Journal of Experimental Psychology: General
1994, Vol. 123, No. 3, 264-283

0096-3445/94/$3.00

Fifty Years of Language Maintenance and Language Dominance in Bilingual Hispanic Immigrants

Harry P. Bahrick, Lynda K. Hall, Judith P. Goggin, Lorraine E. Bahrick, and Stephanie A. Berger

Spanish language tests of 801 Cuban and Mexican immigrants showed no evidence of language loss during 50 years of U.S. residence; a few years after immigration, their English vocabulary approximated that of English monolinguals. The critical-age hypothesis was not supported for the acquisition of English vocabulary when English schooling and language usage were controlled by multiple regression. Most Ss continued to speak about as much Spanish as English; but read, wrote, and heard (on television and radio) far more English than Spanish. Under these conditions, Ss maintained Spanish dominance on tests of vocabulary recognition, lexical decision, and oral comprehension. Dominance was task specific and shifted to English on a category generation task about 12 years after immigration. No evidence of bilingual language interference was found; this is attributed to the strong Spanish foundation of the participants.

We explored first-language (L1) losses associated with living in a bilingual environment over a period of 50 years. We tested 801 Cuban and Mexican immigrants between 4 months and 50 years after their immigration to the United States. Most of the tests measured Spanish language skills. However, several subtests were also given in English to provide information about second-language (L2) acquisition and language dominance. The subjects varied widely in many aspects of their language history, including age at immigration, number of years of U.S. residence, amount of education received in Spanish and English schools, and extent of continued exposure to and use of Spanish after immigration. The influence of these variables on the maintenance of the Spanish language, on the acquisition of English, and on language dominance was also explored.

We address several issues related to bilingualism. A key question is whether bilingual individuals sustain deficits in one or both languages in comparison with monolingual individuals. This issue is of theoretical and practical importance. Psycholinguistic theory must account for facilitative and interfering effects associated with bilingualism, and bilingual education programs should be designed to enhance facilitation and diminish interference. McLaughlin (1977), in a comprehensive review, concluded that there is no proof that interference is an inevitable consequence of bilingualism, a conclusion that has been supported by recent research (e.g., Pearson, Fernandez, & Oller, 1993). However, many studies have reported language deficits (e.g., Berelson & Steiner, 1963; Fritz & Rankin, 1934; Macnamara, 1966; Pan & Gleason, 1986) and interference in the form of hybrid language, for example, Spanglish (Cronnell, 1985; Merino, 1983; Queralt, 1984; Ziros, 1976).

We expected interference between the languages to diminish as a function of the degree of mastery of L1 at the time of L2 acquisition. This hypothesis is consistent with interference theory and has been supported by the results of list-learning experiments (McGeoch & Irion, 1952). It is also consistent with reports of bilingual interference observed in individuals who have limited schooling and in those who acquire two languages early in life, before L1 has been adequately mastered (Cummins, 1979). We investigated maintenance of L1 and compared Spanish and English performance of bilingual subjects with the performance of Spanish and English monolinguals of comparable age and education.

A second controversial issue involves the concept of a critical age (Lenneberg, 1967) or a sensitive period (Oyama, 1976, 1978) for L2 acquisition. Although the conclusion that younger learners have an advantage in mastering the phonology of a second language has been well established

Harry P. Bahrick, Lynda K. Hall, and Stephanie A. Berger, Department of Psychology, Ohio Wesleyan University; Judith P. Goggin, Department of Psychology, University of Texas—El Paso; Lorraine E. Bahrick, Department of Psychology, Florida International University.

The writing of this article was supported by National Science Foundation Grants BNS–8417788 and BNS–9119800, awarded to Harry P. Bahrick; National Institute of Mental Health Grant MH47167 and National Institute of General Medical Science Grant RR08012, awarded to Judith P. Goggin; and National Institute of Child Health and Human Development Grant R01HD25669, awarded to Lorraine E. Bahrick. We thank Guillermo Arango, Jose Ballon, Dolores Smith, Irene Jenkins Weber, and Cara Wellman for contributions to test development and scoring; Soledad Cline, Baltazar Gonzalez, Patricia Gonzalez, Barbara Horowitz, Nancy Marsh, Pastor Jorge Mata, Francisco Medina, and Inese Neiders for assistance with recruiting subjects; Shirley Bodi, Soraya Cadi, Tracy Curtis, Tammy Daily, Diane Gatton Daum, John Dunlosky, Lyn Goff, Jennifer Hlavin, Sylvia Jurado, Marisol Lara, Ileana Llapur, Raquel Llobera, Jodie Gordon Lucas, Teresa Rosello, Teresa Schulze, Carolynn Tomory, Lisa Vandegrift, and John Wiebe for assistance with collecting, scoring, entering, and analyzing data; and Jeff Edwards, Robert MacCallum, and Thomas Nelson for methodological advice. We also thank Morton Ann Gernsbacher and an anonymous reviewer for their comments.

Correspondence concerning this article should be addressed to Harry P. Bahrick, Department of Psychology, Ohio Wesleyan University, Delaware, Ohio 43015.

(Asher & Garcia, 1969; Fathman, 1975; Oyama, 1976; Seliger, Krashen, & Ladefoged, 1975), it is uncertain whether this advantage extends to other aspects of L2 acquisition. Johnson and Newport's (1989) results support the concept of a critical age for phonology and for syntax. Subjects who began learning L2 in early childhood showed superior eventual attainment. Johnson and Newport concluded that findings favoring adult learners (Asher & Price, 1967; Olson & Samuels, 1973; Snow & Hoefnagel-Hohle, 1977, 1978) are based on the early stages of learning, and that the advantage is short-lived.

In this article, we examine the critical-age hypothesis in regard to tasks of lexical decisions, vocabulary recognition, category generation, and oral comprehension. We expected acquisition of these L2 tasks to be facilitated as a function of the degree of mastery of comparable L1 tasks. Contrary to the critical-age hypothesis, we therefore expected older adolescents to be advantaged, not disadvantaged, in comparison with younger adolescents. The rationale for our prediction is that acquisition of L1 vocabulary involves concept formation, and most linguistic concepts are common to both languages. Acquisition of L2 vocabulary for concepts available in L1 requires only the learning of new responses. For concepts not already available, acquisition of L2 vocabulary also requires concept formation and therefore proceeds at a slower pace. Thus, the rate of L2 vocabulary acquisition should increase as a function of the mastery of L1 vocabulary and be superior for individuals who immigrated as older adolescents. Performance on the tasks of vocabulary recognition, category generation, lexical decision, and oral comprehension all depend on the available vocabulary, and the above rationale therefore applies to all four of these tasks. However, phonology, and perhaps syntax, present a different problem. For these domains, transfer is more complex, and more interference and competition are involved (MacWhinney, 1992). An optimum maturational level, that is, a critical age or sensitive period, may therefore have greater influence on L2 acquisition. This hypothesis is consistent with Cummins' (1981) view that the concept of a critical age applies only to oral fluency and phonology but not to cognitive aspects of language acquisition. It is also compatible with Hurford's (1991) inference of several overlapping critical periods; one for phonology and later periods for grammatical structure and vocabulary acquisition.

Finally, we explore the nature of language dominance and the appropriate measurement of dominance. Some researchers (Oyama, 1978; Spolsky, Sigurd, Sako, Walker, & Arterburn, 1968; Whitson, 1972) have suggested that a single indicant, for example, an oral comprehension test, can serve as a valid, general measure of language dominance; others (Cooper, 1969; Fishman, 1968) have contended that global measures inadequately describe bilingual performance and that language dominance is task and context specific. Some investigators have used reaction time or speed of performance as an indicant of dominance (Cooper, 1969; Ervin, 1961; Lambert, 1955; Macnamara, 1967; Mägiste, 1986); in contrast, others have criticized measures based on speed (Fishman, 1968) and have used power tests that disregard response time (Johnson & Newport, 1989; Lemmon & Goggin, 1989).

We believe that the interrelations among indicants of language proficiency must be understood in order to interpret and reconcile findings of language dominance based on various tasks. As a consequence, in the present investigation we determined language dominance on the basis of four tests and established the interrelations among them. We were particularly interested in testing the hypothesis that performance on tasks that require independent and speeded retrieval of words (e.g., category generation) is more strongly affected by the individual's recent language environment than is performance on self-paced tasks that require only comprehension (e.g., vocabulary recognition). Frequency and recency of retrieval strongly affect access to the lexicon (particularly speeded access) on the basis of a cumulative and generalized priming effect but do not affect availability of the lexicon. Therefore, performance on self-paced tasks should reflect the relative size of the L1 lexicon compared with that of L2, with less regard to frequency of recent retrieval. Confirmation of task-specific dominance effects in accord with these predictions provides the basis for a more general understanding of the multifaceted phenomenon of language dominance.

Background and Rationale

This investigation is the sixth in a series concerned with the maintenance of naturalistically acquired knowledge. Previous discussions of the naturalistic approach (Bahrick, 1989; Bahrick & Karis, 1982) have pointed out that this methodology involves trade-offs in which experimental controls over independent variables are sacrificed to explore phenomena of learning and memory that do not lend themselves to laboratory exploration. Investigations of maintenance of knowledge cover long time periods that make variables governing acquisition and rehearsal of semantic memory content extremely difficult to control experimentally. The downside of this trade-off is the potential confounding of variables. In "experiments of nature," it usually is not possible to recruit participants whose characteristics approach an orthogonal design in regard to the variables of interest. For example, in this investigation the age at entry to the United States and the number of years of education before immigration covary, so the effects of one of these variables independently of the other cannot be determined. Because participants are not randomly assigned to groups, there are unavoidable cohort or selection effects, not all of which can be controlled statistically. Despite these limitations, understanding of long-term maintenance of knowledge has progressed during the past 20 years through naturalistic study. We discuss relevant problems of interpretation in the context of the present investigation.

We recruited only Mexican-American and Cuban-American subjects, for several reasons. Literacy levels, school systems, and language usage differ widely among Latin American countries. We wanted to minimize these sources of variation because they were not the focus of our interest.

We would, therefore, have preferred participants from a single country of origin. Cuban-Americans offered the advantages of easy accessibility and the inclusion of many well-educated professionals who usually are underrepresented among first-generation immigrants. However, Cuban arrivals in the United States were concentrated during two discrete time periods (Queralt, 1984), and the intermittent flow of Cuban immigration makes it nearly impossible to recruit representatives of all portions of the 50-year retention interval. In contrast, the flow of Mexican immigration has been continuous, which permitted testing at all retention intervals. Inclusion of Mexican-Americans also enabled us to collect pilot data from Mexican monolinguals who had not experienced a language change. The difficulty of recruiting and testing participants in Cuba prevented us from obtaining data for that group. However, we did test immigrants from Cuba and Mexico within 3 months of their arrival in the United States. These Spanish test scores served as control data to evaluate deficits in Spanish that were due to diminished use of the language during the postimmigration period and that were due to language interference associated with living in a bilingual environment. Our previous findings (Bahrick, 1984) have shown that much of the knowledge of Spanish acquired by native English speakers in U.S. schools is lost within 3–5 years, unless rehearsal continues. One of our goals in the present study was to determine whether such losses also apply to the knowledge of L1 when adolescents move to a different language environment and greatly diminish their use of L1.

