Research Article

CRITICAL EVIDENCE: A Test of the Critical-Period Hypothesis for Second-Language Acquisition

Kenji Hakuta,[1] Ellen Bialystok,[2] and Edward Wiley[1]

[1]Stanford University and [2]York University, Toronto, Ontario, Canada

Abstract—*The critical-period hypothesis for second-language acquisition was tested on data from the 1990 U.S. Census using responses from 2.3 million immigrants with Spanish or Chinese language backgrounds. The analyses tested a key prediction of the hypothesis, namely, that the line regressing second-language attainment on age of immigration would be markedly different on either side of the critical-age point. Predictions tested were that there would be a difference in slope, a difference in the mean while controlling for slope, or both. The results showed large linear effects for level of education and for age of immigration, but a negligible amount of additional variance was accounted for when the parameters for difference in slope and difference in means were estimated. Thus, the pattern of decline in second-language acquisition failed to produce the discontinuity that is an essential hallmark of a critical period.*

The idea that there is a biologically based critical period for second-language acquisition that prevents older learners from achieving native-like competence has appeal to both theorists and social policymakers (Bailey, Bruer, Symons, & Lichtman, 2001). The critical-period hypothesis was originally proposed in the neurolinguistic literature by Penfield and Roberts (1959) and vigorously followed up by Lenneberg (1967), who speculated that maturational aspects of the brain that limited recovery from brain traumas and disorders would extend to second-language acquisition. Subsequent research using behavioral evidence appeared to confirm this hypothesis (Johnson, 1992; Johnson & Newport, 1989; Oyama, 1976; Patkowski, 1980, 1994). The measure of language proficiency in these studies varied (ratings of oral speech, grammaticality judgment tasks), but the typical result was that proficiency scores declined with increases in age of initial exposure to the second language.

The claim that there is an age-related decline in the success with which individuals master a second language is not controversial. The diminished average achievement of older learners is supported by personal anecdote and documented by empirical evidence (Flege, Yeni-Komshian, & Liu, 1999; Stevens, 1999). What *is* controversial, though, is whether this pattern meets the conditions for concluding that a critical period constrains learning in a way predicted by the theory. A critical period minimally entails two characteristics: (a) a high level of preparedness for learning within a specified developmental period to ensure the domain is mastered by the species and (b) a lack of preparedness outside this period (Bornstein, 1989; Colombo, 1982). The consequence of these conditions is that the relation between learning and age is different inside and outside the critical period.

Proponents of a critical-period explanation have attempted to place the description of second-language learning within these parameters. Johnson and Newport (1989, 1991) have argued, for example, that there is a strong age-related decline in proficiency for languages learned prior to puberty (defined as 15 years old) and random variation in achievement among individuals who are exposed to a second language later in life. Such developmental discontinuity at an identifiable maturational time would constitute support for the two conditions of a critical period. The data, however, are controversial because of the difficulty in separating out the effects of age of initial exposure, duration of exposure, and social and linguistic backgrounds of the participants (see the analysis and critique of Johnson and Newport's study in Bialystok & Hakuta, 1994).

Other researchers have argued that the evidence fails to support the interpretation that language-learning potential is fundamentally changed after a critical period (e.g., Epstein, Flynn, & Martohardjono, 1996; Hakuta, 2001). Two kinds of evidence have typically been used in these challenges. The first is the identification of older learners who achieve nativelike competence in the second language (Birdsong, 1992; Bongaerts, Planken, & Schils, 1995; Ioup, Boustagui, El Tigi, & Moselle, 1994). The second is behavioral evidence that fails to reveal a qualitative change in learning outcomes at the close of a critical period (Bialystok & Hakuta, 1999; Bialystok & Miller, 1999; Birdsong & Flege, 2000; Birdsong & Molis, 2001; Flege, 1999; Flege, Munro, & MacKay, 1995; Flege et al., 1999). Whether such evidence is considered damaging to the critical-period hypothesis depends on the stringency of the criteria for defining the boundaries of the critical period (Birdsong, 1999; Harley & Wang, 1997; Singleton & Lengyel, 1995). Nonetheless, both weak and strong interpretations of the critical-period hypothesis require the demonstration of a significant change in learning outcome, not merely a monotonic decline with age.

Defense of the position that language learning is constrained by a critical period requires specifying the maturational stage at which language-learning potential changes, and ideally the reason for the change. However, there has been little consensus about what age constitutes the critical point, and reasons for proposing different ages have rarely been offered. Researchers have variously claimed, for example, that the age at which the critical period terminates is 5 years (Krashen, 1973), 6 years (Pinker, 1994), 12 years (Lenneberg, 1967), or 15 years (Johnson & Newport, 1989).

