Applied Psycholinguistics **25** (2004), 373–396
Printed in the United States of America
DOI: 10.1017.S0142716404001171

Effects of the age of second language learning on the duration of first and second language sentences: The role of suppression

IAN R. A. MACKAY
University of Ottawa

JAMES E. FLEGE
University of Alabama at Birmingham

ADDRESS FOR CORRESPONDENCE
J. E. Flege, Division of Speech & Hearing Sciences, University of Alabama at Birmingham, CH20, Room 119, 1530 3rd Avenue South, Birmingham, AL, 35294-2042. E-mail: jeflege@uab.edu

ABSTRACT
The primary aim of this study was to account for the finding that late bilinguals produce longer English sentences than early bilinguals. In Experiment 1, Italians who immigrated to Canada either between the age of 2–13 years ("early bilinguals") or 15–28 years ("late bilinguals") repeated matched English and Italian sentences following an aural model. The early bilinguals produced shorter English than Italian sentences, whereas the late bilinguals showed the opposite pattern. The same countervailing pattern was evident in Experiment 2, where bilinguals shortened sentences by 20% when instructed to repeat sentences as rapidly as possible. Subgroups of bilinguals who reported using Italian often ($M = 46\%$ Italian use) but not seldom ($M = 8\%$) were found to have produced significantly longer English sentences than native English (NE) speakers did. The results were interpreted to mean that the late bilinguals produced longer English sentences than the early bilinguals because they needed to expend more resources to suppress their Italian subsystem than the early bilinguals. The perceptual effect of sentence duration was evaluated in Experiment 3, where pairs of English sentences differing in duration were presented to NE-speaking listeners for foreign accent ratings. A 10% shortening caused sentences spoken by late bilinguals to sound less foreign accented but it caused sentences spoken by early bilinguals to sound more foreign accented.

It has been known for some time (e.g., Riggenbach, 1991) that nonnative speakers tend to produce sentences in their second language (L2) with longer durations (indicating a slower speaking rate) than native speakers. Guion, Flege, Liu, and Yeni–Komshian (2000) recently found that the duration of L2 sentences is correlated positively with age of arrival (AOA) in a predominantly L2-speaking community. The primary aim of this study was to test competing hypotheses regarding the basis of AOA effects on L2 sentence duration. A secondary aim of the study was to evaluate the role of AOA-related differences in sentence duration on overall degree of perceived foreign accent.

Three possible explanations might be offered to account for the effect of AOA on the duration of L2 sentences. According to the "practice effect" hypothesis, AOA effects are an indirect consequence of differences in how often early and late bilinguals have previously produced the words making up L2 sentences. Words with low text frequencies tend to be produced with longer durations than words with high text frequency (Geffen & Luszcz, 1983; Geffen, Stierman, & Tildesley, 1979; Wright, 1979). Late bilinguals rate L2 words as being less familiar than early bilinguals (Flege, Frieda, Walley, & Randazza, 1998; Flege, Takagi, & Mann, 1996). Text frequency and nonnative speakers' subjective familiarity estimates are correlated (Flege et al., 1998). One might hypothesize, therefore, that late bilinguals produce L2 sentences with relatively longer durations than early bilinguals because they have less prior experience producing L2 words than early bilinguals.

According to the "speaking rate" hypothesis, AOA effects on the duration of L2 sentences are due to differences in volitionally chosen speaking rates (Anderson–Hsieh & Koehler, 1988; see also Derwing & Munro, 1997). Nonnatives' speech is often less intelligible than that of native speakers (e.g., Munro, 1998; van Wijngaarden, Steeneken, & Houtgast, 2001). Individuals who learn an L2 in late adolescence or adulthood ("late bilinguals") generally produce L2 vowels and consonants less accurately than "early bilinguals" who began learning their L2 as children (e.g., Flege, MacKay, & Meador, 1999; Flege, Munro, & MacKay, 1995b; Piske, Flege, MacKay, & Meador, 2002). This contributes to the lower intelligibility of L2 speech produced by late than early bilinguals (e.g., van Wijngaarden, Steeneken, & Houtgast, 2002). Most speakers can augment intelligibility by speaking more slowly (see, e.g., Bradlow & Bent, 2002). It is therefore possible that late bilinguals tend to speak their L2 more slowly than early bilinguals because they feel a greater need than early bilinguals to try to augment their intelligibility.

Third, according to the "suppression" hypothesis, AOA effects on the duration of L2 sentences arise from differences in the resources bilinguals must expend to suppress one language while speaking their other language (Guion et al., 2000). Evidence exists that both the L1 and L2 subsystems of a bilingual always remain activated to some extent (e.g., Soares & Grosjean, 1984) and that a bilingual must expend resources to suppress one language while speaking the other (e.g., Green, 1986). Early bilinguals generally use their L1 less often than late bilinguals (see, e.g., Flege, Yeni–Komshian, & Liu, 1999) and so are more likely than late learners to become dominant in the L2 (Jia & Aaronson, 1999; Flege, MacKay, & Piske, 2002). It seems reasonable to think that bilinguals will need to expend more processing resources to suppress their dominant than nondominant language (Meuter & Allport, 1999) and that sentence duration will increase as a function of the resources needed to suppress the language not currently in use. If both of these assumptions were correct, then late bilinguals would be expected to produce longer L2 sentences than early bilinguals because they must expend more resources to suppress the L1 while speaking the L2 than early learners.

Experiments 1 and 2 of the present study were designed to help chose between these competing accounts of AOA effects on L2 sentence duration. The participants were native speakers of Italian who had emigrated from Italy to Canada at the age

of 2–13 years (early bilinguals) or 15–28 years (late bilinguals). Both experiments examined the production of matched pairs of English and Italian sentences. In Experiment 1, 12 English and 12 Italian sentences were presented for repetition a single time via a loudspeaker. No instruction as to speaking rate was provided. In Experiment 2, the bilinguals repeated just two English–Italian pairs six times each. In the first, "unspeeded" condition, the Italian–English bilinguals produced the sentences without instruction as to speaking rate; in the second, "speeded" condition, they were instructed to produce the sentences as rapidly as possible.

If late bilinguals' relative lack of familiarity with English words were responsible for their production of relatively long English sentences (the practice effect hypothesis), then a difference between early and late bilinguals should be observed in Experiment 1 (where 12 English sentences were repeated a single time) but not in the unspeeded condition of Experiment 2 (where two sentences were repeated six times each). If a difference in volitionally selected speaking rates were responsible for early–late differences (the speaking rate hypothesis), then a difference between early and late bilinguals should be observed in the unspeeded condition of Experiment 2 (where sentences were repeated without instruction as to speaking rate) but not in the speeded condition of that experiment (where sentences were repeated as rapidly as possible).

The results obtained here provided some support for the speaking rate hypothesis. The method used to choose between the speaking rate and suppression hypotheses was to evaluate the relative duration of L1 and L2 sentences produced by the two groups of bilinguals. The suppression hypothesis predicts a countervailing pattern of English and Italian sentence duration values regardless of the number of repetitions of sentences or instructions as to speaking rate. Specifically, it predicts that early bilinguals will produce shorter English sentences than Italian sentences because they need to expend fewer resources to suppress their nondominant language (Italian) when speaking English than to suppress their dominant language (English) while speaking Italian. Conversely, late bilinguals are predicted to produce longer English than Italian sentences because of the need to expend more resources to suppress their dominant language (Italian) while speaking English than to suppress their nondominant language (English) while speaking Italian.

Experiment 3 evaluated the effect of AOA-related variation in sentence duration on overall degree of perceived foreign accent. Listeners' ratings of foreign accent in L2 sentences are determined in large part by nonnative speakers' L1-inspired divergences from L2 prosodic and segmental phonetic norms (e.g., Flege, 1984; Flege & Munro, 1994; Flege et al., 1995a; Flege, Munro, & MacKay, 1995b; Munro, 1995; Munro, Flege, & MacKay, 1996). Correlations have been reported between degree of foreign accent and the number of segmental substitutions in sentences (e.g., Anderson–Hsieh & Koehler, 1992; Brennan & Brennan, 1981; Cunningham–Andersson & Engstrand, 1989; Magen, 1998; Major, 1986; Schairer, 1992). However, phonetic differences that do not result in the perceived substitution of one L2 segment for another (e.g., inappropriately short voice-onset time values) are also known to contribute to the perception of foreign accent (see, e.g., Flege & Munro, 1994; Jonasson & McAllister, 1972).

