Journal of Memory and Language **41,** 78–104 (1999)
Article ID jmla.1999.2638, available online at http://www.idealibrary.com on IDEAL®

Age Constraints on Second-Language Acquisition

James Emil Flege

University of Alabama at Birmingham

Grace H. Yeni-Komshian

University of Maryland

and

Serena Liu

University of Alabama at Birmingham

This study evaluated the critical period hypothesis for second language (L2) acquisition. The participants were 240 native speakers of Korean who differed according to age of arrival (AOA) in the United States (1 to 23 years), but were all experienced in English (mean length of residence = 15 years). The native Korean participants' pronunciation of English was evaluated by having listeners rate their sentences for overall degree of foreign accent; knowledge of English morphosyntax was evaluated using a 144-item grammaticality judgment test. As AOA increased, the foreign accents grew stronger, and the grammaticality judgment test scores decreased steadily. However, unlike the case for the foreign accent ratings, the effect of AOA on the grammaticality judgment test scores became nonsignificant when variables confounded with AOA were controlled. This suggested that the observed decrease in morphosyntax scores was not the result of passing a maturationally defined critical period. Additional analyses showed that the score for sentences testing knowledge of rule based, generalizable aspects of English morphosyntax varied as a function of how much education the Korean participants had received in the United States. The scores for sentences testing lexically based aspects of English morphosyntax, on the other hand, depended on how much the Koreans used English.

Key Words: second language acquisition; phonology; morphosyntax; age; critical period; language use; education.

Many studies examining second language (L2) acquisition have focused on the influence of age. The age variable examined in L2 studies is usually the age of first exposure to the target L2. In studies examining immigrant populations, this is typically indexed by the participants' age of arrival (AOA) in the host country. Previous research has suggested that AOA is apparently an important determinant of overall degree of foreign accent in the L2 (Flege, Munro, & MacKay, 1995a), as well as degree of accuracy in producing particular L2 consonants and vowels (Flege, Munro, & MacKay, 1995b; Munro, Flege, & MacKay, 1996). Age effects have also been reported for the learning of English morphosyntax (Johnson & Newport, 1989).[1] The observation of age effects on the L2

This research was funded by NIDCD Grant DC02892. The authors are indebted to Jackie Johnson, for providing the grammaticality judgment test sentences, and to David Birdsong, for first suggesting the analysis of rule versus lexically based grammaticality judgment test items. The are grateful to E. Bialystok, D. Birdsong, H. Hakuta, M. Mack, T. Piske, and L. White for commenting on earlier drafts of this article, as well as the following University of Maryland students: S. Kim and R. Byun (for help collecting data), J. Kim and E. Kim (for their help preparing stimuli and testing listeners), and M. Cullen, L. Yun, A. Kinn, C. Romero, and M. Robbins (for help with data management).

Editorial correspondence and reprint requests should be addressed to J. E. Flege, Department of Rehabilitation Sciences, University of Alabama at Birmingham, VH 503, Birmingham, Alabama 35294. E-mail: jeflege@uab.edu.

[1] Strong age effects have been also been obtained in

0749-596X/99 $30.00

performance of adults—even those who are experienced in their L2—has suggested to some researchers that the ability to acquire an L2 effectively is limited by a critical period. Such a conclusion is important practically, inasmuch as it might influence decisions regarding educational policy. It is also important theoretically.

The critical period hypothesis rests on the assumption that the age-related effects seen in L2 studies are the result of maturational changes in brain structures that are used to learn and/or to process language. For example, it has been hypothesized that as the brain matures, it becomes less "plastic" and that lost neural plasticity impedes L2 learning (e.g., Scovel, 1988; Patkowski, 1980, 1990). However, others have proposed that age-related changes in L2 performance derive from the nature and extent of the interaction between a bilingual's two language systems (e.g., Oyama, 1979; Flege, 1987, 1988, 1995, 1998b; Bialystok, 1997). This latter approach treats age as an index of the state of development of the L1 system. It assumes implicitly that, all else being equal, the more fully developed the L1 system is when L2 learning begins, the more strongly the L1 will influence the L2.

Choosing between maturational and interactive accounts of age-related effects on L2 performance is difficult. Neural development and native language acquisition are inextricably confounded through much of childhood (Bates & Goodman, 1998). Also, the most common index of age in L2 studies, AOA, is typically confounded with other variables that may influence L2 performance (see Flege, 1987 and 1998a, for discussions). Still another difficulty is that there is no consensus as to how one might test the critical period hypothesis. This is because the critical period hypothesis has been applied with less specificity to the study of age-related changes in L2 performance than it has been applied in ethological studies examining, for example, imprinting behavior in birds (Bornstein, 1989). The lack of specificity is most crucial with respect to a definition of the structures and/or functions that are putatively altered as the brain matures (Bialystok, 1997). For example, those who claim that foreign accents arise due to the passing of a critical period do not specify whether the age-related changes arise from a loss of ability to articulate L2 speech sounds, a loss of ability to auditorily distinguish L2 from L1 sounds or to form perceptual representations for L2 sounds in long-term memory, or a loss of ability to translate such representations into articulatory gestures (see Flege, 1987, 1988, 1995, for discussion).

The aim of this study was to evaluate the critical period hypothesis by examining the effect of AOA on L2 performance. Three methods were used to evaluate the critical period hypothesis. The first method will be called the "discontinuity test." The discontinuity test rests on the assumption that, in an AOA-stratified sample of L2 learners, participants who began learning the L2 before the critical period will perform markedly better than those who began learning their L2 after the end of the critical period. The effect of a critical period could, therefore, be demonstrated by showing a significant departure from linearity in functions relating measures of L2 performance to AOA at an appropriate AOA.

The present study was well suited for the discontinuity test. The 240 native speakers of Korean who participated had arrived in the United States between the ages of 1 and 23 years and had lived there for at least 8 years (*mean* = 15 years).[2] Previous work has shown that as AOA increases, native speakers of Korean make more errors writing down computer-generated English sentences (Bott, 1993), respond more slowly and less accurately to a lexical decision task (Kim, 1996), and make more errors on a grammaticality judgment test (Shim, 1995). Finally, Koreans who learn English in adulthood are known to have difficulty

studies which have examined the acquisition of American Sign Language (e.g., Newport, 1990; Mayberry & Eichen, 1991; Emmorey, Bellugi, Friederici, & Horn, 1995).

[2] The participants were selected on the basis of AOA rather than the age at which they first began to study English at school in Korea. This is because, of the two variables, AOA is the more potent predictor of L2 performance (e.g., Johnson & Newport, 1989).

in accurately producing and perceiving certain English vowels (Flege, Bohn, & Jang, 1997a).

There are, however, at least two drawbacks to using the discontinuity test as a means to evaluate the critical period hypothesis. First, there is disagreement as to when the critical period for L2 acquisition ends. According to some, it ends at 12 years of age (Scovel, 1988). But according to others (e.g., Patkowski, 1990), it ends at 15 years. Second, not everyone would agree that the absence of a discontinuity provides evidence that the critical period hypothesis is incorrect. Many investigators have used the term "sensitive period" and "critical period" interchangeably. The two notions appear to rest equally on the view that diminished L2 performance is the consequence of normal brain maturation (see Bialystok, 1997, and Birdsong, 1998). However, the notion of a sensitive period implies that there will be a gradual, perhaps even linear, decline in L2 performance as AOA increases. Thus, the lack of a discontinuity in AOA-L2 performance functions might be taken as evidence against the existence of a critical period for L2 acquisition. However, it would not disprove the existence of a sensitive period and so would not rule out a maturational account of age-related changes in L2 acquisition.

The second test of the critical period hypothesis employed here will be called the "pre/postcorrelation test." This test involves computing the correlation between AOA and L2 performance for groups of participants thought to have begun learning their L2 before versus after the end of a critical period. According to Johnson and Newport (1989), a significant AOA–performance correlation will be observed for individuals who began learning their L2 before the end of the critical period, because performance declines increasingly as one nears the end of the critical period. However, a significant AOA–performance correlation would not be expected for a group of individuals who had all begun learning their L2 at varying times after the end of the critical period. This is because "postmaturational" learners are all thought to suffer to the same degree from the same deficit, viz. having passed a critical period. The pre/postcorrelation test was implemented by computing a correlation between AOA and each of the outcome measures of the study for the native Korean participants who arrived in the United States before versus after the end of the putative critical period.

A potential problem for the pre/postcorrelation test was that a significant AOA-L2 performance correlation could be obtained for postmaturational learners due to factors unrelated to whether a critical period (should one exist) had been passed. As already mentioned, the Korean participants were selected according to AOA. However, as in previous studies examining large groups of immigrants (see, e.g., Bachi, 1956; Bahrick et al., 1994), the Koreans' AOAs were confounded with other variables that might influence L2 learning, L2 performance, or both. As will be discussed, the variables confounded with AOA in this study were chronological age, the native Koreans' average self-estimated use of English, their self-estimated use of Korean, years of residence in the United States, and the number of years of education they had received in the United States.

In view of these confounds, a third method was used here to assess the critical period hypothesis. The "matched subgroup" method allowed us to test the hypothesis that factors associated with AOA, rather than AOA itself, are responsible for what have been interpreted as "age" effects in previous studies of L2 acquisition. In one set of analyses, subgroups were formed that consisted of native Korean participants drawn from the original group of 240. These subgroups differed in AOA but were matched for variables confounded with AOA. If the AOA-defined subgroups differed significantly, it would demonstrate that age, and possibly age-related maturational changes, was responsible for the difference. However, if the AOA difference disappeared when the confounded variables were controlled, it would show that age was not responsible for the between-group differences and would thus fail to support the existence of a maturationally defined critical period.

As already mentioned, if a critical period does exist for L2 acquisition, it would be necessary to define the structure(s) and/or func-

tion(s) that are altered by maturation. One way to help identify such structures and functions is to determine if AOA effects differ across linguistic domains. For example, if it were shown that native-like performance were possible beyond a certain AOA for syntax but not phonology, this would imply that "phonology and syntax are represented independently, or the processing systems behind phonology and syntax are different" (Bialystok, 1997, p. 120).

There is, in fact, widespread agreement that age constrains the learning of L2 phonology and morphosyntax differently. Snow (1979) showed that phonology and morphosyntax emerged as separate factors in a study examining the naturalistic acquisition of Dutch by native English children and adults over a 1-year period. Some researchers have concluded that a critical period exists only for phonology (Scovel, 1988; Singleton, 1989; Bahrick, Hall, Goggin, Bahrick, & Berger, 1994). Others have concluded that a critical period ends sooner for phonology than for morphology or syntax (Long, 1990; Hurford, 1991). Still others have concluded that morphosyntax is learned more thoroughly, or over a longer period of time, than is phonology (Braine, 1971; MacWhinney, 1992; Snow, 1987).

Despite this convergence of views, there is little empirical evidence for a difference in the phonological and morphosyntactic domains for experienced adult speakers of an L2. Phonology and morphosyntax have been examined concurrently in just four previous studies (Fathman, 1975; Oyama, 1973; Patkowski, 1980, 1990; Snow & Hoefnagel-Höhle, 1982a,b). None of these studies examined an AOA-stratified sample of adults drawn from a single L1 background or assessed phonological and morphosyntactic performance in comparable detail. Fathman (1975) found that older children received higher morphosyntax test scores than did younger children, whereas the reverse held true for pronunciation (see also Olson & Samuels, 1973; Krashen, Long, & Scarcella, 1979; Engstrand, 1982). Snow and Hoefnagel-Höhle (1982a,b) found that adults and older children outperformed younger children in both linguistic domains when tested soon after their first exposure to the L2, whereas the younger participants outperformed older participants 1 year later.

