

AGE AND L1 USE EFFECTS ON DEGREE OF FOREIGN ACCENT IN ENGLISH

Thorsten Piske* and Ian R. A. MacKay*

*The University of Alabama at Birmingham, *University of Ottawa

ABSTRACT

This study¹ assessed the relation between nonnative subjects' age of learning English, self-reported amount of L1 (Italian) use, and overall degree of perceived foreign accent in the L2 (English). Native English listeners evaluated English sentences spoken by English monolinguals and Italian/English bilinguals. The native Italian subjects were assigned to one of two groups of early bilinguals who differed according to their percentage use of Italian, or to one of two groups of late bilinguals differing in the same way. Both early and late bilinguals who spoke Italian frequently had significantly stronger accents in English than age-matched bilinguals who spoke Italian infrequently. This suggests that the degree of activation of the L1 affects L2 production accuracy regardless of a bilingual's age of L2 acquisition.

1. INTRODUCTION

Previous foreign accent (FA) research has shown that the earlier in life people learn a second language, the more likely they are to achieve an accurate, nativelike pronunciation of the L2 [2, 4, 6]. Flege, Munro and MacKay [6], for example, examined short English sentences spoken by 240 native Italian (NI) subjects who had begun learning English in Canada between the ages of 2 and 23 years. The later in life the NI subjects had arrived in Canada, the stronger were their perceived FAs in English. Age of arrival (AOA) in Canada accounted for most of the variance in the ratings of FAs (ca. 59%). However, amount of L2 use was also shown to influence degree of FA (ca. 15% of the variance).

An effect of self-reported L1 use on L2 production accuracy was obtained in a more recent study by Flege, Frieda and Nozawa [5]. These authors selected a subset of early bilinguals from the 240 NI subjects examined previously by Flege, Munro and MacKay [6] and assigned them to two groups. The subjects in these groups were matched for AOA in Canada (5.6 vs. 5.9 years), but differed according to their self-reported use of Italian. The LoUse group ($n = 20$) reported using Italian 3% of the time whereas the HiUse group ($n = 20$) reported using Italian 36% of the time. Native speakers of English (12 Canadians, 12 Americans) were asked to rate three short sentences produced by the Italian/English bilinguals for degree of FA. The subjects in both NI groups were found to have detectable FAs in English. However, the early bilinguals who spoke Italian frequently had significantly stronger FAs than did the early bilinguals who spoke Italian infrequently. This finding was probably not just a "practice" effect, for all of the NI subjects had been living in Canada for at least 18 years ($M = 34$ years).

The findings of [5] bear on some important theoretical issues. Proponents of the Critical Period Hypothesis have often maintained that the complete mastery of a second language is no longer possible after the age of about 12 to 15 [e.g., 8, 10].

Seliger [11] and Walsh and Diller [12] suggested that there are several critical periods, each affecting different linguistic abilities. The first ability to be lost would be the one needed to develop a nativelike pronunciation of an L2. Critical period accounts have usually attributed an age-related change in the ability to attain a nativelike pronunciation of an L2 to loss of neural plasticity, or to some sort of neurofunctional reorganization that occurs at around the onset of puberty [e.g., 7, 9]. The results of Flege, Frieda and Nozawa [5] indicated, however, that such views are too simple. First, the finding that the HiUse subjects had significantly stronger FAs in English than the LoUse subjects supported the results of other studies [2, 6] that had already shown that ultimate attainment in L2 pronunciation is also influenced by factors other than the state of neurological development at the time of first intensive exposure to the L2. Second, the nonnative subjects were found to have detectable FAs even though they had all started to learn English long before the age of 12 to 15, that is well before the proposed end of the critical period.

This latter finding does not mean, however, that it is impossible for L2 learners to achieve a nativelike pronunciation of an L2. Bongaerts, van Summeren, Planken and Schils [1] studied groups of late Dutch/English bilinguals. One group consisted of subjects who had been identified as highly motivated and successful learners of English by university-based EFL experts. The subjects in this group had their first intensive exposure to English at approximately 18 years of age. In an FA rating experiment, five of these 11 individuals received FA ratings from native English listeners that were comparable to the ratings assigned to a group of native English (NE) speakers. This suggested that even people who begin learning an L2 late in life can, in certain circumstances, achieve a fully nativelike pronunciation of this L2.

