesting discussion of the effect on intelligibility of non-segmental phonetic differences between the speech of normal and deaf speakers of English.*

EXPERIMENT 1: ARABIC

Since few previous studies provide data concerning the phonetic basis of the stop voicing contrast in Arabic, it was first necessary to examine stops in the Saudi Arabian dialect of Arabic in order to determine to what extent phonetic characteristics of Arabic-accented English directly result from Arabic-specific patterns of phonetic implementation.

Methods

Six adult male native speakers of Arabic, all university graduates from central or northeastern Saudi Arabia residing in Bloomington at the time of the study, served as subjects. All the speakers reported having a [g] in their native dialect of Arabic. Speakers read randomized lists of the Arabic words listed below from 3 x 5 in. cards, inserting each test word into a constant carrier sentence [ʔagra wamʃılelbeyt] 'I read ___ and then I go home':

	Initial stops		*Final stops*	
labial	**b**aas	'kissed'	ʃaa**b**	'grew old'
dental	**t**aas	'Tass'	gaa**t**	'kat' (tobacco)
	daas	'stepped'	gaa**d**	'led'
velar	**k**aas	'cup'	ʃaa**k**	'encircle'
	gaas	'measured'	ʃaa**g**	'difficult'

The test words were chosen so as to provide word-initial and word-final stop voicing contrast in CV:C minimal pairs. Phonologically long vowels were chosen instead of phonologically short vowels because a pilot study revealed that the duration of the long vowel /aa/ is closer in duration to English /æ/ (the vowel used in test words in the subsequent English experiment) than is short Arabic /a/. Flege (1979) found that in a pre-dental stop environment the duration of Arabic /aa/ was 177 msec, short Arabic /a/, 98 msec. English /æ/ averaged 187 msec when produced in a comparable phonetic context by Americans. Each sentence was produced in colloquial Saudi Arabian (rather than Classical or Standard Arabic)[2] while subjects were seated about 15 in. in front of a microphone (Electrovoice Model 635A) in a sound-proof booth. The experimenter monitored production of each sentence from outside the recording booth to ensure that test words carried main sentence stress and that subjects did not introduce pauses or

[2] *Since reading colloquial Arabic represents an unusual task for speakers of Arabic, we took precautions to insure that our subjects produced the test material in their native dialect. Before the experiment each speaker listened to instructions recorded in colloquial Saudi Arabian Arabic emphasizing the importance of producing the sentence material in colloquial rather than Standard or Classical Arabic. Since there is no /g/ in Classical Arabic, several words that are produced with [g] in Saudi Arabian dialects (e.g., [gamaʕ] v. [qamaʕ] 'full moon') were presented as examples of words produced in colloquial Arabic. An Arabic-speaking linguist later listened to the recordings and confirmed that they had been produced in colloquial Arabic.*

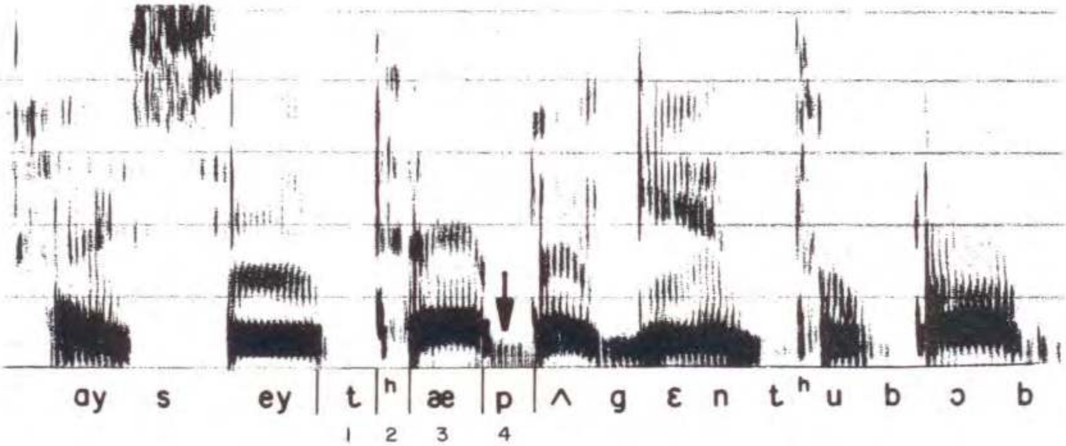

Fig. 1. Four acoustic intervals measured in the Arabic and English experiments: 1) initial stop duration; 2) VOT; 3) vowel duration; 4) final stop duration. Words in the Arabic experiment were preceded by [a] and followed by [w]; those in the English experiment by [ey] and [ʌ]. The final stop in the test word *tap* in this utterance (arrow) was produced by a Saudi Arabian speaker of English with glottal pulsing through the entire closure interval.

noticeable changes in speaking tempo.

The test material was recorded on a Revox (Model A700) tape recorder at 7 1/2 i.p.s. Wideband spectrograms were made of six tokens of each test word (Voice Identification, Model 700) and four contiguous intervals measured by hand to the nearest 5 msec according to the following segmentation criteria: 1) the **closure interval** of word-initial consonants was measured from the offset of the preceding vowel (indicated by a sharp decrease in energy in the region of F1 and F2) to the beginning of the noise burst signalling release of the stop closure interval or, in the case of fricatives, to onset of energy in F1 of the following vowel; 2) **VOT** was measured from the beginning of the release burst to the first visible striation in the region of F1; 3) **vowel duration** was measured from the onset of energy in F1 to the offset of energy in F1 and F2; and 4) the **closure interval** of word-final consonants was measured from offset of the preceding vowel to the beginning of the release burst (for stops) or to onset of energy in F1 in the following vowel (for unreleased stops and fricatives). Segmentation of these intervals is illustrated in Fig. 1 for a sentence from the English experiment (to follow). In addition, a binary judgment of the presence or absence of **glottal pulsing** during the stop closure interval was made. Stops that showed visible periodic striations on the spectrogram through at least half the closure interval were considered to have been produced with glottal pulsing.