Oyama (1973) observed a stronger correlation between AOA and degree of foreign accent than between AOA and the scores on a grammaticality judgment test ($r = .81$ vs. .41). This might be taken as support for the view that age constrains phonology to a greater extent than morphosyntax. However, Patkowski (1980, 1990) obtained equally strong correlations between AOA and measures of English morphosyntax and phonology ($r = -0.74, -0.76$). Moreover, Patkowski observed a sharp decline in performance in both domains at an AOA of 15 years. His finding for phonology agreed with the results of two previous studies (Seliger, Krashen, & Ladefoged, 1975; Tahta, Wood, & Lowenthal, 1981) but diverged from two others (Oyama, 1973; Flege et al., 1995a). The AOA–morphosyntax discontinuity observed by Patkowski agreed with the findings of Johnson and Newport (1989), who concluded that the acquisition of L2 morphosyntax is constrained by the completion of brain maturation at about the age of 15 years (but cf. Bialystok & Hakuta, 1994).

This study's comparison of performance in the phonological and morphosyntactic domains was motivated by the disparate results just reviewed. As advocated by Bialystok and Hakuta, (1998), we used "broadly based" measures of L2 proficiency rather than measures that were designed to test the predictions of a particular linguistic theory (e.g., predictions regarding subjacency or the complex noun phrase constraint). The learning of English morphosyntax was assessed using a 144-item grammaticality judgment test that Johnson and Newport (1989) devised to assess the "most basic aspects of English sentence structure." The learning of English phonology was assessed by having listeners rate a standard set of English sentences spoken by the native Korean participants for overall degree of foreign accent.

The grammaticality judgment test used here was designed to test nine different morphosyntactic structures or "rules" (Johnson & Newport, 1989) but, as discussed below, the sentences testing the nine rule types were

heterogeneous. Therefore, the effect of AOA on two new, functionally specified, sets of sentences was also examined. The sentences comprising both the "rule based" set (e.g., **The man paints his house yesterday*) and the "lexically based" set (e.g., **The farmers were hoping rain*) were drawn from several of the original sentence sets. The distinction between the two new functional sets reflects a distinction drawn in linguistic theory (e.g., Pinker, 1991; Pinker & Prince, 1992) and L2 acquisition research (e.g., Beck, 1997). Correct responses to the rule based sentences required the acquisition of simple rules with widespread application (example: "Add *-ed* to the verb root to form the past tense"). Correct responses to the lexically based sentences, on the other hand, probably required learning that might be characterized as "bottom-up" or "data-driven" or else learning based on the establishment of associative or probabilistic representations (see, e.g., Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1997; Seidenberg, 1997).

The results will be presented in six sections. We began by carrying out ANOVAs that tested for differences between AOA-defined subgroups of native Korean participants and the native English control group. The numbers of Koreans who received foreign accent ratings and morphosyntax scores that fell within the native English range were also determined. Analyses were carried out in the second section to determine if the AOA–performance functions were nonlinear and, if so, where in the AOA continuum the nonlinearity occurred. In the next sections we compared the scores obtained for the original nine sets of grammaticality judgment test sentences, for sentences that were grammatical and ungrammatical, and for the two new functional sets (rule based vs. lexically based, see above). The purpose of the multiple regression analyses presented in the next section was to account for variance in the outcome measures. Finally, a series of matched subgroup analyses was undertaken to test for between-group differences when variables confounded with AOA were controlled.

METHOD

Participants

The 240 native Korean participants arrived in the United States between the ages of 1 and 23 years. Their age at the time of testing ranged from 17 to 47 years (mean = 26). The 24 native English participants ranged in age from 20 to 45 years (mean = 27). All 264 participants passed a pure-tone hearing screening before participating. To be included, native Korean participants had to report speaking no language other than English and Korean and to have lived in the United States for at least 8 years (mean, 14.6 years).

As summarized in Table 1, the native Korean participants were assigned to 1 of 10 subgroups based on AOA. The average AOAs of these subgroups increased in roughly 2-year increments, from 3 years for group NK3 to 21 years for group NK21. Half of the participants in all 11 groups were female. On average, the highest academic grade completed in Korea by the native Korean participants was 5.6 years. The first exposure to English for most participants with AOAs greater than 12 years occurred at school in Korea. On the average, the Koreans had received 10.1 years of education in the United States. All but 1 Korean participant had completed high school in the United States; 156 participants held a bachelor's degree from an American university.

As will be discussed in greater detail below, a number of variables were correlated with AOA. The earlier the native Korean participants had arrived in the United States, the more education they had received in the United States. Also, the younger the native Korean participants were upon arriving in the United States, the longer they tended to have lived in the United States, the more they spoke English, and the less they spoke Korean.

General Procedure

The participants were tested individually by college-aged Korean/English bilingual research assistants in a single 1.5-h session. The research was carried out in a quiet room located either on

TABLE 1

Characteristics of the 12 Male and 12 Female Participants in Each of 11 Groups

	Age	AOA	EXP	LOR	EDUC	KORUSE	ENGUSE
NE	27 (7)						
NK3	23 (3)	3.0 (0.5)	4.5 (1.0)	20.0 (2.3)	15.6 (0.6)	2.3 (0.5)	4.6 (0.3)
NK5	21 (2)	5.0 (0.5)	5.3 (0.7)	16.4 (2.7)	14.8 (1.7)	2.5 (0.6)	4.5 (0.3)
NK7	24 (3)	7.0 (0.5)	7.0 (0.7)	16.9 (3.2)	15.4 (1.5)	2.6 (0.6)	4.3 (0.6)
NK9	24 (3)	9.0 (0.5)	9.0 (0.5)	15.0 (3.2)	13.5 (1.7)	2.9 (0.6)	4.1 (0.6)
NK11	24 (5)	11.0 (0.5)	11.0 (0.7)	13.5 (4.5)	11.1 (1.7)	3.1 (0.6)	4.0 (0.6)
NK13	24 (3)	13.0 (0.5)	12.8 (0.9)	11.7 (3.2)	9.2 (1.6)	3.7 (0.5)	3.6 (0.6)
NK15	27 (5)	15.0 (0.5)	13.2 (0.6)	12.5 (5.4)	7.8 (2.0)	3.5 (0.7)	3.6 (0.7)
NK17	29 (4)	17.0 (0.5)	13.5 (0.8)	12.5 (4.2)	5.8 (2.0)	3.7 (0.5)	3.4 (0.6)
NK19	32 (5)	19.0 (0.5)	13.3 (0.5)	13.7 (5.1)	4.8 (1.8)	3.7 (0.7)	3.5 (0.7)
NK21	34 (5)	21.5 (0.8)	13.7 (1.9)	13.5 (4.5)	2.9 (2.8)	3.8 (0.5)	3.5 (0.7)
	26 (6)	12.0 (5.9)	9.8 (3.6)	14.6 (4.6)	10.1 (4.8)	3.2 (0.8)	3.9 (0.7)

Note. Age, chronological age, in years; AOA, age of arrival in the United States; EXP, age of first exposure to English, either at school in Korea or upon arrival in the United States; LOR, length of residence in the United States, in years; EDUC, years of education in United States; KORUSE, the average of nine 5-point rating scale items pertaining to the use of Korean; ENGUSE, the average of seven similar items pertaining to the use of English. *SD*s are in parentheses.

the campus of the University of Maryland-College Park or in a nearby Korean church.

The participants began by completing a language background questionnaire that assessed their use of both English and Korean, as well as their motivation to learn English and retain Korean. The participants later produced sentences that were rated for foreign accent and responded to a grammaticality judgment test. Steps were taken to reduce processing differences between the online measure of L2 phonology and the offline measure of L2 morphosyntax. The sentences to be rated for foreign accent were repeated following an aural model, and the participants were required to listen to an aural presentation of each grammaticality judgment test sentence before judging its grammaticality.

Foreign accent. The participants repeated 21 English sentences that contained a wide variety of English vowels and consonants. The sentences were each presented twice in a row, over a loudspeaker, in the same order in which they appeared on a written list. A short tone was presented 700 ms after the first presentation of each sentence and 3600 ms after its second presentation. To reduce the likelihood of direct imitations, the participants were required to wait until hearing the tone before repeating each sentence. Just five of the sentences were examined here (*Ron set a thick rug in the sun; Joe will feed the pup who sat by you; You should thank Sam for the food; Fit a ring to the water tap; It is fun to play chess with a rook*), in most instances the second of the two repetitions.[3]

The sentences were digitized at 22.05 kHz, normalized for peak intensity, and later randomly presented three times each in separate, counterbalanced blocks to native English-speaking listeners. The three male and seven female listeners, who had a mean age of 31 years (range = 23–37), were living in Birmingham, Alabama, at the time of testing. All of the listeners had been born and raised in the Washington, DC–Baltimore area, however, and all of them passed a pure-tone hearing screening (at octave frequencies between 500 and 4000 Hz re: 25 dB HL) before participating.

Following practice with 22 sentences spanning a wide range of foreign accents, the listeners rated sentences spoken by the 24 native English and 240 native Korean participants using a scale that ranged from "very strong" for-

[3] A preliminary analysis revealed that the foreign accents in the first and second repetitions of sentences did not differ significantly.

TABLE 2

Number and Examples of the Nine GJT Sentence Types

N	Sentence type	Examples
8	Past tense	A policeman gave Alan a ticket for speeding yesterday. *A policeman gived Alan a ticket for speeding yesterday.
8	Plural	Todd has many coats in his closet. *Todd has many coat in his closet.
8	Third-person singular	Every Friday our neighbor washes her car. *Every Friday our neighbor wash her car.
8	Determiners	The boy is helping the man build a house. *The boy is helping the man build house.
8	Pronouns	Susan is making some cookies for us. *Susan is making some cookies for we.
6	Particle movement	Kevin called up Nancy for a date. *Kevin called Nancy for a date up.
14	Subcategorization	The little boys laughed at the clown. *The little boys laughed the clown.
4	Lexically specified subject/object raising	Larry believed himself to be brave. *Larry believed that himself to be brave.
4	Y/N questions	Should Timothy have gone to the party? *Should have Timothy gone to the party?
4	Wh questions	Where did she put the book? *Why did she put the book?

Note. N, the number of sentence pairs. In the list of examples, the ungrammatical member of each sentence pair is marked by an asterisk.

eign accent (1) to "no accent" (9). The listeners were told to use the whole scale and to guess if unsure. A mean foreign accent rating based on 150 judgments (5 sentences $\times$ 10 listeners $\times$ 3 replicate judgments) was calculated for each participant. Average ratings were examined here for two reasons. First, a similar AOA effect was evident for each of the 10 listeners (averaged over the five sentences) and for each of the five sentences (averaged over the 10 listeners). Second, very high Intraclass Correlations were obtained for the five sentences, $R = .986$, and the 10 listeners, $R = .978$, $P < .001$.

Morphosyntax. The native Korean participants' knowledge of English morphosyntax was assessed using a 144-item grammaticality judgment test. This test was composed of nine sets of sentences, each intended to evaluate a different morphosyntactic structure or rule. Table 2 gives an example of each set, as well as the number of pairs in each set. Most (128) of the sentences were drawn from Johnson and Newport's (1989) test. Three sets of sentences that caused very few errors in previous administrations of the test (present progressive, word order, and auxiliary) were eliminated. Eight of the 16 new sentences tested lexically specified subject/object raising. Half of the 144 sentences used here were grammatical. The other half were ungrammatical sentences created by eliminating a required morpheme or word from a grammatical sentence, by changing a word, or by moving some word(s) to an ungrammatical position. The 144 sentences were printed on an answer sheet. An equal number of exemplars of each sentence type appeared on the first and second halves of the test. The grammatical and ungrammatical versions of each pair always occurred on separate halves of the test.

