COGNITIVE PSYCHOLOGY 21, 60–99 (1989)

Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language

JACQUELINE S. JOHNSON AND ELISSA L. NEWPORT

University of Illinois

Lenneberg (1967) hypothesized that language could be acquired only within a critical period, extending from early infancy until puberty. In its basic form, the critical period hypothesis need only have consequences for first language acquisition. Nevertheless, it is essential to our understanding of the nature of the hypothesized critical period to determine whether or not it extends as well to second language acquisition. If so, it should be the case that young children are better second language learners than adults and should consequently reach higher levels of final proficiency in the second language. This prediction was tested by comparing the English proficiency attained by 46 native Korean or Chinese speakers who had arrived in the United States between the ages of 3 and 39, and who had lived in the United States between 3 and 26 years by the time of testing. These subjects were tested on a wide variety of structures of English grammar, using a grammaticality judgment task. Both correlational and *t*-test analyses demonstrated a clear and strong advantage for earlier arrivals over the later arrivals. Test performance was linearly related to age of arrival up to puberty; after puberty, performance was low but highly variable and unrelated to age of arrival. This age effect was shown not to be an inadvertent result of differences in amount of experience with English, motivation, self-consciousness, or American identification. The effect also appeared on every grammatical structure tested, although the structures varied markedly in the degree to which they were well mastered by later learners. The results support the conclusion that a critical period for language acquisition extends its effects to second language acquisition.

In most behavioral domains, competence is expected to increase over development, whether gradually or in stages. However, in some domains, it has been suggested that competence does not monotonically increase with development, but rather reaches its peak during a "critical period,"[1]

This research was supported in part by NIH Grant NS16878 to E. Newport and T. Supalla, and by NIH Training Grant HD07205 to the University of Illinois. We are grateful to Geoff Coulter, Henry Gleitman, and all of the members of our research group for discussion of the issues raised here, to Lloyd Humphreys for advice on statistical matters, to Marcia Linebarger for the loan of test materials, and to Carol Dweck, John Flavell, Dedre Gentner, Doug Medin, and two anonymous reviewers for helpful comments on earlier drafts of this paper. Reprint requests should be sent to Jacqueline S. Johnson, Department of Psychology, Cornell University, Uris Hall, Ithaca, NY 14850.

[1] In this paper we use the term *critical period* broadly, for the general phenomenon of

0010-0285/89 $7.50

which may be relatively early in life, and then declines when this period is over. For example, in the development of early visual abilities, the development of attachment, or—in the case considered here—the acquisition of language, it has been suggested that learners are best able to achieve the skill in question during a maturationally limited period, early in life. Elsewhere we have presented evidence that first language learning is indeed limited in this way (Newport & Supalla, 1987). The present paper focuses on the acquisition of a *second* language, asking whether this type of learning, undertaken only after a native language is already acquired, is nevertheless still maturationally constrained.

We will begin by reviewing prior evidence on this hypothesis, for both first and second language learning, and will then present a new empirical study which we believe shows evidence for a maturational function in second language learning. Such evidence leaves open, however, whether the underlying maturational change occurs in a specific language faculty, or rather in more general cognitive abilities involved in language learning. We will conclude by considering the types of mechanisms which are consistent with our findings.

Evidence for a Critical Period Effect in First Language Acquisition

The critical period hypothesis, as advanced by Lenneberg (1967), holds that language acquisition must occur before the onset of puberty in order for language to develop fully. As will be detailed in the subsequent section, Lenneberg's hypothesis concerned only first language acquisition; he left open the question of whether this critical period extended to second language acquisition, which would occur after a first language was already in place.

Lenneberg's argument contained two parts. First, he reviewed available behavioral evidence suggesting that normal language learning occurred primarily or exclusively within childhood. At the time his book was written, no direct evidence for the hypothesis (from normal individuals who had been deprived of exposure to a first language for varying lengths of time in early life) was available. His review therefore included

---

changes over maturation in the ability to learn (in the case under consideration in this paper, to learn language). We therefore include within this term maturational phenomena which other investigators have called sensitive, rather than critical, periods. By using the term in this broad fashion, we mean to avoid prejudging what the degree or quality of such maturational change may be (e.g., is it a sharp qualitative change vs. a gradual quantitative one?) and what the nature of the underlying maturational mechanism may be (e.g., is it a change in a special language faculty vs. a more general change in cognitive abilities?). These further questions will be addressed in part by the nature of our findings, and in part by future research.

various types of indirect evidence, for example, differences in recovery from aphasia for children vs. adults, and differences in progress in language acquisition, before vs. after puberty, in the mentally retarded.

Second, he proposed a mechanism which might be responsible for a maturational change in learning abilities. The proposed mechanism was fundamentally neurological in nature. He suggested that the brain, having reached its adult values by puberty, has lost the plasticity and reorganizational capacities necessary for acquiring language. Subsequent research has questioned whether all of the neurological events he cited occur at an appropriate time for them to serve as the basis for a critical period (Krashen, 1975). Nevertheless, the hypothesis that there *is* such a critical period for language learning has remained viable.

Since Lenneberg's writing, behavioral studies approximating a direct test of the critical period hypothesis for first language acquisition have become available. One such study is a well-known case of Genie, a girl who was deprived of language and social interaction until her discovery at the age of thirteen (Curtiss, 1977). Her lack of linguistic competence, particularly in syntax, after seven years of rehabilitation supports the critical period hypothesis. However, the abnormal conditions under which Genie was reared, including nutritional, cognitive, and social deprivation, have led some investigators to question whether her language difficulties have resulted only from lack of linguistic exposure during early life.

More recently, Newport and Supalla (Newport, 1984; Newport & Supalla, 1987) have studied language acquisition in the congenitally deaf, a population in which exposure to a first language may occur at varying ages while other aspects of social and cognitive development remain normal. Their data come from congenitally deaf subjects for whom American Sign Language (ASL) is the first language. However, since 90% of the congenitally deaf have hearing (speaking) parents, only a few deaf individuals are exposed to this language from birth. The majority of deaf people are exposed to ASL only when they enter residential school for the deaf and first associate with other deaf individuals; this can be as early as age four or as late as early adulthood.

Newport and Supalla separated subjects by their age of exposure into three groups: *native learners,* who were exposed to ASL from birth by their deaf parents; *early learners,* who were first exposed to ASL between the ages of 4 and 6; and *late learners,* who were first exposed to ASL at age 12 or later. Wishing to test asymptotic performance (i.e., ultimate command of the language), they chose subjects who had at least 40 years of experience with the language as their primary, everyday communication system. The subjects were tested on their production and comprehension of ASL verb morphology. The results show a linear decline in

performance with increasing age of exposure, on virtually every morpheme tested. That is, native learners scored better than early learners, who scored better than late learners, on both production and comprehension.

This study thus provides direct evidence that there is a decline over age in the ability to acquire a first language. It also tells us, however, that Lenneberg's portrayal is at least partially incorrect in two regards. First, the results show a continuous linear decline in ability, instead of a sudden drop-off at puberty as his hypothesis implies. (This study does not tell us whether the linear function asymptotes or continues to decline after puberty, since separate groups of later learners, before vs. after puberty, were not tested.) Second, it should be noted that, while the postpubescent learners did not reach as high a level of proficiency as the native or early learners, language had not become totally unlearnable for them. This rules out any extreme interpretation of the critical period hypothesis.

In sum, current evidence supports the notion of a maturationally delimited critical period for first language acquisition, with some modifications from Lenneberg's original formulation. However, this evidence is compatible with a number of quite different accounts of the nature of the underlying maturational change. Evidence concerning age effects on *second* language learning can contribute to a further delineation of critical period accounts.

Second Language Acquisition

What it can and cannot tell us about the critical period. Given the early difficulties of performing a direct test of the critical period hypothesis on first language acquisition, many researchers undertook studies of second language acquisition over age as a test of the hypothesis. Some investigators have suggested that a critical period theory must predict that children are better than adults at learning second languages, as well as first languages. Consequently, they have viewed any evidence to the contrary as evidence against the critical period hypothesis (cf. Snow, 1983, for discussion).

In our opinion, data on this issue do have an important consequence for a critical period theory of language acquisition. However, it is not that the critical period hypothesis could be rejected on such evidence, but rather that it can be refined or clarified by such evidence. A critical period theory for language acquisition would have quite a different character depending upon whether second language acquisition were included in its effects.

To capture this distinction there are two different ways we can state the

critical period hypothesis, one which does not include second language acquisition in its effects and one that does:

Version One: The exercise hypothesis. Early in life, humans have a superior capacity for acquiring languages. If the capacity is not exercised during this time, it will disappear or decline with maturation. If the capacity is exercised, however, further language learning abilities will remain intact throughout life.

Version Two: The maturational state hypothesis. Early in life, humans have a superior capacity for acquiring languages. This capacity disappears or declines with maturation.

Notice that, although very different in character, the two versions make the same predictions with regard to first language acquisition. They differ, however, in their predictions for second language acquisition.

The exercise hypothesis predicts that children will be superior to adults in acquiring a first language. By this account, if learners are not exposed to a first language during childhood, they will be unable to acquire any language fully at a later date. However, as long as they have acquired a first language during childhood, the ability to acquire language will remain intact and can be utilized at any age. On such a hypothesis, second language learning should be equivalent in children and adults, or perhaps even superior in adults due to their greater skills in their first language as well as in many related domains.

This hypothesis is not unlike the conception of the visual critical period described for cats (Hubel & Wiesel, 1963), where early visual experience is required to maintain and refine the structure of the visual cortex, or the conception of the critical period described for attachment in dogs (Scott, 1980), where early attachment to one dog is required for subsequently normal socialization and permits unlimited later attachments to other members of the same species. Indeed, as will be discussed below, some of the current evidence on second language learning could be interpreted to support an exercise hypothesis.

In contrast, the maturational state hypothesis claims that there is something special about the maturational state of the child's brain which makes children particularly adept at acquiring *any* language, first as well as second. This hypothesis predicts that language learning abilities decline with maturation, regardless of early linguistic experience: acquiring a first language early in life will not guarantee the ability to acquire a second language later in life. In this version, then, children will be better in second language learning as well as first.

