Second Language Acquisition and the Critical Period Hypothesis

Edited by

David Birdsong
University of Texas

LEA
1999 LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

CHAPTER SEVEN

Confounded Age: Linguistic and Cognitive Factors in Age Differences for Second Language Acquisition

Ellen Bialystok
York University
Kenji Hakuta
Stanford University

THE NATURE OF THE PROBLEM

In spite of what we all learned in our first statistics course, we just cannot resist attributing causality to correlation. We have to remind ourselves every time we see two events contiguously linked in time and space that the most natural explanation for their co-occurrence, namely, that one causes the other, might simply be false. The assumption of causality is one of the basic tenets of commonsense logic: Spring rains lead to flowers, knocking over the juice container results in spilled liquid, and clicking the power button on a small handheld instrument causes pictures to appear on the television screen. We all know, too, that it is counterexamples that compel caution in assuming the interpretation of causality: Superstition notwithstanding, carrying or not carrying an umbrella has no causal consequence for local meteorological conditions.

How are we to discover the correct logical relation between two events that share patterns of occurrence? The simplest explanation, that one event causes the other, is often taken at the expense of details that do not fit easily into the interpretation but are overlooked, set aside, or discounted. Indeed, it was the final effort to deal with the inconsistencies in the Ptolemeic description of planetary motion that led to the overthrow of that explanation, but it took 14 centuries and

countless attempts to patch up the theory before the basic logic was rejected. No doubt one of the reasons that Ptolemy's description endured as long as it did was that prima facie it seemed to be correct. To an observer, it does indeed appear as though the earth is the center of the planetary system. Discovering the correct logical model requires stepping outside of the domain of the immediately perceptual data and imagining alternative explanations that are more subtle, more inaccessible, more indirect.

The problem of discovering the correct explanation for events that appear to have a simple observable relation with each other permeates the inquiry into the relation between age and the ability to learn a second language. Observationally, there is a co-occurrence between two events: The age at which a person starts learning a second language corresponds in some way to the ultimate success that the person will attain after years of having used that language. But are these two events—age and ultimate success—linked causally? Explanations of causality require stronger evidence than co-occurrence.

The critical period hypothesis is a causal explanation for the differential success in acquisition of a second language by younger and older learners. The explanation is causal because the bulk of the variance in achievement as a function of age is attributed to maturational changes in the brain that alter the possibility of successful acquisition. The controversy in the debate over the status of a critical period for second language acquisition has less to do with the documentation of observations than with the interpretation of those data. Are younger learners generally more successful than older ones when ultimate proficiency in a second language is assessed? Yes. Do younger and older learners approach the learning problem differently? Presumably. Are there neurological differences in the brains of younger and older learners? Probably. None of these statements, however, compels the conclusion that there is a critical period for second language acquisition. Similarly, neither the Ptolemeans nor the Copernicans disputed the observation that the sun rose in the east and set in the west; it was their interpretations of those events that were different. To use the terminology of the statistical test, there may well be a correlation between age of initial learning and ultimate achievement, but it does not necessarily follow that age is a causal factor in that relation. It may turn out that it is, but the data would need to show convincing evidence for causality.

Our approach to evaluating the argument for a critical period is to show that age intervenes in the effect that linguistic and cognitive factors have on success in second language acquisition. Therefore, correlations between age and success are spurious because the relation is actually reflecting the effects of these linguistic and cognitive factors. Statistically, this argument could be demonstrated by partialling age

out of the equation and then studying the relation between these linguistic and cognitive factors in the absence of age. If our explanation is correct, then the partial correlations between linguistic and cognitive sources of variance and proficiency should remain significant when age is not included in the equation. Alternatively, if it could be shown that linguistic or cognitive factors (or social, although we do not discuss these) were capable of producing patterns of results that are sometimes attributed to age differences, then the role of age in explaining these effects would need to be reconsidered. Our approach, however, is to offer data that challenge the interpretation that the effects are caused by age by identifying areas in which empirical results contradict predictions from the critical period hypothesis.

The debate over the critical period hypothesis embodies some of the most basic questions about second language acquisition, and indeed, language acquisition in general. These questions permeate the foundations of several disciplines, such as linguistics, cognitive psychology, and neurolinguistics. Is language learning governed by environmental conditions or by an internal bioprogram? Do languages reside in independently constructed mental representations or are they mutually available in processing? Is transfer a legitimate process in language learning or an unwanted symptom of the improper separation of distinct languages? To some extent, the answers to these and other fundamental questions in human language learning rest partly in the role that age plays in acquiring languages. If there is a critical period for second language acquisition, then logically there is also one for first language acquisition, and the answers to questions about language processing take a clear direction. One must be prudent, therefore, in accepting the hypothesis for a critical period in second language acquisition. Methodologically, one must begin with the null hypothesis that no such limitation exists and produce reasons why this hypothesis should be rejected.

CHARACTERIZING CRITICAL PERIODS

What would constitute evidence for a critical period? Consider the following three definitions that have been offered:

> During select times in the life cycle many structures and functions become especially susceptible to specific experiences (or to the absence of those experiences) in a way that alters some future instantiation of that (or a related) structure or function. (Bornstein, 1989, p. 179)

> Certain environmental events must happen at certain times in the development of an organism in order for normal development to

occur. (Gazzaniga, 1992, p. 56)

> Any phenomenon in which there is a maturational change in the ability to learn, with a peak in learning at some maturationally definable period ... and a decline in the ability to learn, given the same experiential exposure, outside of this period. (Newport, 1991, p. 112)

In addition, Colombo (1982) and Bornstein (1989) both identified criteria that need to be specified in learning that is considered to be constrained by a critical period. These criteria include onset and offset times for the period, as well as other factors that characterize the nature of the learning during the critical period. Two points recur throughout all these definitions. First, learning during a critical period is assured, similar across individuals, normatively described, and probably governed primarily by endogenous factors. Exogenous factors, therefore, should have minimal impact on this learning. Second, learning outside of the critical period is different in both form and success, especially in that it would be less certain and more erratic in its outcomes. Therefore, there should be a clear discontinuity between these two types of learning, and the time of that discontinuity should reflect the close of the critical period.