The frequency of L1-inspired divergences from L2 phonetic norms is related in an important way to AOA in a predominantly L2-speaking community (e.g., Flege

et al., 1995b). Degree of perceived foreign accent is correlated with AOA (Flege et al., 1995a, 1999). Guion et al. (2000) found that the duration of L2 sentences is also correlated with AOA. Taken together, these findings suggest that the duration of L2 sentences might serve as a perceptual cue to foreign accent.

Support for this inference was provided by Munro and Derwing (2001), who found that the duration of English sentences produced by nonnative adults accounted for 15% of the variance in foreign accent ratings. In a second experiment, Munro and Derwing (2001) used computer compression–expansion software to shorten and lengthen English sentences produced by nonnative speakers. Native English (NE)-speaking listeners judged sentences shortened by 10% to sound less foreign-accented than the original sentences or sentences lengthened by 10%. The number of segmental errors and sentence duration values accounted for 39% of the variance in the foreign accent ratings. The sentence duration values alone accounted for 6% of the variance.[1]

The two experiments reported by Munro and Derwing (2001) provided converging evidence that relatively long L2 sentence duration values cue foreign accent. However, the technique used in that study to alter sentence duration might be questioned (Picheny, Durlach, & Braida, 1989). The durations of all phonetic segments were shortened by a constant proportion, even though natural variation in speaking rate results in nonuniform changes in the duration of phonetic segments. The procedure adopted in Experiment 3 to vary sentence duration, therefore, was to select pairs of English sentences that had been produced by the same individual at different speaking rates. The question of interest was whether the shorter members of the pairs would be judged to sound less foreign accented than the longer members of the pairs.

EXPERIMENT 1

The aim of this experiment was to measure the duration of matched English and Italian sentences that were repeated following an aural model by early and late Italian–English bilinguals. The sentence duration values examined here were used by Flege et al. (2002) to calculate sentence duration ratios, which were used as an index of bilingual dominance. The duration values were submitted here to an analysis of variance (ANOVA) examining the effect of AOA (early vs. late) and language (Italian vs. English).

Method

Participants. Seventy-two native speakers of Italian were recruited in Ottawa, Ontario. Half of the participants had arrived in Canada from Italy between the ages of 2 and 13 years ($M = 8$ years) and half had arrived between the ages of 15 and 26 years ($M = 20$ years). All but three bilinguals had lived in Canada for more than 20 years at the time of testing. All of the participants passed a hearing screening and responded to a language background questionnaire that included questions pertaining to percentage use of Italian and English and years of education in Italy and Canada.

Table 1. *Mean (standard deviation) and range values of the characteristics of the participants in Experiment 1*

	N (M/F)	Age (years)	AOA (years)	LOR (years)	Use of IT (%)	Educ. in CA	Educ. in IT	EN Ability	IT Ability
Early	16/20	49 (5) 35–61	8 (4) 2–13	41 (4) 33–50	25 (21) 1–80	13 (5) 2–24	2 (2) 0–7	6.6 (0.6) 5–7	5.1(1.0) 2–7
Late	18/18	50 (8) 29–62	20 (3) 15–26	30 (8) 4–42	32 (25) 2–75	2 (2) 0–8	9 (4) 2–18	5.6 (0.7) 4–7	6.6 (0.6) 5–7

Note: Age, chronological age; AOA and LOR, age of arrival and length of residence in Canada, respectively; Use of IT, self-reported percentage use of Italian; Educ. in CA and IT, years of education in Canada and Italy, respectively; EN and IT ability, self-rated ability to speak and understand English and Italian, respectively; $1 = poor$, $7 = good$.

Mean characteristics of the early and late bilinguals are summarized in Table 1. The early and late bilinguals did not differ significantly in age, $F\ (1, 70) = .5$, $p > .10$, or percentage Italian use, $F\ (1, 70) = 1.5$, $p > .10$. However, the early bilinguals had lived longer in Canada than the late bilinguals, $F\ (1, 70) = 48.0$, $p < .05$, and had received more years of education in English-medium schools in Canada, $F\ (1, 70) = 164.4$, $p < .05$. Conversely, the late bilinguals had received more education in Italy than the early bilinguals, $F\ (1, 70) = 90.7$, $p < .05$.

The participants were asked to use 7-point scales ($1 = poor$ to $7 = good$) to rate their ability to speak and understand Italian and English (four items in all). The late bilinguals judged themselves to be less proficient in English than the late bilinguals, $F\ (1, 70) = 41.5$, $p < .05$, but more proficient in Italian, $F\ (1, 70) = 56.8$, $p < .05$. Most (31/36) early bilinguals gave themselves higher ratings in English than Italian, and so might be considered to be English dominant. Conversely, most (31/36) late bilinguals gave themselves higher ratings in Italian than English, and so might be considered Italian dominant (see Flege et al., 2002, for discussion).

Procedure. The participants repeated 12 English and 12 Italian sentences that had similar meanings and were matched for number of words and syllables. The sentences were repeated in a fixed order (English sentence 1 [E1], Italian sentence 1 [I1] I2, E2, E3, I3, I4, etc.), which ensured that English sentences immediately preceded an Italian sentence as often as Italian sentences immediately preceded an English sentence (see Flege et al., 2002, appendix A). The participants were provided with a written list of the sentences to be produced. Each sentence was presented via a loudspeaker twice in a row, followed by a pure tone. The participants were told to repeat the sentences "accurately" a single time after hearing the tone. No instruction was given as to speaking rate.

Measurement. The sentences were recorded using a Sony DAT recorder and a Shure (model SM10A) head-mounted directional microphone. The recordings were digitized at 22.05 kHz with 16-bit amplitude resolution and then normalized for peak intensity. A research assistant measured each sentence to the nearest

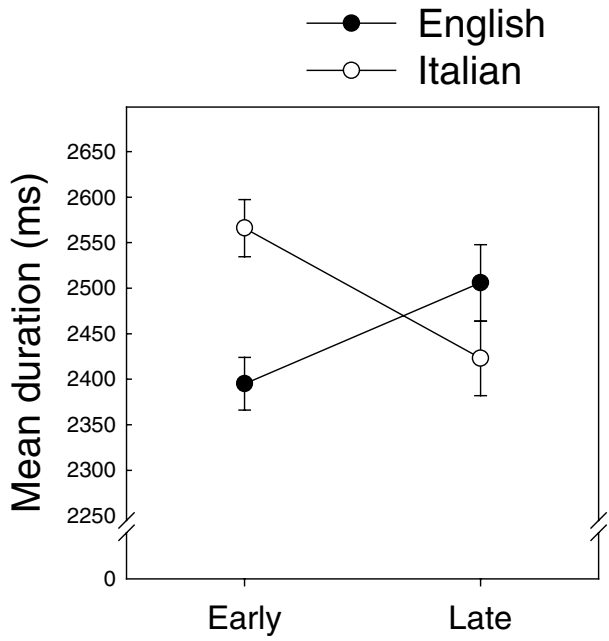

Figure 1. The mean duration of English and Italian sentences produced by Italian–English bilinguals in Experiment 1. The error bars enclose ±1 *SE*.

millisecond from digital spectrograms. Measurement reliability was evaluated by having the assistant remeasure 48 randomly selected sentences. The two sets of measures differed by an average of 20 ms (<1%). The unsigned difference between the two sets of measurements averaged 32 ms (1.3%). The second author's measurements of the 48 sentences differed from the assistant's by an average of 10 ms.

Only fluently produced sentences were measured. A total of 80 (14%) English and 53 (9%) Italian sentences were excluded for one or more of the following reasons: the presence of a part- or whole-word repetition, an extra (or omitted) word or syllable, laughter, the presence of an unfilled pause exceeding 200 ms, or a filled pause exceeding 150 ms. The missing values were replaced using a multiple imputation method (Little & Rubin, 1987). The mean durations of the English and Italian sentences produced by each participant were then calculated. To avoid list series effects, the mean values used in all analyses were based on only the middle eight pairs of sentences.