With certain qualifications, the critical period hypothesis that Lenneberg put forth can be subsumed under either version. In fact, it is not absolutely clear which version he would have favored. Some comments

he made suggest that he thinks the young learner has a superior capacity for acquiring second languages, and therefore that he would favor the maturational state hypothesis:

> . . . the incidence of "language learning blocks" rapidly increases after puberty. Also automatic acquisition from mere exposure to a given language seems to disappear after this age and foreign languages have to be taught and learned through a conscious and labored effort. Foreign accents cannot be overcome easily after puberty. (Lenneberg, 1967, p. 176)

However, other comments within the same paragraph sound as if he would have favored the exercise hypothesis:

> . . . our ability to learn foreign languages tends to confuse the picture. Most individuals of average intelligence are able to learn a second language after the beginning of their second decade . . . a person can learn to communicate in a foreign language at the age of forty. This does not trouble our basic hypothesis on age limitation because we may assume that the cerebral organization for language learning as such has taken place during childhood, and since natural languages tend to resemble one another in many fundamental aspects the matrix for language skills is present. (Lenneberg, 1967, p. 176)

Since Lenneberg's was one of the first proposals in this area, it is not surprising that he did not take a definitive stand on this issue, particularly since there were at that time few data to support either view. Nevertheless, it is a crucial distinction that should be made in any subsequent account of a critical period.

Research on Age Effects on Second Language Acquisition

Is there an age-related limitation on the learning of a second language? A number of studies have investigated this question since the time of Lenneberg's book, focusing particularly on the acquisition of phonology and grammar. Superficially, these studies appear to contradict one another; some have been said to demonstrate an adult advantage, some a child advantage.

This apparent contradiction is resolved when one separates performance in the early stages of learning from eventual attainment in the language. (For a review of these studies, with a conclusion similar to the one presented here, see Krashen, Long, & Scarcella, 1982.) Most of the studies of second language learning have examined just the early stages of learning; these studies tend to show an adult advantage in both phonology (Asher & Price, 1967; Olson & Samuels, 1973; Snow & Hoefnagel-Hohle, 1977) and syntax (Snow & Hoefnagel-Hohle, 1978). Adults thus seem to begin moving toward second language proficiency more quickly. However, this advantage appears to be short-lived.

In contrast, studies of eventual attainment in the language show a su-

periority for subjects who began learning in childhood, both in phonology (Asher & Garcia, 1969; Seliger, Krashen, & Ladefoged, 1975; Oyama, 1976) and in syntax (Oyama, 1978; Patkowski, 1980). However, most of the studies of child–adult differences in ultimate attainment have focused on pronunciation. With anecdotal evidence that late learners do carry an accent and experimental findings that support it, most investigators will concede a child advantage for acquiring phonology (though not necessarily a maturational one; see, for example, Snow & Hoefnagel-Hohle, 1977, Olson & Samuels, 1973).

There is much less available evidence on child–adult differences in the ultimate attainment of grammar. To our knowledge, only two studies have been done. In both, the subjects were U.S. immigrants who were exposed to English upon moving to the United States and who had lived in the United States for at least five years at time of the test.

In one study, subjects' syntactic ability was assessed by trained judges who assigned syntactic ratings to written transcripts of the subjects' speech from tape recorded interviews (Patkowski, 1980). For purposes of analysis, subjects' scores were divided along two variables: age of arrival in the United States (before vs. after age fifteen), and years in the United States (under vs. over 18 years). Additionally, measures of the subjects' exposure to English in both natural and classroom settings were taken. Using either the results from the analysis of variance test or correlations, age of arrival was the only significant predictor of syntactic proficiency, with the prepubescent learners outperforming the postpubescent learners. The correlation of age of arrival with score was $-.74$, which indicates a linear trend; however, the exact shape of the relationship cannot be determined from the reported results.

In the second study mentioned, subjects were measured on their ability to repeat spoken English sentences which had been masked with white noise (Oyama, 1978). This task was meant to tap the ability to integrate different sources of linguistic knowledge including phonology, syntax, intonation, and redundancy patterns. Admittedly this is not a pure measure of syntactic ability; however, it presumably involves syntactic knowledge (along with other factors). This study found the same pattern of results just reported: age of arrival was the only significant predictor of test performance.

In addition, the Oyama study addressed important claims regarding whether children's superiority over adults in final attainment is due to factors other than maturation, which happen to be correlated with age. For example, it has been argued that the adult is less *motivated* than the child to learn the language fully, is more *self-conscious* about speaking (i.e., practicing and making errors), does not have the cultural *identification* with the host country necessary to become fluent, and in general is

less able to achieve the open attitudinal and affective state required for language acquisition to take place (for reviews of this view, see Schumann, 1975; Krashen, 1982). To test these claims, Oyama measured each of these variables, plus other candidate predictors, using interview and questionnaire material. Simple correlations showed a good association between these variables and test score; however, partial correlations removing the effects of age of arrival became essentially zero. In contrast, when the reverse procedure was performed, removing each of these variables from the relationship between age of arrival and test score, the partial correlation remained large and significant. In short, age of arrival, rather than the attitudinal variables, predicted language performance.

These are important findings, for they support the view that age effects are not simply an artifact of child–adult differences in affective conditions of learning. However, a more rigorous test of this question could be performed. Nonmaturational hypotheses do not typically propose that one attitudinal variable, for example, self-consciousness, will alone predict performance; rather, they propose that the combination of all of these variables favors children over adults. Thus a more stringent test would involve partialling out all of the attitudinal variables together from age of arrival, and then determining whether there is any predictive power left.

The study we present in the present paper is an attempt to supplement the findings of these earlier studies. It is similar to the two studies discussed above, in that the focus is on ultimate command of the grammar of the second language as a function of age of exposure to that language. It differs from previous studies, however, in the way subjects' proficiency in the language is assessed and in the types of analyses performed. First, a detailed evaluation of subjects' knowledge of numerous aspects of English morphology and syntax is performed. This allows us to examine the relationship between age of exposure and an overall measure of English proficiency, as well as the possible differential effects of age of exposure on various aspects of grammatical structure. Second, a wide range of ages of exposure is examined, so that the precise shape of the function relating age to proficiency can be determined. Third, multivariate analyses are used to evaluate the relative contributions to proficiency of age as well as a number of affective, sociological, and environmental conditions of learning.

In detail, the primary questions that we address are as follows:

(1) Is there an age-related effect on learning the grammar of a second language?

(2) If so, what is the nature of this relationship? What is the shape of the function relating age to learning and ultimate performance, and where (if anywhere) does the relationship plateau or decline?

(3) Can experimental or attitudinal variables, separately or together, explain the effects obtained for age of learning?

(4) What areas of the grammar are the most and least problematic for learners of different age groups?

In answering these questions we hope to gain a better understanding of the nature of the critical period and, most particularly, to be able to decide between the two versions of the critical period outlined above.

METHOD

Subjects

Subjects were 46 native Chinese or Korean speakers who learned English as a second language. Chinese and Korean were chosen as the native languages because of their typological dissimilarity to English. (For consideration of the effects of the first language on the second, see Discussion.) No differences were found in the results for the two language groups, so they will be presented together throughout the paper.

The primary criterion for selecting subjects was that they vary in the age at which they moved to the United States and thereby first became immersed in English. All subjects were exposed to English by native speakers in the United States. In addition, to be sure that subjects had sufficient experience with English to be considered at their ultimate attainment in the language, every attempt was made to obtain subjects who had lived in the United States for many years. Minimum criteria were as follows: all subjects had to have at least five years of exposure to English and had to have lived in the United States for an unbroken stay of at least three years prior to the time of test. Finally, to ensure ample exposure to English and to ensure some homogeneity of social background, all subjects were selected from the student and faculty population at an American university (University of Illinois). Subjects were recruited through posted sign-up sheets, letters, and by word of mouth.

The resulting 46 subjects varied in age of arrival in the United States from ages 3 to 39; throughout that range there was a fairly even distribution of ages of arrival. Age of arrival was considered the age of first exposure to English. Three additional subjects were tested but eliminated from data analysis when our posttest interview revealed that they did not meet the above criteria: One did not have an unbroken stay in the United States for three years prior to test; the second did not arrive in the United States until adulthood but was immersed in English through attending an all-English-speaking school in a foreign country. For both of these subjects, then, age of immersion could not be determined unambiguously. The third subject was eliminated because her early exposure to English

was from her Chinese parents, who had no prior experience with English but nevertheless decided to speak only English in the home upon their arrival in the United States. Most of her early exposure to English was therefore not to standard English.

Additional experiential characteristics of the subjects varied for subjects arriving in the United States early vs. late in life, and will be discussed separately for these two groups. In all cases, these experiential characteristics, as well as age of arrival, will be evaluated for their relationship to performance in English.

Early arrivals. There were 23 subjects, 12 males and 11 females, who had arrived in the United States before age 15. These early arrivals were, at the time of test, for the most part freshman or sophomore undergraduates who received money or class credit for their participation. All of these subjects, from the time of arrival until college, lived in an environment where their native language was spoken in the home and English spoken outside of the home. Once they entered college, all lived predominantly in an English-speaking environment.

Late arrivals. The remaining 23 subjects were 17 males and 6 females who had arrived in the United States after age 17. Prior to coming to the United States, all of these subjects had had between 2 and 12 years of mandatory formal English instruction in their native country. This raised two possible concerns: One, the classroom experience might reduce the effect of age of arrival on learning, since age of first exposure to English for these subjects is earlier than age of arrival. Two, "age of learning" may turn out to be better defined by age of starting classes rather than age of arrival, which would result in a narrower range of ages than desired. Whether point 2 is true is an interesting question itself and will be examined empirically in the results section.

At the time of test, these subjects were primarily professors, research associates, and graduate students. All subjects, in both the early and late arrivals groups, had at least some years of schooling while in the United States. Within the late arrivals, the smallest number of years of school in the United States was 3 years, the largest 10, with an average of 6 years for the group.

For some of the subjects, the language environment was analogous to that of the early arrivals, in which the native language was spoken in the home and English spoken at school and work; for others, particularly those that were unmarried, the language environment was almost all English. Thus in terms of exposure on a day to day basis, it does not appear that the early arrivals have any advantage over the late arrivals.