TABLE 1

Mean duration of phonetic intervals in Arabic, in msec.
Standard deviations in parentheses.
VOT values were not measured for initial /bdg/ because these voiced stops generally showed continuous glottal pulsing through the stop closure interval

	Phonetic Interval			
	C_1	VOT	V:	C_2
/baas/	85 (14)		189 (26)	124 (28)
/taas/	90 (12)	37 (13)	176 (17)	116 (22)
/daas/	82 (13)		186 (17)	117 (19)
/kaas/	82 (16)	52 (15)	177 (22)	116 (23)
/gaas/	72 (9)		190 (23)	119 (23)
/ʃaab/	134 (18)		170 (21)	101 (29)
/gaat/	77 (13)		177 (24)	72 (13)
/gaad/	75 (13)		183 (28)	72 (17)
/ʃaak/	131 (17)		167 (17)	75 (18)
/ʃaag/	135 (18)		173 (22)	80 (20)

Results and discussion

Word-initial stops. In word-initial position /t/ and /k/ are aspirated stops in Saudi Arabian Arabic, as shown in Table 1. The mean VOT of /t/ was 37 msec (range of all tokens: 20-65 msec) and of /k/, 52 msec (range: 30-85 msec). These values, which are about 20 msec longer than the VOT values reported for utterance-initial stops in Lebanese Arabic by Yeni-Komshian *et al.* (1977), are considerably longer than for the "short-lag" stops found in languages like French or Spanish (Lisker and Abramson, 1964). On the other hand, they are less aspirated (i.e., have shorter VOT values) than "long-lag" stops found in languages such as English or Danish (Lisker and Abramson, 1967; Fischer-Jørgensen, 1968).

In addition to a VOT difference, the durations of the stop closure intervals of voiced and voiceless stops in initial position were also different. Pre-stress /t/ was about 8 msec longer than /d/, and /k/ was about 10 msec longer than /g/. These duration differences of about 12% were significant at the 0.01 level.

Place of articulation was found to exert an effect on the duration of stop closure intervals similar to that found in English and many other languages (Lehiste, 1970). The effect of place on the duration of stops ($p < 0.01$ for both voiced and voiceless stops) was a decrease in duration of the closure interval as place of articulation moved further back in the mouth (cf. Fischer-Jørgensen, 1964).

Vowel duration. The duration of the long vowel preceding voiced stops (/d/ and /g/) was not significantly longer than vowels preceding voiceless stops (/t/ and /k/). The difference in means amounts to only about 3% or 6-7 msec. This seems to violate the claimed universality of the stop voicing effect on preceding vowel duration (Chen, 1970). Our results here are not in agreement with Port *et al.* (1980) who reported a voicing effect on preceding stressed vowels of about 8% or 13 msec in three-syllable words.

Word-final stops. In word-final position the closure interval of voiced and voiceless stops did not show a significant contrast as did the word-initial stops. Our finding of no duration difference as a function of voicing is in agreement with the finding of Port *et al.* (1980) for speakers of several non-Saudi dialects of Arabic. Arabic thus seems to differ from English and at least other Germanic languages in which voiceless stops are longer than voiced stops (/bdg/) in post-stress position (Lisker, 1957; Elert, 1964; Kohler, 1979).

Glottal pulsing. Voiced and voiceless stops were distinguished by the presence or absence of glottal pulsing. Table 2 indicates the percentage of stops in initial and final position that exhibited visible glottal pulsing during at least half the closure interval. Both voiced stops (/d,g/) were produced with glottal pulsing far more frequently than were their voiceless cognates (/t,k/) ($p < 0.01$) in both word-initial and word-final position).

Conclusions

The stop voicing contrast of Saudi Arabian Arabic differs from that of American English in several ways. Word-initial Arabic voiceless stops (/t,k/) seem to be produced with somewhat shorter VOT values than similar stops in English (Lisker and Abramson, 1967). Voiceless stops in Saudi Arabian Arabic are produced with longer closure intervals than homorganic voiced stops in word-initial, pre-stress position. This temporal contrast does not exist in English (Stathopoulos and Weismer, 1979). There does not appear to

TABLE 2

Percentage of stops produced by Saudis with glottal pulsing visible through at least half the closure interval; "n" is number of tokens analyzed

	Initial Position				
	/b/	/d/	/g/	/t/	/k/
%	100	100	92	14	11
n	36	36	36	36	36
	Final Position				
	/b/	/d/	/g/	/t/	/k/
%	89	94	92	6	6
n	36	36	35	36	36

be a temporal contrast either between the closure intervals of voiced v. voiceless stops in word-final (post-stress) position, or in the duration of stressed vowels preceding voiced v. voiceless stops. English possesses both of these inversely related temporal correlates of stop voicing.

Given that previous studies of a number of languages have reported a stop voicing effect on preceding vowel duration (e.g., Chen, 1970), the present finding of no contrast in vowel duration in Saudi Arabian Arabic is somewhat surprising. Studies of other Arabic dialects have reported small or nonsignificant effects (Port *et al.*, 1980; Port and Mitleb, 1980) but Keating (1979) recently reported a similar negative finding for both Czech and Polish. Thus, it appears that this phonetic context effect on vowel duration may not be a phonetic universal as is often supposed.

Based on these findings we may conclude that Saudi Arabians learning English as a foreign language will be faced with a number of clear cross-language phonetic differences. To produce English stops without an Arabic accent a Saudi will need to modify Arabic patterns of phonetic implementation or else acquire novel English-specific patterns beside his existing Arabic patterns. If phonetic interference is direct and persistent, Saudis may be expected to maintain the stop voicing correlates of Saudi Arabian Arabic when producing English stops. In addition, Saudis will also need to learn to produce English /p/, since the phoneme does not exist in their native language.

In the next experiment we directly compared production of English stop voicing by native speakers of English and Arabic in order to determine whether Saudis learn to produce English stops according to English phonetic norms.