The present study replicated and extended the Flege, Frieda and Nozawa [5] study. It replicated the earlier study by examining English sentences spoken by two groups of early Italian/English bilinguals who were matched for their AOA in Canada but differed according to self-reported use of Italian. It extended [5] by also comparing two groups of *late* bilinguals who differed in L1 use but not AOA. The purpose of the study was to determine (a) if the L1 use effect observed for early bilinguals in [5] would also be found for late bilinguals, and (b) if AOA and L1 use interact.

The view that a learner's ultimate attainment in the pronunciation of an L2 is influenced by factors other than the age of L2 learning would be supported if both early and late bilinguals were found to show an L1 use effect. Such a finding would suggest that the degree of activation of the L1 affects L2 production accuracy regardless of a bilingual's age of L2 acquisition. One question not addressed by Flege, Frieda and

Nozawa [5] is whether the subjects in the “LoUse” group examined in their study had actually developed full competence in the L1 (Italian). It is possible that the L1 use differences between the HiUse group and the LoUse group found in [5] were, in fact, functionally equivalent to the presence vs. the absence of active L1 systems. If the present study found a larger L1 use effect for early than for late bilinguals, it would suggest that the effect seen in [5] was not due to differences in amount of L1 use but, rather, arose from the comparison of groups of early learners who did or did not have a functioning L1 system.

2. METHOD

2.1. Talkers

The 43 male and 47 female subjects examined in this study were living in or near Ottawa, Ontario when they participated. They were 49 years old on average ($SD = 6.2$). The 18 NE control subjects had all been born and raised in Canada, whereas the 72 NI subjects had all been born in Italy. The NI subjects had arrived in Canada between the ages of 2 and 26 years, and had lived there for an average of 36 years. They were assigned to four subgroups based on their AOA in Canada and self-reported use of Italian. As summarized in Table 1, the two groups of early Italian/English bilinguals were matched for AOA (EarlyLo-6.9 vs. EarlyHi-8.3 years) but differed significantly in amount of L1 use (EarlyLo-6.8% vs. EarlyHi-42.7%). The two groups of late bilinguals also differed significantly in L1 use (LateLo-9.5% vs. LateHi-53%) but not AOA (LateLo-19.5 vs. LateHi-19.5 years).

NI Groups	AOA	%Italian	LOR
EarlyLo	6.9 (3.3)	6.8 (3.6)	42 (3.7)
EarlyHi	8.3 (3.7)	42.7 (14.5)	40.1 (4.2)
LateLo	19.5 (3.3)	9.5 (4.6)	31.2 (8.3)
LateHi	19.5 (2.9)	53 (13.4)	29.3 (8.7)
Total $M (N = 72)$	13.6 (6.8)	28 (22.7)	35.7 (8.5)

Table 1. Characteristics of the four groups of NI subjects. Each group consisted of 18 subjects. AOA = mean age of arrival in Canada, in years; %Italian = mean self-reported percentage use of Italian; LOR = mean length of residence in Canada, in years. Standard deviations are in parentheses.

2.2. Speech Materials

The 90 subjects were tested in a quiet room in a Roman Catholic church in Ottawa. A delayed repetition technique was used to elicit the three sentences examined here (viz., *Paul ate carrots and peas; I can read this for you; He turned to the right*). Two of these three sentences were drawn from previous studies [2, 3, 4, 5]. The sentences were presented in written form and also aurally, via a tape recording. Each sentence to be produced was preceded and followed on the tape by a context sentence, as in the following mini-dialogue:

- (1) Voice 1: *What did Paul eat?*
- (2) Voice 2: *Paul ate carrots and peas.*
- (3) Voice 1: *What did Paul eat?*
- (4) subject: [repeats Voice 2]

Although the sentences were modeled on the tape, the delay between the model (i.e., Voice 2) and its repetition, as well as the

intervening speech material probably prevented direct imitations from sensory memory. Two tokens of each sentence were elicited. The sentences were not designed to be especially difficult for Italian speakers of English. They do not represent the full range of English allophones, nor do they include the full range of prosodic dimensions that distinguish English from Italian. The 90 subjects’ productions of the three sentences were recorded on DAT tape. The second token of each sentence was digitized at 22.05 kHz on a PC, except in the few instances in which it was judged to be disfluent (in which case the first token was used).

