

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/267553619>

# Language contact in bilingualism: Phonetic system interactions

Article in *Laboratory Phonology* · January 2007

CITATIONS

130

READS

333

1 author:



[James Emil Flege](#)

University of Alabama at Birmingham

227 PUBLICATIONS 13,399 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Flege lab [View project](#)



Cross-language perception/production of vowels [View project](#)

# Language contact in bilingualism: Phonetic system interactions

*James E. Flege*

## Abstract

It is well known that "earlier is better" as far as learning to pronounce an L2 is concerned. Section 1 will show that the age of arrival (henceforth, AOA) of native Korean participants to the United States (US) exerted a powerful effect on their pronunciation of English. However, findings reported in subsequent sections suggest that the AOA effects did not necessarily arise from a loss of neural plasticity. Factors that might influence L2 speech were confounded with AOA (Section 2). Just as the L1 sound system affects L2 production, the reverse held true as well (Section 3). Section 4 offers an account of how the phonetic elements comprising the L1 and L2 phonetic subsystems of bilinguals might interact.

## 1. The effect of AOA

This section will review research examining the production of English by native Korean (NK) participants who had lived in the US for an average of 15 years (range = 7-30). The 240 NK participants, who had arrived in the US between the ages of 1-23 years, were assigned to ten subgroups based on AOA (e.g., the 24 assigned to "AOA-3" arrived in the US at an average age of 3 years, the 24 assigned to "AOA-5" arrived at an average age of 5 years, and so on).

Flege, Yeni-Komshian, and Liu (1999) evaluated overall degree of foreign accent in English sentences. Sentences repeated following an aural model were digitized, then rated by native English (NE) speaking listeners using a scale that ranged from 1 ("strong accent") to 9 ("no accent"). A strong correlation existed between AOA and the average foreign accent ratings obtained from the NK participants,  $r(238) = -.85$ ,  $p < .01$ . Fig. 1 shows the mean ratings obtained for the ten AOA-defined subgroups and the NE controls. A Tukey test revealed that NK participants who arrived in the US prior to the age of 12 years (i.e., those in the AOA-3, AOA-5, AOA-7, AOA-9 and AOA-11 groups) received significantly higher ratings—indicating a better pronunciation of English—than those arriving after the age of 14 years (groups AOA-15, AOA-17, AOA-19, AOA-21). Also NK participants who arrived in the US prior to school age (AOA-3, AOA-5) received higher ratings than those who arrived after the age of 8 years (AOA-9 to AOA-21) ( $p < .05$ ). Importantly, all ten NK groups received significantly lower ratings than the NE controls. This meant that even participants in the AOA-3 group, who had lived in the US for an average of 20 years and had received all of their education in US schools, had detectable foreign accents.

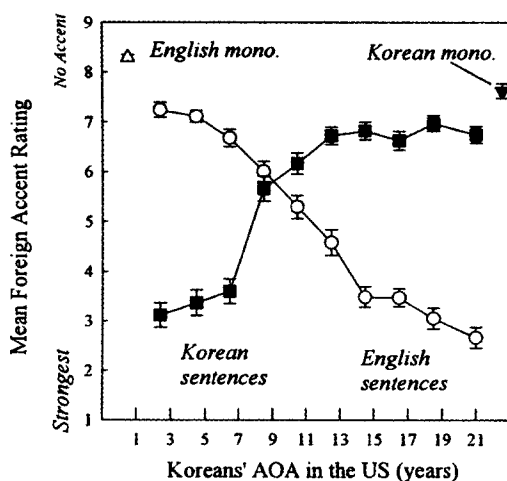


Figure 1. Mean ratings of English and Korean sentences produced by groups of Korean-English bilinguals differing in age of arrival in the US and English and Korean monolinguals ( $n = 24$  per group). The error bars bracket  $\pm 1$  SE. The data are from Flege, Yeni-Komshian, and Liu (1999) and Yeni-Komshian, Flege, and Liu (2000).

The study just reviewed compared groups of NK adults who differed according to AOA. More recently, Flege et al. (2001) directly compared NK adults and children to NE adults and children at two times separated by 1.2 years. English sentences spoken by groups of NK adults who had lived in North America for averages of 3 and 5 years received lower ratings than sentences spoken by NE adults; similarly, sentences spoken by groups of NK children with the same lengths of residence obtained lower ratings than NE children's sentences. This provided additional evidence that children speak their L2 with a detectable foreign accent, even after five years of immersion (see also Asher and Garcia 1969). Importantly, however, the NK children but not the NK adults obtained significantly higher ratings—indicating milder foreign accents—at Time 2 than Time 1.

Flege and Yeni-Komshian (unpublished) examined the production of English vowels and consonants. The 240 adult NK participants repeated sequences of English words following an aural model. Words containing /i/ (feet, sheet), /ɪ/ (fit, ship) /u/ (food, shoot) /ʊ/ (foot, should) and /æ/ (fat, shack) were digitized, intensity normalized, and randomly presented in separate blocks to NE-speaking listeners, who rated vowel pronunciation using a scale that ranged from 1 ("very bad") to 7 ("very good"). An average rating of English vowel production accuracy was obtained for each participant based on 80 ratings (5 vowels x 2 words x 8 listeners). The strong correlation between AOA and these ratings,  $r(238) = -.80$ ,  $p < .01$ , indicated that the later the NK participants arrived in the US, the less English-like their vowels were judged to be.<sup>1</sup> The effect of Group on the ratings was significant,  $F(10,253) = 56.9$ ,  $p < .01$ . A Tukey test revealed that vowels spoken by NK participants who arrived in the US after the age of 8 years (AOA-9 to AOA-21) received significantly lower ratings than vowels spoken by the NE controls ( $p < .05$ ).

Flege and Yeni-Komshian (unpublished) also assessed the NK participants' production of /t/ and /d/. The final VC portions of the words ending in /d/ (/ud/, /ʊd/) and /t/ (/ut/, /ʊt/) were randomly presented to five phonetically trained listeners, who classified the final consonants as "t", "d" or "other". On average, listeners responded "other" 4.5% of the time for /t/ and 4.2% of the time for /d/. The NE controls' stops were heard as intended in 95% of instances. Rates for the NK participants ranged from 96% for AOA-3, to 70% for AOA-21. The correlation between AOA and the correct classification rates was  $r(238) = -.57$ ,  $p < .01$ . A Tukey test revealed that only stops produced by NK participants who arrived in the US after the age

of 12 years (AOA-13 to AOA-21) differed significantly from the NE controls' stops ( $p < .05$ ).

Acoustic analyses suggested that the primary difference in stop production between the NE controls and the late learners was the production of glottal pulsing in /d/. As shown in Fig. 2(a), the participants in all 11 groups produced much the same difference in vowel duration as a function of the voicing feature of the following stop, resulting in a non-significant Group x Stop interaction,  $F(10,253) = 1.3$ . The effect of Group reached significance,  $F(10,253) = 3.1$ ,  $p < .01$ , because vowels spoken by AOA-21 were significantly longer than the NE controls' vowels ( $p < .05$  by Tukey's test). A similar pattern of results was obtained for the duration of the /t/ and /d/ constrictions intervals.<sup>2</sup>

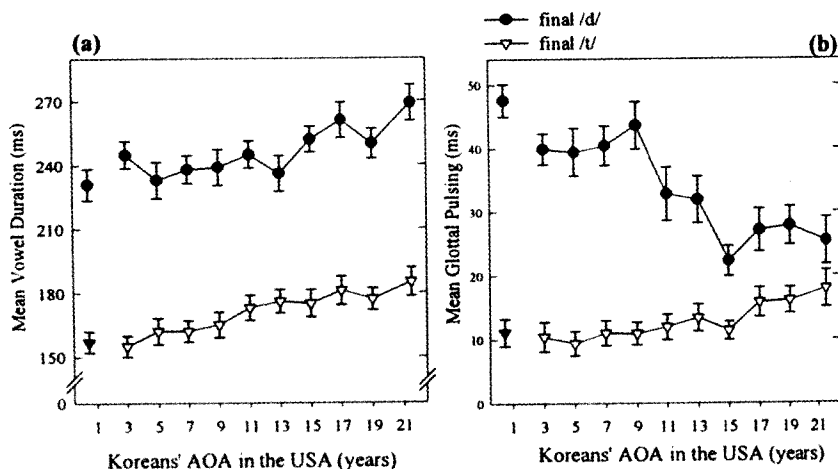


Figure 2. Acoustic measurements of vowels preceding /d/ and /t/ (a) and the duration of closure voicing in /d/ and /t/ tokens (b) produced by groups of Korean-English bilinguals differing in age of arrival in the US and English monolinguals ( $n = 24$  per group). The error bars bracket  $\pm 1$  SE. The data are from Flege and Yeni-Komshian (unpublished).

