Applied Psycholinguistics **27** (2006), 157–183
Printed in the United States of America
DOI: 10.1017.S0142716406060231

Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent

IAN R. A. MACKAY
University of Ottawa

JAMES E. FLEGE and SATOMI IMAI
University of Alabama at Birmingham

Received: October 24, 2004 Accepted for Publication: February 24, 2005

ADDRESS FOR CORRESPONDENCE
James E. Flege, Speech & Hearing Sciences, CH19 Room 301, University of Alabama at Birmingham, 1530 3rd Avenue, South, Birmingham, AL 35294-2042. E-mail: jeflege@uab.edu

ABSTRACT
Immigrants' age of arrival (AOA) in a country where a second language (L2) must be learned has consistently been shown to affect the degree of perceived L2 foreign accent. Although the effect of AOA appears strong, AOA is typically correlated with other variables that might influence degree of foreign accent. This study examined the pronunciation of English by native Italian immigrants to Canada who differed in AOA. As in previous research, those who arrived as young adults (late learners) were somewhat older at the time of testing, and produced somewhat longer English sentences, than those who arrived in Canada when they were children (early learners). The results of Experiment 1 showed that the greater chronological age of early than late learners was not responsible for the late learners' stronger foreign accents. Experiment 2 suggested that the late learners' longer L2 sentences were not responsible for observed early–late foreign accent differences. A principle components analysis revealed that variation in AOA and language use, but not chronological age or sentence duration, accounted for a significant amount of variance in the foreign accent ratings. The findings of the study were interpreted to mean that AOA effects on foreign accent are due to the development of the native language phonetic system rather than to maturational constraints on L2 speech learning.

The overall degree of perceived foreign accent in sentences produced in a second language (L2 foreign accent) has been investigated frequently because it provides a useful index of L2 pronunciation accuracy. Of the several factors known to influence L2 foreign accent, the strongest seems to be the age at which participants were first exposed to the L2. In studies examining immigrants who began learning their L2 upon arrival in a predominantly L2-speaking country, the "age of first exposure" variable has often been indexed by age of arrival (AOA). One robust finding of the literature dealing with L2 speech acquisition is that the later the AOA,

the stronger foreign accents in an L2 tend to be (e.g., Flege, Munro, & MacKay, 1995a; Flege, Yeni-Komshian, & Liu, 1999; Tahta, Wood, & Loewenthal, 1981).

Munro (1998, p. 139) defined foreign-accented speech as "nonpathological speech produced by L2 learners that differs in partially systematic ways from the speech characteristics of native speakers." The degree of foreign accent is usually quantified by having native speakers of the target L2 rate utterances spoken by L2 learners and a smaller number of native speakers using an equal-appearing interval scale (for discussions of methodology, see Flege & Fletcher, 1992; Southwood & Flege, 1999).

Munro and Derwing (1998, p. 160) posited that degree of L2 foreign accent depends on the extent to which an L2 learner's speech is "perceived to differ from native speaker norms." In fact, previous work has established that divergences from L2 phonetic norms for vowels, consonants, and prosody (aspects of rhythm and intonation) contribute to the perception of L2 foreign accent (e.g., Anderson-Hsieh, Johnson, & Koehler, 1992; Cunningham-Andersson & Engstrand, 1989; Flege et al., 1995a; Flege & Munro, 1994; Flege, Munro, & MacKay, 1995b; Jonasson & McAllister, 1972; Magen, 1998; Munro, 1995; Munro, Flege, & MacKay, 1996; Schairer, 1992; Willems, 1982). It is less certain, however, why certain individuals diverge more from L2 phonetic norm, and thus have stronger foreign accents, than others do.

Some researchers maintain that L2 foreign accents get stronger as the age of first exposure to the L2 increases because a critical period for speech learning ends at the age of 12 years (Scovel, 1988) or 15 years (Patkowski, 1990; see also DeKeyser, 2000). On this view, AOA effects are due to a loss of neural "plasticity" that accompanies normal maturation (Scovel, 2000). As discussed by Moyer (1999, p. 82), it is widely believed that individuals who learn an L2 after the end of a critical period are subject to "neurological or motor skill constraints" that will render nativelike speech production "highly unlikely or impossible."

It has proven difficult, however, to establish a clear link between neurological maturation and degree of L2 foreign accent. For one thing, the link between the age of first exposure variable (as indexed by AOA) and the state of neurological development is tenuous. This is because uncertainty exists as to which specific parameters of brain maturation directly influence the speed or completeness of first language (L1) and/or L2 speech learning. A further complicating factor is that parameters of neurological development that might potentially affect speech learning have differing maturational schedules (see, e.g., Bates, 1999).

Another type of explanation for age of learning effects on degree of L2 foreign accent is that the influence of the L1 sound system on subsequent L2 speech acquisition becomes stronger as the language-specific phonetic categories of the native language (L1) become more robust. The speech learning model (SLM) proposes that as L1 phonetic categories develop, they become stronger "attractors" of vowels and consonants encountered in an L2 (see Flege, 1999, 2002, 2003). That is, as L1 phonetic categories develop, L2 learners may become more likely to process instances of L2 vowel and consonant categories as instances of previously established L1 categories, even when physical differences between realizations of the functionally equated L1 and L2 categories can be detected auditorily (Flege, 1995; for supporting evidence, see Baker, Trofimovich, Mack, & Flege, 2002). According

to the SLM, equivalence classification blocks the formation of new categories for certain L2 vowels and consonants which, in turn, precipitates the development of "merged" L1–L2 categories that incorporate the properties of perceptually equated L1 and L2 sound categories. Research carried out within the framework of an Hebbian learning model (McCandliss, Fiez, Protopapas, Conway, & McClelland, 2002) and the native language magnet model (e.g., Iverson et al., 2003) has also led to the proposal that L1 category development may negatively impact L2 speech learning.

A problem for accounts that attribute AOA effects to limitations imposed by maturational constraints, as well as for accounts that attribute AOA effects to the development of L1 phonetic categories, is that the L2 exposure age (i.e., AOA) may covary with other factors that impact L2 speech acquisition (see, e.g., Bahrick, Hall, Goggin, Bahrick, & Berger, 1994; Flege, 1987). For example, in research examining immigrants who learn their L2 in an L2-speaking country, a relatively late AOA is associated with a relatively high self-reported percentage use of the L1 and a correspondingly low percentage use of the L2 (Flege, 1998). Other research has shown that variation in percentage L1/L2 language use affects degree of L2 foreign accent (Flege, Frieda, & Nozawa, 1997; Guion, Flege, & Loftin, 2000).

To determine if AOA and language use affect L2 foreign accent independently, Piske, MacKay, and Flege (2001) tested Italian immigrants who had lived in Canada for many years. These authors recruited four groups of participants who differed orthogonally in AOA (early vs. late) and amount of continued Italian use (high-L1 use vs. low-L1 use). As expected, the participants who arrived in late adolescence or early adulthood (late learners) had stronger foreign accents than those who arrived in childhood (early learners). In addition, both early and late learners who used Italian relatively often had stronger foreign accents in English than AOA-matched participants who used Italian seldom. The aim of the present study was to evaluate the effect of two additional variables that might affect L2 foreign accent but are typically confounded with AOA: chronological age and the duration of L2 sentences.

One might hypothesize that the chronological age of nonnative speakers at the time of test will influence degree of L2 foreign accent. In research examining native Italian (NI) immigrants to Canada (Flege et al., 1995a) and native Korean immigrants to the United States (Flege et al., 1999; $n = 240$ per study), age at test was positively correlated with AOA (Italians, $r = 0.52$; Koreans, $r = 0.68$). Significant correlations existed between age, AOA, and the foreign accent ratings obtained in both studies. The later the participants arrived in an L2-speaking country and the older they were at the time of test, the stronger their foreign accents tended to be. During the process of normal aging, cognitive processing efficiency, including the ability to suppress task-irrelevant information, decreases (for a review, see Bialystok, Craik, Klein, & Viswanathan, 2004). As speakers of an L2 grow older they might, therefore, experience greater difficulty in suppressing the influence of the L1 phonetic subsystem on L2 production. It is therefore possible that the negative effect attributed to a relatively late exposure age (i.e., AOA) in previous L2 speech research might have been due, at least in part, to differences in chronological age.

Two previous studies provided correlational evidence in support of the hypothesis that chronological age affects foreign accent. Tahta et al. (1981) tested 109

immigrants differing in L1 background who had learned English in the United Kingdom. The participants were selected primarily on the basis of their age of first exposure to English, which ranged from 6 to 15+ years. However, the participants also differed widely in age (range = 9–77 years, $M = 27$). In a stepwise multiple regression analysis, both age of first exposure and amount of self-reported English use in the home accounted for a significant amount of variance in foreign accent ratings. The participants' chronological age accounted for a small but significant amount of additional variance when entered following these two variables. More recently, Flege, MacKay, and Imai (2005) obtained a significant correlation between degree of foreign accent and chronological age for 36 NI late learners of English. The correlation remained significant when the effects of AOA and language use were partialed out.

