Running Head: CRITICAL PERIOD IN SECOND LANGUAGE ACQUISITION

Critical Evidence:

A Test of the Critical Period Hypothesis for Second Language Acquisition

Kenji Hakuta

Stanford University

Ellen Bialystok

York University

Edward Wiley

Stanford Unive rsity

Word Count: 3608 text, 149 abstract, 371 footnotes.

Contact: Kenji Hakuta
             CERAS Bldg.
             Stanford University
             Stanford, CA. 94305

             (650) 725-7454
             e-mail: hakuta@stanford.edu

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 there would be a marked difference in the line regressing second language attainment on age of immigration on either side of the critical age point. Predictions tested were that there would be a difference in slope and/or a difference in the mean while controlling for slope. 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 necessary discontinuity that is an essential hallmark of a critical period.

Critical Evidence:

A Test of the Critical Period Hypothesis for Second Language Acquisition [1]

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 these patterns (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 is whether these patterns meet the conditions for concluding that a critical period constrains learning in a way predicted by the theory. A critical period minimally entails two characteristics: (1) high level of preparedness for learning within a specified developmental period to assure the domain is mastered by the species, and (2) lack of preparedness outside of this period (Bornstein, 1989; Columbo, 1982). The

consequence of these conditions is that the relation between learning and age is different inside and outside of 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 those 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 immigration, length of residence, and social and linguistic backgrounds of the participants (see 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 native-like competence in the second language (Birdsong, 1992; Bongaerts, Planken, & Schils, 1995; Ioup, Boustagui, 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, Yeni-Komshian, & Liu, 1999). Whether such evidence is considered damaging to the critical period hypothesis depends on the stringency of the criteria for defining the boundaries (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 a monotonic decline with age.

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

An alternative interpretation to the critical period hypothesis is that second language learning becomes compromised with age, potentially because of factors not specific to language that 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, and 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 either of the language itself or through the language into other educational content (Bialystok & Hakuta, 1994; Flege, Yeni-Komshian, & Liu, 1999). Flege and his colleagues report complex effects of educational programs, including a study in which age of learning effects disappeared when education was controlled (Flege, Yeni-Komshian, and Liu, 1999)

The second group of factors is the changes in cognition that occur with aging. Although most cognitive domains do not posit critical periods in development, there are nonetheless age-related changes in cognitive processing. Some age-related changes in cognitive processes relevant to language learning are a decreased ability to learn paired-associates (Salthouse, 1992), more difficulty encoding new information (Craik & Jennings, 1992; Park, Lautenschlager, Hedden, Davidson, Smith, & Smith, in press; Rabinowitz, Craik & Ackerman, 1982), and less accuracy in recalling detail as opposed to gist (Hultsch & Dixon, 1990). Kemper (1992) points out that older adults' second-language proficiency, like their first-language proficiency, could also be affected by such factors as working-memory limitations, cognitive slowing, or attentional deficits. All these processes decline with age, and the decline is documented across the lifespan. 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.

The present study examines the effect of age of acquisition on second-language proficiency by studying a very large sample of second-language learners who cover a wide range of ages of initial exposure to English. Minimally, a critical period requires evidence for a clear discontinuity in learning outcome around a specified age. Moreover, this pattern must 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, using data made available on CD-ROM that offered detailed cross-tabulations of 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, data were used from California,

Illinois, Texas, and New York, four of the largest states with large populations of Spanish

speakers. The Chinese data used 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, minimum length

of residence in the United States was set at 10 years. Stevens (1999), who analyzed a 1% Public

Use Sample drawn from the same Census, found that her sample of immigrants reached

asympototic 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. Further

breakdown of sample sizes by level of education are reported in Table 7.

Measures

The census form asks respondents to self-describe their English ability into 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) as well as our own analyses of those data have shown an

acceptable level of correlation between these items and objective measures of proficiency,

ranging on the order of $\underline{r} = .52$ to $\underline{r} = .54$.[2] 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 data by important background variables in testing whether there is discontinuity in the age effect.