Results and discussion

As shown in Figure 1, early and late bilinguals showed a countervailing pattern of English and Italian sentence duration values. The ANOVA examining the mean

duration values yielded a significant AOA × Language interaction, $F\ (1, 70) = 54.1$, $p < .05$, which was explored through simple effects tests. The simple effect of language reached significance, but in opposite directions, for the two groups: the early bilinguals produced shorter English than Italian sentences, $F\ (1, 35) = 58.4$, $p < .05$, whereas the late bilinguals produced longer English than Italian sentences, $F\ (1, 70) = 10.1$, $p < .05$.

Simple effects tests also revealed that late bilinguals produced significantly longer English sentences than the early bilinguals, $F\ (1, 70) = 4.8$, $p < .05$. This finding is consistent with the results obtained by Guion et al. (2000) for Italian–English bilinguals living in Ottawa and Korean–English bilinguals living in the Washington, DC, area. The effect of AOA on the mean duration of Italian sentences was opposite to that observed for English sentences. That is, the late bilinguals produced significantly *shorter* Italian sentences than the early bilinguals, $F\ (1, 70) = 7.7$, $p < .05$. This is apparently the first study to have reported an effect of AOA on the duration of L1 sentences. A discussion of how these findings bear on the three hypotheses presented in the Introduction will be deferred until after presentation of the Experiment 2 results.

EXPERIMENT 2

Early and late Italian–English bilinguals again repeated English and Italian sentences following an aural model. However, this experiment differed from Experiment 1 in two crucial respects. First, 12 English and 12 Italian sentences were repeated a single time in Experiment 1 whereas in this experiment, two English and two Italian sentences were repeated six times each. The repetition of sentences was expected to eliminate between-group differences that depended on differences in word familiarity. Second, sentences were repeated in Experiment 1 without instruction as to speaking rate. This experiment included two conditions, one resembling Experiment 1 (the unspeeded condition) and the other in which sentences were repeated as rapidly as possible (the speeded condition). The speeded condition was expected to eliminate between-group differences that depended on differences in volitionally chosen speaking rates.

Method

Participants. Thirty–seven Italian-English bilinguals from Experiment 1 returned to participate in this experiment about 1 year after completing Experiment 1. They were joined by 27 additional Italian–English bilinguals and 16 age-matched NE controls. All 80 participants were tested in Ottawa, Ontario.

As summarized in Table 2, half of the bilinguals had an AOA ranging from 3 to 13 years ($M = 8$ years) and half had an AOA ranging from 15 to 26 years ($M = 19$ years). All but two bilinguals had lived in Canada for more than 20 years. The early and late bilinguals did not differ significantly in age, $F\ (1, 62) = .4$, $p > .10$, or Italian use, $F\ (1, 62) = 2.1$, $p > .10$. However, the early bilinguals had lived longer in Canada than the late bilinguals, $F\ (1, 62) = 42.8$, $p < .05$. The early bilinguals had received more years of education in Canada than the late bilinguals, $F\ (1, 62) = 199.4$, $p < .05$, but fewer years of education in Italy,

Table 2. *Mean (standard deviation) and range values of the characteristics of the bilingual participants in Experiment 2*

	N (M/F)	Age (years)	AOA (years)	LOR (years)	Use of IT (%)	Educ. in CA	Educ. in IT	En Ability	IT Ability
Early	15/17	49 (5) 36–62	8 (3) 3–13	41 (5) 28–51	23 (19) 2–70	12 (4) 2–20	2 (3) 0–10	6.4 (0.7) 5–7	5.1 (1.1) 3–7
Late	14/18	50 (7) 30–63	19 (3) 15–28	31 (7) 9–43	31 (24) 4–75	1 (2) 0–6	9 (3) 5–14	5.5 (0.6) 4–7	6.5 (0.7) 5–7

Note: Age, chronological age; AOA and LOR, age of arrival and length of residence in Canada, respectively; Use of IT, self-reported percentage use of Italian; Educ. in CA and IT, years of education in Canada and Italy, respectively; EN and IT ability, self-rated ability to speak and understand English and Italian, respectively; 1 = *poor*, 7 = *good*.

$F\ (1, 62) = 87.8$, $p < .05$. The late bilinguals judged themselves to be less proficient in English than the early bilinguals, $F\ (1, 62) = 35.5$, $p < .05$, but more proficient in Italian, $F\ (1, 62) = 37.0$, $p < .05$. Most (24/32) early bilinguals gave themselves higher ratings in English than Italian whereas most (28/32) late bilinguals gave themselves higher ratings in Italian than English.

Stimuli. Four sentences from Experiment 1 were examined: *The woman was not in critical condition*, *La donna gli fece una critica severa*, *Un' auto e' meno economica del tram*, and *My Chevrolet is more economical than his*. The selection of these two pairs from the 12 pairs available was based on the mean duration of sentences produced by the 6 (of 72) bilinguals in Experiment 1 who reported being equally proficient in Italian and English. We reasoned that these "balanced" bilinguals would be less likely to show an influence of Italian on their production of English sentences (or vice versa) than English- or Italian-dominant bilinguals. The average unsigned difference between the balanced bilinguals' production of the two selected English–Italian pairs was smaller on average ($M = 22$ ms) than the difference between the remaining 10 English–Italian pairs ($M = 252$ ms).

Procedure. The participants were given a written list of the four test sentences. Native speaker productions of these sentences were then presented via a loudspeaker. The test sentences were presented in the fixed order shown earlier a total of six times. As in Experiment 1, each sentence to be repeated was presented aurally twice in a row followed by a pure tone. The bilingual participants were told to repeat each sentence after hearing the tone. The NE speakers heard the Italian sentences but did not repeat them. No instruction was given as to speaking rate. The elicitation procedure was then repeated a second time with one change: the participants were now told to repeat the test sentences as rapidly as possible after hearing the tone.

Measurement. The 24 sentences produced by each bilingual participant (2 English + 2 Italian sentences × 6 repetitions) and the 12 sentences produced by

each NE speaker were digitized (22.05 kHz, 16-bit amplitude resolution) and then measured from digital spectrograms. The first phonetic segment of sentence 1 (the interdental fricative in *the*) and the last phonetic segment of sentence 3 (the nasal in *tram*) were sometimes low in intensity, creating the possibility that these segments might sometimes fail to be measured. Therefore, to augment measurement reliability, these two segments were excluded from measurement. Sentence 1 (S1) was measured from the onset of periodicity in /ə/ to the offset of frication noise in /ʃ/; S2 from the onset of periodicity in the /l/ to the offset of periodicity in /a/; S3 from the onset of periodicity in /u/ to the offset of periodicity in /a/; and S4 from the onset of periodicity in /m/ to the offset of the frication in /z/.

As in Experiment 1, only sentences produced fluently and correctly were measured. However, fewer sentences were excluded from this experiment than Experiment 1, apparently because just four sentences were repeated multiple times. Less than 1% of sentences produced by the NE speakers in either condition were excluded from analysis. Less than 5% of the bilinguals' unspeeded condition sentences were declared missing (4% early, 5% late) and just 2% of their speeded condition sentences were missing (2% early, 2% late). The bilinguals' missing values occurred with nearly equal frequency for each of the four test sentences.

A series of ANOVAs revealed that in none of the 16 possible instances (2 Groups × 4 Sentences × 2 Conditions) did the bilinguals' sentence duration values change significantly over the six repetitions of each sentence (Bonferroni adjusted $p > .10$). Given this finding and the relatively few tokens declared missing, the decision was made to calculate mean values for each participant based on however many nonmissing tokens were available for each sentence in the two conditions. There were 5–6 tokens available in 490 instances for sentences produced by bilinguals, 3–4 tokens available in 16 instances, and 2 tokens available in 5 instances. One early bilingual did not produce any usable tokens of sentence 1 in the speeded condition.

Results and discussion

Early versus late bilinguals. As shown in Figure 2, a countervailing pattern similar to the one obtained in Experiment 1 was obtained in both conditions of this experiment. The mean duration values obtained for the bilingual participants were submitted to an ANOVA in which AOA (early vs. late bilinguals) served as a between-subjects factor and language (English vs. Italian) and condition (unspeeded, speeded) served as within-subjects factors. Only one of the main effects reached significance, Condition, $F\,(1, 61) = 351.6$, $p < .05$, but not AOA, $F\,(1, 61) = .2$, $p > .10$, or Language, $F\,(1, 61) = .6$, $p > .10$. In addition, two of the two-way interactions reached significance, AOA × Language $F\,(1, 61) = 57.8$, $p < .05$, and Language × Condition, $F\,(1, 61) = 72.8$, $p < .05$, but not AOA × Condition, $F\,(1, 61) = .01$, $p > .10$. The ANOVA also yielded a significant three-way interaction, $F\,(1, 61) = 5.9$, $p < .05$.