The duration of L2 utterances produced by nonnative speakers might also affect degree of L2 foreign accent. Nonnative speakers' speaking rate in an L2 tends to be slower than that of native speakers, meaning that their L2 utterances are typically longer in duration (e.g., Pennington, 1992; Raupauch, 1980; Riggenbach, 1991). Guion, Flege, Liu, and Yeni-Komshian (2000) observed a significant positive correlation between the AOA of Italian and Korean immigrants ($n = 240$ each) and the duration of English sentences the participants repeated following an aural model. Given that strong correlations existed between AOA and the immigrants' degree of foreign accent, significant correlations also existed between sentence duration and degree of foreign accent.

It is possible that sentence duration acts as a perceptual cue to degree of foreign accent independently of variation in phonetic accuracy. This might be true, for example, if a slow perceived rate of speech (i.e., relatively long sentences) conveys hesitancy to listeners. Munro and Derwing (2001) found that variation in sentence duration accounted for 15% of the variance in foreign accent ratings obtained for English sentences spoken by Chinese adults.[1] In a second experiment, these authors found that when sentences were shortened 10% using computer-editing techniques, the sentences were judged to sound less foreign accented than sentences that were unaltered or lengthened by 10%.

There is other reason to think, however, that variation in sentence duration does not cue foreign accent. Adults differ widely in their characteristic speaking rate (see, e.g., Byrd, 1994); and all normal individuals are capable of changing their rate of speech. When an English monolingual chooses to speak slowly, this does not cause her or him to be perceived as a nonnative speaker by native English (NE)-speaking listeners. If individuals who produce relatively long L2 sentences also tend to produce L2 vowels, consonants, and/or prosodic dimensions inaccurately, this could lead to the observed correlations between speaking rate and foreign accent.

THE PRESENT STUDY

Both experiments in this study examined degree of foreign accent in English sentences spoken by NI immigrants to Canada.

Experiment 1 evaluated the influence of chronological age on the English pronunciation of 138 NI participants differing in chronological age (range =

40–71 years) as well as other factors that might influence degree of foreign accent. English sentences produced by the NI participants were rated for degree of foreign accent by NE-speaking listeners. Factors derived from principle components analysis of language background questionnaire variables (including AOA, chronological age, length of residence in Canada) were regressed onto the foreign accent ratings. Factors that had high loadings on AOA and language use were found to predict degree of foreign accent, but not a factor with high loadings on chronological age and length of residence.

The statistical procedure just described obviated the problem of multicolinearity, but did not eliminate interpretative problems. This is because chronological age and length of residence were both correlated in Experiment 1 with variables that had high loadings on the "AOA" factor. An analysis of variance (ANOVA) was therefore carried out to compare the performance of six subgroups of 18 participants each that were matched for AOA but differed significantly in chronological age and length of residence in Canada.

Experiment 2 evaluated the influence of sentence duration on L2 foreign accent. The stimuli for this experiment were drawn from a recent study (MacKay & Flege, 2004) that tested NI and NE participants in Ottawa, ON. These participants first produced English sentences six times each without instruction as to speaking rate (the unspeeded condition), then as rapidly as possible (the speeded condition). Two tokens of an English sentence produced by each participant were selected for analysis. The two tokens (one from each condition) differed by about 10% in duration. The early and late learners were then subdivided according to percentage L1 use, as in the Piske et al. (2001) study, in order to evaluate the influence of AOA and language use.

The English sentence stimuli in Experiment 2 were rated by NE-speaking listeners and by listeners who had learned English as an L2 in Canada. Flege (1988) found that native Chinese adults who spoke English with strong foreign accents were able to scale foreign accent in English sentences spoken by other Chinese speakers much like NE-speaking listeners. The Chinese listeners might have rated Chinese-accented sentences on the basis of how much these sentences resembled their own accented pronunciation of English. Alternatively, the native Chinese listeners might have based their ratings on the detection of divergences from the phonetic norms of English. This second interpretation implies that the Chinese listeners were developing long-term memory representations for English vowels, consonants, and prosodic dimensions.

One aim of Experiment 2 was to help choose between the two interpretations just offered for the findings of Flege (1988). This was accomplished by recruiting a group of nonnative listeners who spoke a different L1 than did the nonnative speakers whose sentences were rated. If the foreign accent ratings obtained from native Arabic (NA)- and NE-speaking listeners were correlated, it would suggest that the NA listeners had developed perceptual representations for English vowels, consonants, and prosodic dimensions, and that they were able to detect diverges from those representations.

The primary aim of Experiment 2, however, was to determine if 10% sentence duration differences would affect the NE- and NA-speaking listeners' ratings differently. Even if the NE and NA listeners' foreign accent ratings were correlated,

Table 1. *Mean characteristics of the 138 native Italian participants in Experiment 1*

	Mean	*SD*	Range
Chronological age	57	6	40–72
Age of arrival in Canada	18	6	7–36
Length of residence in Canada	39	7	13–53
Overall English use (%)	53	18	10–94
Years of education in Canada	3	4	0–17
English proficiency	5.8	1.0	2.5–7
Overall Italian use (%)	46	19	5–90
Years of education in Italy	6	3	0–13
Italian proficiency	6.4	0.8	4–7

Note: English and Italian proficiency, means of self-estimated ability to speak and understand (1 = *poor*, 7 = *good*). The overall percentage of use of English and Italian did not sum to 100% because a few native Italian participants reported using French 0–5% of the time.

the perceptual representations they developed for English vowels, consonants, and prosodic dimensions would probably not be equivalent to the NE listeners'. If the NA-speaking listeners detected fewer divergences from English phonetic norms, then they might show a greater reliance on sentence duration as a cue to foreign accent than the NE listeners did. The inclusion of the NA listener group, therefore, created an additional opportunity to provide evidence that early–late differences in degree of foreign accent can be attributed to confounded differences in sentence duration.

EXPERIMENT 1

Experiment 1 evaluated the effect of chronological age on L2 foreign accent by comparing the pronunciation of English by NI groups whose members differed in age but not AOA.

Method

Participants. A total of 138 NI and 12 NE participants were recruited in the Ottawa, ON, region. The NE participants had all been born and raised in Canada. The NI participants were required to have arrived in Canada after the age of 6 years, to be between the ages of 40 and 75 years at the time of test, and to report using both English and Italian at least 5% of the time. All 150 participants lived in primarily English-speaking communities and passed a pure-tone hearing screening at octave frequencies between 500 and 4000 Hz (35 dB HL). The NI participants responded to a language background questionnaire that assessed places and years of residence, education, and language use. The questionnaire also provided self-assessments of proficiency in Italian and English.

As summarized in Table 1, the NI participants had a mean age of 57 years (range = 40–72) and roughly the same gender makeup as the NE participants.

Table 2. *Between-variable correlations for the 138 Italian–English bilinguals in Experiment 1*

	2	3	4	5	6	7	8	9
1. Chronological age (years)	.39*	.62*	−.13	−.32*	−.15	.15	−.13	.10
2. Age of arrival in Canada (years)		−.49*	−.32*	−.72*	−.30*	.34*	.36*	.39*
3. Length of residence in Canada			.15	.31*	.11	−.14	−.43*	−.24
4. Overall use of English (%)				.43*	.19	−.99*	−.09	−.31*
5. Education in Canada (years)					.40*	−.45*	−.44*	−.48*
6. Proficiency in English						−.21	−.08	.08
7. Overall use of Italian (%)							.06	.30*
8. Education in Italy (years)								.34*
9. Proficiency in Italian								

Note: English and Italian proficiency, means of self-estimated ability to speak and understand (1 = *poor*, 7 = *good*).
*$p < .001$.

The NI participants had arrived in Canada at an average age of 18 years. They had received more education in Italy than Canada, and reported using Italian more than English. Despite their many years of residence in Canada ($M = 39$ years), the NI participants rated themselves as more proficient in Italian than English.

The NI participants' AOA in Canada was correlated with all eight other variables shown in Table 1 ($p < .01$). As summarized in Table 2, an early AOA in Canada was associated with a relatively young chronological age at the time of testing, a long residence in Canada, frequent use of English, infrequent use of Italian, many years of education in Canada, and few years of education in Italy. An early AOA was also associated with a relatively high self-rated proficiency in English and low self-rated proficiency in Italian.[2] A long residence in Canada was associated with a relatively old chronological age at the time of testing and few years of education in Italy. High self-rated proficiency in Italian was associated with relatively many years of education in Italy and frequent use of Italian.