An alternative to the critical-period hypothesis is that second-language learning becomes compromised with age, potentially because of factors that are not specific to language but nevertheless interfere with the individual's ability to learn a new language. These might include social and educational variables that influence learning potential and opportunity, as well as cognitive aging that gradually erodes some of the mechanisms necessary for learning a complex body of knowledge, such as a new language.

Among social factors, education has been most clearly demonstrated to influence second-language acquisition. Learners who arrive as immigrants at different ages have fundamentally different experiences, are exposed to qualitatively and quantitatively different samples of the new language, and have distinctly different opportunities for formal study of the language either directly or through other educational content (Bialystok & Hakuta, 1994; Flege et al., 1999). Flege and his colleagues have re-

Address correspondence to Kenji Hakuta, CERAS Building, Stanford University, Stanford, CA 94305; e-mail: hakuta@stanford.edu.

Table 1. *Regression of English proficiency on education: Spanish- and Chinese-speaking immigrants*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*
		Chinese			
Intercept	1.7431	0.00417	96,590	174,334.0	<.0001
5–8 years education	0.2493	0.00624	884	1,596.0	<.0001
Some high school	0.7324	0.00586	8,659	15,628.1	<.0001
High school graduate	1.0693	0.00548	21,071	38,030.6	<.0001
Some college	1.7398	0.00451	82,450	148,813.0	<.0001
		Spanish			
Intercept	2.0573	0.00136	1,796,840	2,293,711.0	<.0001
5–8 years education	0.3484	0.00184	28,171	35,961.6	<.0001
Some high school	0.8710	0.00196	154,633	197,393.0	<.0001
High school graduate	1.1708	0.00209	244,933	312,664.0	<.0001
Some college	1.4445	0.00198	417,988	533,572.0	<.0001

Note. R^2 = .4221 for Chinese-speaking immigrants and .2622 for Spanish-speaking immigrants.

ported complex effects of educational programs on second-language acquisition, and in one of their studies age-of-learning effects disappeared when education was controlled (Flege et al., 1999).

The second group of factors is the changes in cognition that occur with aging. Although critical periods have not been posited in most cognitive domains, there are nonetheless age-related changes in cognitive processing. Some age-related changes in cognitive processes relevant to language learning are decreased ability to learn paired associates (Salthouse, 1992), increased difficulty encoding new information (Craik & Jennings, 1992; Park et al., in press; Rabinowitz, Craik, & Ackerman, 1982), and reduced accuracy recalling detail as opposed to gist (Hultsch & Dixon, 1990). Kemper (1992) pointed out that older adults' second-language proficiency, like their first-language proficiency, could also be affected by such factors as working memory capacity, cognitive processing speed, and attention. All these factors decline with age, and the decline is documented across the life span. Such a reduction in cognitive resources would surely affect the ability to learn a new language. Older learners would find the task more difficult than younger ones, although no critical period would be involved.

In the present study, we examined the effect of age of acquisition on second-language proficiency by studying a very large sample of second-language learners who covered a wide range of ages of initial exposure to English. Minimally, demonstrating a critical period would require finding evidence for a clear discontinuity in learning outcome around a specified age. Moreover, this pattern would have to be independent of social or educational factors that also impinge on successful second-language acquisition.

METHOD

Participants

Data for this study were derived from the 1990 U.S. Census, which provided detailed data on selected language groups by state (U.S. Department of Commerce, 1995). The participants included for analysis were those respondents identified as native speakers of either Spanish or Chinese. These languages were chosen because they differ in their structural similarity to English. Additionally, speakers from these language groups have a sufficiently long history in the United States that the full range of the parameters in the variables of interest could be investigated. For Spanish speakers, we used data from California, Illinois, Texas, and New York, four of the largest states, with large populations of Spanish speakers. For Chinese speakers, we used data from these same states, plus Florida, Maryland, Massachusetts, New Jersey, Pennsylvania, Virginia, and Washington. These additional states, where concentrated populations of Chinese speakers can be found, were added to increase the sample size. To ensure that English ability reflected a stable level of attainment in the analysis, we set the minimum length of residence in the United States at 10 years. Stevens (1999), who analyzed a 1% public-use sample drawn from the same census, found that her sample of immigrants reached asymptotic levels of self-reported English proficiency after 10 years. The final analysis included data from 2,016,317 speakers of Spanish and 324,444 speakers of Chinese.