Some researchers have tried to take a moderate position by positing a weakened version of the critical period hypothesis. These positions are often signaled by terminological choices, notably, the use of *sensitive period* instead of *critical period*. Colombo (1982) discusses the reasons why this distinction has failed to clarify the issues, primarily because of the difficulty of classifying phenomena as being one or the other and the lack of evidence that the two phenomena were different from each other. Similarly, some attempts have been made to weaken the conditions that make critical periods distinct learning situations. If a critical period is considered to be simply a period of heightened sensitivity that can be overcome outside the period, as some accounts posit, then there is almost no doubt that there is a critical period for language acquisition, but by these standards, there would be a critical period for virtually everything we learn (baseball, music, and calculus being examples).

BUT COMPARED TO WHAT?

Our discussion proceeds by examining the role that some linguistic and cognitive factors play in second language acquisition and considering how age might interact with these factors. But first, we need to know what the rules are. What is it we are trying to explain? What do we mean by proficiency in a second language?

Pinker (1994) recounted the story of Dizzy Dean, a 1950s baseball announcer, who routinely described such plays as, "He slood into second base." Mr. Dean was a native speaker of English, but in his home state of Arkansas, dialectal peculiarities such as these were the standard. What is native speaker proficiency? Although this case may seem extreme, it is only a progression on a continuum of variation in language use.

There is an assumption in all research into second language acquisition that the learner is striving toward some stateable goal, a standard and perfect version of the language that is embodied in the mind of every native speaker. Chomsky (1957) formally acknowledged this idealization as linguistic competence and quickly discounted the likelihood that it would ever be produced by real speakers (Chomsky himself notwithstanding) because of the sobering reality of performance that prevents mortal humans from achieving that level of perfection. For that reason, most linguistic research is based on speaker judgments and not speaker performance because, the argument goes, judgments can be made solely from competence whereas performance cannot. But how would Mr. Dean judge his own sentence describing the runner's arrival at second base? Indeed, native speakers do not perform judgment tasks with 100% accuracy. What do we mean, then, when we speak vaguely of second language learners achieving native-like proficiency? This problem of designating a standard linguistic form is evident at all levels of analysis, but phonology is perhaps the most salient.

In addition to the problem of determining a standard for correctness is the problem of scope and generalizability. On the basis of some local assessment, conclusions are made about general competence, or language proficiency. What kind of assessment legitimately supports such claims? It depends in large measure on the nature of the hypothesis being tested. A theory about the process of second language acquisition, for example, should lead to specific predictions about acquisition that could be tested by detailed analysis of linguistic structures. Such theories, therefore, can be supported through a few discrete linguistic features. A theory about a critical period, however, may require more broadly based evidence covering many aspects of language proficiency. There is an inherent tension between the need to choose measures that are narrowly focused on the theoretical dimension of interest on the one hand, and the need to use measures that are global and ecologically more valid on the other.

Research into the critical period for second language acquisition has made use of a range of outcomes. The most sharply specified are the variables defined by Universal Grammar (UG), the putatively abstract and unlearnable elements of human language, such as subjacency and the complex noun phrase constraint (e.g., Johnson & Newport, 1991 Juffs &

Harrington, 1995; Martohardjono & Flynn, 1995). The idea is that these principles are part of the biological language program that constrain the hypotheses learners are able to construct about grammar. If learners lose access to this bioprogram, then presumably they lose access as well to the specific grammatical hypotheses that follow from these constraints of UG, making it difficult or even impossible to discover such rules naturally. UG, then, is endowed with a level of reality that virtually moves it into the realm of concrete rules rather than abstract constraints. It should be particularly troubling to such theories, then, when a recantation of those constraints is proclaimed, as Chomsky (1995) recently did.

Another kind of outcome is defined by grammatical rules that do not necessarily require formal grammatical theory for explanation. Johnson and Newport (1989), for example, examined 12 rule types, including past tense, plurals, and third-person singular verb. Violations of these grammatical rules were created by omitting the required morpheme, replacing the required morpheme with an inappropriate morpheme, making an irregular item regular, or by attaching a regular marking to an already irregularly marked item. These rules could be abstract in the sense that they are part of a general theory of abstract grammar, but they can also be explained through nonlinguistic models rooted in cognitive analysis. Thus, outcomes defined by these rules are ambiguous with respect to the language specificity of the phenomenon.

A third kind of outcome is global assessment of some aspect of proficiency. For example, Patkowski (1980) asked trained judges to rate the overall syntactic proficiency of transcripts of tape-recorded narratives by second language learners. Oyama (1976) also recruited raters who listened to tape-recorded narratives and gave judgments of fluency. More recently, Bongaerts, Planken, and Schils (1995) elicited raters' judgments of learners' degree of foreign accent. In such studies, it is the overall proficiency that is being judged and as such, probably comes closest to a commonsense definition of language proficiency. Although the reliability of ratings and the criteria used to generate them can be questioned for their scientific authority, the evaluations are high in ecological validity.