Speech elicitation. The speech samples were elicited entirely in English by NE experimenters who did not speak Italian. A series of "questions" and "answers" recorded by two NE speakers (one male, one female) was used to elicit sentence production. The elicitation materials were digitized (22.05 kHz), and then presented to participants via a loudspeaker. The participants heard a female voice ask a question (e.g., *What's that woman's name?*) followed by an answer given by a male voice (e.g., *She's called Patty*). After a second presentation of the question

(i.e., the female voice), the participants repeated the "answer" given by the male voice. Eight different question-answer-question (Q-A-Q) sequences were used to elicit the eight test sentences (see Appendix A). The participants were allowed to rehear a Q-A-Q sequence if they wanted, and to repeat a test sentence if they were not satisfied with their production of it.

The eight English test sentences were elicited three times each. Translation equivalent Italian questions and answers were used to elicit the production of eight Italian sentences by the NI participants. (The NE participants heard, but did not repeat the Italian sentences.) The Italian sentences were elicited between the first and second elicitations of the English test sentences, and then a second time during the third elicitation of the English sentences. During the third elicitation, English sentence 1 was elicited first, then Italian sentence 1, then English sentence 2, and so on. Interspersing English and Italian was intended to increase the influence of the Italian phonetic system on the NI participants' pronunciation of English. However, in a study involving a subset of the NI participants tested here, Flege et al. (2005) found that there was not a significant difference in strength of foreign accents for the three elicitations of the English test sentences used here. Therefore, just the second tokens of the English sentences were examined in this experiment unless there was some specific problem with it.

Stimulus preparation. The sentences were recorded on a portable DAT recorder and later down sampled to 22.05 kHz on a PC. The NI participants, especially older ones, did not always repeat the English target sentences word for word. The least amount of variation existed for sentence 2 (*We were neighbors for 30 years*), sentence 3 (*Billy is Patty's husband*), and sentence 7 (*We'll have pork chops and potatoes*), but some variation existed even for these sentences. For example, some NI participants incorrectly repeated sentence 2 as "We know the neighbors for thirty years" or "We are neighbors for thirty years"; and some participants used the uncontracted form "We will" when repeating sentence 7. It seemed likely that such changes in the target sentences would influence listeners' foreign accent ratings. The decision was made, therefore, to examine only the invariant portions of sentences 2, 3, and 7, which have been underlined above. These phrases were edited out of the original sentence and then normalized for peak intensity.

The procedure used in recent foreign accent research (e.g., Flege et al., 1995a) has been to obtain ratings for multiple sentences that are presented in separate, counterbalanced blocks. In the absence of interactions involving the sentence factor, the standard procedure has been to conduct analyses of ratings averaged across test sentences. A pilot test revealed that, as expected from previous research, much the same results were obtained for the three test phrases described earlier. Therefore, to shorten and facilitate the rating task, the decision was made to concatenate the three phrases produced by each participant into a single waveform, with short pauses between the phrases. This procedure yielded 150 stimuli, one for each participant.

Foreign accent ratings. The stimuli just described were presented to native speakers of Canadian English (nine male, nine female) between the ages of 34 and

56 years ($M = 46$ years) who had received 11–20 years of formal education ($M = 16$ years). The listeners had been raised in households where only English was spoken, attended an English-language elementary and high school in Ontario or a western Canadian province, and reported not speaking any language other than English well or often. All reported having normal hearing, and passed a pure-tone hearing screening prior to participating.

The listeners were tested individually in a sound booth at the Phonetics Laboratory of the University of Ottawa in a single session lasting about 45 min. Three separate randomizations of the 150 stimuli were presented via a loudspeaker at a self-selected comfortable level. The listeners rated each stimulus using a scale that ranged from 1 (*very strong foreign accent*) to 9 (*no foreign accent*). The software ensured that a response would not be accepted until the entire stimulus (i.e., all three phrases) had been played out.

The listeners were urged to use the entire scale in judging the stimuli, and were told to guess if uncertain. The stimuli could be replayed, but ratings could not be changed once given. No instruction or training on use of the rating scale was provided. However, 50 extra stimuli (randomly selected from the set of 150) were presented at the beginning of the test session for familiarization, and were not analyzed.

The median of the three ratings given by each listener to each stimulus was determined. There was a high degree of interrater agreement in the median ratings obtained for the 150 stimuli. The simple correlations among the median ratings obtained for all possible pairings of the 18 listeners averaged $r = .86$ (range = .61–.95). A high intraclass correlation was also obtained for ratings obtained from all 18 listeners, $R = .985$, $F\ (149, 2533) = 100.5$, $p < .01$. The decision was made, therefore, to calculate a single rating for each of the 150 stimuli by averaging over the ratings obtained from the 18 listeners.

Results

Figure 1 plots the average foreign accent ratings as a function of the NI participants' AOA in Canada. As expected from previous work (Flege et al., 1995a, 1999), a strong relationship existed between the NI participants' AOA in Canada and degree of L2 foreign accent ($r = -.75$, $p < .01$). Somewhat more scatter around the line of best fit is evident for the latest-arriving NI participants. The AOA-foreign accent correlation increased to $r = -.80$ when the ratings obtained for five participants with an AOA greater than 28 years were excluded.

The ratings obtained for the 12 NE participants have also been plotted in Figure 1, with AOA assigned a value of "0" years (because the NE participants were born in Canada). The NE participants' ratings ranged from 8.3 to 8.9 ($M = 8.63$, $SD = 0.18$). Just seven of 138 NI participants obtained ratings that fell within the NE range. These same seven NI participants also met a criterion that has been used as an index of "accent-free" L2 pronunciation (e.g., Piske et al., 2001), that is, obtained a rating that fell within 2 *SD*s of the mean rating obtained for the NE group. As expected from previous research (Flege et al., 1995a), all of these NI participants had arrived in Canada prior to the age of 10 years.

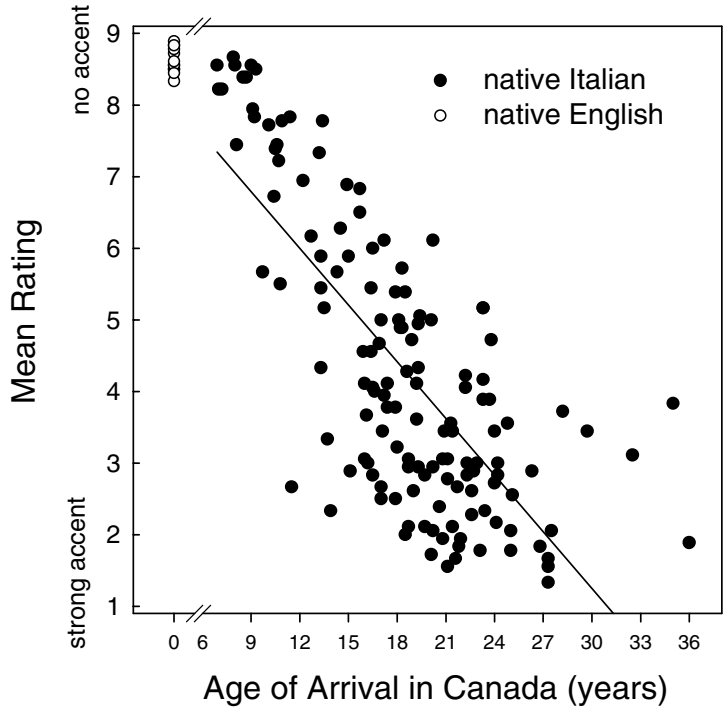

Figure 1. The relationship between the NI participants' AOA in Canada and the overall degree of perceived foreign accent in English. Also included are the ratings for 12 native speakers of English.

Regression analysis. As shown previously in Table 2, a number of language background questionnaire variables that might predict degree of L2 foreign accent were correlated with AOA as well as with one another. A principle components analysis is useful when, as in the present study, there is a high degree of correlation among predictor variables and the internal structure of those variables is unlikely to be apparent in a multiple regression analysis (Metz, Samar, Schiavetti, Sitler, & Whitehead, 1985, p. 351). A principle components analysis was therefore undertaken to identify underlying factors in the questionnaire data. The derived factors were then regressed onto the foreign accent ratings obtained for the NI participants.