2.3. Listeners

Nine native speakers of Canadian English (4 males, 5 females) were asked to auditorily evaluate the three sentences produced by each of the 90 talkers. All nine listeners, who were between 22 and 46 years of age ($M = 31$ years), were living in Birmingham, Alabama at the time of testing. They had grown up in four different Canadian provinces (Ontario-6, Alberta-1, British Columbia-1, Manitoba-1). Six of the listeners had some knowledge of French, and one knew some Japanese. However, none of them was proficient in any language other than English. All of the listeners passed a pure tone hearing screening at octave frequencies between 500 and 4000 Hz (ref: 25 dB HL) before participating in the FA rating experiment.

2.4. Procedures

The listeners were tested individually in a sound booth. A total of 270 sentences (90 talkers x 3 sentences) was presented over loudspeakers at a comfortable level. The three English sentences were presented separately in three randomized blocks, the order of which was counterbalanced across listeners. Each block consisted of 300 trials, of which the first 30 were for practice (and were not analyzed.) The listeners heard each talker’s production of a sentence three times during the following 270 trials. The listeners were told that they would hear sentences spoken by Italian immigrants to Canada or by native speakers of Canadian English. They were asked to rate each sentence for degree of FA by pushing one of nine buttons representing a scale from 1 (very strong FA) to 9 (no FA). During the experiment, the listeners were able to listen to each sentence as often as they wished, and were given the chance to correct their last judgment if they wanted. A 1 sec interval occurred between each rating and the presentation of the next sentence. An average rating was obtained for each sentence, based on the final three judgments.

3. RESULTS

First, inter-listener and inter-sentence consistency was investigated. A preliminary analysis revealed that the mean FA ratings given by the nine listeners were similar. Overall mean ratings were obtained for each listener by averaging across the three sentences. When averaged over all 90 subjects, the mean ratings obtained for the nine listeners (averaged across sentences) ranged from a low of 4.4 to a high of 7.0. The simple correlation was computed for all possible pairwise combinations of two listeners. The lowest correlation was $r = .88$. The Intraclass correlation for the scores obtained for the nine listeners was $\rho = .99$, $F(89,712) = 93.5$, $p < .001$, which indicated that there was a very high degree of inter-listener agreement. Next, an average rating was computed for each of the three sentences as rated by

the nine listeners. The average ratings obtained for the 90 subjects for the three sentences differed little (means of 5.7, 5.5 and 5.8). A high Intra-class correlation for the sentence ratings, $\rho = .97$, $F(89,178) = 34.0$, $p < .001$ indicated that there was also substantial inter-sentence agreement. These results justified using a single average rating for each talker by averaging over listeners and sentences.

The mean ratings obtained for the four groups of NI subjects are shown in Figure 1. Averaged across use of Italian, higher ratings were obtained for sentences spoken by early than late bilinguals (Early = 6.6, Late = 3.5). Averaged across AOA, sentences spoken by native speakers of Italian who used their L1 infrequently received higher ratings than sentences spoken by individuals who used their L1 frequently (LoUse = 5.6, HiUse = 4.5). The scores obtained for all four NI groups were lower than the mean ratings received for sentences spoken by the native speakers of English ($M = 8.1$).