However, the ANOVA examining the duration of glottal pulsing in word-final /d/ and /t/ tokens yielded a Group x Stop interaction,  $F(10,253) = 8.7$ ,  $p < .01$ . All 11 groups, including late bilinguals, produced significantly more glottal pulsing in /d/ than /t/ (Bonferroni  $p < .05$ ). The interac-

NE con-  
 p produc-  
 duction of  
 11 groups  
 tion of the  
 nt Group x  
 gnificance,  
 ere signifi-  
 y's test). A  
 /t/ and /d/



and the dura-  
 by groups of  
 e US and Eng-  
 et  $\pm 1$  SE. The

al pulsing in  
 on,  $F(10,253)$   
 reduced signifi-  
 . The interac-

tion arose because the magnitudes of the /d/-/t/ difference were greater for the NE controls and early learners than for late learners. The simple effect of Group reached significance for /d/,  $F(10,253) = 6.3$ ,  $p < .01$ , but not /t/,  $F(10,253) = 1.7$ . A Tukey test revealed that NK participants who arrived in the US after the age of 12 years (AOA-13 to AOA-21) produced significantly less glottal pulsing than the NE controls ( $p < .05$ ). The reduced duration of glottal pulsing in late bilinguals' /d/ tokens probably affected the listeners' judgments, for a correlation existed between the number of listeners (maximum = 5) who classified the /d/ tokens as "d" and the duration of glottal pulsing,  $r(517) = .37$ ,  $p < .01$ .

In summary, AOA exerted a strong influence on all L2 phonetic dimensions examined: overall degree of foreign accent in sentences, accuracy in producing five English vowels, and the perceptual effectiveness of word-final /t/-/d/ tokens. NK participants who arrived in the US as young children received lower ratings (indicating the presence of detectable foreign accents) in the production of English sentences, but were not found to differ significantly from NE controls in producing specific English vowels or word-final consonants.

## 2. Explaining AOA effects

The AOA effects described in Section 1 could be viewed as support for the hypothesis that L2 learning is limited by maturational constraints. On this view, some mechanism(s) that is needed for successful L1 acquisition works less effectively, or becomes inaccessible for use in L2 acquisition as humans mature. Scovel (1988, 2000) hypothesized that a critical period for L2 speech learning ends at around the age of 12 years because the brain loses plasticity as the result of normal neurological maturation. Patkowski (1990) proposed a critical period ending at 15 years of age. DeKeyser (2000: 518-519) was less specific regarding the end of a critical period, suggesting that between 6-17 years of age humans show a "severe decline" of the ability to "induce abstract patterns implicitly" in L2 speech and language input.

There are three reasons to question the hypothesis that the AOA effects reported in Section 1 were due to a neurologically based critical period. First, there is evidence that some adults are successful in learning to pronounce an L2 (Bongaerts 1999; Bongaerts, Mennen, and Slik 2000). Sec-

ond, as reviewed in Section 1, adults who began learning their L2 in childhood, and also children, may produce their L2 with foreign accents. These differences probably cannot be attributed to the passing of a critical period ending as early as 6 years of age (DeKeyser 2000) because some foreign-accented early learners began to learn their L2 prior to that age.

Third, as discussed by Flege (1987a), it is difficult to interpret age-related effects on L2 speech production because AOA is frequently confounded with factors that might be expected to influence performance. Several such confounds with AOA will be reviewed in this section for the NK participants whose production of English was reviewed Section 1. Confounds with AOA that are similar to those observed for the NK participants have also been reported for native Spanish and native Italian immigrants to North America (Bahrick et al. 1994; Flege 1998). The analyses of AOA "confounds" will focus on three variables derived from a language background questionnaire (LBQ) administered to the NK participants. Two derived variables relate to language use and input, and the third is an index of bilingual dominance.

The derived variable "%Life-US" was the percentage of their lives that the NK participants had spent in the US. The later the NK participants arrived in the US, the fewer years they tended to have lived there,  $r = -.42$ ,  $p < .01$ , and the older they tended to be when tested,  $r = .68$ ,  $p < .01$ . As a result of these countervailing tendencies, a strong correlation,  $r = -.90$ ,  $p > .01$ , existed between AOA and %Life-USA. On average, NK participants who arrived in the US between the ages of 3-7 years had spent 79% of their lives in the US, as compared to just 40% for participants with an AOA of 17-21 years.

The second derived variable, "E/K use", provided an index of the relative frequency with which the NK participants used their two languages. Research with Italian-English bilinguals has shown that self-estimated percentage Italian use exerts an important influence on L2 (English) speech performance. The less the L1 was used (and so also the more English, the L2, was used), the more native-like production and perception of the L2 tended to be (Flege, Munro, and MacKay 1995a, 1995b; Piske et al. 2001, 2002; MacKay et al. 2001; MacKay, Meador, and Flege 2001; Meador, Flege, and MacKay 2000).

The NK participants used a scale (1 = "never", 5 = "all the time") to rate how often they used both of their languages "at home", "at parties and with friends", while watching "TV and videos", and "overall in the past 5

years". AOA was correlated with the average rating of Korean use,  $r = .56$ ,  $p < .01$ , and English use,  $r = -.59$ ,  $p < .01$ . The ratio of the ratings (E/K use) was computed as an overall index of language use. As shown in Fig. 3(a), NK participants with an AOA of 3-11 years indicated using English more than Korean whereas those with an AOA of 13-21 years indicated using English and Korean about equally. A moderate correlation was found to exist between AOA and E/K use,  $r = -.56$ ,  $p < .01$ .

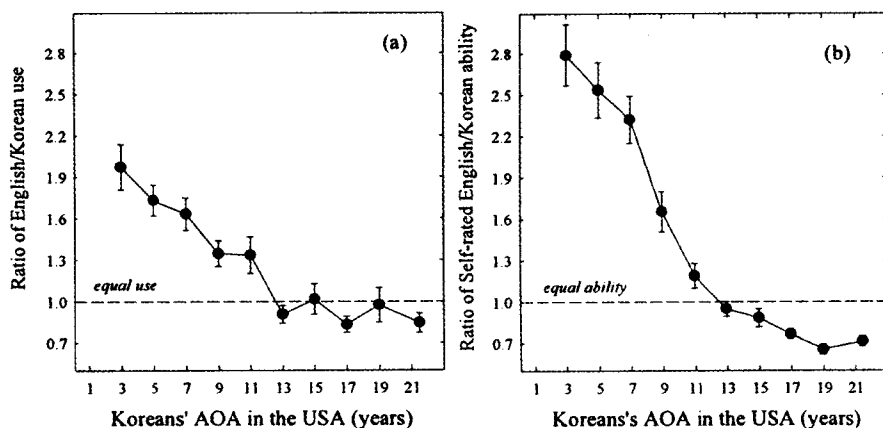


Figure 3. The mean ratio of self-reported English/Korean use (a) and the mean ratio of self-reported English/Korean proficiency (b) for groups of Korean-English bilinguals differing in age of arrival in the US ( $n = 24$  per group). The error bars bracket  $\pm 1$  SE.