Seven of the 10 variables submitted to the principle components analysis have been summarized in Table 1: chronological age, AOA in Canada, length of residence in Canada, overall percentage use of English, overall percentage use of Italian, years of education in Italy, and Italian proficiency.[3] The remaining three variables were self-reported percentage use of Italian at home, percentage use of Italian in social contexts, and a variable that will be referred to as "bilingual dominance." Immigrants' L2 use at home and with friends has accounted for variance in degree of L2 foreign accent in previous research (Flege et al., 1995a,

Table 3. *Loadings on three factors derived in a principle components analysis of 10 language background questionnaire variables*

Variable	Loadings		
	Factor 1	Factor 2	Factor 3
Overall use of Italian (%)	.925	—	—
Overall English use (%)	−.920	—	—
Home use of Italian (%)	.786	—	—
Social use of Italian (%)	.754	—	—
Age of arrival in Canada	—	.802	—
Italian proficiency	—	.682	—
Bilingual dominance	—	.668	—
Chronological age	—	.307	.909
Length of residence in Canada	—	−.393	.848
Years of education in Italy	—	.616	−.423
Variance	3.05	2.29	1.77

Note: For clarity, only loadings greater than .300 are shown.

1995b; Tahta et al., 1981).[4] The index of bilingual dominance used here was the ratio of self-reported proficiency in English and Italian (see Flege, MacKay, & Piske, 2002). Bilingual dominance was included because of preliminary evidence (Piske et al., 2001) that early L2 learners who become L2 dominant, but not those who remain L1 dominant, may evade being detected as nonnative speakers of the L2.[5]

The principle components analysis yielded three factors having eigenvalues greater than 1.0 (range = 1.39–3.88). These three derived factors accounted for 100% of the variance in the 10 language background questionnaire variables.

As summarized in Table 3, four variables had high (>.600) loadings on Factor 1: overall percentage use of English, overall percentage use of Italian, percentage use of Italian at home, and percentage use of Italian in social contexts. Given that all four of these variables pertained to language use, Factor 1 will be designated the "language use" factor.

Four variables had high loadings on Factor 2: AOA in Canada, Italian proficiency, bilingual dominance, and years of education in Italy. Given that AOA had the highest loading on Factor 2, and AOA is widely believed to influence L2 learning, Factor 2 will be designated the "AOA" factor.

Just two variables had high loadings on the third factor: chronological age and length of residence in Canada. Factor 3 will therefore be designated the "age/length of residence" factor. As already mentioned, a strong correlation was found to exist between the NI participants' AOA and degree of foreign accent. However, given that AOA was correlated with both length of residence ($r = -.49$) and chronological age ($r = .39$), effects attributed to the NI participants' age of first exposure to English (i.e., AOA) might actually be due, at least in part, to variation in their age and/or length of residence in Canada. If so, then Factor 3 should account for a significant amount of variance in foreign accent ratings.

Table 4. *Stepwise multiple regression analyses examining the overall degree of foreign accent in the English speech of 138 native Italian speakers of English*

	Adj. R^2	B	β	F	p
Factor 2: AOA	.451	−0.75	−0.37	45.6	.000
Factor 1: language use	.139	−1.36	−0.67	148.3	.000

Note: Factors 1 and 2 were derived by a principle components analysis (see Table 3).

Scores for the three derived factors were obtained for the 138 NI participants by applying scoring coefficients generated by the principal components analysis to standardized values for each variable. The scores were then submitted to a stepwise multiple regression analysis examining the ratings obtained for each participant. As summarized in Table 4, Factor 2 (AOA) accounted for 45% of the variance and Factor 1 (language use) accounted for 14% of the variance. However, Factor 3 (age/length of residence) was not found to account for a significant additional amount of variance in the foreign accent ratings ($F = .12$, $p = .74$).

ANOVA. The analysis just presented suggests that neither chronological age nor length of residence in Canada was predictive of the NI participants' degree of foreign accent. It is noteworthy, however, that significant correlations existed between the NI participants' chronological age and three of the four variables with high loadings on Factor 2 (AOA, $r = .38$; Italian proficiency, $r = .39$; bilingual dominance, $r = .19$; all p values $<.05$), and significant correlations existed between the NI participants' lengths of residence and all four variables with high loadings on Factor 2 (AOA, $r = -.49$; Italian proficiency, $r = -.24$; bilingual dominance, $r = .17$; years of education in Italy, $r = -.43$; $p < .05$). Thus, although the derived factors submitted to the multiple correlation analysis were not correlated, the pattern just described makes it impossible to rule out an influence of chronological age and/or length of residence on degree of foreign accent.

To further evaluate the influence of chronological age and length of residence on the NI participants' degree of foreign accent, a subgroup matching process was used (see Flege et al., 1999). This process involved developing subgroups of the 138 NI participants whose members differed in chronological age but not AOA. To begin, all 138 participants were rank ordered according to chronological age. The 23 participants who were youngest in age (range = 40.4–51.3 years) were assigned to Group 1; the next 23 youngest (range = 51.5–55.1 years) were assigned to Group 2, and so on (Group 3 = 55.1–58.3, Group 4 = 58.7–60.8, Group 5 = 60.9–63.4, Group 6 = 63.5–70.9). However, a preliminary ANOVA indicated that participants in these age-defined subgroups differed in AOA ($p < .01$). A Tukey post hoc test revealed that participants assigned to Group 6 arrived later in Canada (mean AOA = 23 years) than did those in Groups 1–3 (mean AOA = 16 years for all three groups; $p < .05$).

An additional step was taken, therefore, to form subgroups of participants differing in age but not AOA (Table 5). The five participants in Group 1, Group 2,

Table 5. *Characteristic means (SD) of native Italian groups whose English pronunciation was examined in Experiment 1*

	Groups Defined by Chronological Age					
	Age46	Age54	Age57	Age60	Age62	Age65
Chronological age (years)	46 (3)	54 (1)	57 (1)	60 (1)	62 (1)	65 (1)
Age of arrival in Canada (years)	18 (4)	18 (4)	18 (5)	18 (4)	18 (4)	20 (4)
Length of residence in Canada (years)	29 (6)	36 (4)	39 (5)	42 (4)	44 (4)	44 (4)
Overall use of English (%)	52 (15)	56 (20)	54 (14)	48 (20)	47 (16)	52 (20)
Education in Canada (years)	2 (2)	3 (4)	3 (4)	1 (3)	1 (2)	1 (1)
Proficiency in English	5.8 (0.9)	5.7 (0.7)	5.9 (0.7)	5.5 (1.3)	5.4 (1.0)	5.8 (1.2)
Overall use of Italian (%)	47 (15)	43 (21)	44 (16)	51 (21)	53 (16)	48 (20)
Education in Italy (years)	9 (3)	7 (4)	6 (2)	6 (2)	5 (1)	7 (2)
Proficiency in Italian	6.6 (0.5)	6.2 (0.8)	6.6 (0.6)	6.2 (0.6)	6.5 (0.8)	6.6 (0.6)

Note: Group assignment was based on chronological age at the time of test (see text). $N = 18$ per group. Proficiency in English and Italian, average self-reported ability to speak and understand English ($1 = poor$, $7 = good$).

and Group 3 with the earliest AOA (15 participants in all) were removed from these groups, and the five participants each in Group 4, Group 5, and Group 6 with the latest AOA (again, 15 participants) were removed from those groups. This procedure yielded subgroups ($n = 18$ each) whose members did not differ significantly in AOA, $F\ (5, 102) = 1.4$ ($p > .10$). The new age-defined subgroups were named according to the mean chronological ages of their members: Age46 (Group 1), Age54 (Group 2), Age57 (Group 3), Age60 (Group 4), Age62 (Group 5), and Age65 (Group 6).

Additional ANOVAs examining the age-defined subgroups revealed that their members did not differ significantly in overall percentage use of English, $F\ (5, 102) = .7$, overall percentage use of Italian, $F\ (5, 102) = .9$, English proficiency, $F\ (5, 102) = .6$, Italian proficiency, $F\ (5, 102) = 1.7$, or years of education in Canada, $F\ (5, 102) = 1.8$ (all p values $> .10$). The subgroups did differ significantly, however, in length of residence in Canada, $F\ (5, 102) = 35.0$, $p < .01$. A Tukey test revealed that participants in Age46 had a shorter residence (mean = 29 years) than those in Age54 and Age57 ($M = 36$ and 39 years) who, in turn, had a shorter residence than Age60 to Age65 ($M = 42$–44 years, $p < .05$). These differences were expected because, as discussed by Hakuta and Bialystok (1994, p. 66), groups differing in chronological age (say, 50 vs. 60 years) that are matched for AOA (say, 20 years for both) will necessarily differ in length of residence (30 vs. 40 years). This is because the length of residence of each participant is calculated

as chronological age at test minus chronological age at the time of arrival in an L2-speaking country (i.e., AOA).[6]

An ANOVA was carried out to examine the ratings obtained for the six subgroups differing in chronological age and length of residence but not AOA or language use. The foreign accent ratings obtained for the six subgroups were not found to differ significantly, $F\ (5, 102) = .8$, $p > .10$. This finding agrees with the results of the principle components analysis presented earlier.