Measures

The census form asks respondents to self-describe their English ability using one of five categories: "not at all," "not well," "well," "very well," and "speak only English." An independent Census Bureau study to validate the response categories against actual language proficiency measures (Kominski, 1989) and our own analyses of those data have shown an acceptable level of correlation between this item and objective measures, r = .52–.54.[1] Although an objective and more direct measure of English proficiency would be desirable, the strength of the present approach lies in the size of the sample and our ability to disaggregate the data by important background variables in testing whether there is discontinuity in the age effect.

1. To further substantiate the relationship between this census item and objective measures of English proficiency, we obtained the data collected in the National Content Test (NCT) and its reinterview, conducted by the Census Bureau during the spring and summer of 1986 (described in Kominski, 1989). In our analysis of objective and subjective proficiency measures administered to 652 Spanish-background adults sampled as part of NCT, we found substantial correlations between the subjective item and scores from assessments of written (r = .52, p < .001) and oral (r = .54, p < .001) English proficiency. The scores from the written and oral assessments were also correlated, r = .68, p < .001.

Table 2. *Regression of English proficiency on education and age of immigration: Spanish- and Chinese-speaking immigrants*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*	R^2 Partial	R^2 Total
			Chinese				
Intercept	2.69395	0.01185	23,924	51,657.2	<.0001	—	—
5–8 years education	0.03791	0.01731	2	4.8	.0285	—	—
Some high school	0.51324	0.01513	533	1,151.0	<.0001	—	—
High school graduate	0.98867	0.01392	2,337	5,045.6	<.0001	—	—
Some college	1.30098	0.01234	5,144	11,106.3	<.0001	—	.4221
Age of immigration	−0.02186	0.00026	3,325	7,180.0	<.0001	.0932	.5153
			Spanish				
Intercept	2.63091	0.00324	469,497	657,397.0	<.0001	—	—
5–8 years education	0.22956	0.00441	1,939	2,715.5	<.0001	—	—
Some high school	0.88544	0.00434	29,691	41,574.1	<.0001	—	—
High school graduate	1.15842	0.00448	47,812	66,947.3	<.0001	—	—
Some college	1.31456	0.00427	67,572	94,615.9	<.0001	—	.2622
Age of immigration	−0.02022	0.00010	26,566	37,197.7	<.0001	.0632	.3254

Additional census questions included in our analysis ask about present age, year of arrival in the United States, and educational background. The first two allowed us to compute the age of arrival. Independent variables were created from census ordinal variables with 10 to 19 levels. For modeling purposes, we constructed interval-scale approximates by taking the midpoint value for each category. Our analytical goal was to model English proficiency on the following predictors: age of immigration, education, and existence of a critical period. Results for Spanish-speaking and Chinese-speaking immigrants are reported separately.

Years of formal education was determined from Question 12, highest degree of education attained, by assigning year-equivalents to the response categories as follows: "no school or less than kindergarten" = 0, "kindergarten" = 1, "1st to 4th grade" = 3.5, "5th to 8th grade" = 7.5, "9th grade" = 10, "10th grade = 11, "11th grade" = 12, "12th grade, no diploma" = 12, "high school graduate (includes equivalency)" = 13, "some college, no degree" = 15, "associate degree in college (occupational program)" = 15, "associate degree in college (academic program)" = 15, "bachelor's degree" = 17, "master's degree" = 18, "professional school degree" = 18, and "doctoral degree" = 22. In addition to the original 16-level categorical variable and its interval-level approximate, we created a five-category scale consisting of the following levels: less than 5th-grade education = 1, 5th- to 8th-grade education = 2, high school education without diploma = 3, high school graduate = 4, and college = 5.

Length of residence was estimated from Question 10, year of entry, by subtracting the midpoint of each response category from 1990, the year when the census was taken. The response categories (and in parentheses, the derived length-of-residence estimates) considered in this analysis were as follows: 1975–1979 (13 years), 1970–1974 (18 years), 1965–1969 (23 years), 1960–1964 (28 years), 1950–1959 (35.5 years), and before 1950 (>40 years). To ensure that English ability reflected a stable level of attainment, we excluded from the analysis individuals with less than 10 years of residence in the United States.

Age of immigration was estimated by subtracting each individual's length of residence from the midpoint of the response category that individual selected for Question 5, present age. The categories representing age of immigration (and in parentheses, the midpoint in each interval) were as follows: 0–2 years (1 year), 3–4 years (3.5 years), 5–9 years (7 years), 10–14 years (12 years), 15 years (15 years), 16–17 years (16.5 years), 18–19 years (18.5 years), 20–24 years (22 years), 25–29 years (27 years), 30–34 years (32 years), 35–39 years (37 years), 40–44 years (42 years), 45–49 years (47 years), 50–54 years (52 years), 55–59 years (57 years), 60–64 years (62 years), 65–69 years (67 years), 70–74 years (72 years), and 75–115 years (95 years).