The choice between precise specification of learning outcomes and the ecological validity of second language acquisition (L2A) offers an important methodological lesson for researchers. To the extent that a theory has explanatory precision, it is best served by testing for specific structures. For example, if the theory is that UG governs second language acquisition until puberty and then becomes unavailable, then UG-based structures are the prized items to be examined. Discovering age-related effects of non-UG structures may be problematic for a theory of UG but may fuel the development of various alternative

theories. The nature of the linguistic data is critical in setting out the possible interpretations that may follow from those data, regardless of how the results turn out.

LINGUISTIC CONSIDERATIONS

If language is represented as innate abstract principles and there is a critical period for language acquisition, then L2A during the critical period should resemble first language acquisition (L1A) because both processes are governed by the learner's access to those principles. Therefore, L2A during the critical period should show little or no effect of transfer from the first language because direct access to UG should override cognitive intervention in the process of constructing the system of rules for the second language. Learning after the critical period, however, would reflect elements of the first language because general cognitive resources would be recruited to construct the linguistic system, and they would naturally begin with the linguistic structures already in place. Demonstrating different types of language transfer before and after the close of the critical period, therefore, would support the argument for a critical period in L2A.

Historically, evidence for language transfer has been one means of explaining the uniqueness of L2A and was used as the empirical method in early research on this problem to define that difference (Hakuta & Cancino, 1977). If L2A were the same as L1A, it was argued, then the process was largely a linguistic development. Whatever was responsible for the child's assured access into the arcane world of abstract rules and representations would equally guide the second language learner into proficiency. Furthermore, the prevailing linguistic theory that posited universal structures that were wired into the child made neurological factors an essential aspect of first language acquisition. However, if the course and outcomes of L2A were considered to be importantly different from those of L1A, then other kinds of factors, notably cognitive and social ones, needed to be invoked. Empirical evidence attempting to pronounce on this matter turned out to be largely equivocal: L2A was exactly like L1A in some ways and completely unlike it in others. Accordingly, both linguistic-neurological and cognitive-social explanations were going to be needed. The critical period debate entails a return to some of these arguments. If transfer from the first language is discovered to characterize learning even for the youngest learners, then some of the responsibility for second language learning would need to be reassigned to these other factors.

Although it is true that transfer distinguishes L2A from L1A in some respects, it is not clear that the process itself is unique to L2A. A major aspect of children's development consists of their connecting linguistic

competence with conceptual knowledge. In this sense, children's L1A also involves something like transfer from cognitive structures to linguistic ones. Transfer, that is, can be considered to be a much broader process than just the extension of linguistic structures from one language system to another. It also involves the generalization or use of knowledge from one domain into another. To what extent does this cognitive interpretation of transfer in language acquisition apply to the kinds of transfer observed in second language acquisition?

Consider first the kinds of transfer that can be observed in L2A. Transfer has been reported at different levels of linguistic analysis, described earlier as either comprising part of the abstract rules of UG (e.g., subjacency constraint), or surface structure similarity between two languages (e.g., negation, determiners). Transfer has also been detected in semantic interpretations of individual words (e.g., Ijaz, 1986; Kellerman, 1986). These examples fall along a continuum from abstract linguistic structure to cognitive conceptualization. In the first case, the learner is drawing on prewired constraints of UG that characterize the structure of the first language to formulate utterances; in the second, the learner is using knowledge of a structure and applying it to the L2 on the hypothesis that the two will be similar. These processes are different from each other in many respects: They are based on different kinds of mental representations, they involve different degrees of intentionality by the learner, and they are differentially susceptible to variation in the specific language pairs. Nonetheless, they all occur during the construction of an L2. Is transfer, therefore, a linguistic process or a cognitive process?

The important evidence from transfer for the critical period hypothesis, however, comes from the interaction between the type of transfer observed and the maturational stage of the learner. If there is a language learning faculty that undergoes change as a function of maturation, neurological development, or atrophy, then over time the transferred structures would presumably shift away from abstract linguistic principles toward more surface features or cognitively determined structures. This would reflect the move away from the control over language acquisition residing in a specific language center that is both formally (i.e., neurologically) and functionally (i.e., language acquistion device) defined to more general cognitive processes. Consequently, as access to UG becomes weaker, L2 learners' intuitions about the new language will rely less on the constraints of UG that were set for the L1, decreasing transfer of these abstract principles into the L2. This may be compensated by an increasing reliance on transfer effects based on language-specific features. Empirically, the important observation would be a qualitative shift in the extent or nature of transfer from the L1 at different maturational stages of second language acquisition.

The evidence on this point suggests that it is not the case. For example, Juffs and Harrington (1995) found as much transfer of subjacency from older and younger Chinese learners of English. Both groups performed well on a judgment task assessing their mastery of English subjacency, but all the learners took significantly longer to make these judgments than did native speakers. On aspects of linguistic structure that were less constrained by UG, that is, more along the dimension toward surface rules or cognitive regularities, Bialystok and Miller (1998) found no change as a function of the transfer of six structures from Chinese to English. As shown in Fig. 7.1, both younger and older learners made more errors in a sentence judgment task on items containing grammatical features that were different between Chinese and English than on items containing grammatical features that were