The test sentences were recorded by a male native speaker of English, who spoke at a constant moderate rate and took care to articulate all sounds, including word-final consonants. The sentences were digitized and then rerecorded in the same quasi-random order in which they appeared on the answer sheet. The digi-

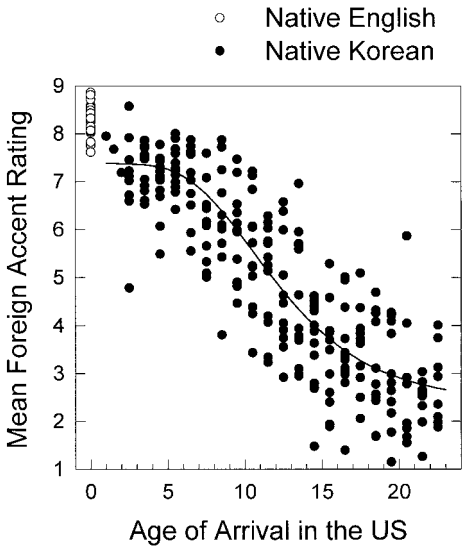

FIG. 1. The mean foreign accent ratings obtained for 24 native English and 240 native Korean participants. The data for the 240 Koreans have been fit to the Gompertz-Makehm distribution (solid line).

tized sentences were presented a single time via a loudspeaker at a comfortable level. The participants were told to wait until they heard the entire sentence before checking "Yes" (grammatical) or "No" (not grammatical) next to the written version of the sentence they had just heard. The test was unspeeded (see Chaudron, 1983). The interval between sentences provided time to respond. (If a participant ever appeared to need more time, the tape recorder was paused.) The terms "grammatical" and "not grammatical" were not defined, but these concepts were made clear to the participants through examples. Also, four practice sentences (two grammatical, two ungrammatical) were presented before the test began.

RESULTS

Effects of AOA

Figure 1 shows that the later the Korean participants had arrived in the United States, the stronger were their foreign accents. The sentences of the 24 native English controls, who were assigned an AOA of "0" years, received higher ratings than those of all but a few early-arriving Koreans. The effect of group was highly significant in a one-way ANOVA, $F(10,253) = 104.4$, $P < .01$. A series of t tests revealed that all 10 native Korean groups, even those composed of individuals who had arrived in the United States as young children, received significantly lower ratings than the native English controls (Bonferroni $P < .01$).

The AOA effect seen here is similar to one observed for certain listeners who rated Italian/English bilinguals in the Flege et al. (1995a) study. However, the foreign accent ratings obtained for certain other native English-speaking listeners did not differentiate groups of early-arriving Italian/English bilinguals from the native English controls. Differences in the ratings obtained from the 10 native English listeners who rated sentences in the present study were not explored, however, because a highly similar overall pattern of ratings was obtained from each of them (see above).

Figure 2 presents the overall grammaticality judgment test scores. As the native Korean participants' AOAs increased, their scores decreased systematically. The scores were arcsine transformed (Kirk, 1968), because variance for

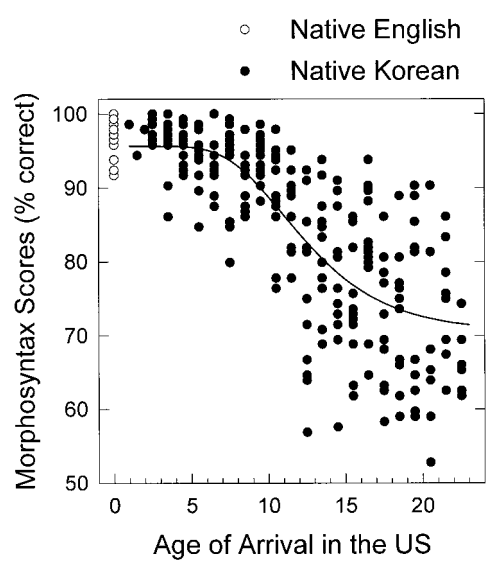

FIG. 2. The grammaticality judgment test scores obtained for 24 native English and 240 native Korean participants. The data for the 240 Koreans have been fit to the Gompertz-Makehm distribution (solid line).

the native Korean groups increased systematically as AOA increased. The significant effect of group obtained in an analysis of the transformed scores, $F(10,253) = 57.4$, $P < .01$, was followed up by a series of *t* tests. These tests revealed that groups NK7–NK21, but not groups NK3–NK5, differed significantly from the native English controls (Bonferroni $P < .01$). The AOA effect on the grammaticality judgment test scores obtained here agreed closely with the results of Johnson and Newport (1989). These authors found that native Korean and Chinese participants with AOAs of 8–39 years, but not participants with AOAs of 3–7 years, received significantly lower grammaticality judgment test scores than did native English controls.

The following procedure was adopted to determine how many native Korean participants might be said to have performed in a "native-like" fashion in the two linguistic domains. The mean and standard deviation (SD) of the 24 native English participants' ratings were calculated. We then determined which native Korean participants obtained a foreign accent rating that fell within two SDs of the native English mean. The same was done for the grammaticality judgment test scores. The number of native Korean participants who met the "two SD" criterion for the foreign accent ratings and the grammaticality judgment test scores (18 vs. 76) differed significantly, $\chi^2(2) = 28.2$, $P < .01$.

In summary, the overall effect of AOA on the foreign accent ratings and grammaticality judgment test scores was similar. However, all 10 native Korean groups differed significantly from the native English comparison group for phonology, whereas just the subgroups with AOAs of 7 to 23 years (not those with AOAs of 2 to 6 years) differed from the native English comparison group for the grammaticality judgment test scores. This suggests that AOA may constrain the learning of L2 phonology to a greater extent than L2 morphosyntax. Also, more of the Korean participants received foreign accent ratings than morphosyntax scores that fell within two SDs of the native English participants' mean values. This finding, too, might be taken as support for the view that age constraints are stronger in the domain of phonology than morphosyntax.

The Relation between AOA and L2 Performance

One aim of the analyses presented here was to determine if a discontinuity existed in AOA-L2 performance functions. First-order (linear) and third-order functions were compared in order to determine if AOA–performance functions were linear. (We reasoned that if significantly more variance was accounted for by a third-order than a first-order function, then that function was nonlinear.) The Gompertz-Makehm distribution, which is used to model aspects of the aging process (Draper & Smith, 1981, pp. 511–513), was also fit to the Koreans' ratings and scores using least-squares estimation. This provided a visual means to organize the individual participants' mean values shown in scattergrams. The other aim of this section was to determine if the correlation between AOA and measures of L2 performance would be significant for participants who began learning English after, as well as before, the end of a critical period.

Foreign accent ratings. A third-order function accounted for significantly more (1.9%) variance in the foreign accent ratings than did a first-order function, $F(2,236) = 8.8$, $P < .01$, indicating the presence of a nonlinearity in the AOA–foreign accent relation. The same procedure was then applied to AOA-defined subsets of participants in order to identify the locus of the nonlinearity. The results suggested that it did not occur near the end of the traditionally defined critical period (i.e., at an AOA of 12 or 15 years). In the analysis of the 193 participants having AOAs ranging from 7 to 23 years, the third-order function accounted for significantly more (2.5%) variance than did the first-order function, $F(2,189) = 6.56$, $P < .01$. However, when the AOA range was restricted further, to an AOA range of 7 to 18 years ($n = 144$ participants), the difference between the third-order and first-order functions (1.0%) was nonsignificant, $F(2,140) = 1.63$, $P > .10$.

The Gompertz-Makehm distribution was fit to the mean foreign accent ratings obtained for

the 240 native Korean participants in Fig. 1. A visual inspection of this function suggests that the Koreans with AOAs of about 1–5 years obtained ratings that were similar to, albeit slightly lower than, the mean rating obtained for the native English controls. The apparent lack of difference between the Koreans with AOAs of 1 to 5 years may have been due to the fact that, for many of these participants, the first extensive exposure to English occurred upon entry to school (see Table 1). There was a roughly linear decrease in ratings in the AOA range of 5–15 years, followed by a slowing in the rate at which the strength of foreign accents increased.

The simple correlations between AOA and the foreign accent ratings were computed for two subsets of the 240 native Korean participants. A significant correlation was found to exist for the native Korean participants having AOAs less than 12 years, $r = -0.62$, $P < .01$, and also for those having an AOA greater than 12 years, $r = -0.50$, $P < .01$.

In summary, the findings presented in this section suggest that the Koreans' degree of foreign accent did not increase sharply near the end of a critical period, that is, at an AOA of 12 or 15 years. The relation between AOA and degree of foreign accent appeared to be linear near the end of the supposed critical period, which fails to provide support for the critical period hypothesis. The critical period hypothesis also led to the expectation that there would be a correlation between AOA and degree of foreign accent for participants who began learning English before the age of 12 years, but not after that age. However, the AOA–foreign accent correlations were significant for both subsets of the 240 Korean participants examined here. Thus, this finding also failed to support the critical period hypothesis.

Grammaticality judgment test. A third-order function accounted for significantly more (1.2%) variance in the overall morphosyntax scores obtained for the 240 native Korean participants than did a first-order function, $F(2,236) = 3.16$, $P < .05$. When just the Koreans with AOAs of 7 to 23 years were considered ($n = 193$), the difference in variance (1.1%) was nonsignificant, $F(2,189) = 2.20$, $P > .05$. However, when those with AOAs of 7 to 18 years were considered ($n = 144$), a third-order function accounted for significantly more (5.0%) variance than did a first-order function, $F(2,140) = 5.91$, $P < .01$. This finding suggests that the relation between AOA and the test scores was nonlinear at an AOA of about 12 years; it agrees with findings reported previously by both Johnson and Newport (1989) and Patkowski (1980, 1990). However, the size of the nonlinearity observed here was much smaller than the one obtained by Patkowski (1980, 1990), probably due to methodological differences. Knowledge of English morphosyntax was assessed here using a 144-item test, whereas Patkowski had two English teachers rate transcripts of conversational speech for syntactic accuracy.

The nonlinearity just reported is not visually apparent in the fit function shown in Fig. 2, which was obtained using the Gompertz-Makehm distribution. A visual inspection of the fit function indicates that the scores declined in a roughly linear fashion between AOAs of about 6 to 15 years.[4] There is no evidence of a nonlinearity at AOAs of 12 or 15 years. However, as can be seen in Fig. 3, there was an increase in the number of participants beyond an AOA of 12 years who gave a large number of incorrect responses to ungrammatical sentences.

Finally, there was a significant correlation between AOA and the Morphosyntax test scores for the 120 native Korean participants having AOAs less than 12 years, $r = -.52$, $P < .01$. A small, but still significant correlation was also obtained for the Koreans having AOAs greater than 12 years, $r = -0.27$, $P < .01$. (The correlations were $r = -0.71$, $P < .01$, and $r = -0.23$, $P < .05$, when the sample of native Korean participants was divided at an AOA of 15 years.) This finding differs from that of Johnson and Newport (1989), but it agrees with the findings obtained in more recent studies by Birdsong (1992; Birdsong & Molis, 1998).

[4] The rate at which the scores decreased seemed to slow for AOA greater than 15 years. The census data examined by Bialystok and Hakuta, 1998, suggest that scores might continue to decline slowly over the entire life span.

