

THE PHONETIC STUDY OF BILINGUALISM¹

James Emil Flege

Individuals who learn a second language (L2) seldom manage to negotiate the full range of differences between the sound systems of their native language (L1) and that of the second language if they begin learning the L2 beyond early childhood. Most speech errors in an L2 can be traced to differences in the inventory of sounds used in the L1 and L2, or to differences in the possible sequences in which those sounds can be arranged to form syllables and words. Also, many of the position sensitive allophones (or "sounds") that any two languages might be said to "share" are likely to differ in their phonetic realization. The nature of the phonetic and phonological differences that distinguish any L1-L2 pair is principally responsible for the kind of foreign accent one will hear. A German accent in English, for example, is very different from an Italian accent because English differs from German and Italian in different ways.

Listeners, once they have acquired an L1, are language-specific perceivers of speech. In early stages of L2 acquisition, at least, bilinguals process L2 phonetic segments through the grid of their L1 phonology. Through a process called "interlingual identification", words heard in the L2 are decomposed into familiar L1 sounds, then produced accordingly. L2 phonology is thus like new wine in old bottles. For example, French /y/ may be produced as /i/ by native Portuguese

Ilha do Desterro	Florianópolis	nº 35	p.017-026	jul./dez. 1998
------------------	---------------	-------	-----------	----------------

speakers, but as /u/ by native English speakers. This is because Portuguese speakers identify French /y/ with Portuguese /i/ whereas English speakers typically identify French /y/ with English /u/. The influence of L1 phonology on L2 production may even be present when one does not hear segmental substitutions. For example, Spanish speakers may produce all four segments of the English word *taco* with Spanish phonetic characteristics (Flege and Munro 1994). Among other things, they are apt to produce /t/ with voice onset time (VOT) values that are too short for English (Spanish /t/ is realized with short-lag VOT), and to produce /A/ with F2 values that are too high (because Spanish /a/ may be used in place of English /A/).

A question of great interest is whether segmental substitutions, and phonetic errors not resulting in perceived sound substitutions, arise from an inability to motorically realize sounds not found in the L1, or whether they have a perceptual basis. It seems reasonable to think that L2 learners will not generally produce what they do not perceive. One might hypothesize that at least certain L2 production errors arise because L2 learners fail to note perceptually certain phonetic differences that distinguish contrastive sound units of the L2, or which distinguish L2 sounds from sounds in the L1. As the result of pre-attentive processes established in childhood, learners of an L2 might simply ignore phonetic differences not directly relevant to contrasts between sounds in the L1.

Fortunately for those who wish to learn an L2, segmental phonetic perception appears to remain malleable across the life span, at least to some extent. A few hours of laboratory feedback training is sufficient to improve the perception of certain unfamiliar L2 speech sound contrasts. Perception also changes slowly during the course of unaided, naturalistic L2 acquisition. At the same time, there is evidence that nearly everyone who learns an L2, adults included, make progress in producing the sounds of their L2. For example, Spanish adults who at first produce English /t/ with Spanish-like short-lag VOT values usually come to produce English /t/ with longer VOT values as they

gain experience in English. Some eventually produce English stops with VOT values that are fully appropriate for English. Others stop short, producing stops with VOT values that are intermediate to those typical for English and the L1. Taken together, such findings raise the following questions: (1) How, and to what extent, does the perceived relation of sounds found in the L1 and L2 change as bilinguals gain experience in the L2? (2) How are L2 sounds represented? As variants of an existing L1 sound, or as separate phonetic entities? And (3) What is the impact of perceptual changes on the *production* of L2 sounds?

An aim of much recent research has been to identify and evaluate factors that predict bilinguals' production and perception of L2 sounds. One such factor is the age of learning (AOL) the second language. Another is amount of L2 experience, often indexed as length of residence (LOR) in a predominantly L2-speaking environment. The relative importance of the AOL and LOR variables depends largely on the population being studied. For a group of adults who have been speaking English between 6 months and 3 years, LOR is likely to account for a great deal of variance in most outcome measures. This is because L2 speech learning proceeds rapidly at first. However, if the study population consists of subjects representing a wide range of AOLs, all of whom have used English for at least five years at the time of the study, LOR is likely to be far less important than AOL.