The third derived variable, "E/K proficiency," provided an index of bilingual dominance. Piske et al. (2001) observed that sentences spoken by groups of Italian adults who learned English as children received significantly lower ratings than sentences spoken by NE controls. However, individual early learners who had become English-dominant appeared to speak English without detectable foreign accents, perhaps because the dominant language of a bilingual develops to a greater extent than the non-dominant language (Grosjean 1982). The NK participants used a scale (1 = "very poor", 5 = "very good") to estimate their own ability to "pronounce" English and Korean, to "read and write" each language, to "remember the pro-



nunciation" of words in each language, and to indicate "overall ability". Average ratings were obtained for English and Korean, each based on four items. AOA was correlated with the average ratings of proficiency in both Korean,  $r = .78$ ,  $p < .01$ , and English,  $r = -.68$ ,  $p < .01$ . Following the practice of Flege and MacKay (2004), language dominance was estimated by computing the ratio of self-rated proficiency in the two languages. As shown in Fig. 3(b), participants with an AOA of 3-11 years reported being more proficient in English than Korean whereas the opposite held true for those having an AOA of 13-21 years. A strong correlation was found to exist between AOA and the E/K proficiency ratios,  $r = -.74$ ,  $p < .01$ .

*Table 1.* Correlations between ratings of English sentences and vowels produced by 240 Korean immigrants to the US, their age of arrival in the US, and three variables derived from a Language Background Questionnaire.

	2. V prod.	3. AOA	4. % Life	5. E/K use	6. E/K prof.
1 Foreign accent in sentences	.89	-.85	.78	.57	.68
2 Vowel production accuracy	--	-.80	.75	.55	.64
3 Age of arrival (AOA) in US	--	--	-.90	-.56	-.74
4 Percentage of life in the US	--	--	--	.57	.76
5 Ratio of English/Korean use	--	--	--	--	.67
6 Ratio of English/Korean proficiency	--	--	--	--	--

*Note:*  $df = 238$ ; all  $p$ -values  $< .0001$ .

Table 1 presents the simple Pearson correlations between ratings of the NK participants' English sentences and vowels (see above), AOA, and the three variables derived from the LBQ. The correlations between the derived LBQ variables and the ratings of L2 pronunciation were only slightly weaker (mean  $r = .66$ ) than were the correlations between AOA and ratings of L2 pronunciation (sentences  $r = -.85$ , vowels  $r = -.80$ ). Not only were the three derived LBQ variables correlated with AOA (mean  $r = -.73$ ), they were also correlated with one another (mean unsigned  $r$ -value  $= .70$ ).

It is not possible at present to explain the causal basis of these correlations, for many basic questions remained unanswered. For example, it is uncertain why participants who arrived in the US as young children ("early" learners) tended to become English-dominant whereas those who arrived in late adolescence or early adulthood ("late" learners) tended to

overall ability".  
 based on four  
 proficiency in both  
 following the prac-  
 as estimated by  
 languages. As  
 reported being  
 ite held true for  
 n was found to  
 $p < .01$ .

vowels produced  
 val in the US, and  
 questionnaire.

6.	
use	E/K prof.
.57	.68
.55	.64
.56	-.74
.57	.76
--	.67
--	--

een ratings of the  
 e), AOA, and the  
 between the derived  
 were only slightly  
 A AOA and ratings  
 Not only were the  
 an  $r = -.73$ ), they  
 value = .70).  
 s of these correla-  
 For example, it is  
 as young children  
 whereas those who  
 learners) tended to

remain Korean-dominant. Perhaps early but not late learners tended to become English-dominant simply because they tended to use English relatively often and Korean relatively seldom. Alternatively, the dominance pattern may have arisen as an indirect consequence of differences in the opportunity and need to use English on a daily basis. Still another possibility is that AOA-related differences in language use were due to the passing of a critical period. If a critical period exists for L2 speech learning, then the late learners may have used English less than the early learners because it was more difficult for them to learn English.

Additional correlation analyses were performed to determine if the three derived LBQ variables remained correlated with the two measures of English pronunciation when variation in AOA was statistically controlled. The correlations between K/E use and the two measures of English pronunciation remained significant when the effect of AOA was partialled out (sentences  $r = .22$ , vowels  $r = .20$ ,  $p < .01$ ). The same held true for the correlations between K/E proficiency and the two measures of English pronunciation ( $r = .13$  for both sentences and vowels,  $p = .05$ ). The correlations between %Life-US and the vowel ratings remained significant when AOA was partialled out ( $r = .14$ ,  $p < .05$ ), but not the correlation between %Life-US and the sentence ratings ( $r = -.05$ ).

The finding for K/E proficiency provided modest support for the inference (Piske, MacKay, and Flege 2001) that bilingual dominance affects overall degree of foreign accent. The "%Life-US" variable may have failed to predict overall degree of foreign accent in English sentences independently of AOA because it provided a poor index of L2 input. Simply living in a predominantly L2-speaking environment does not guarantee L2 use (see, e.g., Flege and Liu 2001).

The basis for the relation between language use and L2 pronunciation is uncertain given the complexity of several relevant variables, and the lack of detailed measurement of those variables. For example, early and late learners tend to differ in amount of L2 (English) use because children who immigrate to North America are usually soon enrolled in an English-medium school where they interact frequently with native English speakers. Also they tend, later in life, to marry native English speakers. Adult immigrants to North America, on the other hand, usually enter the workplace immediately. They tend to interact frequently with fellow L1 native speakers both at work and socially, and are often married (or will marry) a fellow L1 native speaker. Work by Jia and Aaronson (1999) suggested that Chinese

speakers who arrived in the US prior to about the age of 13 years were more likely to seek out native English children as conversational partners and friends than Chinese immigrants who were older upon arrival in the US. Stevens (1999) suggested that patterns of language use and socialization may be an important determinant of success in learning an L2. Success in learning an L2 might, in turn, influence attitudes toward the L2 and the culture associated with the L2. Grosjean (1982: 195ff) suggested that psychosocial factors may exert an effect on L2 learning that is even more important than the age of first exposure to the L2.

The finding that K/E use remained correlated with measures of English pronunciation when variation in AOA was statistically controlled agreed with a matched subgroup analysis presented by Flege, Yeni-Komshian, and Liu (1999) for the NK participants considered here. Two groups of NK participants were selected from the original sample of 240. The "matched subgroups" ( $n = 20$  each) were selected in such a way that they differed significantly ( $p < .01$ ) in their use of both Korean (4.1 vs. 2.1) and English (3.3 vs. 4.5) but were matched for AOA (mean = 11.4 years for both). The process of identifying pairs of participants who differed in language use but not AOA reduced variation in other variables that were confounded with AOA. As a result, the matched subgroups did not differ significantly in length of US residence or years of education in English-medium US schools. The NK participants who used English often and Korean seldom were found to have a significantly better pronunciation of English than those who used English seldom and Korean often,  $F(1,38) = 4.3$ ,  $p < .05$ .

In summary, although the 240 NK participants were chosen according to AOA, they differed according to dimensions other than age of first exposure to English. Two variables were identified that were associated with the pronunciation of English independently of AOA: the relative frequency of L1 and L2 use, and language dominance (operationalized as the ratio of self-reported proficiency in the L1 and L2). These findings do not disprove that age-related effects on L2 learning are due to endogenous factors associated with maturation, such as a loss of neural plasticity. However, the findings suggest that the strong effects of AOA seen in many studies arise, at least in part, from age-related differences in the contexts in which languages are learned and used (Grosjean 1982) and differences in how often the L1 and L2 are used (e.g., Piske, MacKay, and Flege 2001).