Discussion

The aim of this experiment was to determine if chronological age affects degree of L2 foreign accent. When factors derived from a principle components analysis were regressed onto the foreign accent ratings obtained for 138 NI participants, factors with high loadings on AOA and estimates of language use accounted for 59% of the variance. A third factor with high loadings on chronological age and length of residence did not account for a significant additional amount of variance. Moreover, the ratings obtained for subgroups matched for AOA but differing in age and length of residence did not differ significantly.

The results obtained here suggest that age at the time of test does not affect degree of foreign accent. However, two caveats are in order. First, the youngest and oldest groups spanned an age range of just 19 years (means = 46 and 65 years, respectively). The foreign accent ratings obtained for groups representing a wider chronological age range might have differed. Second, the lack of a significant between-group difference does not definitively disprove the existence of an effect. A more sensitive test (e.g., one evaluating longer speech samples) might have yielded an effect of age.

EXPERIMENT 2

As discussed in the introductory section, late learners often produce longer L2 utterances than early learners do. Munro and Derwing (2001) found that when Chinese later learners' sentences were shortened by 10% using a computer-editing technique, the sentences were judged by listeners to be less foreign accented. This led us to hypothesize that if nonnative speakers increase their rate of speech by 10%, they will sound less foreign accented.

The stimuli examined in this experiment were English sentences drawn from a recent study by MacKay and Flege (2004). The early and late NI participants in that study repeated English and Italian sentences following an aural model, first without instruction as to speaking rate (the "unspeeded" condition), then as rapidly as possible (the "speeded" condition). The members of a NE control group produced only English sentences. Two English sentences spoken by each participant (one from the unspeeded condition, the other from the speeded condition) were rated here for foreign accent. A speed-up could not diminish the foreign accent of participants without detectable foreign accents. Thus, our hypothesis predicted that the NI participants' whose sentences were detectably foreign accented in the

unspeeded condition would receive higher ratings (i.e., to be heard as less strongly accented) for sentences they produced in the speeded condition.

A secondary aim of this experiment was to evaluate nonnative speakers' ability to scale foreign accent in L2 sentences produced by nonnative speakers from a different L1 background. Sentences spoken by the NI and NE participants were rated both by NE-speaking listeners and by NA-speaking listeners. Like the NI participants whose sentences were to be rated, the NA-speaking listeners learned English as an L2 in Canada.

Method

Participants. MacKay and Flege (2004) elicited sentence production by 16 NE monolinguals and 64 native speakers of Italian. The NI group consisted of 32 early learners with a mean AOA of 8 years (range = 3–13 years) and 32 late learners with a mean AOA of 19 years (range = 15–28 years). Here, the early and late learners were subdivided according to percentage L1 use (2–15 vs. 30–75%), yielding subgroups designated "early-low" (early learners with a relatively infrequent use of Italian), "early-high" (early learners with a relatively frequent use of Italian), "late-low" and "late-high."

Stimuli. The NI participants produced four sentences[7] a total of six times each without instruction as to speaking rate (the unspeeded condition), and then as rapidly as possible (the speeded condition). Just one sentence, *The woman was not in critical condition*, was examined here. The duration differences between all pairwise combinations of the six tokens produced in the two conditions by each participant were determined. Tokens were excluded if they were produced with a pause and/or a dysfluency, defined as the presence of a prolongation of a phonetic segments(s) or the repetition of part or all of a syllable. This meant that there were fewer than 36 pairings available for some participants.

An attempt was made to find sentence pairs whose durations differed by 9–13%. When more than a single sentence pair was available for a participant, the selection was guided by the desire to reduce overall differences in sentence duration between early and late learners. This involved selecting shorter than average sentences for late learners, and longer than average sentences for early learners. Appropriate sentence pairs could be identified for 49 NI and 11 NE participants. Table 6 summarizes characteristics of these participants. Note that NI participants with low self-reported Italian use were matched fairly well to those with a higher use of Italian for AOA (M = 13 vs. 14 years), chronological age (M = 50 vs. 47 years), and length of residence in Canada (M = 38 vs. 33 years).

Table 7 summarizes the duration of the sentence stimuli. On average, sentences drawn from the speeded condition were 10% shorter than those drawn from the unspeeded condition. There was relatively little overall difference in sentence duration between early and late learners. Sentences produced by the late-low group averaged just 2% longer than those produced by the early-low group; and the late-high groups' sentences averaged 5% longer than the early-high groups'. A Speaking Condition (unspeeded vs. speeded) × Group (early vs. late) ANOVA examining the stimulus duration values yielded a significant main effect of

Table 6. *Characteristic means (SD) of the five groups of participants whose pronunciation of English was examined in Experiment 2*

	Native English	Early-Low	Early-High	Late-Low	Late-High
N (male/female)	5/6	5/8	6/5	3/9	6/7
Chronological age (years)	50 (5)	50 (4)	47 (6)	51 (6)	47 (8)
Age of arrival in Canada (years)	—	8 (3)	8 (4)	18 (3)	19 (4)
Overall use of Italian (%)	—	7 (3)	41 (11)	11 (5)	52 (16)
Length of residence in Canada (years)	—	42 (4)	39 (6)	33 (5)	28 (10)
Italian proficiency		4.8 (1.0)	5.6 (1.0)	6.6 (0.6)	6.5 (0.7)
English proficiency	—	6.6 (0.6)	6.3 (0.8)	5.5 (0.5)	5.3 (0.6)

Note: Group assignment was based on age of arrival in Canada from Italy (relatively early vs. late) and self-reported use of the L1, Italian (relatively high vs. low). English and Italian proficiency, mean rating of ability to speak and understand (1 = *poor*, 7 = *good*).

Table 7. *The mean (SD) duration (ms) of the Experiment 2 sentence stimuli drawn from two speaking conditions*

		Condition		
Group	*N*	Unspeeded	Speeded	Change (%)
Native English	11	1805 (173)	1601 (151)	11.3
Early-low	13	1966 (165)	1758 (150)	10.6
Early-high	11	2059 (145)	1842 (120)	10.5
Late-low	12	2017 (172)	1794 (167)	11.1
Late-high	13	2169 (302)	1933 (254)	10.8

Note: N, number of stimulus sentence pairs.

speaking condition, $F\ (1, 55) = 1848.2$, $p < .01$. However, neither the main effect of group, $F\ (1, 47) = 1.9$, $p > .10$, nor the two-way interaction reached significance, $F\ (1, 47) = 2.9$, $p = .10$.

Listeners. The sentence stimuli were presented to two groups of listeners. All NE- and NA-speaking listeners reported having normal hearing and passed a pure-tone hearing screening prior to participating. The NE listeners (seven male, eight female) ranged in age from 18 to 35 years. All of them were born and raised in Canada, and reported using no language other than English well or frequently. The NA-speaking listeners (eight males, seven female) ranged in age from 18 to 29 years. They had arrived in Canada between the ages of 15 and 27 years ($M =$ 21 years), had lived in Canada for 0.3–10 years ($M = 3$ years), estimated using

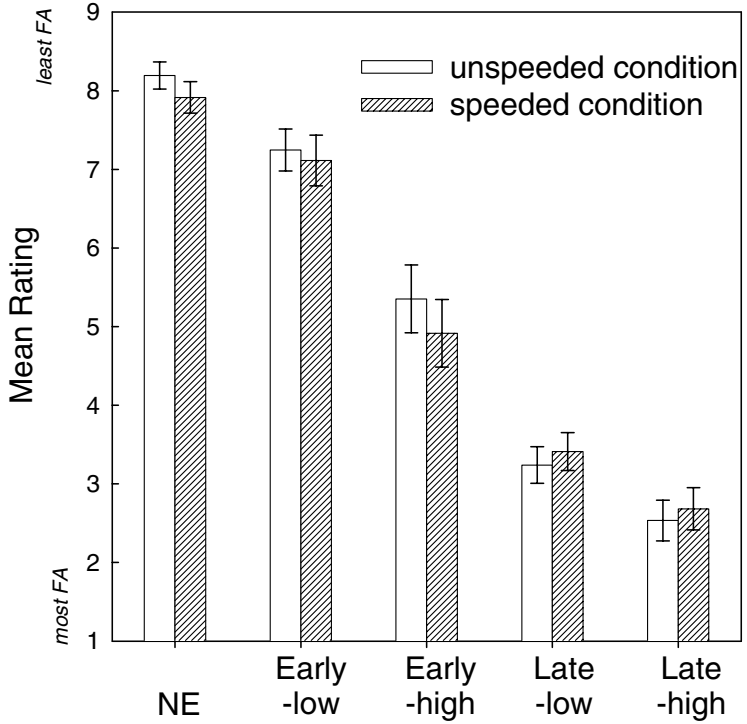

Figure 2. The mean ratings obtained from NE-speaking listeners for relatively long (unspeeded condition) and short (speeded condition) sentences spoken by five groups of talkers. The error bars bracket ±1 *SE*.