In terms of years of exposure in the United States the late and early arrivals also are fairly even. See Table 1. The average number of years in the United States for early and late arrivals is 9.8 and 9.9, respectively.

TABLE 1
The Distribution of Early and Late Arrivals in Terms of the Number of Years They Lived in the U.S.

	Age of arrival	
Years in the U.S.	3—15	17–39
3–6	4	7
7–10	10	11
11–15	9	3
23–26	0	2

The main difference between the two groups is that the late arrivals have a larger range of years in the United States.

To provide a baseline performance on tests of English, 23 native speakers of English were run. Two additional native subjects participated but were not included in the analysis, one because the posttest interview revealed that he acquired English outside of the United States, and one because she spoke a nonstandard dialect of English.

Procedure

The subjects were tested on their knowledge of English syntax and morphology by being asked to judge the grammaticality of spoken English sentences of varying types (see Materials). While such a task, of course, in principle requires metalinguistic skills in addition to knowledge of the language, virtually perfect performance is shown on the same task by 6- and 7-year old native speakers in subsequent studies (Johnson, Newport, & Strauss, in press). This suggests that the metalinguistic skills necessary for our task can only be minimally demanding for an adult and that any variation obtained in performance on the task among adults must be due to variation in knowledge of the language.

The test sentences were recorded on tape by a native American female voice (E.N.). Each sentence was read twice, with a 1–2 second pause separating the repetitions. They were said clearly, with normal intonation at a slow to moderate speed. The ungrammatical sentences were spoken with the intonation pattern of the grammatical counterpart. There was a 3–4 second delay between the different sentences.

Subjects were tested individually in the laboratory. They were instructed to make a grammaticality judgment for each sentence, guessing if they were not sure. It was made clear to the subject that if the sentence was incomplete or otherwise wrong for any reason, they should regard it as ungrammatical. The subject recorded yes/no responses on an answer sheet by circling Y or N. To avoid giving cues to the subject, the exper-

imenter did not face the subject during the testing session while the tape was going. Subjects were given a break halfway through the test, but were told prior to starting that they should tell the experimenter to stop the tape at any time if they need to break sooner, either if the tape was too fast for them or if they were simply getting tired.

Following the grammaticality judgment test, subjects were interviewed for approximately half an hour about their language background. Information was gathered about the type and amount of exposure to English they had, from when they were first learning the language until the time of test. Motivational and additudinal measures were also taken, by having the subjects rate themselves on a scale of 1 to 5 with regard to those measures.

None of the subjects were blind as to the nature of the experiment. They were told prior to participating that we were interested in determining whether children or adults are better at learning second languages; they were not told, however, what type of results were expected.

Materials

The judgments of grammaticality test was modeled loosely after one used by Linebarger, Schwartz, and Saffran (1983) in a study unrelated to the present one. Our test, however, has a different set of English constructions and corresponding test sentences than those of Linebarger et al., with the exception of two rule types which are noted.[2]

Our test was composed of 276 sentences.[3] Of these, 140 were ungrammatical. The other 136 formed the grammatical counterparts of these sentences.[4] The pairs that were formed, between the ungrammatical and grammatical counterparts, were sentences that were exactly the same except for one rule violation contained in the ungrammatical sentence. The pairs of sentences were constructed to test 12 types of rules of English, listed in Table 2. The test contained between 6 and 16 pairs of sentences testing each rule type. The members of a pair were, however, not adjacent to each other, but rather were placed in opposite halves of

[2] We thank Marcia Linebarger for making these and other tests available to us.

[3] An additional six sentences, three ungrammatical and three the grammatical counterparts of these, were included in the test but were eliminated from scoring because native speakers of English made large numbers of errors in judging their grammaticality, due to either auditory problems or dialect variations.

[4] The numbers of ungrammatical and grammatical sentences are unequal because some rule types have more than one grammatical sentence, or more than one ungrammatical sentence, within each set of counterparts (see, for example, the section on particle movement). For the most part, however, the grammatical and ungrammatical sentences form pairs, and for ease of presentation they will be referred to as "pairs" throughout the paper.

TABLE 2
12 Rule Types Tested in Grammaticality Judgment Task

1. Past tense	7. Particle movement
2. Plural	8. Subcategorization
3. Third person singular	9. Auxiliaries
4. Present progressive	10. Yes/no questions
5. Determiners	11. Wh-questions
6. Pronominalization	12. Word order

the test. Within each half, sentences were presented in random order (see Design for further details.)

To ensure as much as possible that the sentences tested the rules under study and not extraneous factors, sentences were constructed to contain only relatively high frequency words, most of which were only one or two syllables in length. The location of the grammatical error (at the beginning, middle, or end of the sentence), the basic phrase structure of the sentence, and the sentence length (ranging from 5 to 11 words per sentence) were balanced across pairs of sentences testing each rule type, so that each rule type was tested by a set of sentences comparable in all of these regards.

The 12 rule types we tested were chosen to represent a wide variety of the most basic aspects of English sentence structure. (Indeed, according to our expectations, native speakers of English found the test very easy, with ungrammatical sentences producing strong feelings of ungrammaticality.) Within the 12 rules types, there were four rule types which dealt specifically with English morphology: past tense, plural, third person singular, and present progressive. They will be discussed together since many of the violations were constructed along similar lines. The other eight types involved various rules of English syntax. Within each rule type, the violations were formed on the basis of a few basic formats, with several pairs of sentences (typically 4) using each format. These are discussed in more detail, with examples of the structure of the pairs, below.

Morphology: Past tense, plural, third person singular, and present progressive. For morphology, the grammatical sentence always contained the target morpheme in a required context, while the grammatical violation was created using one of four formats:

(1) by omitting the required morpheme;

(2) by replacing the required morpheme with an inappropriate morpheme from a different class;

(3) by making an irregular item regular;

(4) by attaching a regular marking to an already irregularly marked item.

The first format was used to make ungrammatical sentences for all four types of morphology. The sentence pairs were constructed so that the grammatical context required the target morpheme, making it a grammatical violation when the morpheme was omitted in one of the sentences of the pair. For example, in sentences (1a) and (1b), a plural marker is required on the noun "pig," and is present in (1a) but is omitted in (1b). In sentences (2a) and (2b), the present progressive ending is required on the verb "speak"; it is present in (2a) but omitted in (2b).

(1a) The farmer bought two pigs at the market.
*(1b) The farmer bought two pig at the market.
(2a) The little boy is speaking to a policeman.
*(2b) The little boy is speak to a policeman.

Sentences were structured similarly for the other classes of morphemes.

The second format applied only to the verb morphology. One sentence of the pair was correct; the other had an inappropriate tense marking for the context. Consider, for example, sentences (3a) and (3b).

(3a) Yesterday the hunter shot a deer.
*(3b) Yesterday the hunter shoots a deer.

In (3a), the verb is in the past tense form as required, while in (3b) the verb "shoot" occurs in present tense form in a past tense context.

The last two formats for creating the ill-formed sentences could be used only for past tense and plural forms. An ill-formed sentence created by making an irregular item regular is exemplified in sentence (4b), with its grammatical counterpart in (4a). Similarly, the ungrammatical sentence (5b) has a regular marking added on an already marked irregular.

(4a) A shoe salesman sees many feet throughout the day.
*(4b) A shoe salesman sees many foots throughout the day.
(5a) A bat flew into our attic last night.
*(5b) A bat flewed into our attic last night.

The test was constructed so that there was an equal number of sentence pairs (4) in each format used for each type of morphology. However, due to the nature of the morphemes, it was impossible for all of the formats to be applied to all of the four rule types. Therefore the past tense and plural are tested by more sentence pairs than are the third person or the present progressive.

Determiners. To test subjects' knowledge of determiners, the grammatical member of the sentence pairs was constructed so that a determiner in a particular position was either necessary or not allowed. The ungrammatical counterparts were then formed by one of three methods: (1) by omitting them in required contexts, as in sentence (6b); (2) by substituting

the indefinite for the definite, as in (7b); and (3) by inserting them where neither article is allowed, see (8b). These examples can be contrasted with their grammatical counterparts (6a), (7a), and (8a), respectively:

(6a) Tom is reading a book in the bathtub.
*(6b) Tom is reading book in the bathtub.
(7a) The boys are going to the zoo this Saturday.
*(7b) A boys are going to the zoo this Saturday.
(8a) Larry went home after the party.
*(8b) Larry went the home after the party.

In many cases, there are other ways of construing the errors; for example, (6b) may be construed as a plural error, instead of a determiner error, for not having the plural marking on the noun "book." In cases like these, where the error classification was ambiguous, the semantic contexts were created to try to bias the listener into the preferred reading. For example, in (6) the reason Tom is in the bathtub is to sway the subject into expecting that he is reading only one book rather than many.

Pronominalization. The sentence pairs for this rule type contain some type of pronominal. The ungrammatical sentences were formed to include one of the following violations: (1) the wrong case marking on the pronoun; (2) an error in gender or number agreement for the pronoun; or (3) an erroneous form of the possessive adjective.

The violations of case involved using nominative pronouns in objective positions (see (9a) and (9b)), and objective pronouns in nominative positions:

(9a) Susan is making some cookies for us.
*(9b) Susan is making some cookies for we.

Gender and number were tested by capitalizing on the fact that reflexive pronouns have to agree with the noun they are coindexed with. Sentence (10a) is an example of correct gender agreement, while (10b) shows a gender agreement violation:

(10a) The girl cut herself on a piece of glass.
*(10b) The girl cut himself on a piece of glass.

For possessive adjectives, the error is in the form the word takes. So, for example, some ungrammatical items have a possessive adjective with the possessive marker added, as in (11b). Compare this to the correct form in (11a):

(11a) Carol is cooking dinner for her family.
*(11b) Carol is cooking dinner for hers family.

