The Phonological Basis of Foreign Accent: 
A Hypothesis*

James Emil Flege

Foreign accent is often thought to be the result of an age-related diminution in the ability to learn to pronounce languages. Existing studies of L2 pronunciation, however, do not seem to support the claim that there is some fundamental difference between children and adults in phonetic learning ability. The continued presence of foreign accent may instead be a consequence of the establishment of stable phonological representations for sounds and words in the native language. Language learners who perceive sounds in the target language to be phonologically identical to native-language sounds (despite possible phonetic differences between the two languages) may base whatever phonetic learning that does occur during the acquisition process on an acoustic model provided by pairs of similar sounds in two languages, rather than on a single language-specific acoustic model as in first-language acquisition. Thus an adult or child learner of a foreign language may retain the same kind of phonetic learning ability evident in early childhood and yet still speak with an accent because phonological translation provides a two-language source of phonetic input that may ultimately limit progress in learning to pronounce a foreign language.

It has been observed that adults learning a foreign language "never seem capable of ridding themselves entirely of foreign accent" (Scovel 1969:245) whereas most normally developing children "learn to recognize and pronounce the sounds of . . . their speech community so well that . . . their speech lacks any trace of the foreign accent of people who learn the language later" (Ferguson and Garnica 1975:154). The observation that children and adults differ fundamentally in their ability to learn to pronounce a foreign language has led to the formulation of a critical period hypothesis. The continued presence of foreign accent in the speech of post-pubescent language learners has often been specifically linked to neurophysiological maturation and the establishment of cerebral lateralization for language functions (Penfield and Roberts 1966, Lenneberg 1967, Scovel 1969). According to a critical period hypothesis for second language learning, adults usually cannot learn to speak a foreign language without accent because the central nervous system undergoes some permanent reorganiza-

* The author would like to thank Roy Koenigsknecht for comments on a previous version of this paper. This work was supported by NINCDS grant NS07100 to the Department of Communicative Disorders at Northwestern University.

Mr. Flege is NIH Post-doctoral Fellow in the Department of Communicative Disorders at Northwestern University. He has published in Language and Speech and Language Learning on second language acquisition.
tion at about puberty, thereby distinguishing adults and adolescents from younger language learners.

In a recent phonetic study of English pronunciation by non-native speakers, Flege (1980) noted a direct influence of phonetic characteristics of Arabic on the English stops produced by Saudi Arabians. Even the more experienced of two adult speaker groups (3 years residence in America) continued to produce stop consonants much as if they were Arabic sounds. There was, however, some evidence of phonetic learning. The more experienced Saudis produced a durational contrast between word-final /p-b/, /t-d/, and /k-g/. However, the magnitude of the durational contrast produced by the Saudis was much smaller than the one produced by Americans, seemingly because of the absence of a similar duration contrast between word-final stops in the Saudis' native language. Moreover, it was noted that experienced Saudi speakers of English more nearly approximated native English pronunciation than did newly arrived Saudis. For both groups, however, it seemed that phonetic implementation of the stop voicing contrast was characteristic neither of Arabic nor of English. Instead, values for the phonetic parameters fell between those measured for stops in Arabic and stops in English.

Findings such as this raise two questions about how adults learn a second language. First, why did the Saudis not simply use Arabic sounds when producing English words if, in fact, they had passed a critical period for learning to pronounce foreign languages? Second, since the Saudis did modify their pronunciation of English stops, why wasn’t their modification complete? In other words, why did the Saudis’ production of English stops seem to represent a compromise between the phonetic characteristics of Arabic and those of English?

It is unlikely that small phonetic differences such as those noted between the Saudis and Americans would in themselves be sufficient to cause the perception of segmental substitutions by American listeners. Yet such phonetic differences are of interest because they may contribute to what is perceived as a foreign accent. Moreover, an understanding of factors that tend to limit the extent of progress in foreign language pronunciation at any level of analysis may offer insight into why language learners often seem to show a distinct limit on the extent to which L2 pronunciation ordinarily improves (Selinker 1972, Nemser 1971).

In this article we shall consider foreign language pronunciation from several perspectives. First, we will discuss some of the dimensions that may form the acoustic basis of foreign accent. Second, we will review evidence concerning the claim that children and adults differ fundamentally in terms of phonetic learning ability. And, third, we will propose an alternative to the critical period hypothesis to account for the continued presence of a foreign accent. The phonological translation hypothesis proposed here starts...
with the assumption that neither physiological maturation nor neurological reorganization renders an adult incapable of speaking a foreign language without an accent.