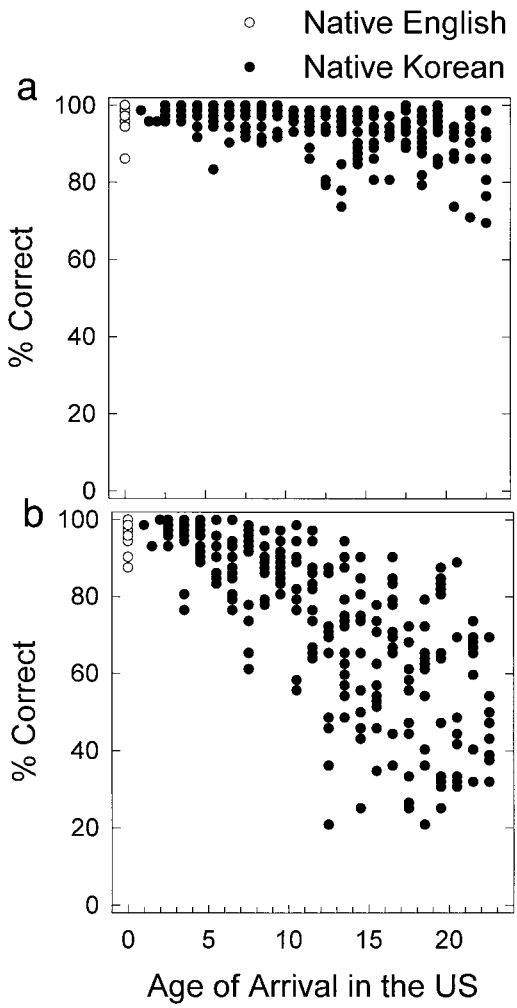

FIG. 3. The mean percentage of correct scores obtained for (a) 72 grammatical and (b) 72 ungrammatical items on a grammaticality judgment test.

In summary, when the discontinuity test was applied to the grammaticality judgment test scores, it supported the critical period hypothesis. A nonlinearity was detected in the AOA region of 12–15 years. This was apparently due to an increase in the number of participants with AOAs greater than 12 years who accepted a large number of ungrammatical sentences as grammatical (see Figs. 2 and 3). (The basis for this increase is uncertain but, as will be discussed later, it may have been related to changes in language use.) However, the pre/postcorrelation test did not support the critical period hypothesis. The correlation between AOA and the scores was significant for the participants who had begun learning English both before the age of 12 years (or 15 years) and after the age of 12 years (or 15 years).

Sentence Types

The native English controls' and early-arriving Koreans' score for grammatical and ungrammatical items on the grammaticality judgment test did not differ (because they were at ceiling for both), but later arriving native Korean participants did differ. As shown in Fig. 3, the native English controls obtained a high percentage of correct scores for both grammatical ($M = 98.3\%$, range, 86–100%) and ungrammatical ($M = 97.4\%$, range = 88–100%) sentences. However, as in previous research with nonnative speakers of English (e.g., Murphy, 1997), the native Korean participants obtained higher scores for the grammatical ($M = 94.3\%$, range = 69–100%) than ungrammatical ($M = 74.0\%$, range = 21–100%) sentences. The mean scores for grammatical and ungrammatical sentences were submitted to an (11) Group by (2) Grammaticality ANOVA, which yielded a significant interaction, $F(10,253) = 22.7$, $P < .01$. A series of t tests revealed that eight native Korean groups differed significantly from the native English controls for the ungrammatical sentences (groups NK7–NK21), whereas just six (NK11–NK21) did so for the grammatical sentences (Bonferroni $P < .01$). More importantly, the grammatical versus ungrammatical difference was nonsignificant for the native English controls and the first two native Korean groups (NK3, NK5), whereas it was significant for the remaining eight native Korean groups (NK7–NK21) (Bonferroni $P < .01$).

The basis for the difference between grammatical and ungrammatical sentences is uncertain. It may have arisen from a bias by the Korean participants to respond "grammatical" (White, 1989). It might also mean that some later arriving native Korean participants' knowledge of English morphosyntax was "fragmentary or fluctuating in its accessibility" or that their grammars of English were less "determinate" than those of the native English con-

TABLE 3

Correlations between Age of Arrival and the Percentage of Correct Responses Obtained for Nine Types of GJT Sentences

Sentence type	*r*(238)	NK	NE	*F*(10,253)	Differences
Past tense	−0.49	90.2	98.2	12.8	K13–K21
Plural	−0.64	76.9	97.4	30.0	K5–K21
Third-person singular	−0.44	89.2	99.5	11.4	K11–K21
Determiners	−0.74	78.3	98.2	49.9	K7–K21
Pronominalization	−0.51	93.2	99.7	15.0	K15–K21
Particle movement	−0.63	84.7	96.9	21.7	K13–K21
Subcategorization	−0.71	81.4	97.2	40.0	K7–K21
Subject/object raising	−0.63	77.5	93.2	19.9	K11–K21
Questions (Wh and Y/N)	−0.71	85.2	98.2	35.8	K13–K21

Note. NK and NE, the mean scores obtained for native Korean and native English participants. The *F* values are for one-way ANOVAs testing the effect of group (11 levels); all were significant at the .001 level. Differences, the NK groups which differed significantly from the native English comparison group (Bonferroni $P < .01$).

trols (Johnson et al., 1996). Still another possibility is that certain native Korean participants incorrectly judged some aspects of English morphosyntax that were tested here to be optional (Johnson et al., 1996). Additional research will be needed to choose among these interpretations.

As mentioned earlier, the sentences which comprised the grammaticality judgment test were intended to test knowledge of nine grammatical structures or "rules" (past tense, plural, Wh and Y/N questions, third-person singular, determiners, pronouns, particle movement, verbal subcategorization frames, and lexically specified subject/object raising; see Table 2). The effect of AOA on the mean scores obtained for the nine sets of sentences was examined in a series of one-way ANOVAs. As summarized in Table 3, the effect of group was significant in all nine instances ($P < .01$), but the strength of the simple correlations between the Korean' AOAs and scores for the nine sets of sentence varied considerably. A series of *t* tests revealed that nine native Korean subgroups differed from the native English controls for the plural sentences ($P < .01$). A difference was noted between the native English controls and eight Korean groups for the determiner and subcategorization sentences, seven groups for the subject/object raising and third-person singular sentences, six groups for the past tense, question, and particle movement sentences; and just the last four Korean groups for the pronoun sentences (Bonferroni $P < .01$).

The finding just presented might be taken as evidence that age constrains the learning of various aspects of L2 morphosyntax in different ways. However, such a conclusion may be unwarranted given that the sentences representing the various "rule" types were heterogeneous. As discussed by Kellerman (1995), some sentences may not have tested the intended grammatical structure or rule. Consider, for example, these two ungrammatical determiner sentences:

(1) **Tom is reading book in the bathtub.*

(2) **A boys are going to the zoo this Saturday.*

The ungrammaticality of both sentences might be attributed to the presence/absence of the determiner "a." However, the ungrammaticality of the first sentence might also be attributed to the lack of the plural marker "-s" on "book," whereas this alternative interpretation is not possible for the second sentence because "boys" must have a plural marker to agree with the verb.

An item analysis revealed a great deal of heterogeneity within the nine sentence sets. For example, there were far more errors for **The girl's swimsuit is full of sands* and **Two mouses ran into the house this morning* than for other plural sentences (e.g., **The farmer bought two*

pig at the market). The first two sentences may have been especially difficult because "sand" is a mass noun that is not pluralized by adding "-s," and "mouse" has an irregular plural ("mice"). To take another example, there were far more errors for **Yesterday the baby throwed a cat into the bathtub* and **A bat flewed into our attic last night* than for other past tense sentences (e.g., **Sandy fill a jar with cookies last night*). The first two sentences may have been especially difficult because participants lacked knowledge of which English verbs have an irregular past tense, not because they did not know how to form the regular past tense.

Rule Based versus Lexically Based Sentences

Given the heterogeneity just discussed, two functionally defined sets of sentences were established for further analysis. The sentences comprising these two new sets were drawn from multiple sentence sets. In a series of preliminary factor analyses, certain third-person singular, past tense, and determiner sentences that involved regular rules of verb and noun inflection were found to have high loadings on the same factor(s). The 22 grammatical and 22 paired ungrammatical sentences that were deemed to best reflect this functional similarity were included in a "rule based" set (see Appendix 1). These sentences tested knowledge of regular, productive, and generalizable rules of the surface morphology of English. They all involved case or number assignment on nouns or person or tense markers on verbs (e.g., regular past tense on plural formation, third-person singular morphology on present tense verbs, or case assignment on personal pronouns).

The 22 grammatical and 22 ungrammatical sentences in the "lexically based" set (see Appendix 2) were also drawn from several sentence sets (subcategorization, question, particle movement). They, too, tended to have high loadings on the same factor(s) in preliminary factor analyses. The lexically based sentences tested irregular and ungeneralizable aspects of English morphosyntax involving the proper assignment of particles or prepositions with verbs or knowledge of idiosyncratic features of particular English verbs. For example, some sentences tested which preposition should precede a nominal complement (e.g., **The farmers were hoping rain*), the use of a particle in phrasal verbs (e.g., **The little boys laughed the clown*), or the placement of particles in phrasal verbs (e.g., **The man climbed the ladder up carefully*). All ungrammatical lexically based sentences could be made grammatical by replacing the verb (for example, changing "lets" to "permits" in **The man lets his son to watch TV*). The ungrammatical rule based sentences could not be corrected in this way, however.

In Fig. 4, the 240 native Korean participants' mean scores for ungrammatical rule based and lexically based sentences have been fit to the Gompertz-Makehm distribution. The fit function for the rule based sentences showed a plateau up to an AOA of about 5 years and then a gradual decline to the end of the AOA range sampled here. For the lexically based sentences, however, the initial plateau extended to an AOA of about 8 years and then decreased rapidly to a chance level at an AOA of about 16 years.

Figure 5 shows the mean percentage of correct responses obtained for the rule based and lexically based sentences as a function of group. The scores were higher for grammatical than ungrammatical sentences (means, 95% vs. 79%). However, as AOA increased the scores for both sentence types decreased systematically, especially for the lexically based sentences.

The scores obtained for each subject were submitted to a mixed-design (11) Group $\times$ (2) Functional Type $\times$ (2) Grammaticality ANOVA, which yielded a three-way interaction, $F(10,253) = 2.34$, $P < .05$. A series of t tests revealed that the native Korean subgroups with AOAs greater than 11 years (NK11–NK21) received significantly lower scores than the native English controls for the ungrammatical sentences, but only those with AOAs greater than 13 years (NK13–NK21) did so for the grammatical sentences (Bonferroni $P < .01$). More importantly, just the five Korean subgroups with AOAs greater than 13 years (NK13–NK21) received significantly lower scores for the lexically based than the rule based sentences (Bonferroni $P < .01$).

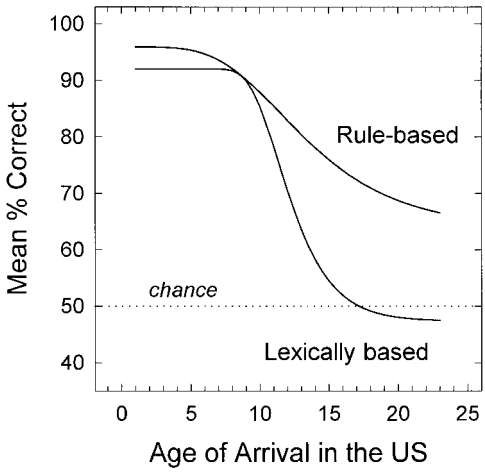

FIG. 4. Curves obtained by fitting the Gompertz-Makehm distribution to the percentage of correct scores obtained for the ungrammatical rule based and lexically based items from the grammaticality judgment test (see text).