The partial correlations obtained between K/E use and K/E proficiency and the measures of L2 pronunciation were considerably smaller when

AOA was statistically controlled than the partial correlations obtained between AOA and the L2 pronunciation measures when effects due to variation in K/E use and K/E proficiency were statistically controlled. This suggests that AOA exerted a stronger influence on L2 speech learning than the factors confounded with AOA. It is important to note, however, that it is not possible at present to quantify the effect on L2 performance of variables confounded with AOA because the confounded variables have yet to be measured precisely. For example, no L2 study has as yet successfully measured the L2 input received by groups of L2 learners who differed in AOA and were shown to differ in L2 speech production. A proper assessment would of course include measures of how much the L1 and L2 were being used at the time of test. It would also include an assessment of how often, and in what contexts, the L1 and L2 had been used over the entire period of L2 use. It is also essential to evaluate who has modeled L2 speech patterns for the L2 learners. Individuals who are exposed only to native-speaker models are likely to have a better L2 pronunciation of the L2 than individuals who have received a substantial amount of foreign-accented input. Also, it seems plausible to think that individuals who have lost the ability to speak their L1 through disuse will pronounce their L2 better than those who remain bilingual.

### 3. Phonetic interference

Prior to the suggestion by Lenneberg (1967) that a critical period might exist for L2 speech acquisition, native versus non-native differences in L2 performance were usually attributed to cross-language "interference", that is, an influence of prior learning on subsequent learning. Although most L2 research has focused on the effect of the L1 on L2 performance, interference appears to be bi-directional. The aim of this section is to present research showing an influence of learning English on the production of Korean by the 240 NK participants considered in Sections 1 and 2.

A recent study examining the duration of L1 and L2 sentences spoken by Italian-English bilinguals (MacKay and Flege 2004) suggested that the dominant language might influence the non-dominant language to a greater extent than the reverse. As already discussed, most of the NK participants who arrived in the US as children seem to have become English-dominant whereas many of those who arrived later in life may have remained Ko-

rean-dominant (see also Jia and Aaronson 1999). One would therefore expect a stronger L2 (English) effect on L1 (Korean) production for early than late bilinguals. The results obtained by Yeni-Komshian, Flege, and Liu (2000) supported this prediction.

Yeni-Komshian, Flege, and Liu (2000) examined Korean sentences spoken by the NK participants whose pronunciation of English was described in Section 1. The procedures used to elicit and rate Korean sentences were similar to those used for the English sentences. The primary difference was the replacement of the monolingual NE control group by Korean monolinguals who were recorded in Seoul, and the replacement of NE-speaking by NK-speaking listeners. Fig. 1 shows that, as predicted, lower ratings were obtained for Korean sentences spoken by early than late bilinguals. The effect of Group was significant,  $F(10,253) = 60.9$ ,  $p < .01$ . A Tukey test indicated that significantly lower ratings were obtained for sentences produced by groups AOA-3 to AOA-11 than by the Korean monolinguals ( $p < .01$ ). This indicated NK participants who arrived in the US before the age of 12 year produced the Korean sentences with an "American" foreign accent.

The finding just reported supported the view that phonetic interference is bi-directional. To further assess the influence of L2 learning on L1 production, each of the 240 NK participants was classified as either having or not having a detectable foreign accent in Korean. "No accent" for individual participants was operationally defined as a foreign accent rating that fell within 2 SDs of the mean rating that was obtained for the Korean monolingual control group ( $n = 24$ ). The 106 participants meeting the no-accent criterion for Korean sentences had arrived in the US at an average age of 16 years ( $range = 9-23$ ). They had spent 45% of their lives in the US, reported using English less than Korean ( $mean$  E/K use ratio = .91), and being less proficient in English than Korean ( $mean$  E/K proficiency ratio = .85). The same procedure was also applied to the production of English sentences. The 16 NK participants who met the no-accent criterion for English had arrived in the US at an average age of 5 years ( $range = 1-9$ ), had spent 79% of their lives in the US, used English more than Korean ( $mean$  E/K use ratio = 1.87), and reported being far more proficient in English than Korean ( $mean$  E/K proficiency ratio = 2.67).

A visual inspection of Fig. 1 reveals that the NK participants who had a relatively good pronunciation of English generally had a relatively poor pronunciation of Korean, and vice versa. Just one individual—a male who

therefore ex-  
on for early  
age, and Liu

n sentences  
ish was de-  
Korean sen-  
The primary  
ol group by  
placement of  
s predicted,  
arly than late  
0.9,  $p < .01$ .  
obtained for  
the Korean  
arrived in the  
ces with an

interference  
g on L1 pro-  
er having or  
for individ-  
ting that fell  
an monolin-  
ne no-accent  
ge age of 16  
US, reported  
d being less  
= .85). The  
h sentences.  
English had  
d spent 79%  
ean E/K use  
than Korean

ts who had a  
atively poor  
-a male who

arrived in the US at the age of 9 years—met the no-accent criterion for both Korean and English sentences. That is, just one of the 240 NK participants was credited with having produced both L1 and L2 sentences without a detectable foreign accent. This finding suggests that bi-directional phonetic interference effects are likely to be evident in the great majority of bilinguals. Put another way: they are the “rule” rather than an “exception” characterizing the speech of a few aberrant individuals.

Yeni-Komshian and Flege (unpublished) further evaluated the effect of L2 learning on L1 production by examining the NK participants’ production of four Korean consonants (/s/, /s’/, /t<sup>h</sup>/, /t’/) in word-initial position. Two tokens of each consonant were randomly presented in separate, counterbalanced blocks (one for each consonant) to NK-speaking listeners who had lived in the US for less than 1.5 years. The listeners rated each consonant (1 = “wrong consonant”, 2 = “distorted”, 3 = “okay”, 4 = “very good”), and a mean rating was obtained for each participant by averaging over 64 judgments (4 consonants x 2 tokens x 8 listeners).<sup>3</sup>

Fig. 4 shows that lower ratings were obtained for all ten AOA-defined subgroups than for the Korean monolinguals. The effect of Group was

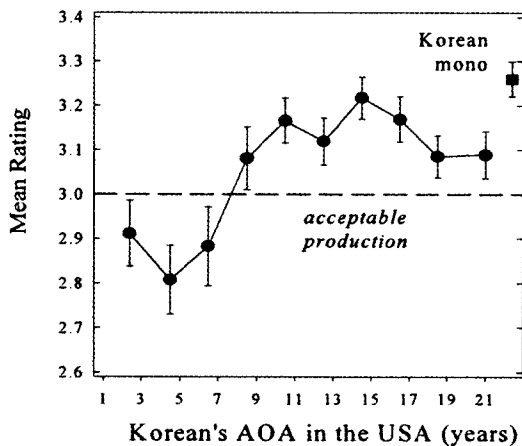


Figure 4. Mean ratings of four word-initial Korean consonants produced by groups of Korean-English bilinguals differing in age of arrival in the US and Korean monolinguals ( $n = 24$  per group). The error bars bracket  $\pm 1$  SE. The data are from Yeni-Komshian and Flege (unpublished).

significant,  $F(10,253) = 5.6$ ,  $p < .01$ . A Tukey test revealed that only consonants produced by participants in AOA-3 to AOA-7 received significantly lower ratings than the Korean monolinguals' consonants ( $p < .01$ ). Although fewer groups differed from the Korean monolinguals for Korean consonants than sentences (3 vs. 5), the overall pattern of results was similar, and a significant correlation was obtained for ratings of Korean word-initial consonants and Korean sentences,  $r(238) = -.49$ ,  $p < .01$ .

#### 4. L1-L2 phonetic interactions

The findings reported in Section 3 suggested that the L1 and L2 phonetic subsystems of bilinguals mutually influence one another, and that the relative strength of L1→L2 and L2→L1 interference effects may depend on bilingual balance. Individuals who remain L1-dominant (mostly late learners; see MacKay and Flege 2004) seem to show stronger L1→L2 than L2→L1 effects whereas the reverse seems to hold true for individuals who become dominant in the L1 (mostly early learners). At issue is how the L1 and L2 phonetic subsystems affect one another. The purpose of this section is to discuss interactions between L1 and L2 phonetic segments from the perspective of the Speech Learning Model, or SLM (Flege 1995, 1999, 2002, 2003), which is apparently the only theoretical account predicting bidirectional L1-L2 interference effects.