Arabic more than English ($M = 52$ vs. 35%), and reported greater proficiency in Arabic than English.

Procedure. The 120 stimuli (60 participants × 2 speaking conditions) were randomly presented four times each via a loudspeaker to the listeners, who were tested individually in a sound booth. The listeners rated each sentence using a scale that ranged from 1 (*most foreign accented*) to 9 (*least foreign accented*). The interval between each rating and presentation of the next sentence stimulus was fixed at 1.0 s. A rating was required for each sentence. The listeners were told to make their best guess if uncertain, and to use the entire scale. The median of the final three ratings given by each listener to each sentence was determined. Average values were computed for each stimulus by averaging over the median values obtained from the 15 NE-speaking listeners and from the 15 NA-speaking listeners.

Results

Figure 2 shows the mean ratings obtained from the NE-speaking listeners for sentences produced in the two speaking conditions. As expected from previous

research (Piske et al., 2001), sentences produced by the two groups of late learners (late-low, late-high) received lower ratings, indicating stronger foreign accents, than sentences produced by the two groups of early learners (early-low, early-high). In addition, the NI participants who used Italian often (late-high, early-high) received lower ratings than those who used Italian relatively seldom (late-low, early-low). Finally, the average ratings obtained for all four NI groups were lower than the NE group's ratings.

Averaged over all five groups, there was little difference in the ratings obtained for sentences drawn from the unspeeded and speeded conditions ($M = 5.3$ vs. 5.2). However, the effect of speaking rate varied across groups. Somewhat higher ratings (indicating less foreign accent) were accorded sentences produced in the speeded than unspeeded condition by the two groups of late learners. However, the opposite held true for sentences produced by the two groups of early learners and the NE speakers.

The mean ratings obtained from the NE-speaking listeners were submitted to an ANOVA in which Group (five levels) served as a between-subjects factor and Speaking Condition (unspeeded, speeded) served as a repeated measure. This analysis yielded a significant main effect of Group, $F\ (4, 55) = 135.1$, $p < .001$, a nonsignificant main effect of Speaking Condition, $F\ (1, 55) = 3.3$, $p = .078$, and a significant two-way interaction, $F\ (4, 55) = 4.1$, $p = .006$. The simple effect of Group was significant in both conditions: unspeeded $F\ (4, 55) = 76.5$, $p < .0001$; speeded $F\ (4, 55) = 57.5$, $p < .0001$. Tukey tests revealed that, for both conditions, all pairwise differences between the five groups were significant ($\alpha = .05$) except those between the NE and early-low groups, and between the late-low and late-high groups.

The results of the post hoc tests indicated that, of the four groups of NI participants, only the early learners who seldom used Italian evaded detection as nonnative. That being the case, the hypothesis that speeding up will cause nonnative speakers to sound less foreign accented predicts significantly higher ratings for sentences produced in the speeded than unspeeded condition by the early-high, late-low, and late-high groups. This prediction was not supported, however, inasmuch as the simple effect of speaking condition was not significant for any of the five groups at a Bonferroni corrected .05 α level. The two-way interaction seems to have arisen, therefore, from the opposite direction of the speaking rate effect for the late learners compared to the early learners and NE speakers.

The purpose of the next analysis was to determine if the NA-speaking listeners would show the predicted effect of a speedup in speaking rate. The ratings accorded the 120 stimuli by the NE-speaking and NA-speaking listeners were highly correlated ($r = .94$, $p < .01$). Moreover, as shown in Figure 3, ratings obtained for the 60 pairs of sentences drawn from the unspeeded and speeded conditions were highly correlated for both groups of listeners, NE, $r\ (59) = .979$, $p < .0001$; NA, $r\ (59) = .968$, $p < .0001$.

The ratings obtained from the NA-speaking listeners were submitted to a two-way ANOVA, which yielded a significant main effect of group, $F\ (4, 55) = 111.8$, $p < .0001$, a nonsignificant effect of speaking condition, $F\ (1, 55) = .1$, $p > .10$, and a nonsignificant two-way interaction, $F\ (4, 55) = 2.5$, $p = .056$. A Tukey test examining ratings averaged over the two conditions revealed that the pattern of

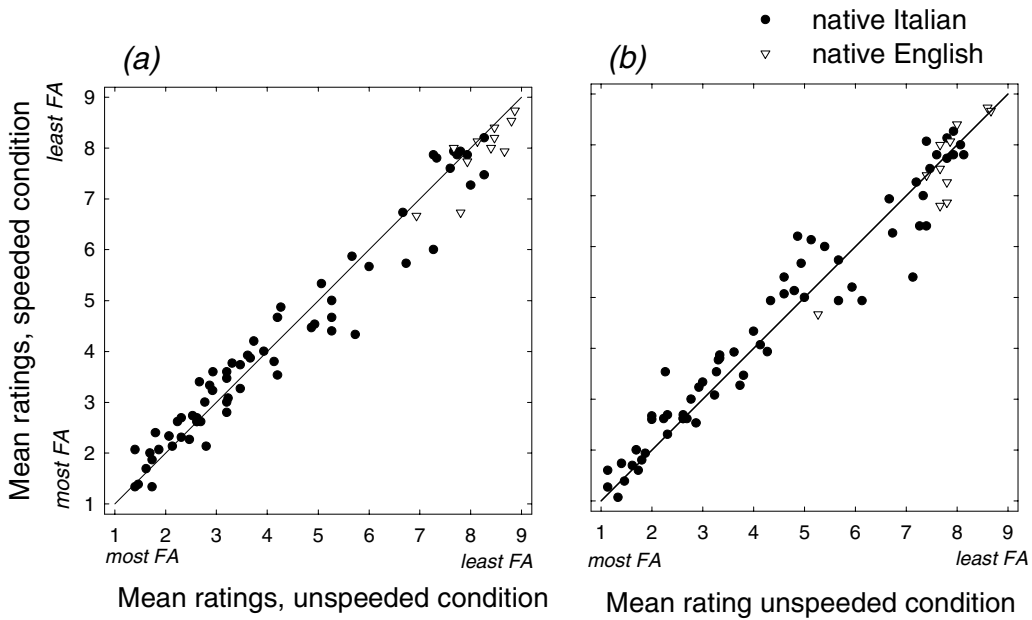

Figure 3. The mean foreign accent ratings obtained for pairs of sentences differing by about 10% as spoken by native speakers of English ($n = 11$) and Italian ($n = 49$). The sentences were rated by groups of listeners whose native language was (a) English or (b) Arabic.

significant pairwise between-group differences was the same for the NA listeners as for the NE listeners. That is, all between-group differences were significant except those between the NE and early-low groups and the late-low and late-high groups.

Discussion

We hypothesized that a speedup in speaking rate would cause foreign-accented nonnative speakers' to sound less accented when speaking their L2. This hypothesis was not supported by an ANOVA, however. Moreover, a very strong correlation was observed to exist between the 60 pairs of sentences that differed, on average, by 10% in duration. This held true for ratings obtained from both NE-speaking and NA-speaking listeners.

The results obtained here appear to diverge from the results obtained by Munro and Derwing (2001). When these authors shortened Chinese later learners' English sentences by 10% using a computer editing technique, the sentences were judged to be less foreign accented. Perhaps the 10% speed up in this experiment reduced foreign accent as hypothesized, but the ameliorative effect of a speeding up was offset by the presence of more (or greater) phonetic errors in the speeded condition. Alternatively, an effect of speeding up might be evident only for L2 learners who have been using their L2 for shorter periods of time than those examined here ($M = 31$ years for the NI late learners). Whatever the explanation for the apparent divergence of results, the present results suggest that the effect of AOA on degree of L2 foreign accent can not be attributed to AOA-related differences in the duration of L2 sentences.

The purpose of including a group of non-NE-speaking listeners was to determine if they might show a greater effect of the 10% sentence-duration difference than the NE-speaking listeners had. This was not the case, however. The NE- and NA-speaking listeners' ratings were highly correlated, and the NA listeners showed the same pattern of between-group differences as the NE-speaking listeners. The results obtained here for the NA-speaking listeners agreed with the results of Flege (1988) in showing that nonnative speakers of English are able to gauge foreign accent in English sentences much like NE-speaking listeners. However, whereas the nonnative listeners in the Flege (1988) study were asked to rate sentences produced by nonnative speakers who shared their L1 (Chinese), the NA listeners in this experiment evaluated nonnatives who spoke a different L1 (Italian).