Particle movement. With some minor changes, all of the items in this

rule type are from Linebarger et al. (1983). Here the sentences take advantage of the differences between particles and prepositions. The ill-formed sentences were created by treating prepositions as particles, that is, by moving the preposition to the right of the object NP as in (12b), as compared to the correct form in (12a). These were contrasted with grammatical sentences with particles in their moved and unmoved positions, as in (13a) and (13b). Additionally, other sentences were ill-formed by moving the particle outside its own clause as in (13c). Notice that, for this rule type, the sets of counterpart sentences are not pairs but triples:

(12a) The man climbed up the ladder carefully.
*(12b) The man climbed the ladder up carefully.
(13a) Kevin called up Nancy for a date.
(12b) Kevin called Nancy up for a date.
*(13c) Kevin called Nancy for a date up.

Subcategorization. The items in this rule type are also from Linebarger et al. (1983). These items test subjects' knowledge of the subcategorization frames of various verbs. In English, individual verbs determine the type of syntactic frames that may follow them. For example, some verbs require a direct object, while others require prepositional phrases. Because the details of these frames are lexically determined, ill-formed sentences could be created by changing the structure of the required frame for a particular verb while keeping the meaning intact. Thus, the change in these sentences involved using the subcategorization frame of a semantically similar verb. See, for example, the contrasts below.

(14a) The man allows his son to watch T.V.
*(14b) The man allows his son watch T.V.
(15a) The man lets his son watch T.V.
*(15b) The man lets his son to watch T.V.

The ungrammatical sentences were formed by exchanging the different subcategorization frames of the two semantically similar verbs "allow" and "let."

Auxiliaries. In this rule type, the affix requirements for different auxiliary verbs were tested. In particular, the ungrammatical sentences were formed by violating three rules of auxiliaries. Each rule, with an example of the correct and incorrect forms, is given below:

"Have" requires a past participle.
(16a) The baby bird has fallen from the oak tree.
*(16b) The baby bird has fall from the oak tree.

Following any form of "be," the main verb must take the progressive.

(17a) Fred will be getting a raise next month.
*(17b) Fred will be get a raise next month.

Only the first element of Aux is tensed.

(18a) Leonard should have written a letter to his mother.
*(18b) Leonard should has written a letter to his mother.

Yes/no questions. For this rule type, the ungrammatical sentences contain primarily errors in subject–aux inversion. The errors are of three types. In one, two auxiliaries are moved in front of the subject, as in (19b). In another, both the auxiliary and the verb are fronted (20b); and in the third, the verb is fronted in a sentence where do-insertion would normally occur, as in (21b). The grammatical counterparts are (19a), (20a), and (21a), respectively.

(19a) Has the king been served his dinner?
*(19b) Has been the king served his dinner?
(20a) Can the little girl ride a bicycle?
*(20b) Can ride the little girl a bicycle?
(21a) Did Bill dance at the party last night?
*(21b) Danced Bill at the party last night?

Additionally, there were some ungrammatical sentences formed by copying, instead of moving, the auxiliary verb, the difference being shown in (22a) and (22b):

(22a) Can the boy drive a tractor?
*(22b) Can the boy can drive a tractor?

Wh-questions. The ungrammatical wh-questions have three forms, two of them also dealing with aux. In one form, no subject–aux inversion occurs, as in (23b) as compared with (23a); in the other, do-insertion is omitted, as in (24b) compared to (24a):

(23a) When will Sam fix his car?
*(23b) When Sam will fix his car?
(24a) What do they sell at the corner store?
*(24b) What they sell at the corner store?

The third form of the ungrammatical wh-questions was lexical. A question was ill-formed by substituting an incorrect wh-word for a correct one. In sentence (25b), for example, "why" cannot be used unless the subcategorization frame of the verb "put" is satisfied by supplying a locative. Sentence (25a) satisfies this restriction by replacing the locative with a locative wh-word.

(25a) Where did she put the book?
*(25b) Why did she put the book?

Word Order. In this last rule type, basic word order rules are tested. Sentences of three types were used: intransitive (NP–V), transitive (NP–V–NP), and dative (NP–V–NP–NP). Within each type, the ungrammatical sentences were formed by systematically rearranging the verbs and noun phrases so that all of the possible orders of constituents occurred. Thus the simplest ill-formed sentence involves the reversal of an NP and intransitive verb, as in (26a) versus (26b); the most complex involves the rearrangement of NPs and V in double-object structures, as in (27a) versus (27b).

(26a) The woman paints.
*(26b) Paints the woman.
(27a) Martha asked the policeman a question.
*(27b) Martha a question asked the policeman.

Design

The test was divided into two halves. An equal number of exemplars of each rule type and subrule type were represented in each half. The grammatical and ungrammatical members of a pair were in opposite halves of the test. Within each half, sentences were randomized in such a way that no rule type was concentrated in one section of the test, and no run of grammatical or ungrammatical sentences was longer than four.

Results

Age of Acquisition

Age of acquisition and ultimate performance. The primary question of this study involved examining the relationship between age of learning English as a second language and performance on the test of English grammar. The results show a clear and strong relationship between age of arrival in the United States and performance. Subjects who began acquiring English in the United States at an earlier age obtained higher scores on the test than those that began later, $r = -.77$, $p < .01$.

A more detailed understanding of this relationship can be gained from Table 3 and Fig. 1. Subjects were grouped by age of arrival into categories similar to those used in past research (e.g., Snow & Hoefnagel-Hohle, 1978). Table 3 presents the mean score, standard deviation, and the ranges of the number of correct responses and the number of errors for each group and for the native English comparison group. The means are also presented graphically in Fig. 1. The adjacent age groups were com-

TABLE 3
Mean Scores of Nonnative and Native Speakers of English

		Age of arrival			
	Natives (n = 23)	3–7 (n = 7)	8–10 (n = 8)	11–15 (n = 8)	17–39 (n = 23)
Means	268.8	269.3	256.0	235.9	210.3
SD	2.9	2.8	6.0	13.6	22.8
Range	275–265	272–264	263–247	251–212	254–163
(Errors)	(1–11)	(4–12)	(13–29)	(25–64)	(22–113)

Note. Maximum score = 276.

pared, two at a time, by a set of two-sample t tests using separate variance estimates.[5]

The first comparison involved determining whether there was any difference between the age 3–7 group and the native group in their performance in English. The two groups were not significantly different from each other, $t(10.4) = 1.28, p > .05$; indeed, the two groups were entirely overlapping in performance. In contrast, all of the other age groups performed significantly below the natives (for natives vs. the next closest group (8–10), $t(8.1) = 6.67, p < .01$). This suggests that, if one is immersed in a second language before the age of 7, one is able to achieve native fluency in the language;[6] however, immersion even soon after that age results in a decrement in ultimate performance.

Given that the 3–7 group is the only group that reached native performance, it is perhaps not surprising that the difference between the means of the 3–7 and 8–10 age groups is significant, $t(10) = 5.59, p > .01$. As can be seen in Table 3, while the absolute difference between the means of these two groups is small, both groups have very small SDs, and the range of scores for the 3–7 group is entirely nonoverlapping with the 8–10 group. All of the later adjacent age groups are also significantly different from each other. The age 8–10 group obtained higher scores than the 11–15 group, $t(9.7) = 3.83$, $p < .01$, with almost nonoverlapping distributions between the two groups, and the age 11–15 group obtained higher scores than the 17–39 (adult) group, $t(21) = 3.78, p < .01$.

In sum, there appears to be a strong linear relationship between age of exposure to the language and ultimate performance in that language, up to

[5] Using a two-sample t statistic where the variance of each group is estimated separately is appropriate whenever the population variances are not assumed to be equal, as is the case here.

[6] It is always possible, however, that the equivalence in performance between natives and the 3–7 group is due to a ceiling effect on our test, and that tests of more complex aspects of English syntax would show differences even between these groups.

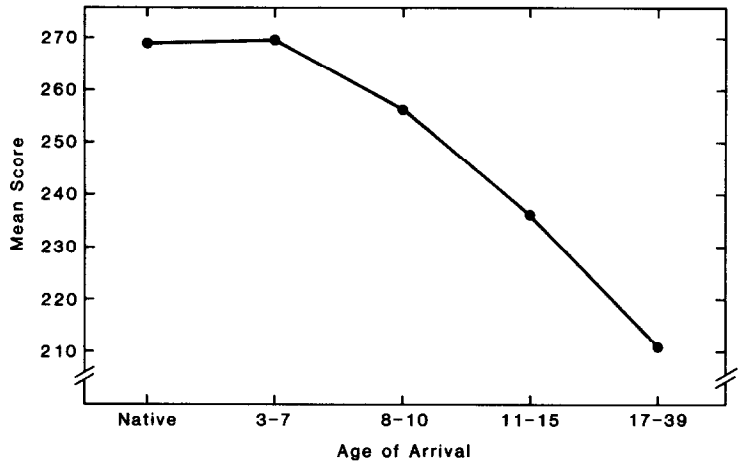


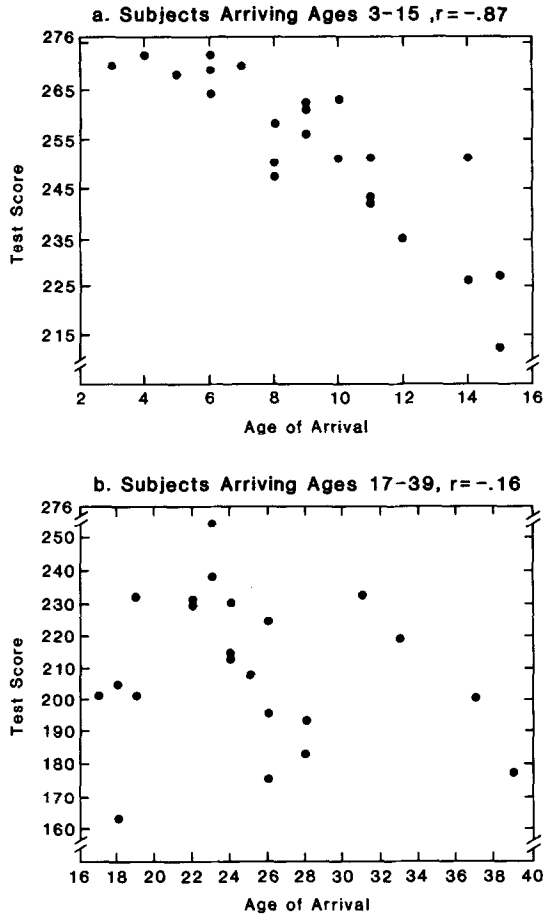
FIG. 1. The relationship between age of arrival in the United States and total score correct on the test of English grammar.

adulthood. In the next section we examine the shape of this function in more detail.

