Factors affecting degree of perceived foreign accent in English sentences

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

(Received 14 July 1987; accepted for publication 16 March 1988)

This study used interval scaling to assess degree of perceived foreign accent in English sentences spoken by native and non-native talkers. Native English listeners gave significantly higher (i.e., more authentic) pronunciation scores to native speakers of English than to Chinese adults who began learning English at an average age of 7.6 years. The results for the "child learners" suggest that a sensitive period for speech learning is reached long before the age of 12 years, as commonly supposed. Adults who had lived in the U.S. for 5 years did not receive higher scores than those who had lived there for only 1 year, suggesting that amount of unaided second-language (L2) experience does not affect adults' L2 pronunciation beyond an initial rapid stage of learning. Native speakers of Chinese who rated the sentences for foreign accent showed the same pattern of between-group differences as the native English listeners. The more experienced of two groups of Chinese listeners differentiated native and non-native talkers to a significantly greater extent than a less experienced group, even though the subjects in both groups spoke English with equally strong foreign accents. This suggests that tacit knowledge of how L2 sentences "ought" to sound increases more rapidly than the ability to produce those sentences.

PACS numbers: 43.71.Es, 43.71.Hw

INTRODUCTION

Those who learn a second language (L2) are often perceived to speak it with a foreign accent. Degree of perceived foreign accent increases along with the number, and perhaps severity, of segmental misarticulations (Gatbonton, 1975; Ryan *et al.*, 1975; Major, 1987; Flege and Eefting, 1987b). It may also be influenced by divergences from L2 phonetic norms for implementation of stress and emphasis, and divergences from L2 rhythmic and intonational patterns (Varonis and Gass, 1982; Willems, 1982; Fokes and Bond, 1984; Bond and Fokes, 1985). Foreign accents are often associated with low intelligibility and negative personal evaluations of non-native speakers (Flege, 1987b).

It is not known how much experience listeners need in order to recognize specific foreign accents, but even inexperienced listeners can detect small departures from the phonetic norms of their native language (L1). For example, Flege (1984) found that native English listeners could identify non-native speakers at above-chance rates in a paired-comparison task that involved presenting release bursts edited from /t/ tokens spoken by native speakers of English and French.

This study used interval scaling to assess the degree of perceived foreign accent in English sentences spoken by four groups of native Chinese speakers. Previous studies have shown that degree of accent can be estimated validly and reliably by trained and untrained listeners (Brennan *et al.*, 1975; Ryan *et al.*, 1975; Brennan and Brennan, 1981; Varonis and Gass, 1982; Major, 1987). The study aimed to determine: (1) to what extent L2 pronunciation is influenced by the age at which L2 learning begins; (2) whether amount of L2 experience influences adults' L2 pronunciation; (3) whether non-natives who themselves speak L2 with a foreign accent can gauge degree of foreign accent accurately; and (4) whether removing pauses makes L2 sentences sound less foreign accented.

A. Age of learning

It is widely believed that a "critical" period exists for human speech learning. Sapon (1952) thought that a foreign accent in adults' speech stems from a "loss of flexibility" of the speech organs. According to Lamendella (1977), the "immature neurolinguistic system" of children facilitates their learning of L2. Lenneberg (1967) believed that adults "inevitably" speak L2 with an accent if learning begins after a neurologically based critical period. Similarly, Scovel (1969) claimed that the "nature of the human brain," not environmental experience, determines how L2 is pronounced.

If a critical period for human speech learning exists, L2 learning should be ultimately less successful if it begins after this period has been passed (Flege, 1987a). The finding of numerous studies that L2 pronunciation becomes progressively less authentic as the age of the onset of L2 learning increases is consistent with a critical period hypothesis (e.g., Asher and Garcia, 1969; Fathman, 1975; Patkowski, 1980; Oyama, 1982).

Some researchers (e.g., Lenneberg, 1967; Scovel, 1969, 1988) have inferred that the critical period is reached at about the age of 12 years. Individuals who begin learning L2 before this age would, therefore, be expected to speak L2 without an accent, provided, of course, that they had sufficient experience speaking and hearing L2. There is, nevertheless, evidence that adults who began learning L2 as chil-

 0001-4966/88/070070-10$00.80

dren ("child learners"), and even children themselves, may speak with a measurable foreign accent (Asher and Garcia, 1969; Tahta *et al.*, 1981; Oyama, 1982; Thompson, 1984).

It is possible, however, that listeners' identification of child learners and children as non-native in previous studies was sometimes due to false alarms, not misses. In the Asher and Garcia (1969) study, for example, 23% of native English children were incorrectly identified as non-native. The present study, therefore, tested the hypothesis that child learners will speak L2 without a foreign accent. Sentences produced by native English adults were compared to sentences produced by Chinese adults who began learning English at an average age of 7.6 years. If a critical period for speech learning occurs at about the age of 12 years, their pronunciation scores should not differ from those of English native speakers. If a critical period occurs before the age of 12 years (or does not exist), on the other hand, the child learners might receive significantly lower (i.e., less authentic) scores.

B. Length of residence

The critical period hypothesis leads to the expectation that the amount of L2 experience will have little effect on L2 pronunciation for individuals who begin learning L2 after the critical period has been passed. In keeping with this, the L2 pronunciation of adults is often said to "fossilize" (Selinker, 1972), that is, be resistant to further change after an initial period of rapid improvement (Scovel, 1988). However, several studies examining the effect of length of residence have shown that L2 pronunciation improves with L2 experience. The dependent variable in these studies is "length of residence," usually quantified as the number of years spent in a place where L2 is the predominant language. It is generally assumed that the amount of L2 input increases linearly with years in an L2-speaking environment.[1]

It has been shown that degree of foreign accent, or the frequency with which non-native talkers are identified as such, decreases as length of residence increases (Asher and Garcia, 1969; Purcell and Suter, 1980; Snow and Hoefnagel-Höhle, 1982). A closer examination suggests, however, that the effect of length of residence may be confined to individuals who begin learning L2 before the age of 12 years. Suter (1976) and Purcell and Suter (1980) found equally strong simple correlations between age of learning and degree of accent as between length of residence and degree of accent. (Length of residence and age of learning were inversely correlated, as in many other studies.) However, while length of residence accounted for a significant amount of variance in the pronunciation scores in a multiple regression analysis, age of learning did not.