1. Foreign Accent

Perception of a foreign accent derives from differences in pronunciation of a language by native and non-native speakers. The most readily apparent basis for a foreign accent are mispronunciations that lead to the perception of a segmental sound substitution, such as in French-accented I sink so or Arabic-accented I put my car in the barking lot. One recent study indicated that the frequency with which segmental substitutions were noted in short excerpts of speech produced by non-native speakers was highly correlated with native-speaker judgments of accentedness (Brennan, Ryan, and Dawson 1975).

However, this does not necessarily mean that perception of a foreign accent is based just on overtly detectable mispronunciations of sound segments. Listeners are more likely to base a judgment of foreign accent on some combination of segmental, subsegmental, and suprasegmental differences which distinguish the speech of native from that of non-native speakers. Some mispronunciations which depart from target-language phonetic norms may merely sound distorted, such as the underaspirated /ptk/ that can often be heard in the English produced by native speakers of Romance languages. An important question is whether every audible acoustic difference is weighted equally when a foreign accent is perceived. One study (Flege and Hammond 1980) suggests that at least not all segmental sound substitutions produce a similar effect. Monolingual English speakers were asked to try to read English sentences with a Spanish accent. Their imitations of Spanish-accented English included a number of the sound substitutions one might associate with a stereotypic Spanish accent. Some sound substitutions were produced much more frequently than others, suggesting that they may figure prominently in what Americans perceive as a Spanish accent in English.

2. Cross-language phonetic differences

Children not only learn to produce sounds intelligibly in their native language, they also learn to produce them according to the language-specific phonetic norms of the surrounding speech community. The monolingual child’s attainment of accent-free speech in the course of normal speech development is by no means a trivial achievement, for human languages manifest considerable variation at the phonetic level. Language-specific phonetic characteristics must be acquired by the child if s/he is to sound like a native speaker. At the same time, the many phonetic differences which distinguish two languages represent a potential source of foreign accent in the speech of a second language learner since interference
during second language acquisition appears to be prominent at the level of phonetic implementation (Flege and Port 1981).

Some phonetic differences between languages may be localized at the level of phonetic segments. The fact that phonetically similar sounds in two languages might be transcribed with the same IPA symbol should not obscure the fact that sounds may be realized differently at the phonetic level. For example, the Danish vowel /i/ is higher than its phonological counterpart in English. Both Hausa and Kalabari, to take another example, have a voiced bilabial implosive in the phonetic surface but the same sound is produced with a very different kind of voicing in the two languages (Ladefoged 1980). Vowels preceding a nasal consonant may be nasalized to a greater extent in some languages than others (Clumeck 1976).

Phonetic differences between languages that may potentially contribute to accentedness may be suprasegmental. For example, average speaking fundamental frequency may differ cross-linguistically (Majewski, Hollein, and Zalewski 1972) as may the fundamental frequency variations which underlie similar intonation contours in two language (Willems 1978).

Cross-language differences may also be thought of as subsegmental, such as the speech timing differences between languages which together affect the perceived rhythmic qualities of speech and may carry over from the native to the target language. For example, the extent to which vowels are lengthened at the end of an utterance may be much greater in some languages than others (Delattre 1966, Oller 1977). In languages where the voicing characteristic of a consonant influences the duration of a preceding vowel, the relative magnitude of this effect appears to differ substantially across languages (Chen 1970). Even the phonetic domain within which successive phonetic segments have a temporal influence on one another seems to vary across languages. In English, for example, a consonant affects only the duration of a preceding vowel, while in Japanese the duration of a consonant seems to affect the duration of both preceding as well as following vowels (Port, Al-Ani, and Maeda 1980).

3. Child vs. adult phonetic learning

Considering the many phonetic differences that might distinguish the learner’s native language from a foreign language, one might wonder if any learner of a second language—child or adult—will manage to pronounce a foreign language completely without accent. At this point, we simply don’t know if a speaker can be bilingual at the phonetic level, or if native-like pronunciation at the level of phonetic implementation is even necessary for accent-free speech. If children are more capable than adults of learning to correctly pronounce foreign languages, they will of course be more likely than adults to escape the onus often associated with ac-
Foreign Accent

Foreign Accent

447

centedness. Existing empirical evidence, however, does not seem to support the notion that there is a true qualitative difference in how well children as opposed to adults pronounce foreign languages. A foreign accent can sometimes be detected in the speech of children (Vallette 1964, Asher and Garcia 1969). At the same time, some adults appear capable of producing a foreign language without accent (William 1980, Neufeld 1980).

