Factors affecting strength of perceived foreign accent in a second language

James Emil Flege
Department of Biocommunication, University of Alabama, Birmingham, Alabama 35294

Murray J. Munro
Department of Linguistics, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Ian R. A. MacKay
Department of Linguistics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada

This study assessed the relation between non-native subjects' age of learning (AOL) English and the overall degree of perceived foreign accent in their production of English sentences. The 240 native Italian (NI) subjects examined had begun learning English in Canada between the ages of 2 and 23 yr, and had lived in Canada for an average of 32 yr. Native English-speaking listeners used a continuous scale to rate sentences spoken by the NI subjects and by subjects in a native English comparison group. Estimates of the AOL of onset of foreign accents varied across the ten listeners who rated the sentences, ranging from 3.1 to 11.6 yr ($M = 7.4$). Foreign accents were evident in sentences spoken by many NI subjects who had begun learning English long before what is traditionally considered to be the end of a critical period. Very few NI subjects who began learning English *after* the age of 15 yr received ratings that fell within the native English range. Principal components analyses of the NI subjects' responses to a language background questionnaire were followed by multiple-regression analyses. AOL accounted for an average of 59% of variance in the foreign accent ratings. Language use factors accounted for an additional 15% of variance. Gender was also found to influence degree of foreign accent.

INTRODUCTION

Critical periods have been observed in a wide range of animal behaviors, including vocal learning in birds (e.g., Gould and Marler, 1989). Lenneberg (1967) introduced the critical period concept to research in native language acquisition. He also observed that "foreign accents [in an L2] cannot be overcome easily after puberty" (p. 176). Patkowski (1990) defined the critical period for L2 speech learning as an "age based constraint on the acquisition of full native fluency" in the pronunciation of individuals who have reached their ultimate attainment in the L2 (p. 74). Indeed, many previous studies have shown that earlier is usually better as far as the pronunciation of an L2 is concerned (Suter, 1976; Asher and Garcia, 1969; Seliger *et al.*, 1975; Oyama, 1976; Tahta *et al.*, 1981; Purcell and Suter, 1980; Patkowski, 1990; Thompson, 1991; Flege, 1988b; Flege and Fletcher, 1992). Both the *proportion* of individuals observed to speak their L2 with a detectable accent, as well as the *strength of perceived foreign accents* among individuals with detectable foreign accents have been found to increase as the age of learning an L2 increases.

The foreign accent studies just cited are consisted with the claim that a critical period exists for speech learning, although the actual cause(s) of foreign accents remain uncertain (Flege, 1987a; Long, 1990). Some believe that age-related changes in L2 pronunciation are due to a loss of neural plasticity, or to neurofunctional reorganization which affects the processing and/or storage in long-term memory of information pertinent to the L2 sound system (e.g., Penfield and Roberts, 1959; Lenneberg, 1967; Lamendella, 1977; Neville *et al.*, 1992; Paradis, 1995). Although never demonstrated directly, neurological changes might diminish the ability to add or modify sensorimotor programs for producing the vowels and consonants of an L2 (Sapon, 1952; McLaughlin, 1977), or to establish perceptual representations for new vowels and consonants (Flege, 1992a,b; Rochet, 1995). However, many other explanations for foreign accents have been proposed in the literature, including attitudinal and psychosocial factors (for reviews, see Flege, 1988a; Bialystock and Hakuta, 1994). This suggests that multiple factors may contribute to the presence of foreign accent.

Based on a literature review, Long (1990) concluded that the L2 is generally spoken without accent up to an age of learning (AOL) of 6 yr; with a foreign accent by nearly all subjects having AOLs greater than 12 yr; and either with or without foreign accent by subjects in the intervening AOL range. However, despite the many foreign accent studies that have been undertaken, three important questions remain unanswered: (1) What is the *earliest* AOL at which persistent foreign accents become common? (2) What is the *latest* AOL at which accent-free pronunciation of an L2 remains possible? (3) Does the critical period for speech learning affect *all* individuals? As to the "AOL onset" question, Flege and Fletcher (1992) provided indirect evidence that foreign accents may be evident in the speech of adults who began learning their L2 as early as 7 yr of age. English sentences

spoken by Chinese adults who first arrived in the United States (US) at an average age of 7.6 yr received significantly lower ratings than did sentences spoken by native English subjects. The ratings obtained for sentences spoken by native Spanish subjects who had arrived in the US by the age of 6 yr, on the other hand, did *not* differ significantly from the ratings obtained for native-produced sentences. As to the "AOL offset" question, Scovel (1988) claimed that foreign accents are present in the speech of most if not all individuals who began learning their L2 after the age of 12 yr. Patkowski (1990) concluded that a critical period ends somewhat later, at the age of 15 yr.

The foreign accent studies cited earlier have provided converging evidence concerning age effects on L2 pronunciation, but each study has suffered from one or more methodological limitations. A comprehensive study is needed to resolve the AOL onset, AOL offset, and universality questions. The present study met this need by (1) examining a population of non-native subjects who had reached their ultimate attainment in English, (2) sampling over the entire range of AOLs in which L2 pronunciation accuracy might be expected to vary, (3) examining speakers of a single native language, and (4) using a scaling technique that was sufficiently sensitive to reveal even small divergences from the phonetic norms of English. Native English-speaking listeners used a continuous scale to rate sentences spoken by 240 native Italian (NI) adults who differed primarily according to the age at which they had first begun learning English. Two recent studies showed that the Italian subjects' accuracy in producing vowels (Munro *et al.*, 1995) and consonants (Flege *et al.*, 1995) in isolated English words decreased systematically as AOL increased. We therefore expected the degree of perceived foreign accent in English sentences spoken by the NI subjects to increase as a function of AOL. To evaluate factors in addition to AOL that might affect degree of accent, the NI subjects' responses to a language background questionnaire were submitted to principal components analyses followed by multiple-regression analyses.

I. METHOD

A. Subjects

All subjects examined in this study were living in Ottawa, Ontario when they were recorded. The native English subjects (10 males, 14 females) had all been born in Canada. (They will be designated in figures as having an age of arrival in Canada of "0" yr.) None of the native English (NE) subjects had been exposed to Italian in childhood or had learned Italian later in life. The 110 male and 130 female NI subjects were all born in Italy. The NI subjects were slightly older than the NE subjects (*M* =44 vs 39 yr). They had arrived in Canada between the ages of 2 and 23 yr (*M* =13 yr), and had lived in Canada for 15 to 44 yr (*M* =32 yr). The NI subjects reported using English more often than Italian on a daily basis (69% vs 28%), and said they pronounced Italian better than English (5.5 vs 5.2 on a 7-point scale). A few NI subjects reported speaking French, but self-reported use of French averaged less than 3%.