Thompson (1984), on the other hand, found that age of learning but not length of residence accounted for a significant amount of variance in a multiple regression analysis. The length of residence variable was probably significant in the Suter (1976) but not the Thompson (1984) study because most of Suter's subjects began learning L2 after the age of 12 years, while many of Thompson's subjects began learning L2 as young children. Results obtained by Oyama (1982) support this interpretation. English listeners judged degree of perceived foreign accent in the English spoken by subjects who arrived in the U.S. between the ages of 6 and 20 years. Oyama found that a significant correlation existed between age of arrival and degree of foreign accent when the confounding effect of variations in length of residence was partialed out, but not the reverse.

The second aim of this study was, therefore, to test the hypothesis that, beyond an initial stage of rapid improvement, amount of L2 experience does not affect how well adult learners pronounce L2. Foreign accent was evaluated in sentences spoken by two groups of adults who had lived for about 1 and 5 years in the U.S. Since both groups began learning English L2 as adults, the critical period hypothesis would be supported if no difference in pronunciation scores between the two groups was observed.

C. Non-native's ability to gauge foreign accent

Scovel (1988) described a series of unpublished experiments examining listeners' ability to detect foreign accent (see also Neufeld, 1979, 1980). The English sentences examined were spoken by ten native speakers of American English, eight non-natives whose English pronunciation was "excellent," and two native speakers of non-American English (Irish, South African). The task was to identify which samples were spoken by someone "not American." As shown in Fig. 1, native English children's ability to detect non-native speakers, and to avoid calling natives foreign, did not reach adultlike levels until about the age of 9 years. It is not surprising that non-native adults performed the task less well than native English adults. The rate at which they detected foreign accent correctly was correlated with their degree of proficiency in L2. As shown in Fig. 1, percent correct detection increased slightly across the three proficiency subgroups (elementary, intermediate, advanced) of non-native subjects examined.

Flege (1984, 1987b) hypothesized that native speakers develop detailed phonetic category prototypes against which to judge the goodness of phones (see also Johansson, 1978;

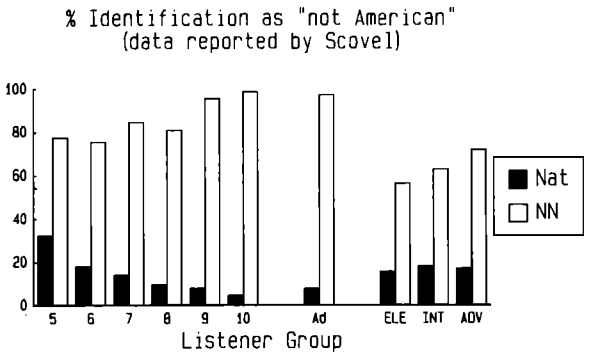

FIG. 1. The results of experiments by Scovel (1988) examining the ability of native English children aged 5–10 years of age, native English adults (Ad), and non-native adults differing in English-language proficiency (ELE, INT, ADV) to identify speech samples produced by native speakers (Nat) and by individuals who were not native (NN) speakers of American English. Based on Scovel (1988).

Samuel, 1982). Once a phone has been identified as being the realization of a phonetic category, its distance in the phonetic space from the category's center is gauged. If a sentence contains a phone judged to fall outside a "tolerance region" centered on the prototype, it might be heard to be foreign accented.[2]

Scovel's (1988) results suggest that the ability to gauge degree of foreign accent in English sentences is a skill that develops slowly with English-language experience. Perhaps adults who learn an L2 become better able to detect foreign accent—and to gauge its strength—by establishing prototypes for phones in L2. This was tested in the present study by examining the pronunciation scores given to native and non-native talkers by groups of native Chinese speakers who had lived in the U.S. for about 1 and 5 years. If the prototype hypothesis is correct, the experienced Chinese listeners should distinguish the native and non-native speakers to a greater extent than relatively inexperienced Chinese listeners.

D. The effect of pauses

The oral interview test used by the Foreign Service Institute to assess foreign language proficiency makes use of five variables: grammar, vocabulary, comprehension, accent, and fluency. "Accent" is likely to be related to details of segmental articulation, intonation, and rhythm. These are all dimensions along which an L2 learner's performance might be compared to L2 phonetic "norms." The term "fluent" is often used to describe L2 production (e.g., Elman *et al.*, 1977). It is uncertain upon what acoustic dimensions the perception of degree of fluency depends, but the dimensions are likely to include the number, location, and duration of pauses, prolongations, and repetitions in sentences.

Non-native speakers seem to produce sentences containing more, and perhaps longer, pauses than native speakers (James, 1988). Removing pauses might make sentences spoken by non-natives sound more fluent, and thus lead to higher global foreign accent scores. This hypothesis assumes that, as implied by the nominal components of the Foreign Service test, accent and fluency represent separate (or separable) perceptual dimensions. It may be, however, that the perceived degree of fluency does not contribute to foreign accent judgments, or that accent and fluency judgments interact in some complex fashion so that removing pauses from sentences spoken by non-native speakers would not improve the foreign accent scores accorded to their sentences. This was tested in the present study by presenting sentences twice to listeners, once with the original pauses present and once with all auditorily detectable pauses removed.

I. Methods

A. Talkers

Table I summarizes several characteristics of the five groups of talkers who produced the test sentences that were rated for foreign accent. There were three groups of Chinese adults (MA, T1, T2) who began learning English L2 in adulthood. The subjects in MA and T1 were differentiated primarily by native-language background. Whereas the subjects in MA were native speakers of Mandarin from Beijing, those in T1 were native speakers of Taiwanese from Taiwan. The average length of residence in the U.S. was 1.1 years for the subjects in both groups. It was not known whether the few differences in the segmental inventories of Mandarin and Taiwanese (Maddieson, 1984) would lead the subjects in MA and T1 to pronounce English differently. Group MA was included primarily to permit a test of the ability of non-native listeners to gauge degree of foreign accent, as explained below.

TABLE I. Characteristics of the talkers in five groups. The subjects in T1, T2, and CL were native speakers of Taiwanese differentiated according to age of learning and length of residence in the U.S. (see text); those in MA were native speakers of Mandarin who were matched to the talkers in T1; those in EN were native speakers of English. The mean values are number of years; standard deviations are in parentheses.

	Talker group				
	MA	T1	T2	CL	EN
Gender	5M/5F	9M/1F	7M/3F	4M/3F	7M/3F
Birthplace	Beijing	Taiwan	Taiwan	Taiwan[a]	U.S.
Mean age	35.0(5.2)	28.4(1.6)	30.5(2.4)	19.4(1.7)	25.6(4.1)
Years in U.S.	1.1(0.9)	1.1(0.7)	5.1(2.7)	12.0(2.6)	25.6(4.1)
Age of arrival	34.0(6.1)	27.6(2.0)	25.4(2.3)	7.6(3.4)	···

[a] One talker was born in Hong Kong.

