TESOL QUARTERLY, Vol. 21, No. 4, December 1987

Age and Rate of Acquisition of Second Language for Academic Purposes

VIRGINIA P. COLLIER
George Mason University

The study reported in this article analyzed the length of time required for 1,548 advantaged limited English proficient (LEP) students to become proficient in English for academic purposes while receiving instruction in English in all subject areas. Variables included were age on arrival, English proficiency level upon arrival, basic literacy and math skills in the native language upon arrival, and number of years of schooling in English. Second language and content-area achievement were measured by students' performance on the Science Research Associates tests in reading, language arts, mathematics, science, and social studies. The results indicated that LEP students who entered the ESL program at ages 8-11 were the fastest achievers, requiring 2-5 years to reach the 50th percentile on national norms in all the subject areas tested. LEP students who entered the program at ages 5-7 were 1-3 years behind the performance level of their LEP peers who entered the program at ages 8-11, when both groups had the same length of residence. Arrivals at ages 12-15 experienced the greatest difficulty and were projected to require as much as 6-8 years to reach grade-level norms in academic achievement when schooled all in the second language. Whereas some groups may reach proficiency in some subjects in as little as 2 years, it is projected that at least 4-8 years may be required for all ages of LEP students to reach national grade-level norms of native speakers in all subject areas of language and academic achievement, as measured on standardized tests.

Acquiring a second language is never easy. Some people would like to think it is fairly simple for young children, but second language acquisition (SLA) researchers have documented a very complex process that occurs over a long period of time (McLaughlin, 1984). First language acquisition begins at birth and

continues through at least age 12, with continuing acquisition of new vocabulary and subtleties of the language throughout our adult lives (de Villiers & de Villiers, 1979).

Second language is acquired to varying degrees of proficiency depending on the context in which the acquirer needs to use it. Immigrants of school age who must acquire a second language in the context of schooling need to develop full proficiency in all language domains (including the structures and semantics of phonetics, phonology, inflectional morphology, syntax, vocabulary, discourse, pragmatics, and paralinguistics) and all language skills (listening, speaking, reading, writing, and metalinguistic knowledge of the language) for use in all the content areas (language arts, mathematics, science, and social studies). Language used in school is sometimes unique to that context, and it becomes increasingly abstract as students move from one grade to the next. Language is the focus of every content-area task, with all meaning and all demonstration of knowledge expressed through oral and written forms of language.

Cummins (1979, 1980, 1981a, 1981b) proposed one of the first theoretical models for SLA that distinguished between two basic types of language proficiency. In early formulations of his theory, Cummins labeled these basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP). The use of these acronyms has been questioned by some researchers (see Cummins & Swain, 1983; Edelsky et al., 1983; Rivera, 1984) as possibly leading to misinterpretation of the complex concepts that they actually represent, but the terms have become symbolic and meaningful for many people in our field as a way of distinguishing between face-to-face conversational proficiency (BICS) and context-reduced, cognitively demanding aspects of language proficiency (CALP).

In context-embedded, face-to-face communication, meaning can be negotiated and is enhanced with a wide range of paralinguistic and situational cues. Context-reduced oral and written language, on the other hand, relies primarily on linguistic cues to meaning. Cummins (1981b) elaborates his conception of these terms by creating four quadrants which best illustrate the range of possibilities in the BICS and CALP distinction. The quadrants are divided by a horizontal continuum from context-embedded to context-reduced and by a vertical continuum from cognitively undemanding to cognitively demanding.

Language proficiency required for school tasks can incorporate the whole range of skills in all four quadrants, but it is especially in school that students need to develop context-reduced and

cognitively demanding aspects of language in order to function successfully in the classroom. In his continuing refinement of the BICS/CALP distinction, Cummins (1984) defines CALP as aspects of language that involve cognitive processes at the higher levels of Bloom's taxonomy of educational objectives for the cognitive domain: analysis, synthesis, and evaluation (see Bloom & Krathwohl, 1977).

Analyzing rate of attainment of CALP in the second language, Cummins (1981a) found that whereas it generally takes students 2 years to master BICS in the L2, young children with little or no formal schooling in their L1 require approximately 5-7 years to reach the level of native speakers in CALP in the L2, as measured on standardized tests. Cummins (1981b) emphasizes that older children's common underlying proficiency in their first and second language assists with the process of SLA. Thus, for older students, many academic skills and concepts acquired in the L1 transfer to the L2, and the process of SLA occurs at a faster rate than for younger children.