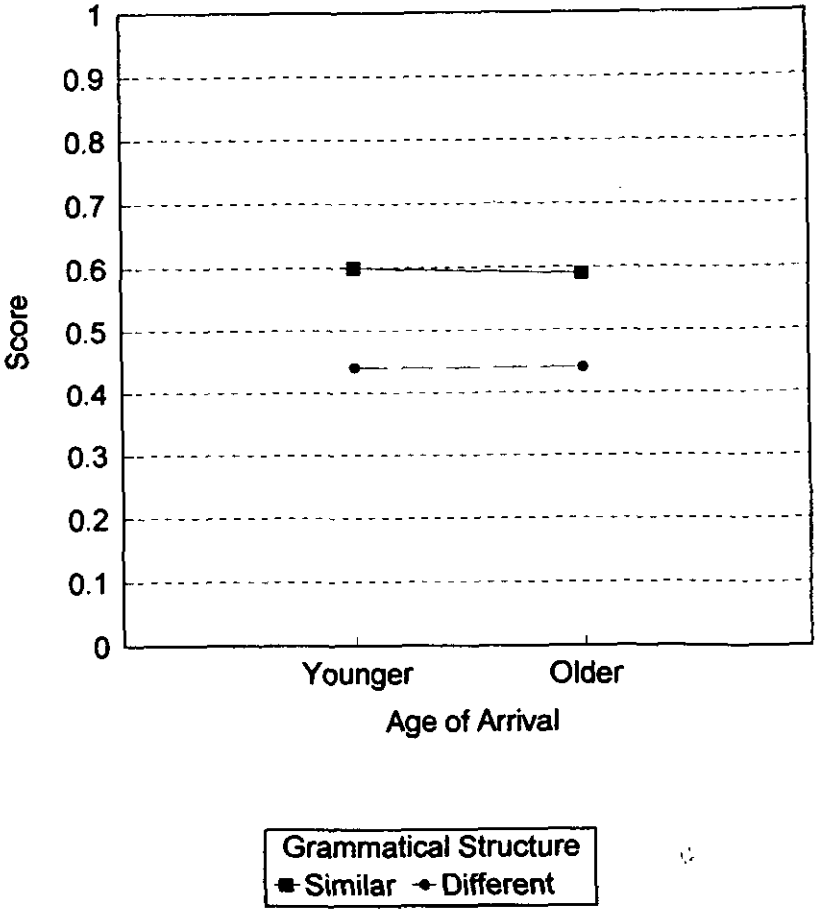

FIG. 7.1. Mean score for Chinese bilinguals by age of arrival.

similar in the two languages. Similarly, older and younger Spanish-speaking learners of English had more difficulty in judging sentences containing an error in a structure that was different between the languages than in judging sentences with errors that were common. The younger learners performed at a higher level than older ones, but the pattern was exactly the same. In other words, even though the amount of first language interference was different for younger and older learners, the nature of the interference was the same. These data are plotted in Figure 7.2. The results of the study by Johnson and Newport (1989) also support the position that older learners transfer more than younger ones in absolute terms. However, accepting the experimental hypothesis for a critical period requires evidence of a discontinuity in the quality of rules that are transferred within and outside of that

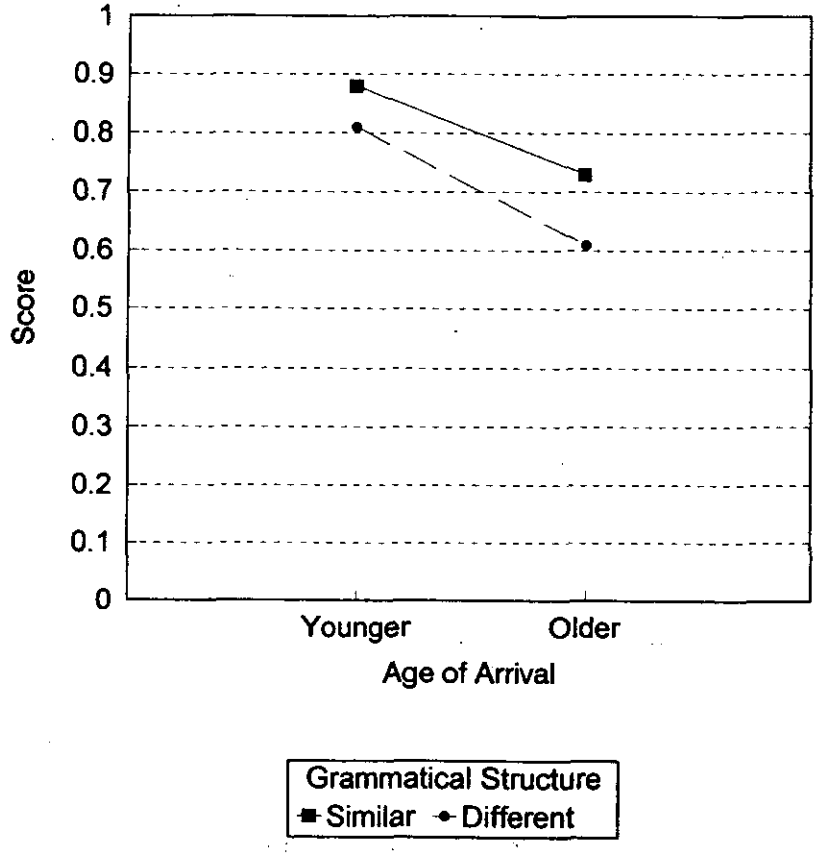

FIG. 7.2. Mean score for Spanish bilinguals by age of arrival.

period. No such discontinuity has been found (Bialystok & Hakuta, 1994).

COGNITIVE CONSIDERATIONS

Even for those theorists who view language as an independently functioning module, governed by domain-specific principles and acquired through dedicated mechanisms, it overflows at every turn into the realm of cognition. Indeed, it was Chomsky (1957) who made the study of language a cognitive problem and unleashed a shift in psychological theorizing that has come to be known as the "cognitive revolution." But how is cognition implicated in the debate over the existence of a critical period for second language acquisition?