The SLM begins with two controversial assumptions. The first is that L2 speech learning is not constrained by a critical period. The second is that the capacities needed by monolingual children to learn the language-specific properties of their L1 are preserved across the life span, and remain accessible to L2 learners of all ages. These starting assumptions do not represent the implicit claim that age-related effects do not exist, which is clearly wrong (see Section 1). As already discussed, age-related differences in L2 speech may be due in part to differences in input. However, according to the SLM, the primary basis of age-related effects on L2 speech is a change in how the L1 and L2 phonetic subsystems of bilinguals interact.

The SLM posits that the categories making up the L1 and L2 subsystems of a bilingual exist in a "common phonological space", and so will mutually influence one another. The SLM proposes two mechanisms through which the phonetic categories comprising the L1 and L2 phonetic subsystems interact: "phonetic category assimilation" and "phonetic cate-

gory dissimilation". Both mechanisms are thought to affect the categories used in producing the vowels and consonants ("sounds" for short) of the L1 and L2. Which of the two mechanisms operates is thought to depend on whether category formation has or has not taken place for L2 sounds.

The SLM proposes that individuals who learn an L2, regardless of their age when first exposed to the L2, are able to establish new categories for L2 sounds. It is implausible, however, to think that L2 learners (even young children) will establish new phonetic categories for all L2 sounds that can be shown to differ statistically from the closest L1 sound. For example, one would not expect native English speakers to establish new categories for the voiceless aspirated stops of Danish, for Danish /p t k/ are realized with only slightly longer voice onset time (VOT) values than English /p t k/ are.

By hypothesis, an important determinant of whether a new category will be established for an L2 sound is the state of development of neighboring L1 sounds. According to the SLM, L1 categories become more likely to subsume L2 sounds as they develop through childhood and into adolescence (e.g., Hazan and Barrett 1999). An L2 sound is said to have been "equated" with an L1 sound if it continues indefinitely to be processed as an instance of an L1 category. (Importantly, equivalence classification does not prevent L2 learners from auditorily detecting phonetic differences between an equated L2 sound and its L1 counterpart.)

The SLM also hypothesizes that the perceived phonetic dissimilarity of an L2 sound from the closest L1 sound is a determinant of whether a new phonetic category will or will not be established for the L2 sound. The more distant from the closest L1 sound an L2 sound is judged to be, the more likely it is that L2 learners—regardless of age—will establish a new category for the L2 sound.<sup>4</sup>

#### 4.1. Phonetic category assimilation

The SLM proposes that the L1 and L2 phonetic subsystems of a bilingual will interact through the mechanism of category assimilation when phonetic category formation has been blocked by equivalence classification. By hypothesis, equivalence classification is most likely to prevent a new category from being established for an L2 sound when neighboring L1 categories are fully developed and when the L2 sound is perceived to be phonetically similar to a neighboring L1 sound.



The SLM predicts that an L2 learner will at first use the closest L1 sound to produce L2 sounds, without evidence of modification or learning. However, equivalence classification does not prevent L2 learners from auditorily detecting cross-language phonetic differences. The SLM predicts that, when L2 category formation is blocked, production of an L2 sound will be modified slowly over time if the L2 sound differs audibly from the L1 sound with which it has been equated. The modification will be limited, however, because a single long-term memory representation will be used to process instances of the L2 sound and its L1 counterpart. This long-term memory representation ("phonetic category") will gradually evolve over time, yielding a "merged" category that reflects both L1 and L2 phonetic input. Depending on the nature of the input received, the merged category may resemble more closely the long-term representation of L1 or L2 monolinguals.

Evidence of the operation of category assimilation was obtained by Flege (1987b). This study examined the production of /t/ in French and English words by native French women who had learned English in adulthood and had lived in Chicago for 12 years, and American women who had learned French in adulthood and had lived in Paris for about the same length of time. Both groups of bilinguals read phrase lists of the form "Tous les\_" and "Two little\_" (where the blanks were filled with variable French or English nouns). VOT was measured in the utterance-initial /t/ tokens. English monolinguals were found to produce /t/ with longer VOT values than French monolinguals. The question of interest was whether the bilinguals would resemble monolinguals in both their languages.

Both groups of bilinguals showed evidence of phonetic learning in their L2. The French-English bilinguals in Chicago produced English /t/ with longer average VOT values than the French monolinguals produced in French /t/. The English-French bilinguals in Paris produced French /t/ with shorter VOT values than English monolinguals produced in English /t/.

The fact that neither group of bilinguals produced L2 words with an unmodified L1 /t/ suggested that they had detected at least some aspect of the phonetic difference between French and English /t/. However, the French-English bilinguals in Chicago produced English /t/ with shorter VOT values than English monolinguals; and the English-French bilinguals in Paris produced French /t/ with longer VOT values than French monolinguals. In summary, both bilingual groups phonetically approximated but did not achieve the VOT norm for /t/ in their L2, producing L2 stops with

the closest L1  
tion or learning.  
L2 learners from  
the SLM predicts  
of an L2 sound  
audibly from the  
n will be limited,  
n will be used to  
. This long-term  
ally evolve over  
and L2 phonetic  
merged category  
L1 or L2 mono-

was obtained by  
/t/ in French and  
English in adult-  
women who had  
about the same  
of the form "Tous  
h variable French  
initial /t/ tokens.  
nger VOT values  
whether the bilin-  
es.

c learning in their  
l English /t/ with  
uals produced in  
ed French /t/ with  
in English /t/.  
L2 words with an  
ast some aspect of  
/t/. However, the  
h /t/ with shorter  
-French bilinguals  
n French monolin-  
approximated but  
ing L2 stops with

VOT values that were intermediate to the means obtained for French and English monolinguals.

Most importantly, Flege (1987b) obtained evidence that L2 learning influenced L1 production. The French-English bilinguals in Chicago produced French /t/ with longer average VOT values than French monolinguals, thereby showing an influence of English on their production of the L1, French. Conversely, the English-French bilinguals in Paris produced English /t/ with shorter VOT values than English monolinguals, thereby showing an influence of French on their production of the L1, English.

The results obtained by Flege (1987b) suggested that both groups of late bilinguals detected cross-language phonetic differences between French and English /t/ but nevertheless did not establish a new phonetic category for the /t/ of their L2. As a result, their pre-existing (L1) category representation for /t/ may have evolved to reflect a two-language source of input. The exact nature of these representations presumably reflected the array of input values they had experienced over their lives, perhaps with more recently encountered tokens being weighted more heavily than tokens encountered in the distant past (Sancier and Fowler 1997).

MacKay et al. (2001) obtained evidence of phonetic category assimilation for both early and late bilinguals. This study focused on the production of English /b d g/ in word-initial position by four groups of native Italian (NI) participants differing in AOA in Canada (early learners 2-13 years; late learners 15-26 years) and percentage Italian use (low-L1-use 1-15%; high-L1-use 25-80%). The NI participants were deemed unlikely to establish new categories for English /b d g/, which are usually realized with short-lag VOT values rather than with lead VOT values (i.e., pre-voicing), as in Italian. This is because English /b d g/ are sometimes realized with lead VOT values, and because the NI participants already possessed phonetic categories defined by short-lag VOT values (viz., those for Italian /p t k/).<sup>5</sup> Three NI groups (Early-high, Late-high, Late-low) showed an influence of Italian, producing English /b/ with pre-voicing significantly more often than NE monolinguals. However, all three of these groups pre-voiced English /b/ less often than Italian monolinguals pre-voiced Italian /b/, indicating that Italian continued to influence their production of English.

Most importantly, MacKay et al. (2001) obtained a correlation between the frequency with which the bilinguals realized /b d g/ with full (uninterrupted) pre-voicing in their two languages. The less frequently they fully pre-voiced English /b/, the less frequently they did so for Italian /b d g/.

The group whose productions of English /b/ were most English-like (viz., the Early-low group) also showed the greatest influence of English on production of Italian /b d g/. The authors interpreted these results to mean that the NI participants used a single phonetic category for producing and perceiving the /b/ of English and Italian, and suggested that differences between the NI groups arose not from the application of different mechanisms, but from differences in phonetic input.