TABLE 3

Mean duration of English stop closures and vowels produced by three speaker groups, in msec. Standard deviations in parentheses

		Initial Stop Closure					
		/b–p/		/t–d/		/k–g/	
		pat	*bat*	*tab*	*dab*	*cab*	*gab*
Am	mean:	94	97	83	87	77	76
	S.D.:	(12)	(19)	(13)	(15)	(10)	(13)
Ar2	mean:	114	92	98	80	90	72
	S.D.:	(17)	(13)	(13)	(10)	(9)	(12)
Ar1	mean:	116	94	96	90	93	80
	S.D.:	(26)	(17)	(14)	(17)	(15)	(9)
		Vowel					
		tap	*tab*	*bat*	*bad*	*back*	*bag*
Am	mean:	134	163	174	199	162	204
	S.D.:	(19)	(26)	(26)	(38)	(22)	(35)
Ar2	mean:	138	139	151	153	147	160
	S.D.:	(24)	(27)	(21)	(24)	(23)	(22)
Ar1	mean:	133	135	135	150	138	146
		(28)	(24)	(27)	(33)	(31)	(27)
		Final Stop Closure					
		tap	*tab*	*bat*	*bad*	*back*	*bag*
Am	mean:	82	59	43	30	65	51
	S.D.:	(16)	(10)	(24)	(8)	(9)	(9)
Ar1	mean:	78	71	75	63	67	55
	S.D.:	(19)	(16)	(16)	(12)	(12)	(10)
Ar2	mean:	77	72	71	67	65	59
	S.D.:	(17)	(17)	(14)	(10)	(17)	(11)

EXPERIMENT 2: ENGLISH

Methods

Procedures for the English experiments were as similar as possible to those of the Arabic experiment. As in the Arabic experiment, subjects read a randomized list of minimal-pair test words that differed according to the voicing of word-initial or word-final stops produced at all three places of articulation, as shown below:

Initial		*Final*	
pat	bat	tap	tab
tab	dab	bat	bad
cab	gab	back	bag

The vowel /æ/ was chosen because it most nearly resembles the Arabic vowel /aa/ found in the test words of the Arabic experiment. The carrier sentence used in the English experiment ("I say ____ again to Bob"), was chosen to approximate the syllabic structure of the carrier sentence used in the Arabic experiment.

Three groups of speakers (six in each) served as subjects. One group consisted of Americans (Group Am), and two groups consisted of Saudi Arabian students at Indiana University (Groups Ar1 and Ar2). Three speakers in both Saudi groups had previously served as subjects in the Arabic experiment. Speakers in the two Saudi groups were male university graduates ranging in age from 24 to 32. Those in Ar1 had lived less than one year (mean: 8 months) in the U.S. at the time of the study, while speakers in Ar2 had lived in the U.S. for over two years (mean: 39 months). A preliminary questionnaire indicated that speakers in both groups had received comparable English language training in Saudi Arabia and had similar career objectives. Thus, any phonetic difference between the two groups of Saudis should be due primarily to learning based on experience speaking English.

The same acoustic correlates of stop voicing examined in the Arabic experiment – segment duration, VOT, and glottal pulsing – were measured in this experiment according to the same criteria. Measurement reliability was estimated by making a separate set of duration measurements from 32 duplicate spectrograms (198 acoustic intervals) produced by one speaker.[3] The average error was found to be 2.5 msec (range: 0-20 msec). Computer-implemented data analysis was conducted as for the Arabic experiment.

Results and discussion

Results, presented in Table 3, indicate that phonetic differences between Arabic and English lead to non-English phonetic characteristics in the English produced by Saudi Arabians.

[3] *The intervals measured were closure of initial and final stops, pre- and post-stress VOT, vowel duration, and utterance duration.*

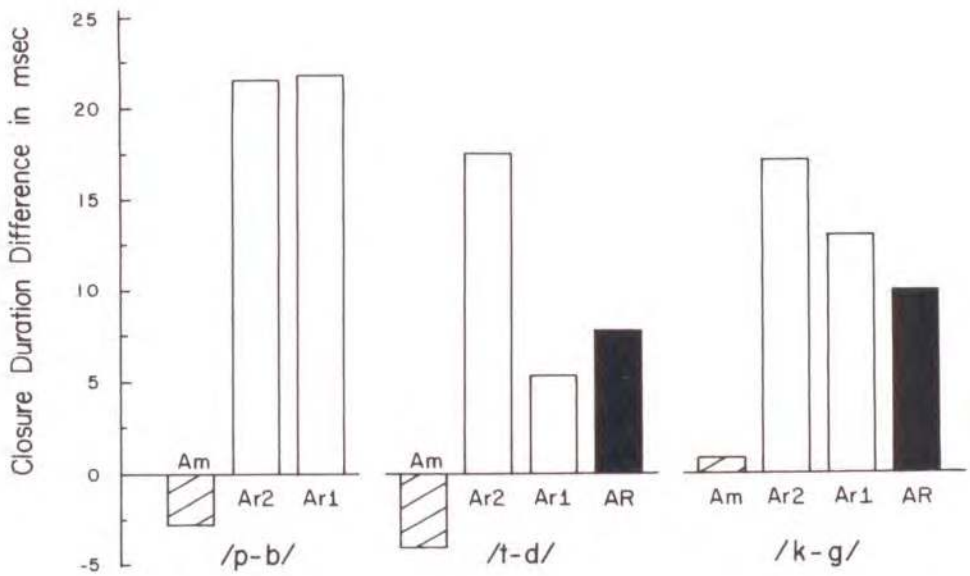

Fig. 2. Mean closure duration differences between word-initial voiced and voiceless stops produced by four speaker groups, in msec. The mean durations of voiced stops are subtracted from those of homorganic voiceless stops. Results from the Arabic experiment (Group AR) are juxtaposed to those of the English experiment (Groups Am, Ar1, Ar2).

Word-initial stops. The Saudi speakers (Ar1, Ar2) produced a temporal correlate of stop voicing for word-initial stops which was not produced by Americans. This temporal contrast between /ptk/ and /bdg/ is displayed in Fig. 2, where the mean durations of voiced stop closures are subtracted from the mean durations of homorganic voiceless stops. Here we see that the Saudis made the closure intervals of voiceless stops longer than those of voiced stops in word-initial position, a contrast which was significant ($p < 0.01$) in all but one case (the /t–d/ contrast produced by Ar1). The Americans, on the other hand, either produced no temporal contrast or else made voiced stops slightly (but non-significantly) longer than voiceless stops.

In order to display the influence of Arabic on the Saudis' English, results from the Arabic experiment (marked at AR in Fig. 2) are juxtaposed to results from the English experiment.