In summary, the rule based and lexically based scores obtained for participants having AOAs less than 12 years did not differ significantly, whereas participants having AOAs greater than 12 years obtained significantly lower scores for the lexically based than rule based sentences. This finding suggests that AOA might influence the learnability of lexically based aspects of English morphosyntax to a greater extent than rule based aspects. To respond correctly to the rule based sentences, the Koreans had to acquire a simple rule with widespread application in English, such as "add *-ed* to the verb root to form past tense." The relatively good performance by late-arriving participants on rule based items may have reflected deductive (top-down) learning, the use of rule based mental representations in processing, or both.

On the other hand, success on the lexically based items may have required a kind of learning that one might characterize as "bottom-up" or "data-driven" or else learning based on the establishment of associative or probabilistic representations (see, e.g., Elman et al., 1997). As will be discussed further in the next section, the later the native Korean participants arrived in the United States, the less English-language input they were likely to have received. This may have contributed to the greater difficulty that late-arriving native Korean participants had in learning aspects of English morphosyntax tested by the lexically based than rule based sentences.

Factor Analyses

The purpose of the analyses presented here was to identify factors that might account for

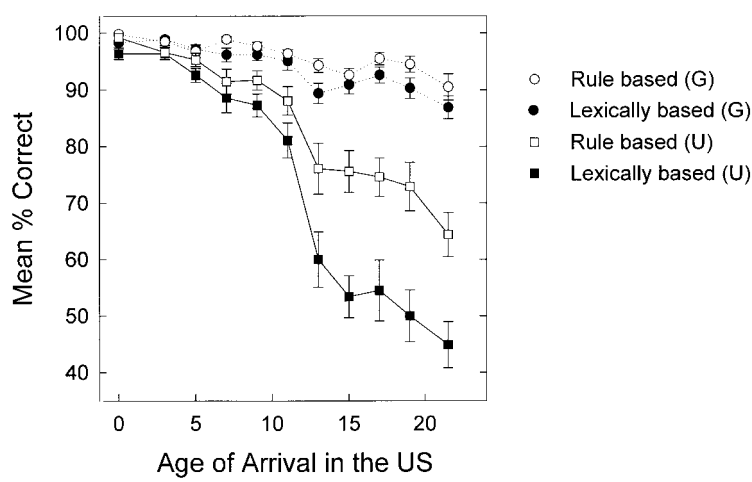

FIG. 5. The mean percentage of correct scores obtained for rule based and lexically based grammaticality judgment test sentences that were grammatical ("G") or ungrammatical ("U"). The error bars enclose ±1 standard error.

TABLE 4

Factors Identified in a Principal Components Analysis of 39 Questionnaire Items

Factor	Questionnaire items
F1: Age of L2 learning	Age of arrival in the United States (.912); estimated age of speaking English comfortably (.892); years of education in the United States (−0.856)
F2: English media input	Frequency of watching TV in English (.852); frequency of watching movies/videos in English (.848); frequency of listening to radio in English (.691)
F3: Instrumental motivation-1	Will get respect for correct English grammar and vocabulary (.861); will get respect for good pronunciation of English (.851); judged importance of English for success at work/school (.694)
F4: Judged importance of Korean	Judged importance of correct Korean pronunciation (.811); judged importance of correct Korean grammar (.788); enjoyment of learning new Korean words and expressions (.623)
F5: Judged importance of English	Judged importance of correct English grammar (.819); enjoyment of learning new English words and expressions (.701); judged importance of correct English pronunciation (.684)
F6: Languages used at work	Use of English at work (.841); use of Korean at work (−0.777)
F7: Home use of Korean	Frequency of use of Korean at home (.674); frequency of use of Korean with parents (.501)
F8: Sound processing ability	Ability to imitate foreign accents and dialects (.783); musical ability (.611); ability to remember how English words are pronounced (.532)
F9: Integrative motivation	Try to have as many American friends as possible (.741); pay attention to how E is pronounced (.724)
F10: Length of residence	Length of residence in the United States (.827)
F11: Instrumental motivation-2	Judged importance of English for getting a job (.699)

Note. The loadings for each questionnaire item are in parentheses.

variance in the outcome measures. The native Korean participants' responses to 39 language background questionnaire items were submitted to a principal components analysis with varimax rotation.[5] The resulting factors were then regressed onto the outcome measures. The results for the two primary variables (the foreign accent ratings and overall morphosyntax test scores) will be presented in the first section, and the results for the rule based and lexicon based sentences in the following section.

Principal components analysis. As summarized in Table 4, the principal components analysis identified 11 factors with eigenvalues greater than 1.0. These factors accounted for 69.4% of the variance in the questionnaire items. Just four of these factors accounted for variance in the outcome measures (see below). Factor 1 was designated "Age of L2 Learning," because the items with the highest loadings on it were AOA and the participants' estimates of the age at which they could first speak English "comfortably." Factor 2 was named "English Media Input," because the items with high loadings on it pertained to how much the native Korean participants watched movies, videos, and TV and listened to the radio in English (as opposed to Korean). Factor 8 was called "Sound Processing Ability." The items having the highest loadings on it pertained to self-estimated ability to imitate foreign accents and dialects, "musical ability," and ability to "remember how English words are pronounced." Factor 10 was designated "Length of Residence," because the only item with a high loading on it was years of residence in the United States.

Primary variables. The regression analyses examining the foreign accent ratings and overall grammaticality judgment test scores are summarized in Tables 5 and 6, respectively. More variance was accounted for in the foreign accent ratings than in the morphosyntax test scores

[5] A total of 82 missing responses comprising less than 1% of the data were replaced with group mean values.

TABLE 5

Regression Analysis Examining the Native Korean Participants' Degree of Foreign Accent

Step	Variable	R^2	Change	F value	Probability
1	F1: Age of L2 learning	0.677	0.677	747.5	.001
2	F10: Length of residence	0.713	0.036	40.0	.001
3	F8: Sound processing ability	0.737	0.024	26.7	.001
4	F2: English media input	0.761	0.024	26.3	.001
5	F9: Integrative motivation	0.774	0.014	15.0	.001
6	F11: Instrumental motivation	0.786	0.011	12.6	.001

Note. The principal components factors that were regressed onto the dependent variables are listed in Table 4. Probability, the probability of a significant increase in variance. Only factors accounting for at least 1.0% of variance are shown.

(68% vs. 49%). (This was expected because the simple correlation between AOA and the foreign accent ratings was significantly stronger than the correlation between AOA and the grammaticality judgment test scores, $X(1) = 30.9$, $P < .001$.) Other principal components factors accounted for substantially less variance in both outcome variables. Factor 10 (Length of Residence) accounted for 4 and 5% additional variance in the ratings and scores. Factor 2 (English Media Input) accounted for 2% additional variance in both outcome variables, as did Factor 8 (Sound Processing Ability). The factors that pertained to motivation (F9, F11) accounted for less than 3% of additional variance in the two outcome measures. The factors that pertained to language use (F6, F7) were not entered into either model.

The effect of length of residence (Factor 10) on the foreign accent ratings agrees with the results of Flege et al. (1995a) for Italian/English bilinguals. However, a similar finding for the morphosyntax test scores diverged from the findings of Johnson and Newport (1989; see also Patkowski, 1980, but cf. Cummins, 1981). This may have been due to the larger number of participants examined here (240 vs. 46 participants) or to a differing distribution of the length of residence variable ($M = 15$, range = 8–30 vs. $M = 10$ years, range = 3–26).

In summary, Factor 1 (Age of Learning) accounted for more variance than any other factor, and it accounted for more variance in the foreign accent ratings than grammaticality judgment test scores. This might be taken as support for the view that AOA is the most important determinant of overall success in L2 learning but, at the same time, AOA is more important for phonology than morphosyntax learning. Indeed, such conclusions are supported by partial correlation analyses. These analyses showed that the simple correlation between AOA and

TABLE 6

Regression Analysis Examining the Native Korean Participants' Grammaticality Judgment Test Scores

Step	Variable	R^2	Change	F value	Probability
1	F1: Age of L2 learning	0.494	0.494	313.7	.001
2	F10: Length of residence	0.547	0.053	33.6	.001
3	F2: English media input	0.570	0.023	14.5	.001
4	F8: Sound processing ability	0.593	0.023	14.5	.001
5	F9: Integrative motivation	0.616	0.023	14.4	.001
6	F11: Instrumental motivation	0.625	0.010	6.0	.015

Note. The principal components factors regressed onto the dependent variables are shown in Table 4. Probability, the probability of a significant increase in variance. Only factors accounting for at least 1.0% of variance are shown.

both the foreign accent ratings and morphosyntax test scores remained significant when the effects of variation in length of residence, use of English, and use of Korean were removed ($P < .01$).

However, the problem of multi-collinearity may have led to an overestimation of the importance of AOA. As mentioned earlier, the two questionnaire items with the highest loadings on Factor 1 pertained to age. However, many variables were correlated with AOA. As a result, other items also had high loadings on Factor 1 (i.e., years of education in the United States, −0.856; use of Korean with a spouse, .786; use of Korean with close friends, .729; use of Korean at social gatherings, .737; use of English at social gatherings, −0.712; and age, .700). For example, this pattern of intercorrelations might have been responsible for the surprising absence of a language use effect on the foreign accent ratings (see Flege et al., 1995a). That is, a somewhat stronger correlation between AOA and the foreign accent ratings may have obscured a weaker relation between language use variables and the foreign accent ratings. This and similar issues will be addressed in the next section, where matched subgroup analyses were performed to assess the effect of two variables correlated with AOA.

Secondary variables. Regression analyses were also carried out to examine the rule based and lexicon based morphosyntax test scores. The analysis of the rule based scores accounted for 45% of variance (F1, 32%; F8, 4%; F10, 3%; F11, 2%; F6, 1%; F5, 1%; F9, 1%), and the analysis of the lexicon based scores accounted for 61% of variance (F1, 48%; F10, 5%; F2, 4%; F9, 3%; F4, 1%; F6, 1%). It is noteworthy that Factor 8 (Sound Processing Ability) accounted for more variance in the rule based than lexically based scores (4% vs. 1%).

In posthoc analyses, we discovered that Factor 8 accounted for 10% of the variance in the rule based scores obtained for Koreans with AOAs of 14–23 years ($n = 96$), but no variance for their lexically based scores. It accounted for no variance in the rule based or lexically based scores for Koreans having AOAs of 6–13 years ($n = 96$). This finding may help explain why late-arriving native Korean participants made errors on sentences such as **Last night the old lady die in her sleep.* Phonologically nonsalient morphological markers at the end of words, which pose problems for certain children with language disorders (Leonard, 1982), may also be difficult for L2 learners (see, e.g., Mochizuki-Sudo, Susuki, Matsuno, & Kiritani, 1990) whose auditory skills are limited. If so, it would suggest that phonological and morphosyntactic learning interact in L2 learning in a way that is analogous to the interaction seen in early stages of L1 acquisition (Camarata & Gandour, 1985; Camarata & Schwartz, 1985).

Matched Subgroup Analyses

The results obtained in this study have been plotted as a function of AOA because the native Korean participants were selected according to AOA. However, this does not mean necessarily that the "age" effects presented so far can be attributed exclusively, or even primarily, to AOA. As mentioned earlier, AOA is typically confounded with other variables in studies examining large immigrant populations (see, e.g., Bachi, 1956; Bahrick et al., 1994). This study was no exception, for there was multi-collinearity among variables associated with the Korean participants' AOAs. AOA was correlated with chronological age, $r = .68$; self-estimated use of English and Korean, $r = -0.56$ and .66; years of residence in the United States, $r = -0.42$; and years of education in the United States, $r = -0.92$. Further, these variables were all correlated significantly with one another ($P < .01$).