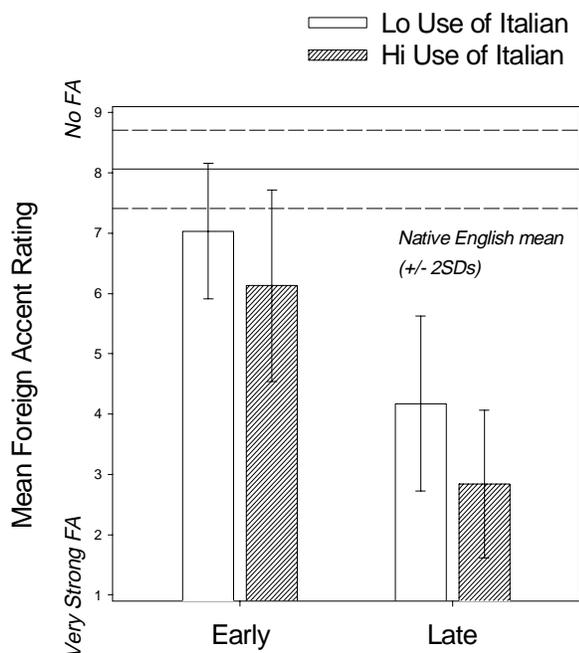


Figure 1. The mean ratings given by Canadian listeners to English sentences spoken by NE subjects or by NI subjects who spoke Italian frequently or infrequently. The solid horizontal line indicates the mean ratings obtained for sentences spoken by the NE speakers ($M = 8.1$). The dashed horizontal lines indicate the NE range (NE mean ± 2 SDs).

A series of F tests were carried out to determine if the scores obtained for any of the four NI groups were significantly lower than those obtained for the NE speakers' sentences. In all four instances, the NI groups' ratings were significantly lower, EarlyLo: $F = 13.9$, EarlyHi: $F = 25.7$, LateLo: $F = 123.2$, LateHi: $F = 307.3$, $df = 1,34$, Bonferroni $p < .001$.

The mean ratings obtained for the 72 NI subjects were submitted to a (2) Age of Arrival x (2) L1 Use ANOVA. The effect of AOA was significant, $F(1,68) = 92.4$, $p < .001$, and also L1 use, $F(1,68) = 12.2$, $p < .01$. However, the two-way interaction was non-significant, $F(1,68) = 0.45$, $p > .10$, indicating that the effect of L1 use was much the same for the early and the late bilinguals.

The effect of AOA on the FA ratings seems to have been greater than the effect of L1 use. The simple correlation between AOA and the FA ratings was larger ($r = -.83$) than that between the L1 use estimates and the FA ratings ($r = -.41$). This was also true when the correlation between AOA and FA was re-computed after the effect of L1 use was partialled out ($r = -.84$ vs. $r = -.47$). AOA accounted for more variance in the FA ratings in a step-wise multiple regression analysis (68.4% of the variance at Step 1) than did L1 use (7.1% more variance at Step 2).

4. DISCUSSION

The results obtained here replicated the finding by Flege, Frieda and Nozawa [5] that native speakers of Italian who continue to speak their L1 frequently have significantly stronger FAs in English than do individuals who speak their L1 infrequently. The L1 use effect obtained here was much the same for early and late bilinguals. Another result of [5] replicated here is that, on the whole, AOA had a greater effect on L2 production accuracy than L1 use. Future research is needed to determine if amount of L1 use will account for more variance than AOA when the differences in AOA are smaller than in the present study (as, for example, would be expected for groups having AOAs of 10 and 16 years).

There are at least two possible explanations for why Italian/English bilinguals who often speak Italian have stronger FAs in English than bilinguals who seldom speak Italian. One might argue that subjects in the HiUse groups did not get enough "practice" in English. This, however, is very unlikely, because subjects in all four NI groups had been living in Canada for an average of 36 years. Alternatively, the L1 systems of the HiUse subjects may have been relatively strong and, as a result, had a greater influence on the L2 than was the case for the LoUse subjects. Unfortunately, this hypothesis cannot be directly assessed, for the NI subjects' performance in the L1 was not investigated. Moreover, as in the study by Flege, Frieda and Nozawa [5], one might question if subjects in the LoUse groups examined here actually had a functioning L1 system. To reliably answer this question, performance in the L1 would also have to be investigated. It appears reasonable to assume that all of the late L2 learners will have had an active L1 system at some point, even if they stopped using the L1 frequently. The subjects in the EarlyLo group, on the other hand, may never have developed full competence in the L1. It is striking, however, that the early bilinguals did not show a greater effect of L1 use than the late bilinguals. This seems to indicate that the two groups of early bilinguals did *not* differ drastically from the two groups of late bilinguals in terms of the absence vs. the presence of active L1 systems. In any case, to determine conclusively (a) whether bilingual subjects who have been living in an L2 environment for many years still have a functioning L1 system and (b) to which extent a strong L1 system may have an influence on the L2 system, it would be necessary to assess performance in both the L2 and the L1. Guion, Flege, and Loftin [this volume] found an

L1 use effect and assessed performance in both the L2 and the L1. However, their results are not really comparable to the results of this study due to differences in the subjects examined. In contrast to the subjects of the present study, who were living in a predominantly monolingual L2 environment, Guion et al.'s subjects were living in a truly bilingual community and therefore at least used their L1 on a regular basis.