In the El Paso, Texas area we tested primarily Mexican participants, and in the Miami, Florida area we tested primarily Cubans. The main advantage of these sites—that is, the availability of subjects—was accompanied by an inevitable disadvantage. Most of these subjects were exposed to and used Spanish extensively. Because one of our goals was to explore the long-term maintenance of a native language when that language is rarely used, we needed to go elsewhere to obtain information about the effects of living in an English language environment in relative isolation from Spanish. Our solution was to send research assistants to midwestern cities that had small Hispanic communities, on the assumption that the exposure to Spanish, particularly in work and social settings, would be more limited. The most useful sources of information for locating potential participants were the naturalization records of U.S. district courts. These archival records list the country of origin and the address (at the time of naturalization) of naturalized citizens, who were then traced through telephone directories. People who agreed to be tested frequently referred us to others who met our criteria. However, even among the participants in the Midwest, few were totally isolated from the Spanish language. The relatively high level of continuing postimmigration use of Spanish is therefore a limiting parameter to this investigation.

Our large sample size constrained the nature of the test we used, most importantly, its length. Testing time in excess of 2 hr would have substantially increased the difficulty of recruiting participants and might also have impaired the validity of the test by overtaxing participants' patience and motivation. In addition, we sought a test that could be scored objectively and administered by individuals without extensive language training.

The above considerations led us to test some aspects of language proficiency and to neglect others. The choices reflect our theoretical interests, the limited length of the test, and the potential for obtaining reliable measurements. Proficiency in spelling and punctuation were not tested because these skills are tangential to our theoretical interests. We chose not to assess level of written or oral expression, because the testing procedures would have been too time consuming and the scoring too unreliable. We did not test phonology, because we assumed it to be well maintained for L1, and the critical-age hypothesis is well established for L2. Our emphasis was on comprehension of the written language; but subtests measuring oral comprehension, recognition and retrieval of vocabulary, and knowledge of grammar were also included. We administered separate tests for vocabulary recognition and for category generation to test our predictions in regard to the differential effects of bilingualism on accessibility versus availability of the lexicon. In addition, there was a subtest for identifying anglicisms to explore certain aspects of language interference associated with bilingualism.

We administered four subtests, selected on the basis of theoretical and practical considerations, in both Spanish and English, to measure language dominance. The tests involved making lexical decisions, recognizing vocabulary, generating exemplars of semantic categories, and answering questions about the content of passages presented auditorily in the presence or absence of noise. In accord with our hypothesis that comprehension of L1 vocabulary would be maintained more effectively than speeded retrieval of L1 vocabulary, we expected language dominance to shift to L2 relatively soon after immigration for the category generation task and much later for the vocabulary recognition and lexical decision tasks. The Spanish version of the subtests served two purposes: to indicate the level of maintenance of Spanish and to measure bilingual dominance for that aspect of language.

Although we tried to hold mean testing time to 90 min, our goal was to develop a power test rather than a speed test. The focus of our inquiry was on maintenance of knowledge, not on variations of processing or retrieval speed among participants, whose ages ranged from 13 to 70 years. Therefore, with the exception of the category generation task, the subtests were self-paced without imposition of time limits.

Method

Test Construction

Test construction required 18 months. The purpose of the preliminary work was to ensure that the subtests were appropriately difficult, had adequate reliability and internal consistency, and could be performed within the available time. Furthermore, we attempted to equate the difficulty of the English and Spanish versions of each dominance test by (a) choosing tasks that seemed to be inherently equated (e.g., generating exemplars for the same

semantic categories in both languages) and (b) testing monolingual pilot subjects. To be sure, neither method assures absolute equality. Cultural variables may result in differential exposure to exemplars of the same categories for Mexican or Cuban vs. U.S. natives, and the two groups of monolingual control subjects may differ in various ways. Hamers and Blanc (1989) suggested that differential cultural effects are best avoided by relying on monolingual control data in each language rather than by directly comparing performance in the two languages. However, we believe that the effect of cultural differences can be minimized if tasks are chosen with care, and that lack of comparability of monolingual control groups is then a greater source of error than are general cultural differences. Inferences of language dominance based on superior performance on the English or Spanish version of a particular subtest require careful evaluation; we address these concerns later. However, we can compare the magnitude of the effects of several independent variables on indicants of language dominance without concern about having achieved complete equality in the two versions of a test.

Pilot Subjects

A total of 384 pilot subjects were tested at five sites. Mexican monolingual subjects were tested in Mexico City, Mexico; Ciudad Juárez, Mexico; and Miami, Florida; English monolingual subjects were tested at Ohio Wesleyan University; bilingual subjects were tested at the Ohio State University.

Individual Subtests

Text Comprehension

This subtest evaluated both paragraph comprehension and sentence comprehension in Spanish. Items were taken from the *Preparación Para el Examen de Equivalencia de la Escuela Superior (en Español;* High School Equivalency Examination; Lanzano & Sagrista, 1984). The pilot version of the subtest included 13 paragraphs and 49 questions. Participants were classified into terciles based on the total subtest score, and the percentage of people who passed each question was established separately for each tercile. Whole paragraphs and individual questions were then retained or eliminated on the basis of their contribution to tercile discrimination and their difficulty level.

The paragraph comprehension items tested understanding of gist, rather than verbatim memory of the narrative's surface structure. Thus, most questions addressed the theme of the entire paragraph or the meaning of large segments of text. Questions that entailed more logical analysis or cultural interpretation rather than language comprehension per se were also avoided. However, these variables could not be separated completely, because tests of language comprehension inevitably reflect general intellectual functioning.

Grammar Recall and Recognition

Grammar skills were assessed by both recall and recognition, with questions from the Spanish Placement Examination of the University of Texas at El Paso (1980); other questions were written by the experimenters. The recall format used a modified cloze technique (Oller, 1973). A sentence was presented, along with the stem of an adjective or the infinitive form of a verb, and subjects were required to supply the adjective or verb form that would fit the sentence. The recognition format included two types of questions: (a) identification of 5 grammatically correct sentences from a list of 10; and (b) identification of the 1 of 5 alternative words that could not complete a sentence grammatically. The method of analysis and the criteria for item selection on the basis of pilot data were the same as for the text comprehension subtest.

Anglicisms

This test was included because we expected that continuing residence in an English language environment might not only arrest or impair knowledge of various aspects of Spanish but also lead to failures to discriminate Spanish modifications of English words (e.g., factorias, defrostiar, correctar) from authentic Spanish words and expressions. Anglicisms supplied by Hispanic collaborators in Miami and El Paso were selected on the basis of a consensus among collaborators. Because the anglicisms used by Cuban and Mexican immigrants differ to some extent, we developed two versions of this test and administered the version appropriate to the nationality of each participant.

The final version of the Cuban anglicism test consisted of two parts, in which subjects were required to identify: (a) 15 anglicisms from a list of 30 words, half of which were authentic Spanish words from an unabridged Spanish dictionary (Garcia-Pelayo & Durand, 1983) and half of which were anglicisms, and (b) 5 anglicisms from among 10 underlined passages embedded in five sentences. The Mexican anglicism test followed the same format, except the first section contained 10 anglicisms in a list of 22 words. For both tests, appropriate difficulty level was established on the basis of pilot data.

Category Generation Dominance Test

From among the semantic categories birds, fruits, articles of clothing, and body parts in the pilot test, clothing and body parts were selected for the final test because the response rate for these categories was quite low after 3 min. These two categories also appeared to be least affected by cultural variables; that is, the hierarchy of exemplars generated by English monolingual and Hispanic subjects was quite similar.

Subjects generated exemplars for both categories in both languages; however, to control sequence effects, we used an ABBA order for the two languages. Spanish was the first language used for generating exemplars of body parts, and English was the first language for generating the names of articles of clothing.

Lexical Decision and Vocabulary Recognition Dominance Tests

The lexical decision portion of this subtest consisted of 40 words and 40 nonwords. For the English version of the test, the 40 words were selected from Thorndike and Lorge (1952). A frequency range of 4–23 occurrences per million was chosen to avoid floor and ceiling effects. Pronounceable nonwords (e.g., forate) were generated by adding suffixes to or changing letters in actual words, but the letter sequences characteristic of English words were maintained. The final test included 18 real words and the 18 nonwords most frequently mistaken for English words in forced-choice lexical decisions by pilot subjects. For the Spanish version of this subtest, words were chosen to match the frequency of the English words as closely as possible (Eaton, 1961). However, the Eaton frequency norms do not include the lowest frequency words; consequently, low frequency Spanish words were added to the

pool on the advice of expert consultants, and final selections were made so as to achieve comparable performance on the English and Spanish versions by the monolingual control groups. Spanish nonwords (e.g., infloso) were generated by expert consultants in accordance with the procedures used for generating English nonwords.

The vocabulary recognition subtest contained 20 words. Subjects identified a synonym for each target word from among 5 alternatives with the same part of speech. The frequency of occurrence of the alternatives was equal to or greater than that of the target, to diminish the likelihood of errors stemming from a failure to recognize the correct synonym when the meaning of the target word was known.

Oral Comprehension Dominance Test

This subtest was the only one that used auditory rather than visual presentation. Half of the test measured the participant's ability to understand spoken language when it was presented against a background of white noise, a procedure developed by Spolsky and his associates (Gradman & Spolsky, 1975; Spolsky, 1971; Spolsky et al., 1968). Although this procedure has been criticized (Johansson, 1973), others have found it useful for measuring language dominance (Whitson, 1972) and for integrating different sources of linguistic knowledge (Oyama, 1978).

The subjects were presented with short passages in English or Spanish (68–102 words) through earphones. The passages were obtained from the Spanish language editions of the *Reader's Digest* (Hornblower, 1985; Pearson, 1985; Van der Ley, 1985) and from the Stanford Achievement Test (Gardner, Rudman, Karlsen, & Merwin, 1982). Because of differences in pronunciation, expert native readers prepared separate tapes for Cuban and Mexican participants. Subjects first heard a short practice passage and were instructed to adjust the volume for optimum comprehension. They were told that they would hear each passage only once and then were to answer questions about each. Subjects were also warned that some of the passages would be presented against a background of noise that would make them difficult to understand.

Two passages were presented in each language. Language order effects were controlled by following a Spanish–English–English–Spanish sequence. The first passage in each language was presented without noise, the second one with noise. In the latter case, the signal-to-noise ratio was approximately 2:1. The specific passages used in each noise–language condition were counterbalanced across subjects. After listening to each passage, subjects answered four questions about the content. The noise ratio, passages, and questions were selected on the basis of pilot data so as to avoid floor and ceiling effects and to achieve adequate reliability.

Questionnaire

Before taking the test, participants completed a questionnaire regarding their age, gender, birthplace, occupation, date of immigration, and parents' level of education. They also provided detailed information about their education in Cuban or Mexican and U.S. schools, in particular, the number of years and the level of pre- and postimmigration schooling with emphasis on courses in the two languages.

Of special interest was the continued use of Spanish after immigration. Participants estimated the overall percentage of time they used Spanish in conversation at home, at work, at school, in their neighborhood, and with friends. They also estimated their frequency of reading, writing, and listening to the radio and television in Spanish and English by estimating the number of hours they spent per week in each activity for each of five equal postimmigration periods.

Finally, participants rated their language dominance on a 7-point scale on which a rating of 4 designated equal competence in both languages. Separate ratings were obtained for speaking, understanding, reading, and writing.

Participants

Over a period of 3 years, we tested 348 men and 453 women, most of whom were between the ages of 10 and 26 years when they came to the United States. Of these, 264 were tested in the Midwest, 261 in Miami, Florida; and 276 in El Paso, Texas. Table 1 gives a breakdown of age at immigration and years of postimmigration residence in the United States at the time of testing.