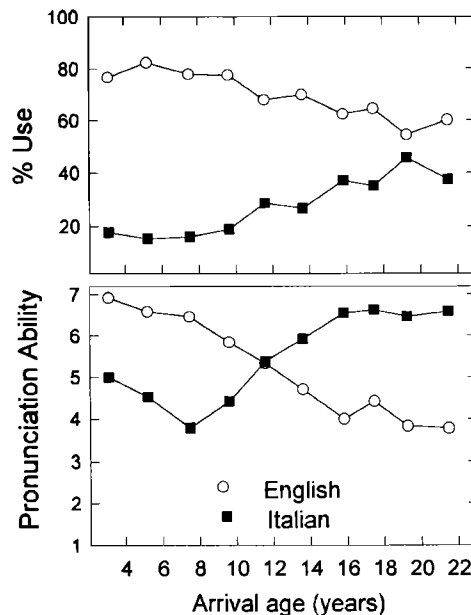As is unavoidable in studies of large immigrant popula-

FIG. 1. (top) The native Italian subjects' mean estimates of their relative use of English and Italian. (bottom) The native Italian subjects' self-estimated ability to pronounce English and Italian on a scale ranging from "poor" (1) to "good" (7).

tions (e.g., Bachi, 1956; Bahrick *et al.*, 1994), a number of "natural" confounds existed between the primary independent variable of the study, AOL, and a number of other variables that might be related to L2 pronunciation. As shown in Fig. 1 (top), the NI subjects' estimates of how frequently they used English and Italian varied as a function their age of arrival in Canada. Those who arrived in early childhood reported using English more than Italian. The gap between English and Italian narrowed as the NI subjects' arrival age increased. However, even those subjects who arrived in Canada as adults reported using English more than Italian.

The NI subjects' self-reports concerning how well they pronounced English and Italian also varied with age of arrival. As shown in Fig. 1 (bottom), subjects who arrived before the age of 12 yr reported pronouncing English better than Italian, whereas the reverse held true for those who arrived *after* that age. Significant correlations existed between self-reported frequency of use of English and ability to pronounce English [$r = 0.348$, $df = 238$, $p < 0.01$], and between self-reported use and pronunciation of Italian [$r = 0.468$, $df = 238$, $p < 0.01$]. A significant inverse correlation existed between age of arrival and length of residence (LOR) in Canada [$r = -0.437$, $p < 0.01$]. This confounded the design. The earlier the NI subjects began learning English, which should contribute to accurate L2 pronunciation, the longer they had spoken English, which might also contribute to accurate L2 pronunciation. However, as will be shown below, LOR contributed little to variation in degree of perceived foreign accent.

The NI subjects were asked which of their two languages they spoke best. Of those who arrived in Canada by the age of 12 yr, 90% indicated that English was their better language, whereas only 26% of those who arrived later in life gave that response. The subjects were also asked which of their two languages they would be most reluctant to lose

TABLE I. Characteristics of subgroups of native Italian subjects ($n = 24$ per subgroup).

Group	Age of arrival[a]		First exposure[b]		Age of acquisition[c]	
	M	Range	M	Range	M	Range
1	3.1	1.9–4.1	4.3	2.6–6.5	5.6	4.5–7.0
2	5.2	4.2–6.4	5.4	4.2–6.4	6.3	5.0–9.0
3	7.5	6.5–8.6	7.5	6.5–8.6	8.3	7.0–10.0
4	9.6	8.7–10.6	9.6	8.7–10.6	10.6	9.0–15.0
5	11.6	10.6–12.6	11.6	10.6–12.6	12.6	11.0–15.0
6	13.6	12.7–14.8	13.6	12.7–14.8	14.6	13.0–16.0
7	15.8	15.0–16.8	15.9	15.0–16.8	17.6	15.0–25.0
8	17.5	16.8–18.5	17.5	16.8–18.5	19.1	17.0–25.0
9	19.3	18.5–20.2	20.2	18.0–28.0	22.1	19.0–30.0
10	21.5	20.2–23.2	21.3	16.0–23.2	23.5	21.0–28.0

[a]Chronological age of arrival in Canada, in years.
[b]Age of first exposure to English in Canada, in years.
[c]Age at which the NI subjects estimated they could first speak English comfortably, in years.

through injury or illness. Despite differences in self-reported ability to speak English, roughly the same percentages of early- and late-arriving subjects reported being less willing to lose English than Italian (86% vs 79%). This suggests that English was very important to *all* of the NI subjects, not just the early arrivals.

The NI subjects were assigned to subgroups based on their age of arrival in Canada. As shown in Table I, the average arrival age of the 24 subjects in ten successive NI subgroups increased in roughly 2-yr increments. Also as shown in the table, the NI subjects reported having first been exposed to English 0.2 yr after their arrival in Canada, on average. They estimated being able to speak English "comfortably" 1.6 yr after their arrival in Canada, on average. Of these three age variables (i.e., age upon arrival, first exposure, and acquisition), age of arrival has been used most often in previous research as an index of the age of learning (AOL) a second language. It will therefore be the index of AOL used here.

B. Procedure

The subjects were tested individually by a single experimenter (MJM) in a quiet room of a Catholic church in Ottawa, Ontario. A delayed repetition technique was used to elicit both isolated words and sentences. Only the sentences will be examined in the present study. The sentences examined (*I can read this for you; The red book was good; Paul ate carrots and peas; The good shoe fit Sue; He turned to the right*) did not sample the full range of English allophones, nor did they represent the full range of prosodic dimensions that distinguish English from Italian. These sentences were not designed to be especially difficult for Italian speakers of English; three of them were drawn from previous studies (Flege and Eefting, 1987; Flege, 1988b; Flege and Fletcher, 1992). The sentences were presented in written form and also aurally, via a tape recording, to reduce the possibility that difficulty in reading might masquerade as foreign accent. Each sentence to be produced was preceded and followed on the tape by a context sentence, as in the following mini-dialogue:

(1) Voice 1: *In which direction did he turn?* (pause)
(2) Voice 2: *He turned to the right.* (pause)
(3) Voice 1: *In which direction did he turn?* (longer pause)

Although the sentences were modeled on the tape by voice 2, the delay between the model and its repetition, as well as the intervening speech material, probably prevented direct imitations from sensory memory. Three tokens of each sentence were elicited. The five sentences were expected to yield comparable foreign accent ratings.