The effects of age of acquisition before vs. after puberty. An important question to answer is whether, throughout adulthood, performance continues to decline as a function of age of exposure or whether it plateaus at some point (H. Gleitman, personal communication). If the explanation for late learners' poorer performance relates to maturation, performance should not continue to decline over age, for presumably there are not many important maturational differences between, for example, the brain of a 17-year old and the brain of a 27-year old. Instead, there should be a consistent decline in performance over age for those exposed to the language before puberty, but no systematic relationship to age of exposure, and a leveling off of ultimate performance, among those exposed to the language after puberty. This is precisely what was found.

Subjects were divided into two groups in terms of age of exposure, from age 3–15 versus 17–39, with an equal number of subjects ($N = 23$) in each group. The correlations between age of exposure and performance for these two groups were strikingly different. For the group first exposed to English between the ages of 3 and 15, the correlation was $-.87$, $p < .01$. Note that this correlation is even more substantial than that for the subjects as a whole. In contrast, for the group first exposed to English between the ages of 17 and 39, there is no significant correlation, $r = -.16$, $p > .05$. Scatterplots demonstrating this effect are presented in Figs. 2a and b.

Age of acquisition and variance in ultimate performance. Another age-related result, which is obvious from inspecting the scatterplots of Fig. 2



*Note: The Y-axes are on different scales.

FIG. 2. Scatterplots of test score in relation to age of arrival for subjects arriving in United States before vs. after puberty.

and the SDs in Table 3, is the heterogeneous variance. For groups who acquired English at early ages, the variance is very small; with increasing age of exposure, variance gets larger, creating a megaphone shape, so that for subjects exposed to English after 15 the variance is very large. Note that it would have been quite possible to find that the means of these groups increased but the variance stayed constant over the age groups. The heterogeneity of variance obtained, and the relation between age of acquisition and variance, is an independent result.

This heterogeneity of variance underscores two simple but important points:

(1) Before age 15, and most particularly before age 10, there are very

few individual differences in ultimate ability to learn language within any particular age group; success in learning is almost entirely predicted by the age at which it begins.

(2) For adults, later age of acquisition determines that one will not become native or near-native in a language; however, there are large individual variations in ultimate ability in the language, within the lowered range of performance.

Age of exposure to formal instruction. It has been assumed thus far that age of arrival in the United States is the best measure of age of exposure to the language. For early arrivals it is the only measure available, since these subjects had no prior experience with English at all. However, for the late arrivals there are two measures possible: age of arrival in the United States, or age of beginning English instruction in school within the native country. There is already a high correlation between age of arrival and test score for the subjects as a whole, $r = -.77$; if age of classes is a better measure of first exposure for the late arrivals, then the correlation should be even higher when using that as a measure of time of exposure. This is not what was found. The correlation for the subjects as a whole between age of exposure, defined as classes or immersion (whichever came first), and test score is $-.67$. These correlations, however, are not statistically different from each other, $t(43) = 1.26$, $p > .05$. This is not surprising since, due to the early arrivals, half of the measurements are exactly the same; moreover, most of the late arrivals are defined as later learners (pubescent or postpubescent) either way they are measured. Because of this overlap in measurement, the best way to evaluate the effect of age of classes is to do so using only the subjects who had classroom instruction. For these subjects alone ($N = 23$), the correlation between age of classes and test score is $-.33$, which is not significant, $p > .05$.

This result has two implications. First, it means that we are using the right measure for "age of exposure"; age of arrival in the United States, with its resulting immersion in English, is more strongly related to ultimate performance in English than is age of beginning formal English instruction. More profoundly, it means that the learning which occurs in the formal language classroom may be unlike the learning which occurs during immersion, such that early instruction does not necessarily have the advantage for ultimate performance that is held by early immersion. It should be noted, however, that this last conclusion may be limited by the relatively narrow age range for formal instruction found in our subjects: our subjects all began their English classes between the ages of 7 and 16, with most subjects beginning at ages 12–15. This conclusion may also be restricted to the type of formal instruction received in Chinese and Korean schools (and, of course, any other schools in which the instruction is

similarly formal), and should be less true the more formal instruction approximates immersion in the United States. In any event, age of arrival in the United States appears to be the better measure of age of acquisition for the population we studied.

Experiential and Attitudinal Variables

Experiential variables. Years of exposure in the United States was also a variable of interest in this study. First, careful attention was paid to balance the years of exposure between early and late learners. This was done in order to avoid the possibility that obtained age effects would be due to differences in years of exposure, rather than to true differences in age of exposure. That we were successful in controlling for years of exposure between the early and late learners is apparent from the lack of correlation between age of arrival and years in the U.S., $r = -.09$, $p > .05$.

Beyond controlling for this potential confound, it is also important to ask what effect years of exposure has on learning, independent of the age effects. It is known that number of years has some effect on subjects' competence during the initial stages of learning a second language (see, for example, Snow & Hoefnagel-Hohle, 1978). At an extreme, people who have been in a host country for 1½ years must perform better than those who have only been there half a year. The question here is, however, do people continue to improve over time through continued exposure to the language, or do they reach an asymptote after a certain number of years? To answer this question, a correlation coefficient was computed between years of exposure in the United States and test performance. The resulting correlation, $r = .16$, is not significant, $p > .05$ (see also Table 4). This is in agreement with other studies (Oyama, 1978; Patkowski, 1980), also showing no significant effect of the number of years of exposure on language performance for learners beyond the first few years of exposure.

In addition to years of exposure, Table 4 also presents other variables which we considered possible experiential correlates with ultimate per-

TABLE 4
Correlation Coefficients of Experiential Variables with Score

Interview variable	*Correlation w/score*
Length of exposure (years in the U.S.)	.16
Amount of initial exposure (first year or two in U.S.)	.03
Age of English classes[a]	−.33
Years of English classes[a]	.25
Motivation to learn in classes[a]	.05

[a] Correlations for late learners only; measure not applicable to other subjects.

formance, such as amount of initial exposure to English, classroom experience, and attitude. Most of these variables were computed from information provided by the subjects; amount of initial exposure (measured as the percentage of time English was used during the first year or two in the United States) and motivation to learn in English classes (rated 1 to 5) were estimates provided by the subjects. None of the correlations are significant.

Regarding amount of initial exposure, the mean percentage for the group is 51.4%, with a standard deviation of 20.2%. Unless subjects' estimates are inaccurate, it appears that ultimate performance is not sensitive to fairly large differences in amount of initial exposure to the language, at least not after the subjects have been immersed in the language for a number of years.

The classroom variables include the age at which the subjects began English classes in their native country (already discussed in the previous section), the number of years they took English classes, and their ratings of how motivated they were to learn English in the classroom. Again, none of these variables correlate significantly with performance. It may be of interest for future research, however, that age of starting English classes is the highest of the (nonsignificant) experiential correlations. This may suggest some benefit of early classroom exposure, if classroom exposure occurred earlier than in the population we studied, and particularly if the classroom were more like immersion.

Attitudinal variables. Some investigators (see Schumann, 1975, and Krashen, 1982, for reviews) have suggested that age effects are secondary by-products of changes in people's level of self-consciousness, in their cultural identification, and in their motivation to learn a second language well, rather than maturational changes in learning. To address this claim, correlation and regression analyses were performed. Table 5 presents correlations of such attitudinal variables with test score as well as with age of arrival. These variables were measured by asking subjects to rate themselves according to the questions presented at the bottom of Table 5.

The correlations show a strong relationship between these attitudinal variables and both test score and age of arrival. Higher ratings of American identification and increased measures of motivation were associated with better performance in English and with younger age of arrival, while higher ratings of self-consciousness were associated with poorer performance and with later age of arrival. Both of these sets of results would be predicted by a theory which attempted to explain age differences in language learning as a function of attitudinal variables correlated with age, rather than a function of maturation. The other possibility is, of course, the reverse; the attitudinal variables may have obtained their correlations with test score as a result of the correlation with age of arrival. Thus it

TABLE 5
Correlation Coefficients of Attitudinal Variables with Test Score and Age of Arrival

Attitudinal variables	Test score	Age of arrival
Identification	.63**	−.55**
Self-consciousness	−.36*	.19
Motivation	.39**	−.48**

* $p < .05$.
** $p < .01$.
Questions:
1. How strongly would you say you identify with the American culture? (subjects reply) If 5 means you strongly identify with the American culture, that is, you feel like a complete American, and 1 means not at all, how would you rate your identification?
2. Did you feel self-conscious while learning English in the United States? (Most often an explanation was needed here). How would you rate that on a scale from 1 to 5, where 5 is very self-conscious and 1 is not at all?
3. Motivation is a composite of two questions: (a) Is it important to you to be able to speak English well? (subject's reply) On a scale of 1 to 5, where 5 means very important and 1 means not at all, how would you rate it? (b) Do you plan on staying in the United States? The composite was formed by adding one point to their importance rating if they planned on staying in the United States, and by subtracting one point if they did not.

becomes a question as to which is the better measure: age of arrival or attitudinal variables?

It is clear that age of arrival is the better measure over any of the attitudinal variables considered alone. The correlation between age of arrival and test score ($r = -.77$) surpasses the correlation between any of the attitudinal variables and test score. Furthermore, the attitudinal variables are more adversely affected when age of arrival is partialled out than is age of arrival when each of the attitudinal variables is partialled out, as shown in Table 6. This is in complete agreement with Oyama's (1978) results.

As stated earlier, however, the most powerful evidence against this alternative hypothesis is to show that age of arrival can account for vari-

TABLE 6
Partial Correlations of Age of Arrival and Attitudinal Variables with Test Score

	Attitudinal variables w/age of arrival removed	Age of arrival w/attitudinal variables removed
Identification	.39*	−.65**
Self-consciousness	−.34*	−.76**
Motivation	−.04	−.72**

* $p < .05$.
** $p < .01$.

ance not accounted for by the attitudinal variables combined. To test this, a regression analysis was performed using the three attitudinal variables together, which resulted in a regression coefficient of .47. This was compared to the .69 regression coefficient obtained with the three attitudinal variables plus age of arrival. The contribution made by age of arrival is statistically significant $F(1,41) = 28.1$, $p < .01$. This shows that, independent of any possible attitudinal effects, age of arrival has an effect on learning a second language.