The finding obtained in a vowel perception experiment by Flege, MacKay, and Meador (1999) also suggested the operation of category assimilation. These authors compared the categorial discrimination of English vowels (/æ/-/ʌ/, /ʌ/-/a/, /i/-/ɪ/, /u/-/ʊ/) by late bilinguals (*mean* AOA = 19 years) and early bilinguals (*mean* AOA = 7 years). The late bilinguals discriminated the English vowels less accurately than the early bilinguals and a NE control group (see also Flege and MacKay 2004). The English vowels examined in all four contrasts occupied a portion of vowel space in which Italian has only one vowel (e.g., Italian has only /i/ in the portion of vowel space occupied by English /i/ and /ɪ/). The early-late difference in the discrimination of English vowels was therefore attributed to the early bilinguals' greater likelihood of establishing new categories for English vowels.

The late bilinguals were also found to discriminate Italian vowels (/u/-/o/, /e/-/a/, /u/-/i/) less accurately than the early bilinguals and NE controls. If the late bilinguals were less likely than the early bilinguals to establish new categories for English vowels then, according to the SLM, they should have been more likely to show the influence of phonetic category assimilation. Their representation for Italian /i/, for example, might have evolved to reflect the properties of Italian /i/, English /i/, and English /ɪ/. If so, the "broadening" of an L1 category representation may have impeded the recognition, storage, retrieval, or comparison in short-term memory of the Italian /i/ tokens to other vowel stimuli.

#### 4.2. Phonetic category dissimilation

When someone learning an L2 establishes a new category for an L2 sound, their phonetic space becomes more crowded. The SLM posits that the L1 and L2 phonetic categories of a bilingual tend to disperse in order to maintain phonetic contrast, in much the same way that the vowels of a language disperse in vowel space.<sup>6</sup> According to the SLM, a newly established L2

sh-like (viz.,  
English on pro-  
to mean that  
ing and per-  
ferences be-  
erent mecha-

nt by Flege,  
category as-  
on of English  
n AOA = 19  
ilinguals dis-  
ilinguals and  
English vowels  
pace in which  
tion of vowel  
ce in the dis-  
e early bilin-  
English vowels.  
n vowels (/u/-  
d NE controls.  
ls to establish  
M, they should  
gory assimila-  
ave evolved to  
/I/. If so, the  
ped the rec-  
emory of the

r an L2 sound,  
its that the L1  
order to main-  
s of a language  
established L2

category may dissimilate from a neighboring L1 category (or vice versa) when implementations of the L2 and L1 categories might otherwise be confusable. When this happens, a bilingual's productions of the L2 and/or the L1 sound would be expected to differ those of monolingual.

Flege and Eefting (1987) obtained evidence of category dissimilation in a study examining the production of Spanish /p t k/ by native speakers of Spanish who had learned English as children. In a preliminary study, Flege and Eefting (1986) found that English monolinguals produced /p t k/ with longer VOT values than age-matched Spanish monolinguals ( $p < .01$ ). Within both languages, VOT values were longer for adults than children ( $p < .01$ ). The lack of a Language x Age interaction in the analysis of the monolingual data indicated that the adult-child differences were comparable in Spanish and English despite the fact that /p t k/ are realized with substantially longer VOT values in English than Spanish. The age effect was consistent with the view that the representations used to guide the production of L1 phonetic segments continue to develop slowly over time.

Flege and Eefting (1987) examined the production of word-initial tokens of /p t k/ in Spanish and English words by Spanish-English bilingual children and adults. All of the early bilinguals (both children and adults) were first exposed to English in a bilingual school located in Mayaguez, Puerto Rico.<sup>7</sup> The bilingual children's production of Spanish words were compared to that of the monolingual Spanish children mentioned earlier (who attended a public school in Mayaguez). Their production of English stops were compared to those of monolingual English children (who attended a parochial school in Alabama). Similarly, the bilingual adults' productions were compared to that of monolingual Spanish adults (students at the University of Mayaguez who had studied English at school but were unable to carry on a rudimentary conversation in English) and monolingual English adults (students in Birmingham, Alabama).

Flege and Eefting (1987) found that the early bilinguals produced /p t k/ with substantially longer VOT values in English than Spanish words ( $p < .01$ ). The lack of a significant Age x Language interaction suggested that child and adult bilinguals produced comparable English-Spanish VOT differences. Most importantly, both groups of early bilinguals produced Spanish /p t k/ with significantly shorter VOT values than age-matched Spanish monolinguals. The lack of a significant Age x Language Status interaction in the analysis of Spanish VOT values produced by monolinguals and bilinguals suggested that both groups of early bilinguals (children, adults)

shortened VOT in Spanish /p t k/ to a comparable extent. Flege and Eefting (1987) concluded that the early bilinguals established new phonetic categories for English /p t k/ based on the phonetic input they had received in Puerto Rico, and that their Spanish /p t k/ categories dissimilated from these new English categories for /p t k/ in order to maintain phonetic contrast with English /p t k/. Dissimilation at the phonetic category level resulted in a shortening of VOT values in Spanish /p t k/.

The explanation just offered assumed that the early bilinguals established new phonetic categories for English /p t k/. Independent evidence for this was obtained by Flege and Eefting (1988). The same participants examined by Flege and Eefting (1987) were asked to imitate the members of a synthetic /da/-to-/ta/ continuum whose VOT values ranged from lead (typical for Spanish /d/) to short-lag (typical for English /d/, Spanish /t/) to long-lag (typical for English /t/). None of the participants accurately reproduced the stimulus VOT values even though they had been instructed to imitate the CV stimuli "as accurately as possible". The nature of stimulus-response VOT differences depended on how participants classified the stimuli, which was examined in a separate identification experiment.

Flege and Eefting (1988) found that Spanish monolinguals tended to produce imitation responses having lead or short-lag VOT values similar to those that are typical for Spanish /d/ and Spanish /t/. English monolinguals tended to produce short-lag and long-lag VOT values similar to those that are typical for English /d/ and /t/. Both groups of early bilinguals, on the other hand, produced imitation responses having VOT values that fell into all three modal VOT ranges. Importantly, they tended to imitate short-lag stimuli with short-lag VOT values and long-lag stimuli with long-lag VOT values. Given the short time interval available for imitating the CV stimuli, this finding supported the view that the early bilinguals had established new phonetic categories for English /p t k/.

Flege, Schirru, and MacKay (2003) obtained acoustic evidence of phonetic category dissimilation in a study examining native Italian (NI) participants' production of English /e<sup>l</sup>/. Italian adults with little English-language experience tend to judge English /e<sup>l</sup>/ tokens as being instances of Italian /e/ (Flege and MacKay 2004) even though English /e<sup>l</sup>/ is produced with considerably more tongue movement than Italian /e/. Flege, MacKay, and Meador (1999) found that just one of four groups of Italian-English bilinguals (*viz.*, Early-low) were able to discriminate English /e<sup>l</sup>/ and Italian /e/ tokens at a significantly above-chance rate. This suggested that early

bilinguals who seldom used Italian were more likely than early bilinguals who used Italian often, or late bilinguals, to establish a new phonetic category for English /e<sup>l</sup>/.

Flege, Schirru, and MacKay (2003) tested this by comparing the production of English /e<sup>l</sup>/ by four NI groups differing in AOA in Canada (early 2-13 years, late 15-26 years) and percentage Italian use (low-L1-use 1-15%, high-L1-use 25-85%). Linear predictive coding analysis was used to estimate the F1 and F2 frequencies at 20% and 80% locations in each /e<sup>l</sup>/ token. Rough estimates were then obtained of the amount of tongue movement from the 20% to the 80% location. The formant frequency values were first converted from Hertz to bark units, and the position of the tongue in a high-low (B0-B1) and front-back dimension (B1-B2) were computed for both measurement locations. Finally, the Euclidean distance between points defined by B1-B0 and B2-B1 values at the two measurement locations was computed.