A potent and long-lasting effect of AOL was demonstrated in a recent series of studies examining the production of English by 240 Italian adults. These subjects had first begun to learn English when they immigrated to Canada between the ages of 2 and 23 years. All of them had lived in Canada for at least 15 years. In a study by Flege, Munro, and MacKay (1995), native English listeners used a continuous scale to rate sentences spoken by the native Italian subjects and a group of native English speakers. AOL accounted for 60% of the variance in the Italian subjects' foreign accent ratings, whereas LOR accounted for very little variance (< 2%). The later the Italian subjects had begun to learn English, the stronger were their foreign accents. Strength of foreign

accent increased linearly with AOL. There was not a dramatic increase in foreign accent at the age of 15 years, as has been claimed. However, none of the subjects who had begun learning English after the age of 15 managed to speak English without a detectable foreign accent. Somewhat surprisingly, some native Italian subjects who had begun to learn English as young children were found to speak English with slight but detectable foreign accents. Thus, if foreign accents arise because a critical period has been passed, one might suppose that the critical period ends earlier in life than is commonly supposed.

Strong age effects were also observed in studies examining the Italian/English bilinguals' productions of English vowels and consonants. Munro, Flege, and MacKay (1995) examined 11 English vowels spoken in a /b_ɹ/ context. Vowels spoken by most Italian subjects, even those who had begun to learn English as adults, were identified correctly in a forced-choice test. Given that Standard Italian has far fewer vowels than English (7 vs. 15), this suggested that a great deal of learning had taken place. However, another experiment revealed that the vowels were more strongly foreign-accented as AOL increased. This held true for English vowels without an obvious Italian counterpart (e.g., those in *bat* and *Bert*) as well as English vowels that differ just slightly from a roughly comparable Italian vowel. These two putatively different types of vowels are called "new" and "similar". (It should be emphasized, however, that to be used in testing hypotheses, classification requires an empirical evaluation of the perceived relation of L1 and L2 vowels. This has not as yet been done for the Italian/English interface.) Given the intelligibility findings mentioned earlier, it seems unlikely that the native Italian subjects simply substituted the closest Italian vowel for each of the "new" English vowels. Foreign accent in new English vowels spoken by many Italian subjects with AOLs greater than 10 years disconfirmed the claim that such vowels will eventually be mastered by all learners of an L2, regardless of AOL.

An important question is why foreign accent in the "new" English vowels grew increasingly strong as AOL increased beyond about the

age of 10 years. An equally important question, perhaps, is how some "late learners" managed to produce the new vowels without foreign accent. One possibility is that, as AOL increases, L2 learners become less likely to form phonetic categories for new L2 vowels. On this view, production of an L2 vowel will be completely nativelike if an independent category is established for it. Given listeners' perceptual sensitivity to small differences between vowels, it seems likely that all of the native Italian subjects detected differences between the new English vowels and the closest vowel from the Italian inventory. The inter-subject variability may have depended on whether or not subjects treated such differences as phonetically relevant. In instances where this did not happen, one might suppose that the native Italian subjects continued to treat the new English vowels as distorted variants of an Italian category.

Given this scenario, one might expect the single phonetic representation used by Italian/English bilinguals to produce both a "new" English vowel and its Italian counterpart to change over time so as to reflect the acoustic properties of the many tokens (English and Italian) judged to be realizations of it. On this view, as a bilingual approximates L2 phonetic norms for a new L2 vowel, productions of the L1 vowel with which it has been linked perceptually should shift in the direction of the L2 vowel. This hypothesized process, called cross-language phonetic assimilation, has been observed in studies of stop production and perception. The relevant studies compare speakers of an L1 in which /p t k/ are realized with short-lag VOT values who later learn an L2 in which the stops /p t k/ are realized with long-lag VOT (or vice versa). In a forced-choice identification experiment, such subjects may show a cross-over point (phoneme boundary) that is intermediate to the ones observed for monolingual speakers of the L1 and L2. This strongly suggests that a single perceptual representation is being used to process corresponding L1 and L2 stops. In production, bilinguals may produce /p t k/ in both the L1 and L2 with VOT values that are intermediate to monolingual speakers' VOT values.