Of independent interest is whether the attitudinal variables can account for any of the variance not accounted for by age of arrival. Even though it is clear that age of exposure to a language is an important variable for predicting ultimate performance, other variables may contribute to this as well. Unlike previous studies (e.g., Oyama, 1978), we did find added predictive value with two attitudinal variables: self-consciousness and American identification. Each of the two makes a significant contribution to a regression model including only age of arrival ($F(1,43) = 5.6$, $p < .05$., for self-consciousness, and $F(1,43) = 7.5$, $p < .05$., for identification), as well as a significant contribution to a regression model including age of arrival and the other attitudinal variable ($F(1,42) = 5.0$, $p < .05$, for the addition of self-consciousness to age plus identification, and $F(1,42) = 6.9$, $p < .05$., for the addition of identification to age plus self-consciousness). Motivation, whether analyzed separately or in conjunction with the other two variables, failed to add significantly to the regression coefficient. Thus it appears at first glance that a model of second language learning would have to include both age effects and the effects of self-consciousness and identification, though not the effects of motivation. Such a model might argue, for example, that while age of arrival affects language learning, so does the self-consciousness and the cultural identification of the learner.

At this time one might, however, be cautious about inferring a direct causal link between self-consciousness and cultural identification to language learning, until this result is corroborated in future studies. Not only are the effects of self-consciousness and cultural identification not supported in other studies, but also possible mediating variables have not been ruled out. For example, language performance may be correlated with subjects' evaluation of their performance, which may in turn affect how self-conscious they are and how much they identify with the host country. Thus poorer learners may, as a result of their performance problems, become more self-conscious and identify less with the United States. In this account, greater self-consciousness and less identification would be the result rather than the cause of the performance problems. In any case, apart from whether attitudinal variables do or do not play a role, there is a clear independent effect of age of arrival on ultimate performance.

Age of Acquisition and Rule Type

The results show a striking effect of age of acquisition on performance in our test of English syntax and morphology. It is of interest to know what particular areas of the grammar create the most and least problems for second language learners. Are the errors random, with an even dispersal across rule type, or do late learners err more frequently on a particular type of rule? To answer this question, an analysis was performed on age of learning in relation to the differing types of rules evaluated on the test. This analysis used only the ungrammatical items, since it is only the ungrammatical items which can be said to be testing any particular rule type. That is, when a subject marks a grammatical sentence as ungrammatical, it is unclear what part of the sentence, or grammar, (s)he is having problems with. In contrast, when a subject marks an ungrammatical sentence as grammatical, (s)he must have failed to represent just that structure under test as a native speaker would. For purposes of this analysis, the age groups were the same as those used previously, except that the late learner group was further divided into two groups, (17–24) and (25–39), with an approximately equal number of subjects in each. This was done to reach a more nearly equal number of subjects in each of the age of learning groups. A two-way analysis of variance was performed, using the 12 rule types (outlined in the methods section above) and six ages of acquisition.

The results of the anova showed a significant effect of rule type $F(11,693) = 53.2$, $p < .01$, a significant effect of age of acquisition, $F(5,63) = 32.3$, $p < .01$, and an interaction between rule type and age of acquisition, $F(55,693) = 8.3$, $p < .01$. The age effect here is simply a reproduction of the finding that early learners perform better than the late learners; apparently there is no reduction of this effect when scoring only the ungrammatical test items. The effect of rule type shows that subjects made more errors on certain rule types than on others. Finally, the interaction appears mainly to be the result of late learners making proportionately more errors on some rule types, and proportionately fewer on others. Thus, many of the late learners' errors do not appear to be random; rather, there are particular parts of the grammar that seem more difficult.

The pattern of errors for each age group across the 12 rule types can be seen in Fig. 3. In Fig. 3, rule types are ordered along the x-axis in decreasing order of difficulty for later learners.[7] As can be seen, determiners and plural morphology appear to be the most difficult for the two latest groups of learners, with scores significantly worse than chance for determiners ($t = 3.35$, $p < .01$), and no different from chance for plurals ($t =$

[7] This ranking of rule type difficulty remains the same when using other criteria, for example, ordering rule type according to the number of subjects who score almost perfectly on that rule (that is, 0 or 1 item wrong, out of 6 to 16 possible, depending on the rule type).

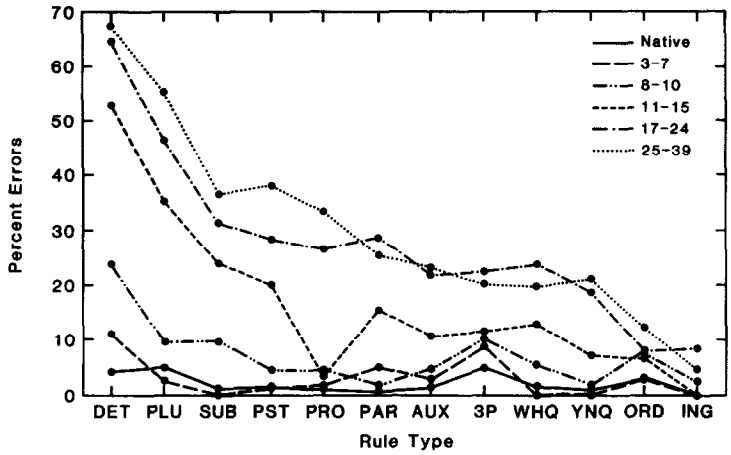


FIG. 3. Mean percentage of errors on 12 types of English rules.

.16, $p > .05$). While all of the remaining rule types receive scores significantly better than chance (t ranges from 3.46 to 26.1, $p < .01$), they vary widely in level of performance. Most notably, basic word order rules and the present progressive are giving very few problems, with most subjects getting virtually all of the items of these types correct.[8]

Why are subjects performing better on some rule types and worse on others? One uninteresting possibility is that the items testing some rule types are inherently easier than those testing other rule types, since in some cases different rules are tested by quite different sentential variations. On the other hand, it is clear that this is not the whole account of our effects. Rule types tested in very similar ways on our test (e.g., the various rule types involving morphology) did not show similar degrees of difficulty for late learners, suggesting that these rule type effects are not due to the difficulty of the format by which we tested the rules.

A second possibility is that the subjects suffered from phonological difficulties which made the items for that rule type difficult to process. Again, although we cannot definitely eliminate this possibility, we do not believe it is the whole account of the rule type effects. Rule types with exactly the same phonological form (e.g., plural and third person singular, both -s) did not show similar degrees of difficulty for late learners. Also, rule types testing forms which were phonologically more substantial and

[8] Some other rule type scores also benefited from subjects' apparent ease with basic word order rules. For example, those items testing yes/no question formation by presenting questions in a V-N-N order (e.g., "Learns Jane math from Mr. Thompson?") were particularly easy for subjects. This pattern fit in with a general tendency for V-first items to be easily judged ungrammatical.

therefore easier to hear (e.g., rule types with whole words reversed or eliminated) were not necessarily easier for late learners than those that involved smaller phonological units.

A third possibility is that subjects suffered from interference from the nature of their first language (Chinese or Korean), and so should show special difficulty with rule types most different from the first language. Once again this did not appear to be the full account of our effects. Rule types equally absent from Chinese and Korean (e.g., past tense and present progressive) did not show similar degrees of difficulty for late learners.

Most important, our rule type ordering corresponds in certain striking ways to the order of difficulty obtained in studies of second language learners from other first language backgrounds, as well as in studies of the isolated girl Genie. In particular, the relative ease of word order and the present progressive show up in all of these studies. We believe, then, that the rule type effects we obtained are at least in part reflections of what is generally difficult or easy for a late learner. We will return to this issue in the Discussion section below.

One final question involved the relationship between age of arrival and each of the individual rule types. Given that late learners' competence varies over rule types, it is of interest to know whether age of arrival predicts performance on only certain selected rules of the second language. The data show, however, that this is not the case. Table 7 presents the correlations between age of arrival and the scores on each of the 12

TABLE 7
Correlation Coefficients between Age of Arrival and Rule Type[a]

Rule type	Correlations
Determiners	.64**
Plural	.75**
Subcategorization	.53**
Past tense	.79**
Pronouns	.73**
Particles	.44**
Auxiliaries	.45**
Third person singular	.29*
Wh-questions	.39**
Yes/no questions	.50**
Word order	.48**
Present progressive	.32*

[a] These correlations, unlike others with age of arrival, are positive correlations, since they relate age of arrival to number of errors.

* $p < .05$.

** $p < .01$.

rule types. Despite late learners' proficiency on some rule types, all of the rule types showed significant correlations with age of arrival. This result shows that age of exposure to the language affects all of the structures we examined, despite variations across rule types in the absolute level of performance late learners achieved.

DISCUSSION

This study was designed to answer certain empirical questions about critical period effects in second language learning, and thereby to clarify and refine theoretical proposals regarding a critical period for language acquisition more generally. We will begin our discussion by reviewing the empirical findings, and then turn to the general theoretical issues.

The Basic Empirical Findings

Age of acquisition and ultimate performance. The first question we asked was whether there was a relation between age of acquisition and ultimate performance in the grammar of a second language. The results of this study clearly show such a relation, and therefore support the notion that children have an advantage over adults in acquiring a second language. The overall correlation between age of arrival in the United States and performance on our test of English grammar was −.77; and, for those subjects arriving in the United States before puberty, this correlation was −.87. Indeed, there was a significant correlation between age of arrival and performance on every type of syntactic and morphological rule we tested.

These findings are in accord with the results of the previous studies which have tested asymptotic performance, despite the fact that these studies used very different measures of English proficiency. (Oyama, 1978, measured number of words detected through white noise; Patkowski, 1980, measured syntactic ratings of production). The present study enhances the previous studies' findings by providing a much more detailed examination of English syntax and morphology. The three studies, however, complement each other well, for each emphasizes a different aspect of language use. Oyama's study, for example, taps some aspect of on-line processing in comprehension, Patkowski's measures free production, and in our study we presume to be measuring underlying grammatical competence via sentence judgments. Because these studies complement each other, the compatibility of the results is all the more impressive. Together they provide a strong case for the conclusion that children are indeed better than adults in their ultimate attainment in a second language.