VOT values of the Saudis' English stops also closely resemble values found in Arabic. As shown in Table 4, the VOT of /pkt/ produced by the Saudis (Ar1, Ar2) averaged about 25 msec less than VOT values produced by the Americans (Am). Both the effect of place of articulation on VOT, and the difference in VOT between Americans and both Saudi groups were significant ($p < 0.01$). In Fig. 3 we have cumulatively plotted

TABLE 4

Mean voice-onset times (VOT) produced by three speaker groups, in msec. Standard deviations in parentheses

	Voice-Onset Time (VOT)		
	/p/ *pat*	/t/ *tap*	/k/ *cab*
Am	46 (10)	56 (11)	67 (14)
Ar2	21 (18)	30 (15)	47 (20)
Ar1	14 (14)	35 (12)	41 (12)

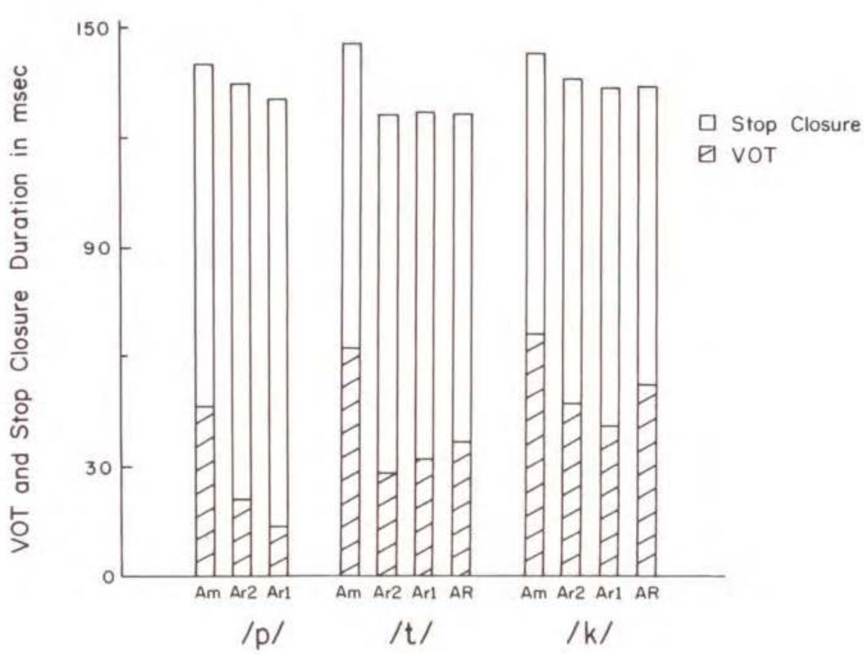

Fig. 3. Mean voice-onset time (VOT) and stop closure duration lined up at the onset of the following vowel, in msec. Results from the Arabic experiment (Group AR) are juxtaposed to those of the English experiment (Groups Am, Ar1, Ar2).

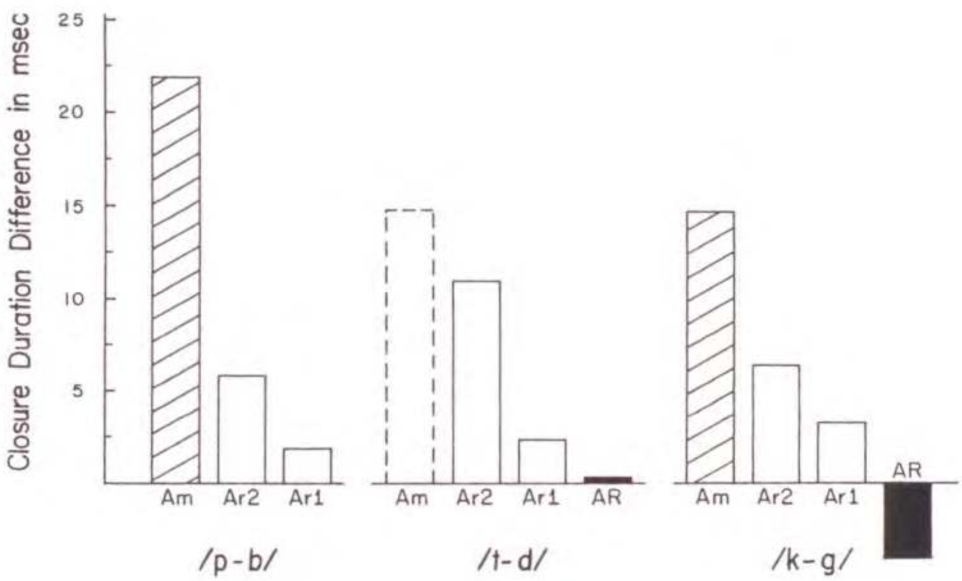

Fig. 4. Mean closure-duration difference between word-final voiced and voiceless stops produced by four speaker groups, in msec. The mean durations of voiced stops are subtracted from those of homorganic voiceless stops. Results from the Arabic experiment (Group AR) are juxtaposed to those of the English experiment (Groups Am, Ar1, Ar2). The histogram for the Americans' (Group Am) /t–d/ contrast represents 12 alveolar stops (of 72 tokens) that were not flapped.

VOT and the duration of the closure interval of /ptk/ produced by the three speaker groups alongside similar results for Arabic from the Arabic experiment (AR). We see here that the Americans (Am) produced longer VOT but shorter stop closure intervals in English than the Saudis (Ar1, Ar2) ($p < 0.01$). It is interesting to note that the *sum* of the VOT and stop closure intervals for word-initial /t/ and /k/ remain fairly constant for Saudi speakers in both the English and Arabic experiments (AR, Ar1, Ar2). Since the two experiments were designed to be as similar as possible,[4] it is surprising to see that the Saudis speaking English (Ar1, Ar2) do not approximate the longer VOT of English /t/ and /k/, but instead tend to slightly *shorten* VOT (vis-à-vis Arabic values (AR) from Experiment 1) and to lengthen the closure intervals of initial stops relative to Arabic values. This suggests a compensatory relation between the closure intervals of voiceless