The Census questions included in our analysis were age, year of arrival in the U.S., and educational background. The first two allow the computation of age of arrival and length of residence. Independent variables were created from census ordinal variables with 10-19 levels. For modeling purposes, we constructed interval-scale approximates by taking midpoint values for each category. Where appropriate, we tested the robustness of our analyses to the substitution of interval approximates. 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 by assigning the following year-equivalent to the response categories to "Highest Degree" (Q12): "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), "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: (1) Less than 5th grade education, (2) 5th to 8th grade education, (3) High school education without diploma, (4) High school graduate, and (5) College.

Length of residence was estimated from Q10, "Year of Entry" by subtracting the midpoint of each category from 1990, the year when the census was taken. The "Year of Entry"

categories (and in parentheses, the derived length of residence estimate) considered in this analysis are: 1975-1979 (13 years), 1970-1974 (18 years), 1965-1969 (23 years), 1960-1964 (28 years), 1950-1959 (35.5 years), before 1950 (>40 years). To ensure that English ability reflected a stable level of attainment, this analysis excluded individuals with less than ten years of residence in the United States.

Age of immigration was estimated by taking the present Age (Q5), calculating the midpoint of each category, and subtracting each individual length of residence value from it. The Age categories, representing the age at which English language learning began (and in parentheses, the midpoint in each interval) are: 0-2 years (1 year), 3-4 (3.5 years), 5-9 (7 years), 10-14 (12years), 15 (15 years), 16-17 (16.5 years), 18-19 (18.5 years), 20-24 (22 years), 25-29 (27 years), 30-34 (32 years), 35-39 (37 years), 40-44 (42 years), 45-49 (47 years), 50-54 (52 years), 55-59 (57 years), 60-64 (62 years), 65-69 (67 years), 70-74 (72 years), 75-115 (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 the partial $\underline{R}^2$ value. 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, et. al, 1996, p. 339). Regardless of statistical significance, variables added to the regression model must have large partial $\underline{R}^2$ values (that is, they must account for

substantial proportions of variability in the outcome variable) in order to be considered practically significant.

<div align="center">Results</div>

Education, Age of Immigration, and Cohort Effects

To begin, we considered the simple model of English proficiency as a function of education. Four fits were carried out to determine whether English proficiency was best modeled on (a) dummy variables for the 16 categories in the census, (b) dummies for the simpler 5-level categorical variable, (c) a linear term for the derived interval-level variable, or (d) both a linear and quadratic interval-level education term. The five-level education variable provided the best balance between parsimony and model fit (Chinese: $\underline{R}^2 = 0.4221$; Spanish: $\underline{R}^2 = 0.2622$)[3] and so was used in all subsequent analyses including education. The coefficients for this model are shown in Table 1.

In the second step, age of immigration was added to the model with the education dummies. 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 $\underline{R}^2 = 0.0932$; Spanish: Marginal $\underline{R}^2 = 0.0632$). No interaction term added more than 0.0016 to the model $\underline{R}^2$, providing very little evidence for an interaction between these two variables. These results are reported in Table 2.

A cohort variable representing English proficiency differences between those entering the United States in the 1960s and those entering in the 1970s was included next. This analysis was conducted in part to test the validity of our assumption that we would be sampling $\underline{S}$s at their

asymptotic levels of English proficiency by selecting out only those with length of U. S. residence greater than 10 years. The length of residence of the two cohorts differs by an average of 10 years, allowing us to test for the effect of length of residence within the range of the study (i.e., >10 years). There was little indication of either cohort main effects or interaction effects between the cohort variable and age of immigration and education; none of the terms added more than 0.0032 to the model $\underline{R}^2$. 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 includes simple additive effects for the 5-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, Kutner, Nachtsheim, & Wasserman (1996, pp. 474-478). If there is a critical period, then there will be a discontinuity in the regression of English proficiency on age of immigration at the point that marks the end of the critical period (hereafter referred to as the "critical point"). As Neter, et. al, point out, a regression line might be discontinuous at a point X_c due to 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: points outside the critical period can be marked by (1) a drop in English proficiency, (2) a change in slope between age of immigration and English proficiency, or (3) both a mean drop and a slope change. Note that in the slope change model, there are alternative projections for the discontinuity as shown by the two lines labelled (a) and (b). Johnson and Newport's data as re-analyzed in Bialystok & Hakuta (1994), for example, resemble model (b) with the critical point being at age 20.