Summaries of the literature on age and rate of attainment of the L2 (Krashen, Long, & Scarcella, 1979; Krashen, Scarcella, & Long, 1982) confirm that older children and adults initially acquire many aspects of the L2 faster than younger children. However, with acquisition of pronunciation and influence of the socioaffective filter (Dulay & Burt, 1978; Krashen, 1982), adults sometimes experience problems with SLA, so that overall, with time, younger acquirers tend to attain higher levels of proficiency in second languages than those who begin SLA as adults.

This study was designed to follow up Cummins's research and Krashen, Scarcella, and Long's literature synthesis on age variables, rate of attainment, and influence of L1 CALP development on the process of L2 CALP development. In this study, Cummins's theoretical framework was used as a basis for analysis of the type of L2 proficiency needed for academic purposes. The measures available for this study, however, were not tests that assessed all aspects of language proficiency.

The standardized tests required only the language skill of reading to be able to answer the questions. Metalinguistic knowledge of the language was assessed in the language arts test, and ability to classify, generalize, manipulate ideas, problem solve, and apply knowledge in each of the content areas was assessed in the reading, social studies, science, and mathematics tests. Measures used thus assessed some aspects of L2 CALP development as well as content-area achievement of students. Since Cummins's (1981a) data were based on 1,210 limited English proficient (LEP) students in Grades

K-9 in Canada, this study sought to extend the literature by analyzing 1,548 LEP students in Grades K-11 in a U.S. context.

METHOD

Sample and Setting

Cross-sectional data from 1977 to 1986 were gathered on language minority students attending a large U.S. public school system on the East Coast. At the time of the study, language minority students represented 11% of the total student population in this school district. Predominantly an affluent suburban area connected to a large metropolitan hub, the district also included a few pockets of low-income families. Almost all the language minority students in the district were relatively recent immigrants to the United States, with over 75 different languages and over 100 different countries represented.

Approximately 65% of the subset of language minority students who received special ESL instruction qualified for free or reduced-price lunches, indicating that upon entry, a majority of these students came from low-income families, as measured by U.S. standards. However, a large percentage of the immigrant families who settle in the district come from an upper or middle-income background in their country of origin, and they bring strong aspirations of upward mobility to their new home, with many achieving a more middle-class standard of living in the United States within the first 10-15 years of their arrival. Thus, the language minority population could best be categorized as lower to middle income, with strong middle-class aspirations.

In educational background, the large majority of language minority students in the district entered school at grade level, with parents having come from middle-class or upper class backgrounds in their home countries. In just the last 3 years, the district has experienced a small but increasing influx of language minority students with little or no formal schooling in their native language.

Subjects for this study, a total of 1,548 students, included all language minority students who were placed in beginning-level ESL classes upon entry and remained in the school system for several years. One subset of this population was not included in the study, those students who tested below grade level in L1 skills during placement testing upon entry, as well as older students with little or no formal schooling in L1.

This study was restricted to a group of LEP students with these particular characteristics for two reasons. First, it was assumed that

an "advantaged" group of LEP immigrants, those with a middle- to upper class background in their home country and a strong educational background in their L1, would be more likely to reach the L2 proficiency and content-area achievement of native English speakers faster than LEP immigrants who had a lower class background or were below grade level in L1 skills. Second, since the amount of time required to reach L2 proficiency can vary significantly depending on level of English proficiency at which a student begins study all in English and level of formal schooling in the L1, it was decided to control for these two variables.

Assessment for controlled variables. Upon entry, a placement-testing procedure determined students' level of English proficiency and basic L1 literacy and math skills. After a placement staff member conducted an interview with the student and his or her parents or relatives, the student was given a locally developed placement test to measure listening, speaking, reading, and writing skills in English and ability to do math computation. The math tests in basic computation, decimals, and fractions were written using six of the major variations in world notations of math symbols. For measurement of basic literacy in the L1, the student was asked to read a short paragraph and to write a short language sample in the native language. The placement center staff had materials in each language to make a rough judgment that the student had had at least some minimal formal training in the L1. Bilingual staff were available to analyze in more depth Spanish, Vietnamese, and Korean language samples, which were among the largest minority language groups of the district.

A decision to place a student below grade level was based on ability to produce a short L1 writing sample and ability to perform math calculations at grade level for the student's age. A third criterion for grade placement was a student's transcripts, which would indicate interrupted or little formal schooling or very low grade point average. Placement staff had extensive references to conduct transcript analysis. Students were rarely placed more than one grade level below their age-appropriate grade.