The AOA × Language × Condition interaction was explored by first determining if the AOA × Language interaction reached significance in both conditions. This was indeed the case, unspeeded $F\,(1, 62) = 33.7$, $p < .05$; speeded $F\,(1, 61) = 56.8$, $p < .05$. Accordingly, the simple effects of AOA and Language

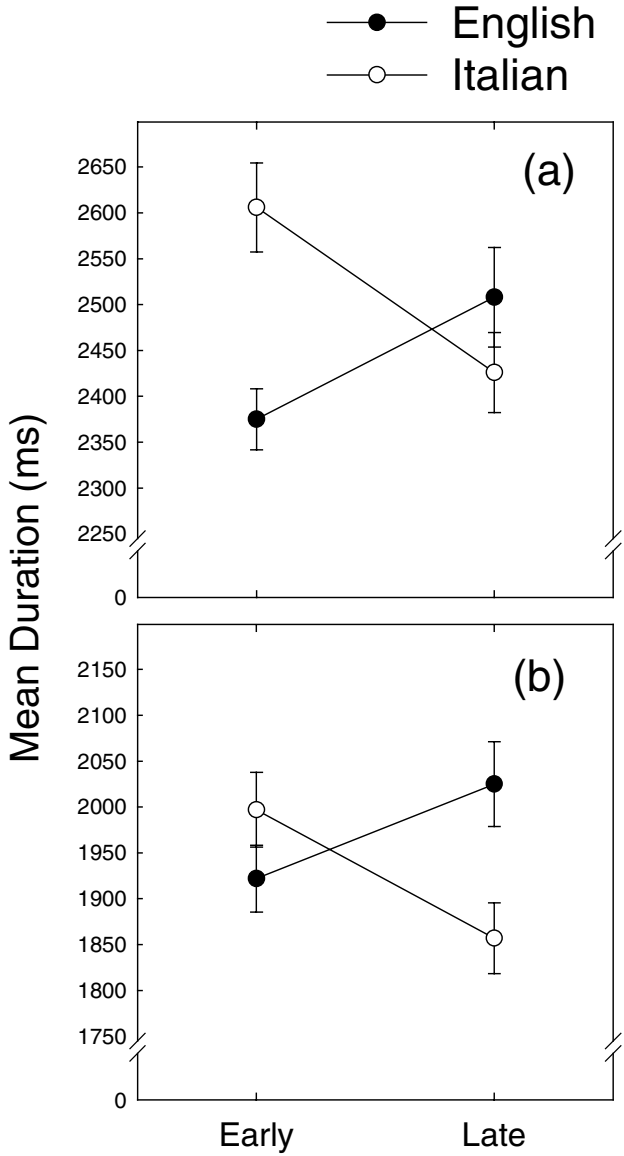

Figure 2. The mean duration of English and Italian sentences produced by Italian–English bilinguals in the (a) unspeeded and (b) speeded conditions of Experiment 2. The error bars enclose ±1 *SE*.

were tested separately for both conditions. These tests indicated that the three-way interaction arose because one of the eight simple effects was nonsignificant.

The simple effect of Language was significant in all four possible instances. The early bilinguals produced shorter English than Italian sentences in both conditions, unspeeded $F\ (1, 30) = 47.9$, $p < .05$; speeded $F\ (1, 30) = 9.2$, $p < .05$, whereas the late bilinguals produced longer English than Italian sentences in both conditions, unspeeded $F\ (1, 31) = 7.3$, $p < .05$; speeded $F\ (1, 31) = 65.5$, $p < .05$. The late bilinguals produced significantly shorter Italian sentences than the early bilinguals in both condition, unspeeded $F\ (1, 62) = 8.5$, $p < .05$; speeded $F\ (1, 62) = 7.2$, $p < .05$. In agreement with the results obtained in Experiment 1 and by Guion et al. (2000), the late bilinguals produced significantly longer English sentences than the late bilinguals in the unspeeded condition, $F\ (1, 61) = 4.3$, $p < .05$. However, the difference between the late and early bilinguals for speeded-condition sentences, although tending in the same direction as for the unspeeded condition, was nonsignificant, $F\ (1, 61) = 3.0$, $p = .09$.

These results bear on the three possible explanations of AOA effects on L2 sentence duration that were outlined in the Introduction. The practice effect hypothesis led to the expectation that late bilinguals would produce longer English sentences than the early bilinguals in Experiment 1 (where 12 English sentences were repeated a single time), but not in the unspeeded condition of this experiment (where two sentences were repeated six times each). However, significant early–late differences were observed in both instances, thereby failing to support the practice effect hypothesis. Another finding that was inconsistent with the practice effect hypothesis (see the Method section) was that the late bilinguals' English sentence duration values did not decrease systematically over the six repetitions of English sentences in either condition of this experiment.

The speaking rate hypothesis led to the expectation that the late bilinguals would produce significantly longer English sentences than the early bilinguals in the unspeeded condition of this experiment (where sentences were repeated without instruction as to speaking rate) but not in the speeded condition (where sentences were repeated as rapidly as possible). The predicted pattern was obtained, thereby providing support for the speaking rate hypothesis. It should be noted, however, that support for the speaking rate hypothesis was weak inasmuch as the difference in English sentences produced by the early and late bilinguals in the speeded condition narrowly missed reaching significance ($p = .09$).

Stronger support was provided for the suppression hypothesis. This hypothesis correctly predicted that the late bilinguals would produce significantly longer English than Italian sentences in Experiment 1 as well as in both conditions of this experiment. The suppression hypothesis also correctly predicted that the early Italian–English bilinguals would produce significantly shorter English than Italian sentences in all three instances.

An effect of percentage Italian use. The aim of the analyses presented here was to provide a further test of the suppression hypothesis. It did so by evaluating the influence of percentage Italian use on the duration of English sentences. We reasoned that Italian–English bilinguals who continued to use Italian frequently would need to expend more resources to suppress Italian while speaking English

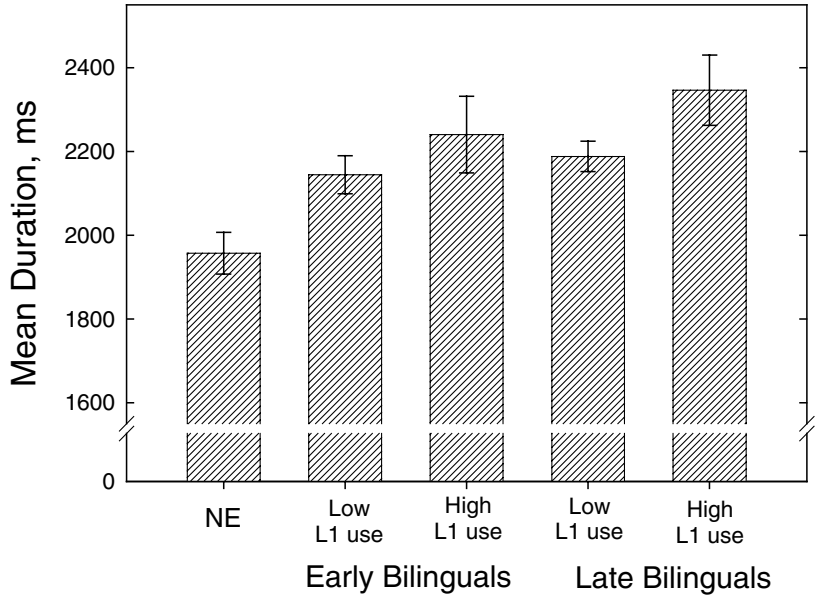

Figure 3. The mean duration of English sentences produced by native English speakers and Italian–English bilinguals differing in age of arrival in Canada (early, late) and amount of continued Italian use (low L1 use, high L1 use). The error bars bracket ±1 *SE*.

than bilinguals who used Italian seldomly. If so, then "high L1 use" bilinguals should produce longer English sentences than "low L1 use" bilinguals.