One of the benefits of using census data is the availability of extremely large samples for analysis. Because statistical significance reflects sample size as well as effect size, statistical significance can be misleading in analyses based on these large samples. More important in these analyses is the practical significance of any tested effects. The interpretation of effect sizes provides insight into the magnitude of tested effects (independent of sample-size considerations). In regression-based modeling techniques, one appropriate effect-size measurement is partial R^2. This statistic provides a measure of the increased proportion of variability in an outcome variable that can be explained by the inclusion of an additional independent variable in the regression model (Neter, Kutner, Nachtsheim, & Wasserman, 1996, p. 339). Regardless of statistical significance, variables added to the regression model must have large partial R^2 values (i.e., they must account for substantial proportions of variability in the outcome variable) in order to be considered practically significant.

RESULTS

Education, Age of Immigration, and Cohort Effects

To begin, we considered the simple model of English proficiency as a function of education. We tested whether English proficiency was best modeled on (a) dummy variables for the 16 categories in the census, (b) dummies for the simpler five-level categorical variable, (c) a linear term for the derived interval-level variable, or (d) both a linear and a quadratic interval-level education term. The five-level education variable provided the best balance between parsimony and model fit (Chinese: R^2 = .4221;

Table 3. *Regression of English proficiency on education, age of immigration, and critical-period variables: Chinese-speaking immigrants (critical point = age 15)*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*	R^2 Partial	R^2 Total
Intercept	2.76989	0.00923	41,863	90,139.4	<.0001	—	—
5–8 years education	0.10036	0.00576	141	303.8	<.0001	—	—
Some high school	0.44701	0.00551	3,061	6,591.1	<.0001	—	—
High school graduate	0.73278	0.00521	9,201	19,812.2	<.0001	—	—
Some college	1.26844	0.00455	36,074	77,674.7	<.0001	—	.4221
Age of immigration	−0.02640	0.00067	712	1,533.3	<.0001	.0932	.5153
Change in mean	0.05804	0.00424	87	187.8	<.0001	.0003	.5156
Change in slope	0.00227	0.00068	5	11.1	.0009	0	.5156

Spanish: $R^2 = .2622$)[2] and so was used in all subsequent analyses including education. The results of this analysis are shown in Table 1.

In the second step of the analysis, age of immigration was added to the model with the education dummies (see Table 2). In addition to testing the linear main effect of age of immigration, we tested the interaction between age of immigration and education. An interaction between these variables would suggest that the relation between age of immigration and English proficiency changes with different levels of education. There was a moderate effect for the age-of-immigration linear term (Chinese: marginal $R^2 = .0932$; Spanish: marginal $R^2 = .0632$). No interaction term added more than .0016 to the model R^2, providing very little evidence for an interaction between these two variables.

A cohort variable representing differences in English proficiency between individuals who entered the United States in the 1960s and those who entered in the 1970s was included next. This analysis was conducted in part to test the validity of our assumption that we would be sampling immigrants at their asymptotic levels of English proficiency by selecting only those who had lived in the United States at least 10 years. The length of residence of the two cohorts differed by an average of 10 years, allowing us to test for the effect of length of residence within the range of the study. There was little indication of either a main effect of cohort or interactions of cohort with age of immigration and education; none of the terms added more than .0032 to the model R^2. Thus, in our sample of individuals who had 10 or more years of U.S. residence, there is no evidence for an effect of length of residence on English proficiency.

Testing the Critical-Period Hypothesis

The model so far included simple additive effects for the five-category education variable and age of immigration. To test for evidence of a critical period, we followed the procedures for modeling regression discontinuities suggested by Neter et al. (1996, pp. 474–478). If there is a critical period, then there would be a discontinuity in the regression of English proficiency on age of immigration at the point marking the end of the critical period (hereafter referred to as the *critical point*). As Neter et al. pointed out, a regression line might be discontinuous at a point X_c because of a change in mean (i.e., a break in the regression line), a change in slope, or both. Figure 1 represents these alternatives for the critical-period hypothesis. Note that in the two panels that incorporate the slope-change model, there are alternative projections for the discontinuity, as shown by the two lines labeled (a) and (b). Johnson and Newport's (1989) data as reanalyzed in Bialystok and Hakuta (1994), for example, resemble model (b) with the critical point being at age 20.