C. Foreign accent ratings

The second token of each sentence was low-pass filtered at 10.0 kHz and digitized at 22.05 kHz on a PC. The 1320 sentences were normalized for peak intensity, then presented to ten NE-speaking listeners (four males, six females) residing in Birmingham, Alabama. The listeners, whose mean age was 27 yr (range: 20–38), came from towns in Ontario (Windsor-2, Brantford, Georgetown, Kenora, Hamilton) and elsewhere in Canada (Calgary, Alberta-2; Edmonton, Alberta; Sherbrook, Quebec). None had training in a speech-related discipline. Each listener passed a pure-tone hearing screening before participating.

The listeners were tested individually in a sound booth. The 264 tokens of each sentence were randomly presented via headphones four times each in separate sessions. A 1-s interval occurred between each rating and presentation of the next sentence. The order in which the five sentences were tested was counterbalanced across listeners. The listeners were told that most but not all sentences had been spoken by Italian immigrants to Canada, and that the first of the four ratings they gave to each sentence would be "just for practice." The listeners positioned the lever on a response box at some point along a range defined by the labels "native speaker of English—no foreign accent" (top), "medium foreign accent" (middle), and "native speaker of Italian—strongest foreign accent" (bottom). The listeners were told to reserve the top of the range for sentences spoken by native speakers of English, and to reserve the very bottom of the range for the *one talker* they judged to have the strongest foreign accent of all talkers being considered. Depending on
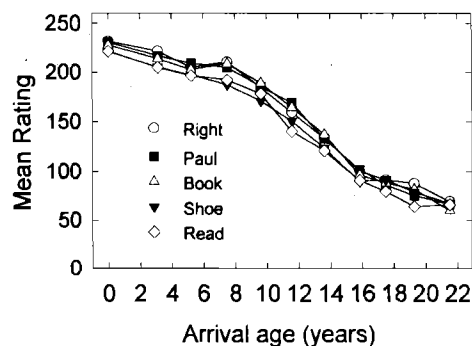
FIG. 2. Mean ratings obtained for five sentences spoken by native English subjects (AOL=0) and ten subgroups of native Italian subjects differing in age of learning English.

FIG. 3. Mean ratings obtained for ten native English-speaking listeners of sentences spoken native English subjects (AOL=0) and ten subgroups of native Italian subjects differing in age of learning English.

the lever's position, a value ranging from 0 (strongest foreign accent) to 255 (no foreign accent) was stored on the PC used for testing. A mean was calculated for the final three ratings obtained for each sentence.

The lowest mean ratings given by the ten listeners to any individual subject (averaged over sentences) ranged from 0 to 22. The lowest mean rating given to any *subgroup* of 24 subjects was 66. This indicates that the listeners followed the instruction to reserve their lowest rating for the one subject who had the strongest foreign accent. Preliminary analyses revealed that the order in which the listeners heard the five sentences did not exert a significant effect on between-group differences, nor were important variations found to exist over the five days of testing. These factors will therefore not be discussed further.

II. RESULTS

A. Effects of AOL

We began by evaluating our assumption that the five English sentences would receive similar foreign accent ratings. A mean rating for each subject's production of the five sentences was calculated, averaged over the ten listeners. The mean values obtained for the 11 subgroups are shown in Fig. 2 as a function of sentence. Visual inspection of this figure reveals that the NI subjects' degree of perceived foreign accent increased as AOL increased. The later the NI subjects began learning English, the lower were the foreign accent ratings for each sentence. The sentence-based scores were submitted to an (11) group×(5) sentence ANOVA, with repeated measures on sentence. As expected, the two-way interaction was nonsignificant, $[F(40,1012)=1.09, p=0.322]$. The ratings considered in all subsequent analyses will therefore be averaged over sentences.

The foreign accent ratings given by each listener were plotted as a function of group, then inspected visually. The overall pattern for all ten listeners appeared similar. Of the 11 groups of 24 subjects each, the NE group received the highest mean rating from all ten listeners (range: 199–252). The NI subjects who arrived last in Canada were accorded the lowest group mean average by each listener (range: 34–88). The listeners did differ somewhat, however, in their use of the rating scale. Some tended to avoid using the lowest por-
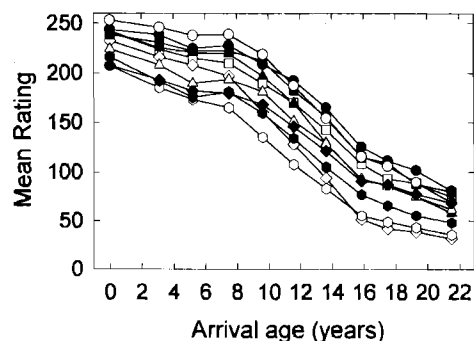
tion of the scale (see above) whereas others tended to avoid using the highest portion of the scale, which may have indicated a lack of confidence that NE subjects were indeed native speakers of English. When the ranges of mean ratings given by the ten listeners to the 264 subjects were considered, it was found that only one listener's ratings spanned the entire available range (L1: 8–255; L2: 22–244, L3: 10–247, L4: 13–255, L5: 2–254, L6: 1–244, L7: 9–255, L8: 18–253, L9: 0–255, L10: 6–244). Accordingly, the 264 mean ratings obtained from the other nine listeners were normalized so that they too would span the entire possible range, viz., 0–255.

The normalized ratings are shown in Fig. 3. The overall effect of AOL seemed to be much the same for each listener. To test this, the 264 mean ratings obtained from each listener were submitted to an (11) group×(10) listener ANOVA, with repeated measures on listener. Contrary to our expectation, this analysis yielded a significant two-way interaction $[F(90,2277)=6.71, p<0.01]$. The simple main effect of group was found to be significant for all ten listeners ($p <0.01$).

Williams' test ($\alpha=0.005$) was used to determine, for each listener, which of the NI subgroups received significantly lower ratings than did the NE subjects.[1] The AOL of the first foreign-accented NI subgroup varied across the ten listeners. One listener gave significantly lower ratings to the first NI subgroup than to the NE subjects. This group consisted of subjects with a mean AOL of 3.1 yr. For three other listeners, the first foreign-accented subgroup was the group consisting of NI subjects with an average AOL of 5.2 yr. The remaining listeners first detected foreign accent in the NI subgroups made up of subjects having average AOLs of 7.5 (two listeners), 9.6 (three listeners), and 11.6 yr (one listener). Averaging over the ten listeners, 7.4 yr was the AOL of the first foreign-accented NI subgroup.