Characteristics of sample chosen. Students in the advantaged LEP sample chosen for this study exhibited the following major characteristics upon arrival and entry into schooling all in English: (a) They were from over 100 different countries and spoke over 75 different languages, with no single language predominant, although Spanish, Korean, and Vietnamese speakers represented the largest language groups; (b) the students were of lower to middle-class background, as measured by U.S. economic standards, but they had

strong middle-class aspirations and had come from middle- or upper class backgrounds in their home countries; (c) they had little or no proficiency in English; and (d) they were at or close to grade level in academic skills in their L1. In social class and educational background in their home country, these LEP students would be expected to have an advantage over their LEP peers who came from lower class backgrounds or those who had had interrupted or little formal schooling.

ESL program characteristics. In this school system, once each LEP student was tested and placed in the appropriate grade level, the student was given special assistance from ESL teachers, who provided English language arts instruction appropriately structured and sequenced to build a student's proficiency level in English. Since there were no self-contained ESL classes, students spent only part of their day with specialized ESL teachers and the rest of the day in the mainstream classroom. ESL staff assisted with the development of BICS in English, as well as with CALP development with some instruction in the content areas. A few ESL content-area classes (ESL algebra, ESL biology, etc.) were taught at the secondary level. Students did not receive any formal instruction in their L1 at school.

Students were taught by ESL staff until staff members felt they could function full-time in the mainstream. For exit from the ESL program, students were administered the California Achievement Test (CTB/McGraw-Hill, 1986), and ESL staff rated them on a locally developed scale focusing on students' development in English of oral comprehension, oral expression, reading and writing skills, and study habits. Most students were generally mainstreamed from the ESL program within 2-3 years of entry into the school system. Mainstreaming did not imply that the ESL staff believed that students had achieved CALP in English but that they were sufficiently far along in their growth in CALP skills in English to continue their development in a mainstream class.

Research Questions

1. How many years of schooling all in English are required for LEP students' achievement in reading, language arts, social studies, science, and mathematics to reach national average scores of native English speakers at each grade level?
2. How strongly does age on arrival of LEP students influence the rate of acquisition of cognitive academic second language proficiency and content-area achievement?

Data Collection and Analysis

Cross-sectional data from the years 1977-1986 for all students exited from the ESL program were collected to analyze age on arrival and rate of attainment of some aspects of CALP in English and content-area achievement. Dependent variables were scores for Grade Levels 4, 6, 8, and 11 on the Science Research Associates (SRA) Achievement Series (Science Research Associates, 1978) tests of reading, language arts, social studies, science, and mathematics. Testing was done only in English. Independent variables included age on arrival and number of years of schooling all in English, which for this sample was equivalent to length of residence in the United States. L1 literacy and math skills on arrival and English proficiency level on arrival were controlled variables.

School records, available on machine-readable media, were used to construct the data set. The initial data tape consisted of information on all LEP students who were placed in ESL, including their age on arrival, length of residence, sex, primary language, placement scores, and grade level in which they were initially placed. These records were checked for errors by running computer programs that noted unusual and obviously incorrect entries whose values exceeded the normal upper and lower bounds for that variable (e.g., an age of 25). In addition, the computer programs noted inconsistencies among the data for a particular student (e.g., a student of age 16 in Grade 2). These errors were manually corrected.

Files of SRA test data for Grades 4, 6, 8, and 11 were available from the years 1982-1986. Other files were available from 1977-1981, but since different forms of the SRA test were used in those years, the scores could not be compared with those from the new form. A set of relational data-base computer programs was employed to find test score matches from approximately 160,000 testing records for each of approximately 14,000 students who had registered for ESL classes. After only those students who entered at the beginning level of ESL proficiency and who were at grade level in L1 literacy and math skills had been selected and those students who had not remained in the school system had been eliminated, 1,548 LEP students remained in the data file with longitudinal records (4-6 years) of school performance.

With this data file, data sets were constructed consisting of all the students who had been tested in a given grade for a given number of years after ESL registration, resulting in a total of 17 different groups. Each of the 17 groups varied according to the three variables of length of residence, age on arrival, and grade level

when the SRA test was taken. The variable of number of years of schooling all in English was defined by the following grouping of months: 12-23 months (labeled 1-2 years in the following figures), 24-35 months (2-3 years), 36-47 months (3-4 years), 48-59 months (4-5 years), and 60-71 months (5-6 years). The number of years of schooling all in English included time spent in the ESL program as well as time spent in the mainstream.