[4] *Experiments in two languages must be compared with great caution unless identical phonetic material and procedures are used in both (see, e.g., Barry, 1974). Unfortunately, it was impossible to find a full list of minimal pairs that are real words in both Arabic and English. Use of nonsense CVCs seemed inadvisable because the focus of study was the learning of English stop voicing rather than some hypothetical phonetic ability.*

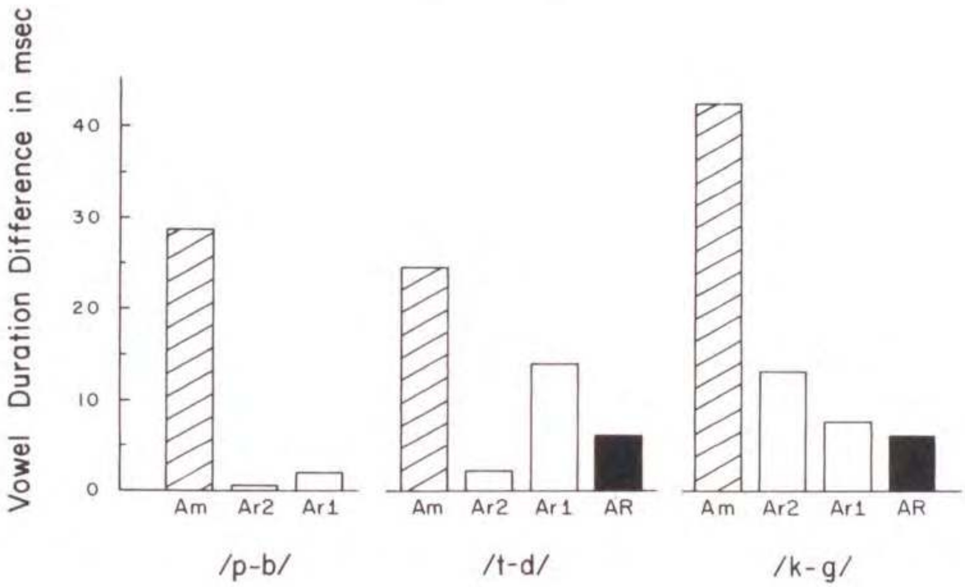

Fig. 5. Mean vowel-duration contrast produced by four speaker groups, in msec. The mean durations of vowels preceding voiceless stops are subtracted from the durations of vowels preceding homorganic voiced stops. Results from the Arabic experiment (Group AR) are juxtaposed to those from the English experiment (Groups Am, Ar1, Ar2).

stops and VOT. It is reminiscent of Weismer's claim (1980) that there may be a constant-duration gesture of devoicing in English, and implies that the duration of VOT and an adjacent closure interval may not be independently controlled.

Vowel duration. The effect of stop voicing on vowel duration is much smaller in the Saudis' than in the Americans' English. As displayed in Fig. 4, the mean durations of vowels preceding voiceless stops are subtracted from those of vowels preceding homorganic voiced stops. We see that the Americans (Am) made vowels longer before voiced than voiceless stops at all three places of articulation, a finding reported for English in many previous studies (e.g., House and Fairbanks, 1953; Peterson and Lehiste, 1960). The Saudis (Ar1, Ar2), on the other hand, produced a much smaller vowel duration contrast than the Americans; their differences reached significance in only three of six minimal pairs (/t–d/ for Ar1; /k–g/ for Ar1 and Ar2). Note that the relatively small effect of stop voicing on vowel duration in Arabic-accented English is closely comparable to the small and nonsignificant effect found in Arabic in Experiment 1 (AR) and plotted in Fig. 4 for /t–d/ and /k–g/.

Word-final stops. The closure-duration contrast between word-final voiced and voiceless stops produced by Saudis (Ar1, Ar2) was much smaller than that produced by native English speakers (Am). In Fig. 5 these contrasts are displayed by subtracting the mean durations of voiced stops from those of homorganic voiceless stops. We see here that the Americans (Am) made /ptk/ longer than /bdg/ in final position, as expected from previous studies of English (e.g., Lisker, 1957).[5] The apparent duration contrast between the Americans' /t/ and /d/ is based on the few tokens (12 of 72) of /t/ and /d/ that were not flapped (where a flap was operationally defined as having a closure interval of 40 msec or less). The flapped /t/s and /d/s were about equal in duration. The Saudis (Ar1, Ar2), on the other hand, produced much smaller duration contrasts between final voiced and voiceless stops than the Americans. The newly-arrived Saudis (Ar1) produced no significant difference in any pair of word-final voiced-voiceless stops, but the relatively more experienced Saudi speakers of English (Ar2) *did* make the closure intervals of voiceless stops longer than those of voiced stops at all three places of articulation.

The relatively small magnitude of the Saudis' contrast between word-final stops compared to the Americans' is clearly related to the absence of a duration contrast between final voiced and voiceless stops in Arabic. In Experiment 1 (marked as AR in Fig. 5) we found that the durations of the unflapped /t/ and /d/ of Saudi Arabian Arabic were about equal, while the closure interval of /g/ was actually somewhat *shorter* than that of /k/ in final position. Note that Saudis did not flap English /t/ or /d/. This is somewhat surprising in view of the recent finding by Port and Mitleb (1980) that speakers of Jordanian Arabic who had lived in the U.S. for about the same length of time as speakers in our group Ar2 flapped word-final post-stress alveolar stops (in phrases like "bat again") in a similar experimental context.

Just as for stops in word-initial position, the effect of place of articulation on the duration of final stops was significant for all three speaker groups ($p < 0.01$), the closure interval shortening as the place of articulation moved further back in the mouth.

Glottal pulsing. Both Americans (Am) and Saudis (Ar1, Ar2) produced the phonologically voiced stops /bdg/ with glottal pulsing, as seen in Table 5. (Note that the word-initial stops being analyzed were intervocalic since they occurred sentence-medially after the word *say.*) The native and non-native speakers of English differed, however, in their production of voiceless stops. The Americans (Am) generally kept the closure intervals of /ptk/ free of glottal pulsing (except for the normally flapped /t/ which we did not attempt to measure and have left out of the table). Both groups of Saudis (Ar1, Ar2), however, produced a larger percentage of /p/s with glottal pulsing than did the Americans in both word-initial and word-final position ($p < 0.01$). The relatively less experienced Saudi speakers of English (Ar1) produced /p/ with glottal pulsing more frequently than the Saudis (Ar2) who had lived for several years in the U.S. ($p < 0.01$).