A histogram with a bin size of ten was prepared to provide insight into how the first two NI subgroups differed from the NE subjects. A total of 1200 mean ratings were available for each of these groups (24 subjects×10 listeners×5 sentences). As shown in Fig. 4, the NI and NE subjects differed in two respects. The NI subjects received the highest possible rating less often than did the NE subjects. They also received more ratings of 100–200 than did the NE subjects. As noted earlier, however, these differences
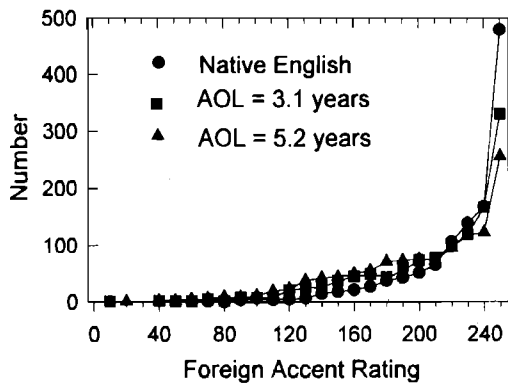
FIG. 4. The frequency of foreign accent ratings obtained for 24 native English subjects and for 24 native Italian subjects each who began learning English at average ages of 3.1 and 5.2 yr. A total of 1200 mean ratings were available for each of the three groups.

FIG. 5. Separate estimates, for each of ten listeners, of how many subjects in each native Italian subgroup pronounced English sentences "authentically" (see text).

between NE and NI subjects were due in large part to a subset of the listeners.

Patkowski (1990) claimed that a step function best describes the relation between AOL and degree of perceived foreign accent. More specifically, he claimed that foreign accent increases precipitously after a critical period has been passed at the age of 15 yr. Visual inspection of the data in Fig. 3 fails to support this claim. The r^2 values obtained when first-, second-, and third-order equations were fit to the overall mean foreign accent ratings obtained for the 240 NI subjects differed little (0.714, 0.716, and 0.742, respectively). The r^2 values obtained when functions were fit to mean ratings obtained for just the 63 NI subjects with AOLs of 13–17 yr were virtually identical (0.174, 0.175, and 0.179). Similar results were obtained when these analyses were applied to mean values obtained for each of the ten listeners. If degree of perceived foreign accent increased precipitously at the age of 15 yr, one would have expected the third-order functions to account for more variance in the foreign accent ratings than did the first-order functions.

We used the following statistical procedure to address the AOL offset question. Subjects who received a mean rating that fell within 2.0 standard deviations (s.d.s) of the mean rating obtained for the 24 NE subjects were considered to have spoken the English sentences authentically. Subjects who received a rating that fell more than 2.0 s.d.s below the NE mean were considered to have spoken the sentences with a foreign accent. Separate classifications were performed for each of the ten listeners.

The 24 NE subjects could have been misclassified as "foreign accented" in 240 instances (24 subjects×10 listeners). However, misclassifications occurred in just seven (2.9%) instances, which is a low rate for a study of this kind (see, e.g., Asher and Garcia, 1969). The ratings given by five listeners to one of the NE subjects resulted in a misclassification, and the ratings given by two other listeners to another NE subject resulted in misclassifications. It is at least possible that these instances were not actually *misclassifications*. The two NE subjects just mentioned were brothers aged 40 and 43 yr. In screening NE subjects, we were careful to exclude anyone who had been exposed to Italian in childhood,
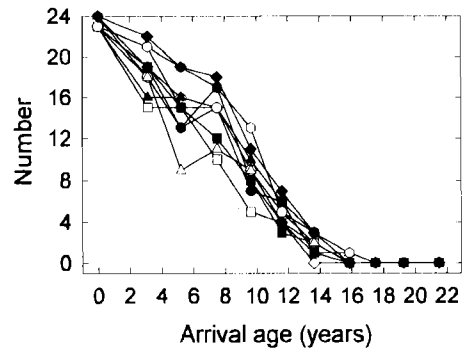
or who had learned Italian later in life. Neither of the two NE subjects just mentioned had been exposed to Italian. However, subsequent inquiry revealed that, as young children, they had been spoken to in French by their mother and maternal grandmother. Thus the "misclassifications" may actually have been correct detections of very subtle *French* accents in sentences spoken by the two NE subjects.

Figure 5 shows the number of subjects in each subgroup (maximum=24) who met the statistical criterion for "authentic" pronunciation of the English sentences. The number of NI subjects decreased steadily as AOL increased. Similar results were evident for each of the listeners. Averaged over listeners, 78% of the NI subjects having AOLs of less than 4 yr met the criterion, as did 61% of the subjects with AOLs of 4–8 yr and 29% of the subjects with AOLs of 8–12 yr. Just 6% of the 120 NI subjects who began learning English after the age of 12 yr met the criterion for authentic pronunciation. No subject who began learning English after the age of 16 yr met the authenticity criterion for any of the ten listeners.

B. Gender

Previous studies have provided divergent results concerning the effect of gender on L2 pronunciation (Asher and Garcia, 1969; Tahta *et al.*, 1981; Thompson, 1991; Purcell and Suter, 1980), probably because the results obtained for subjects differing in AOL were averaged. As shown in Fig. 6, female NI subjects with AOLs of less than 12 yr received higher ratings than did males matched for AOL, whereas
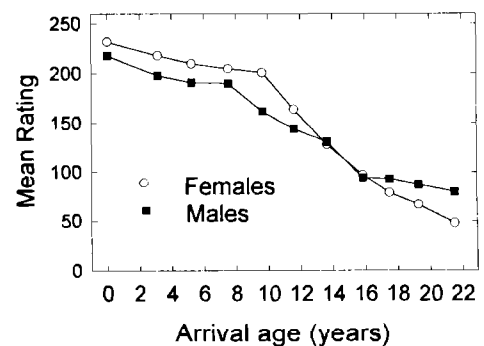

FIG. 6. Mean ratings obtained for English sentences spoken by male and female native speakers of English and Italian (see text).

TABLE II. Factors with eigenvalues greater than 1.0 that were identified in principal components analyses examining responses to 34 questionnaire items, in descending order of importance. The two or three items with the highest loading are shown in parentheses.