Individuals who entered the United States before age 10 were excluded, because we considered a fourth grade literacy level in Spanish necessary to comprehend the written instructions. The upper entry age of 26 served as a general guide to recruiting, but 5 participants who immigrated when they were older than 26 were included because they possessed characteristics that were underrepresented in the sample. This upper age limit reflected our desire to bracket the so-called critical age for language acquisition (Johnson & Newport, 1989; Lenneberg, 1967; Patkowski, 1980) without sacrificing the goal of assessing language maintenance and dominance over a full 50-year postimmigration period.

Testing Procedure

Subjects were informed about the purpose of our investigation during the recruiting interview (which was usually conducted by telephone). They were paid $7–$10 for participating and were given debriefing information at the conclusion of the test. Subjects were tested individually or in groups of up to 3 at diverse locations, including the psychology laboratories of the University of Texas at El Paso, Florida International University, and Ohio Wesleyan University, in space provided by neighborhood churches or public libraries; or in subjects' homes.

Participants first completed the questionnaire with instructions provided in Spanish or English at the subject's option. The subtests were then administered in the following order: text comprehension, grammar, 15-min break, anglicisms, lexical decision, vocabulary recognition, oral comprehension, and category generation. The experimenter timed the initiation and completion of each half of the test but imposed no time limits except in the category generation task. The average completion time was approximately 90 min.

Table 1
Number of Subjects by Age at Immigration and Length of U.S. Residence

Years in United States	Age at immigration			Total
	10–13	14–17	18+	
0–2	12	32	72	116
3–6	28	44	31	103
7–15	92	48	33	173
16–25	86	64	50	200
26–37	58	57	52	167
38+	14	20	8	42
Total	290	265	246	801

Results and Discussion

We report the analysis of data in the following sequence: (a) transformations of independent variables and interrelations among dependent variables; (b) retention functions for Spanish tests, acquisition functions for English, and a comparison of these functions with those of monolingual controls; (c) determination of language dominance for each participant based on each of four subtests, and the relation of dominance to demographic and language usage variables; and (d) multiple regression analyses of the effect of demographic and language usage variables on the maintenance of Spanish and the acquisition of English, and the implications of these effects for changes in language dominance and the critical-age hypothesis.

Analysis of Independent and Dependent Variables

Principal Components Analysis of Independent Variables

The major objective of this analysis was to reduce the large number of independent variables (obtained from questionnaire data) to a more manageable number that could represent the data parsimoniously and could be used to predict performance on the language subtests. Four language activities (reading, writing, listening to the radio, and watching television) were scored in both Spanish and English, but speaking was scored in Spanish only. Speaking in English versus Spanish was represented by one score, because we assumed that this activity is complementary for the two languages. This yielded nine usage estimates, which were obtained for each of five equal postimmigration time periods for a total of 45 language usage variables. Separate principal component analyses for sets of variables pertaining to a single activity were used to reduce and linearly transform each set into a single score that best represented the variance of the original scores. Derived scores accounted for 62%–86% of the variance in the original scores, with 73% as the median value. The derived scores were used in subsequent regression analyses to determine the effects of language usage on performance on each subtest.

Table 2 shows the intercorrelations among the nine transformed usage variables as well as the mean and standard deviation for each variable. For all variables except speaking, mean hours per week spent in the activity are specified separately in each language. Speaking is reported as the mean percentage of time Spanish was spoken.

Several trends are evident from the data. First, English was used three to four times more frequently than Spanish in all activities except speaking. Second, correlations tended to be positive among activities in which the same language is used and negative when different languages were involved. The highest correlations were obtained for writing and reading in the same language (.53 for English and .52 Spanish). The correlations between several other usage variables are statistically significant, but they generally account for less than 10% of relevant variance.

We also compared Mexican and Cuban subjects on demographic and language usage variables. Differences on demographic variables were small but statistically significant because of the large number of subjects. Cuban participants averaged more English education both before immigration (2.28 years vs. 1.80 years) and after immigration (5.47 years vs. 4.06 years). Mexican subjects reported higher amounts of reading and radio and television exposure in both languages. However, the overall pattern of demographic and language usage variables is generally consistent for the two groups.

Interrelations Among Dependent Variables

Data pertaining to performance on each English and Spanish subtest are presented in Table 3. Note that the scores for the oral comprehension test are collapsed across the noise variable. The combined noise and no-noise condition scores yielded the highest reliability, and separate noise deficit scores were not significantly correlated with other independent and dependent variables. At the bottom of this table are the Spearman reliability coefficients for

Table 2
Intercorrelations Among Language Maintenance Activities and Mean Number of Hours per Week Devoted to Each Activity

Language activity	1	2	3	4	5	6	7	8	9
English									
1. Video	—								
2. Radio	.35	—							
3. Reading	.27	.24	—						
4. Writing	.21	.16	.53	—					
Spanish									
5. Video	−.18	−.06	−.10	−.02	—				
6. Radio	−.02	.07	−.07	−.03	.37	—			
7. Reading	−.01	−.03	.09	.05	.31	.19	—		
8. Writing	.01	.02	.08	.17	.31	.18	.52	—	
9. Speaking	−.35	−.22	−.33	−.15	.38	.35	.26	.16	—
Mean hours	12	9	11	6	3	3	4	2	52
SD	8.7	11.5	9.9	8.12	5.5	5.6	5.1	2.9	24.70

Note. Speaking was measured as percentage of Spanish spoken at home, work, and school.

Table 3
Reliability, Median Correlation, and Intercorrelations of Test Scores

Test	1	2	3	4	5	6	7	8	9	10	11	12
English												
1. Lexical decision	—											
2. Category generation	.46	—										
3. Vocabulary	.68	.58	—									
4. Oral comprehension	.36	.35	.46	—								
Spanish												
5. Lexical decision	.53	.16	.27	.15	—							
6. Category generation	.33	.66	.31	.19	.38	—						
7. Vocabulary	.54	.34	.50	.32	.59	.43	—					
8. Oral comprehension	.37	.35	.41	.32	.28	.35	.44	—				
9. Text comprehension	.42	.35	.46	.27	.41	.46	.63	.41	—			
10. Grammar	.41	.35	.41	.28	.47	.45	.55	.41	.70	—		
11. Mexican anglicisms	.53	.39	.46	.29	.51	.52	.52	.41	.61	.58	—	
12. Cuban anglicisms	.34	.18	.23	.17	.53	.33	.44	.22	.37	.48	—	—
Median correlation	.42	.35	.46	.29	.41	.38	.50	.37	.42	.45	.52	.34
Spearman reliability	.89	.89	.89	.98	.76	.83	.77	.98	.87	.92	.82	.86

each subtest. Because the tests were derived using pilot data, these values are reasonably high, ranging from .76 for the Spanish lexical decision test to .98 for the oral comprehension test in both Spanish and English. Also presented are the intercorrelations among the subtests and the median intercorrelation of each test with all other tests.

Table 3 shows that scores on subtests pertaining to vocabulary (category generation, lexical decision, and vocabulary recognition) not only intercorrelate within each language but also correlate with scores on the same test in the other language. Carrell (1991) reported similar cross-language correlations for indicants of reading proficiency; these findings reflect individual differences in performance that are specific to various aspects of language but generalize from L1 to L2.

Separate principal factor analyses within each language further clarified the structure of intercorrelations among the four dominance tasks. For the Spanish tests, one factor with an eigenvalue of 1.5 was extracted. The factor yielded weights of .83 for vocabulary recognition, .67 for lexical decision, .53 for category generation, and .48 for oral comprehension. A similar factor structure was obtained for the intercorrelations among the English dominance tests. A single factor with a final eigenvalue of 1.9 had weights of .90 for vocabulary recognition, .73 for lexical decision, .63 for category generation, and .49 for oral comprehension. Comparable interrelations among scores on the Spanish and English test versions support the prior conclusion that individual differences in language performance generalize from L1 to L2.

Retention and Acquisition Functions

Retention of Spanish

We first present performance in the form of conventional retention functions. Figure 1 shows the mean scores on all subtests obtained by participants belonging to each of six retention intervals. The intervals, shown on a natural logarithmic scale, specify the number of years of residence in the United States. These retention functions give an overall impression of how well our participants preserved the aspects of Spanish language measured by our test. The functions do not reveal the extent to which performance is affected by any of the other independent variables (e.g., continued use of Spanish); rather, the systematic effects of the retention or maintenance interval are confounded with the increasing age of the individual and with the net effect of all other uncontrolled variables.

Figure 1 shows that, overall, performance on the Spanish subtests remained quite stable during years of residence in the United States. This finding holds true even for the anglicism test; prolonged residence in the United States did not significantly diminish ability to identify misuses of Spanish stemming from the influence of English words or expressions. Performance of the oldest group of participants was lower, particularly on tests of vocabulary recognition and category generation. However, performance on comparable English subtests was also lower for these participants. The decline is therefore a function either of aging or of other uncontrolled cohort effects; it does not reflect the effects of relative language isolation. The stability of these functions indicates that, for individuals who continue to use Spanish to the extent typical of our participants, performance on the aspects of Spanish language we tested shows little decline. Note that these retention functions obscure improvements in performance during adolescence and early adulthood, because individuals ranging in age from 10 to 26 years contributed to each data point. To reveal the effects of independent variables on performance, we later examine retention functions projected on the basis of multiple regression analyses.

Another way to evaluate bilinguals' Spanish language deficits is to compare their performance with that of monolingual individuals of the same age and educational level. This comparison can reveal language arrest or stagnation (Pan & Gleason, 1986) by identifying language improvements of monolinguals that are not shared by bilinguals.