The Taiwanese subjects in group T2 were differentiated from the subjects in T1 primarily according to English-language experience. They had lived in the U.S. for 5.1 years at the time of the study. The "child learners" in group CL were native Taiwanese speakers who differed from the subjects in T1 and T2 according to the age of L2 learning. Whereas the subjects in groups T1 and T2 were first massively exposed to native-produced English upon their arrival in the U.S. sometime after the age of 20 years, those in CL arrived in the U.S. at an average age of 7.6 years. The subjects in CL had also lived much longer in the U.S. (12 years on average) than the subjects in T1 and T2.

The native speakers of American English in group EN were monolinguals who were roughly matched for gender and age to talkers in the Chinese groups. Like the Chinese talkers, they were all affiliated with the University of Alabama at Birmingham. There were ten talkers in all groups except CL, which included seven talkers.

B. Stimuli

The talkers each read the following English sentences (along with two others) five times each from a randomized list:

The good shoe fits Sue.
I can read this for you.
The red book was good.

These sentences contained vowels (/æ/, /ə/, /ɪ/, /ʊ/ and consonants (/ð/, /ʃ/, /ɹ/) not found in Mandarin or Taiwanese, and word-final obstruents (/d/, /k/, /s/, /z/,

/ts/) which do not occur in Chinese. They were used in a previous study examining degree of foreign accent in English sentences spoken by native speakers of Dutch (Flege and Eefting, 1987a). The sentences, which will be referred to as the "Sue," "read," and "book" sentences, were recorded (Sony TCD5M) in a sound booth with a microphone (Nakamichi CM300) placed about 6 in. from the mouth.

The third token of the "Sue," "read," and "book" sentences spoken by each talker was low-pass filtered at 8 kHz before being digitized at 20 kHz with 12-bit resolution. Copies were made of each digitized sentence and pauses were edited from the copies whenever the author and an assistant both heard a pause, and when the perceived pause was visually evident in a display of rms amplitude. Table II presents the mean number and total duration of pauses removed from sentences spoken by talkers in the five groups. Pauses were removed from 62% of the "Sue" sentences, 43% of the "read" sentences, and 70% of the "book" sentences. Most pauses removed were shorter than 200 ms (see footnote 3). Slightly more pauses, and pauses of longer duration, were removed from sentences spoken by the talkers in groups T1 and T2 than in group MA. This last group, in turn, had more and longer pauses than groups CL and EN.

The intensities of the sentences were normalized by determining the peak intensity in each sentence to the nearest 1.0 dB(A). Sentences with relatively low intensities were multiplied by a weighting function greater than 1.0, and those with relatively high intensity by a weighting function less than 1.0. A one-way ANOVA showed that the peak intensity differences between the sentences in the six blocks (three sentences $\times$ two editing conditions) were nonsignificant after normalization.

C. Listeners

Three groups of listeners rated the sentences for foreign accent. The listeners differed in native language and/or English-language experience. The listeners in group T1-L (eight males, one female) were drawn from talker group T1. They had a mean age of 28 years (s.d. = 2) and had lived in the U.S. for an average of 1.5 years (s.d. = 9). Eight listeners in group T2-L were drawn from talker group T2. These listeners (six males, three females) had a mean age of 30 years (s.d. = 3) and had lived in the U.S. for an average of 5.3 years (s.d. = 30). The native English listeners in group EN-L (five males, four females) had a mean age of 27 years (s.d. = 4). Two of these listeners had previously participated as talkers in group EN.

The listeners in all three groups were students or staff members at the University of Alabama at Birmingham. It is not known what effect, if any, allowing listeners to evaluate their own sentences might have had.[4]

TABLE II. The mean number (M) and total duration (TD) in ms of pauses removed from sentences spoken by talkers in five groups.

Talker group	Sentence "Sue" M	"Sue" TD	"read" M	"read" TD	"book" M	"book" TD	Average M	Average TD
MA	0.9	207	0.4	27	0.8	139	0.7	124
T1	1.5	316	1.0	88	2.6	267	1.7	224
T2	2.7	398	1.1	163	1.9	287	1.9	282
CL	0.4	53	0.1	11	0.7	78	0.4	47
EN	0.2	36	0.0	0	0.1	19	0.1	18

D. Procedures

The listeners were told they would hear sentences spoken by an unspecified proportion of native and non-native talkers. They were told to estimate the degree of foreign accent in each sentence by moving a lever on a response box over a 10-cm range. The range was defined by the labels "no foreign accent" (at the top of the range), "medium foreign accent" (at the middle), and "strong foreign accent" (at the bottom). The lever activated a potentiometer connected to an 8-bit A/D converter. Sentences judged to have been produced with the greatest authenticity could receive a rating of 256. Those produced with the least authenticity (i.e., with the "strongest" possible foreign accent) could receive a rating as low as 1.

The sentences were presented binaurally (TDH-49) at a peak syllable intensity of 73 dB (A) ($\pm$ 3 dB). The subjects were told to use the whole range. They pressed a button after positioning the lever at a position they deemed appropriate. Each sentence was presented 1.0 s after a response was received for the preceding sentence. The "Sue," "read," and "book" sentences were presented in separate blocks, the order of which was counterbalanced across listeners. The order of condition (unedited versus edited) was counterbalanced within the three sentence types. Each of the three sentences $\times$ two editing conditions = six blocks lasted about 10 min.

Each separately randomized block contained three randomizations of 47 sentences. Responses to the first randomizations, which were presented to familiarize listeners with the range of foreign accents they would hear, were not analyzed.

E. Analysis

The mean pronunciation scores given to the 47 talkers by the listeners in three groups were calculated for the edited and unedited versions of the "Sue," "read," and "book" sentences. Each mean was based on 18 judgments (nine listeners $\times$ two presentations). These talker-based scores were submitted in an ANOVA in which talker group and listener group served as between-subjects factors, and sentence and editing condition served as within-subjects factors. An alpha level of 0.01 was used to test main effects and simple main effects. An alpha level of 0.05 was used for *post hoc* tests (Newman–Keuls).