The effects of age of acquisition before vs. after puberty. The second question we asked concerned the shape of the relationship between age of

acquisition and ultimate performance. Due to the large range of ages in the learners we tested, and our division of the early learners into small age groups, we are able to make a fairly good generalization about the shape of this relationship. Subjects who arrived in the United States before the age of seven reached native performance on the test. For arrivals after that age, there was a linear decline in performance up through puberty. Subjects who arrived in the United States after puberty performed on the average much more poorly than those who arrived earlier. After puberty, however, performance did not continue to decline with increasing age. Instead, the late arrival group, while performing on the whole more poorly than the younger arrivals, distinguished itself by having marked individual differences in performance, something which was not found in the earlier arrivals.

The pattern of this relationship supports a maturational account of the age effects found. It does this by the fact that the age effect is present during a time of ongoing biological and cognitive maturation and absent after maturation is complete (i.e., at puberty). Thus it appears as if language learning ability slowly declines as the human matures and plateaus at a low level after puberty. The precise level of this plateau differs between individuals.

Again, these findings are in line with previous studies, although no previous study has asked this question in detail. Both Oyama (1978) and Patkowski (1980) reported only overall correlations and grouped means, with groupings which were larger and slightly different from our own and with a more limited range of ages of acquisition than our own. Both studies found the general linear decline of performance with age of acquisition that we found, but the groupings of their subjects make it difficult to tell whether the precise ages at which we found changes in the function are supported by their results as well. In addition, in a study of age of acquisition in relation to first (rather than second) language acquisition, Newport and Supalla (1987; Newport, 1984) found a linear decline in ultimate performance over three age groups: subjects exposed to American Sign Language from birth vs. at age 4–6 vs. after age 12.[9] In short, the surrounding literature on both first and second language acquisition appears to be generally consistent with the more detailed results obtained in the present study.

[9] One discrepancy between the Newport and Supalla results for first language acquisition and the present results for second language acquisition is in the level of performance attained by subjects who began learning the language between the ages of 3 and 7. In the Newport and Supalla data, the 4–6 age group performed consistently, although not always significantly, below natives. In the present study, the 3–7 age group was entirely within native performance. This difference will be discussed below, in the section entitled "The age at which a decline in performance is first detected."

Experiential and attitudinal variables. The third question we asked was whether the effects of age of acquisition could be due to experiential or attitudinal variables coincidentally related to age, rather than to maturational changes in language learning. Our results suggest that entirely nonmaturational explanations for the age effects would be difficult to support. Certainly the attitudinal variables (motivation, American identification, and self-consciousness) were unable to explain away the age effects, in accord with Oyama's (1978) study. This held true in the present data even when all three variables together were pitted against age.

It is also doubtful that the age effects are the result of differences in the amount of English exposure between the younger and older arrivals. This is true for several reasons. First, the younger arrivals did not differ significantly, if at all, from the adult arrivals in the amount of English they were exposed to during learning (see the Method section for description of the subjects' experiential characteristics). Second, the nonsignificant correlation between amount of initial exposure and performance suggests that second language learning is not particularly sensitive to differences in the amount of exposure, at least when that exposure has occurred over a number of years and is fairly high in the first place.

Some researchers have claimed that there are differences in the quality of the exposure that adults and children receive, rather than in the mere quantity, and that this difference may account for the age differences found in language learning success. According to this view, children receive the ideal type of input for successful language learning, while adults do not. Many have said, for example, that children receive "simple," reduced input which refers to concrete objects, existing in the here and now. Adults, on the other hand, are exposed to syntactically more complex input which most often refers to abstract concepts and events that are displaced in space and time. The simple concrete input of the child is thought to be helpful for language acquisition, while the complex input of the adult is thought to interfere with language acquisition (Dulay, Burt, & Krashen, 1982).

Both the empirical and theoretical assumptions underlying this approach have been disputed. First, the assumption that language is easier to learn from limited simple input has been questioned (Wexler & Culicover, 1980; Gleitman, Newport, & Gleitman, 1984). Second, the empirical evidence for this claim has also been brought into question. Freed (1980) performed a study which compared the type of input given to adult and child second language learners and found that adults and children actually receive comparable input in terms of syntactic complexity (as measured by the number of clause boundaries). Interestingly, however, the adult-directed input contained a more limited range of constructions than the child-directed input. Adults received input which tended to main-

tain the canonical shape of a sentence, while children received sentences with more deformations. Thus in terms of transformational complexity, adults received the simpler input. From this it would be just as reasonable to argue that adults learn less well because their input is not as complex and varied as the child's. In any case, the role of input in second language learning needs to be better formulated before we can decide whether children have any advantage in learning a language due to the type of input they receive.

Age of acquisition and rule type. The fourth question we asked concerned the nature of the effects of age on the attained grammar of the second language. Our results suggested that, although there was an effect of age of acquisition on every rule type we examined, some rules of English grammar were more profoundly affected by age of acquisition than others. In particular, knowledge of the basic word order of the language was acquired by all of our subjects, regardless of their age of learning. Similarly, knowledge of the present progressive (-ing) was acquired by all of our subjects. These areas of competence likewise appear in other studies of second language learning (see Krashen, 1982, for a review of the order of morpheme difficulty in second language learning). Perhaps even more striking, they are the only two aspects of English which were successfully acquired by Genie, who was exposed to English as a first language only after puberty (Curtiss, 1977; Goldin-Meadow, 1978). In contrast, other aspects of English syntax and morphology gave late learners much more difficulty.

We believe that these rule type results are at least in part reflections of universal factors in learnability, and not merely the result of item difficulty or transfer from the first language. Newport, Gleitman, and Gleitman (1977; Gleitman, Newport, & Gleitman, 1984) and Goldin-Meadow (1978) have suggested that basic word order is a highly resilient property of languages, appearing in the acquisition of a first language under widely varying conditions of both input and age of exposure. The present results on the acquisition of a second language under varying conditions of age of exposure are in accord with these claims. However, accounting for why word order and -ing are particularly easy for learners remains for future research.

Before turning to a more general discussion of critical period hypotheses in language learning, we must consider whether the set of results we have obtained will be replicable on other second language learning groups or whether they are confined in any way to the particular second language learners (Chinese and Koreans) we have studied.

Possible effects of the first language on second language learning. We have thus far presented our results as though the findings were generalizable to second language learning, regardless of the nature of the first

language or the relationship between the grammar of the first language and that of the second language. Indeed we believe this is the case, although we also recognize that certain aspects of the structure of one's first language are likely to have some effects on the learning of the second language (see, for example, Zehler, 1982, and Hakuta & Cancino, 1977, for a review of transfer effects in second language learning). Here we wish to raise two points of relevance to the question of whether our results are limited in any way to the Chinese and Korean speakers we studied.

First, we do not believe that the relationship found here between age of exposure and ultimate performance in the second language is unique to the circumstances where Chinese or Korean is the first language and English is the second. We did purposely choose to concentrate on first and second languages where the grammars were sufficiently different that a significant second language learning problem would arise. Chinese and Korean are relatively more isolating languages than English and have syntaxes which are different in many ways from that of English. However, studies currently underway, as well as certain details of our present results, suggest that the basic findings do not depend on these particular language combinations.

Several studies in progress (Johnson & Newport, in press) examine performance on our test by subjects with a wide variety of first languages. It is too early to say from these data whether there is any effect of the nature of the first language (we expect that there might be); however, it is already clear that the strong correlation between age of arrival and test performance replicates with subjects from these other first-language backgrounds.

In addition, the detailed results of the present study suggest that the nature of the first language cannot fully explain the difficulties of the second language learner. The examination of performance on the 12 rule types reveals relationships to age of arrival on every structure we examined, regardless of how similar or different these structures were to ones in the first language. For example, determiners and plural inflection, which gave late learners their most serious difficulties on our test, are notably lacking in Chinese and Korean; but so are inflections for the present progressive, on which late learners performed exceptionally well. A more detailed understanding of which of our effects, if any, may arise from first language characteristics should emerge from our studies in progress.

Second, we do not believe that our results derive in any important way from the input or cultural circumstances which characterize Chinese and Korean speakers. The Chinese and Korean speakers we tested were perhaps unusual, compared with many second language learners of English, in that they often continued close associations with other speakers of their

first language. One might wonder, therefore, whether their exposure to English or their maintenance of their first language somehow influenced their second language learning. Again, this is an empirical question which is best resolved by the results of our studies in progress, which include many speakers isolated from their first language group as well as speakers of first languages with large communities. Within the present study, all of our subjects (both early and late learners) continued speaking their first language with their families and others into adulthood, and all were exposed to English from native English speakers. In addition, all had a significant amount of exposure to English, since they were all active members of an English-speaking community (that is, American schools and universities). These factors therefore could not be responsible for the differences we found between early and late learners of English. Whether these factors have an additional effect on learning, beyond the effect of age of exposure, was not the focus of our study, although some of our results do bear on this question.

In sum, we believe that in other language groups the strong effects of age of acquisition may be accompanied by effects of input, first language typology, or other variables that do not appear in our data on Chinese and Korean learners. Most importantly, however, we have reason to expect, on the basis of our data, that these effects of age of acquisition will persist.

Theoretical Conclusions for a Critical Period Hypothesis in Language Acquisition

The present study was performed primarily for the purpose of understanding the nature of the critical period for language acquisition. In particular, we wanted first to discover whether the critical period occurs at all in second language acquisition or whether it is exclusively a first-language phenomenon. To delineate this distinction we began by presenting two possible versions of a critical period hypothesis. They are repeated here for convenience.

Version One: The exercise hypothesis. Early in life, humans have a superior capacity for acquiring languages. If the capacity is not exercised during this time, it will disappear or decline with maturation. If the capacity is exercised, however, further language learning abilities will remain intact throughout life.

Version Two: The maturational state hypothesis. Early in life, humans have a superior capacity for acquiring languages. This capacity disappears or declines with maturation.