As already mentioned, many native Italian subjects, including some who began learning English in childhood, produced "similar" English vowels (e.g., those in *beat*, *boot*, *boat*, *bet*) with a foreign accent. Perhaps the corresponding Italian vowels were sufficiently different that, when used in English, the substitution was auditorily evident to native English-speaking listeners. Alternatively, certain native Italian subjects, especially those who began to learn English early in life, may have established categories for the similar English vowels. Their newly formed English vowel categories may have been deflected by a cross-language dissimilatory process. For example, a category formed for English /o/ may have been altered to preserve phonetic contrast between English /o/ and Italian /o/. Note that this hypothesis assumes that, for bilinguals, phonic elements of the L1 and L2 exist in a common phonological space.

Cross-language dissimilation of the type just mentioned was evident in the speech of an 11-year-old French/English bilingual examined by Mack (1990). This boy's renditions of French /p t k/ had VOT values that were about 20 ms too long for French. (He may have learned to produce French stops in this way from his mother, who learned English as an adult. His mother could therefore be expected to have produced English /p t k/ with "compromise" VOT values as the result of cross-language assimilation.) The boy's renditions of English /p t k/ also had VOT values that were too long for English. He was unlikely to have heard English /p t k/ produced with the kind of VOT values he himself produced. The extra-long VOT values he produced in English /p t k/ may have served to contrast these stops from his realizations of French /p t k/.

Another potential source of error in L2 production and perception may be age limits on the use of features that distinguish sounds of the L2 but not the L1. Fox, Flege, and Munro (1995) had English monolinguals and Spanish-English bilinguals rate pairs of English and Spanish vowels for degree of dissimilarity. The English subjects used three features in perceiving the vowels, whereas the native Spanish

subjects used just two. The native English subjects may have used more features because the 15 vowels of English are differentiated along a wider range of dimensions than are the five vowels of Spanish. Productions of the vowel in *Bert* by the Italian subjects discussed earlier were nearly always identified correctly. However, degree of foreign accent in /ô/ increased with AOL. Acoustic analysis suggested that the Italian subjects who produced /ô/ with a foreign accent failed to produce this vowel with second and third formants that were close together in frequency. They may have failed to use the rhotic feature (which distinguishes /ô/ from all other English vowels) because such a feature is not used to distinguish Italian vowels.

AOL also strongly affected the Italian subjects' production of word-initial English consonants, such as those in *read*, *think*, and *they*. (The consonants initiating these words are not found in Italian.) Subjects who had begun to learn English by about the age of 10 years produced these consonants correctly as often as did native English subjects. As AOL increased beyond 10 years, fewer Italian subjects were judged to have produced the English consonants correctly. The Italian subjects' most typical error in producing interdental fricatives was to substitute a stop. Perhaps those subjects who had passed a critical period were unable to establish the new motoric patterns needed to produce /T/ and /D/. Alternatively, the subjects who erred in producing English interdental fricatives may have failed to treat the difference between /t/-/T/ and /d/-/D/ as phonetically relevant.

Evidence that adults can learn to produce accurately certain L2 sounds not found in the L1 provides a strong argument against the view that humans lose the ability to motorically implement sounds not found in the L1. One such example was provided by a recent study examining native Japanese subjects who had begun to learn English as adults (Flege, Takagi and Mann, 1995). The study evaluated the Japanese subjects' productions of liquids in English words such *read* and *lead*. The liquids in such words do not occur in Japanese. As expected, liquids spoken by subjects who had lived in the United States

for 2 years were frequently misidentified. However, liquids spoken by subjects who had lived in the US for 20 years were always identified correctly. Moreover, liquids spoken by ten of the 12 experienced Japanese subjects received ratings that fell within the range of ratings obtained for 12 native English speakers.