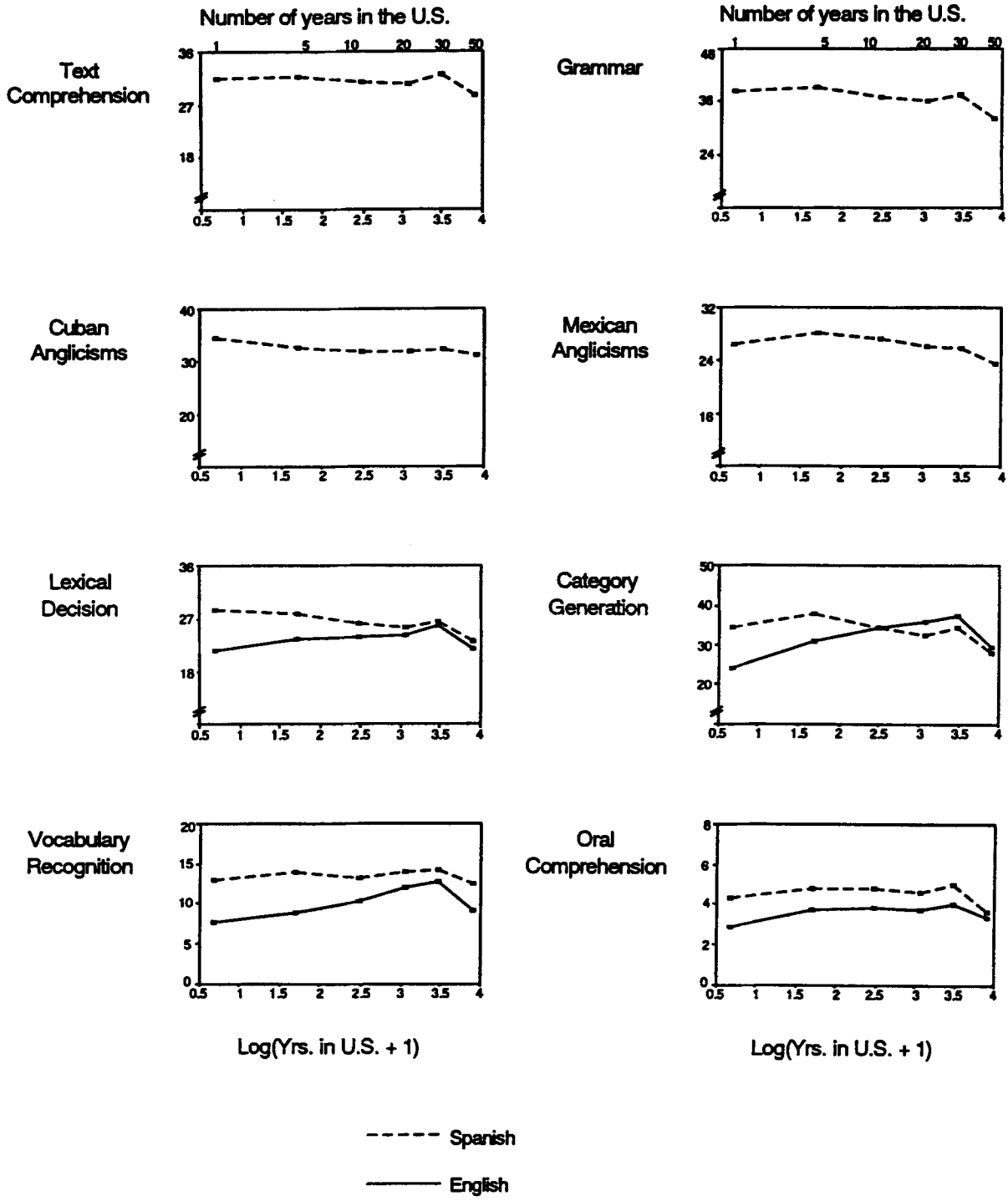

Figure 1. Unadjusted mean scores based on length of residence in the United States.

Figure 2 compares the bilingual subjects with 60 monolingual Hispanic immigrants ranging in age from 10 to 70. The monolinguals were slightly older than the bilinguals (35 years vs. 33 years), and the groups were equivalent in the years of education prior to immigration (11) and in the years of English courses taken in their native country (2). The groups also had comparable grade point averages, 2.5 (on a 4-point scale) for monolinguals and 2.2 for bilinguals. These control subjects were tested within 3 months of their arrival in the United States, so they had experienced no prolonged diminution in the use of Spanish or exposure to a bilingual environment. Figure 2 shows no systematic dif-

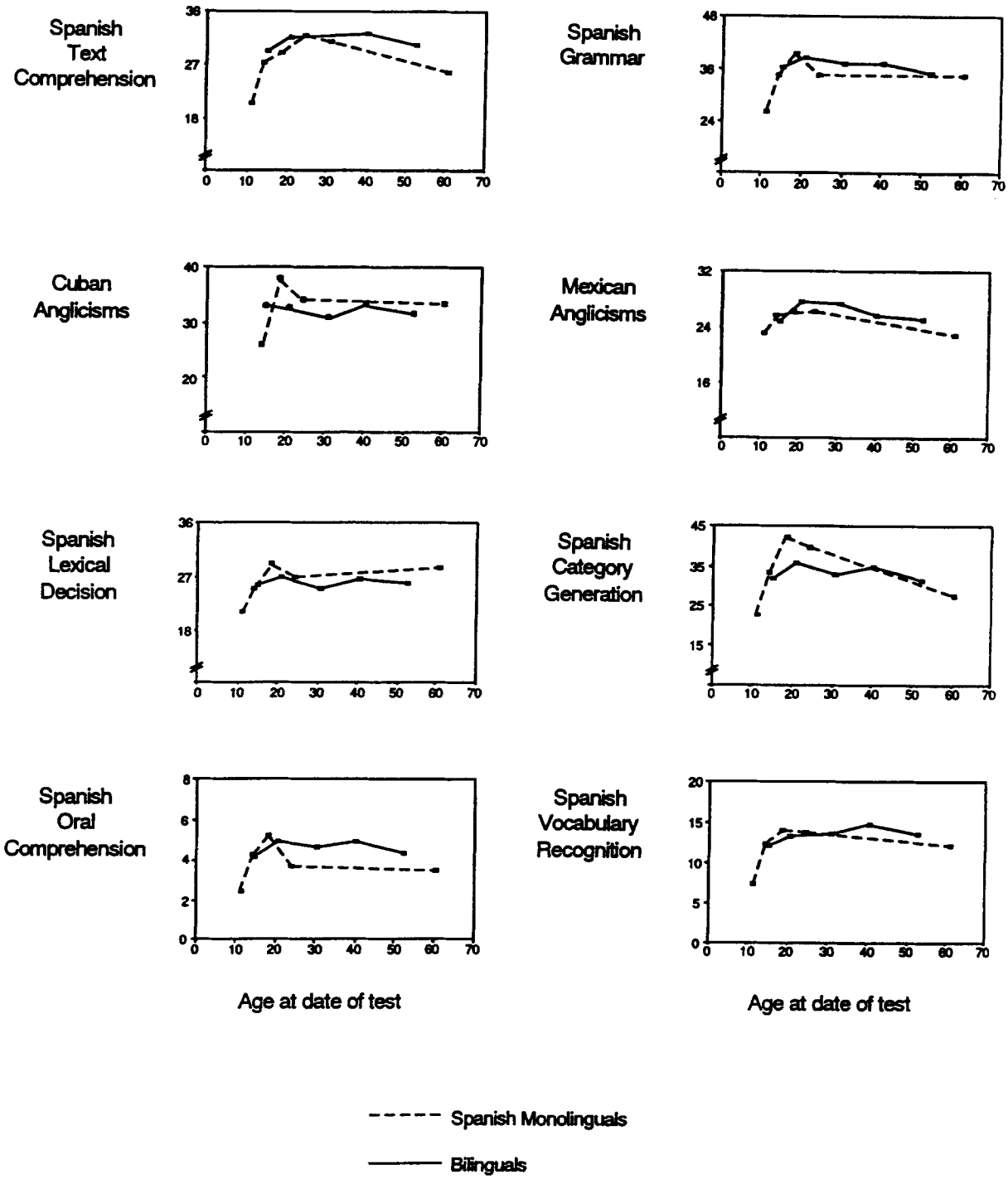

Figure 2. Mean test scores as a function of age for bilingual subjects and monolingual Spanish control subjects.

ference favoring the control subjects. The long-term performance of the bilingual subjects was actually somewhat better on the vocabulary recognition, text comprehension, and oral comprehension tests; and differences between the groups are generally small. The comparison group is not large; but the results give no evidence of progressive deficits, language interference, or language stagnation associated with living in a bilingual environment for up to 50 years.

These findings contrast sharply with those obtained for maintenance of Spanish learned in U.S. high school or

college courses (Bahrick, 1984). In Bahrick's (1984) study, a large amount of that knowledge was lost during the first 3–5 years of the retention interval, after which retention stabilized for 25–30 years. Two conditions account for the differences between the present findings and those of Bahrick's study. The uniformly high degree of original learning of L1 in the present study is the key condition preventing early retention losses. Early losses are observed only when large portions of knowledge are acquired late and do not share the benefit of extended, spaced practice. The retention functions obtained by Bahrick, Bahrick, and Wittlinger (1975) for the recognition of names and faces of high school classmates showed no early decline for the same reason. The high level of maintenance activities typical of the present subjects is, of course, a second variable that prevents early forgetting.

We compared performance of Cuban and Mexican subjects on all subtests. Differences are slight, but nearly all differences are statistically significant because of the large sample size. Cuban subjects performed somewhat better than Mexican subjects on all Spanish and English tests, except for category generation, on which Mexican subjects performed better than Cuban subjects.

Acquisition of English

The unadjusted acquisition functions for the English language subtests (lower panels of Figure 1) are negatively accelerated; the largest improvements occurred during the first postimmigration year, an interval during which few participants were tested. Assuming that most of our subjects would have obtained low scores on the English language tests at the time of immigration, these functions show that English lexical decisions and oral comprehension improved only slightly after the first year in the United States. However, performance on category generation and vocabulary recognition tests continued to improve for at least 30 postimmigration years, declining finally for the oldest group of participants. It is important to remember that these findings are for young immigrants; acquisition functions for other age groups may be somewhat different. As we noted for the Spanish maintenance functions, the unadjusted English acquisition functions confound systematic effects of length of residence in the United States with the effects of aging and with the net effects of all uncontrolled variables. We used projections based on multiple regression analyses to reveal the effects of independent variables on acquisition.

We compared the performance of the bilingual subjects on the English tests with the performance of 61 monolingual English-speaking subjects (see Figure 3). The mean ages of the English controls and the bilinguals were comparable (35 years vs. 33 years), as were the years of education (10 for the monolinguals vs. 11 for the bilinguals) and grades (2.1 for the monolinguals vs. 2.0 for the bilinguals).

Bilinguals equaled or exceeded the performance of monolinguals by age 40, except on the oral comprehension task, and differences between the groups were generally small. Performance of the oldest bilinguals was lower on the category generation and vocabulary recognition tasks, as was the case for the Spanish tasks (see Figure 1). Because the monolingual English subjects showed no decline at that age, the effect is probably a cohort effect rather than a general function of aging. The decline could also reflect a long-delayed consequence of bilingualism, but that interpretation is less likely.

The English monolinguals we tested may have been intellectually and linguistically less talented than the bilingual subjects, and the relatively high performance of the bilinguals may reflect this difference. Immigrants may lack the ready access to higher education typical of native-born citizens, so control of the level of education does not assure control of general intellect. This interpretation remains speculative, because we have no indication of general intelligence. Selection effects are much less likely to affect the comparison with Spanish monolinguals, because that sample was drawn from the same Cuban and Mexican immigrant population as the bilinguals.

Indicants of Language Dominance

The lower panels of Figure 1 show that, except for category generation, unadjusted performance was consistently higher on Spanish subtests than it was on English subtests throughout the postimmigration period. In contrast, category generation in English began to exceed that in Spanish after approximately 11 years of residence in the United States.

Apparently, Spanish remains the dominant language for most participants, but this conclusion is subject to caveats. As we discuss later, we must assume that the English and Spanish subtests are of comparable difficulty. Furthermore, the functions are based on the *mean* performance of the various time groups, and the mean does not reveal the effects of individual differences.

Analysis of Language Dominance

Dominance Scores for Individual Participants

In the following analysis we determine whether each participant's score is higher on the English or on the Spanish version of a subtest or whether the two scores are equal. This analysis disregards the size of the difference. The trichotomized scores offer the same advantages as nonparametric statistics, that is, they involve no assumptions about scaling or the distribution of scores. However, because they are derived from difference scores, changes in one or both of the scores being compared are reflected to an indeterminate degree. Thus, an increase in English dominance as a function of time lived in the United States or any other variable could be due to improvement in English scores, decline in Spanish scores, or some combination of both. Separate regression analyses of the influence of the same independent variable on the English and on the Spanish test scores generally permit a resolution of these ambiguities.