Among the 17 groups, there were no significant differences in educational background, in level of English proficiency upon arrival, or in proportional representation of sex and language background. Table 1 presents the general pattern of representation of sex and language background within the sample across the four grade levels tested.

TABLE 1

Percentage of Sample in Each Grade by Sex and Language Background

	Grade			
	4	6	8	11
	Language background			
Spanish	23	21	20	18
Vietnamese	17	20	19	20
Korean	16	21	27	24
Farsi	5	7	6	6
Chinese	5	3	4	7
Khmer	4	4	3	6
Lao	4	3	2	5
Urdu	3	3	4	2
Japanese	3	3	3	1
Arabic	3	2	1	2
Other	17	13	11	9
Total	100	100	100	100
	Sex			
Female	53	49	48	48
Male	47	51	52	52
Total	100	100	100	100

For each of these groups, the means and standard deviations of the scaled standard scores in the achievement test areas of reading, language arts, social studies, science, and mathematics were computed. In addition, confidence intervals for the means were calculated. The means are displayed in graphs for each subtest and for each grade tested (school system mean scores are also given).

The means were graphically represented by converting the mean standard scores into normal curve equivalents (NCEs), for

appropriate interval-level scaling. NCEs are the appropriate equivalent of percentiles to use when displaying the results in graph form because a NCE is a conversion of the percentile into equal-interval data. Unlike percentiles, NCEs are preferred for statistical analysis because arithmetical operations can be performed only on equal-interval scales (Tallmadge, 1976).

The number of cases for each group ranged from 21 (for 3 out of 17 groups) to 151, or an average of 74 per group, providing a large enough number for testing for statistical significance in the differences between groups. The school system comparison group (which included both native speakers and ESL graduates) consisted of an average of 9,258 students each year for each grade level.

As a general guideline in this study, by computing confidence intervals, it was found that a difference between group means of 6-7 NCEs was sufficient for significance at the .05 level even in the smallest of groups. In some cases, group differences of 3-4 NCEs were significant at the .01 level. Because of the large number of groups to be compared (17 groups of LEP students, plus 4 groups of native speakers, times 5 subject areas, or 105 groups total), significance tests for each possible comparison were not performed due to the high probability of making Type I errors. Only important comparisons were made and are reported here.

RESULTS

Several fairly consistent patterns in scoring among the groups of ESL graduates emerged as analysis of the data was conducted. These patterns are reported by focusing on differences among groups by (a) length of residence, (b) age on arrival, (c) grade-level achievement (by grade when tested), and (d) subject-area achievement.

Length of Residence

For LEP students in this study, the number of years of their all-English schooling was equal to their length of residence in the United States. Results are reported using the term *length of residence* (LOR) to be consistent with other studies using similar variables.

For over half of the comparisons between groups by LOR, the more years of all-English schooling they had, the higher LEP students in the same grade scored on the SRA tests. The increase in scores that each group achieved with each additional year's LOR was very small and in most cases not statistically significant. However, in 15 out of 16 comparisons (with the exception of certain

groups, to be discussed), groups with LOR of 4-5 years achieved 2-8 NCEs higher than groups with LOR of 1-2 years. In 10 out of 16 of these comparisons, groups scored at least 4 NCEs higher, a significant increase overall. Exceptions to this general pattern were 5-year-old arrivals tested in fourth grade, 6- and 7-year-old arrivals tested in sixth grade, 12-year-old arrivals tested in eighth grade, and mathematics achievement across all grades, which are all discussed in later sections.

Figures 1-4 illustrate ESL graduates' academic achievement on the SRA tests, comparing subject-area achievement by length of residence and age on arrival for each grade at which students were tested. Reported below each figure are school system means, ranging from the 59th to the 71st NCEs, with a mean of 64 NCEs for the school system across all subject areas and grades.

An example of the pattern of slightly higher achievement with each group's added year of LOR can be seen in Figure 1 in fourth graders' reading, language arts, and social studies scores, for students with LOR of 1-4 years. The 5-year-old arrivals, with LOR of 4-5 years, did not maintain the pattern and were a special case, to be discussed shortly. Fourth-grade LEP students scored at the 46th NCE in reading when first administered the SRA test at 1-2 years' LOR, and the group with 3-4 years' LOR had reached the 51st NCE, slightly above the national average. Social studies and language arts score comparisons between the same two LOR groups were higher by two NCEs (at the 52nd NCE in social studies and the 57th NCE in language arts), not a statistically significant difference.