The possibility of such discontinuities was tested by the following variables in our regression model (Neter, et. al, 1996, p. 478):

CHANGE IN MEAN
(To test for change in $\quad = \begin{cases} 1 \text{ if AGE OF IMMIGRATION} \geq \text{CRITICAL POINT} \\ 0 \text{ if AGE OF IMMIGRATION} < \text{CRITICAL POINT} \end{cases}$
mean of regression line)

CHANGE IN SLOPE
(To test for change in $\quad = (\text{CHANGE IN MEAN}) * \begin{pmatrix} \text{AGE OF IMMIGRATION -} \\ \text{CRITICAL POINT} \end{pmatrix}$
slope of regression line)

Two different ages were used to define the critical point: ages 15 and 20 at the time of immigration. The first point at age 15 corresponds to the typical onset of puberty. This age has become the standard empirical cutoff, following the influential study by Johnson and Newport (1989). The second point at age 20 was based primarily on visual inspection of the local regression curves reported later in this paper (in Figures 2 and 3) suggesting that if discontinuities existed, they would be at an age later than puberty (cf., Bialystok & Hakuta, 1994). Model estimates were performed separately for each of these putative critical points.

Evidence for either a significant break in mean score 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 similar findings in tests for a critical period ending at age 20. In no case does either the change in mean or change in slope add more than 0.0043 to the model $\underline{R}^2$.

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 $\underline{R}^2$ 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[4] 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 fit a straight line or parametric curve to the data at hand, local

regression provides an individual model fit for each point in the dataset. 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 to the question of the existence of a critical period.

In local regression modeling we specify a "smoothing parameter" to determine the size of

a "neighborhood"[5] of nearby data used to determine the value of the regression line at each

point.[6] As the smoothing parameter increases, a larger local neighborhood is used for

determining the fit at each data point, therefore increasing the "smoothness" of the regression

line.[7] Typical values chosen for smoothing parameters range from .25 to .75. Local regression

models reported here were run using both these values in order to test both extremes of

smoothness. As shown in Table 7, using the smaller smoothing parameter (.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. Although we report

quantitative results using both smoothing parameters, we provide figures only for the model

estimated with the larger value.

Local regression plots of English proficiency on age of immigration, separately for each education group, appear in Figures 2 and 3. The curves indicate 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 proposed as the close of the putative critical period, nor is there evidence suggesting the variation in older learners to be random – proficiency continues to decline into adulthood.

The apparent linearity of these plots is confirmed by considering the gain in $\underline{R}^2$ that is obtained by including a nonparametric form to model the relationship between English proficiency and age of immigration for each education group. Table 5 contains $\underline{R}^2$ figures 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.

<u>Discussion</u>

The critical period is a popular way of explaining differences between the apparent success of children and failure of adults at second-language acquisition. In the U.S., 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 this hypothesis, particularly searching for evidence of discontinuity in English proficiency attainment across a large sample of participants. Using both 15 years and 20 years as hypothesized cut-off points for the end of the critical period, we found no evidence of a change in language learning potential at those times. Instead, the most compelling finding was that the degree of success in second language acquisition steadily declines throughout the life span.

In addition to the effect of age of immigration, these data show the importance of socioeconomic factors, in particular, the amount of formal education, in predicting the learning

of English by immigrants. The years of formal education added substantial amounts to the explanation of variance in both groups and did not interact with other factors. The linear decline across age 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 the decline in second-language proficiency with increasing age of initial exposure is a real function that describes performance in a large population base. The pattern of decline, however, failed to produce the necessary discontinuity that is the essential hallmark of a critical period.

References

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

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.), Second language acquisition and the critical period hypothesis. Mahwah, NJ: Erlbaum.

Birdsong, D., & Flege, J.E. (2000). 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 (pp. 30-50). Clevedon: 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., and 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. Bailey, J. T. Bruer, F. J. Symons & J. W. Lichtman (Eds.), Critical thinking about critical periods. (pp. 193-205). Baltimore: Brookes Publishing Co.