Given the pattern of intercorrelations just mentioned, one aim of the matched subgroup analyses presented in this section was to reexamine the effect of AOA when other variables were controlled. Another aim was to assess the influence of two other variables (language use and education) independently of variation in AOA.

Variables. The Koreans' L1 use was estimated by averaging their responses to nine questions pertaining to the use of Korean at home, at work or school, in social settings, with close friends, and with a spouse. The response to each item was a number ranging from 1

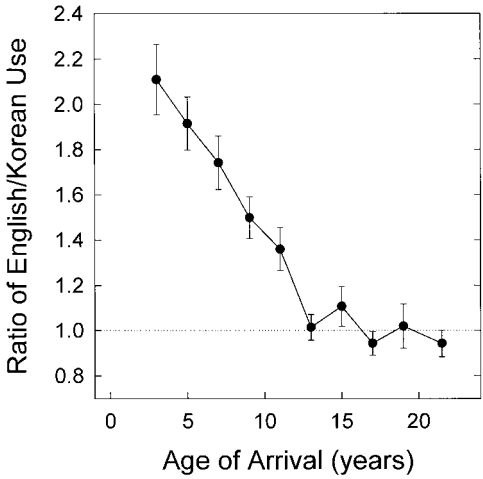

FIG. 6. Relative language use. The dashed line indicates a equal use of Korean and English according to self report. The error bars enclose ±1 standard error.

("never") to 5 ("all the time"). English use was estimated by calculating the mean ratings given to seven similar questions pertaining to English. The averages for the two sets of ratings were inversely correlated, $r = -0.76$, $P < .01$. Both the average Korean use ratings and the English use ratings were correlated with AOA (see Table 1).

An examination of language use was motivated by several observations. First, there was a striking correspondence between language use patterns and the age that is widely believed to mark the end of the critical period. Figure 6 shows the ratio of English use to Korean use. The ratios obtained for the native Korean groups having AOAs of 3–11 years were greater than 1.0, indicating more English than Korean use. One might speculate that the participants who used English more than Korean were "dominant" in English. However, the ratios were close to 1.0, indicating approximately equal use of English and Korean, for groups having AOAs of 13–21. Second, it was shown recently that the frequency with which Italian/English bilinguals spoke Italian affected their performance in English in the phonological domains (Flege, 1998a,b; Flege, Frieda & Nozawa, 1997). Finally, partial correlation analyses suggested that variations in the Koreans' use of their two languages were correlated with the outcome measures of this study independently of AOA.[6]

The other variable examined here was education. One might reasonably expect performance in some aspects of English to vary as a function of how much education the native Korean participants had received in English. However, Johnson and Newport (1989) observed a nonsignificant correlation ($r = .25$) between Morphosyntax test scores and the number of years of English classes their participants had taken before arriving in the United States.

We focused here on how many years of education the native Korean participants had completed in English-speaking United States schools, rather than on how long they had studied English in Korea. There was a correlation between the Koreans' AOA and how many years of education they had received in the United States, $r = .92$, $df = 238$, $P < .01$. The actual variable examined here was called "total years of education in the United States" (or "U.S. education," for short). It was computed by adding the number of years of special English classes the native Korean participants had taken, if any, to years of formal education in the United States.[7]

Hypotheses. The first hypothesis tested here was that when AOA was controlled, the native Korean participants who used English relatively often (and Korean seldom) would have a better pronunciation of English and receive higher morphosyntax test scores than those who used English relatively seldom (and Korean often).

[6] The foreign accent ratings reported earlier were correlated with the Koreans' self-reported use of both English, $r = .61$, and Korean, $r = .70$, $P < .01$. These correlations remained significant when variations in AOA and length of residence were removed, $r = .30, -0.29$, $P < .01$. Similarly, the correlations between the overall grammaticality judgment test scores and both English use, $r = .54$, and Korean use, $r = -0.60$, $P < .01$, were significant. These correlations remained significant when variations in AOA and length of residence were partialled out, $r = .20$ and -0.22, $P < .01$.

[7] The number of years of special English classes, $M = 1.6$ years, range = 0–4 years, was not correlated significantly with AOA, $r = .07$, $P > .10$, which reduced the correlation between total years of education and AOA, $r = .90$.

TABLE 7

Comparisons of Two Groups of 20 Native Korean Participants Each Who Differed in Amount of Education in the United States but Were Matched for AOA

Outcome variable	More years of education	Fewer years of education	$F(1,38)$
Foreign accent	5.0 (1.8)	5.0 (2.0)	0.00
Overall GJT score	87% (8)	80% (14)	3.29
Lexicon based GJT	84% (13)	79% (18)	1.39
Rule based GJT	93% (4)	85% (12)	9.15*

Note. The two groups differed significantly in U.S. education (8.9 vs. 15.1 years) but had the same AOA (12.3 years for both groups). Standard deviations are in parentheses. *$P < .01$.

The second hypothesis was that when AOA was controlled, U.S. education would affect at least some of the outcome measures. More specifically, it was hypothesized that education would have a greater effect on morphosyntax than phonology, and a greater effect for rule based than lexically based aspects of morphosyntax. The final hypothesis was that "age" effects on phonology, but not morphosyntax would remain significant after variables confounded with AOA were controlled. These hypotheses were tested by establishing matched subgroups of 20 native Korean participants, each consisting of participants drawn from the original sample of 240 without regard for the scores or ratings they obtained.

Results. The first set of matched subgroup analyses compared groups that differed in U.S. education (15.1 vs. 8.9 years, $P < .01$) but were matched for AOA (*mean* = 12.3 years for both subgroups). The matching process reduced variation for variables in addition to AOA. The matched subgroups did not differ significantly in terms of their self-reported use of Korean (3.2 vs. 3.1, $P > .10$) or English (3.9 vs. 4.0, $P > .10$). They did differ, however, in length of residence in the United States (17.0 vs. 13.6 years, $P < .05$). As summarized in Table 7, the "Much U.S. Education" group had significantly higher rule based morphosyntax scores than did the "Little U.S. Education" group ($P < .01$). However, the two matched groups did not receive significantly different foreign accent ratings, overall morphosyntax scores, or lexically based morphosyntax scores ($P > .10$).

The comparison of the two matched groups was accompanied by a set of "control" analyses comparing groups that differed in U.S. education but were not matched for AOA. The participants in the two unmatched control groups of 20 participants each were randomly selected to have the same mean number of years of U.S. education as did the two matched groups compared earlier (viz. 15.1 vs. 8.9 years, $P < .01$). Given that the participants in the control groups were randomly selected, and given that AOA was correlated with years of U.S. education in the original sample of 240 participants, the control groups differed significantly in terms of AOA (7.2 vs. 16.2 years, $P < .01$). Like the matched groups, the unmatched control groups differed significantly in terms of their rule based scores (93% vs. 85%; $P < .01$). However, unlike the matched groups, they also received significantly different foreign accent ratings (6.6 vs. 3.1), overall morphosyntax scores (94% vs. 78%, $P < .01$), and lexically based morphosyntax scores (92% vs. 73%, $P < .01$).

These results indicate that the amount of U.S. education had a significant and independent influence on just one of the four outcome measures considered here: the rule based morphosyntax scores. It is not certain whether explicit or implicit instruction affected the Koreans' learning of rule based aspects of English morphosyntax (see Winitz, 1996). Whatever the kind of instruction (or input) it was, it does not seem to have augmented the Koreans' knowledge of ungeneralizable aspects of English morphosyntax or to have improved their pronunciation of English. This finding might, therefore, be taken as support for the view that an important difference exists in the learning of rule based versus lexically based aspects of English morphosyntax (Pinker, 1991).

The second set of matched subgroup analyses compared groups that differed in AOA (9.7 vs. 16.6 years, $P < .01$) but were matched for U.S. education (mean = 10.5 years for both groups). The matching process reduced variation in variables other than just U.S. education. The two

TABLE 8

Comparisons of Two Groups of 20 Native Korean Participants Each Who Differed in AOA but Were Matched for Amount of Education in the United States

Outcome variable	Earlier AOA	Later AOA	$F(1,38)$
Foreign accent	5.2 (2.1)	3.6 (1.4)	8.22*
Overall GJT score	83% (13)	81% (9)	0.24
Lexicon based GJT	81% (17)	78% (12)	0.39
Rule based GJT	87% (11)	89% (8)	0.43

Note. The two groups differed in AOA (9.7 vs. 16.2 years) but had the same number of years of education in the United States (10.8 years). Standard deviations are in parentheses. *$P < .01$.

matched subgroups did not differ significantly in years of residence in the United States (13.9 vs. 14.7 years), Korean use (3.2 vs. 3.4), or English use (3.9 vs. 3.6, all P values $> .10$). As summarized in Table 8, the "earlier" bilinguals received significantly higher foreign accent ratings (i.e., pronounced English better) than did the "later" bilinguals. However, the earlier bilinguals did not differ significantly from the later bilinguals for any of the morphosyntax scores (overall, lexically based or rule based; $P > .10$).

The accompanying control analyses yielded different results. These analyses compared randomly selected subgroups of 20 native Korean participants each who had the same AOAs as the matched groups compared earlier (viz. 9.7 vs. 16.6 years) but were not matched for U.S. education. Given that the amount of U.S. education was correlated with AOA, the two unmatched control groups differed significantly in U.S. education (14.4 vs. 8.0 years, $P < .01$). The two unmatched groups were found to differ significantly not only in terms of foreign accent (5.9 vs. 3.4, $P < .01$) but also in terms of their overall morphosyntax scores (92% vs. 79%), lexically based morphosyntax scores (92% vs. 76%), and rule based morphosyntax scores (94% vs. 85%) ($P < .01$).

These results indicate that AOA had a significant, independent effect on just one of the four outcome measures considered here: the foreign accent ratings. From this, one might conclude that age constrains the learning of phonology but not the learning of L2 morphosyntax. The difference across linguistic domains that was observed here can be interpreted in at least two different ways. It might derive from the use of different neural substrates for phonological versus lexical-semantic and syntactic learning and/or processing (Warrington, 1975; Schwartz, Marin, & Saffran, 1979; Berndt, Caramazza, & Zurif, 1983; Mateer, 1983; Liberman & Mattingly, 1985; Keller, 1987; Gracco & Abbs, 1987), or it might arise from the use of different "modules" (Forster, 1979; Garrett, 1980; Fodor, 1983).

The native Korean participants compared in the final set of matched subgroup analyses differed significantly in their use of Korean (4.1 vs. 2.1, $P < .01$) and English (3.3 vs. 4.5, $P < .01$) but were matched for AOA (mean = 11.4 years for both). The matching process reduced variation in variables in addition to AOA. The matched subgroups did not differ significantly in years of residence in the United States (14.5 vs. 15.1 years, $P > .10$) or U.S. education (12.6 vs. 12.5 years, $P > .10$). As summarized in Table 9, the Koreans who used English relatively often and Korean relatively seldom had a significantly better pronunciation of English than did those who used English relatively seldom (and Korean often) ($P < .05$). They also had higher lexically based scores ($P < .05$). However, the two matched subgroups did not

TABLE 9

Comparisons of Two Groups of 20 Native Korean Participants Each Who Differed in Self-Reported Language Use but Were Matched for AOA

Outcome variable	Little L1/much L2	Much L1/little L2	$F(1,38)$
Foreign accent	5.6 (1.7)	4.4 (1.9)	4.27*
Overall GJT score	89% (10)	83% (12)	2.45
Lexicon based GJT	89% (12)	80% (15)	4.14*
Rule based GJT	92% (8)	88% (11)	1.32

Note. The two groups differed significantly in self-reported Korean use (4.1 vs. 2.1) but were matched for AOA (11.4 years). Standard deviations are in parentheses. *$P <$ .05.

differ significantly in terms of their overall morphosyntax or rule based morphosyntax scores ($P > .10$).