Figure 2 illustrates sixth graders' scores, which demonstrate the pattern even more consistently in reading, language arts, science, and social studies achievement. The 6- and 7-year-old arrivals with LOR of 4-5 years and 5-6 years respectively did not maintain the pattern and are discussed shortly. The group in the sixth grade with LOR of 3-4 years increased their scores over the group with LOR of 1-2 years by 5 NCEs on the reading test (reaching the 51st NCE), 6 NCEs in language arts (at the 62nd NCE), 4 NCEs in social studies (at the 59th NCE), and 4 NCEs in science (at the 58th NCE).

Figure 3 illustrates a similar pattern for eighth graders, with the exception of the 12-year-old arrivals, another special case to be discussed later. In comparing the eighth-grade group with LOR of 2-3 years with the group with LOR of 4-5 years, test scores increased by 3 NCEs in reading (reaching the 47th NCE), 4 NCEs in language arts (at the 54th NCE), 6 NCEs in social studies (at the 58th NCE), and 3 NCEs in science (at the 51st NCE).

Figure 4 shows little difference among groups tested in the 11th grade with LOR of 1-2, 2-3, and 3-4 years, but students in the group

with LOR of 4-5 years increased their scores significantly in comparison with those of 1-2 years' LOR, with an increase of 8 NCEs in reading (reaching the 31st NCE), 7 NCEs in language arts (at the 42nd NCE), 3 NCEs in social studies (at the 38th NCE), and 7 NCEs in science (at the 37th NCE).

When compared with national averages at the 50th NCE overall, LEP students with age on arrival below age 12 appeared to be making good progress within their first 2 years of all-English schooling. By the end of 2 years' LOR, all groups with age on arrival of 6-11 had reached at least the 50th NCE on the language arts, social studies, and mathematics tests. (Although only data for 8- and 10-year-old arrivals are shown in the figures for LOR of 1-2 years, this statement is based on the assumption that if the pattern of increases in scores remained consistent, 6-, 7-, 9-, and 11-year-old arrivals would also have reached the 50th percentile within 1-2 years.) The 6- to 11-year-old arrivals tested in the fourth and sixth grades had also reached the 50th NCE in science within their first 2 years' LOR, whereas those tested in eighth grade reached the 51st NCE by 4-5 years' LOR. On the reading test, those tested in fourth and sixth grades reached the 51st NCE at the end of 3-4 years' LOR, but those tested in eighth grade had only reached the 47th NCE after 4-5 years' LOR.

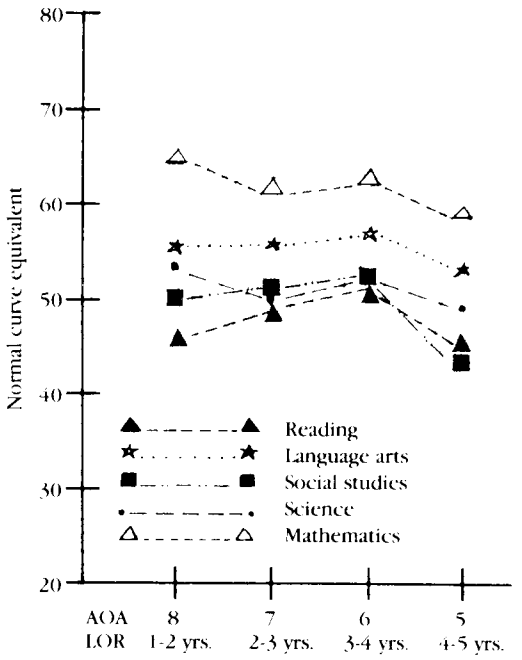
Although some groups of LEP students at all grade levels had reached the 50th NCE (the national average) on some subject-area tests within the 4-5 years measured in this study, when compared with the achievement levels of native speakers in their local school district, LEP students had not yet begun to reach the school system means across grade levels of 62-64 NCEs in reading, 62-64 NCEs in language arts, 60-65 NCEs in social studies, and 59-64 NCEs in science. They had, for the most part, met and excelled native speakers' achievement in mathematics, in which the school system mean ranged from 62 to 71 NCEs across all grade levels.