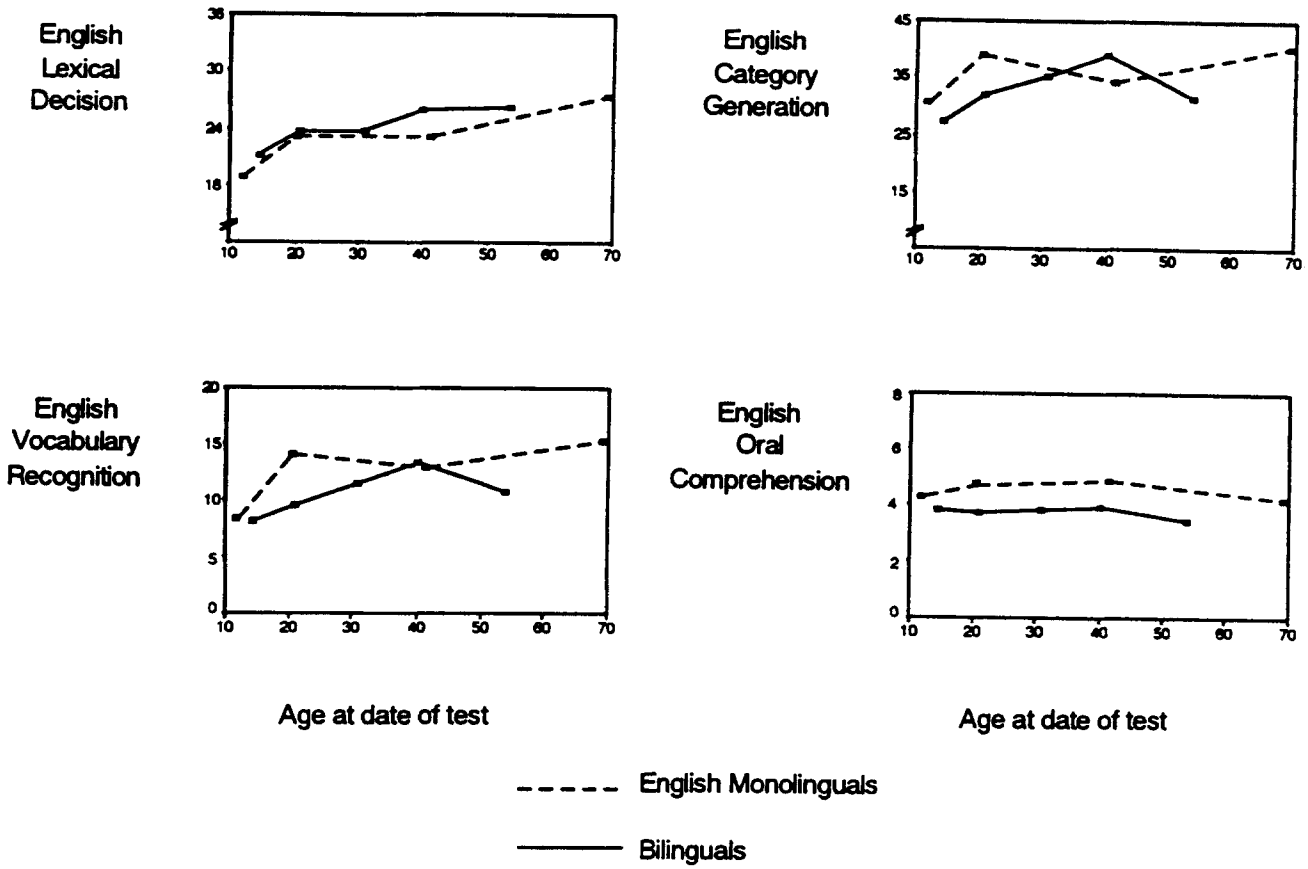

Figure 3. Mean test scores as a function of age for bilingual subjects and monolingual English control subjects.

Interrelations Among Indicants of Dominance

To what extent does dominance in one aspect of language predict language dominance in other aspects? Table 4 gives mean gamma values for all combinations of tasks. The table shows that language dominance is quite task specific.

Dominance in lexical decisions, category generation, and vocabulary recognition are moderately interrelated; however, none of these tests allows even a moderately accurate prediction of dominance on the oral comprehension test. This finding does not support the contention that oral comprehension tasks provide a valid, general measure of language proficiency (Gradman & Spolsky 1975; Oyama, 1978; Spolsky, 1971; Spolsky et al., 1968; Whitson, 1972; Zirkel, 1976), rather, the results support Fishman's (1968) and Cooper's (1969) conclusion that global measures of the degree of bilingualism inadequately describe bilingual performance.

Table 4
Gamma Correlations Among Indicants of Language Dominance

Dominance test	1	2	3	4
1. Lexical decision	—	.39	.48	.07
2. Category generation		—	.41	.10
3. Vocabulary			—	.08
4. Oral comprehension				—

Self-Assessments of Language Dominance

We calculated the percentage of participants whose performance was better on the English version of a task, given that they rated themselves as English dominant separately for each self-rating scale. For the category generation and vocabulary recognition tasks, these values ranged from 81% to 95%, indicating highly valid assessments regardless of whether the self-ratings were based on reading, writing, understanding, or speaking. Dominance self-ratings are somewhat less valid for lexical decisions (proportions ranged from 70%–78%) and least valid for the oral comprehension task. However, even for oral comprehension, 60% of participants who judged their understanding of English as better than their understanding of Spanish scored higher on the English test; 22% scored lower, and 18% had equal scores. These data show that the oral comprehension task has face validity, that is, individuals who rated their understanding of English as better than their understanding

of Spanish obtained a higher English comprehension score nearly three times as often as they obtained a higher Spanish score.

Individual Differences in Language Dominance

Effect of length of residence. The percentage of English-dominant subjects increased with length of residence in the United States for all four tasks, though the increments for oral comprehension were small. Notwithstanding these increments, average English performance eventually exceeded Spanish performance only for category generation, as is evident in the lower panels of Figure 1. Sixty percent of participants eventually become English dominant on that task; versus 35%, 29%, and 25% of participants on the lexical decision, vocabulary recognition, and oral comprehension tasks, respectively.

Effect of age at entry. Age at entry to the United States strongly influenced language dominance. Participants who were 13 or younger when they entered the United States were more likely to become English dominant and to exhibit English dominance after a shorter residence period. This effect was particularly pronounced in regard to lexical decision and vocabulary recognition. On the lexical decision test, a majority of the youngest immigrants (52%) eventually became English dominant, but of those who were older than 13 at immigration, only 14% did. On the vocabulary recognition test, 46% of the youngest group became English dominant, versus 10% of those who were older than 13 at immigration. For the oral comprehension test this effect was less pronounced; the proportion of participants who eventually achieved English dominance on this test is 25% of those 13 years of age or less at immigration and 13% of those who were older.

Effect of schooling. The effect of the level of schooling before immigration is quite similar to the effect of age of immigration. This result was expected, because the variables are inextricably confounded. Participants who attended only elementary school in their native country were much more likely to achieve English dominance than were those who attended secondary or high schools. The oral comprehension test was the exception; the level and type of prior schooling had little impact on ultimate language dominance.

We also assessed the influence of academic achievement on likelihood to achieve English dominance. The effect was indicated in one of two ways. For participants who attended U.S. schools after they immigrated ($n = 475$), self-ratings of competence as students (excellent = A, good = B, or fair/poor = C/D) were used. For the 79 participants who had not attended U.S. schools, this variable reflects the grade level achieved in Cuban or Mexican schools considering the age of immigration. For example, high achievers were individuals who had reached a high grade level in relation to their age. We compared language dominance for participants who reported that they were A students with those who reported that they were B or C–D students. The effects are minor, but the C–D students were somewhat more likely to remain Spanish dominant.

The number of years of English language schooling before immigration had no obvious effect on subsequent language dominance. Participants with 4 or more years of prior education were no more likely to achieve English dominance than were those who had none. The age at immigration is a confounding variable here, because individuals with 4 or more years of prior English instruction also tended to be older at the time of immigration.

Effect of language usage. Table 2 shows the participants' estimates of their postimmigration English and Spanish usage. As we noted earlier, our participants generally reported speaking Spanish slightly more than English, but approximately 3 times as many weekly hours were devoted to English as were devoted to Spanish reading, writing, and video or radio exposure. Our analyses reflect the effects of variations among participants within this general pattern.

We calculated point-biserial correlations between each usage variable and the language dominance dichotomy on each subtest. The few individuals who obtained equal scores on both the Spanish and English versions of a test were omitted from each subtest. Separate correlations were obtained for Spanish and English usage (except for speaking), because these activities are statistically independent (see Table 2). Correlation coefficients ranged from $-.22$ to .44 and were all in the expected direction. Higher usage of English was associated with English dominance, and higher usage of Spanish was associated with Spanish dominance. However, the reported variations in English usage generally had less effect on language dominance than did variations in comparable Spanish use. This finding held true in spite of the fact that the reported means and standard deviations of usage time were much higher for English activities than for Spanish activities. The reported percentage of Spanish speaking yielded the highest correlations with language dominance, particularly for the category generation task ($r = .44$). For the oral comprehension test, the effect of English use was negligible, and the effect of Spanish use was small. All correlations of usage variables with the lexical decision and category generation tasks were statistically significant.

Results of the Regression Analyses

The purpose of the regression analyses was to examine the effects of independent variables (including the usage variables derived from the principal components analysis) on maintenance of Spanish and on acquisition of English (cf. Bahrick, 1984; Bahrick & Hall, 1991). By sorting out the effects of the same set of independent variables on performance in the two languages separately, the regression analysis clarified the nature of previously discussed changes in language dominance. The analyses show the extent to which changes in dominance reflect improvement in English performance, decline in Spanish performance, or some combination of both, for a given set of independent variables.

We subjected the data for each subtest to eight hierarchical multiple regression analyses, one to evaluate the effect

Table 5
Multiple Correlations at Three Steps in Hierarchical Multiple Regression Analyses of Subtests for Four Sets of Variables of Interest

Variable of interest	Subtest											
	COMP	GRAM	ANM	ANC	LDE	LDS	CGE	CGS	VCE	VCS	OCE	OCS
Age and prior education												
Controls	.509[a]	.518[a]	.632	.530	.596	.593	.640	.590[a]	.740	.528	.505	.468
Age and prior education	.556[a]	.567[a]	.658[a]	.597[a]	.619	.636[a]	.653	.615[a]	.760[a]	.597[a]	.508	.483[a]
Interaction with retention interval	.561	.569	.662	.626[a]	.623	.641	.657	.619	.761	.600	.515	.485
Postimmigration education												
Controls	.556[a]	.558[a]	.653	.635[a]	.610	.636	.638	.611[a]	.713[a]	.597	.500	.482
Years of education in U.S.	.561[a]	.568[a]	.655	.635	.623	.641[a]	.656	.618[a]	.741	.600	.510[a]	.484
Interaction with retention interval	.561	.569	.662	.637	.623	.641	.657	.619	.761[a]	.600	.515	.485
Academic achievement												
Controls	.545[a]	.555[a]	.656	.607[a]	.611	.630	.648	.599[a]	.756	.513	.513	.483
Student type (A to C–D)	.560[a]	.565[a]	.662	.626[a]	.623	.640[a]	.657[a]	.615[a]	.761[a]	.599[a]	.512	.483
Interaction with retention interval	.561	.569	.662	.632	.623	.641	.657	.619	.761	.600	.515	.485
Speaking												
Controls	.549[a]	.563[a]	.660	.626[a]	.617	.637	.648	.618[a]	.743	.598	.509	.476
Percent of Spanish spoken	.561[a]	.566	.660	.626	.622[a]	.638	.657[a]	.618	.760[a]	.600	.514	.483[a]
Interaction with retention interval	.561	.569	.662	.637	.623	.641	.657	.619	.761	.600	.515	.485

Note. See text for explanation of the steps in the regression analyses. COMP = text comprehension; GRAM = grammar; ANM = Mexican anglicisms; ANC = Cuban anglicisms; LDE = English lexical decision; LDS = Spanish lexical decision; CGE = English category generation; CGS = Spanish category generation; VCE = English vocabulary; VCS = Spanish vocabulary; OCE = English oral comprehension; OCS = Spanish oral comprehension.
[a] Denotes a significant increase in the multiple correlation.

of each of the following sets of variables: (a) age at immigration and preimmigration education, (b) postimmigration education, (c) general academic achievement, (d) Spanish speaking postimmigration, (e) listening to radio postimmigration, (f) watching television and movies postimmigration, (g) reading in the language corresponding to each subtest postimmigration, (h) total amount of reading in both languages postimmigration.