The possibility of such discontinuities was tested by two variables in our regression model (Neter et al., 1996, p. 478). One allowed us to test for a change in the mean of the regression line:

$$\text{change in mean} = \begin{cases} 1 \text{ if age of immigration} \geq \text{critical point} \\ 0 \text{ if age of immigration} < \text{critical point} \end{cases}$$

The other allowed us to test for a change in the slope of the regression line:

$$\text{change in slope} = (\text{change in mean}) * (\text{age of immigration} - \text{critical point})$$

Two different ages were used to define the critical point: ages 15 and 20. The first point, age 15, corresponds to the typical onset of puberty. This age

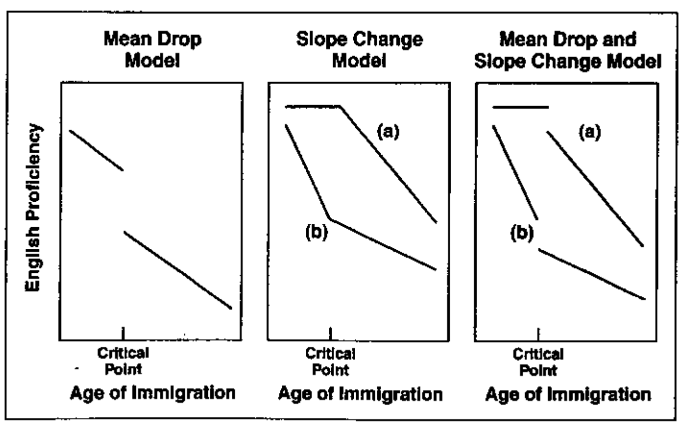

Fig. 1. Three alternative predictions of the critical-period hypothesis. The alternative lines (a) and (b) represent two logically possible ways in which the slope may change about the putative critical point.

2. The 16-level categorical education variable provided the best fit (Chinese: $R^2 = .4389$; Spanish: $R^2 = .2667$), followed by the 5-level variable (Chinese: $R^2 = .4221$; Spanish: $R^2 = .2622$), the linear and quadratic fit (Chinese: $R^2 = .4096$; Spanish: $R^2 = .2556$), and the simple linear fit (Chinese: $R^2 = .4023$; Spanish: $R^2 = .2479$). Although the 16-level variable provided the best fit, it accounted for only a slightly greater proportion of variance in English proficiency than its 5-level counterpart (Chinese: R^2 difference = .0168; Spanish: R^2 difference = .0045).

Table 4. *Regression of English proficiency on education, age of immigration, and critical-period variables: Spanish-speaking immigrants (critical point = age 15)*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*	R^2 Partial	R^2 Total
Intercept	3.02532	0.00361	498,299	700,445.0	<.0001	—	—
5–8 years education	0.26340	0.00177	15,802	22,213.1	<.0001	—	—
Some high school	0.67604	0.00192	88,362	124,209.0	<.0001	—	—
High school graduate	0.94236	0.00206	148,943	209,366.0	<.0001	—	—
Some college	1.19965	0.00196	265,741	373,544.0	<.0001	—	.2622
Age of immigration	−0.04573	0.00027	21,033	29,565.9	<.0001	.0632	.3254
Change in slope	0.02730	0.00028	7,004	9,844.7	<.0001	.0043	.3297
Change in mean	−0.05045	0.00185	531	745.8	<.0001	.0002	.3300

Table 5. *Regression of English proficiency on education, age of immigration, and critical-period variables: Chinese-speaking immigrants (critical point = age 20)*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*	R^2 Partial	R^2 Total
Intercept	2.72891	0.00750	61,569	132,559.0	<.0001	—	—
5–8 years education	0.09922	0.00576	138	296.4	<.0001	—	—
Some high school	0.44600	0.00551	3,045	6,556.2	<.0001	—	—
High school graduate	0.73139	0.00521	9,156	19,713.1	<.0001	—	—
Some college	1.26715	0.00455	35,962	77,427.8	<.0001	—	.4221
Age of immigration	−0.02206	0.00038	1,558	3,353.6	<.0001	.0932	.5153
Change in mean	0.03465	0.00374	40	85.9	<.0001	.0002	.5155
Change in slope	−0.00245	0.00040	17	37.5	<.0001	.0001	.5156

has become the standard empirical cutoff, following the influential study by Johnson and Newport (1989). The second point, age 20, was based primarily on visual inspection of the local regression curves (discussed later in this section), which suggested that if discontinuities existed, they would be at an age later than puberty (cf. Bialystok & Hakuta, 1994). Model parameters were estimated separately for each of these putative critical points.

Evidence for either a significant break in the mean or a change in slope of the regression line would support the existence of a critical period in second-language acquisition. Tables 3 (Chinese speakers) and 4 (Spanish speakers) report the results of testing for a critical period ending at age 15; Tables 5 and 6 report the results of testing for a critical period ending at age 20. In no case does either the change in mean or the change in slope add more than .0043 to the model R^2.