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.), <u>Handbook of the psychology of aging</u> (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. <u>Studies in Second Language Acquisition</u>, <u>16</u>, 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. <u>Language Learning</u>, <u>42</u>, 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. <u>Cognitive Psychology</u>, <u>21</u>, 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. <u>Cognition</u>, <u>39</u>, 215-258.

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

Kominski, R. How Good is "How Well"? (1989). An Examination of the Census English-Speaking Ability Item. Paper presented at the Annual Meeting of the American Statistical Association, August 6-11, Washington, D.C.

Krashen, S. (1973). Lateralization, language learning, and the critical period. <u>Language Learning</u>, <u>23</u>, 63-74.

Lenneberg, E.H. (1967). <u>Biological foundations of language</u>. New York: Wiley.

Neter, J., Kutner, M., Nachtsheim, C., & Wasserman, W. (1996). <u>Applied Linear Statistical Models</u>. Chicago: Irwin.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. <u>Journal of Psycholinguistic Research</u>, <u>5</u>, 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. <u>Psychology and Aging</u>.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. <u>Language Learning</u>, <u>30</u>, 449-472.

Patkowski, M. (1994). The critical age hypothesis and interlanguage phonology. In M. Yavas (Ed.), <u>First and second language phonology</u> (pp. 205-221). San Diego: Singular Publishing Group.

Penfield, W., & Roberts, L. (1959). <u>Speech and brain mechanisms</u>. Princeton, NJ: Princeton University Press.

Pinker, S. (1994). <u>The language instinct</u>. New York: W. Morrow and Co.

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

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

Singleton, D. & Lengyel, Z. (Eds.). (1995). <u>The age factor in second language acquisition: a critical look at the critical period hypothesis</u>. Clevedon: 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). Detailed cross-tabulations of selected language groups for states: 1990. [CD-ROM] 1990 Census of Population and Housing. Washington, DC: Bureau of the Census, Population Division, Decennial Programs Coordination Branch.

Figure 1. Three alternative predictions of the critical period hypothesis.
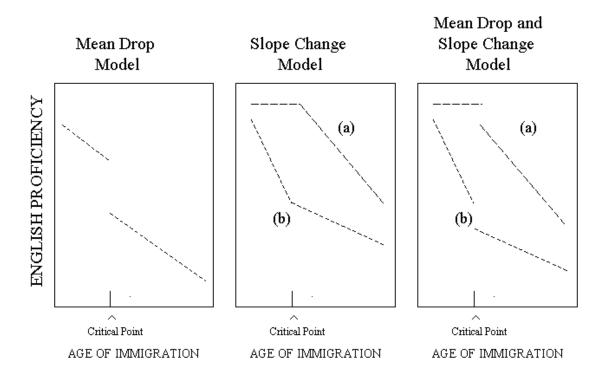
Figure 2. Loess fits (span = .75) for English proficiency by age of immigration, separately by education level. Data are for Chinese-speaking immigrants.
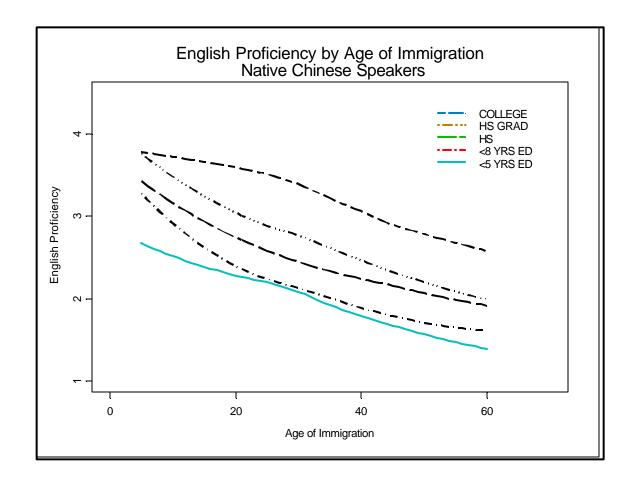
Figure 3. Loess fits (span = .75) for English proficiency by age of immigration, separately by education level. Data are for Spanish-speaking immigrants.
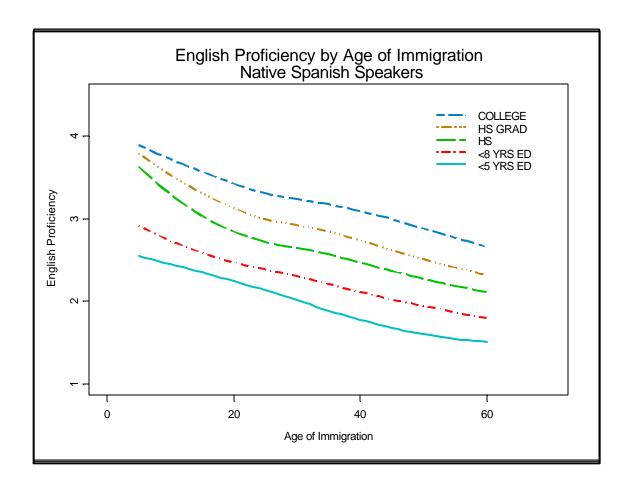