II. Results

A. Differences between talker groups

Removing pauses had little effect on the pronunciation scores (see below), and there was little overall difference in the scores obtained for the "Sue," "read," and "book" sentences (120, 136, 130). The mean scores shown in Fig. 2

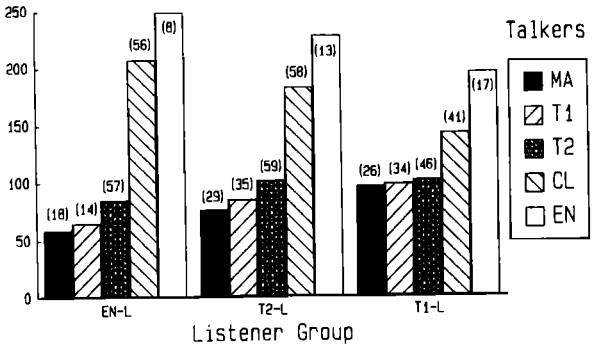

FIG. 2. The mean pronunciation scores given by listeners in three groups to English sentences spoken by native speakers of Mandarin (MA), native speakers of Taiwanese who learned English as adults (T1, T2) or as children (CL), and native speakers of English (EN). The sentences judged to be most authentic could receive a score of 256, while those judged to have the strongest foreign accent could receive a score of 1. Intertalker standard deviations are in parentheses.

have, therefore, been averaged across the three sentences and two editing conditions. Note that the higher the score, the more authentic (hence less foreign-accented) the sentences were judged to sound.

Not surprisingly, the native English talkers (EN) received scores near the top of the 256-point scale (224), indicating a high degree of perceived authenticity. The child learners in group CL received somewhat lower scores than the native English speakers (177), but considerably higher scores than the adult learners in groups MA, T1, and T2 (77, 83, and 96, respectively).

The ANOVA yielded a significant talker group $\times$ sentence interaction [$F(8,252) = 8.50$]. Tests of simple main effects indicated that the group effect was significant for all three sentence types [$F(4,136) \geqslant 58.5$ in each instance]. The *post hoc* tests revealed that, for all three sentences, higher scores were obtained for the native English talkers (EN) than for all four Chinese groups (CL, MA, T1, and T2), and for the child learners (CL) than for the adult learners (MA, T1, and T2). The interaction arose because the group MA talkers received significantly lower scores than the talkers T1 and T2 for just the "Sue" sentences (53 vs 76 and 89).[5]

B. Differences between listener groups

Inspection of Fig. 2 reveals differences in how the three groups of listeners rated sentences spoken by native and nonnative talkers. The extent to which listeners differentiated the proficient speakers of English (groups CL and EN) from the nonproficient speakers of English (groups MA, T1, and T2) varied as a function of English-language experience. The average scores given to the proficient talkers decreased from group EN-L (228) to T2-L (205) to T1-L (168). That is, the listeners who were least experienced in English (T1-L) judged sentences spoken by the proficient talkers to be less authentic than the listeners who were most experienced in English (EN-L). Conversely, the average scores given to the nonproficient talkers increased from listener group EN-L (69) to T2-L (88) to T1-L (98). This means that the least experienced listeners (viz., group T1-L) judged sentences produced by the nonproficient talkers to be less accented than those who were most experienced in English (viz., group EN-L).

These differences lead to a significant talker group $\times$ listener group interaction [$F(8,126) = 4.10$]. Tests of simple main effects indicated that the talker group effect was significant for listener groups EN-L, T2-L, and T1-L [$F(4,42) \geqslant 14.7$ in each instance]. The *post hoc* tests revealed that the listeners in all three groups gave the native English speakers (EN) significantly higher scores than the talkers in all four Chinese groups (CL, MA, T1, and T2), and the child learners (CL) higher scores than all three adult learner groups (MA, T1, and T2).

The simple main effect of listener group was significant only for the talkers in groups MA and EN [$F(2,27) \geqslant 5.73$ in both instances]. The *post hoc* tests revealed that talkers in MA received significantly higher scores from the listeners in group T1-L (96) than EN-L (59). The talkers in EN received significantly *lower* scores from the listeners in T1-L (195) than from the listeners in T2-L and EN-L (229 and 249, respectively).

One reason for caution in concluding that listeners' ability to gauge degree of foreign accent in English increases with English-language experience is that some of the listeners had previously served as talkers (see Sec. I A). Both children learning L1 and adults learning L2 seem at times to be unaware of their own—but not others—pronunciation errors. Another reason for caution is that the mean values submitted to the ANOVA were the average scores given to each talker by all nine listeners in each listener group (Sec. I E). This makes it risky to generalize the results to populations of *listeners*.

Accordingly, a second analysis was performed. The mean scores given to the ten talkers each in groups MA and EN by each listener was computed. This yielded six averaged scores (three sentences $\times$ two talker groups) for the nine listeners each in groups T1-L and T2-L. Scores were computed only for the seven listeners in EN-L who had *not* served as talkers in group EN. The 150 scores that resulted were submitted to an ANOVA in which listener group and talker group were between-subjects factors, and sentence was a within-subjects factor.

Figure 3 shows the mean values given by the listeners in the three groups to the talkers in groups MA and EN. The scores given to the non-natives (MA) increased from listener group EN-L (60) to T2-L (78) to T1-L (98). This means that the Mandarin talkers sounded more authentic to the listeners with the least experience in English (T1-L) than to those with the most (EN-L). The scores given to the native speakers (EN) *decreased* from listener group EN-L (248) to T2-L (230) to T1-L (182). The native speakers sounded *less* authentic to listeners who were inexperienced in English than to relatively experienced listeners.

These differences lead to a significant listener group $\times$ talker group interaction [$F(2,44) = 15.5$]. Tests of simple main effects indicated that the effect of listener group was significant for talker groups MA and EN [$F(2,22) \geqslant 5.62$ in both instances]. The *post hoc* tests revealed that the talkers

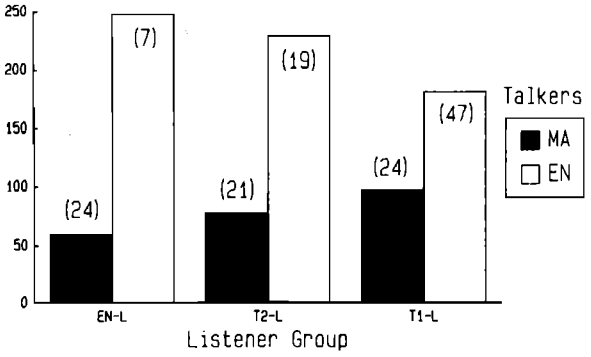

FIG. 3. The mean pronunciation scores given to sentences spoken by native speakers of Mandarin (MA) and English (EN) by native speakers of English (EN-L), native speakers of Taiwanese who had lived for an average of 5.3 years in the U.S. (T2-L), and native speakers of Taiwanese who had lived for 1.5 years in the U.S. (T1-L). Interlistener standard deviations are in parentheses.

in MA received significantly higher scores from the listeners in T1-L than EN-L, and the talkers in EN received significantly *lower* scores from the listeners in T1-L than those in EN-L or T2-L.