All four NI groups and the NE controls produced /e<sup>l</sup>/ with similar vowel qualities at the 20% location. However, vowel qualities at the 80% location differed substantially across the five groups. The Late-high group produced an ending vowel quality that was lower and farther back in vowel space than the NE controls' whereas the Early-low group produced an ending vowel quality that was higher and farther forward. A Tukey test revealed that participants in the Early-low group produced /e<sup>l</sup>/ with significantly larger Euclidean distance values—indicating more tongue movement—than participants in the NE and Late-high groups ( $p < .05$ ). No other between-group differences reached significance.

Flege, Schirru, and MacKay (2003) interpreted these results to mean that participants in the Early-low groups were more likely than participants in the other groups to have established a new phonetic category for English /e<sup>l</sup>/, and so produced it with exaggerated tongue movement in order to differentiate it from their previously established Italian /e/ category. This conclusion is consistent with the results of a vowel perception study by Flege and MacKay (2004). These authors compared the categorial discrimination of nine pairs of English vowels by the four groups of Italian-English bilinguals who participated in the Flege, Schirru, and MacKay (2003) study. Both AOA and L1 use affected the bilinguals' discrimination of English vowels. The Early-low group obtained the highest discrimination scores and the Late-high group obtained the lowest scores.

## 5. Summary and conclusions

Sections 1-3 of this chapter focused on 240 native speakers of Korean who immigrated to the US between the ages of 1-23 years. Section 1 showed that AOA exerted a strong effect on their overall degree of foreign accent, as well as their production of specific English vowels and consonants. The superior L2 production of early than late learners, but not the finding that early learners produced English with a foreign accent, was consistent with the hypothesis that a critical period exists for L2 speech learning. Section 2 revealed that, as for other groups of immigrants to North America, factors likely to influence L2 pronunciation independently of AOA were confounded with AOA. Section 3 showed that learning English as an L2 influenced the NK participants' production of their L1, Korean. Early learners were found to speak Korean with an "American" foreign accent, and their production of word-initial Korean consonants were judged to diverge from the phonetic norms of Korean.

These results did not disprove the hypothesis that age-related effects on L2 speech production are due to maturational constraints on speech learning. However, they demonstrated, at the very least, that factors other than neurological development at the time of first exposure to an L2 exert an important influence on L2 speech production. For example, language use patterns affect how well an L2 will be pronounced. Also, cross-language phonetic interference, which was shown to be bi-directional, affects L2 speech production.

The aim of Section 3 was to illustrate hypotheses of the SLM pertaining to how phonetic elements of bilinguals' L1 and L2 phonetic subsystems interact. Examples were given of phonetic category assimilation, which occurs when learners are unable to establish a new phonetic category for an L2 sound that differs audibly from the closest L1 sound (either because the L2 sound is perceptually subsumed by a fully-developed L1 category, because it is perceived to be too similar to an L1 sound, or both). Phonetic category assimilation may result in the "non-nativelike" pronunciation of both L1 and L2 sounds. L2 sounds may show the influence of the L1 phonetic system (L1→L2 phonetic interference) and L1 sounds may show the influence of the L2 phonetic system (L2→L1 phonetic interference). Examples were also given of phonetic category dissimilation, which occurs when learners are able to establish a new phonetic category for an L2 sound. In one example, L1 speech sounds were produced in such a way as

to differ more from neighboring L2 sounds; in the other example, production of an L2 sound was exaggerated so as to differentiate it from a neighboring L1 sound.

Several limitations pertaining to the treatment of age effects on L2 speech production should be mentioned briefly. First, no "litmus test" now exists for the formation of new categories for L2 speech sounds. Inferences were drawn as to whether new L2 phonetic categories had or had not been established. However, before the conclusions presented earlier with respect to the operation of phonetic category assimilation and dissimilation can be accepted with confidence, it will be necessary to replicate the findings reported here for groups of L2 learners who have been independently demonstrated to have either formed, or not formed, new categories for particular L2 sounds.

Second, at least one previous study provided results that run counter to the SLM claim that phonetic category dissimilation occurs when new L2 categories are created. Flege (1991) examined the production of English stop consonants by early Spanish-English bilinguals who were living in the US. These early bilinguals produced English stops with VOT values that were virtually identical to NE monolinguals'. This might be interpreted as evidence that the early bilinguals had established new phonetic categories for English /p t k/. However, an analysis of the early bilinguals' production of Spanish /p t k/ revealed no difference from the VOT values produced by Spanish monolinguals.

Why did the early Spanish-English bilinguals living in Puerto Rico (Flege and Eefting 1987) shorten VOT values in Spanish /p t k/ whereas those living in the US (Flege 1991) did not? One possible explanation is that the two groups received different English phonetic input. The early bilinguals in the US matched English monolinguals, which provided indirect evidence that they had received appropriate English phonetic input. The fact that early bilinguals in Puerto Rico produced English /p t k/ with shorter VOT values than English monolinguals suggested that they may have based their new phonetic categories for English /p t k/ on "compromise" VOT values they heard in the speech of native speakers of Spanish who learned English in adulthood. If this explanation is true, it might mean that phonetic category dissimilation occurs only when a newly established L2 phonetic category is relatively close in phonetic space to a pre-existing L1 category. More work is clearly needed to establish the conditions under which cross-language phonetic dissimilation occurs in bilinguals.



Frequent mention has been made in this contribution of the importance of input in L2 speech learning. This brings us again to a caveat that has already been mentioned. DeKeyser (2000: 519) claimed that "input differences are not a good explanation for age effects, because it is precisely in the linguistic domains where input varies least—phonology—that the age effects are most readily apparent." Contrary to this observation, L2 input has been shown to vary considerably as a function of AOA, and variation in language use has been shown to exert an important influence on L2 speech production (e.g., Flege, Munro, and MacKay 1995a). At present, it appears that language use exerts a less powerful effect on L2 pronunciation than AOA does. However, until researchers can find precise ways to quantify the overall amount—and quality—of L2 input, the relative importance of input in comparison to other factors (e.g., state of neurological development or the state of development of L1 categories at the time of first exposure to the L2) will remain uncertain.

Other issues that are crucial to an understanding of L2 phonetic and phonological development have scarcely been addressed in L2 speech research. For example: Are new categories more likely to be formed by individuals who have a large lexicon containing many potentially confusable minimal pairs than by individuals with restricted L2 lexicons? Does the size of L1-L2 phonetic differences that are produced by bilinguals depend on the frequency with which they code switch in conversations with other bilinguals?

In summary, a great deal of research has shown that "earlier is better" as far as the pronunciation of an L2 is concerned. However, the underlying basis (or bases) of age-related effects is by no means clear. The most productive research strategy in the coming years will be to devise studies that aim to determine the amount of variance in L2 speech performance that can be accounted for by degree of perceived cross-language phonetic distance, cross-language perceptual assimilation patterns, amount and quality of L2 input, the patterns and context of L1 and L2 use, and L2 lexical development.

### **Acknowledgements**

This study was supported by grant DC00257 from the National Institute for Deafness and Other Communicative Disorders. The author thanks Ian MacKay and Grace Yeni-Komshian for their valuable contributions to studies cited here.