Several studies have investigated how English native speakers rate the English produced by individuals who began learning English as a foreign language at various ages. Some such studies seem to indicate that the speech of someone who begins learning a foreign language at an early age will be perceived as less accented than that of someone who begins learning at a relatively later age (Asher and Garcia 1969, Fathman 1975, Oyama 1976). But at least one study indicates that older children and adults may be, at least initially, more successful than young children in pronouncing a foreign language (Snow and Hoefnagel-Hohle 1978). Still other studies indicate that older children and adults can imitate words in an unfamiliar foreign language better than young children, and that ability to discriminate unfamiliar sounds in a foreign language—seemingly a prerequisite for the development of correct pronunciation—may actually improve rather than diminish with age (Politzer and Weiss 1969, Snow and Hoefnagel-Hohle 1977, Winitz 1981). Finally, it appears that pronunciation of a foreign language by both adolescents and adults will improve with additional exposure (Asher and Garcia 1969, Snow and Hoefnagel-Hohle 1977, 1978), at least within limits (Oyama 1976).

Thus at present there does not appear to be evidence of a fundamental difference between children and adults in their ability to learn to pronounce a second language. There will admittedly be instances where a child's accent appears to diminish more rapidly or thoroughly than an adult's. However, such disparities might be due to differences in linguistic input (Burling 1981), cultural expectations (Hill 1970), or to the period of

1In several experiments, listeners have been asked to evaluate the personality of a speaker after listening to short samples of speech. Unknown to the listeners in such evaluational experiments, a single talker has produced a single speech sample twice, once in an accented guise and another time with no accent (Anisfeld, Bogo, and Lambert 1962, Arthur, Farrar, and Bradford 1974). In other experiments, speech samples produced by larger numbers of both native and non-native speakers have been evaluated (Ryan and Carranza 1975). The consistent finding of such studies seems to be that a person whose speech is accented will be rated less favorably (along subjective rating scales such as intelligent vs. stupid or kind vs. cruel) than speakers who talk without a foreign accent. Other studies have established the validity of using such rating scales to measure subjective reactions to a complex phenomenon like accentedness. Measures of accentedness resulting from psychophysical scaling techniques (magnitude estimation, sensory modality matching) show a high correlation with measures derived from rating scales. Such studies indicate that not only are native speakers without special phonetic training quite capable of distinguishing native from non-native speakers, but also that they can reliably distinguish between degrees of accentedness (Brennan, Ryan, and Dawson 1975, Ryan, Carranza, and Moffie 1977, cf. Giles 1972). Unfortunately, for most foreign language learners, the greater the degree of perceived accent, the more unfavorably a speaker will be evaluated.
time over which pronunciation continues to show improvement for learners of different ages (Snow and Hoefnagel-Höhle 1977) rather than neurophysiological maturation or reorganization.

4. Simultaneous vs. successive learning

Perhaps the best way to learn two languages without accent is to learn both in early childhood. There is some tentative evidence from single subject observation to suggest that young children who acquire two languages simultaneously before about age 3;0 may produce both languages with a native-like accent (McLaughlin 1978), at least insofar as judged by the standards of articulation that adults apply to child speech. For example, a two-year-old exposed to English and German produced the word ball (Ball) differently in those two languages. In English she said [baU] and in German [baI], thereby approximately the dark word-final [1] of English and the light (1) of German (Leopold 1947). Another two-year-old, who was learning English and Portuguese, produced voiceless stops with aspiration when speaking English but without aspiration in Portuguese (Major 1977). Given the scarcity of empirical evidence, it would be premature to assume this is typical of early childhood bilingualism. But if such a pattern is characteristic, it might provide some insight into why it seems young bilingual children are more apt than others to pronounce both languages with native-like accents. Perhaps such children learn the sounds occurring in the phonetic surface of their two languages independently, much as if each sound were a separate phoneme in an enriched, pan-language system.

5. The phonological translation hypothesis

It may perhaps be more useful to ask why anyone should speak a foreign language with an accent than to ponder the source of possible differences in degree of accent between children and adults. One possible approach to this issue would be to determine why children acquiring a second language before about age 3;0 may be more likely than older learners to speak the second language without accent.

One possibility is that there is a critical period for phonetic learning and that it occurs long before puberty (Krashen 1973). If so, age rather than simultaneity of learning is the crucial factor. However, a more probable cause, in my estimation, is the important difference in phonological development between the young child who simultaneously learns two languages and the older child or adult who begins learning a second language after the establishment of the first. One important difference may be that only relatively mature speakers are apt to interpret sounds in a second language in terms of sounds found in another language.