Factor	LBQ items with highest loadings	Factor name
	Male subjects	
1	AOE (0.947), AOA (0.944), AOAcq (0.936)	age of learning
2	I4 (0.906), E1 (−0.830), I2 (0.723)	home use
3	E2 (−0.864), I5 (0.779), PEN (−0.667)	work use
4	I6 (0.826), E3 (−0.670), I1 (0.495)	social use
5	ABIL5 (0.822) I12 (−0.703), ABIL4 (0.505)	imitation ability
6	LOR (0.849), AGE (0.805)	length of residence
7	E6 (0.776), E5 (0.581), ABIL3 (−0.542)	concern for pronunciation
8	E14 (0.722), E15 (0.577), E5 (0.434)	instrumental motivation-1
9	E8 (0.779), E7 (0.560)	integrative motivation
10	E9 (0.682), E12 (0.682), ABIL6 (0.513)	instrumental motivation-2
11	I11 (−0.736), TII (0.550)	use of Italian
	Female subjects	
1	AOA (0.931), AOE (0.923), AOAcq (0.914)	age of learning
2	I4 (0.783), E1 (−0.756), I2 (0.684)	home use
3	E5 (0.822), I11 (−0.716), E6 (0.696)	concern for pronunciation
4	I6 (0.855), E3 (−0.675), I1 (0.556)	social use
5	PEN (−0.800), PIT (0.757), I7 (0.446)	overall language use
6	E12 (0.744), E14 (0.700), E15 (0.637)	instrumental motivation
7	ABIL5 (0.817), I12 (−0.698), ABIL6 (0.451)	imitation ability
8	AGE (0.854), LOR (0.853)	length of residence
9	E2 (0.865), I5 (−0.604)	work use
10	E13 (0.791), I5 (0.403)	integrative motivation
11	TII (0.649), ABIL4 (−0.438)	use of Italian

ratings were lower for females than males with AOLs greater than 16 yr. To assess the effect of gender, the average ratings obtained for the 264 subjects were submitted to an (11) group×(2) gender ANOVA. The significant two-way interaction yielded by the ANOVA [$F(10,242)=2.80$, $p<0.01$] was explored by tests of the simple effect of gender. Female NI subjects with an average AOL of 9.6 yr received significantly higher ratings than did males matched for AOL. Female NI subjects with an average AOL of 21.5 yr, on the other hand, received significantly *lower* ratings than did AOL-matched males ($p<0.01$). No other difference reached significance. Additional ANOVAs were carried out in an attempt to account for the gender–AOL interaction. These analyses focused on 24 male and 24 female subjects with AOL of 4–9 or 17–22 yr (96 subjects in all). The ANOVAs examined responses to questionnaire items pertaining to the frequency of use of English and Italian, attitudes towards pronunciation, and motivation. These analyses failed to provide a straightforward account of the interaction, and so will not be presented here.[2]

C. Other factors

1. Questionnaire

The language background questionnaire administered to the NI subjects consisted of 34 items. (A copy will be furnished upon request.) Multiple items were included to provide information pertinent to factors identified in the literature as potentially important to L2 pronunciation, including amount of formal instruction in the L2 (Purcell and Suter, 1980; Flege and Fletcher, 1992), musical and mimicry ability (Purcell and Suter, 1980; Thompson, 1991), strength of con-

cern for pronunciation (Purcell and Suter, 1980), motivation (Gras, 1983), and overall amount of L2 input or use (Purcell and Suter, 1980; Tahta *et al.*, 1981). For example, the subjects were queried as to how often they used both English and Italian in a number of specific situations, as well as how often in general they used their two languages. As expected, significant negative correlations were obtained for all five pairs of questions dealing with frequency of use of Italian and English [$M=-0.665$; $p<0.001$].

2. Regression analyses

The NI subjects' responses to the questionnaire were analyzed to identify factors that might influence degree of perceived foreign accent. Separate principal components analyses were carried out to examine the male and female subjects' responses to the questionnaire. The purpose was to identify common underlying factors. A varimax rotation was performed, which retained the orthogonality of the factors while optimally separating items related to one another from those that were not (see Metz *et al.*, 1985). The analysis for male subjects yielded 11 factors which accounted for 75% of variance in the 34 questionnaire items. The analysis for females subjects accounted for 74% of variance. Some of the underlying factors, which are summarized in Table II, were similar for the male and female subjects, and thus might be considered to be the "same" factor. For example, the same three questionnaire items had the highest loadings on an "AOL" factor for both genders (viz., age of arrival in Canada, age of self-reported first exposure to English, and age of self-reported acquisition of English; see Table I). For

3130   J. Acoust. Soc. Am., Vol. 97, No. 5, Pt. 1, May 1995

Flege *et al.*: Perceived foreign accent   3130

TABLE III. Results of stepwise multiple-regression analyses examining the relation between sentence ratings and factors derived by principle components analyses (see Table II).

Step	Variable	R-square[a]	Change[b]	F value[c]	Prob.[d]
		Male subjects			
1	F1: age of L2 learning	0.547	0.547	220.09	0.000
2	F3: language used at work	0.634	0.087	35.15	0.000
3	F4: languages used socially	0.681	0.046	18.66	0.000
4	F6: length of residence	0.700	0.019	7.72	0.006
5	F2: languages used at home	0.719	0.019	7.66	0.007
6	F7: concern for pronunciation	0.735	0.016	6.51	0.012
7	F9: motivation-2	0.746	0.011	4.33	0.039
		Female subjects			
1	F1: age of L2 learning	0.633	0.633	408.5	0.000
2	F4: language used socially	0.690	0.056	36.5	0.000
3	F2: languages used at home	0.743	0.053	34.2	0.000
4	F5: overall language use	0.785	0.041	26.9	0.000
5	F8: length of residence	0.801	0.016	10.8	0.001
6	F11: native language loyalty	0.809	0.007	4.7	0.032

[a]Cumulative variance accounted for at each step.
[b]Increase in variance from the preceding step.
[c]F-value testing the significance at each step.
[d]Probability that adding the last variable accounted for significantly more variance.

both genders, the same three questionnaire items had the highest loadings on a factor designated home use (viz., frequency of Italian use in the home, frequency of English use in the home, and frequency of Italian use with children or younger relatives). Factors designated work use, social use, imitation ability, length of residence in Canada, and use of Italian were also identified for both genders.

Other factors identified for the male and female subjects differed, however. A factor designated overall language use, which had high loadings on items dealing with the overall use of both English and Italian, was identified for the female but not the male subjects. Two factors identified for the female subjects, and three factors for male subjects, had high loadings on items that dealt with motivation. Of these, just the factor designated instrumental motivation seemed to be shared by the two genders. It had high loadings on items that dealt with the practical benefits of learning English well. Two factors identified for just the male subjects were designated instrumental motivation and integrative motivation. The former had high loadings on items that seemed to reflect a desire to achieve competence in English in order to achieve specific (often economic) goals; the latter factor had high loadings on items that might be thought to reflect a desire to master English for social and cultural reasons.

Factor scores were calculated for each of the 240 NI subjects by applying the scoring coefficients generated by the principal components analysis to standardized values for subjects' responses to the 34 questionnaire items. The 11 factor scores were then submitted to stepwise multiple-regression analyses, one for each gender. The dependent variable in these analyses was the overall average foreign accent ratings obtained for the 110 male and 130 female NI subjects. The regression model accounted for slightly more variance in the female than male subjects' ratings (81% vs 75%). As summarized in Table III, the AOL factor accounted for more variance than did any other factor (females: 63%, males: 55%).