In spite of the degree to which language acquisition may be governed by innate principles, aspects of language learning and use are clearly beyond the reach of such dedicated modules. The acquisition of literacy, for example, inflicts permanent change on children's conceptions of language. We know that children who learn to read in alphabetic scripts develop more sophisticated conceptions of phonological structure, and all children, irrespective of the language they read, advance rapidly in their metalinguistic concepts as literacy is established (review in Adams, 1990). To the extent that literacy is a factor in second language acquisition, that aspect of the process must be considered to be controlled by cognitive and not purely linguistic mechanisms. Although there is little research into the role that literacy plays in second language acquisition, some inferences are possible. For example, certain forms of instruction are possible with literate, that is older, learners that are unavailable to preliterate or younger learners. Different instructional forms could lead to differences in proficiency. The literacy factor might also influence the outcomes of language acquisition in situations of immigration, a common population for critical period studies, where such differences as literacy of the learners, availability of written texts, opportunity for instruction, and other such factors influence the proficiency of the learners. In general, younger immigrants would likely attend schools in the host language and learn the literate grammatical forms through texts as part of their curriculum. Some older learners, especially those without strong cultures of literacy, may not have access to these standard written forms. It would not be surprising if the eventual attainment of those immigrants who had attended schools in the target language surpassed in large measure that of their parents. However, these factors are rarely discussed in the literature, and so demonstrations of simple age-related differences in ultimate proficiency do not determine the cause of those differences.

Another example of the cognitive influences on assessments of second

language proficiency, if not the process of acquisition itself, can be seen through differences in performance that are attributable to testing methods. In a replication of the study by and Newport (1989), Johnson (1992) gave the same sentences to the same participants but used written presentation instead of the original oral format. Not only did participants perform at a higher level overall, but there were fewer structures for which differences in age were relevant. In other words, evidence for age-related differences in learning depended not only on which structures were being examined but also on the testing modality. This difference in modality, in which written presentation elicits higher levels of performance than oral ones, was also found in the studies by Bialystok and Miller (1998). Why would this be the case? It is possible that here, too, some effects of literacy emerge in the results. Again, for age to be a main effect and be credited with the explanatory power in these results, the role of testing method needs to be clarified.

If second language acquisition is under the control of cognitive processes that are not unique to a language learning module, then the age-related changes in ultimate proficiency must be explained to some extent by changes in these general cognitive mechanisms. Because ultimate proficiency declines with age of initial acquisition, these general cognitive mechanisms must also deteriorate in their efficiency or effectiveness to serve as part of the explanation for changes in proficiency. There is evidence from studies in lifespan cognition that exactly this sort of deterioration takes place (Schaie & Willis, 1991). In a paired-associate task (like vocabulary acquisition), older learners were more sensitive to timing factors in the presentation of the material and required longer intervals than younger learners to recall the same pairs (Craik, 1986). Older learners were also more cautious and unlikely to venture a response if they were unsure of its correctness (Birkhill & Schaie, 1975). The encoding stage of establishing long-term memory also took longer for older learners, and they required more trials to learn the list (Rabinowitz & Craik, 1986). There is also a decline across the lifespan in the ability to recall details, and as learners aged they increasingly remembered only the gist (Hultsch & Dixon, 1990). These are examples of declining cognitive functions that take place across the lifespan. All of these abilities are involved in learning and using language, so their decline would adversely affect the ability to learn a new language. However, the decline of these functions is gradual and constant. No one has ever suggested that there is a critical period for memory and cognition. Therefore, if age-related changes in ultimate language proficiency are to be attributable to these cognitive changes and not to a specific language module that is constrained by a maturational schedule, then the decline in ultimate proficiency in a second language should also be gradual and constant. Conversely, if the age-related changes in ultimate proficiency are

reflections of a critical period for second language acquisition, proficiency should show a discontinuity at a certain point in time, probably around puberty. Such a discontinuity is the minimal essential evidence needed to reject the null hypothesis of no critical period.

The empirical issues that reflect these concerns are the shape of the function that relates proficiency to age of language learning and the role that other factors play in this relation. If there is a critical period, then the relation between age of learning and proficiency will be nonlinear because of a sharp break at the critical period; if there is no critical period, the relation will be linear. Regarding other factors, if there is a critical period, then age will be the exclusive or primary factor accounting for proficiency; if there is no critical period, then other factors will be significant.

We conducted a preliminary analysis of data from the 1990 U.S. population census (U.S. Department of Commerce, 1995) to test these two hypotheses. The data set includes information on a large number of population characteristics, such as home language background, age of immigration to the United States, level of formal education, and English ability. Population data of this sort have both advantages and disadvantages. The advantages are (a) the sample is close to the universe of the population and relatively free from bias; (b) the numbers are large enough that parameter estimates are highly reliable; and (c) the data have already been collected, and the empirical properties of many of the demographic variables are well understood. The major weakness is that the measure of English proficiency is obtained through self-report, which is susceptible to various forms of corruption. However, a number of studies have compared self-report on English proficiency with behaviorally measured proficiency and report reasonable positive relationships between these two measures (Hakuta & D'Andrea, 1992). Kominski, 1989, cited in McArthur, 1993; McArthur & Siegel, 1983;

The present analysis is based on data from New York State, which, along with California, Florida, Illinois, and Texas, has among the largest language minority numbers in the United States. From the New York population, individuals were selected whose home language was either Spanish or Chinese. The following variables were estimated:

1. Length of Residence in the United States (based on year of entry)
2. Current Age (as of 1990)
3. Age of Arrival (subtraction of Length of Residence from Current Age)
4. Years of Formal Education
5. English Proficiency ("Not at all", "Not well", "Well", "Very Well", "Speak only English").