English Proficiency by Age of Immigration
Native Spanish Speakers

Table 1

Regression coefficients for English Proficiency on Education, Spanish- and Chinese-speaking Immigrants

Chinese[1]

Variable	Parameter Estimate	Standard Error	SS	F	p
Intercept	1.7431	0.00417	96590	174334.0	<.0001
ED(5-8 years)	0.2493	0.00624	884	1596.0	<.0001
ED(some HS)	0.7324	0.00586	8659	15628.1	<.0001
ED(HS grad)	1.0693	0.00548	21071	38030.6	<.0001
ED(Some College)	1.7398	0.00451	82450	148813.0	<.0001

[1] $R^2 = 0.4221$

Spanish[1]

Variable	Parameter Estimate	Standard Error	SS	F	p
Intercept	2.0573	0.00136	1796840	2293711.0	<.0001
ED(5-8 years)	0.3484	0.00184	28171	35961.6	<.0001
ED(some HS)	0.8710	0.00196	154633	197393.0	<.0001
ED(HS grad)	1.1708	0.00209	244933	312664.0	<.0001
ED(Some College)	1.4445	0.00198	417988	533572.0	<.0001

[1] $R^2 = 0.2622$

Table 2

Regression coefficients for English Proficiency on Education and Age of Immigration, Spanish- and Chinese-speaking Immigrants

Chinese

Variable	Parameter Estimate	Standard Error	SS	F	p	R-squared Partial	Total
Intercept	2.69395	0.01185	23924	51657.2	<.0001		
ED(5-8 years)	0.03791	0.01731	2	4.8	0.0285		
ED(some HS)	0.51324	0.01513	533	1151.0	<.0001		
ED(HS grad)	0.98867	0.01392	2337	5045.6	<.0001		
ED(Some College)	1.30098	0.01234	5144	11106.3	<.0001		0.4221
Age of Immigration	-0.02186	0.00026	3325	7180.0	<.0001	0.0932	0.5153

Spanish

Variable	Parameter Estimate	Standard Error	SS	F	p	R-squared Partial	Total
Intercept	2.63091	0.00324	469497	657397.0	<.0001		
ED(5-8 years)	0.22956	0.00441	1939	2715.5	<.0001		
ED(some HS)	0.88544	0.00434	29691	41574.1	<.0001		
ED(HS grad)	1.15842	0.00448	47812	66947.3	<.0001		
ED(Some College)	1.31456	0.00427	67572	94615.9	<.0001		0.2622
Age of Immigration	-0.02022	0.00010	26566	37197.7	<.0001	0.0632	0.3254

Table 3

Regression coefficients for English Proficiency on Education, Age of Immigration, and Critical Period Variables, Chinese-speaking Immigrants (Critical Period = Age 15)