Three "language use" factors accounted for 15% of variance in the foreign accent ratings obtained for both genders. However, the nature of these factors differed across genders. The factors identified for males, designated work use, social use, and home use, accounted for 9%, 5%, and 2% of variance, respectively. For females, factors designated social use, home use, and overall language use accounted for 6%, 5%, and 4% of the variance. Factors designated integrative motivation and concern for pronunciation accounted for 3% of variance in the sentence ratings obtained for the male subjects. No motivational factor accounted for a significant amount of variance for the female subjects.

A length of residence factor was identified as a significant predictor of the foreign accent ratings obtained for both genders. A common observation in the L2 literature is that phonology learning proceeds rapidly in early stages, then stabilizes. The shortest amount of time any NI subject had resided in Canada was 15 yr. The LOR factor accounted for little variance for either the male (1.9%) or the female subjects (1.6%), probably because of a ceiling effect.

III. DISCUSSION

The aims of this study were to determine when foreign accents first emerge, the latest age at which accent-free pronunciation of a second language is possible, and whether a critical period affects all individuals who learn a second language. Foreign accent ratings were obtained for English sentences spoken by native English (NE) and native Italian (NI) subjects. The NI subjects differed according to their age of learning (AOL) English. The first NI subgroup to receive significantly lower ratings than the NE subjects varied across the ten NE-speaking listeners who rated the subjects' productions of English sentences. The most discerning listener gave significantly lower ratings to the first Italian subgroup than to the NE subjects. This NI subgroup consisted of individuals with a mean AOL of 3.1 yr. For the least discerning listener,

the first NI subgroup to receive significantly lower ratings had a mean AOL of 11.6 yr. Averaged over all ten listeners, the average AOL for the first foreign-accented subgroup was 7.4 yr.

No previous study has tested for differences among individual listeners in ability to detect or scale foreign accent. Scovel (1981) found, however, that a group of 5- to 10-yr-old children were less able to detect foreign accent than were groups of older children and adults. The basis for the listener effect obtained in the present study is unknown. It may have been due to small differences in the native dialect of Canadian English spoken by the ten listeners (see Sec. I), to idiolectal differences in their representations for segmental and prosodic characteristics of English, or to differences in "tolerance" ranges for English phonetic structures. Long (1990) speculated that individuals who have encountered many varieties of English, including foreign-accented English, are less willing to judge a speech sample to be foreign accented than are individuals who have been exposed to fewer varieties of English. We did not anticipate the listener effect obtained here, and so did not take the precaution of obtaining detailed information pertaining to each listener's language history. Thus further research will be needed to determine the basis for differences between listeners in foreign accent perception.

Whatever the explanation, the results obtained here demonstrated that foreign accents may be present in the L2 speech of individuals who began learning their L2 in childhood (Long, 1990). The earliest of the estimates of the AOL of onset of foreign accents, 3.1 yr, may be the most valid of the estimates obtained. The speech samples evaluated here consisted of five short English sentences. These sentences were not intentionally loaded with vowels, consonants, and consonant clusters known to be difficult for native Italian speakers. Longer and/or more difficult speech samples might have yielded a more consistent identification of the first NI subgroup as foreign accented (Osberger et al., 1993). A similar outcome might have been obtained had the sentences been spoken spontaneously rather than elicited using a delayed repetition technique. Finally, the first NI subgroup might have been identified as foreign accented by more listeners had sentences spoken by just the first three Italian subgroups been presented along with those of the NE subjects. (See Flege and Fletcher, 1992, for a discussion of range effects.)

The results did not support the claim (Patkowski, 1990) that degree of perceived foreign accent increases precipitously after a critical period for speech learning has been passed at the age of 15 yr. However, the results support the view that, after a certain age, very few if any individuals will manage to speak their L2 without a trace of foreign accent (Scovel, 1988). We determined how many of the 24 subjects in the NE group and in the ten NI subgroups received a foreign accent rating that fell more than 2.0 s.d.s below the mean rating obtained for the NE group. Virtually none of the NE subjects were incorrectly classified as "foreign accented" using this statistical criterion, whereas virtually all of the NI subjects who began learning English after the age of 15 yr were so classified. It is important to recognize, how-

ever, that the statistical criterion used here did not provide direct evidence of foreign accent detection. The most direct and sensitive method for doing so is a paired comparison task (see Flege, 1984) but, unfortunately, such a task requires an extremely large number of trials.

Another aim of the study was to identify factors that might influence degree of perceived foreign accent. Female subjects who began learning English as children were found to pronounce English somewhat better than did males matched for AOL, whereas males who began learning English in late adolescence pronounced English somewhat better than did their female cohorts. The male and female subjects' responses to a language background questionnaire were submitted to separate principal components analyses. The factors yielded by these analyses were submitted to multiple-regression analyses. Somewhat more variance in the foreign accent ratings obtained for female than male subjects was accounted for (81% vs 75%). A factor designated age of learning accounted for the largest amount of variance for both genders. Language use factors accounted for a total of 15% of variance for both genders, but the nature of those factors varied somewhat.

A factor designated length of residence (LOR) in Canada accounted for a very small, albeit significant, amount of variance in the foreign accent ratings. As mentioned in Sec. I, LOR was confounded with the NI subjects' age of arrival in Canada. The correlation between the NI subjects' arrival age and the foreign accent ratings was far stronger when variations in LOR were partialled out ($r = -0.837$, $p < 0.01$) than was the correlation between LOR and foreign accent ratings when variations in *arrival age* were partialled out ($r = -0.193$, $p < 0.01$). Still, given that all of the NI subjects had lived in Canada for at least 15 yr, even a small LOR effect is of interest. Apparently, L2 pronunciation does not "fossilize" completely within the first few years of L2 learning (Selinker, 1972).