I would like to propose that a tendency by mature speakers to interpret sounds occurring in a foreign language in terms of sounds found in their native language may be a more important cause of foreign accent than
any limitation on phonetic learning imposed by neurophysiological matura-
tion (Lenneberg 1967). It has been observed that before about the age of
three bilingual children often do not fully realize they are speaking two
different languages (McLaughlin 1978), a confusion which is most readily
apparent from language mixing at the lexical level (Redlinger and Park
1980). Moreover, young pre-literate children may not yet be capable of
reliably identifying phone-sized units in speech (Morais, Cary, Alegría,
and Bertelson 1979). Somewhat older children, however, clearly do distin-
guish between languages and seem to be aware of the similarity of sounds
found in two languages (see Tervoort 1979). For example, a young
American boy speaking French used English words spoken with a French
accent when he didn’t know the French equivalent (Valette 1964). English-
speaking children have been observed trying to “speak” Spanish by produc-
ing English words with a Spanish accent (Hernández-Chávez, cited in
Ervin-Tripp 1974). American children exposed to French for only a month
showed in an informal experiment that they were able to produce English
words with French sounds (Ervin-Tripp 1974). This kind of linguistic be-

The phonological translation hypothesis can be illustrated by consider-
ing how French native speakers produce English words. **VOT (voice onset
time)** is a well known acoustic measure of the laryngeal timing differen-
tes which may serve to distinguish a voiceless stop (like /p/) from its voiced
homorganic cognate (/b/). A VOT difference between voiced and voice-
less stops can be observed in the speech of English-learning children as
early as age two (Macken and Barton 1980). Figure 1 displays VOT
values reported by Caramazza for stops in isolated English words produced
by monolingual and bilingual speakers of English and French. Children
acquiring English as a first language will eventually learn to produce
stops in a similar phonetic context with VOT values like those seen here
for adult English monolinguals in Figure 1a (about 60-90 msec). Children
learning French, in contrast, will eventually produce stops with the rela-
tively short VOT values (0-30 msec) typical of French, such as those seen
here for French Canadian adults in Figure 1d.

Since corresponding French and English stops are realized differently
at the level of phonetic implementation, how will a French Canadian
produce English stops? Previous research and common experience tell us
that the English produced by French Canadians is likely to be accented.
Moreover, there is reason to think that subsegmental phonetic differences
such as the VOT difference between English and French stops may be
detected by listeners experiencing a foreign accent (Flege and Hammond
1981). Thus we might expect the VOT values in English words produced
FIGURE 1
Mean VOT in msec of /ptk/ in English and French words produced by 
(A) monolingual English speakers; (B) French Canadian speakers of 
English; (C) bilingual French Canadians; (D) monolingual French 
Canadians; and (E) monolingual French speakers in France.

This figure is based on data reported by Caramazza, Yeni-Komshian, Zurif, 

by French Canadians to either match or resemble the VOT values measured in phonologically similar French words. In fact, as seen in Figure 1b, it appears that French Canadian speakers produce English words with VOT values between those typical of English and French.

Two aspects of these data are important to the present discussion. First, the French Canadians produced English words with shorter VOT values than monolingual English speakers. And second, although the English stops of French Canadians were not completely English-like, they were nonetheless different from the stops produced in French words by these same speakers.

These data can be interpreted in two radically different ways. One can infer that the French Canadians, who began learning English at about six, did not possess the same phonetic learning ability as somewhat younger children acquiring English as a first language. But one can instead interpret these data to mean that the French Canadians did not differ from young
children in terms of phonetic learning ability. The French-Canadian bilinguals studied by Caramazza et al. may have produced English stops with smaller VOT values than the English monolingual precisely because the capacity of older children and adults for learning to pronounce languages remains essentially unchanged from early childhood.

It seems reasonable to think that the French-speaking child who is exposed to English will consider English /p/, for example, to be the same as the /p/ found in French words despite VOT and other acoustic differences between French and English stops. Phonologists since Trubetzkoy have recognized that a listener interprets sounds in a foreign language in terms of phonological categories of the native language (see Lisker 1978). One important consequence of phonological translation may be that a French-speaking child (or adult) learning English as a second language will base production of English /p/ on the acoustic phonetic model provided by voiceless bilabial stops occurring in both English and French words.