Half of the early bilinguals and half of the late bilinguals reported using Italian 1–15% of the time; the remaining bilinguals in each group reported using Italian 25–75% of the time. Bilingual participants were assigned to one of four groups based on AOA and percentage Italian use. These groups will be referred to as early-low (i.e., early bilinguals with a relatively infrequent use of Italian), early-high, late-low, and late-high. The mean English sentence duration values obtained for the four groups of bilinguals ($n = 16$ per group) and the NE speakers ($n = 16$) were then submitted to a 5 × 2 Group × Condition ANOVA.

The ANOVA revealed that the 483 ms (20.1%) difference between sentences produced in the unspeeded and speeded conditions was significant, $F\ (1, 74) = 382.1$, $p < .05$. The main effect of Group reached significance, $F\ (4, 74) = 6.9$, $p < .05$, but did not enter into a significant two-way interaction, $F\ (4, 74) = .4$, $p > .10$. The values obtained for each participant in the two conditions were therefore averaged and submitted to a Tukey test. Neither the two groups of early bilinguals differing in percentage Italian use (early-low, early-high) nor the two groups of late bilinguals (late-low, late-high) differed significantly ($p > .10$). This failed to support the predicted effect of L1 use on the duration of L2 sentences.

One other finding, however, was consistent with the predicted effect of L1 use. As shown in Figure 3, the early-high and late-high groups differed more

from the NE group ($M = 15\%$ and 20%) than the early-low and late-low groups ($M = 10\%$ and 12%). Both groups of high L1 use bilinguals were found to have produced significantly longer English sentences than the NE speakers ($p < .05$). However, neither group of low L1 use bilinguals differed significantly from the NE group. A significant correlation was found to exist between the average duration of English sentences and percentage Italian use, $r\ (62) = .34$, $p < .01$. This correlation remained significant when the effect of the following variables were partialed out: AOA, $r\ (61) = .32$; chronological age, $r\ (61) = .37$; years of residence in Canada, $r\ (61) = .34$; years of education $r\ (61) = .26$; self-rated ability to speak and understand Italian, $r\ (61) = .31$.[2]

EXPERIMENT 3

This experiment evaluated the effect of L2 sentence duration on nonnative speakers' overall degree of perceived foreign accent. As mentioned in the Introduction, Munro and Derwing (2001) found that decreasing the duration of English sentences spoken by late Chinese–English bilinguals caused their sentences to sound less foreign accented. The first aim of this experiment was to attempt to replicate this finding for late bilinguals and extend it to early bilinguals.

The computer expansion–compression software that Munro and Derwing (2001) used to vary sentence duration (and thus speaking rate) altered the duration of all phonetic segments within test sentences by a constant proportion. However, when speaking rate is varied naturally, the duration of phonetic segments change to varying extents. For example, Miller, Green, and Reeves (1986) found that voice onset time (an acoustic phonetic interval that affects both segmental identity and foreign accent ratings) shortens more when speaking rate is increased for /p t k/ than /b d g/. Acoustic analysis has revealed that when talkers increase their speaking rate, vowels typically shorten more than consonants (Gaitenby, 1965; see also Guion et al., 2000, for English sentences spoken by nonnative speakers). The technique used here to manipulate sentence duration, therefore, was to select pairs of sentences produced by the same individual at different speaking rates.

A literature review by Piske, MacKay, and Flege (2001) suggested that the two most important predictors of foreign accent are nonnative speakers' age of first exposure to their L2 (often indexed by AOA) and percentage L1 use. Munro and Derwing (2001) noted the need for caution in interpreting foreign accent ratings inasmuch as sentence duration may be confounded with other variables thought to predict degree of foreign accent. In fact, Guion et al. (2000) reported a significant correlation between Italian–English bilinguals' AOA in Canada and the duration of their English sentences, $r\ (238) = .33$, $p < .01$. In Experiment 2 of this study, a significant correlation was found to exist between the average duration of English sentences and percentage Italian use, $r\ (62) = .34$, $p < .01$. Thus, the second aim of this experiment was to determine if the correlation between foreign accent and two variables often identified as predictors of foreign accent (AOA, percentage L1 use) remained significant when the effect of variation in L2 sentence duration was controlled statistically.

Table 3. *Mean duration (standard deviation) and speaking rate of the sentence stimuli used in Experiment 3*

		Condition		
Group	*N*	Unspeeded	Speeded	Difference
Native English	11	1805 (173) 6.70 syll/s	1601 (151) 7.55 syll/s	11.3% 0.85 syll/s
Early bilinguals	24	2008 (160) 6.01 syll/s	1796 (141) 6.72 syll/s	10.6% 0.71 syll/s
Late bilinguals	25	2096 (255) 5.80 syll/s	1867 (224) 6.52 syll/s	10.8% 0.72 syll/s

Note: The sentences were drawn from the two conditions of Experiment 2 (see text). *N*, number of stimulus sentence pairs; syll/s, syllables per second.

Method

Stimuli. Tokens of one sentence from Experiment 2 (*The woman was not in critical condition*) were selected as stimuli. The results obtained by Munro and Derwing (2001) suggested that foreign accents will be reduced to the greatest extent when L2 sentences produced by nonnative speakers are shortened by 10%. We therefore attempted to find pairs of sentences differing by this percentage.

Sentence pairs were selected solely on the basis of duration without reference to segmental articulation or prosodic features. To begin, the percentage differences in the duration of the 36 possible pairs of English sentences produced by each participant in the unspeeded and speeded conditions of Experiment 2 were calculated. (Fewer than 36 pairing were available if one or more sentence tokens was missing; see Experiment 2.) The percentage differences were rank ordered, and the pair coming closest to the 10% target value was selected. Only pairs of sentences differing by 9–13% were selected. This meant that sentences spoken by just 11 NE speakers, 24 early bilinguals and 25 late bilinguals could be included as stimuli. Sentences drawn from the unspeeded and speeded conditions will be referred to as the "longer" and "shorter" sentences, respectively.

As summarized in Table 3, duration differences between the longer and shorter stimulus sentences were comparable for all three groups. The results of a pilot experiment indicated that these long–short differences were detectable auditorily. The 60 pairs of sentences were presented to three adult listeners with training in phonetics. The shorter (speeded-condition) sentences occupied the first position of one-half of the pairs. The listeners identified the shorter member of the pairs in 96% of instances (*range* = 95–98%). No more than one listener erred for any one pair.

Procedure. Seven male and eight female native speakers of Canadian English rated the stimulus sentences for foreign accent. The listeners had a mean age of 25 years and reported an average of 18 years of formal education. All reported having normal hearing and passed a pure-tone hearing screening at octave frequencies

between 500 and 4000 Hz before being tested at the Phonetics Laboratory of the University of Ottawa.

The 120 sentence stimuli were randomly presented four times each via a loudspeaker. The listeners rated each stimulus using an equal-appearing interval scale that ranged from 1 (*most foreign accent*) to 9 (*least foreign accent*). The interval between each rating and presentation of the next stimulus was 1.0 s. A response was required for each stimulus. The listeners were told to make their best guess if uncertain and were urged to use the entire scale. No definition of foreign accent was provided, and the listeners were not told they would hear pairs of sentences spoken by the same individual.

The median of the final three ratings obtained from each listener for each sentence was determined. An examination of the median scores revealed that all of the listeners followed the instruction to use the whole scale. A high intraclass correlation was obtained for the median ratings obtained from the 15 listeners, longer sentences, $R = .988$, $p < .001$, shorter sentences, $R = 982$, $p < .001$. The decision was made, therefore, to average over the median ratings obtained from the 15 listeners for each stimulus sentence.

Results and discussion

The mean ratings obtained for the longer and shorter sentences are shown in Figure 4. Both sets of sentences spoken by the late bilinguals received lower ratings (indicating stronger foreign accents) than sentences spoken by the early bilinguals, whose sentences received lower ratings than the NE speakers' sentences. The late bilinguals' shorter sentences received higher ratings (indicating milder foreign accents) than their longer sentences whereas the reverse held true for the early bilinguals and NE speakers' sentences. The mean foreign accent ratings were examined in a Group (NE, early, late) × Stimulus Duration (longer, shorter) ANOVA, which yielded two significant main effects, Group, $F\ (2, 57) = 84.6$, $p < .05$, and Stimulus Duration, $F\ (1, 57) = 4.3$, $p < .05$; and a significant interaction, $F\ (2, 57) = 6.6$, $p < .05$.