The glottal pulsing observed during the closure interval of the Saudis' /p/s was stronger than the "edge" vibrations noted by Lisker and Abramson (1967) as can be seen in Fig. 1. Moreover, it was generally audible when isolated by electronic gating and would therefore probably contribute to the perception of these stops as voiced. The glottal pulsing we observed may have resulted from an insufficiently wide abduction of the vocal folds,

[5] *The unstressed syllable immediately after the keyword in the carrier sentence seems to have made the "word-final" stops of this study comparable to the "intervocalic" stops of Lisker's (1957) work.*

TABLE 5

Percentage of stops produced by three speaker groups with glottal pulsing visible through at least half the closure interval; "n" is number of tokens analyzed. The alveolar stops produced by group Am were not analyzed because these stops were generally produced as flaps

		/b/	/d/	/g/	/p/	/t/	/k/
				Initial Position			
Am	%	100	97	100	3	3	0
	n	36	36	35	36	36	36
Ar2	%	94	100	100	11	0	3
	n	36	36	36	36	36	36
Ar1	%	97	91	100	47	6	9
	n	36	35	36	36	35	36
				Final Position			
Am	%	97	–	94	0	–	0
	n	36		36	36		36
Ar2	%	97	97	94	39	0	3
	n	36	36	36	36	35	36
Ar1	%	97	97	76	83	0	3
	n	36	36	36	36	36	36

or else initiation of abduction which occurred too late to insure voicelessness during the closure interval of /p/ (see Weismer, 1980). We cannot be entirely certain, of course, that the Saudis' /t/ and /k/ were not also voiced because of the limited dynamic range of a sound spectrograph. But since only the Saudis' /p/ frequently exceeded our criterion, we can probably conclude that the Saudis' laryngeal control differed during their production of /p/ as compared to /t/ and /k/. Future research using other instrumental techniques should establish in greater detail whether glottal pulsing observed during a /p/ produced by Saudis differs from that seen in /b/ (and other stops) since this question bears directly on the issue of how speakers learn to control laryngeal timing during stop production.

We have seen that phonetic differences between Arabic and English seem to have a direct influence on Saudis' production of English stops. The question remains, however, whether the acoustic differences between stops produced by Americans and Saudis noted here – as well as other acoustic differences we have not examined – will lead to perceptual confusions for English-speaking listeners. The next experiment addresses this issue.

EXPERIMENT 3: INTELLIGIBILITY OF STOPS PRODUCED BY SAUDIS

The English experiment showed that the stop voicing contrast produced by Americans and Saudis differed along several phonetic dimensions. Some such differences might only contribute to the perception of foreign accent, while others might result in misperception. This experiment was designed to test our impression that many of the Saudis' intended /p/s were perceivable as /b/, even though they had been produced under fairly ideal conditions. This finding would not be surprising since English /p/ is widely considered to pose a "problem" for Arabs learning English by those who teach them (e.g., Aziz, 1974).

Methods

The English sentences produced by both groups of Saudis (Ar1, Ar2) were dubbed from the master tape onto listening tapes using a matched pair of Revox Model A700 tape recorders. Extraneous sounds and repeated utterances were deleted, and pauses inserted where necessary to yield 2.5 sec intervals between utterances. Care was taken to insure that any variation in signal strength on the original recording was equalized on the listening tapes.

Two randomizations of the test sentences were presented free-field to seven native American graduate students in linguistics, none of whom had studied Arabic. They were selected because of their experience in phonetics and in transcribing sounds of foreign languages. These listeners heard the tapes at a comfortable level while seated equidistant from an Advent loudspeaker in a quiet room. Although the listeners knew which English words the Saudis had intended to produce, they were instructed to transcribe any real or possible English word they heard.

A confusion matrix was prepared from listener responses to initial and final stops in minimal pairs (**p**at/**b**at, ta**p**/ta**b**; **t**ab/**d**ab, ba**t**/ba**d**; **c**ab/**g**ab, ba**ck**/ba**g**) representing 478 responses to each English stop (6 tokens, 12 speakers, 6 listeners plus 3 tokens each for a seventh listener who was interrupted after one randomization). Chi-square tests were performed to determine the effect of place of articulation, phonological voicing, position within the syllable, and speaker group on intelligibility.

Results

As seen in the confusion matrices in Table 6, the American listeners had difficulty identifying some stops produced by the Saudis. In word-initial position there were seven times as many confusions based on voicing than on place of articulation ($p < 0.01$). About 2/3 of the voice confusions were between /p/ and /b/. In this regard it is important to note that the expected confusion pattern for English speakers and listeners (Miller and Nicely, 1955) is to find many place-of-articulation confusions but relatively few confusions based on voicing.

We found twice as many confusions in word-final than word-initial position, a difference which was significant at the 0.01 level. Here, too, errors due to voicing greatly outnumbered those due to place ($p < 0.01$). And, finally, the relatively more experienced

TABLE 6

Percent identification of word-initial stops (above) and word-final stops (below) produced by 12 Saudi speakers as /ptkbdg/. Each stop was presented for identification 72 times to 7 listeners. Not included are 15 non-stop responses

Initial Position

Intended \ perceived	/p/	/t/	/k/	/b/	/d/	/g/
/p/	77%	1		22		
/t/	1	99				
/k/		1	95			4
/b/	10			89		
/d/	5	3			93	
/g/			6			93

Final Position

Intended \ perceived	/p/	/t/	/k/	/b/	/d/	/g/
/p/	50%	1		49		
/t/		86	10		2	
/k/		4	93			3
/b/	22			78		
/d/		6		1	90	2
/g/			4	1	8	86

Saudi speakers of English (Ar2) produced fewer /p/s that were heard as [b] (16%) than did the less experienced Saudis (36%) ($p < 0.01$).