Recall that the French Canadians pronounced /p/ differently in English and French. If we assume that the acoustic model upon which they based production of English was just the acoustic substance of sounds in English words, we must necessarily conclude that their phonetic learning was only partially successful. This would of course imply that these language learners differed in some important way from somewhat younger children acquiring English as a first language. But if, on the other hand, the French Canadians based their production of English words on a composite of corresponding sounds (and other nonsegmental phonetic dimensions) occurring in both their languages, one might reasonably conclude that their phonetic learning was completely successful.

The VOT of English stops produced by the French Canadians was intermediate between monolingual French and English values. This seeming compromise between French and English patterns of phonetic implementation may reflect a restructuring of the phonetic space so that it encompasses two languages (Williams 1980). Such a convergence can be described in Piagetian terms as the accommodation of existing French patterns of stop production to the English pattern, together with an assimilation of the English pattern of stop production to the French Canadians' French phonetic pattern.

6. Testing the hypothesis

It has been postulated here that children and adults possess the same general capability for learning to pronounce foreign languages and that one important cause of foreign accent is phonological translation between languages by speakers who already speak a first language. According to the phonological translation hypothesis, an individual may be completely successful in his/her phonetic learning of a second language and yet still
retain an accent because pronunciation of the foreign language is based on pairs of corresponding sounds (or non-segmental phonetic dimensions) found in the native and target language. The hypothesis rests on the assumption that both children and adult language learners modify native-language patterns of phonetic implementation, and that superordinate acoustic models based on pairs of corresponding sounds or phonetic dimensions in two languages serve as input for phonetic learning in second language acquisition. This hypothesis seems to be consistent with the results of existing phonetic studies showing that the values of phonetic parameters measured in the speech of second language learners are often intermediate to those typical of monolingual speakers of the native and target language (Pinkerton 1972, Suomi 1976, Niemi 1979, Flege 1980, Port and Mitleb 1980, Elsendoorn 1980, Williams 1980, Flege and Port 1981, Mitleb 1981). However, further investigation is clearly needed to evaluate the importance of phonological translation as a mechanism leading to foreign accent.

The hypothesis as it is formulated leads to two predictions through which it is potentially falsifiable. First, it predicts that bilingualism is not possible at the phonetic level since it posits that pronunciation of a foreign language is, for most speakers, eventually shaped to conform to an acoustic model provided jointly by the native and foreign languages. An individual might learn to speak a foreign language without apparent accent, but a fine-grained acoustic analysis should reveal differences between the language learner and native speakers of the target language along those parameters where phonetic differences exist between the native and target language. The expected pattern is for the language learner to produce foreign language sounds in such a way that they are phonetically intermediate to those found in the native and target languages.

Second, one would expect phonetic learning in a second language to affect pronunciation of a learner's first language. If the acoustic model provided by native language sounds can influence foreign language pronunciation, then foreign language sounds should, by virtue of their identification with native language sounds, exert an influence in the opposite direction. The extent of such influence might be determined by such factors as the discriminability of cross-language phonetic differences, an individual's ability to transform an auditory image abstracted from an acoustic model into a stable articulatory representation, or the amount of use and/or exposure to the foreign language (see Rosen 1979). Some support for this prediction is evident in Figure 1, which shows that the VOT values in French words produced by French-English bilinguals are slightly less French-like (i.e., longer, in the direction of English values) than are stops.

2 We have not discussed here the perceptual modification which may occur as the result of exposure to a foreign language. There is evidence to suggest that both discrimination (MacKain, Best, and Strange 1980) and identification (Williams 1980) of foreign language speech sounds may change during second language learning. Such changes might well influence speech production.
produced by French monolinguals in either Canada or France. Somewhat more substantial support for this prediction is provided by Williams' (1980) study of Puerto Rican children learning English. Williams found that native Spanish-speaking children learned to produce English stops with increasingly large (i.e., English-like) VOT values and also to produce Spanish stops more like English stops (i.e., with more aspiration) than monolingual Spanish speakers.

The hypothesis proposed here will not, of course, provide an explanation for cross-language interference where translation pairs of sounds or non-segmental dimensions in two languages cannot be identified by theoretical or empirical means. Nor does it attempt to account for the carry-over of phonological processes from the native to the target language. However, it does provide an explanation for the important interference that exists at the level of phonetic implementation during second language acquisition (Flege and Port 1981). If it is substantiated by further research, the phonological translation hypothesis offered here may provide insight into why learners typically continue to show interference at the phonetic level when other aspects of their control of L2 are more nearly native-like.

REFERENCES