Because census data are categorical, models are best tested through log-linear analysis. However, one of our goals is to ask whether the data are linear, so the categorical data were converted into individual scores through interpolation, and some assumptions had to be made to make the data interpretable through linear analysis.

Because we are interested in asymptotic effects that reveal ultimate proficiency rather than the learning curve, we assumed that length of residence of 10 years would be ample time for most individuals to reach stable proficiency in English. Therefore, we eliminated participants who had length of residence of 10 years or less. This left us with a sample that included 24,903 speakers of Chinese and 38,787 speakers of Spanish. The initial analysis plots English proficiency as a function of Age of Arrival. The question of linearity can be answered by fitting a locally weighted, nonlinear function to the data using the LOWESS procedure available through SYSTAT (Wilkinson, 1996). The linear trend in these data is shown in Fig. 7.3 for Chinese ($r = -.52$) and Fig. 7.4 for Spanish ($r = -.44$). Superimposing the two curves

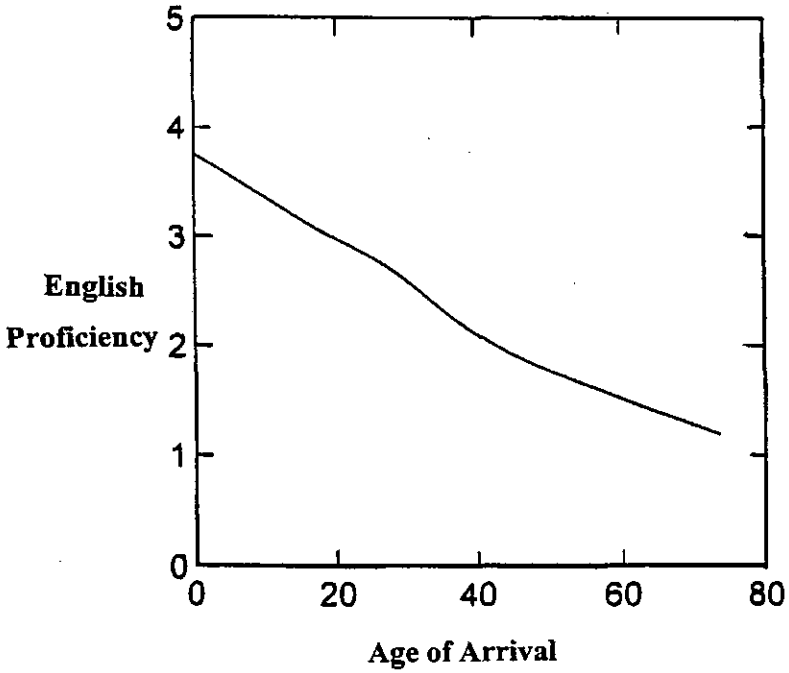

FIG. 7.3. Proficiency by age of arrival for Chinese speakers.

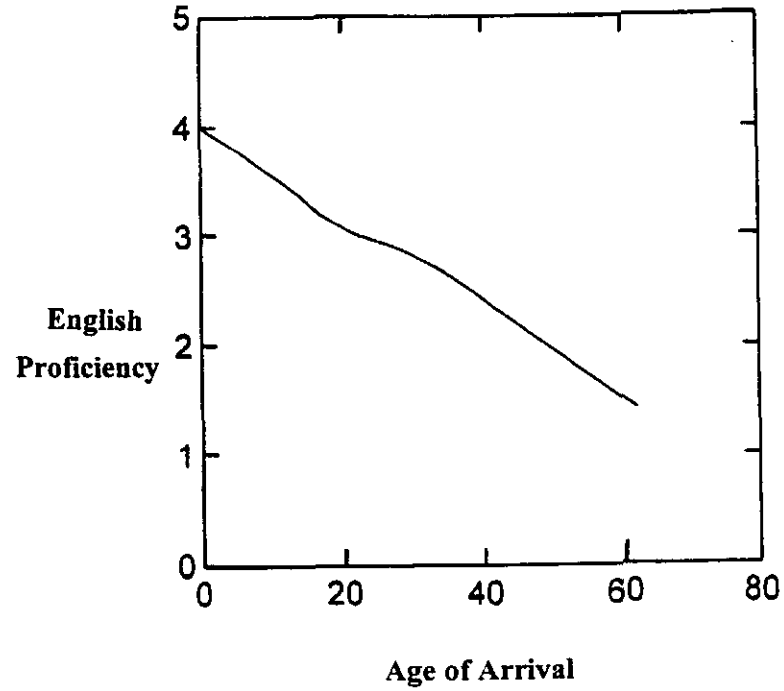

FIG. 7.4. Proficiency by age of arrival for Spanish speakers.

on each other shows how similar the slopes are, although there is a slightly higher mean score for Spanish than for Chinese. Most important, there appears to be nothing special about the age range before puberty. The decline in proficiency remains constant across the ages and is similar for both Spanish and Chinese.

To separate out the effects of cognition, the data were disaggregated by the educational level of the participants. Three categories were created: (a) less than 9 years of formal education; (b) between 9 and 13 years of formal education; and (c) more than 13 years of formal education. The graphs are shown in Fig. 7.5 for Chinese and Fig. 7.6 for Spanish. Schooling was positively related to proficiency, independently of age of arrival or language. These data should be interpreted carefully with respect to cause and effect: for those participants who immigrated as children, increased English proficiency could just as easily lead to more formal education as the other way around.