Discussion

The confusion of /p–b/ is readily interpretable in terms of acoustic measurements made in the English experiment. The /p/ in *pat* produced by the Saudis may have often been heard as [b] both because its VOT was very short and because glottal pulsing frequently occurred during the closure interval. Word-initial /t/ and /k/ were probably seldom identified as voiced stops because they were not produced with glottal pulsing and because the VOT of /t/ and /k/ was proportionally closer to English values than that of /p/. It is surprising that the Saudis' initial /b/ was sometimes heard as [p] since it was nearly always produced with glottal pulsing through the entire closure interval and without aspiration at stop release, both of which should support perception of a *voiced* stop by American listeners. Perhaps the American listeners, hearing too many /b/s, randomly identified some as /p/.

In final position several acoustic dimensions seem to have led to confusions of /p/ and /b/. The relative shortness of the vowel in *tab* produced by the Saudis probably led American listeners to hear some of the final stops in that word as [p] (see, e.g., Raphael, 1972). Both the frequent presence of glottal pulsing during the closure interval and the lack of a temporal contrast between final /p/ and /b/ probably led the American listeners to hear some of the final stops of *tap* as [b].

We cannot be certain, of course, that the acoustic dimensions we examined in the English experiment are alone responsible for these perceptual confusions, nor adequately assess the effect on intelligibility of the acoustic differences we discovered between the Saudis' and Americans' stops. Still, it seems likely that both glottal pulsing and the articulatory timing variables we noted did contribute to a deficit in intelligibility. Moreover, this experiment verifies the existence of a serious intelligibility problem for /p/ and /b/ produced by Arab learners of English, as would be predicted by a contrastive analysis (Lado, 1957).

GENERAL DISCUSSION

Phonetic level interference

Our results show that phonetic implementation of stop voicing in Saudi Arabian Arabic directly influenced Saudis' production of English stop voicing. Having found a small and non-significant effect of stop voicing on preceding vowel duration in Arabic, we observed a correspondingly small effect in Arabic-accented English. The Saudis produced a duration contrast between the closure intervals of word-initial voiced and voiceless stops in English similar to that observed in Arabic, but the Saudis tended *not* to produce a similar duration contrast between word-final English stops. This seemed to reflect the lack of such a contrast in Arabic. Finally, the Saudis' VOT for English stops was shorter than the VOT values produced by Americans, corresponding fairly closely to Arabic VOT values. Thus the Saudis were clearly generalizing temporal correlates of stop voicing found in Arabic to their production of English stops.

This means we cannot account for foreign accent strictly in terms of segmental phonemic or phonetic feature differences between languages (see Kenstowicz and Kisseberth, 1979, p. 154), for even non-segmental differences in temporal implementation carry over from one language to another. Since the temporal specification of speech sounds can apparently vary in unpredictable ways from language to language (Lehiste, 1970; Port *et al.,* 1980; Keating, 1979; Kohler, 1979), they must be learned and, in this sense, be considered part of the linguistic knowledge of speakers. Yet it is often assumed that the linguistic control of speech is restricted to segmental units of phonetic transcription (Chomsky and Halle, 1968; cf. Lisker and Abramson, 1971; Lisker, 1974). Our results tend to undermine this notion of a linguistic phonetic space restricted to a fixed universal set of segmental elements.

The special case of /p/

The Saudis' inability to produce a good English /p/ provides further evidence of interference at the level of phonetic implementation. Both groups of Saudis made the closure interval of /p/ in *pat* much longer than that of /b/ in *bat,* even though no such duration contrast was produced by Americans and no such /p–b/ contrast exists in Arabic. This contrast in Arabic-accented English seems to represent an **extrapolation** by the Saudis of the duration contrast between word-initial /t–k/ and /k–g/ which exists in Arabic to the /p–b/ contrast of English. For word-final stops we found that the more experienced Saudi speakers of English (Ar2) produced a small duration contrast between /bdg/ and /ptk/, while the less experienced Saudis (Ar1) did not. Thus the Saudis produced qualitatively similar (non-English) duration contrasts between *all three* stop voicing pairs (/p–b/, /t–d/, /k–g/) even though the first doesn't exist in Arabic and the second is ordinarily neutralized by flapping in American English. The fact that the Saudis generalized a pattern of temporal correlates of stop voicing evenly across all three places of articulation suggests, first, that they recognized the functional similarity of the voicing contrast in English and Arabic and, second, were aware that /p/ is equivalent to a "voiceless /b/" or a "labial /t/." That is, they seemed to treat voicing and place as commutable at a featural level.

Although the Saudis generalized the Arabic pattern of timing of articulatory closure to a new segment (/p/), the same cannot be said of their control of glottal pulsing. The difference in glottal pulsing that characterized the Saudis' contrast between /t–d/ and /k–g/ (both in English and Arabic) was not clearly present in the Saudis' English /p–b/ contrast (many of their /p/s were produced with glottal pulsing during the closure interval). Several studies of first and second language also suggest that it may be more difficult to learn to control a new pattern of glottal-supraglottal timing than one involving purely supraglottal timing. Smith (1979) reports that American children as old as 4½ produced a clearer contrast in duration than glottal pulsing between post-stress voiced and voiceless stops (cf. Kewley-Port and Preston, 1974). Fischer-Jørgensen (1968) found that an adult French-Danish bilingual contrasted voiced and voiceless stops in French and Danish by means of language-specific temporal correlates of stop voicing that are appropriate to those two languages. However, this speaker did *not* maintain a phonetic distinction between her French and Danish stops with respect to glottal pulsing.

She produced /bdg/ in French without glottal pulsing during the closure interval as if they were *Danish* voiced stops. Finally, data reported by Suomi (1976) suggest that Finnish learners of English succeeded better in learning temporal correlates of the English stop voicing contrast that do not appear in Finnish than in learning to contrast English stops by means of glottal pulsing. It may be, then, that coordination of laryngeal control with particular supraglottal articulatory gestures is an especially difficult articulatory skill for both first and second language learners to acquire.

CONCLUSIONS

Although the universality of a voicing effect on preceding vowels almost has the status of a truism of phonetic timing, it cannot be maintained as a strong universal. There may be a tendency for vowels to be relatively longer before voiced consonants than before voiceless consonants, but such a contrast is not uncontrollable by phonological factors. It is possible that Arabic and English differ in the internal structure of their syllables in ways that result in language-specific timing differences, such as the voicing effect on the duration of closure intervals in initial v. final stops.