We entered predictor variables into each regression analysis according to the following hierarchy of steps:

1. The control variables were entered first, including: (a) all predictor variables except those whose independent effects were being evaluated (the variables of interest), (b) the retention interval and the retention interval squared, (c) interactions of control predictors with the retention interval and retention interval squared.
2. The set of variables of interest was entered second.
3. The interactions of the variables of interest with the retention interval and the retention interval squared were entered last.

This approach enabled us to determine if each of the eight sets of variables of interest (or their interactions with the retention interval, or both) accounted for a significant proportion of variance over and above the variance accounted for by the control variables. Step 2 determined whether the variables of interest had a significant effect on performance on the subtest being analyzed, and Step 3 determined whether this effect changed significantly during the 50-year retention interval.

Table 5 reports the multiple R for each step of the regression analyses of (a) age at immigration and preimmigration education, (b) postimmigration education, (c) general academic achievement, and (d) speaking Spanish after immigration for each subtest. Total variance accounted for ranged from 24% (for Spanish oral comprehension) to 58% (for English vocabulary recognition). Statistically significant increases in R[1] are identified with a superscript "a". When the variables of interest or their interactions with the retention interval significantly improved prediction, we created plots of projected scores to illustrate the form of the effects.

For example, for Spanish text comprehension, age of immigration and prior education increase the total variance accounted for (R changes from .509 to .556), and this change is statistically significant. Adding the interactions of these variables with time does not significantly increase accounted for variance (R = .561), indicating that the independent contribution of age and prior education does not change significantly as a function of the retention interval. The regression coefficients and intercept constant from the final step of this regression analysis, which we used to project text comprehension scores, are presented in Table 6.

[1] The final value of R varies slightly from one analysis to the next for the Cuban anglicism subtest, because some predictors were excluded from the analysis when they failed to meet minimum tolerance standards. The particular variables excluded varied depending on the order in which they were entered.

Table 6
Final Regression Coefficients for Prediction of Text Comprehension From Age and Education

Regression variables	Main effect	Interactions	
		Time	Time squared
Step 1: Control variables			
Mother's occupation	3.11	−2.48	0.43
Father's occupation	−1.00	1.42	−0.35
Number of Spanish courses, U.S.	−9.20	6.76	−1.20
Months since last Spanish course	−0.15	−0.48	0.01
Number of English courses, U.S. college	1.96	−1.43	0.27
Months since last English course, U.S. college	−6.93	6.03	−1.18
Number of English courses, U.S. jr. & sr. high	3.67	−2.61	0.45
Years of education after immigration	0.44	−0.14	0.02
Academic achievement	0.27	−1.65	0.40
Amount of English video	−4.13	2.51	−0.37
Amount of English radio	−2.86	2.69	−0.61
Amount of English reading	8.82	−7.29	1.43
Amount of English writing	−3.08	1.87	−0.29
Amount of Spanish speaking	1.11	−1.55	0.29
Amount of Spanish video	1.54	−0.58	0.01
Amount of Spanish radio	6.76	−5.73	1.17
Amount of Spanish reading	−4.18	4.86	−1.05
Amount of Spanish writing	1.98	−1.32	0.22
Code for nationality and test site	0.47		
LN (number of years in U.S. + 1)	19.41		
LN (# years in U.S. + 1) squared	−3.58		
Step 2: Variables of interest			
Age of immigration	0.74		
Education level before immigration	3.25		
Years of English before immigration	0.08		
Step 3: Interactions of time with variables of interest			
Age of immigration		−0.60	0.12
Education level before immigration		−1.99	0.33
Years of English before immigration		0.24	−0.04
Constant	4.30		

Note. LN = natural log.

This equation was evaluated nine times, each yielding a point on the first graph in Figure 4. For each solution, mean values were entered for all control predictors except the retention interval. The criterion for selecting levels of the retention interval and the variables of interest was whether they were realistic. For example, there were a number of participants at each of five retention intervals who were 12 years of age at immigration, had a fifth grade education, and no English courses. Thus, the projected functions do not extrapolate beyond the parameters of our subject pool, and they assume only those combinations of values that actually existed. The effects of specific variables of interest are discussed in the following sections.

Effects of Age and Prior Education

Because age at immigration, the number of years of Spanish education, and the number of years of English education before immigration are naturally confounded, these variables were entered simultaneously in Step 2 of these regression analyses. The analyses show that age and prior education significantly affected performance on all subtests except English oral comprehension. The direction of the effect is one of better performance for *older* immigrants on both the Spanish (Figure 4) and the English (Figure 5) subtests. The absence of significant interactions with time indicates that the effects remained relatively stable throughout the postimmigration period.

The fact that participants who were older and had more Spanish schooling at the time of immigration performed better on tests of their native language than did those who were younger and had less Spanish schooling when they immigrated was not unexpected. However, the stronger English performance of those who were older at immigration, especially its continuation for many years, was less predictable. This finding shows that the critical-age hypothesis is not supported for the aspects of L2 acquisition we examined, that is, for domains that depend directly or indirectly on the acquisition of L2 vocabulary. The concept of a critical age is therefore not valid for L2 acquisition in general. It applies only to acquisition of those aspects of L2 that benefit least from L1 transfer. Phonology is the only well-documented example. Johnson and Newport (1989) also found support for application of the concept to the acquisition of syntax, but that conclusion will remain con-

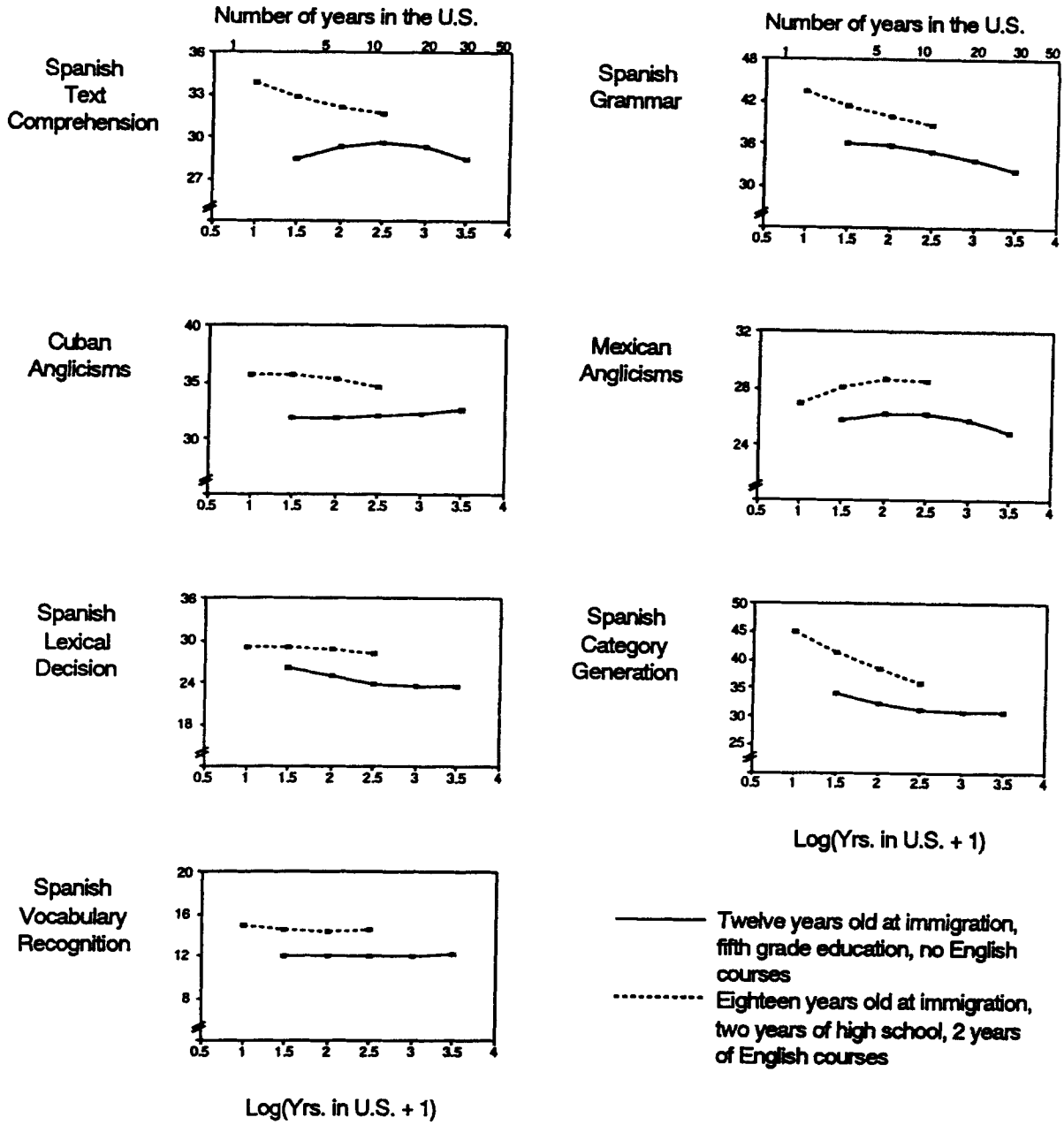

Figure 4. Projected scores on Spanish subtests based on age at immigration and amount of education prior to immigration.

troversial until postimmigration language usage and education variables can be well controlled.