One final test examining differences between the listener groups was performed. The difference in scores given by each of the 25 listeners to the talkers in groups EN and MA was calculated and submitted to a one-way randomized block ANOVA. The effect of listener group on the magnitude of the difference in scores for native and non-native talkers was highly significant [$F(2,22) = 15.9$]. The *post hoc* tests indicated that the listeners in EN-L differentiated the talkers in MA and EN significantly more than the listeners in T2-L and T1-L (189 vs 152 and 84), and the listeners in T2-L showed significantly greater difference scores than those in T1-L.

Each analysis supported the hypothesis that the ability to gauge foreign accent in English sentences increases with English-language experience. It is possible, however, that the observed differences between listener groups resulted from some kind of response bias rather than from a genuine perceptual difference. This would be true, for example, if the listeners in TL-1 did not follow the instruction to use the whole 256-point scale to the same extent as listeners in the other two groups. Perhaps the listeners in T1-L used a smaller part of the scale (one centered near the middle of the range) because they lacked confidence in their ability to gauge foreign accent in English sentences, not because they were less able to determine how authentically the English sentences had been produced.

The "lack of confidence" hypothesis was tested by noting the highest and lowest mean scores accorded the unedited "Sue" sentences by each of the 27 listeners. If the listeners in T1-L made relatively less use of the entire 256-point scale, then the difference between the high and low scores should be smaller for listeners in T1-L than in T2-L and EN-L. A one-way ANOVA showed that listeners in groups EN-L, T1-L, and T2-L did not follow the instruction to use the whole scale differentially. The difference between groups EN-L, T1-L, and T2-L (255, 237, and 245, respectively) approached significance [$F(2,24) = 3.46$, $p = 0.047$], but *post hoc* tests showed that no between-group difference was significant at the 0.05 level.

C. Effect of removing pauses

Removing pauses had little effect on the pronunciation scores. The small overall difference in scores for the edited and unedited sentences (128 vs 129) was nonsignificant [$F(1,126) = 1.51$]. As expected, the edited sentences produced by the talkers in T1 and T2 received slightly higher (that is, more authentic) scores than the unedited versions of the same sentences (85 vs 81, 97 vs 95). For talkers in MA, CL, and EN, on the other hand, the edited sentences received slightly *lower* (more foreign-accented) scores (77 vs 78, 174 vs 180, 223 vs 226). The resulting significant condition $\times$ talker group interaction [$F(4,126) = 5.55$] probably derived from the fact that more and longer pauses were removed from sentences spoken by the talkers in groups T1 and T2 than by talkers in the other three groups. Tests of simple main effects revealed, however, that the effect of condition was nonsignificant for the talkers in all five groups.

As expected, the listeners in EN-L gave slightly higher (more authentic) scores to the edited than unedited sentences (130 vs 127). The native Chinese listeners in T1-L and T2-L, on the other hand, gave slightly *lower* scores to the edited sentences (125 vs 126, 130 vs 133), which lead to a significant condition $\times$ listener group interaction [$F(2,126) = 4.98$]. Tests of simple main effects revealed that the effect of condition was significant for EN-L but not T2-L or T1-L [$F(1,46) = 6.74$, 4.07, and 0.91, respectively].[6]

Some sentences presented in the "edited" condition did not actually have pauses removed, while others had only very short pauses removed (see Table II). To further test the effect of removing pauses, the 15 "Sue" and 13 "book" sentences from which at least 200 ms of silence was removed were examined.[7] The scores given to the edited and unedited versions of these sentences were submitted to separate listener group $\times$ condition ANOVAs. These tests indicated that removing pauses did not affect listeners' foreign accent judgments. The difference between the edited and unedited "Sue" sentences (68 vs 65), and between the edited and unedited "book" sentences (81 vs 75), were nonsignificant. Neither the effect of listener group, nor the listener group $\times$ condition interaction, reached significance in either analysis.

One last test of the effect of removing pauses was performed. Of the sentences spoken by the native speakers of Chinese, 28 "Sue," 20 "read," and 32 "book" sentences actually had pauses removed. The average difference in scores given by the native English listeners (EN-L) to the unedited and edited versions of these sentences was calculated. Separate multiple regression analyses were performed to determine if the number, average duration, and total duration of the pauses removed were related to the edited–unedited difference scores. The independent variables did not predict a significant amount of the variance in the pronunciation scores for either the "Sue," "read," or "book" sentences.

Finally, presentation in the "edited" condition of 18 "Sue," 27 "read," and 14 "book" sentences from which no pauses were actually removed provided a way to test the reliability of the procedure used for estimating degree of perceived foreign accent. If the procedure is reliable, then the listeners should have given the same scores to these sentences in both conditions since they were, in fact, judging identical stimuli. One-way ANOVAs showed that the scores obtained in the edited and unedited conditions for the "Sue" (181 vs 174), "read" (180 vs 179), and "book" (208 vs 207) sentences were not significantly different.

III. Discussion

A. Differences between talker groups

Two hypotheses were generated from the assumption that a critical period exists for human speech learning. The hypothesis that adults who begin learning L2 before the age of 12 years will speak L2 without a measurable foreign accent was not supported. English sentences spoken by native Taiwanese adults who began learning English at an average age of 7.6 years received significantly lower (that is, more accented) pronunciation scores than sentences spoken by native English adults, even though these "child learners" had lived for an average of 12.0 years in the U.S. This finding agrees with previous studies showing that individuals who begin learning L2 as children, as well as children themselves, often speak L2 with a detectable foreign accent (Asher and Garcia, 1969; Tahta *et al.*, 1981; Oyama, 1982).

The second hypothesis tested was that, after a rapid initial phase of learning, length of residence does not affect the L2 pronunciation of individuals who began learning L2 as adults. Length of residence was quantified as the number of years spent in the U.S. It was assumed that length of residence would be closely related to the amount of phonetic input received from English native speakers. The hypothesis received support from the finding that the pronunciation scores received by groups of native speakers of Taiwanese who had lived in the U.S. for an average of 5.1 and 1.1 years did not differ significantly.