Simple effects tests suggested that the Group × Stimulus Duration interaction arose because the sentence duration manipulation had different effects on ratings obtained for sentences produced by the three groups. The simple effect of Group was significant for both sets of sentences, longer, $F\ (2, 57) = 99.1$, $p < .05$; shorter, $F\ (2, 57) = 67.6$, $p < .05$. Tukey tests revealed that all between-group differences were significant for both sets of sentences ($p < .05$). This agrees with the results of previous studies in which early bilinguals' sentences were accorded higher ratings (indicating milder foreign accents) than late bilinguals' sentences but lower ratings than NE speakers' sentences (see Piske et al., 2001, for review).

The late bilinguals' shorter sentences received significantly higher ratings (indicating milder foreign accents) than their longer sentences, $F\ (1, 24) = 4.2$, $p = .05$. This agrees with the findings of Munro and Derwing (2001), who used a different technique to evaluate the effect of English sentence duration on late Chinese–English bilinguals' degree of foreign accent. However, the effect of duration was significant, but in the opposite direction, for the early bilinguals', $F\ (1, 23) = 6.0$, $p < .05$, and NE speakers' sentences, $F\ (1, 10) = 6.3$, $p < .05$.

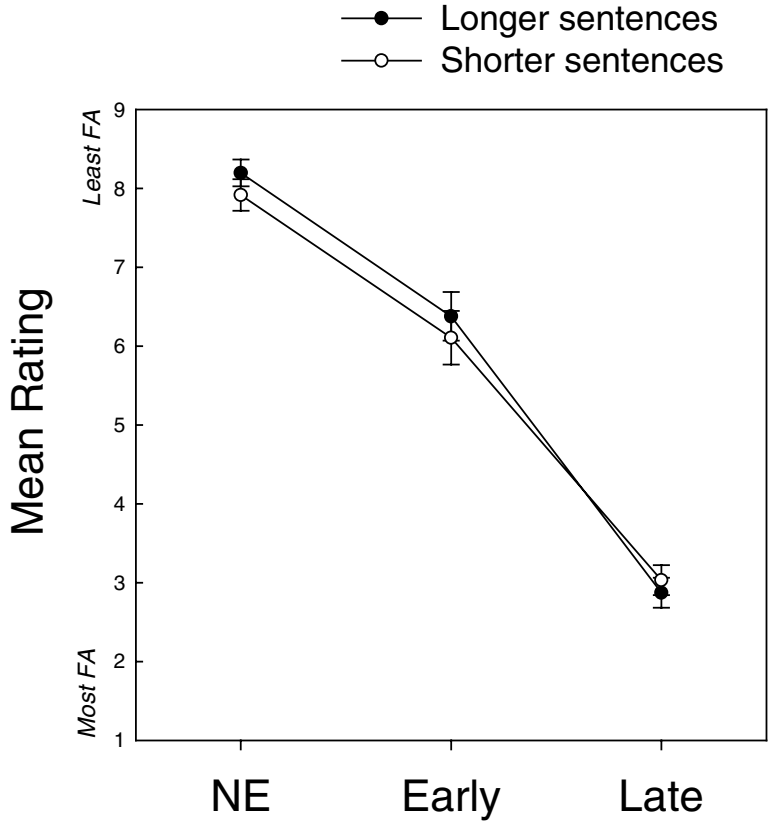

Figure 4. The mean foreign accent ratings in two sets of sentences produced by native English (NE) speakers and two groups of Italian–English bilinguals.

What accounted for the differing effect of stimulus duration for sentences produced by the early and late bilinguals? One possible explanation is that the late bilinguals made fewer segmental errors when they sped up, whereas the early bilinguals made more errors. An analysis of segmental errors ruled out this possible explanation, however.

A broad phonetic transcription of the stimulus sentences was performed. As expected from previous research, moderate correlations were found to exist between the number of segmental errors and strength of foreign accent in both sets of sentences, longer $r\ (58) = -.66$, $p < .001$; shorter $r\ (58) = -.53$, $p < .001$. The error scores were submitted to a Group (NE, early, late) $\times$ Stimulus Duration (longer, shorter) ANOVA, which yielded a significant main effect of Group, $F\ (2, 57) = 19.8$, $p < .05$. A Tukey test revealed that the late bilinguals produced significantly more errors ($M = 4.2$) than the early bilinguals and NE speakers ($M = 2.0$ and 1.6), who did not differ significantly ($p < .05$). Significantly more

errors were observed in the shorter than the longer sentences ($M = 3.3$ vs. 2.4), $F\ (1, 57) = 8.5$, $p < .05$. The lack of a significant Group × Stimulus Duration interaction, $F\ (2, 57) = 1.2$, $p > .10$, suggested that the instruction to speak as rapidly as possible resulted in a comparable increase in the number of segmental errors for all three groups. Two other analyses, which focused on error subtypes, yielded the same conclusion.[3]

Another possible explanation of the difference between the early and late bilinguals was suggested by the findings of Munro and Derwing (2001). The results obtained by these authors in a study examining normal and rate-altered English sentences spoken by late Chinese–English bilinguals suggested that the relationship between speaking rate and perceived foreign accent might be curvilinear (i.e., best described by a U-shaped function). Munro and Derwing (2001) inferred that when nonnative speakers speed up, their phonetic errors might be detected less readily, leading to an apparent improvement in L2 pronunciation. However, if nonnative speakers speed up too much, the "processing" burden on listeners may increase, creating the impression of stronger foreign accents.

The differing effect of the difference in sentence duration values on the foreign accent ratings obtained in this experiment for early and late bilinguals might be explained by the existence of an optimal speaking rate. On this view, the late bilinguals' shorter sentences sounded less foreign accented than their longer sentences because they approximated the optimal speaking rate when asked to speed up. The early bilinguals, who normally spoke at faster rates than the late bilinguals, may have exceeded the optimal rate when they were asked to speed up. However, an examination of the duration values of the stimulus sentences examined in this experiment (see Table 3) does not support an "optimal rate" explanation.

The procedures used to select stimulus sentences from the unspeeded condition (see the Method section) led to the selection of somewhat longer than average sentences for early bilinguals and somewhat shorter than average sentences for late bilinguals. As a result, an (2) AOA × (2) Stimulus Duration ANOVA examining stimulus duration values indicated that the early and late bilinguals' sentences did not differ significantly, $F\ (1, 47) = 1.9$, $p > .10$, and that AOA did not interact significantly with stimulus duration, $F\ (1, 47) = 2.9$, $p > .10$. It is unlikely that the late but not the early bilinguals' shorter sentences were produced an at optimal rate, because these two sets of sentences did not differ significantly.

Munro and Derwing (2001) suggested that the relationship between degree of foreign accent and speaking rate is best described by a U-shaped function. To evaluate this conclusion, the duration values of stimulus sentences spoken by early and late bilinguals were converted to a measure of speaking rate and plotted as a function of the foreign accent ratings obtained for these sentences. As shown in Figure 5, there was no evidence of a curvilinear relation for either group. Similar amounts of variance in foreign accent ratings were accounted for by linear and quadradic functions for both the early bilinguals ($R^2 = .16$ vs. .17) and the late bilinguals sentences ($R^2 = .17$ vs. .26).