To reiterate the differences between these two versions, the exercise hypothesis only requires that a first language be acquired during child-

hood; as long as this occurs, the capacity for successful language learning will remain intact. Thus it predicts no differences between child and adult learners, due to maturation itself, in the ability to acquire a second language to native proficiency. In contrast, the maturational state hypothesis says that any language, be it first or second, must be acquired during childhood in order for that language to develop fully. Our results support the maturational state hypothesis, and not the exercise hypothesis. Human beings appear to have a special capacity for acquiring language in childhood, regardless of whether the language is their first or second.

The maturational state hypothesis is, however, not itself an explanation of critical period phenomena in language; rather, it merely outlines a class of explanations which would be compatible with our results (namely, those which posit maturational changes in general language learning abilities). In order to approach a more precise theoretical account of the phenomena, our study has also provided additional information which should aid in understanding the nature of the critical period: namely, information about the shape of the function relating age of acquisition and ultimate performance. Our results provide three sets of facts which any theory regarding critical periods would have to account for: the gradual decline of performance, the age at which a decline in performance is first detected, and the nature of adult performance.

The gradual decline of performance. Lenneberg's original proposal of a critical period in language acquisition seemed to predict a rectangular function in the relationship between age of acquisition and ultimate performance. That is, Lenneberg hypothesized that "normal" language learning was possible during the period from infancy to puberty, with a loss of abilities after puberty. However, the data on second language learning in the present study did not have this shape. We did not find a flat relationship between performance and age of learning throughout childhood, with a sudden drop in performance marking the end of the critical period; instead, performance gradually declined from about age seven on, until adulthood. Insofar as such data are available from other studies of first and second language acquisition, the same linear trend seems to appear (Oyama, 1978; Patkowski, 1980; Newport, 1984; Newport & Supalla, 1987).

Although this gradual decline is not in accord with Lenneberg's implied function, it is in accord with results from other behavioral domains in which critical periods have been hypothesized. As research accumulates on critical periods, whether it be in imprinting in ducks (Hess, 1973), socialization in dogs (Scott, 1978), or song learning in birds (Kroodsma, 1981), it is becoming apparent that most, if not all, critical periods conform to the more gradual function. This point has recently been noted by several investigators (Tanner, 1970; Immelman & Suomi, 1981).

> . . . usually these periods consist of . . . beginning and end parts . . . [during] which the organism is slightly sensitive to the specific influence, with a period of maximum sensitivity in the middle. It is not as a rule an all-or-none phenomenon. (Tanner, 1970, p. 131).

Whatever mechanisms underlie a critical period effect in language learning, then, must be compatible with this gradual decline of performance over age.

The age at which a decline in performance is first detected. Lenneberg's proposal also seemed to imply that a decline in performance should first appear at puberty. Instead of puberty, we found a small but significant decline in performance in subjects who had arrived in the United States as early as age 8 to 10. Indeed, the only discrepancy we know of between our results and other data is that, in first language acquisition, this decline may occur even earlier (Newport, 1984; Newport & Supalla, 1987); in the Newport and Supalla data, a 4–6 age group scored consistently, although not always significantly, below native performance. It is possible that a similarly early decline may occur in second language acquisition as well on a test that included more complex aspects of syntax than our own; on our present test, the age 3–7 group scored at ceiling.

Further research is therefore necessary to determine with certainty the exact point at which a decline in learning begins for second language acquisition. It is clear from the present data, however, that this decline begins well before puberty. It also appears that this early decline is small, and that another more major change occurs around puberty. Proposed mechanisms underlying a critical period effect in language learning must therefore account for the details of timing of these changes and, particularly, for the fact that the decline in learning ability begins earlier than initially thought by most researchers.

The nature of adult performance. There are two aspects of adult performance with which any theoretical account of the critical period must be compatible. The first is that language does not become totally unlearnable during adulthood. This has held true in all of the studies which have tested age differences in asymptotic performance, including both first and second language learning. In the present study, late learners scored significantly above chance on all of the rule types tested except for determiners and plurals. It appears to be the case, then, that quite a few aspects of language are learnable to a fair degree at any age, even though deficiencies in this learning occur.

The second aspect of adult performance with which any theory must be compatible is the great variability found among individuals. For adult learners, age does not continue to be a predictor of performance; thus any proposed mechanism accounting for adult performance likewise cannot be correlated with age. Moreover, while early learners are uniformly

successful in acquiring their language to a high degree of proficiency, later learners show much greater individual variation (see also Patkowski, 1980, for related comments). A theoretical account of critical period effects in language learning must therefore consider whether the skills underlying children's uniformly superior performance are similar to those used by adult learners, or rather whether adult language learning skill is controlled by a different set of variables.

Final remarks on a critical period theory of language acquisition. In sum, we now have a number of findings which should be accounted for in any explanation of a critical period. There is the nature of the relationship between age of arrival and performance: a linear decline in performance up through puberty and a subsequent lack of linearity and great variability after puberty. There is also the pattern of errors found for the wide range of aspects of syntax and morphology of English studied: age effects were found for every rule type, with low levels of performance on every rule type except word order and present progressive. The primary and most general finding to accommodate for any critical period theory, of course, is that the critical period is not just a first language phenomenon, but extends to a second language as well.

These findings rule out certain types of accounts of a critical period for language acquisition and make other types of accounts more plausible. We have suggested that our results are most naturally accommodated by some type of maturational account, in which there is a gradual decline in language learning skills over the period of on-going maturational growth and a stabilization of language learning skills at a low but variable level of performance at the final mature state. This leaves open, however, the precise explanation of such a phenomenon. The traditional view of critical period effects in language learning has been that there is maturational change in a specific language acquisition device (Lenneberg, 1967; Chomsky, 1981). Such a view, with some modifications to incorporate the detailed points of maturational change, is consistent with our results. Also consistent with our results are views which hypothesize more general cognitive changes over maturation (see, for example, Newport, 1984). From this view, an increase in certain cognitive abilities may, paradoxically, make language learning more difficult. We are hopeful that future research will provide more detailed results which may differentiate these views from one another. In any event, the present study makes clear that some type of critical period account for language acquisition is necessary and that the proper account of a critical period will include both first and second language in its effects.

REFERENCES

Asher, J., & Garcia, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal,* 53, 334–341.

Asher, J., & Price, B. (1967). The learning strategy of total physical response: Some age differences. *Child Development*, **38**, 1219–1227.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris.

Curtiss, S. (1977) *Genie: A psycholinguistic study of a modern day "wild child."* New York: Academic Press.

Dulay, H., Burt, M., & Krashen, S. (1982) *Language two*. New York: Oxford University Press.

Freed, B. (1980). Talking to foreigners versus talking to children: Similarities and differences. In R. Scarcella and S. Krashen (Eds.), *Research in second language acquisition*. Rowley, MA: Newbury House.

Goldin-Meadow, S. (1978). A study in human capacities. *Science*, **200**, 649–651.

Gleitman, L. R., Newport, E. L., & Gleitman,H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, **11**, 43–79.

Hakuta, K., & Cancino, H. (1977). Trends in second language acquisition research. *Harvard Educational Review*, **47**, 294–316.

Hess, E. H. (1973). *Imprinting*. New York: Van Nostrand.

Hubel, D., & Weisel, T. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, **26**, 994–1002.

Immelmann, K., & Suomi, S. J. (1981). Sensitive phases in development. In K. Immelmann, G. W. Barlow, L. Petrinovich, & M. Main (Eds.), *Behavioral development: The Bielefeld Interdisciplinary Project*. Cambridge: Cambridge University Press.

Krashen, S. (1975). The development of cerebral dominance and language learning: More new evidence. In D. Dato (Ed.), *Developmental psycholinguistics: Theory and applications: Georgetown Round Table on Language and Linguistics*. Washington, DC: Georgetown University.

Krashen, S. (1982). Accounting for child–adult differences in second language rate and attainment. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child–adult differences in second language acquisition*. Rowley, MA: Newbury House.

Krashen, S., Long, M., & Scarcella, R. (1982). Age, rate, and eventual attainment in second language acquisition. In S. Krashen, R. Scarcella, & M. Long (Eds.), *Child-adult differences in second language acquisition*. Rowley, MA: Newbury House.

Kroodsma, D. E. (1981). Ontogeny of bird song. In K. Immelmann, G. W. Barlow, L. Petrinovich, & M. Main (Eds.), *Behavioral development: The Bielefeld Interdisciplinary Project*. Cambridge: Cambridge University Press.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called a grammatic aphasics. *Cognition*, **13**, 361–392.

Newport, E. L. (1984). Constraints on learning: Studies in the acquisition of American Sign Language. *Papers and Reports on Child Language Development*, **23**, 1–22.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

Newport, E. L., & Supalla, T. (1987) *A critical period effect in the acquisition of a primary language*. University of Illinois, manuscript under review.

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

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, **5**, 261–285.

Oyama, S. (1978). The sensitive period and comprehension of speech. *Working Papers on Bilingualism*, **16**, 1–17.

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

Schumann, J. (1975). Affective factors and the problem of age in second language acquisition. *Language Learning*, **2**, 209—235.

Scott, J. P. (1978). Critical periods for the development of social behavior in dogs. In J. P. Scott (Ed.), *Critical periods*. Stroudsburg, PA: Dowden, Hutchinson, & Ross.

Scott, J. P. (1980). The domestic dog: A case of multiple identities. In M. A. Roy (Ed.), *Species identity and attachment: A phylogentic evaluation*. New York: Garland STPM Press.

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

Snow, C. (1983). Age differences in second language acquisition: Research findings and folk psychology. In K. Bailey, M. Long, and S. Peck (Eds.), *Second language acquisition studies*. Rowley MA: Newbury House.

Snow, C., & Hoefnagel-Hohle, M. (1977). Age differences in pronunciation of foreign sounds. *Language and Speech*, **20**, 357–365.

Snow, C., & Hoefnagel-Hohle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, **49**, 1114–1128.

Tanner, J. M. (1970). Physical growth. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*. New York: Wiley.

Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition*. Cambridge, MA: The MIT Press.

Zehler, A. M. (1982). *The reflection of first language-derived experience in second language acquisition*. Unpublished doctoral dissertation, University of Illinois.

(Accepted April 22, 1988)