In the accompanying control analyses, 20 native Korean participants each were randomly selected to create groups that had the same mean Korean use ratings as did the matched subgroups compared earlier (viz. 4.1 vs. 2.1). Given that AOA was correlated with amount of Korean use, the two unmatched control groups in the control analyses differed significantly in AOA (16.2 vs. 7.0 years; $P < .01$). The two unmatched groups received significantly different foreign accent ratings (6.5 vs. 3.4, $P < .01$) and lexically based morphosyntax scores (93% vs. 73%, $P < .01$). Unlike the matched subgroups, they also received significantly different overall morphosyntax scores (93% vs. 76%, $P < .10$) and rule based morphosyntax scores (94% vs. 83%, $P < .01$).

These results indicate that the Koreans' pattern of language use exerted a significant, independent effect on their degree of foreign accent in English and their lexicon based morphosyntax scores, but not on their overall or rule based morphosyntax scores. The conclusion that language use affected the Koreans' pronunciation of English independently of AOA agrees with the results of a regression analysis examining the pronunciation of English by Italian/English bilinguals (Flege et al., 1995a), as well as analyses examining other aspects of Italian/English bilinguals' performance in the phonological domain (Flege, MacKay, & Meador, under review; Mackay, Meador, & Flege, under review). However, this is apparently the first time that an effect of language use has been reported for knowledge of any aspect of L2 morphosyntax.

The fact that language use affected the Koreans' learning of lexically based aspects of morphosyntax suggests that the learning of phonology and irregular, ungeneralizable aspects of morphosyntax have something in common. The physical realization of consonants and, especially, vowels varies as a function of many factors (e.g., neighboring context, speaking rate, degree of stress). The perception of speech is shaped by what one hears. As a result, the long-term memory representations developed for language-specific speech sounds depend on experience with a wide range of tokens over many years of exposure. Thus, the commonality shared by the learning of phonology and lexically based aspects of morphosyntax may be that both require a bottom-up, data-driven type of learning with associative or probabilistic representations (Elman et al., 1997). This type of learning implies that the more input an L2 learner receives from native speakers, the more native-like their representations or processing in the L2 will be.

GENERAL DISCUSSION

Two outcomes of this study were expected. First, the native Korean participants' strength of foreign accent in English grew stronger and the scores they received on a 144-item grammaticality judgment test decreased as their age of arrival (AOA) in the United States increased. Second, more individual native Korean participants, and more AOA-defined Korean subgroups differed from the native English controls in the phonological than morphosyntactic domain.

The second set of findings might be taken as support for the view that age constrains the learning of L2 phonology to a greater extent than it does the learning of L2 morphosyntax (e.g., Braine, 1971; Bever, 1981; Long, 1990; Hurford, 1991; MacWhinney, 1992). It is uncertain, however, which of several possible explanations provides the best account for the difference. Bever (1981) proposed that the difference arises because phonological learning in the L1 reaches completion sooner than does morphosyntactic learning (so that a critical period for phonology ends sooner than does one for morphosyntax). According to Cook (1992), bilinguals have more difficulty separating the phonological than morphosyntactic systems of their two languages. Others have cited the greater overall perceived similarity of phonological structures in the L1 and L2 than of corresponding morphosyntactic structures (Felix, 1980; Ioup, 1984; MacWhinney, 1987) or claimed that the role of the motor cortex in speech articulation fundamentally distinguishes

phonological learning from the learning of morphosyntax (Zatorre, 1989).

The primary aim of this study, however, was to provide a better understanding of the underlying basis for the AOA effects observed here and in previous studies. We did this by evaluating the validity of the critical period hypothesis for L2 acquisition, and with it the claim that age-related declines in L2 performance are due to a diminished ability to learn language that results from brain maturation (e.g., Scovel, 1988).

Three methods were used to test the critical period hypothesis. The discontinuity test was applied to functions relating the 240 native Korean participants' AOA to their foreign accent ratings and to scores obtained on the grammaticality judgment test. There was no evidence of a nonlinearity for the foreign accent ratings near the end of the putative critical period. However, there was a nonlinearity for the grammaticality judgment test scores, which supported the existence of a critical period for morphosyntax. The second method applied here was the pre/post-correlation test. The critical period hypothesis leads to the expectation that there will be a correlation between AOA and L2 performance for individuals who began learning their L2 before the age of 12 years, but not for those who began learning their L2 later in life. However, the AOA–foreign accent correlations and the AOA–morphosyntax correlations were significant both for Koreans with AOAs of 2–12 years and those with AOAs of 13–23 years. These findings, therefore, failed to support the existence of a critical period for the learning of either phonology or morphosyntax.

A matched subgroup analysis confirmed that the AOA effect on foreign accent ratings was not due to factors that were confounded with AOA as in previous research (e.g., Bachi, 1956; Bahrick et al., 1994; see Flege, 1998a). Two matched subgroups of 20 Korean participants each were established by selecting participants who differed in AOA but did not differ significantly in terms of how much education they had received in the United States, their length of residence in the United States, or their use of English and Korean. The later arriving subgroup (mean AOA = 16.6 years) had significantly stronger foreign accents than did the earlier arriving subgroup (mean AOA = 9.7 years), even though the other variables were controlled.

Recall that the AOA–foreign accent function was essentially linear. One might, therefore, hypothesize that L2 phonology learning is constrained by a sensitive period (Oyama, 1973, 1979; Bornstein, 1989), perhaps one that follows from, or is shaped by, brain maturation. Based on their review of a large body of relevant literature, Bates et al. (1992) noted that there is a slow, monotonic decline in synaptic density and overall levels of brain metabolism between the age of 4 years and the end of the second decade of life. These authors posited that a connection exists between the rate and extent of human neural development and the "slow decrease in capacity for second-language learning" that one sees through childhood and adolescence (1992, p. 102).

There is an alternative interpretation that we prefer, however. It is that the age-related decline in L2 pronunciation accuracy derives from the fact that, as AOA increases, the state of development of the L1 phonetic system also increases, thereby changing the way in which the L1 and L2 phonological systems interact (Flege, 1995, 1998a,b). More specifically, age-related changes in the pronunciation of an L2 may derive from differences in how, or if, L2 learners perceptually relate L2 sounds to the sounds making up the L1 phonetic inventory. This, in turn, may lead to age-related differences in whether new phonetic categories are or are not established for sounds in the L2.

The results summarized earlier provided mixed support for the existence of a critical period in the domain of morphosyntax. Given this, as well as the ambiguity that exists with respect to the discontinuity test (see the Introduction), the crucial test for morphosyntax was the matched subgroup test. The results of this test differed from the one obtained for phonology. The scores obtained for the earlier arriving subgroup were not significantly higher than those of the later arriving subgroup, even though, in a control analysis, subgroups having the same mean AOAs (9.7 vs. 16.6 years) that

were not matched on the confounded variables were found to differ significantly. This suggested that the native Korean participants' knowledge of English morphosyntax did not decrease as the result of an increase in AOA, as reported by Johnson and Newport (1989). The apparent AOA effect observed by Johnson and Newport may have been the result of factors confounded with AOA. If this conclusion is correct, then the AOA effect observed here and in previous studies cannot be ascribed to the passing of a critical period for language learning that arises inevitably from normal brain maturation.

Additional analyses provided insight into the factors that might actually have been responsible for the previously reported AOA effects on L2 morphosyntax. Two functionally defined subsets of grammaticality judgment test sentences were examined. The "rule based" sentences were characterized as testing the participants' knowledge of regular, productive, and generalizable rules of the surface morphology of English. The "lexically based" sentences, on the other hand, were characterized as testing knowledge of irregular and ungeneralizable aspects of English morphosyntax. The two sets of sentences patterned quite differently. The scores for the lexically based sentences decreased more dramatically as AOA increased than did the scores for the rule based sentences.

Even more importantly, matched subgroup analyses showed that the rule based and lexically based scores were influenced by different variables. One matched subgroup analysis compared subgroups of native Korean participants who differed in self-reported use of English and Korean but were matched for AOA. The subgroup consisting of participants who used English often obtained higher lexically based scores (and also had a better pronunciation of English) than those who used English relatively seldom. The two subgroups' rule based scores did not differ significantly, however. Another matched subgroup analysis compared subgroups of Koreans who differed in years of U.S. education but were matched for AOA. The participants with more U.S. education received higher rule based scores than those with less U.S. education. However, the two subgroups' lexically based scores did not differ significantly.

The differing effect of AOA on the rule based and lexically based morphosyntax scores, when taken together with the results of the matched subgroup analyses, bear on a conclusion that Bates and Goodman (1998) drew from their extensive review of evidence from L1 acquisition, language breakdown, and real-time processing. These authors concluded that the case for a modular distinction between grammar and the lexicon has been "overstated" in the literature. While this may be so, the evidence obtained here suggests that such a distinction is operative in L2 acquisition. The results summarized above suggest that knowledge of ungeneralizable aspects of L2 morphosyntax (as well as the ability to pronounce an L2) improves gradually as a function of experience using the L2. Knowledge of generalizable aspects of English morphosyntax, on the other hand, may be influenced more importantly by amount of formal education.

Of course, the more the native Korean participants used English, the less they used Korean. One might, therefore, hypothesize that the language use effect observed here was due to variations in how much the native Korean participants continued to speak Korean, not to the frequency with which they used English. It may be that the more the L1 is used, the more it will influence the kind of knowledge that develops for lexically based aspects of L2 morphosyntax (as well as L2 pronunciation). Still another hypothesis that might be examined in future research is that a relatively infrequent use of the L2 is an effect of poor performance in the L2, not its cause.

In summary, foreign accents grew stronger and scores on the grammaticality judgment test decreased as the Korean participants' AOAs increased. However, the underlying bases of these effects differed importantly. The AOA effect on phonology but not morphosyntax remained significant when variables confounded with AOA were controlled. The AOA effect on phonology may have been due to a sensitive period arising from brain maturation or, more

likely, from changes in how the L1 and L2 phonological systems interact as the L1 system develops. The apparent AOA effects on morphosyntax seem to have arisen from variations in education and language use that were correlated with AOA and so were unlikely to have arisen from a maturationally defined critical (or sensitive) period.

APPENDIX 1

The 44 Grammatical and Ungrammatical "Rule-Based" GJT Sentences. The Grammatical Version of Each Sentence Is Specified by the Word(s) in Parentheses

The girl cooks (cooked) dinner for her family last night.
Last night the old lady die (died) in her sleep.
Last night Mary walks (walked) to the store.
The man paints (painted) his house yesterday.
Sandy fill (filled) a jar with cookies last night.
Every Friday our neighbor wash (washes) her car.
John's dog always wait (waits) for him at the corner.
Every day Terri talk (talks) to her Mom on the phone.
Mrs. Sampson clean (cleans) her house every Wednesday.
Many house (houses) were destroyed by the flood last week.
Three boy (boys) played on the swings in the park.
Todd has many coat (coats) in his closet.
A (The) boys are going to the zoo this Saturday.
Mary opens a (the) windows in her room every night.
Him (He) is fixing the tire on Jamie's bicycle.
Them (They) worked on the project all night.
A snake bit she (her) on the leg.
Susan is making some cookies for we (us).
We ate the whole pizza by themselves (ourselves).
The girl cut himself (herself) on a piece of glass.
They took theirs (their) children to the theater.
Tom drove himselves (his) sister to the concert.