Age on Arrival

As can be seen in Figure 1, a gradual upward trend in scores for each group with an additional year's LOR was evident on the reading, language arts, and social studies tests. Only the 5-year-old arrivals, who had been in the U.S. longer than the other fourth-grade groups, did not achieve at a level expected for their length of residence. These students scored 6 NCEs lower on the reading test than the 6-year-old arrivals, who had 1 year less LOR, reaching only the 45th NCE, 4 NCEs lower in language arts at the 53rd NCE, 9

FIGURE 1

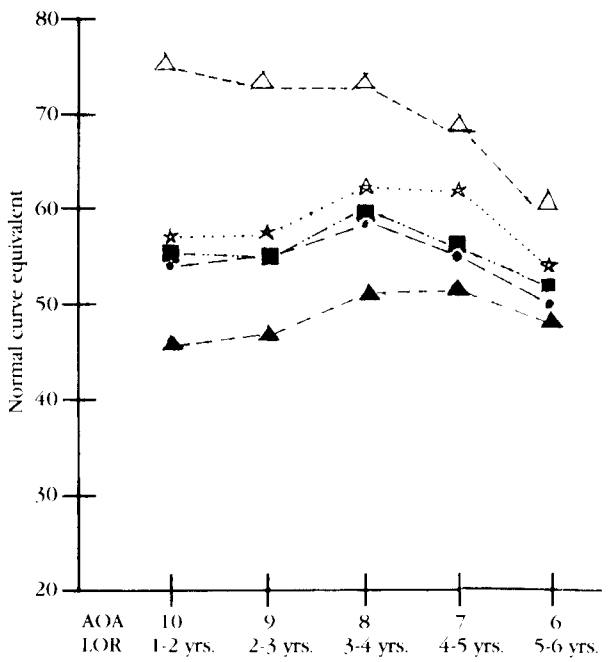
Age on Arrival (AOA), Length of Residence (LOR), and Subject-Area Achievement for 4th-Grade SRA Scores



Note: School system means for 4th grade were as follows: reading—62; language arts—63; social studies—60; science—59; mathematics—62.

FIGURE 2

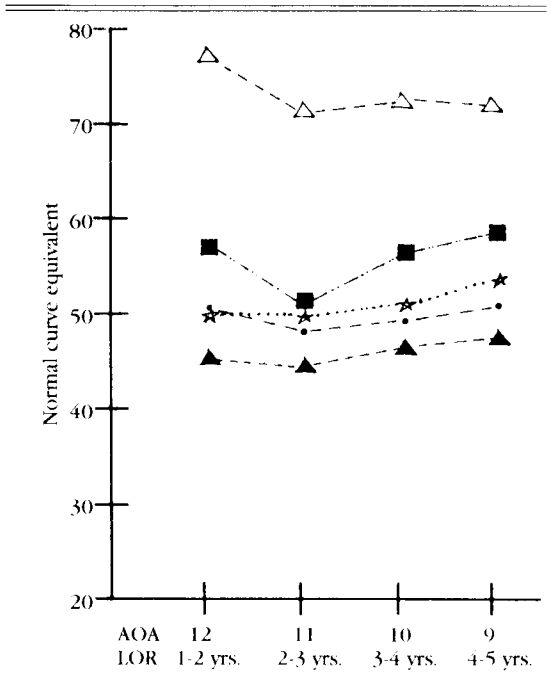
Age on Arrival (AOA), Length of Residence (LOR), and Subject-Area Achievement for 6th-Grade SRA Scores



Note: School system means for 6th grade were as follows: reading—62; language arts—64; social studies—65; science—64; mathematics—70.

FIGURE 3

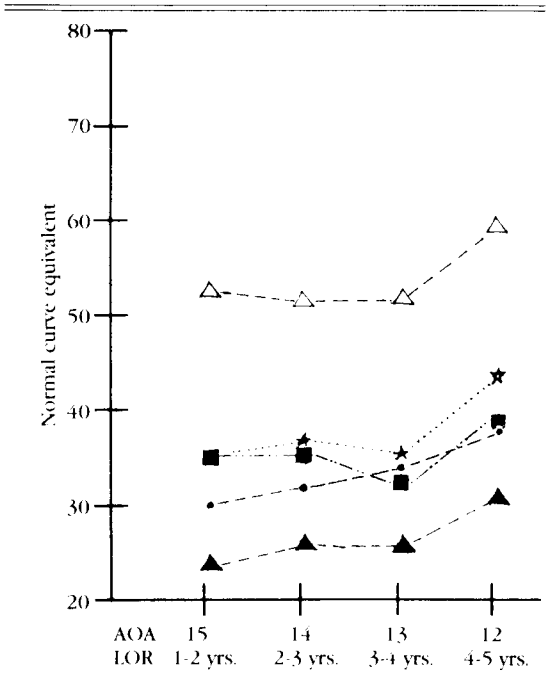
Age on Arrival (AOA), Length of Residence (LOR), and Subject-Area Achievement for 8th-Grade SRA Scores



Note: School system means for 8th grade were as follows: reading—64; language arts—64; social studies—65; science—63; mathematics—71.