Finally, it is of interest to address the question of whether at least some of the NI subjects tested in this study had achieved a nativelike level of performance in the pronunciation of English. Flege, Munro and MacKay [6] and Bongaerts et al. [1] considered their bilingual subjects to have spoken English sentences without a foreign accent if they received a mean rating that fell within 2.0 standard deviations of the mean rating obtained for the subjects in the NE control groups. If this criterion of nativelike performance is adopted, it turns out that 11 of the 72 NI subjects tested in this study had actually received scores comparable to the ones given to the native speakers of English. Nine of these belonged to the EarlyLo group and two to the EarlyHi group. A Chi-square test revealed that significantly more people in the EarlyLo group fell within the NE range than in the EarlyHi group at the $p < .02$ level. In contrast to the subjects tested by Bongaerts et al. [1], however, none of the late learners received scores that fell within the NE range.

There are at least two explanations for the finding that none of the late learners received ratings that fell within the NE range. First, the late Italian/English bilinguals who participated in this study may not have been as motivated as the late learners of the Bongaerts et al. [1] study to learn their L2. The subjects tested here were immigrants engaged in a wide range of professions, whereas all but two of the subjects tested by Bongaerts et al. [1] were university-level teachers of English who considered it very important to speak English without a noticeable Dutch accent. Second, Dutch is typologically more closely related to English than is Italian. Therefore the relatively low ratings given to the sentences spoken by the late learners in this study may also have been due to the fact that the L1/L2 differences were greater in this study than in the Bongaerts et al. [1] study. Therefore, another question that should be addressed in future FA research is to what extent typological differences between an L1 and an L2 can account for differences in FA ratings obtained for groups of learners with different L1 backgrounds.

In summary, the results of this study do not disprove the existence of a critical period for L2 speech learning. However, they strongly corroborate the view that ultimate attainment in the pronunciation of an L2 is dependent on various factors, not just on the state of neurological development at the age of first intensive exposure to the L2. It has now been shown several times that amount of L1 use and, to a lesser extent, length of residence in an L2 environment [2, 6] also affect L2 production accuracy. According to the findings of Bongaerts et al. [1], other factors that might have an influence on the strength of perceived FAs include (1) the amount of intensive training learners have received in the production and the perception of an L2 and (2) the degree of motivation learners have to speak an L2 well.

ACKNOWLEDGMENTS

This research was funded by a grant from the National Institute of Deafness and Communication Disorders (DC00257). The authors would like to thank Anna Boron for help obtaining the recordings used here and Susan Guion for comments on a previous version of this paper.

NOTES

1. Jim Flege could not be listed among the authors because of Congress restrictions.

REFERENCES

- [1] Bongaerts, T., van Summeren, C., Planken, B. and Schils, E. 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447-465.
- [2] Flege, J. E. 1988. Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America*, 84, 70-79.
- [3] Flege, J. E. and W. Eefting. 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185-202.
- [4] Flege, J. E. and K. L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91, 370-389.
- [5] Flege, J. E., Frieda, E. M. and Nozawa, T. 1997. Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25, 169-186.
- [6] Flege, J. E., Munro, M. J. and MacKay, I. R. A. 1995. Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97, 3125-3134.
- [7] Lenneberg, E. 1967. *Biological foundations of language*. New York: Wiley.
- [8] Patkowski, M. 1990. Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11, 73-89.
- [9] Scovel, T. 1969. Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19, 245-253.
- [10] Scovel, T. 1988. *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House/Harper & Row.
- [11] Seliger H. 1978. Implications of a multiple critical periods hypothesis for second language learning. In Ritchie, W. (ed.), *Second language acquisition research. Issues and implications*. New York: Academic Press.
- [12] Walsh, T. and K. Diller 1981. Neurolinguistic considerations on the optimum age for second language learning. In Diller, K. (ed.), *Individual differences and universals in language learning aptitude*. Rowley, MA: Newbury House.