An effect of experience might have been noted if the talkers in group T2 had had more English-language experience, or experience of a qualitatively different nature from the less experienced talkers in group T1. However, the finding that groups T1 and T2 did not differ is consistent with the widespread belief that the L2 pronunciation of adults ceases to improve ("fossilizes") at a relatively early stage of L2 learning (Selinker, 1972). It also agrees with previous empirical studies that, when considered together, suggest that length of residence is a significant predictor of L2 pronunciation success only for individuals who begin learning L2 before about the age of 12 years (Asher and Garcia, 1969; Suter, 1976; Purcell and Suter, 1980; Oyama, 1982; Snow and Hoefnagel-Höhle, 1982; Thompson, 1984).

The present results seem to diverge from those obtained by Flege and Eefting (1987b) using the same protocol and sentences as in the present study. Their study examined English sentences spoken by two groups of Dutch university students thought to differ principally according to English-language experience. The subjects in both groups began learning English in school in the Netherlands at the age of 12 years. The subjects in one group, who were majoring in English, continued to study English after high school. The subjects in the other group, who were majoring in engineering, did not study English after high school and had much less need and opportunity to speak English than the English majors. As in the present study, the more experienced of two groups (the English majors) received significantly lower pronunciation scores than native speakers of English. However, native English listeners gave the students of English significantly higher pronunciation scores than the engineering students (178 vs 86).

Why was an effect of L2 experience noted by Flege and Eefting (1987b) but not in the present study? Perhaps the apparent divergence was due to the fact that, whereas the two Taiwanese groups examined in the present study were differentiated only by the amount of L2 input from native speakers, the two groups of subjects examined by Flege and Eefting (1987a) may have differed in other ways also. The talkers in the Taiwanese groups T1 and T2 began to study English as an academic subject in junior high school. They did not arrive in the U.S. until after the age of 20 years, and were all graduate students or professors at the University of Alabama at Birmingham (mostly in computer science or engineering). The subjects in both groups used English during the workday, but continued to speak Chinese regularly.

Perhaps the English majors examined by Flege and Eefting (1987a) had more aptitude and/or greater motivation for L2 learning than the engineering majors in that study. *Prima facie* support for this hypothesis comes from the fact that the English majors were specializing in a foreign language at the university level. The training in English phonetics that the Dutch students of English but not engineering received as part of their post-secondary education may also have been important.

Taken together with previous findings (e.g., Flege and Hillenbrand, 1984), the present results suggest an upper limit on how well L2 can be pronounced. It is not necessary to conclude, however, that the basic ability (or abilities) that permits children to learn L1 without accent—viz., the ability to translate the sensory input that accompanies articulation into gestures—is lost or even attenuated after a critical period is passed. As noted by Flege (1987b), many factors in addition to age and possible neurological organization differentiate young children who learn L2 from adult learners of L2. These include the nature of L2 input received by children and adults, social and psychological factors, and differing communicative needs. It is also possible that L2 learners may succeed in producing some but not all L2 phones authentically (Flege, 1987c).

The notion that a critical period for human speech learning stems from neurological maturation or reorganization (Lenneberg, 1967) implies an all-or-none phenomenon with well-defined temporal boundaries. However, the critical period for human speech learning differs from behaviors such as imprinting in ducks, where the behavior may be acquired (or at least stimulated) only during a brief and well-defined interval. This is clearly not the case for human

speech learning. For example, the adult learners in the present study had all succeeded to some extent in learning English pronunciation. It therefore seems more reasonable to speak of a "sensitive" period for human speech learning (Oyama, 1979).

Scott's (1978) review of behavioral development in a wide range of species suggested a number of broad characteristics of sensitive periods. First, organizational processes are modified most easily at the time they are proceeding most rapidly. Second, behavioral development is cumulative in the sense that, as more behaviors are added, they are "integrated into specialized systems and subsystems" that may interfere with the acquisition of subsequent new behaviors (p. 3). Third, change becomes progressively more difficult as organizational processes become more stable.

In keeping with this general outline, Flege (1987a–c; Flege and Eefting, 1987a–c) proposed that the seeming upper limit on the learning of L2 pronunciation derives from the development of phonetic systems that underly the production and perception of speech. Phones in L2 that do not have a direct L1 counterpart (called "new" sounds) may be produced authentically by learners of all ages if sufficient native-speaker input is received. However, after about the age of 5 years, it may become impossible for learners to produce "similar" L2 sounds (i.e., phones that differ physically from L1 counterparts) authentically. It is at about this age that awareness of phonetic segments begins to emerge, and the number and nature of (L1) phonetic categories become defined. Equivalence classification limits learning for "similar" sounds. This basic mechanism may lead L2 learners to regard an L2 phone that differs physically from phones in L1 as being the realization of an already-established (L1) phonetic category.[8]

B. Listeners' ability to gauge foreign accent

Two groups of native Taiwanese speakers who judged English sentences for degree of foreign accent showed the same pattern of differences between the five groups of talkers as native speakers of English. This agrees with Neufeld's (1979) finding that native English subjects who spoke French with an accent were able to detect an English accent in French passages. The Taiwanese subjects who had lived for an average of 5.3 years in the U.S. were able to distinguish native and non-native speakers better than those who had lived in the U.S. for 1.5 years. This agrees with the finding by Neufeld (1980) that proficient English-speaking listeners were better able than relatively nonproficient ones to detect an English accent in French.

The results obtained here support the hypothesis that, as the native speakers of Taiwanese became more experienced in English, they gained more accurate information concerning how the phonetic segments in English sentences "ought" to sound (Linell, 1982). This interpretation assumes, of course, that an important mechanism in making global foreign accent judgments as well as detecting foreign accent (Flege, 1984) is an evaluation of the extent to which the acoustic properties of particular phonetic segments diverge from internalized phonetic norms. An alternative hypothesis is that ability to gauge foreign accent improved because the listeners in T2-L were better able to note the presence of L1 phones in the English sentences than those in group T1-L (Neufeld, 1979).

Most Taiwanese subjects in listener T1-L and T2-L had served as talkers in group T1 or T2. As mentioned earlier, the sentences produced by the talkers in T1 and T2 were judged to be equally accented. The present results therefore suggest that the ability of non-natives to perceive the phonetic properties of L2 sentences accurately exceeds their ability to produce those sentences authentically. This conclusion is consistent with the results of studies showing that the identification of L2 phones may be more nativelike than is the production of those phones (e.g., Carramazza *et al.*, 1973; Flege and Eefting, 1987b). The finding is reminiscent of studies examining vocal learning in birds, which have shown that different critical periods may exist for motor and auditory learning (Nottebohm, 1969).