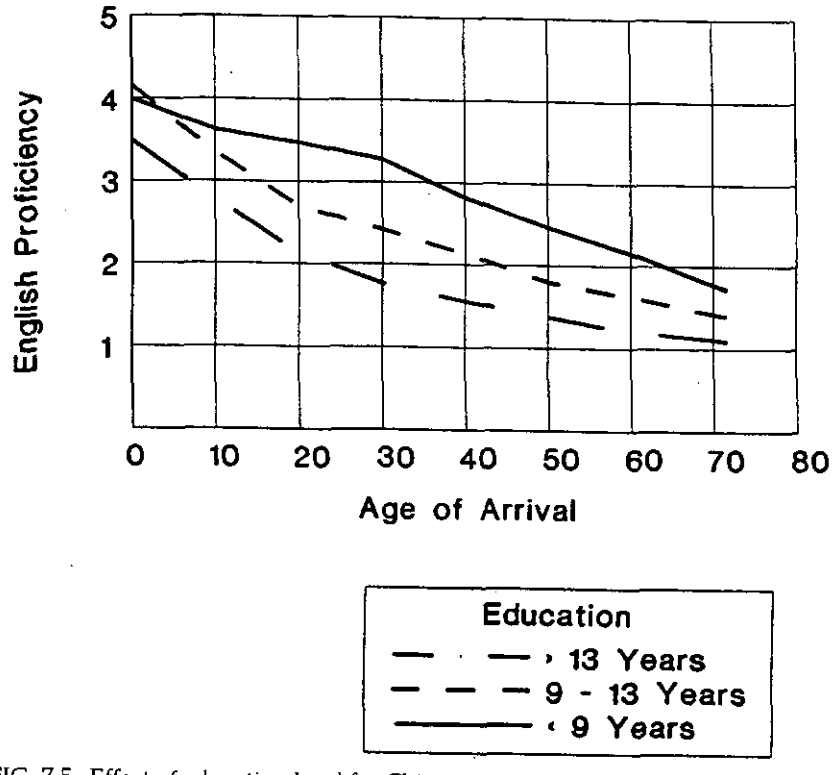

FIG. 7.5. Effect of education level for Chinese speakers.

We will conduct further analyses to separate out those individuals who were educated in the United States from those who were educated prior to immigration. Nevertheless, the graphs reveal systematic effects for educational level for both groups of participants.

CONCLUSION

It is tempting to believe that children are better second language learners than adults because their brains are specially organized to learn language, whereas those of adults are not. This is the explanation of the critical period hypothesis. The evidence for it comes from several sources. Informal observation irrefutably shows children to be more successful than adults in mastering a second language. Empirical studies confirm this pattern by demonstrating performance differences between children and adult learners on various tasks and measures. Yet both informal observation and empirical testing also yield exceptions to this rule. Late learners are sometimes

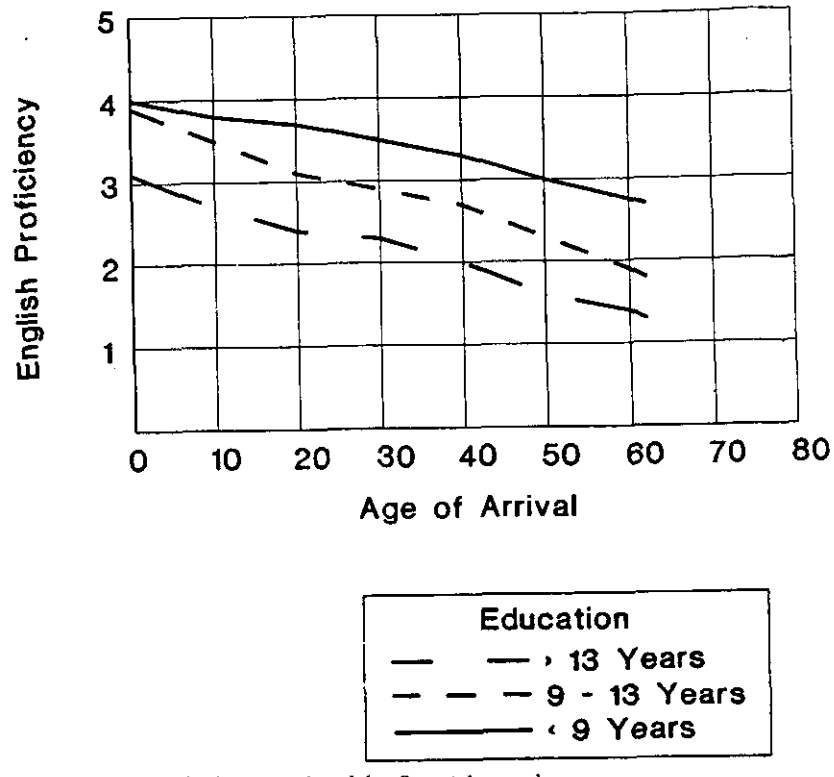

FIG. 7.6. Effect of education level for Spanish speakers.

able to achieve native-like perfection in a second language (e.g., Ioup, Boustagui, El Tigi, & Moselle, 1994) and experimental results sometimes show late learners performing just as well as early learners, even though the older group on average performs worse (e.g., Birdsong, 1992). Biological restrictions such as brain maturation should not be so easily overturned.