Additional analyses confirmed that AOA and percentage of L1 use predicted degree of foreign accent when the effect of variation in sentence duration was

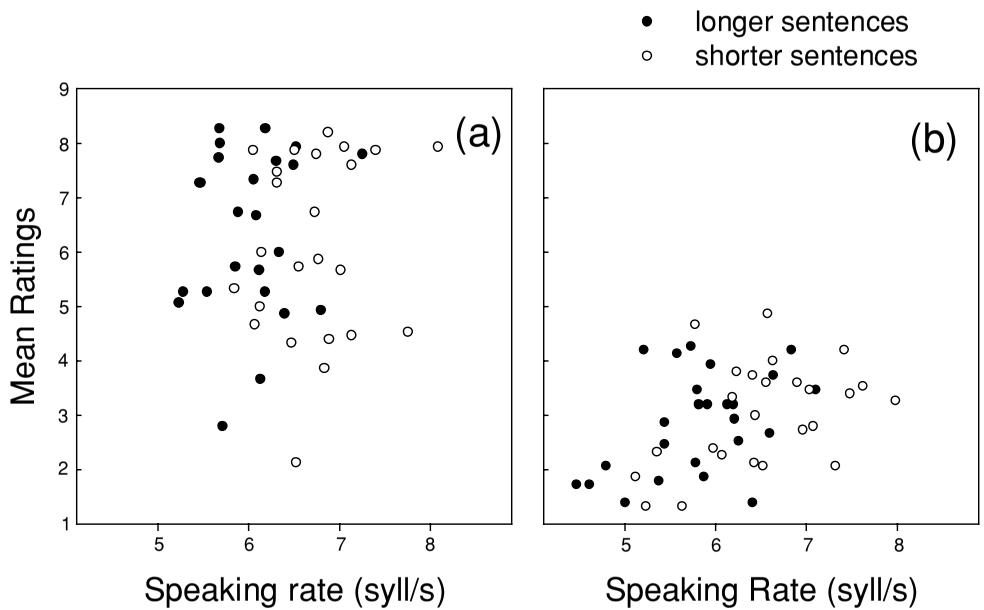

Figure 5. The relation between the mean foreign accent ratings obtained for sentences produced by (a) early and (b) late Italian–English bilinguals and the speaking rate of the sentences that were rated.

Table 4. *Simple and partial correlation analyses between the foreign accent ratings obtained for two sets of sentences produced by Italian–English bilinguals and selected variables*

Set	Variable Number/Name	Simple Corr.	Number of Variables Partialed Out		
			1	2	3
Longer	1 AOA	−.79**	−.77** (2)	−.68** (2, 3)	−.67** (2, 3, 4)
Longer	2 Use of Italian (%)	−.45**	−.38** (1)	−.37* (1, 3)	−.35* (1, 3, 4)
Longer	3 Italian ability	−.56**	−.13 (1)	−.06 (1, 2)	−.05 (1, 2, 4)
Longer	4 Stimulus dur.	−.32*	−.12 (1)	−.05 (1, 2)	−.04 (1, 2, 3)
Shorter	1 AOA	−.77**	−.75** (2)	−.67** (2, 3)	−.66** (2, 3, 4)
Shorter	2 Use of Italian (%)	−.47**	−.41** (1)	−.40** (1, 3)	−.39** (1, 3, 4)
Shorter	3 Italian ability	−.54**	−.11 (1)	−.02 (1, 2)	.01 (1, 2, 4)
Shorter	4 Stimulus dur.	−.31*	−.16 (1)	−.09 (1, 2)	−.09 (1, 2, 3)

Note: The numbers in parentheses indicate the variables(s) that were partialed out. The longer and shorter sentences were drawn from the unspeeded and speeded conditions of Experiment 2. Italian ability, self-rated ability to speak and understand Italian (average of two 7-point scales); Stimulus dur., the duration of the English sentences being rated.
$^{*}p = .05$. $^{**}p = .01$.

controlled. Stepwise multiple regression analyses were carried out to identify variables that predicted degree of foreign accent. The variables examined in the analyses of foreign accent in the longer and shorter sentences were AOA, percentage of Italian use, self-rated ability to speak and understand Italian, and duration of the stimulus sentences that were rated. AOA and Italian use account for 67.2 and 65.8% of the variance in foreign accent ratings obtained for the longer and shorter sentences, respectively ($p < .01$). Sentence duration was not identified as a significant predictor in either analysis.

Correlations between the four variables just mentioned and the foreign accent ratings obtained for the two sets of sentences are summarized in Table 4. Significant simple correlations were found to exist between the ratings obtained for both sets of sentences and AOA, percentage Italian use, Italian ability, and sentence duration. The foreign accent–AOA correlation remained significant when the effects of percentage Italian use, Italian ability, and sentence duration were partialed out. The foreign accent–Italian use correlation remained significant when the effects of AOA, Italian ability, and sentence duration were partialed out. However, the relation between foreign accent and stimulus duration became nonsignificant when the effect of one or more of the other variables was partialed out.

GENERAL DISCUSSION

The primary aim of this study was to account for why late bilinguals produce longer L2 sentences than early bilinguals do (Guion et al., 2000). Three possible

accounts of the AOA effect on L2 sentence duration values were evaluated by measuring the duration of Italian and English sentences. In Experiment 1, early and late Italian–English bilinguals repeated 12 matched pairs of Italian and English sentences a single time, without instruction as to speaking rate. In Experiment 2, bilinguals repeated two pairs of Italian and English sentences six times each without instruction as to speaking rate (the unspeeded condition), and then a second time as rapidly as possible (the speeded condition).

The hypothesis that AOA effects result from differences in previous practice producing specific words in L2 sentences was evaluated by determining if multiple repetitions of a sentence eliminated early–late differences in the duration of English sentences. This hypothesis was not supported inasmuch as between-group differences were obtained both when 12 English sentences were repeated a single time (Experiment 1) and when just two English sentences were repeated multiple times (Experiment 2). Moreover, the bilinguals' English sentence duration values did not decrease over the multiple repetitions of English sentences in Experiment 2.

The hypothesis that early–late differences are the result of different volitionally selected speaking rates received some support. We reasoned that if late bilinguals speak more slowly than early bilinguals because they are more desirous than early bilinguals of augmenting their intelligibility, then early–late differences should disappear when participants spoke as rapidly as possible. A significant early–late difference was indeed observed in the unspeeded but not the speeded condition of Experiment 2. However, the results provided only weak support for the speaking rate hypothesis inasmuch as the early–late difference narrowly missed reaching significance in the speeded condition ($p = .09$).

The study provided stronger support for the hypothesis that early–late L2 sentence duration differences are due to AOA-related differences in the relative strength of bilinguals' L1 phonetic subsystem (Guion et al., 2000). According to the suppression hypothesis, sentence duration is affected by the processing resources a bilingual must expend to suppress the language not currently in use. The greater the strength of the phonetic subsystem not in use, the more resources will be needed to suppress it and the longer sentences in the language currently in use will be. Most of the late bilinguals were Italian dominant (i.e., reported being better able to speak and understand Italian than English) whereas most early bilinguals were English dominant. The late bilinguals may therefore have needed to expend greater resources than the early bilinguals to suppress Italian while speaking English, thereby increasing the duration of their English sentences. Conversely, the early bilinguals may have needed to expend greater resources than the late bilinguals to suppress English while speaking Italian, thereby increasing the duration of their Italian sentences.

Three patterns of data were interpreted as providing support for the suppression hypothesis. First, in Experiment 1, early bilinguals produced significantly longer Italian than English sentences whereas late bilinguals produced significantly longer English than Italian sentences. This countervailing pattern was replicated in Condition 1 of Experiment 2 (where English and Italian sentences were repeated six times each) as well as in Condition 2 of that experiment (where the bilinguals shortened sentences by an average of 20% under the instruction to speak as rapidly as possible). Second, bilinguals who reported using Italian 25–75% but not 1–15%

of the time were found to have produced significantly longer English sentences in Experiment 2 than NE speakers. Finally, a significant correlation was found to exist between the average duration of English sentences in Experiment 2 and the bilinguals' percentage use of Italian, even when the effects of other variables (AOA, chronological age, length of residence in Canada, years of education in Canada, self-rated ability to speak and understand Italian) were partialed out.

An anonymous reviewer suggested that the late bilinguals may have produced longer English sentences than the early bilinguals because the late bilinguals' production of English had not yet become "fully automatic" or because of greater "articulatory difficulties" for late than early bilinguals. These explanations seem unlikely to account for the results obtained here. Nearly all of the late bilinguals had lived in Canada for more than 20 years and used English on a regular basis. Their articulatory patterns, including errored productions, were therefore likely to have become stabilized. It is uncertain how many years of L2 use is needed by bilinguals to establish automaticity in L2 speech production. However, the results obtained in naming tasks by Mägiste (1992) suggests that less than 20 years is needed. Moreover, the results of Favreau and Segalowitz (1982, 1983) suggested that between-group differences in text reading rates were not the result of differences in low-level motor skills. These authors found that the rate at which bilinguals read L1 and L2 texts corresponded to the presentation rates the bilinguals selected in another experiment as being optimal for their aural comprehension of L1 and L2 texts.