Interactions between both the mean and slope-change variables and the education variables were also tested; sizable effects would be evidence for regression discontinuities (and therefore critical periods) specific to certain educational groups. Again, there was little evidence for such discontinuities (no change in R^2 of more than .0018).

To this point, we have reported tests of parametric models accounting for variability in English proficiency. To better understand the data, we tested a model that relaxed the parametric form to create a local regression[3] fit. Local regression models provide greater flexibility than their parametric counterparts by allowing the specification of relationships that may not adhere to a parametric form. Rather than fitting a straight line or parametric curve to the data at hand, local regression provides an individual model fit for each point in the data set. Because of this nonparametric flexibility, local regression models generally are more sensitive to relationships between variables. In our analysis, local regression models contribute visual as well as quantitative evidence regarding the existence of a critical period.

In local regression modeling, a smoothing span specifies the size of a neighborhood[4] of nearby data used to determine the value of the regression line at each point.[5] As the smoothing span increases, a larger local neighborhood is used for determining the fit at each data point, therefore increasing the smoothness of the regression line.[6] Typical values chosen for smoothing spans range from .25 to .75. The local regression models reported here were run using both these values in or-

3. All local regression modeling was carried out using releases 3.4 and 4.0 of S-Plus Advanced Data Analytic Software (Insightful Corp., Seattle, Washington). Local regression fits utilized the loess function; loess curves were plotted using predicted values from loess models.

4. One typically specifies a probability distribution to weight the individual data points within this neighborhood.

5. An intuitive way to think about this neighborhood is to consider a window (with length equal to the smoothing span) centered around one specific data point. The data within that window are used to estimate the model fit for that data point. The window then slides to the next data point to estimate model fit for that point, and so forth.

6. In terms of the trade-off between bias and variance of fit, larger smoothing spans decrease model bias and increase model variance. Choosing an extremely small value for the smoothing span can result in bias due to overfitting the model to the data in hand.

Table 6. *Regression of English proficiency on education, age of immigration, and critical-period variables: Spanish-speaking immigrants (critical point = age 20)*

Variable	Parameter estimate	*SE*	*SS*	*F*	*p*	R^2 Partial	R^2 Total
Intercept	2.96103	0.00295	716,088	1,006,113.0	<.0001	—	—
5–8 years education	0.26273	0.00177	15,723	22,090.5	<.0001	—	—
Some high school	0.67541	0.00192	88,201	123,923.0	<.0001	—	—
High school graduate	0.94321	0.00206	149,239	209,683.0	<.0001	—	—
Some college	1.20114	0.00196	266,544	374,498.0	<.0001	—	.2622
Age of immigration	−0.03913	0.00016	43,806	61,548.7	<.0001	.0632	.3254
Change in slope	0.02061	0.00018	9,116	12,807.4	<.0001	.0042	.3296
Change in mean	0.02030	0.00188	83	116.9	<.0001	.0000	.3296

der to test both extremes of smoothness. As shown in Table 7, the smaller smoothing span (.25) brings only marginal improvement over the larger value (.75) in terms of standard errors. Furthermore, these trivial improvements come at the substantial cost of increasing the effective number of parameters in the model from 4 (representing a cubic fit) to nearly 12.

Figures 2 (Chinese) and 3 (Spanish) show the local regression plots of English proficiency on age of immigration when the larger smoothing span is used. A separate curve is plotted for each education group. The curves show essentially smooth declines in English proficiency as a function of age of immigration for all the education groups. There is no evidence for discontinuity in the function around any of the ages

Table 7. *Model summaries: Nonlinear and linear models of English proficiency on age of immigration by educational level, Spanish- and Chinese-speaking immigrants*