We stress that our *unadjusted* acquisition data favor younger participants on all tests, and the analysis of language dominance also confirms that older immigrants were more likely to remain Spanish dominant. The superior English language performance of older participants becomes evident only when multiple regression analysis is used to control postimmigration English schooling and related language usage variables. The number of years of postimmigration full-time English schooling is, of course, negatively correlated with age at immigration, especially for participants of school age. Language usage variables (particularly writing) are also closely related to the number of years of English schooling. Johnson and Newport (1989) concluded, on the basis of a nonsignificant correlation between amount of initial exposure and performance, that L2 learning is not particularly sensitive to differences in the amount of exposure when it is spread over a number of years and is fairly high. In contrast, we found that the benefit associated with

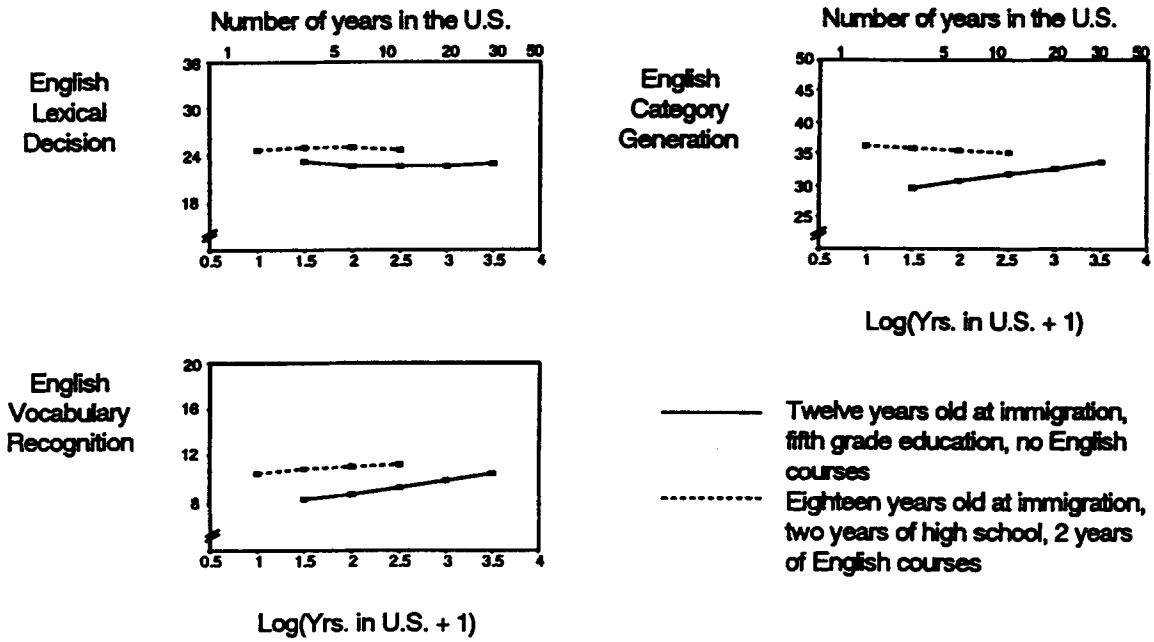

Figure 5. Projected scores on English subtests based on age at immigration and amount of education prior to immigration.

years of formal English education is substantial. The significant effects of postimmigration English schooling on performance are shown in Figure 6; the largest effects were on the English vocabulary recognition and category generation tasks.

To sort out the effect of age per se on L2 acquisition and to answer questions about a critical age for L2 acquisition, one must control the effects of education and usage variables that are confounded with age in the present study. When this is done through multiple regression analysis, the advantage shifts to older learners, and their advantage is maintained for many years. In a study of Japanese–English bilinguals, which used multiple-regression techniques, Walberg, Hase, and Rasher (1978) also concluded that the hypothesis of special competence of young children for language acquisition is unsupported and that the advantage, if any, lies with older children. We believe that prior conclusions supporting a critical age, beyond the acquisition of phonology, need to be reexamined with adequate control of confounded variables.

Effect of Academic Achievement

As expected, high academic achievement facilitated performance on all Spanish and English subtests except for oral comprehension and Mexican anglicisms. No interactions with the retention interval were found, indicating that the effects were of long duration. Thus, good students tended to perform well in both languages. However, as indicated by the dominance analysis, academic achievement had only minor effects on dominance. Poor students were somewhat more likely to remain Spanish dominant, suggesting that the adverse effects are larger on English performance than on Spanish performance.

Effect of Spanish and English Usage Activities

Participants who reported speaking predominantly Spanish showed adverse effects on category generation, lexical decisions, and vocabulary recognition in English. The effect on vocabulary was large (see Figure 7) and appeared to increase during the 50-year period, although the interaction with time was not significant. Speaking predominantly Spanish was also associated with lower performance on the Spanish subtests of reading comprehension and oral comprehension, and no Spanish test showed significant positive effects. The previously reported Spanish dominance of participants who spoke predominantly Spanish therefore is best explained on the basis of adverse effects on English, not beneficial effects for Spanish. The negative effects on Spanish subtest performance are counterintuitive, and they are probably due to confounds with other variables. If participants who continue to favor speaking Spanish are, on average, linguistically less talented, such differences could account for the adverse effect on Spanish tests. We also believe that this interpretation applies primarily to young immigrants. With a wider age range, the degree of continued use of the native language is likely to be more closely related to the age at immigration and less to linguistic talent.

The effects of other usage variables are relatively small and in the expected direction. Exposure to video and radio enhanced performance in each language. Reading enhanced

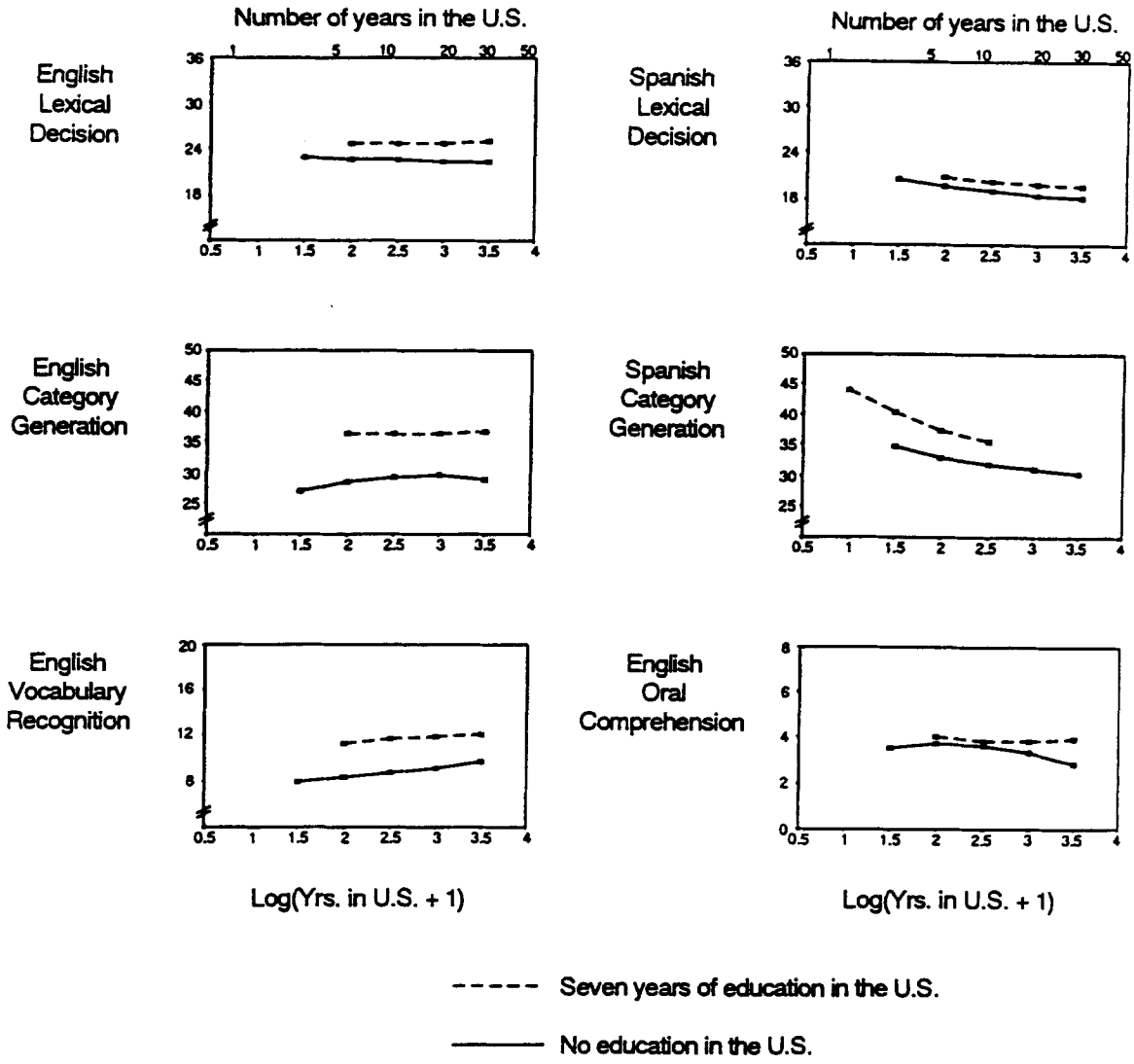

Figure 6. Projected scores as a function of years of education in the United States.

test performance in each language, and predominance of reading in one or the other language affected dominance in the expected direction.

Summary and Conclusions

We must limit interpretation to the parameters that characterize our tests and our participants. The participants were adolescents or young adults at the time of their immigration. The majority continued to speak about as much Spanish as English during postimmigration years, although their other language-related activities (reading, writing, exposure to radio and video) were predominantly in English. Given these circumstances, performance on Spanish text comprehension, vocabulary, grammar, and oral comprehension was unimpaired for most of the 50 postimmigration years. We found that older immigrants performed better than younger ones, that they maintained this advantage for many postimmigration years, and that there was little indication of forgetting the native language for any age group. Test performance was lower during the fifth postimmigration decade, but the decline reflects changes associated with aging or other cohort effects, not conditions of relative language isolation.

Our tests focused on maintenance of knowledge, not on speed of processing or retrieval of information. Reports of decline in native language performance during the first few years in a bilingual environment have been based on speeded tasks or young children (e.g., Mägiste, 1986). We believe that indicants of linguistic knowledge that are based on retrieval or processing speed are more sensitive to the most recent language environment than are indicants of knowledge that place no premium on speed of response.

Notwithstanding reports of language interference and of hybrid language (Spanglish) usage by Hispanic-Americans,

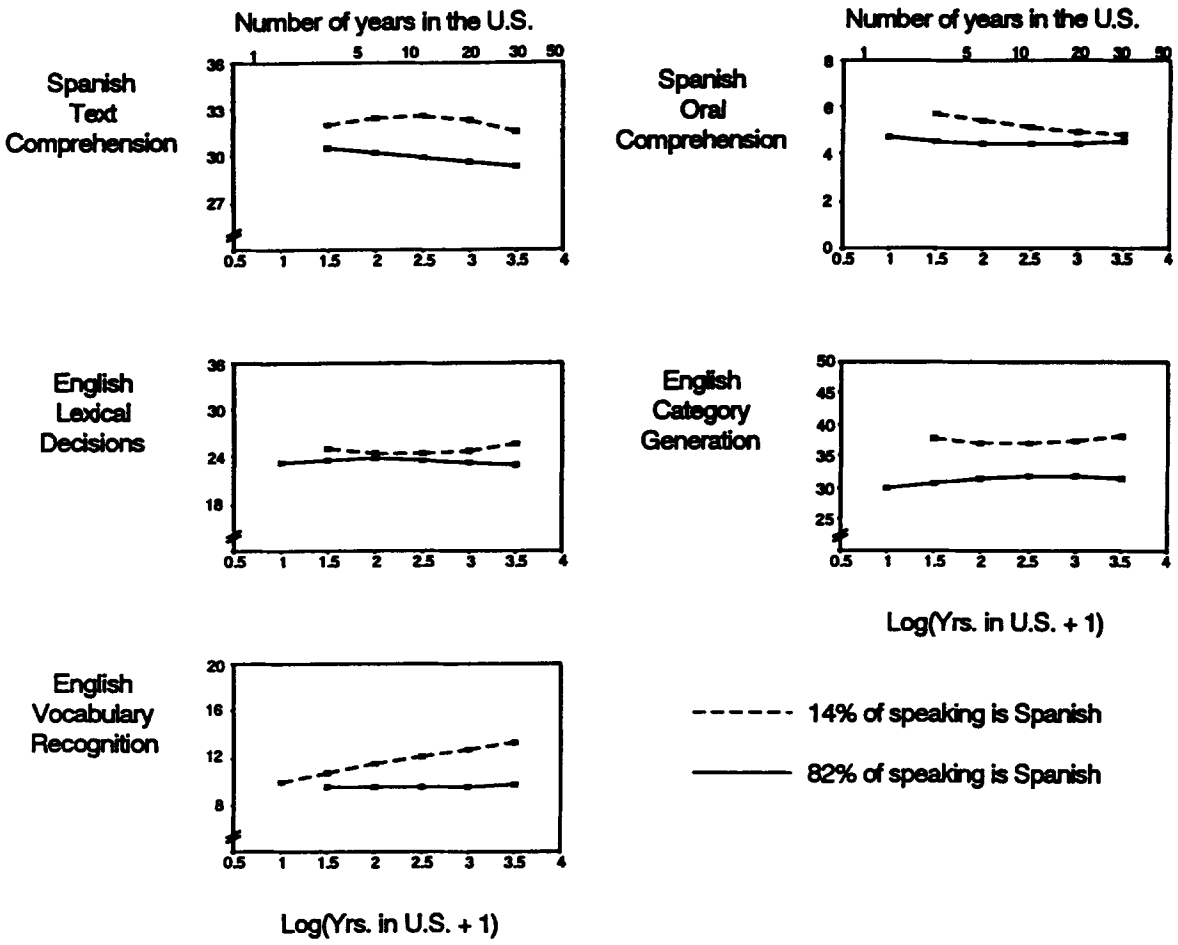

Figure 7. Projected scores as a function of proportion of time spent speaking Spanish.