The present study focused on variations in degree of perceived foreign accent without attempting to establish *which dimensions* caused listeners to hear a foreign accent. However, a recent study by Flege and Hillenbrand (mss in preparation) suggested that the foreign accent was cued by segmental, subsegmental, and prosodic divergences from the phonetic norms of English. That study focused on a subset of the sentences ("The red book was good," "I can read this for you," "Paul at carrots and peas") and of the subjects examined here. A disproportionately later number of segmental errors were transcribed in sentences spoken by 18 NI males (those who had the strongest foreign accents in the present study) than in sentences spoken by nine NE males (88 vs 4 errors). Many NI subjects devoiced word-final consonants (in "peas" and "good") and realized /ð/ as /d/ (in "the" and "this"). A few NI subjects made errors affecting syllable structure (omitting the final /d/ in "good" or the final /s/ in "carrots," inserting a schwa-like vowel following "red"). In other instances, consonant errors and vowel quality errors were heard that could not be characterized as segmental *substitutions* (see also Munro et al., 1995). Two forms of sentences spoken by the 27 subjects were randomly presented to NE listeners for foreign accent ratings. The unprocessed sen-

tences were identical to those examined here. The processed version of the sentences preserved only amplitude and F0 variations. As expected, the unprocessed version of sentences spoken by the NE subjects received much higher ratings than did the NI subjects' sentences. A smaller but nevertheless significant difference between the NE and NI subjects was also observed for the *processed* sentences. This indicated that prosodic dimensions in the NI subjects' production of English sentences were sufficient to cue foreign accent (see also Willems, 1982).

The second-language Speech Learning Model developed by Flege (1995) attempts to account for segmental aspects of foreign accent. The inaccurate production of position-sensitive L2 vowel and consonant allophones (or "sounds") is seen to arise from an age-related decline in L2 learner's recognition that certain auditorily detectable differences between L1 and L2 sounds are *phonetically relevant*. This, it is claimed, impedes the formation of new phonetic categories. When a category cannot be established for an L2 sound, the model predicts limits on the accuracy with which the L2 sound can be produced. Also, the L1 sound to which it is linked perceptually is predicted to gradually resemble the L2 sound in production, as has been shown for L1 stop consonants (Flege, 1987b). The model claims that new phonetic categories can be established for L2 sounds under certain circumstances. By hypothesis, the likelihood of category formation varies directly as a function of the perceived dissimilarity of an L2 sound from the closest L1 sound, and inversely as a function of AOL. Establishment of categories for an L2 sound does not guarantee that an L2 sound will be produced exactly as by monolingual native speakers of the L2, however. The model postulates that L1 and L2 sounds exist in a common phonological space. By hypothesis, the phonetic categories for neighboring L1 and L2 sounds may be deflected away from one another so that phonetic contrast might be maintained both within *and across* languages (see also Watson, 1991). Also, the use of new features in categories that are formed for L2 sounds may be age limited.

The model just described, if supported, would provide an account for the positive correlation between AOL and strength of perceived foreign accent, as well as the presence of foreign accents in the speech of individuals who began learning their L2 in childhood. It also predicts that L2 learning will adversely affect bilinguals' pronunciation of their L1. When categories are added for L2 sounds, their presence in phonological space may cause neighboring L1 categories to be "deflected away" so as to maintain phonetic contrast in a common L1–L2 phonological space. In instances where category formation is blocked by equivalence classification, the prediction is that neighboring L1 and L2 sounds will come to resemble one another. It is hypothesized that, in such instances, a single category is used to process a diverse range of L1 and L2 phones. The extent to which L2 phonetic norms are approximated reflects the overall balance of L1 and L2 phones identified in terms of that single category. Thus the more accurately the L2 phones are pronounced (from a monolingual perspective), the *less* accurately the corresponding L1 phones will be pronounced (from the perspective of monolingual speakers of the L1).

One incidental finding of the present study was that NI subjects' self-estimates of their ability to pronounce English and Italian were inversely related. Subjects who reported pronouncing English well said they pronounced Italian poorly, and vice versa (Fig. 1). Such a pattern might be the result of the assimilatory and dissimilatory effects predicted by the Speech Learning Model (Flege, 1995). We hasten to add, of course, that the data shown in Fig. 1 were *self-estimates* which will need to be verified experimentally. However, Yeni-Komshian and Flege (1994) obtained preliminary ratings by listeners of words and sentences spoken in their two languages by Korean–English bilinguals. The pattern of ratings corresponded to the self-ratings just mentioned. Also, a mutual influence of the L1 and L2 on one another has been observed for processing in the semantic domain (e.g., Lambert and Rawlings, 1969; Obler and Albert, 1978; Magiste, 1979; Mack, 1986), in the syntactic domain (e.g., Blair and Harris, 1981), and in the phonological domain (e.g., Altenberg and Cairns, 1983; Cutler et al., 1989).

In summary, native Italian subjects' age of learning (AOL) English was found to exert a powerful influence on how well they pronounced English sentences even though they had all spoken English for many years. Some native English listeners were able to detect a foreign accent in sentences spoken by subjects who had begun learning English as young children. Virtually no Italian subjects who began learning English after the age of 15 yr obtained sentence ratings that fell within the native English range. Several factors in addition to AOL were found to exert effects on degree of perceived foreign accent: gender, length of residence in Canada, and especially the subjects' relative frequency of use of English and Italian.

ACKNOWLEDGMENTS

[1]This *post hoc* test assumes that the first level of a factor is a comparison group (here, the NE subjects), and that all remaining levels of the factor are ordered with respect to some logical parameter (here, AOL). By making use of such ordered information, Williams' test is more powerful than other *post hoc* tests in finding the first level of a factor whose mean differs significantly from that of the comparison group (Williams, 1971).

[2]Although significant two-way interactions were obtained for certain questionnaire items, they provided contradictory evidence. For example, early-arriving male NI subjects estimated using Italian more at home than did female NI subjects, whereas late-arriving males estimated using Italian at home *less* than did late-arriving females. Just the opposite pattern of results was obtained for the items that dealt with the frequency of use of Italian in *social settings*, however.

Altenberg, E., and Cairns, H. (1983). "The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects," J. Verb. Learn. Verb. Behav. 22, 174–188.

Asher, J., and Garcia, R. (1969). "The optimal age to learn a foreign language," Mod. Lang. J. 53, 334–341.

Bachi, R. (1956). "Statistical analysis of the revival of Hebrew in Israel," Scr. Hierosolymitana 3, 179–247.

Bahrick, H., Hall, L., Goggin, J., Bahrick, L., and Berger, S. (1994). "Fifty years of language maintenance and language dominance in bilingual Hispanic immigrants," J. Exp. Psychol. Gen 123, 264–283.

Bialystok, E., and Hakuta, K. (1994). In Other Words, The Science and Psychology of Second-Language Acquisition (Basic Books, New York).

Blair, D., and Harris, R. (1981). "A test of inter lingual interaction in comprehension by bilinguals," J. Psycholinguist. Res. 10, 457–467.

Cutler, A., Mehler, J., Norris, D., and Segui, J. (1989). "Limits on bilingualism," Nature 340, 229–230.

Flege, J. (1984). "The detection of French accent by American listeners," J. Acoust. Soc. Am. 76, 692–707.

Flege, J. (1987a). "A critical period for learning to pronounce foreign languages?," Appl. Linguist. 8, 162–177.