As discussed in the introductory section, foreign accent arises in large part from cross-language phonetic interference. This is why, for example, German- and French-accented English sound so different. The NA-speaking listeners could probably not have resembled the NE-speaking listeners so closely simply by gauging how much the Italians' sentences matched their own pronunciation of English. This is because the NI participants were likely to produce the English test sentences with a different kind of L1-inspired foreign accent than the NA-speaking listeners would. The present findings imply, therefore, that the NA-speaking listeners were developing long-term memory representations for English phonetic segments and prosodic dimensions, and were able to detect divergences from those developing representations. This conclusion agrees with the inference drawn from previous research examining late learners' perception of L2 phonetic

segments (e.g., Flege & MacKay, 2004) and recognition of L2 words (e.g., Imai, Walley, & Flege, 2005).

The findings obtained for the NA-speaking listeners have one other important implication. The NA-speaking listeners' own pronunciation of English was not assessed. However, informal observation suggested that all of them spoke English with readily detectable foreign accents. That being the case, the present findings support the view (Flege, 1995, 1999, 2002, 2003) that the ability to perceive phonetic characteristics of an L2 may develop more rapidly or fully than the ability to pronounce the L2.

GENERAL DISCUSSION

The aim of this study was to provide insight into the basis of age-related effects on L2 speech acquisition. Previous research has examined the overall degree of foreign accent in the L2 speech of individuals who differ in age of first exposure to the L2. Age of exposure has often been indexed by AOA in a predominantly L2-speaking environment. Many studies have yielded a significant correlation between degree of accent and AOA: the later the arrival, the stronger foreign accents tend to be. Although robust, the effect of AOA has been difficult to interpret, however. This is because AOA is typically confounded with other potentially important variables. This study evaluated the effect on L2 foreign accent of two such confounded variables: age at the time of test, and the duration of L2 sentences. The question of interest was whether late learners' relatively greater age and longer L2 utterances contribute to their being perceived as more strongly foreign accented than early learners.

The results suggested that the AOA effects obtained in previous research (e.g., Flege et al., 1995a) were not due to the two confounded variables examined here. The results of Experiment 1 suggested that chronological age had little effect on the foreign accent ratings obtained for NI speakers of English in Canada. Experiment 2 showed that a 10% difference in the duration of English sentences, brought about by asking participants to speak rapidly, had little effect on degree of perceived foreign accent.

Although this study confirmed previous research in showing that AOA exerts a robust effect on degree of L2 foreign accent, the present findings do not explain the basis of AOA effects. As summarized in the introductory section, AOA might index L2 learners' state of neurological development at the time they are first exposed to an L2. With this view, AOA effects provide evidence that the capacity for L2 learning declines as the brain matures. Differences in the performance of early and late learners who are highly experienced in an L2 might also be taken as evidence that late but not early learners have passed a critical period for L2 learning (DeKeyser, 2000; Scovel, 2000).

One problem with a maturational constraint (or "critical period") explanation of AOA effects is evidence that foreign accents appear in the speech of individuals exposed to their L2 before the supposed end of a critical period at the age of 12 or 15 years (e.g., results obtained in this study as well as by Flege et al., 1995a, and Piske et al., 2001). Another problem is evidence that some late learners manage to speak their L2 without detectable foreign accents (e.g., Bongaerts, Mennen, & van der Slik, 2000).

Three alternative explanations for AOA effects have been offered in the literature. First, Hakuta, Bialystok, and Wiley (2003; see also Bialystok & Hakuta, 1999) hypothesized that L2 learning becomes less successful as the age of first exposure to an L2 increases because cognitive processes that are relevant to L2 learning deteriorate with increasing age. According to Hakuta et al. (2003), the impact of normal cognitive aging is "played out over a protracted period," yielding a slow decline in ultimate L2 proficiency across the life span, not just a dramatic decline near puberty. This account correctly predicts the gradual increase in degree of foreign accent beyond an AOA of 20 years (Flege, MacKay, & Imai, 2005; see also Figure 1). However, it does not predict the presence of foreign accent in some individuals whose cognitive systems have not yet begun to show age-related declines, that is, early learners who continue to use their L1 frequently.

Second, AOA effects on L2 speech acquisition might be attributed to differences in language use and input (see, e.g., Flege et al., 1997; Flege & Liu, 2001). A late AOA is typically associated with a relatively frequent continued use of the L1 and corresponding infrequent use of the L2 (Flege, 1999). The results of previous research (e.g., Flege et al., 1997; Guion et al., 2000; Piske et al., 2001) as well as the results of Experiment 2 of this study have shown that language use exerts an effect on L2 foreign accent that is independent of AOA. In addition, late-arriving immigrants tend to congregate with compatriots from their home country to a greater extent than early-arriving immigrants (Jia & Aaronson, 1999). As a result, late learners may hear their L2 spoken with an L1 foreign accent more often than early learners do. A problem with the input/use account is that measures of L2 input and language use remain imprecise, making it impossible to accurately assess the amount of variance in L2 performance that such measures predict.

Third, an account is provided by the SLM (Flege, 1995, 1999, 2002, 2003) and other theoretical approaches that specify a negative impact of L1 category development on L2 phonetic learning (e.g., Iverson et al., 2003; McCandliss et al., 2002). According to the SLM, AOA effects are primarily due to variation in the state of development of the L1 phonetic system at the time L2 learning begins. On this view, the L1 phonetic subsystem will exert a stronger influence on L2 pronunciation if it is fully developed when L2 learning begins than if it is still developing.

The SLM hypothesizes that as the phonetic categories used to produce and perceive L1 vowels and consonants develop through childhood and into adolescence (see, e.g., Hazan & Barrett, 2000; Johnson, 2000; Lee, Potamianos, & Narayanan, 1999), they are more likely to subsume L2 phonetic categories (i.e., be used to produce and perceive L2 speech sounds), and thus to block the formation of new phonetic categories for L2 sounds. Several problems exist with this account. The end point of L1 category development has not yet been established. Foreign accents continue to get stronger beyond the point when L1 categories could reasonably be expected to continue developing (Flege et al., 2005). Finally, and most importantly, no study has as yet demonstrated that groups of individuals differing in the state L1 category development exhibit differential success in learning the sound system of an L2.

Additional research will be needed to choose between these competing accounts of AOA effects on L2 speech acquisition. It seems likely that an optimal account will eventually incorporate several factors. It is worth noting, however, that the

results of Experiment 1 of this study provided indirect support for the SLM account.

Ten variables drawn from a language background questionnaire administered to the NI participants were submitted to a principle components analysis. When three factors derived from this analysis were regressed onto the foreign accent ratings, a factor with high loadings on four variables (AOA in Canada, Italian proficiency, bilingual dominance, years of education in Italy) accounted for 45% of the variance in foreign accent ratings. This factor was designated the "AOA" factor, but it might be reasonably designated the "L1 development" factor. This is because the NI participants who arrived in Canada relatively late in life tended to have attended school in Italy longer than those who arrived earlier in life. The late-arriving participants judged themselves to be more proficient in Italian than the early-arriving participants, so their L1 phonetic systems may well have been more developed than the early-arriving participants' at the time of immigration to Canada. The more frequent Italian use by the late-arriving than by the early-arriving participants probably ensured that they maintained their already greater L1 proficiency better than the early-arriving participants. This may have engendered a stronger influence of the L1 phonetic system on L2 pronunciation for the late than the early learners.

In summary, this study examined degree of foreign accent in English sentences produced by Italian immigrants who were long-time residents of Canada. As in previous research, the NI participants who arrived in Canada relatively late in life had stronger foreign accents than those who arrived earlier in life. Early learners managed to evade foreign accent detection, but only if they used their L1 infrequently. Two experiments showed that foreign accent differences between the early and late learners were not due to the late learners' tendency to produce longer sentences, or to their greater age at test. The results of a principle components analysis were interpreted as providing support for the view that AOA exerts a strong effect on degree of L2 foreign accent because it provides an indirect index of the strength of the L1 phonetic system when L2 learning begins.

APPENDIX A

The English test sentences and answers examined in Experiment 1

Questions	Answers
1. What's that woman's name?	She's called Patty.
2. Do you know her well?	We were neighbors for 30 years.
3. Who's she married to?	Billy is Patty's husband.
4. Does she have a sister?	Barbara is Patty's sister.
5. Who's taller?	Barbara is a little taller than Patty.
6. Who did you invite for dinner?	Barbara and Patty will be coming.
7. What do you plan to serve?	We'll have pork chops and potatoes.
8. Anything else?	We'll have mozzarella cheese.

Note: Translation equivalent Italian sentences (not shown) were used to elicit the production of Italian sentences by native Italian-speaking participants.

ACKNOWLEDGMENTS

This research was supported by Grant DC00257 from the National Institute for Deafness and Other Communicative Disorders. The authors thank Anders Højen and Murray Munro for comments on a previous version.