The Japanese subjects' perception of English liquids was tested by having them identify, as "R" or "L", words spoken by native speakers of English. As expected, the inexperienced Japanese subjects frequently misidentified liquids whereas the experienced Japanese subjects identified liquids at the same near-perfect rates as native English speakers, at least for liquids found in minimally paired words that were of equal subjective familiarity (e.g., *read*, *lead*). The experienced Japanese subjects sometimes misidentified liquids, however, in words that were less familiar than their minimal pairs. For example, they sometimes responded "L" to *rook*, which was less familiar subjectively to them than was *look*. For English monolinguals, lexical frequency effects are usually confined to phonetic segments that are ambiguous. Thus, it is possible that the experienced Japanese subjects established phonetic categories for English /r/ and /l/, but their categories differed from English monolinguals', perhaps because they specified their categories using different features or feature weights.

At present it is uncertain whether foreign accent in production and perception arises because a biologically scheduled capacity for speech learning diminishes with maturation, or whether age effects on L2 production and perception arise because L1 representations become better established with age and thus exert a stronger influence on L2 sounds. The key to distinguishing between a "maturational state" hypothesis and an "interference" hypothesis is to determine whether, or to what extent, the acquisition of L2 vowels and consonants influences L1 vowels and consonants (either through assimilatory or dissimilatory processes). Two recent findings suggest that, as implied by the interference hypothesis, pronunciation of the L1 and L2 are interrelated. The native Italian subjects mentioned earlier were asked to rate their

ability to pronounce English and Italian. A criss-cross pattern was obtained, with very few subjects giving themselves the highest possible ratings in both English and Italian. Subjects who began to learn English relatively late in life said they spoke Italian better than English, whereas subjects who began to learn English as children said they spoke English better than Italian. Yeni-Komshian and Flege (1995) had monolingual listeners (Korean or English) rate Korean-English bilinguals' production of Korean and English. This study also yielded a criss-cross pattern.

Additional research is needed that considers pronunciation of both the L1 and L2. Evidence of linked, bi-directional errors would support the interference hypothesis. Such evidence would call into question the prevailing view that segmental production and perception errors arise from a diminished ability to learn speech. In such research, it will be important to consider the bilingual subjects' history of language acquisition and language use. For example, effects of learning an L2 on L1 production may be most readily evident in situations where L1 is clearly non-dominant, or in which the L1 has seldom been used in the recent past. If the interference hypothesis is supported, it will become important to establish measures that can be reliably used to assess the extent to which bilinguals separate the phonological and phonetic systems they use to pronounce the L1 and L2.

Note

- 1 This article was published previously as follows: Flege, J.E. (1995) The phonetic study of bilingualism. In G. Bloothoof, V. Hazan, D. Huber, J. Llisterrri (Eds.) *European Studies in Phonetics and Speech Communication* (pp. 98-103). Utrecht: OTS. The guest editors are grateful to James E. Flege, to the editors of the above collection, and to OTS, Utrecht for their permission to reprint the article.

References

- Flege, J.E., & Munro, M.J. (1994). The word unit in L2 speech production and perception. *Studies in Second Language Acquisition*, 16, 381-411.
- Flege, J.E., Munro, M.J., & MacKay, I. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.
- Flege, J.E., Munro, M.J., & MacKay, I. (1995). The effect of age of second language learning on the production of English consonants. *Speech Communication* 16, 1-26.
- Flege, J.E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language & Speech*, 38, 25-55.
- Flege, J.E., Takagi, N., and Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of the Acoustical Society of America* 99, 1161-1173.
- Fox, R.A., Flege, J.E., & Munro, M.J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America* 97, 2540-2551.
- Mack, M. (1990). Phonetic transfer in a French-English bilingual child. In P.H. Nelde (Ed.), *Language Attitudes and Language Conflict*. Bonn, Germany: Dümmler.
- Munro, M., Flege, J.E., & MacKay, I. (1995). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics*, 17, 313-334.
- Yeni-Komshian, G., & Flege, J.E. (1994). *Maturational limitations to pronunciation accuracy in a second language*. Paper presented at the Acoustical Society of America Meeting, Austin, Texas, November, 1994.