FIGURE 4

Age on Arrival (AOA), Length of Residence (LOR), and Subject-Area Achievement for 11th-Grade SRA Scores



Note: School system means for 11th grade were as follows: reading—64; language arts—62; social studies—60; science—61; mathematics—67.

NCEs lower in social studies at the 43rd NCE, 3 NCEs lower in science at the 49th NCE, and 4 NCEs lower in mathematics at the 59th NCE.

A similar pattern was found in the scores on the sixth-grade test for 6- and 7-year-old arrivals, who were also significantly below the appropriate performance level for their LOR (see Figure 2). Comparing 7-year-old arrivals with 8-year-old arrivals, who had 1 year less LOR, both groups reached the same level on reading (the 51st NCE) and language arts (the 62nd NCE). In social studies, the 7-year-old arrivals were 3 NCEs lower at the 56th NCE, in science 3 NCEs lower at the 55th NCE, and in mathematics 6 NCEs lower at the 68th NCE. Even more dramatic drops occurred in the scores of 6-year-old arrivals, who had 2 years more LOR than 8-year-old arrivals, with scores in reading 3 NCEs lower (at the 48th NCE), in language arts 8 NCEs lower (at the 54th NCE), in social studies 6 NCEs lower (at the 53rd NCE), in science 8 NCEs lower (at the 50th NCE), and in mathematics 14 NCEs lower (at the 60th NCE).

In the eighth grade (see Figure 3), 12-year-old arrivals with only 1-2 years' LOR achieved at a higher level than 11-year-old arrivals with 2-3 years' LOR. In reading, the 12-year-old arrivals achieved 1 NCE higher, not a significant difference. In language arts, the two groups reached the same level. Twelve-year-old arrivals' scores in social studies were 5 NCEs higher, in science 3 NCEs higher, and in mathematics 5 NCEs higher.

Grade-Level Achievement

Another consistent pattern in scores appeared in comparisons of achievement across grade levels. As can be seen in Figure 4, the 11th-grade LEP students' test scores were dramatically lower than LEP students' performance in 4th, 6th, and 8th grades (see Figures 1-3). The 11th graders tested after 1-2 years of English schooling scored from the 23rd to the 35th NCE on reading, language arts, social studies, and science. After 3-4 years, they had made meager progress as measured on the SRA. Finally, the group with 4-5 years' LOR increased their scores by 4-8 NCEs, in comparison with the group with 3-4 years' LOR, a significant increase. Eleventh-grade ESL graduates still appeared to need several more years of schooling in English beyond the 4-5 years measured here in order to reach native-speaker levels (50th NCE nationwide and 60th-64th NCE for the local district).

Math achievement of ESL graduates in the 11th grade was much higher than 11th-grade achievement in other subject areas, reaching above national averages (53rd-59th NCEs) but still lower than 4th-, 6th-, and 8th-grade LEP students' math achievement. The 11th-grade school system mean in mathematics was at the 67th NCE.

Subject-Area Achievement

When comparing subject-area performance, Figures 1-4 illustrate high mathematics achievement, with LEP students scoring 3-6 NCEs above native speakers even in their first 2 years of all-English schooling, with the exception of 11th graders, who scored 14 NCEs below the school system mean. Scores in mathematics generally did not vary significantly across groups in each grade level, remaining basically at the same high level of achievement with each additional year of LOR. Exceptions were the 5-year-old arrivals' drops in scores on the 4th-grade test, 6- and 7-year-old arrivals' drops in scores on the 6th-grade test, 12-year-old arrivals' better performance in comparison with the other 8th-grade groups, and the increased achievement of the group in 11th grade with LOR of 4-5 years. All groups with age on arrival of 8-11 years scored above the school system mean in mathematics, regardless of LOR.

Whereas mathematics represented ESL graduates' highest performance, their lowest scores were in reading. Reading and language arts were the two subject areas directly focused on testing knowledge of L2. Performance in these two areas differed significantly, with students scoring from 5-11 NCEs higher on the language arts test, which measured punctuation, grammar, and spelling. Differences in performance among the social studies, science, and language arts tests were not for the most part significant.