NOTES

1. The variable examined by Munro and Derwing (2001) was speaking rate, expressed in syllables per second. When sentences are produced without pauses, prolongations, or part/whole syllable repetitions, as was the case for the sentences examined in the present study, speaking rate is functionally equivalent to sentence duration.
2. Years of education in English-medium schools in Canada was strongly correlated with AOA because the NI participants who arrived in Canada at school age were soon enrolled in school. As a result, the pattern of correlations just summarized for AOA also existed for years of Canadian education.
3. Italian proficiency was the average self-rated ability to speak and understand Italian. English proficiency was not included in the analysis because it was not of interest to assess the NI participants' ability to self-assess their pronunciation of English. Years of formal education in English-medium schools in Canada was also excluded because a strong correlation existed between this variable and AOA, and between these two variables and other participant variables. AOA seemed likely to be a more important predictor of foreign accent than years of education. Admittedly, however, years of education might be related to the amount of native-speaker input and motivation to pronounce the L2 well which, in turn, might influence L2 pronunciation.
4. For NI participants in this study, percentage use of Italian in the home and in social settings were correlated with overall percentage use of Italian, r (136) = .64 and .56, $p < .01$, as well as AOA, r (136) = .37 and .28, $p < .01$. Home use of Italian was correlated with chronological age, r (136) = .27, $p < .01$, but not social use of Italian, r (136) = .10, $p > .10$.
5. The principle components analysis was based on a correlation matrix containing all pairwise correlation coefficients among the 10 variables, with unity retained along the diagonal. This approach was used because the 10 variables differed in terms of unit of measurement (e.g., self-reported percentage use of English, 7-point equal appearing interval scales). A varimax rotation retained the orthogonality of the derived factors while optimally separating the variables that were related to one another from those that were not (that is, deriving a simple structure).
6. Unexpectedly, the age-defined subgroups also differed significantly in years of education in Italy, F (5, 102) = 4.6, $p < .01$. A Tukey test revealed that participants in the Age46 group had received more education in Italy (M = 9 years) than those in the Age57, Age60, and Age62 groups (M = 5–6 years). An inspection of individual data revealed that four participants in both the Age54 and Age46 groups had completed 12–13 years of formal education in Italy prior to immigrating to Canada, whereas just one participant in the remaining four groups had done so. Most participants in Age54 and Age46 arrived in Canada after 1965, whereas most participants in the remaining groups had arrived prior to 1965. The between-group difference in years of education in Italy might therefore be due to a shift over time in the average education level of Italian immigrants to Canada.

7. The sentences were *The woman was not in critical condition*, *La donna gli fece una critica severa, Un' auto è meno economica del tram*, and *My Chevrolet is more economical than his*. The English sentence models were produced by an adult female native speaker of English, and Italian sentence models were produced by an adult female native speaker of Italian.

REFERENCES

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*, 529–555.

Bahrick, H., Hall, L., Goggin, J., Bahrick, L., & Berger, S. (1994). Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants. *Journal of Experimental Psychology: General, 123*, 264–283.

Baker, W., Trofimovich, P., Mack, M., & Flege, J. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In B. Skarabela, S. Fish, & A. Do (Eds.), *Proceedings of the 26th annual Boston University Conference on Language Development* (pp. 36–47). Somerville, MA: Cascadilla Press.

Bates, E. (1999). Plasticity, localization and language development. In S. Broman & J. M. Fletcher (Eds.), *The changing nervous system: Neurobehavioral consequences of early brain disorders* (pp. 214–253). New York: Oxford University Press.

Bialystok, E., Craik, F., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*, 290–303.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Bongaerts, T., Mennen, S., & van der Slik, F. (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.

Byrd, D. (1994). Relations of sex and dialect to reduction. *Speech Communication, 15*, 39–54.

Cunningham-Anderson, U., & Engstrand, O. (1989). Perceived strength of foreign accent and identity of foreign accent in Swedish. *Phonetica, 46*, 138–154.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499–534.

Flege, J. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8*, 162–177.

Flege, J. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*, 70–79.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.

Flege, J. (1998). The role of subject and phonetic variables in L2 speech acquisition. In M. Gruber, D. Higgins, K. Olsen, & T. Wysocki (Eds.), *Papers from the 34th annual meeting of the Chicago Linguistic Society: Vol. II. The Panels* (pp. 213–232). Chicago: Chicago Linguistic Society.

Flege, J. (1999). Age of learning and second-language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–132). Hillsdale, NJ: Erlbaum.

Flege, J. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217–244). Trier, Germany: Wissenschaftlicher Verlag.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production; Differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.

Flege, J., & Fletcher, K. (1992). Talker and listener effects on the perception of degree of foreign accent. *Journal of the Acoustical Society of America, 91*, 370–389.

Flege, J., Frieda, A., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169–186.

Flege, J., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23*, 527–552.

Flege, J., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26*, 1–34.

Flege, J., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Journal of Applied Psycholinguistics, 23*, 567–598.

Flege, J., MacKay, I. R. A., & Imai, S. (2005). *Accounting for age effects on overall degree of foreign accent in a second language*. Manuscript submitted for publication.

Flege, J., & Munro, M. (1994). The word unit in L2 speech production and perception. *Studies in Second Language Acquisition, 16*, 381–411.

Flege, J., Munro, M., & MacKay, I. R. A. (1995a). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125–3134.

Flege, J., Munro, M., & MacKay, I. R. A. (1995b). The effect of age of second language learning on the production of English consonants. *Speech Communication, 16*, 1–26.

Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language, 41,* 78–104.

Guion, S., Flege, J., Liu, S., & Yeni-Komshian, G. (2000). Age of learning affects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21*, 205–228.

Guion, S., Flege, J., & Loftin, J. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28*, 27–42.

Hakuta, K., & Bialystok, E. (1994). *In other words: The science and psychology of second-language acquisition*. New York: Basic Books.

Hakuta, K., Bialystok, E., & Wiley E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science, 14*, 31–38.

Hazan, V., & Barrett, S. (2000). The development of phonemic categorisation in children aged 6 to 12. *Journal of Phonetics, 28*, 377–396.

Imai, S., Walley, A., & Flege, J. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America, 117*, 896–907.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47–B57.

Jia, G., & Aaronson, D. (1999). Age differences in second language acquisition: The dominant language switch and maintenance hypothesis. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd annual Boston University conference on language development* (pp. 301–312). Somerville, MA: Cascadilla Press.

Johnson, C. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research, 43*, 129–143.

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

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455–1468.

MacKay, I. R. A., & Flege, J. (2004). Effects of the age of second-language learning on the duration of first-language and second-language sentences: The role of suppression. *Applied Psycholinguistics, 25*, 373–396.

Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics, 26*, 381–400.

McCandliss, B., Fiez, J., Protopapas, A., Conway, M., & McClelland, J. (2002). Success and failure in teaching the [r]–[l] contrast to Japanese adults: Predictions of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective and Behavioral Neuroscience, 2*, 89–108.

Metz, D., Samar, V., Schiavetti, N., Sitler, R., & Whitehead, R. (1985). Acoustic dimensions of hearing-impaired speakers' intelligibility. *Journal of Speech and Hearing Research, 28*, 345–355.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*, 81–108.

Munro, M. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition, 17*, 17–34.
Munro, M. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition, 20*, 139–154.
Munro, M., & Derwing, T. (1998). The effects of speaking rate on listener evaluation of native and foreign-accented speech. *Language Learning, 48*, 159–182.
Munro, M., & Derwing, T. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. *Studies in Second Language Acquisition, 23*, 451–468.
Munro, M., Flege, J., & MacKay, I. R. A. (1996). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics, 17*, 313–334.
Patkowski, M. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics, 11*, 73–89.
Pennington, M. (1992). Discourse factors related to L2 phonological proficiency: An exploratory study. In J. Leather & A. James (Eds.), *Proceedings of the 1992 Amsterdam symposium on the acquisition of second-language speech* (pp. 137–155). Amsterdam: University of Amsterdam, Department of English.
Piske, T., MacKay, I. R. A., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191–215.
Raupach, M. (1980). Temporal variables in first and second language speech production, In H. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 263–270). The Hague: Mouton.
Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes, 14*, 423–441.
Schairer, K. (1992). Native speaker reaction to non-native speech. *Modern Language Journal, 76*, 309–319.
Scovel, T. (1988). *A time to speak. A psycholinguistic inquiry into the critical period for human speech*. Cambridge: Newbury House Publishers.
Scovel, T. (2000). A critical review of critical period research. *Annual Review of Applied Linguistics, 20*, 213–223.
Southwood, H., & Flege, J. (1999). The validity and reliability of scaling foreign accent. *Clinical Linguistics and Phonetics, 13*, 335–349.
Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech, 24*, 265–272.
Willems, N. (1982). *English intonation from a Dutch point of view*. Dordrecht: Foris.