Statistic	Chinese Loess Span = .25	Chinese Loess Span = .75	Chinese Linear	Spanish Loess Span = .25	Spanish Loess Span = .75	Spanish Linear
Less than 5 years education						
n	31,790	31,790	31,790	424,554	424,554	424,554
Equivalent number of parameters	11.9	4.1	2.0	11.9	4.1	2.0
Residual *SE*	0.7712	0.7727	0.7736	0.9203	0.9203	0.9213
R^2	.1542	.1507	.1488	.0706	.0706	.0687
5–8 years education						
n	25,757	25,757	25,757	511,865	511,865	511,865
Equivalent number of parameters	11.9	4.1	2.0	11.9	4.1	2.0
Residual *SE*	0.7471	0.7477	0.7523	0.9051	0.9052	0.9060
R^2	.1223	.1207	.1097	.0467	.0465	.0449
High school, no graduation						
n	32,786	32,786	32,786	392,147	392,147	392,147
Equivalent number of parameters	11.9	4.1	2.0	11.9	4.1	2.0
Residual *SE*	0.7714	0.7715	0.7754	0.8418	0.8420	0.8506
R^2	.1540	.1536	.1449	.1361	.1357	.1180
High school, graduation						
n	43,848	43,848	43,848	308,507	308,507	308,507
Equivalent number of parameters	11.9	4.1	2.0	11.9	4.1	2.0
Residual *SE*	0.7462	0.7475	0.7514	0.7779	0.7783	0.7839
R^2	.2597	.2570	.2493	.1547	.1538	.1415
College						
n	190,263	190,263	190,263	379,244	379,244	379,244
Equivalent number of parameters	11.9	4.1	2.0	11.9	4.1	2.0
Residual *SE*	0.6126	0.6129	0.6160	0.6988	0.6989	0.7008
R^2	.1602	.1593	.1508	.1193	.1191	.1142

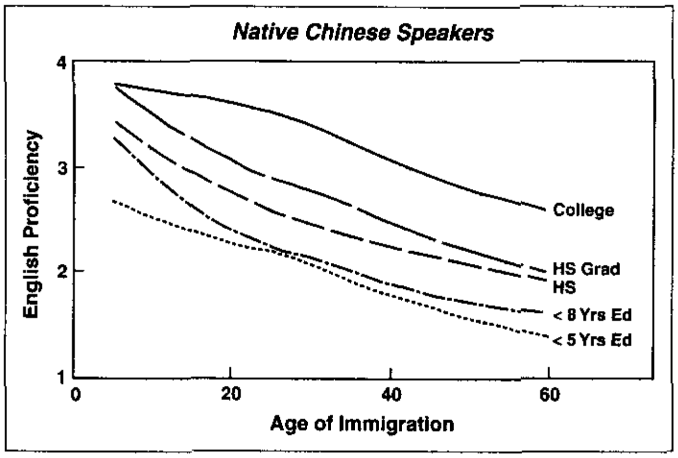

Fig. 2. Loess fits (span = .75) for English proficiency by age of immigration among Chinese immigrants. Results are shown separately for different education levels: less than 5 years ("<5 Yrs Ed"), less than 8 years ("<8 Yrs Ed"), some high school ("HS"), high school graduate ("HS Grad"), and some college ("College").

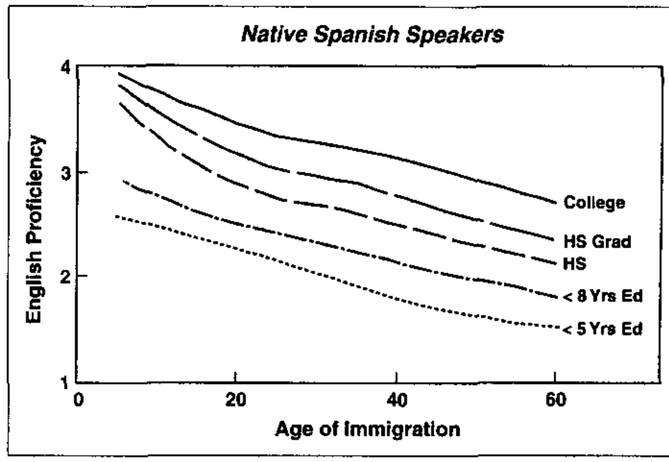

Fig. 3. Loess fits (span = .75) for English proficiency by age of immigration among Spanish-speaking immigrants. Results are shown separately for different education levels: less than 5 years ("<5 Yrs Ed"), less than 8 years ("<8 Yrs Ed"), some high school ("HS"), high school graduate ("HS Grad"), and some college ("College").

proposed as the close of the putative critical period, nor is there evidence suggesting the variation in older learners is random—proficiency continues to decline into adulthood.

The apparent linearity of these plots is confirmed by considering the gain in R^2 that is obtained by including a nonparametric form to model the relationship between English proficiency and age of immigration for each education group. Table 7 contains R^2 values for both linear and nonparametric fits of English proficiency on age of immigration for each education group. Little is gained by including an assumption of nonlinearity.

DISCUSSION

The critical period is a popular way of explaining the reason for the apparent success of children and failure of adults in learning a second language. In the United States, it has even been used in policy debates on how early to introduce immigrant children to English and when to teach foreign languages in school. We tested the critical-period hypothesis, and in particular searched for evidence of discontinuity in the level of English proficiency attained across a large sample of participants. Using both 15 years and 20 years as hypothesized cutoff points for the end of the critical period, we found no evidence of such a discontinuity in language-learning potential. Instead, the most compelling finding was that the degree of success in second-language acquisition steadily declines throughout the life span.