_____

| | | | | | | R-squared | |
Variable	Parameter Estimate	Standard Error	SS	F	p	Partial	Total
Intercept	2.76989	0.00923	41863	90139.4	<.0001		
ED(5-8 years)	0.10036	0.00576	141	303.8	<.0001		
ED(some HS)	0.44701	0.00551	3061	6591.1	<.0001		
ED(HS grad)	0.73278	0.00521	9201	19812.2	<.0001		
ED(Some College)	1.26844	0.00455	36074	77674.7	<.0001		0.4221
Age of Immigration	-0.02640	0.00067	712	1533.3	<.0001	0.0932	0.5153
Change in Mean	0.05804	0.00424	87	187.8	<.0001	0.0003	0.5156
Change in Slope	0.00227	0.00068	5	11.1	0.0009	0	0.5156

_____

Table 4

Regression coefficients for English Proficiency on Education, Age of Immigration, and Critical Period Variables, Spanish-speaking Immigrants (Critical Period = Age 15)

_____

| | | | | | | R-squared | |
Variable	Parameter Estimate	Standard Error	SS	F	p	Partial	Total
Intercept	3.02532	0.00361	498299	700445.0	<.0001		
ED(5-8 years)	0.26340	0.00177	15802	22213.1	<.0001		
ED(some HS)	0.67604	0.00192	88362	124209.0	<.0001		
ED(HS grad)	0.94236	0.00206	148943	209366.0	<.0001		
ED(Some College)	1.19965	0.00196	265741	373544.0	<.0001		0.2622
Age of Immigration	-0.04573	0.00027	21033	29565.9	<.0001	0.0632	0.3254
Change in Slope	0.02730	0.00028	7004	9844.7	<.0001	0.0043	0.3297
Change in Mean	-0.05045	0.00185	531	745.8	<.0001	0.0002	0.3300

_____

Table 5

Regression coefficients for English Proficiency on Education, Age of Immigration, and Critical Period Variables, Chinese-speaking Immigrants (Critical Period = Age 20)

Variable	Parameter Estimate	Standard Error	SS	F	p	R-squared Partial	Total
Intercept	2.72891	0.00750	61569	132559.0	<.0001		
ED(5-8 years)	0.09922	0.00576	138	296.4	<.0001		
ED(some HS)	0.44600	0.00551	3045	6556.2	<.0001		
ED(HS grad)	0.73139	0.00521	9156	19713.1	<.0001		
ED(Some College)	1.26715	0.00455	35962	77427.8	<.0001		0.4221
Age of Immigration	-0.02206	0.00038	1558	3353.6	<.0001	0.0932	0.5153
Change in Mean	0.03465	0.00374	40	85.9	<.0001	0.0002	0.5155
Change in Slope	-0.00245	0.00040	17	37.5	<.0001	0.0001	0.5156

Table 6

Regression coefficients for English Proficiency on Education, Age of Immigration, and Critical Period Variables, Spanish-speaking Immigrants (Critical Period = Age 20)

_____

| | | | | | | R-squared | |
Variable	Parameter Estimate	Standard Error	SS	F	p	Partial	Total
Intercept	2.96103	0.00295	716088	1006113.0	<.0001		
ED(5-8 years)	0.26273	0.00177	15723	22090.5	<.0001		
ED(some HS)	0.67541	0.00192	88201	123923.0	<.0001		
ED(HS grad)	0.94321	0.00206	149239	209683.0	<.0001		
ED(Some College)	1.20114	0.00196	266544	374498.0	<.0001		0.2622
Age of Immigration	-0.03913	0.00016	43806	61548.7	<.0001	0.0632	0.3254
Change in Slope	0.02061	0.00018	9116	12807.4	<.0001	0.0042	0.3296
Change in Mean	0.02030	0.00188	83	116.9	<.0001	0.0000	0.3296

_____

Table 7

Model Summaries: Nonlinear and Linear Models of English Proficiency on Age of Immigration by Educational Level, Spanish- and Chinese-speaking Immigrants.