Neurological evidence has also been amassed to support claims for a critical period in second language acquisition. Neville (1995; Neville & Weber-Fox, 1994; see also Weber-Fox & Neville, chap. 2, this volume), for example, demonstrated event-related brain potential differences that show that neural organization is different for early and late language learners. Again, however, correlation is not causality. Researchers remind us that neural organization can reflect different kinds of experiences without being abnormal or supporting inferior performance (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995; Locke, 1993; Merzenich et al., 1984). Special experiences, in other words, may influence neural organization without affecting performance. As Gazzaniga (1992) pointed out, neural configurations

are just as likely to be altered by cognitive processes as cognitive processes are to be determined by neurological structures. It is not surprising that the experience and knowledge we accumulate as we grow changes the way in which new information, including new languages, will be represented and that these differences can be detected as different patterns of neural organization in the brain. Indeed, brain patterns vary in the population: In some people, language is lateralized to the right hemisphere instead of the left, but they can still write, draw, and throw a baseball. The only issue is whether or not learning is impaired by these differences and whether the critical variable in determining the difference is age of first exposure. Here, only behavioral evidence is relevant, and the behavioral evidence does not make a sufficiently compelling case.

A more unusual argument for a critical period in language acquisition (but not specifically second language acquisition) was offered by Hurford (1991; see also Hurford & Kirby, chap. 3, this volume). Using computer modelling to simulate population growth and evolution, he demonstrated how a critical period for language acquisition is an adaptive feature in population terms. His explanation was that there is no selective pressure to keep the capacity for language learning available after puberty, so it turns off. The argument is interesting, but the amount of conjecture in the discussion is staggering.

Our discussion described some linguistic and cognitive factors involved in the language learning process that both contradict specific claims from the critical period hypothesis and offer an alternative means of explaining the advantage younger learners normally enjoy in second language acquisition. In addition, social factors conspire to ease the effort for young children by providing a nurturing environment, simplified input, educational opportunities, cooperative peers, and other supporting aspects of a social context that facilitate the acquisition of any language. Armed with these problems in the experimental studies designed to support a critical period, unconvinced that performance differences for younger and older learners reflect more than simple correlation, and given alternative explanations for the patterns of data that do occur, we see no reason to reject the null hypothesis that there is no critical period for second language acquisition.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

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

Bialystok, E., & Miller, B. (1998). *The problem of age in second language acquisition: Influences from language, task, and structure.* Unpublsihed manuscript, York University.

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

Birkhill, W. R., & Schaie, K. W. (1975). The effect of differential reinforcement of cautiousness in the intellectual performance of the elderly. *Journal of Gerontology, 30,* 578–583.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30–50). Clevedon, England: Multilingual Matters.

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

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Colombo, J. (1982). The critical period concept: Research, methodology, and theoretical issues. *Psychological Bulletin, 91,* 260–275.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities* (pp. 409–422). Amsterdam: Elsevier.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science, 270,* 305–306.

Gazzaniga, M. S. (1992). *Nature's mind: The biological roots of thinking, emotions, sexuality, language, and intelligence.* New York: Basic Books.

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

Hakuta, K., & D'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13,* 72–99.

Hultsch, D., & Dixon, R. (1990). Learning and memory and aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd edition, pp. 258–274). New York: Academic Press.

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

Ijaz, H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning, 36,* 401–451.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition, 16,* 73–98.

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

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

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

Juffs, A., & Harrington, M. (1995). Parsing effects in L2 sentence processing: Subject and object asymmetries in Wh-extraction. *Studies in Second Language Acquisition, 17,* 483–516.

Kellerman, E. (1986). An eye for an eye: Crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 35–48). New York: Pergamon.

Locke, J. L. (1993). *The child's path to spoken language.* Cambridge, MA: Harvard University Press.

Martohardjono, G., & Flynn, S. (1995). Is there an age-factor for universal grammar? In D. Singleton & Z. Lengyel, (Eds.), *The age factor in second language acquisition* (pp. 135–153). Clevedon, England: Multilingual Matters.

McArthur, E. (1993). *Language characteristics and schooling in the United States: A changing picture, 1979 and 1989* (NCES 93–699). U.S. Department of Education: National Center for Education Statistics.

McArthur, E., & Siegel, P. (1983). *Developments in the measurement of English language proficiency.* American Statistical Association, Proceedings of the Social Statistics Section (pp. 373–378).

Merzenich, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., & Zook, J. M. (1984). Somatosensory cortical map changes following digit amputation in adult monkeys. *Journal of Comparative Neurology, 224,* 591–605.

Neville, H. J. (1995). Developmental specificity in neurocognitive development in humans. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 219–231). Cambridge, MA: MIT Press.

Neville, H. J., & Weber-Fox, C. M. (1994). Cerebral subsystems within language. In B. Albowitz, K. Albus, U. Kuhnt, H.-Ch. Norhdurft, & P. Wahle (Eds.), *Structural and functional organization of the neocortex* (pp. 424–438). Berlin: Springer-Verlag.

Newport, E. (1991). Contrasting conceptions of the critical period for language. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 111–130). Hillsdale, NJ:

Lawrence Erlbaum Associates.

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

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

Pinker, S. (1994). *The language instinct.* New York: Morrow.

Rabinowitz, J. C., & Craik, F. I. M. (1986). Prior retrieval effects in young and old adults. *Journal of Gerontology, 41,* 368–375.

Schaie, K. W., & Willis, S. L. (1991). *Adult development and aging* (3rd ed.). New York: Harper Collins.

U. S. Department of Commerce (1995). *Census of Population and Housing: Detailed Cross-tabulations of Selected Language Groups for States: 1990.* CD-ROM, Bureau of the Census, Population Division, Decennial Programs Coordination Branch.

Wilkinson, L. (1996). *Systat 6.0 for Windows: Graphics.* Chicago: SPSS Inc.