APPENDIX 2

The 44 Grammatical and Ungrammatical Lexically Based GJT Sentences. The Grammatical Version of Each Sentence Is Specified by the Word(s) in Parentheses

The farmers were hoping (hoping for) rain.
Why (What) did the company send?
The policeman was talking (talking to) a woman.
Larry went the (omit "the") home after the party.
Jenny set the book that was (omit "that was") on the bed.
The man lets his son to (omit "to") watch TV.
The girls enjoy to feed (feeding) the ducks.
Kevin called Nancy for a date up (up for a date).
The man looked the new cars yesterday over (over the new cars yesterday).
I hope you to go to (will go to) the store now.
Mrs. Johnson went to (the) library yesterday.
The man climbed the ladder up (up the ladder) carefully.
The horse jumped the fence over (over the fence) yesterday.
He came (to) my house at six o'clock.
She let the cat very quickly in (in very quickly).
The man allows his son watch TV (to watch TV).
The little boys laughed (laughed at) the clown.
The girls want feeding (want to feed) the ducks.
George says (says his prayers) much too softly.
The boy put the bowl (add: "in the kitchen").
Why (Where) did she put the book?
Nancy put the dishes last night away (away the dishes last night).

REFERENCES

Bachi, R. (1956). Statistical analysis of the revival of Hebrew in Israel. *Scripta Hierosolymitan,* **3,** 179–247.

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

Bates, E., & Goodman, J. (1998). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes,* **12,** 507–586.

Bates, E., Thal, D., & Janowsky, J. (1992). Early language development and its neural correlates. In S. J. Segalowitz & I. Rapin (vol. Eds.), *Handbook of neuropsychology: Vol. 7, Child neuropsychology* (pp. 69–100) Amsterdam: Elsevier.

Beck, M. L. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence. *Second Language Research,* **13,** 93–115.

Berndt, R., Caramazza, A., & Zurif, E. (1983). Language functions: Syntax and semantics. In Segalowitz (Ed.), *Language functions and brain organization* (pp. 5–28). New York: Academic Press.

Bever, T. (1981). Normal acquisition processes explain the critical period for language learning. In K. Diller (Ed.), *Individual differences in language learning aptitude* (pp. 176–198). Rowley, MA: Newbury House.

Bialystok, E. (1994). Towards an explanation of second language acquisition. In G. Brown, K. Malmkjaer, A. Pollitt, & J. Williams (Eds.), *Language and understanding* (pp. 115–138). Oxford: Oxford Univ. Press.

Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research,* **13,** 116–137.

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

Bialystok, E., & Hakuta, K. (1998). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. P. Birdsong (Ed.), *New perspectives on the critical period hypothesis for second language acquisition.* Hillsdale, NJ: Erlbaum.

Birdsong, D. P. (1992). Ultimate attainment in second language acquisition. *Language,* **68,** 706–755.

Birdsong, D. P. (Ed.) (1998). *New perspectives on the critical period hypothesis for second language acquisition.* Hillsdale, NJ: Erlbaum.

Birdsong, D. P., & Molis, M. (1998). Limits on attainment in second language acquisition: A replication study of Johnson & Newport (1989). [unpublished manuscript, University of Texas]

Bongaerts, T., Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition,* **19,** 447–466.

Bornstein, M. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin,* **105,** 179–197.

Bott, S. M. (1993). *Speech intelligibility and bilingualism: The effects of age of acquisition.* [unpublished Ph.D. dissertation, Univ. of Illinois, Urbana-Champaign]

Braine, M. (1971). The acquisition of language in infant and child. In C. Reed (Ed.), *The learning of language* (pp. 7–95). New York: Academic Press.

Camarata, S., & Gandour, J. (1985). Rule invention in the acquisition of morphology by a language-impaired child. *Journal of Speech and Hearing Disorders,* **50,** 40–45.

Camarata, S., & Schwartz, R. (1985). Production of object words and action words: Evidence for a relationship between phonology and semantics. *Journal of Speech and Hearing Research,* **28,** 323–330.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning,* **33,** 343–377.

Cook, V. (1992). Evidence for multicompetence. *Language Learning,* **42,** 557–591.

Coppieters, R. (1987). Competence differences between native and fluent non-native speakers. *Language,* **63,** 544–573.

Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics,* **2,** 131–149.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge: Cambridge Univ. Press.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis.* New York: Wiley.

Emmorey, K., Bellugi, U., Friederici, A., & Horn, P. (1995). Effects of age of acquisition on grammatical sensitivity: Evidence from on-line and off-line tasks. *Applied Psycholinguistics,* **16,** 1–24.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1997). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Engstrand, L. (1982). Age and length of residence as variables related to the adjustment of migrant children, with special emphasis to second language learning. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 15–123). Rowley, MA: Newbury House.

Fathman, A. (1975). The relationship between age and second language learning ability. *Language Learning,* **25,** 245–253.

Felix, S. (1980). Interference, interlanguage, and related issues. In S. Felix (Ed.), *Second language development: Trends and issues.* Tübigen, Germany: Gunter Narr.

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

Flege, J. E. (1988). The production and perception of speech sounds in a foreign languages. In H. Winitz (Ed.), *Human communication and its disorders: A review, 1988* (pp. 224–401). Norwood, NJ: Ablex.

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

Flege, J. E. (1998a). The role of subject and phonetic variables. In K. D. Gruber, D. Higgins, K. Olsen, & T. Wysocki (Eds.), *Papers from the 34th annual meeting of the Chicago Linguistic Society: Vol. II. The panels.* Chicago: Chicago Linguistic Soc.

Flege, J. E. (1998b). Age of learning and second-language speech. In D. P. Birdsong (Ed.) *New perspectives on the critical period hypothesis for second language acquisition.* Hillsdale, NJ: Erlbaum.

Flege, J. E., Bohn, O-S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics,* **25,** 437–470.

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

Flege, J. E., MacKay, I. R. A., & Meador, D. (1998). Native Italian speakers' production and perception of English vowels. [article under review]

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

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

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Forster, K. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & R. Walker (Eds.), *Sentence processing* (pp. 27–85). New York: Wiley.

Garrett, M. (1980). Levels of sentence production. In B. Butterworth (Ed.), *Language production* (pp. 177–221). New York: Academic Press.

Gracco, V., & Abbs, J. (1987). Programming and execution

processes of speech movement control: Potential neural correlates. In E. Keller & M. Gopnick (Eds.), *Motor and sensory processes of language* (pp. 163–202). Hillsdale, NJ: Erlbaum.

Hurford, J. (1991). The evolution of the critical period for language acquisition. *Cognition,* **40,** 159–201.

Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: North-Holland.

Ioup, G. (1984). Is there a structural foreign accent? *Language Learning,* **34,** 1–17.

Johnson, J. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning,* **42,** 217–248.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* **21,** 60–99.

Johnson, J., & Newport, E. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition,* **39,** 215–258.

Johnson, J., Shenkman, K., Newport, E., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language,* **35,** 335–352.

Keller, E. (1987). The cortical representation of motor processes of speech. In E. Keller & M. Gopnick (Eds.), *Motor and sensory processes of language* (pp. 125–162). Hillsdale, NJ: Erlbaum.

Kellerman, E. (1995). Age before beauty: Johnson and Newport revisited. In L. Eubank, L. Selinker, & M. Sharwood Smith (Eds.), *The current state of interlanguage* (pp. 219–231). Amsterdam: Benjamins.

Kim, E. J. (1996). *The sensitive period for second-language acquisition: An experimental study of a lexical-decision task with semantic priming and a grammaticality judgment test.* [unpublished doctoral thesis, Univ. of Illinois, Urbana-Champaign]

Kirk, R. (1968). *Experimental design: Procedures for the behavioral sciences.* Belmont, CA: Brooks/Cole.

Krashen, S., Long, M., & Scarcella, R. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly,* **13,** 573–582.

Leonard, L. (1982). Phonological deficits in children with developmental language impairment. *Brain and Language,* **16,** 73–86.

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition,* **21,** 1–36.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition,* **12,** 251–285.

MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics,* **8,** 315–327.

MacWhinney, B. (1992). Transfer and competition in second language learning. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371–390). Amsterdam: North-Holland.

Mateer, C. (1983). Motor and perceptual functions of the left hemisphere and their interaction. In S. Segalowitz (Ed.), *Language functions and brain organization* (pp. 145–170). New York: Academic Press.

MacKay, I., Meador, D., & Flege, J. (1999). Factors affecting the identification of consonants in a second language. [article under review]

Mayberry, R., & Eichen, E. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language,* **30,** 486–512.

Mochizuki-Sudo, M., Susuki, H., Matsuno, K., & Kiritani, S. (1990). The perception of articles in spoken English by Japanese college students. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics,* **24,** 163–170.

Munro, M., Flege, J., & MacKay, I. (1996). The effect of age of second language learning on the production of English vowels. *Applied Psycholinguistics,* **17,** 313–334.

Murphy, V. A. (1997). The effect of modality on a grammaticality judgment task. *Second Language Research,* **13,** 34–65.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science,* **14,** 147–172.

Olson, L., & Samuels, S. (1973). The relationship between age and accuracy of foreign language pronunciation. *Journal of Psycholinguistic Research,* **66,** 263–267.

Oyama, S. (1973). A sensitive period for the acquisition of a second language. [unpublished Harvard Univ. Ph.D. dissertation]

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

Oyama, S. (1982). A sensitive period for the acquisition of a nonnative phonological system. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 20–38). Rowley, MA: Newbury House.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning,* **30,** 449–472.

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

Pinker, S. (1991). Rules of language. *Science,* **253,** 530–535.

Pinker, S., & Prince, A. (1992). Regular and irregular morphology and the psychological status of the rules of grammar. In L. Sutton, C. Johnson, & R. Shields (Eds.), *Proceedings of the 17th annual meeting of the Berkeley Linguistics Society: General session and parasession on the grammar of event structure.* Berkeley, CA: Berkeley Linguistics Soc.

Schwartz, M. F., Marin, O., & Saffran, E. (1979). Dissociations of language function in dementia: A case study. *Brain and Language,* **7,** 277–306.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* Cambridge, MA: Newbury House.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science,* **275,** 1599–1604.

Seliger, H., Krashen, S., & Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences,* **36,** 20–22.

Shim, R. (1994). *The sensitive period for second-language acquisition: An experimental study of age effects on universal grammar.* [unpublished Ph.D. dissertation, Univ. of Illinois, Urbana-Champaign.

Singleton, D. (1989). *Language acquisition: The age factor.* Clevedon: Multilingual Matters.

Snow, C. (1979). Individual differences in second-language ability: A factor-analytic study. *Language and Speech,* **22,** 151–162.

Snow, C. (1987). Relevance of the notion of a critical period to language acquisition. In M. Bornstein (Ed.), *Sensitive periods in development: Interdisciplinary perspectives* (pp. 183–209). Hillsdale, NJ: Erlbaum.

Snow, C., & Hoefnagel-Höhle, M. (1982a). Age differences in the pronunciation of foreign sounds. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 84–92). Rowley, MA: Newbury House.

Snow, C., & Hoefnagel-Höhle, M. (1982b). The critical period for language acquisition. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 93–112). Rowley, MA: Newbury House.

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

Warrington, E. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology,* **27,** 635–657.

White, L. (1989). *Universal grammar and second language acquisition.* Amsterdam: Benjamins.

Winitz, H. (1996). Grammaticality judgment as a function of explicit and implicit instruction in Spanish. *The Modern Language Journal,* **80,** 32–46.

Zatorre, R. (1989). On the representation of multiple languages in the brain: Old problems and new Directions. *Brain and Language,* **36,** 12–147.

(Received February 9, 1998)
(Revision received January 11, 1999)