	Chinese			Spanish		
	LOESS[1] span=0.25	LOESS span=0.75	Linear[2]	LOESS span=0.25	LOESS span=0.75	Linear
Less than 5 years Education						
<u>n</u>	31790	31790	31790	424554	424554	424554
Equivalent # 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-squared	0.1542	0.1507	0.1488	0.0706	0.0706	0.0687
5-8 years Education						
n	25757	25757	25757	511865	511865	511865
Equivalent # 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-squared	0.1223	0.1207	0.1097	0.0467	0.0465	0.0449
HS w/o Graduation						
n	32786	32786	32786	392147	392147	392147
Equivalent # 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-squared	0.1540	0.1536	0.1449	0.1361	0.1357	0.1180

[1]Local regression model of English proficiency on age of immigration (span = 0.25).
[2]Linear regression model of English proficiency on age of immigration.

Table X7, con't

Model Summaries: Nonlinear and Linear Models of English Proficiency on Age of Immigration by Educational Level, Spanish- and Chinese-speaking Immigrants

	Chinese			Spanish		
	LOESS[1] span=0.25	LOESS span=0.75	Linear[2]	LOESS span=0.25	LOESS span=0.75	Linear
HS Graduation						
n	43848	43848	43848	308507	308507	308507
Equivalent # 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-squared	0.2597	0.2570	0.2493	0.1547	0.1538	0.1415
College						
n	190263	190263	190263	379244	379244	379244
Equivalent # 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-squared	0.1602	0.1593	0.1508	0.1193	0.1191	0.1142

[1]Local regression model of English proficiency on age of immigration (span = 0.25).
[2]Linear regression model of English proficiency on age of immigration.

Notes

---

[1] This study was supported in part by a grant from the Spencer Foundation to the first author.

[2] To further substantiate the relationship between the census item and objective measures of English Proficiency, we obtained the data collected in the National Content Test (NCT) and its re-interview, 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, $r = .68$, $p < .001$.

[3] The 16-level categorical education variable provided the best fit (Chinese: $R^2 = .4389$; Spanish: $R^2 = 0.2667$), followed by the 5-level variable (Chinese: $R^2 = 0.4221$; Spanish: $R^2 = 0.2622$), the linear and quadratic fit (Chinese: $R^2 = 0.4096$; Spanish: $R^2 = 0.2556$), and the simple linear fit (Chinese: $R^2 = 0.4023$; Spanish: $R^2 = 0.2479$). 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 = 0.0168; Spanish: R^2 difference = 0.0045).

[4] All local regression modeling was carried out using releases 3.4 and 4.0 of S-Plus Advanced Data Analytic Software. Local regression fits utilized the LOESS function; Loess curves were plotted using predicted values from Loess models.

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

[6] An intuitive way to think about this neighborhood is to consider a window (with length equal to the smoothing parameter) 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.

[7] In terms of the tradeoff between bias and variance of fit, larger smoothing parameters decrease model bias and increase model variance. Choosing an extremely small smoothing parameter value can result in bias due to overfitting the model to the data in hand. This study was supported in part by a grant from the Spencer Foundation to the first author.

[7] To further substantiate the relationship between the census item and objective measures of English Proficiency, we obtained the data collected in the National Content Test (NCT) and its re-interview, 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, $r = .68$, $p < .001$.

[7] The 16-level categorical education variable provided the best fit (Chinese: $R^2 = .4389$; Spanish: $R^2 = 0.2667$), followed by the 5-level variable (Chinese: $R^2 = 0.4221$; Spanish: $R^2 = 0.2622$), the linear and quadratic fit (Chinese: $R^2 = 0.4096$; Spanish: $R^2 = 0.2556$), and the simple linear fit (Chinese: $R^2 = 0.4023$; Spanish: $R^2 = 0.2479$). 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 = 0.0168; Spanish: R^2 difference = 0.0045).

[7] All local regression modeling was carried out using releases 3.4 and 4.0 of S-Plus Advanced Data Analytic Software. Local regression fits utilized the LOESS function; Loess curves were plotted using predicted values from Loess models.

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

[7] An intuitive way to think about this neighborhood is to consider a window (with length equal to the smoothing parameter) 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.

[7] In terms of the tradeoff between bias and variance of fit, larger smoothing parameters decrease model bias and increase model variance. Choosing an extremely small smoothing parameter value can result in bias due to overfitting the model to the data in hand.