Results of these experiments have important implications for a theory of interference in second language acquisition. A difference in the phonemic inventories of Arabic and English did not seem to be the principal cause of the Saudis' difficulty in producing a perceptually effective English /p/. The timing of the labial articulation of /p/ was just what one would expect if the Saudis were producing a voiceless /b/, a finding which demonstrates their awareness of the phonological and phonetic features of /p/. The Saudis' primary difficulty was in adjusting the glottis in such a way as to prevent glottal pulsing from occurring during the closure interval of /p/. Although this instance of non-commutability suggests an interdependence between features (those defining voicelessness and labiality), it would be viewed by many as part of the implementation rules applied to a matrix of phonetic features, and thus peripheral to the phonetic segments themselves. Temporal correlates of the stop voicing contrast produced by Saudi Arabians exhibited –even after several years in an English-speaking environment – only a modest amount of modification in the direction of the English pattern of phonetic implementation (cf. Flege, 1980). Such timing effects are also currently considered to be a part of a sub-segmental level of phonetic implementation. Thus, our conclusion must be that the most important interference from a first to a second language during the process of foreign language acquisition occurs at the level of phonetic implementation rather than at an abstract level of organization based on features.

REFERENCES

AL-ANI, S. (1970). *Arabic Phonology* (The Hague).

AZIZ, Y. (1974). Some problems of English consonant sounds for the Iraqi learner. *English Language Teaching Journal*, **28**, 166-168.

BARRY, W. (1974). Language background and the perception of foreign accent. *Journal of Phonetics*, **2**, 65-89.

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

CHOMSKY, N. and HALLE, M. (1968). *The Sound Pattern of English* (New York).

CLUMECK, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, **4**, 337-351.

ELERT, C.-C. (1964). *Phonologic Studies of Quantity in Swedish* (Stockholm).

FISCHER-JØRGENSEN, E. (1964). Sound duration and place of articulation. *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung*, **17**, 175-207.

FISCHER-JØRGENSEN, E. (1968). Les occlusives françaises et danoises d'un sujet bilingue. *Word*, **24**, 112-152.

FLEGE, J. (1979). Phonetic interference in second language acquisition. Ph.D. dissertation, Indiana University.

FLEGE, J. (1980). Phonetic approximation in second language acquisition. *Language Learning*, **30**, 117-134.

HOUSE, A. and FAIRBANKS, G. (1953). The influence of consonant environment upon the secondary characteristics of vowels. *Journal of the Acoustical Society of America*, **25**, 105-113.

HUGGINS, A. (1976). Speech timing and intelligibility. In J. Requin (ed.), *Attention and Performance, VII* (Hillsdale, N.J.).

JAKOBSON, R. (1962). On the identification of phonemic entities. In *Roman Jakobson, Selected Writings, I* (The Hague), p. 418-426.

JONASSON, J. and MCALLISTER, R. (1972). Foreign accent and timing: An instrumental phonetic study. *Papers from the Institute of Linguistics, University of Stockholm*, **14**, 11-40.

KEATING, P. (1979). A Phonetic Study of a Voicing Contrast in Polish. Ph.D. dissertation, Brown University.

KENSTOWICZ, M. and KISSEBERTH, C. (1979). *Generative Phonology, Description and Theory* (New York).

KEWLEY-PORT, D. and PRESTON, M. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, **2**, 195-210.

KLATT, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, **59**, 1208-1221.

KOHLER, K. (1979). Dimensions in the perception of fortis and lenis plosives. *Phonetica*, **36**, 332-343.

LADO, R. (1957). *Linguistics Across Cultures: Applied Linguistics for Language Teachers* (Ann Arbor).

LEHISTE, I. (1970). *Suprasegmentals* (Cambridge, Mass.).

LISKER, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language*, **33**, 42-49.

LISKER, L. (1974). On time and timing in speech. In T.A. Sebeok, A.S. Abramson, D. Hymes, H. Rubenstein, E. Stankiewicz and B. Spolsky (eds.), *Current Trends in Linguistics, 12* (The Hague), pp. 2387-2418.

LISKER, L. (1978). *Rabid* vs. *rapid:* A catalogue of acoustic features that may cue the distinction. *Haskins Laboratories Status Report on Speech Research*, SR-54.

LISKER, L. and ABRAMSON, A.S. (1964). A cross-language study of voicing in initial stops: Acoustical measurement. *Word*, **20**, 384-422.

LISKER, L. and ABRAMSON, A.S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, **10**, 1-28.

LISKER, L. and ABRAMSON, A.S. (1971). Distinctive features and laryngeal control. *Language*, **47**, 767-785.

MILLER, G. and NICELY, P. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, **27**, 338-352.

MONSEN, R. (1976). The production of English stop consonants in the speech of deaf children. *Journal of Phonetics*, **4**, 29-41.

PETERSON, G. and LEHISTE, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, **32**, 693-703.

PORT, R., AL-ANI, S. and MAEDA, S. (1980). Temporal compensation and universal phonetics. *Phonetica*, **37**, 235-252.

PORT, R. and MITLEB, F. (1980). Phonetic and phonological manifestations of the voicing contrast in Arabic-accented English. *Research in Phonetics, 1* (Dept. of Linguistics, Indiana University), 137-165.

RAPHAEL, L. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America*, **51**, 1296-1303.

SMITH, B. (1979). A phonetic analysis of consonant devoicing in children's speech. *Journal of Child Language*, **6**, 19-28.

STATHOPOULOS, E. and WEISMER, G. (1979). The duration of stop consonants. In J. Wolf and D. Klatt (eds.), *Speech Communication Papers: 97th Meeting of the Acoustical Society of America* (New York), pp. 197-201.

SUOMI, K. (1976). English Voiced and Voiceless Stops as produced by Finnish and Native Speakers. *Jÿvaskkyla Contrastive Studies, 2,* (University of Jÿvaskkyla, Finland).

WEINREICH, U. (1953). *Languages in Contact: Findings and Problems* (The Hague).

WEISMER, G. (1980). Control of the voicing distinction for intervocalic stops and fricatives: Some data and theoretical considerations. *Journal of Phonetics*, **8**, 427-438.

YENI-KOMSHIAN, G., CARAMZZA, A. and PRESTON, M. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, **5**, 35-49.