Two cautionary notes must be sounded with respect to the present findings. The first is that we have not identified the locus of the sentence duration differences that were observed to exist between early and late bilinguals. It is uncertain if the differences arose during the formulation of the syntactic and/or semantic plans for sentences, during lexical access, or while sounds (or words) were being articulated motorically. Second, although the results seem to have provided stronger support for the suppression hypotheses than for the practice effect or speaking rate hypotheses, the data in hand are insufficient to reject the latter two hypotheses. It would be useful, for example, to directly manipulate word familiarity in a future experiment and to manipulate the speaking rate in a more ecologically valid manner (e.g., asking participants to speak more rapidly under some real-world time pressure).

The second aim of this study was to evaluate the influence of sentence duration on overall degree of perceived foreign accent. When late bilinguals' sentences were shortened by 10%, they sounded less foreign accented to NE-speaking listeners. This replicated the finding obtained for late Chinese–English bilinguals by Munro and Derwing (2001). However, when early bilinguals' sentences were shortened, they sounded more foreign accented.

Several possible accounts of the differing effect of sentence duration on early and late bilinguals' degree of perceived foreign accent were considered, but none explained the results that were obtained. It was not the case, for example, that an increase in speaking rate led to fewer errors by late bilinguals but more errors by early bilinguals. The equal duration of the early and late bilinguals' stimulus sentences seems to have ruled out differences in the ability of the NE-speaking listeners to detect phonetic errors in the sentences. Importantly, all of the stimulus sentences were produced at a speaking rate that exceeded the optimal speaking rate that Munro and Derwing (2001) identified for nonnative speakers (viz., 3.76

syllables/s). More research will therefore be needed to understand how variation in L2 sentence duration affects degree of perceived foreign accent. However, the present findings showed clearly that confounds with the duration of L2 sentences were not responsible for the findings of previous research (see Piske et al., 2001, for review) that AOA and L1 use are important predictors of degree of foreign accent.

In summary, late Italian–English bilinguals produced longer English sentences than early Italian–English bilinguals. The opposite pattern was observed when matched Italian sentences were measured. The results were interpreted as providing support for the hypothesis that differences in the duration of L2 sentences spoken by early and late bilinguals are due to differences in the processing resources that they must expend to suppress their L1. On this view, the countervailing pattern of English and Italian sentence duration values observed here for early and late bilinguals was an indirect consequence of the fact that, as reported by Flege et al. (2002), most of the early bilinguals had become dominant in English whereas most late bilinguals remained dominant in their L1, Italian.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute for Deafness and other Communicative Disorders (DC00257). The authors are grateful to L. Gunnin and N. Gribben for their measurement of sentences and to Y. Hirata, A. Højen, S. Imai, M. Munro, and K. Tsukada for comments on a previous version of this article.

NOTES

1. The variable examined by Munro and Derwing (2001) was speaking rate, expressed in syllables/s. When sentences are produced without pauses, as was the case for the sentences examined here, speaking rate and sentence duration are functionally equivalent variables.
2. The 32 high L1 use and 32 low L1 use bilinguals did not differ significantly for any of these variables: AOA, $F\ (1, 62) = 0.7$; age, $F\ (1, 62) = 0.7$; years of residence, $F\ (1, 62) = 1.9$; years of education, $F\ (1, 62) = 0.9$; Italian ability, $F\ (1, 62) = 3.0$ (all p values $> .05$).
3. One analysis examined errors affecting the shape of syllables. The most frequent of these errors was the omission of /l/ in the word *critical* ($n = 14$ of 98 possible occurrences), the omission of /n/ in *in* ($n = 15$), and the rendition of the final syllable of *woman* as a syllabic nasal ($n = 14$). The other analysis examined errors not affecting syllable shape. The most frequent of these errors was the substitution of /d/ for /ð/ in *the* ($n = 56$), the substitution of /s/ for /z/ in *was* ($n = 30$), and production of the final vowels in *woman* and *condition* as full rather than reduced vowels ($n = 22$ and 16, respectively). As in the main analysis, neither analysis of error subtypes yielded a significant two-way interaction, syllable structure errors, $F\ (2, 57) = 1.8$, $p > .10$; phonetic quality errors, $F\ (2, 57) = 0.4$, $p > .10$.

REFERENCES

Anderson–Hsieh, & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, *38*, 561–614.

Bradlow, A., & Bent, T. (2002). The clear speech effect for nonnative listeners. *Journal of the Acoustical Society of America*, *112*, 272–284.
Brennan, E., & Brennan, J. (1981). Measurement of accent and attitude toward Mexican American speech. *Journal of Psycholinguistic Research*, *10*, 487–501.
Cunningham–Andersson, U., & Engstrand, O. (1989). Perceived strength and identity of foreign accent in Swedish. *Phonetica*, *46*, 138–154.
Derwing, T., & Munro, M. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, *19*, 1–16.
Favreau, M., & Segalowitz, N. (1982). Second language reading in fluent bilinguals. *Applied Psycholinguistics*, *3*, 329–341.
Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition*, *11*, 565–574.
Flege, J. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76*, 692–707.
Flege, J., Frieda, E., Walley, A., & Randazza, L. (1998). Lexical factors and segmental accuracy in second-language speech production. *Studies in Second Language Acquisition*, *20*, 155–188.
Flege, J., MacKay, I., & Meador, D. (1999). Native Italian speakers' production and perception of English vowels. *Journal of the Acoustical Society of America, 106,* 2973–2987.
Flege, J., MacKay, I., & Piske, T. (2002). Assessing bilingual dominance. A*pplied Psycholinguistics*, *23*, 567–598.
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. (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. (1995b). The effect of age of second language learning on the production of English consonants. *Speech Communication*, *16*, 1–26.
Flege, J., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America, 99*, 1161–1173.
Flege, J., Yeni–Komshian, G., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language, 41*, 78–104.
Gaitenby, J. (1965). The elastic word. *Haskins Laboratories Status Report on Speech Research*, *2*, 1–12.
Geffen, G., & Luszcz, M. (1983). Are the spoken durations of rare words longer than those in common words? *Memory & Cognition*, *11*, 13–15.
Geffen, G., Stierman, I., & Tildesley, P. (1979). The effect of word length and frequency on articulation and pausing during delayed auditory feedback. *Language and Speech*, *22*, 191–199.
Green, D. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, *27*, 210–223.
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.
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.
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.
Little, R., & Rubin, D. (1987). *Statistical analysis with missing data*. New York: Wiley.
Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, *26*, 381–400.
Mägiste, E. (1992). Second language learning in elementary and high school students. *European Journal of Cognitive Psychology*, *4*, 355–365.
Major, R. (1986). Paragoge and degree of foreign accent in Brazilian English. *Second Language Research*, *2*, 53–71.
Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.
Miller, J., Green, K., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica*, *43*, 106–115.

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. (2001). Modeling perceptions of the accentedness and comprehensibility. *Studies in Second Language Acquisition*, *23*, 451–468.
Munro, M., Flege, J., & MacKay, I. (1996). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics*, *17*, 313–334.
Picheny, M., Durlach, M., & Braida, L. (1989). Speaking clearly for the hard of hearing: III. An attempt to determine the contribution of speaking rate differences in intelligibility between clear and conversational speech. *Journal of Speech and Hearing Research*, *32*, 600–683.
Piske, T., Flege, J., MacKay, I., & Meador, D. (2002). The production of English vowels by fluent early and late Italian–English bilinguals. *Phonetica*, *59*, 49–71.
Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*, 191–215.
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.
Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, *12*, 380–386.
van Wijngaarden, S., Steeneken, H., & Houtgast, T. (2001). The effect of a non-native accent in Dutch on speech intelligibility. *Journal of the Acoustical Society of America*, *109*, 2473(A).
van Wijngaarden S., Steeneken H., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *Journal of the Acoustical Society of America*, *112*, 3004–3013.
Wright, C. (1979). Duration differences between rare and common words and their implication for the interpretation of word frequency effects. *Memory & Cognition*, *7*, 411–419.