C. Effect of removing pauses

All audible pauses were removed from sentences and the edited sentences were presented in separate blocks. This was done to help determine if an assessment of fluency affects global foreign accent judgments. Removing pauses might have influenced the foreign accent judgments of the non-native listeners to a greater extent than the native English listeners. This should have occurred if foreign accent judgments are based primarily on the detection of divergences from internalized norms for segments (and prosodic dimensions) in L2, and if the L2 learners had not internalized such auditory-based norms (or did so to a lesser extent than native speakers). It seems reasonable to assume that ability to estimate fluency will not depend on L2 experience (insofar as it is cued by the detection of pauses, repetitions, and prolongations).

Contrary to expectation, the native but not the non-native listeners showed a significant effect of pause removal. It should be noted, however, that the effect of removing pauses was minuscule (3 points on a 256-point scale) even for the native English listeners. This finding suggests either that fluency judgments do not influence degree of perceived foreign accent, or that fluency cannot be perceived independently from the segmental and suprasegmental dimensions which determine accent. The present results do not allow us to choose between these two possible explanations.

One additional difficulty concerns the status of pauses in sentence production. One might reasonably regard pauses as an integral part of the sentences in which they are found. If so, the removal of pauses in the present study might have had a generally salutary effect on degree of perceived accent, but it may nevertheless have lead to a countervailing *decrease* in the perceived goodness of individual syllables or phonetic segments. For example, placing syllables that occurred originally in a prepausal position in nonprepausal positions might make them seem too long. Similarly, the sentences might have seemed less well intonated after pauses were removed. This is because the voiced portions of the sentence, which have fundamental frequency contours, would be closer together.

It would be useful to determine in future research whether listeners can differentiate groups of native and non-native talkers solely on the basis of fluency. Fluency and accent judgments may interact with one another, and with other dimensions such as grammar and word choice (Varonis and Gass, 1982). Perhaps removing pauses would have had a significant effect in the present study had the non-native speakers' segmental articulation been more authentic, had more or longer pauses been removed, or both.

IV. SUMMARY

This study examined factors that might influence global foreign accent scores accorded English sentences spoken by several groups of native Chinese adults. The first aim was to determine to what extent L2 pronunciation is influenced by the age at which L2 learning begins. Adults who began learning English L2 at an average age of 7.6 years received significantly higher pronunciation scores than adults who began learning English L2 as adults, but significantly lower scores than native speakers of English. This suggested that the early learning of L2 represents an important advantage, although it should be noted that the "child learners" had lived for a considerably longer time in the U.S. than the "adult learners." The finding demonstrated that, if a sensitive period for human speech learning exists, it occurs well before the age of 12 years.

The second aim was to determine whether the amount of L2 experience influences adult learners' L2 pronunciation. The sentences spoken by two groups of Chinese subjects who had lived in the U.S. for about 1 and 5 years did not differ significantly. This suggested that, beyond an initial stage of rapid learning, additional unaided experience with an L2 does not produce dramatic results. The finding did not mean necessarily that specific training would be futile, or that the seeming limitation on the improvement of pronunciation extends to all phones and phonetic dimensions.

The third aim was to determine whether non-native speakers who themselves speak with a foreign accent could gauge degree of foreign accent accurately. A group of Chinese subjects who had lived in the U.S. for about 5 years were better able to gauge degree of foreign accent in English sentences spoken by other native speakers of Chinese than a group who had lived in the U.S. for about 1 year only. The more experienced Chinese listener group had been shown not to differ from the less experienced group in ability to pronounce English. The finding therefore suggests that amount of L2 experience has a greater effect on ability to perceive L2 accurately than to produce it authentically. Perhaps the refinement of internal auditory perceptual representations for phones and phonetic dimensions in L2 continues over longer periods than do motor speech abilities.

Finally, the study provided a preliminary attempt to determine whether an assessment of fluency affects global foreign accent judgments. Somewhat surprisingly, removing pauses from the sentences spoken by the Chinese subjects did not result in an appreciable increase in foreign accent scores. This suggested that foreign accent judgments depend only (or mainly) on segmental and suprasegmental articulation, that "fluency" is not affected importantly by the presence of pauses, or that the perception of disfluency might play a role only for sentences without obvious segmental or suprasegmental errors.

ACKNOWLEDGMENTS

This study was funded by NIH Grant NS20963. The author thanks Elizabeth James for contributing to all phases of the research, and Ocke Bohn and Winifred Strange for comments on a previous draft of the article.

[1]This is unlikely to be true, of course, for all L2 learners. Personality factors and various social exigencies may influence the amount and quality of contact that learners have with native speakers of L2.

[2]The tolerance region for segmental or suprasegmental norms may vary as a function of factors such as the social context in which the phone was spoken (Carranza and Ryan, 1985), the listener's attitude (Brennan and Brennan, 1981), or the listener's and/or talker's social background (Ryan and Sebastian, 1980).

[3]It was uncertain what part of the silent interval was associated with stop closure and what part was the pause when an unreleased stop abutted a perceived pause. In these instances, all but a portion of the silent interval appropriate for stop closure was removed. A 64-ms silent interval was left for /d/, 80 ms for /b/, /g/, and /k/, and 125 ms for the /d#b/ cluster in "red book." These durations were based on values observed in sentences spoken without pauses by the same or other talkers.

[4]Some listeners said after the experiment that they had heard their own voice, but others did not. In other experiments of this type, we have found that listeners may erroneously report hearing their own voice. The listeners were not told the language background of the non-native talkers whose speech they were to assess. Some said they heard only Chinese accents, while others said they heard other kinds of foreign accents, especially Japanese-accented English.

[5]Tests of simple main effects indicated that the effect of sentence was significant only for groups MA and T1 [$F(2,58) \geqslant 3.64$ in both instances]. *Post hoc* tests revealed that group T1's "book" sentences received significantly higher scores than their "Sue" sentences (92 vs 76). Group MA's "read" sentences received higher scores than their "book" sentences which, in turn, received higher scores than their "Sue" sentences (104 vs 75 vs 53).

[6]Slightly higher scores were obtained for the edited than unedited versions of the "Sue" sentences (121 vs 119). Slightly lower scores, on the other hand, were obtained for the "read" and "book" sentences (135 vs 137; 129 vs 131), which lead to a significant condition×sentence interaction [$F(2,252) = 3.62$]. There were no obvious differences in terms of the number, average duration, or total duration of pauses removed from the three sentences that might explain this interaction.