These data show that in addition to age of immigration, socioeconomic factors, and in particular the amount of formal education, are important in predicting how well immigrants learn English. Number of years of formal education added substantial amounts to the explanation of variance in both language groups and did not interact with other factors. The linear decline in proficiency across age of immigration was similarly confirmed in both groups. Although we could not directly test an explanation for this decline, the factors implicated in normal cognitive aging appear to be plausible sources of this effect.

Our conclusion from these models is that second-language proficiency does in fact decline with increasing age of initial exposure. The pattern of decline, however, failed to produce the discontinuity that is the essential hallmark of a critical period.

Acknowledgments—This study was supported in part by a grant from the Spencer Foundation to the first author. We thank Edith McArthur for bringing the data set to our attention, and Dorothy Waggoner for providing us with data on the National Content Test that enabled the analysis reported in footnote 1.

REFERENCES

Bailey, D.B., Bruer, J.T., Symons, F.J., & Lichtman, J.W. (Eds.). (2001). *Critical thinking about critical periods*. Baltimore: Paul Brookes.

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

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Bialystok, E., & Miller, B. (1999). The problem of age in second language acquisition: Influences from language, task, and structure. *Bilingualism: Language and Cognition, 2*, 127–145.

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

Birdsong, D. (Ed.). (1999). *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Erlbaum.

Birdsong, D., & Flege, J.E. (2000, October). *Regular-irregular dissociations in the acquisition of English as a second language*. Paper presented at the 25th Boston University Conference on Language Development, Boston, MA.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44*, 235–249.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters 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: A critical look at the critical period hypothesis* (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.

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

Craik, F.I.M., & Jennings, J.M. (1992). Human memory. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Epstein, S.D., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences, 19*, 677–758.

Flege, J.E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–131). Mahwah, NJ: Erlbaum.

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

Flege, J.E., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language, 41,* 78–104.

Hakuta, K. (2001). A critical period for second language acquisition? In D.B. Bailey, J.T. Bruer, F.J. Symons, & J.W. Lichtman (Eds.), *Critical thinking about critical periods* (pp. 193–205). Baltimore: Paul Brookes.

Harley, B., & Wang, W. (1997). The critical period hypothesis: Where are we now? In A.M.B. de Groot & J.F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 19–51). Mahwah, NJ: Erlbaum.

Hultsch, D.F., & Dixon, R.A. (1990). Learning and memory in aging. In J.E. Birren & K.W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 259–274). San Diego, CA: Academic Press.

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 in 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.

Kemper, S. (1992). Language and aging. In F.I.M. Craik & T.A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 213–270). Hillsdale, NJ: Erlbaum.

Kominski, R. (1989, August). *How good is "how well"? An examination of the census English-speaking ability item.* Paper presented at the annual meeting of the American Statistical Association, Washington, DC.

Krashen, S. (1973). Lateralization, language learning, and the critical period. *Language Learning, 23,* 63–74.

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

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). *Applied linear statistical models.* Chicago: Irwin.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5,* 261–283.

Park, D.C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A.D., & Smith, P.K. (in press). Models of visuospatial and verbal memory across the adult lifespan. *Psychology and Aging.*

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

Patkowski, M. (1994). The critical age hypothesis and interlanguage phonology. In M. Yavas (Ed.), *First and second language phonology* (pp. 205–221). San Diego, CA: Singular Publishing Group.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms.* Princeton, NJ: Princeton University Press.

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

Rabinowitz, J.C., Craik, F.I.M., & Ackerman, B.P. (1982). A processing resource account of age differences in recall. *Canadian Journal of Psychology, 36,* 325–344.

Salthouse, T.A. (1992). *Mechanisms of age-cognition relations in adulthood.* Hillsdale, NJ: Erlbaum.

Singleton, D., & Lengyel, Z. (Eds.). (1995). *The age factor in second language acquisition: A critical look at the critical period hypothesis.* Clevedon, England: Multilingual Matters.

Stevens, G. (1999). Age at immigration and second language proficiency among foreign-born adults. *Language in Society, 28,* 555–578.

U.S. Department of Commerce. (1995). *1990 census of population and housing: Detailed cross-tabulations of selected language groups for states: 1990* [CD-ROM]. Washington, DC: Bureau of the Census, Population Division, Decennial Programs Coordination Branch.

(RECEIVED 9/18/01; REVISION ACCEPTED 1/27/02)