## Notes

1. To prevent variation in the final consonant from affecting the vowel judgments, everything following the complete constriction of the final consonant was edited out. The listeners, who were told the identity of the target word before each block began, were not trained on the rating task; however, each block began with 22 extra stimuli for familiarization that were not analyzed. The AOA effect was similar for all five vowels. Correlations with AOA ranged from a low of  $r = -.61$  for /t/ to a high of  $r = -.74$  for /u/ and /o/ ( $p < .01$ ).
2. The size of the duration difference between the constriction intervals of /t/ and /d/ (means = 111 vs. 66 ms) was similar across groups, resulting in a non-significant Group x Stop interaction,  $F(10,253) = 1.1$ ,  $p > .10$ . The effect of Group was significant,  $F(10,253) = 9.0$ ,  $p < .01$ , because six Korean groups (AOA-9, AOA-11, AOA-15 to AOA-21) produced significantly longer stop closure intervals than the NE controls ( $p < .05$ ).
3. Participants repeated series of Korean words, including 1- and 2-syllable words beginning with /sa/, /s'a/, /t<sup>h</sup>a/, and /t'a/, following an aural model. The initial syllable of the 2-syllable words were edited out, and all 528 tokens of each syllable were intensity normalized. A preliminary ANOVA indicated that neither Word Length (i.e., whether the syllables were derived from 1- or 2-syllable words) nor Consonant interacted with Group.
4. Cross-language phonetic distance has usually been measured perceptually (e.g., Schmidt 1996). Participants typically identify L2 speech sounds in terms of the abstract phonetic categories of their L1, then rate the L2 speech sounds for degree of goodness of fit to the selected L1 categories. Unfortunately, L2 speech researchers have not yet converged on the most appropriate method for gauging cross-language phonetic distance. One complicating factor is that certain L1 and L2 speech sounds that are likely to be judged to be "the same" may not occur in the same range of phonetic contexts in both languages. Even if they did, patterns of allophonic variation may differ cross-linguistically. It remains uncertain at what level of analysis (e.g., phonemes, position-sensitive allophones, "major" allophones, context-sensitive allophones) the phonic elements of a bilingual's two languages are perceptually related. It is also uncertain whether contextual effects on cross-language perceptual assimilation patterns influence L2 speech learning.
5. In fact, an identification experiment showed that the NI participants tended to mistakenly identify short-lag English /b d g/ tokens as /p t k/.
6. Bilinguals are apt to use both of their languages in rapid succession (Grosjean 1982). Distinguishing phones generated by categories in the L1 and L2 phonetic subsystems would be expected to facilitate lexical access, especially

when bilinguals search both the L1 and L2 lexicons for cognate words that have different pronunciations in the L1 and L2.

7. English was used as a language of instruction in this school, but many teachers and most students were native speakers of Spanish.

## References

- Asher, James and R. Garcia  
1969 The optimal age to learn a foreign language. *The Modern Language Journal* 53: 334-341.
- Bahrack, Harry, L. Hall, J. Goggin, L. Bahrack, and S. Berger  
1994 Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants. *Journal of Experimental Psychology: General* 123: 264-283.
- Bongaerts, Theo  
1999 Ultimate attainment in L2 pronunciation: The case of very advanced late learners. In *Second Language Acquisition and the Critical Period Hypothesis*, David Birdsong (ed.), 133-160. Mahwah NJ: Lawrence Erlbaum Associates.
- Bongaerts, Theo, S. Mennen, and F. van der Slik  
2000 Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54: 298-308.
- DeKeyser, Robert  
2000 The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition* 22: 499-534.
- Flege, James  
1987a A critical period for learning to pronounce foreign languages? *Applied Linguistics* 8: 162-177.  
1987b The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15: 47-65.  
1991 Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89: 395-411.  
1995 Second-language speech learning: theory, findings, and problems. In *Speech Perception and Linguistic Experience*, Winifred Strange (ed.), 233-277. Timonium, MD: York Press.

- 1998 The role of subject and phonetic variables in L2 speech acquisition. In *Papers from the 34th Annual Meeting of the Chicago Linguistic Society, Volume II, The Panels*, M. Gruber, D. Higgins, K. Olsen, and T. Wysocki (eds.) 213-232. Chicago: Chicago Linguistic Society.
- 1999 Age of learning and second-language speech. In *Second Language Acquisition and the Critical Period Hypothesis*, David Birdsong (ed.), 101-132. Mahwah NJ: Lawrence Erlbaum Associates.
- 2002 Interactions between the native and second-language phonetic systems. In *An Integrated View of Language Development: Papers in Honor of Henning Wode*, Petra Burmeister, Thorsten Piske, and Andreas Rohde (eds.), 217-244. Trier: Wissenschaftlicher Verlag.
- 2003 Assessing constraints on second-language segmental production and perception. In *Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities*, Antje Meyer and Niels Schiller (eds.) 319-355. Berlin: Mouton de Gruyter.
- Flege, James, David Birdsong, Ellen Bialystok, Molly Mack, and Hyekyung Sung  
2001 The production of English sentences by native English and native Korean adults and children. *Journal of the Acoustical Society of America* 110: 2684-2685.
- Flege, James and Wieke Eefting  
1986 Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica* 43: 155-171.
- Flege, James and Wieke Eefting  
1987 The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics* 15: 67-83.
- Flege, James and Wieke Eefting  
1988 Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America* 83: 729-740.
- Flege, James and Serena Liu  
2001 The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23: 527-552.
- Flege, James and Ian MacKay  
2004 Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26: 1-34.
- Flege, James and Grace Yeni-Komshian  
Unpublished The production of English vowels and consonants by Koreans differing in age of arrival in the United States.

- Flege, James, Grace Yeni-Komshian, and Serena Liu  
1999 Age constraints on second language acquisition. *Journal of Memory and Language* 41: 78-104.
- Flege, James, Ian MacKay, and Diane Meador  
1999 Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America* 106: 2973-2987.
- Flege, James, Murray Munro, and Ian MacKay  
1995a Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97: 3125-3134.
- Flege, James, Murray Munro, and Ian MacKay  
1995b The effect of age of second language learning on the production of English consonants. *Speech Communication* 16: 1-26.
- Flege, James, Carlo Schirru, and Ian MacKay  
2003 Interaction between the native and second language (L2) phonetic systems during L2 speech acquisition. *Speech Communication* 40: 467-491.
- Grosjean, François  
1982 *Life with Two Languages*. Cambridge, MA: Harvard University Press.
- Hazan, Valerie and S. Barrett  
1999 The development of phonemic categorization in children aged 6-12. *Journal of Phonetics* 28: 377-396.
- Jia, Gisella and D. Aaronson  
1999 Age differences in second language acquisition: The dominant language switch and maintenance hypothesis. In *Proceedings of the 23rd Annual Boston University Conference on Language Development*, 301-312. Somerville, MA: Cascadilla Press.
- Lenneberg, Eric  
1967 *Biological Foundations of Language*. New York: John Wiley and Sons, Inc.
- MacKay, Ian and James Flege  
2004 Effects of age of second language learning on the duration of first and second language sentences: The role of suppression. *Applied Psycholinguistics* 25: 373-396.
- MacKay, Ian, James Flege, Thorsten Piske, and Carlo Schirru  
2001 Category restructuring during second-language (L2) speech acquisition. *Journal of the Acoustical Society of America* 110: 516-528.
- MacKay, Ian, Diane Meador, and James Flege  
2001 The identification of English consonants by native speakers of Italian. *Phonetica* 58: 103-125.

Meador, Diane, James Flege, and Ian MacKay

- 2000 Factors affecting the recognition of words in a second language. *Bilingualism: Language and Cognition* 3: 55-67.

Patkowski, Mark

- 1990 Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics* 11: 73-89.

Piske, Thorsten, James Flege, Ian MacKay, and Diane Meador

- 2002 The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica* 59: 49-71.

Piske, Thorsten, Ian MacKay and James Flege

- 2001 Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191-215.

Sancier, M. and Carol Fowler

- 1997 Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25: 421-437.

Schmidt, Anna

- 1996 Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustical Society of America* 99: 3201-3211.

Scovel, Thomas

- 1988 *A Time to Speak. A Psycholinguistic Inquiry Into the Critical Period for Human Speech.* Cambridge, MA: Newbury House Publishers.  
2000 A critical review of critical period research. *Annual Review of Applied Linguistics* 20: 213-223.

Stevens, Gillian

- 1999 Age at immigration and second language proficiency among foreign-born adults. *Language in Society* 28: 555-578.

Yeni-Komshian, Grace and James Flege

- Unpublished. The production of Korean consonants by Koreans differing in age of arrival in the United States.

Yeni-Komshian, Grace, James Flege, and Serena Liu

- 2000 Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition* 3: 131-150.