[7]Five "Sue" sentences meeting this criterion were spoken by talkers in T1, six by talkers in T2, and four by talkers in MA. Of the "book" sentences, five came from group T1, five from T2, and three from MA. The "read" sentences were not examined because only five met the 200-ms criterion.

[8]Individuals who begin learning L2 by about the age of 5–6 years but not those who learn L2 as adults may form new phonetic category for similar L2 sounds. This hypothesis was based on the observation that L2 learning has different effects on the production of L1 stop consonants for subjects who learn L2 before as opposed to after the age of 5 years (Flege, 1987c; Flege and Eefting, 1987a). Early L2 learners showed dissimilation of corresponding L1 and L2 stops, while adult learners showed *assimilation*. The dissimilation suggested the formation of a new phonetic category for the L2 stops (see, also, Flege and Eefting, 1987c). The inferred age of the sensitive period is also consistent with Thompson's (1984) finding that adults who began learning L2 between the ages of 5–10 years, but not those who learned L2 before the age of 5 years, were detected as non-native.

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

Bond, Z., and Fokes, J. (**1985**). "The timing of English words by non-native speakers," J. Acoust. Soc. Am. Suppl. 1 **77**, S53.

Brennan, E., and Brennan, J. (**1981**). "Accent scaling and language attitudes: Reactions to Mexican–American English speech," Lang. Speech **24**, 207–221.

Brennan, E., Ryan, E., and Dawson, W. (**1975**). "Scaling of apparent accentedness by magnitude estimation and sensory modality matching," J. Psychol. Res. **4**, 27–36.

Carramaza, A., Yeni-Komshian, G., Zurif, E., and Carbone, E. (**1973**). "The acquisition of a new phonological contrast: The case of stop consonants in French–English bilinguals," J. Acoust. Soc. Am. **54**, 421–428.

Carranza, M., and Ryan, E. (**1975**). "Evaluative reactions of bilingual Anglo and Mexican American adolescents towards speakers of English and Spanish," Int. J. Sociol. Lang. **6**, 83–104.

Elman, J., Diehl, R., and Buchwald, E. (**1977**). "Perceptual switching in bilinguals," J. Acoust. Soc. Am. **62**, 971–974.

Fathman, A. (**1975**). "The relationship between age and second language productive ability," Lang. Learn. **25**, 245–253.

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

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

Flege, J. E. (**1987b**). "The production and perception of speech sounds in a foreign language," in *Human Communication and Its Disorders, Vol. III*, edited by H. Winitz (Ablex, Norwood, NJ).

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

Flege, J. E., and Eefting, W. (**1987a**) "The production and perception of English stops by Spanish speakers of English," J. Phonet. **15**, 67–83.

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

Flege, J. E., and Eefting, W. (**1987c**). "Imitation of a VOT continuum by native speakers of Spanish and English: Evidence for phonetic category formation," J. Acoust. Soc. Am. **83**, 729–740.

Flege, J. E., and Hillenbrand, J. (**1984**). "Limits on pronunciation accuracy in adult foreign language speech production," J. Acoust. Soc. Am. **76**, 708–721.

Fokes, J., and Bond, Z. (**1984**). "Patterns of English word stress by native and non-native speakers," in *Proceedings of the Tenth International Congress of Phonetic Sciences*, edited by M. Van den Broecke and A. Cohen (Dordrecht, Foris), pp. 682–686.

Gatbonton, E. (1975). "Systematic variations in second language speech: A sociolinguistic study," unpublished Ph.D. thesis, McGill University.

James, A. (**1988**). "Phonological structure and the contextual variability of interlanguage segments," in *Further Insights into Contrastive Linguistics*, edited by J. Fisiak (Benjamins, Amsterdam).

Johansson, S. (**1978**). "Studies in Error Gravity," Gothenburg Stud. Eng. **44** (Univ. of Gothenburg, Sweden).

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).

Linell, P. (**1982**). "The concept of phonological form and the activities of speech production and speech perception," J. Phonet. **10**, 37–72.

Maddieson, I. (**1984**). *Patterns of Sound* (Cambridge U.P., Cambridge, England).

Major, R. (**1987**). "English voiceless stop production by speakers of Brazilian Portuguese," J. Phonet. **15**, 197–202.

Neufeld, G. (**1979**). "Towards a theory of language learning ability," Lang. Learn. **29**, 227–241.

Neufeld, G. (**1980**). "On the adult's ability to acquire phonology," TESOL Q. **14**, 285–298.

Nottebohm, F. (**1969**). "The 'critical period' for song learning," Ibis. **111**, 386–387.

Oyama, S. (**1979**). "The concept of sensitive period in developmental studies," Merrill-Palmer Q. **25**, 83–102.

Oyama, S. (**1982**). "The sensitive period for the acquisition of a nonnative phonological system," in *Child–Adult Differences in Second Language Acquisition*, edited by R. Scarcella and M. Long (Newbury House, Rowley, MA), pp. 20–38.

Patkowski, M. (**1980**). "The sensitive period for the acquisition of syntax in a second language," Lang. Learn. **30**, 449–471.

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

Ryan, E., Carranza, M., and Moffie, R. (**1975**). "Reactions toward varying degrees of accentedness in the speech of Spanish–English bilinguals," Lang. Speech **20**, 267–273.

Ryan, E., and Sebastian, R. (**1980**). "The effects of speech style and social class background on social judgments of speakers," B. J. Soc. Clin. Psychol. **19**, 229–233.

Samuel, A. (**1982**). "Phonetic prototypes," Percept. Psychophys. **31**, 307–314.

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

Scott, J. P. (**1978**). *Critical Periods* (Dowden, Hutchinson & Ross, Stroudsburg, PA).

Scovel, T. (**1969**). "Foreign accents, language acquisition, and cerebral dominance," Lang. Learn. **19**, 245–253.

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

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

Snow, C., and Hoefnagel-Höhle, M. (**1982**). "Age differences in the pronunciation of foreign sounds," in *Child-Adult Differences in Second Language Acquisition*, edited by S. Krashen, R. Scarcella, and M. Long (Newbury House, Rowley, MA), pp. 84–92.

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. (**1984**). "Experimental study of foreign accents," unpublished Ph.D. Thesis, George Washington University.

Varonis, E., and Gass, S. (**1982**). "The comprehensibility of non-native speech," Studies Sec. Lang. Acquis. **4**, 114–136.

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