our subjects sustained no deficit in their ability to identify anglicisms, or adulterated forms of Spanish words or usage. We conclude that using hybrid language does not necessarily mean failure to discriminate hybrid language. We also conclude that 50 years of bilingualism and diminished use of Spanish has not led to "language stagnation." The performance of our participants did not differ systematically from that of control subjects of comparable age and education who were tested within 3 months of arrival in the United States and who had no prolonged experience of bilingualism. We conclude that previous findings of damaging effects of bilingualism in the form of language interference and language stagnation are based on younger immigrants who lack a solid foundation in their first language because of a paucity of formal schooling before immigration, or on Chicanos born in the United States who learned Spanish primarily at home and from peers who have poor command of the native language.

Performance on English tests improved rapidly during the first postimmigration year and more slowly thereafter. Performance on oral comprehension and lexical decision tests increased only slightly after the first year, but scores on category generation and vocabulary recognition continued to improve substantially for 30 postimmigration years.

Our participants' performance on English tests was below that of monolingual English control subjects of comparable age and education during early postimmigration years. By age 40, however, the mean scores of our bilingual subjects approached or exceeded those of monolingual controls, except for the oral comprehension test. Thus, there is no evidence of a permanent language handicap that is due to bilingualism in the nonphonological aspects of language we tested.

In agreement with prior findings, English test performance of older immigrants was generally below that of younger immigrants. However, the reverse result is obtained when postimmigration language usage and years of postimmigration English education are controlled through multiple regression analyses. Our results challenge the concept of a critical or sensitive age for second-language acquisition in regard to vocabulary and related nonphonological aspects of language. Contrary earlier findings should be reexamined with adequate control of education and language usage variables. We argue that the rate of acquisition of L2 vo-

cabulary is facilitated by the available L1 lexicon, because most language concepts are common to L1 and L2. Our data support the conclusion that older adolescents with larger L1 lexicons acquire L2 vocabulary at a faster rate and maintain their advantage for many years. Acquisition of L2 phonology does not benefit from comparable positive transfer so that a critical developmental stage can have more influence on L2 acquisition.

Our assessments of language dominance are based on four subtests that were administered in both languages. The difficulty level of the Spanish and English versions was well controlled by several methods. Results show that dominance is task specific. Dominance in oral comprehension is largely independent of dominance measured by tests of lexical decision, vocabulary recognition, and category generation, and the interrelations among the other three tests are only moderately high. Changes in language dominance can be due to performance changes in one or both languages. A separate evaluation of the effects of independent variables on performance in each language is therefore needed in order to interpret dominance changes. A majority of our participants maintained Spanish language dominance on tests of vocabulary recognition, lexical decision, and oral comprehension. In contrast, a majority became English dominant in generating category exemplars by the twelfth postimmigration year. Performance on the category generation test appears to have been more influenced by recent language usage than has performance on the other tests. This is the only task that is speeded and that requires independent retrieval rather than recognition or comprehension. We conclude that the more recent language environment facilitates rapid access to the lexicon by a cumulative generalized priming effect. As a result, speeded tasks of retrieval of vocabulary, for example, category generation, will show dominance of L2, whereas self-paced tasks that reflect the size of the available lexicon may continue to show dominance of L1. Findings showing more rapid shifts of language dominance (e.g., Mägiste, 1986) have generally been based on tasks that emphasize speed over power, or they are based on children with a less solid foundation in their native language (Merino, 1983), or both.

A large minority of participants eventually achieved English language dominance in lexical decisions and oral comprehension; only a few participants become English dominant on vocabulary recognition. The likelihood of English dominance is greatest for individuals who are younger than 13 years at the time of immigration; have relatively little Spanish schooling; are good students; and speak, read, write, and listen to more English than Spanish.

References

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

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

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General, 113,* 1–29.

Bahrick, H. P. (1989). The laboratory and the ecology: Supplementary sources of data for memory research. In L. Poon, D. Rubin, & B. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 73–83). Hillsdale, NJ: Erlbaum.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104,* 54–75.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General, 120,* 20–33.

Bahrick, H. P., & Karis, D. (1982). Long-term ecological memory. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 427–465). San Diego: Academic Press.

Berelson, B., & Steiner, G. A. (1963). *Human behavior: An inventory of scientific findings.* New York: Harcourt, Brace & World.

Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics, 12,* 159–179.

Cooper, R. L. (1969). Two contextualized measures of degree of bilingualism. *Modern Language Journal, 53,* 172–178.

Cronnell, B. (1985). Language influence in the English writing of third and sixth grade Mexican-American students. *Journal of Educational Research, 78,* 168–173.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222–251.

Cummins, J. (1981). Age on arrival and immigrant second language-learning in Canada—A reassessment. *Applied Linguistics, 2,* 132–149.

Department of Modern Languages. (1980). *Examen de colocacion. Español* [Spanish Placement Examination]. El Paso: University of Texas Press.

Eaton, H. S. (1961). *Semantic frequency list for English, French, German and Spanish: An English–French–German–Spanish word-frequency dictionary.* New York: Dover.

Ervin, S. M. (1961). Semantic shift in bilingualism. *American Journal of Psychology, 74,* 233–241.

Fathman, A. (1975). The relationship between age and second language productive ability. *Language Learning, 25,* 245–266.

Fishman, J. A. (1968). Sociolinguistic perspective on the study of bilingualism. *Linguistics, 29,* 21–49.

Fritz, R. A., & Rankin, N. R. (1934). The English handicap of junior high school pupils from foreign speaking homes, and remedial suggestions. *Journal of Educational Research, 27,* 412–421.

Garcia-Pelayo, R., & Durand, M. (1983). *Gran diccionario moderno (espanol/ingles) Larousse* [Large modern dictionary Spanish/English Larousse]. Mexico, D. F.: Ediciones Larousse.

Gardner, E. F., Rudman, H. C., Karlsen, B., & Merwin, J. C. (1982). *Stanford Achievement Test: Advanced Forms E/F.* San Diego: Harcourt Brace Jovanovich.

Gradman, H. L., & Spolsky, B. (1975). Reduced redundancy testing: A progress report. In R. L. Jones & B. Spolsky (Eds.), *Testing language proficiency* (pp. 59–70). Arlington, VA: Center for Applied Linguistics.

Hamers, J. F., & Blanc, M. H. A. (1989). *Bilinguality and bilingualism.* Cambridge, England: Cambridge University Press.

Hornblower, M. (1985). Instantaneas personales [Personal instances]. *Selecciones del Reader's Digest, 539,* 94.

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

Johansson, S. (1973). An evaluation of the noise test: A method for testing overall second language proficiency by perception under masking noise. *IRAL, 11,* 107–133.

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.

Lambert, W. E. (1955). Measurement of the linguistic dominance of bilinguals. *Journal of Abnormal and Social Psychology, 50,* 197–200.

Lanzano, S., & Sagrista, A. (Eds.). (1984). *Preparación para el examen de equivalencia de la escuela superior (en Español)* [Preparation for the Spanish high school equivalency diploma test]. New York: Arco.

Lemmon, C. R., & Goggin, J. P. (1989). The measurement of bilingualism and its relationship to cognitive ability. *Applied Psycholinguistics, 10,* 133–155.

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

Macnamara, J. (1966). *Bilingualism and primary education.* Edinburgh, Scotland: Edinburgh University Press.

Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior, 6,* 729–736.

MacWhinney, B. (1992). Transfer and competition in second language learning. In R. J. Harris (Ed.), *Cognitive processes in bilinguals* (pp. 371–390). Amsterdam: Elsevier.

Mägiste, E. (1986). Selected issues in second and third language learning. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives* (pp. 97–122). Hillsdale, NJ: Erlbaum.

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning.* New York: Van Rees Press.

McLaughlin, B. (1977). Second-language learning in children. *Psychological Bulletin, 84,* 438–459.

Merino, B. J. (1983). Language loss in bilingual Chicano children. *Journal of Applied Developmental Psychology, 4,* 277–194.

Oller, J. W., Jr. (1973). Cloze tests of second language proficiency and what they measure. *Language Learning, 23,* 105–118.

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

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

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

Pan, B., & Gleason, J. (1986). The study of language loss: Models and hypotheses for an emerging discipline. *Applied Psycholinguistics, 7,* 193–206.

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

Pearson, B. Z., Fernandez, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43,* 93–120.

Pearson, H. (1985). Instantaneas personales [Personal instances]. *Selecciones del Reader's Digest, 539,* 93.

Queralt, M. (1984). Understanding Cuban immigrants: A cultural perspective. *Social Work, 29,* 115–121.

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

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

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

Spolsky, B. (1971). Reduced redundancy as a language testing tool. In G. E. Perren & J. L. M. Trim (Eds.), *Applications of linguistics: Selected papers of the Second International Congress of Applied Linguistics, Cambridge 1969* (pp. 383–390). Cambridge, England: Cambridge University Press.

Spolsky, B., Sigurd, B., Sako, M., Walker, E., & Arterburn, C. (1968). Preliminary studies in the development of techniques for testing overall second language proficiency. *Language Learning, 18,* 79–101.

Thorndike, E. L., & Lorge, I. (1952). *The teacher's word book of 30,000 words.* New York: Teachers College.

Van der Ley, S. (1985). Reloj delator [Tattle tale clock]. *Selecciones del Reader's Digest, 540,* 21.

Walberg, H., Hase, K., & Rasher, S. (1978). English acquisition as a diminishing function of experience rather than age. *TESOL Quarterly, 12,* 427–437.

Whitson, V. (1972). The correlation of auditory comprehension with general language proficiency. *Audio-Visual Language Journal: Organ of the British Association for Language Teaching, 10,* 89–91.

Zirkel, P. A. (1976). A method for determining and depicting language dominance. In J. Alatis & K. Twaddell (Eds.), *English as a second language in bilingual education* (pp. 133–142). Washington, DC: Teachers of English to Speakers of Other Languages.

Ziros, G. I. (1976). Language interference and teaching the Chicano to read. *Journal of Reading, 19,* 284–288.

Received August 17, 1993
Revision received April 8, 1994
Accepted April 11, 1994 ■