Flege, J. (1987b). "The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification," J. Phon. 15, 47–65.

Flege, J. (1988a). "The production and perception of speech sounds in a foreign language," in Human Communication and Its Disorders, A Review—1988, edited by H. Winitz (Ablex, Norwood, NJ), pp. 224–401.

Flege, J. (1988b). "Factors affecting degree of perceived foreign accent in English sentences," J. Acoust. Soc. Am. 84, 70–79.

Flege, J. (1992a). "The intelligibility of English vowels spoken by British and Dutch talkers," in Intelligibility in Speech Disorders: Theory, Measurement, and Management, edited by R. Kent (John Benjamins, Amsterdam), pp. 157–232.

Flege, J. (1992b). "Speech learning in a second language," in Phonological Development, Models, Research, and Applications, edited by C. Ferguson, L. Menn, and C. Stoel-Gammon (York, Parkton, MD), pp. 565–604.

Flege, J. (1995). "Second-language speech learning: Findings, and problems," in Speech Perception and Linguistic Experience: Theoretical and Methodological Issues, edited by W. Strange (York, Timonium, MD, in press).

Flege, J., and Eefting, W. (1987). "Cross-language switching in stop consonant perception and production by Dutch speakers of English," Speech Commun. 6, 185–202.

Flege, J., and Fletcher, K. (1992). "Talker and listener effects on the perception of degree of foreign accent," J. Acoust. Soc. Am. 91, 370–389.

Flege, J., Munro, M., and MacKay, I. (1995). "The effect of age of second language learning on the production of English consonants," Speech Commun. 16, 1–26.

Gould, J., and Marler, P. (1989). "Learning by instinct," Sci. Am. 255, 74–85.

Gras, M. (1983). "Integrative attitude and difficulties second language learners encounter in comprehending formal speech," Ph.D. dissertation, University of Texas.

Lambert, W., and Rawlings, C. (1969). "Bilingual processing of mixed-language associative networks," J. Verb. Learn. Verb. Behav. 8, 604–609.

Lamendella, J. (1977). "General principles of neurofunctional organization and their manifestation in primary and non-primary language acquisition," Lang. Learn. 27, 155–196.

Lenneberg, E. (1967). Biological Foundations of Language (Wiley, New York).

Long, M. (1990). "Maturational constraints on language development," Stud. Sec. Lang. Acquis. 12, 251–285.

Mack, M. (1986). "A study of semantic and syntactic processing in monolinguals and fluent early bilinguals," J. Psycholinguist. Res. 15, 463–488.

Magiste, E. (1979). "The competing language systems of the multilingual: A developmental study of decoding and encoding processes," J. Verb. Learn. Verb. Behav. 18, 79–89.

McLaughlin, B. (1977). "Second-language learning in children," Psychol. Bull. 84, 438–459.

Metz., D., Samar, V., Schiavetti, N., Sitler, R., and Whitehead, R. (1985). "Acoustic dimensions of hearing-impaired speakers' intelligibility," J. Speech Hear. Res. 28, 345–355.

Munro, M., Flege, J., and MacKay, I. (1995). "The effect of age of second language learning on the production of English vowels," Appl. Psycholinguist. (in press).

Neville, H., Mills, D., and Lawson, D. (1992). "Fractionating language: Different neural subsystems with different sensitive periods," Cerebr. Cortex 2, 244–258.

Obler, L., and Albert, M. (1978). "A monitor system for bilingual language processing," in Aspects of Bilingualism, edited by M. Paradis (Hornbeam Columbia, SC), pp. 156–164.

Osberger, M., Maso, M., and Sam, L. (1993). "Speech intelligibility of children with cochlear implants," J. Speech Hear. Res. 36, 186–203.

Oyama, S. (1976). "The sensitive period for the acquisition of a nonnative phonological system," J. Psycholinguist. Res. 5, 261–285.

Paradis, M. (1995). "Neurolinguistic aspects of 'native speaker,'" in The Native Speaker, edited by R. Singh (Sage, Beverly Hills, CA, in press).

Patkowski, M. (1990). "Age and accent in a second language: A reply to James Emil Flege," Appl. Linguist. 11, 73–89.

Penfield, W., and Roberts, L. (1959). Speech and Brain Mechanisms (Princeton U.P., Princeton, NJ).

Purcell, E., and Suter, R. (1980). "Predictors of pronunciation accuracy: A reexamination," Lang. Learn. 30, 271–287.

Rochet, B. (1995). "Perception and production of L2 speech sounds by adults," in Speech Perception and Linguistic Experience: Theoretical and Methodological Issues, edited by W. Strange (York, Timonium, MD, in press).

Sapon, S. (1952). "An application of psychological theory to pronunciation problems in second language learning," Mod. Lang. J. 36, 111–114.

Scovel, T. (1981). "The recognition of foreign accents in English and its implications for psycholinguistic theories of language acquisition," in Proceedings of the 5th Congress of the AILA (University of Laval, Laval, Canada), pp. 389–401.

Scovel, T. (1988). A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech (Newbury House/Harper & Row, New York).

Seliger, H., Krashen, S., and Ladefoged, P. (1975). "Maturational constraints in the acquisition of second language accent," Lang. Sci. 36, 20–22.

Selinker, L. (1972). "Interlanguage," Int. Rev. Appl. Linguist. 10, 209–231.

Suter, R. (1976). "Predictors of pronunciation accuracy in second language learning," Lang. Learn. 26, 233–253.

Tahta, S., Wood, M., and Lowenthal, K. (1981). "Foreign accents: Factors relating to transfer of accent from the first language to the second language," Lang. Speech 24, 265–272.

Thompson, I. (1991). "Foreign accents revisited: The English pronunciation of Russian immigrants," Lang. Learn. 4, 177–204.

Watson, I. (1991). "Phonological processing in two languages," in Language Processing in Bilingual Children, edited by E. Bialystok (Cambridge U.P., Cambridge), pp. 25–48.

Willems, N. (1982). English Intonation from a Dutch Point of View (Foris, Dordrecht).

Williams, D. A. (1971). "A test for differences between treatment means when several dose levels are compared with a zero dose control," Biometrics 27, 103–117.

Yeni-Komshian, G., and Flege, J. E. (1994). "Accuracy and fluency of pronunciation in L2 and L1 as a function of age of L2 acquisition," Fourth International Congress of Psycholinguistics as a Multidisciplinary Connected Science, University of Bologna, June 1994 (unpublished).

3134   J. Acoust. Soc. Am., Vol. 97, No. 5, Pt. 1, May 1995

Flege et al.: Perceived foreign accent   3134