DISCUSSION

Age on Arrival: 5-7

The data in this study on younger arrivals (ages 5-7) appear to support Cummins's threshold hypothesis (1976) and his interdependence hypothesis (1981b), which describes the "common underlying proficiency" of a student's two languages: "To the extent that instruction in Lx is effective in promoting proficiency in Lx, transfer of this proficiency to Ly will occur provided there is adequate exposure to Ly (either in school or environment) and

adequate motivation to learn Ly" (p. 29). Cummins argues that common underlying proficiency makes possible the transfer of cognitive academic proficiency from one language to another. He suggests that there must be some minimal literacy development in the L1 for cognitive development to transfer readily to the L2 and that this minimal "threshold" level significantly aids the process of CALP development in the L2.

The data in this study suggest that this threshold involves a minimum of 2 years of L1 schooling for students' most rapid progress in CALP development in the L2. Among first through eighth graders here for several years, only arrivals at age 5 who were tested in fourth grade and arrivals at ages 6 and 7 who were tested in sixth grade did not achieve at a rate equal to arrivals at ages 8-11, when both groups had the same LOR.

Conservatively assuming at least a minimal increase of 1 NCE for each additional year of LOR, the 5-year-old arrivals tested in fourth grade scored significantly below their predicted level of achievement, at 7 NCEs below their predicted score on reading, 5 NCEs below on language arts, 10 NCEs below on social studies, 4 NCEs below on science, and 5 NCEs below on mathematics, or an average, on all subject-area tests combined, of 6 NCEs below expected levels of performance for their LOR. Using the same measure, 6-year-old arrivals tested in the sixth grade were 5, 10, 8, 10, and 16 NCEs below their predicted scores, or an average of 10 NCEs below. Similarly, 7-year-old arrivals tested in the sixth grade were 1, 1, 4, 4, and 7 NCEs below their predicted scores, or an average of 3 NCEs below expected levels of performance for their LOR.

These 5-, 6-, and 7-year-old arrivals received the least amount of L1 schooling in comparison with all other older arrivals in the study. This was the only known variable that differentiated them from older LEP arrivals. One might question why the students who took the 4th-grade test and who had an age on arrival of 6 or 7 years did not seem to experience the same lower levels of achievement of the 6th-grade test takers whose age on arrival was 6 or 7 years. However, it is important to remember that the test at each succeeding grade level becomes cognitively more complex. Apparent lags in mastery of the content areas become more visible in the upper grades. This is especially evident in the data from the 11th-grade test.

Age on Arrival: 12-15

At first glance, the data in this study on adolescent arrivals (ages 12-15) appear to contradict Cummins's (1981b) interdependence hypothesis, which predicts that "older learners who are more cognitively mature and whose L1 proficiency is better developed would acquire cognitively demanding aspects of L2 proficiency more rapidly than younger learners" (p. 29). Even with a strong academic background in their L1, students in this study who arrived in 7th grade at age 12 and were tested 4 years later in 11th grade were substantially below national norms in achievement in all subject areas except mathematics.

In this researcher's opinion, however, these dramatic drops in adolescents' scores cannot automatically be attributed to the critical period hypothesis (Lenneberg, 1967), or, to use Oyama's (1976) term, the sensitive period hypothesis. Instead, the major difference in academic performance may be a result of the schools' greater demands on students at the secondary level and the limited length of time LEP secondary students have to reach those levels. The 8th-grade and 11th-grade versions of the SRA differ dramatically, and these differences are strongly reflected in the high school curriculum. In this study, 12-year-old arrivals taking the 8th-grade SRA test were scoring around the 50th NCE after just 2 years' schooling in English, whereas 12-year-old arrivals taking the 11th-grade SRA test after 5 years' schooling in English had only reached the 31st NCE in reading, 42nd NCE in language arts, 38th NCE in social studies, 37th NCE in science, and 59th NCE in mathematics.

It must be kept in mind that these LEP students were not being provided with any L1 content instruction to help them continue cognitive and academic subject mastery at grade level while they were acquiring beginning levels of BICS and CALP in English. By the time they had acquired enough proficiency in English to receive meaningful instruction in content-area classes, they had in the meantime lost 2-3 years of CALP development and content knowledge in mathematics, science, and social studies at their age-grade level. This put them significantly behind in mastery of the complex material required for high school students. Between their third and fourth years of schooling all in English, they began to increase their achievement levels, but even projecting this increased rate (an average of 6 NCEs per year), it might require 6-8 years' LOR for them to reach national averages of native-speaker achievement across all the subject areas. Cummins et al. (1984)

discuss this possibility in summarizing the literature on older versus younger arrivals:

> The findings of Cummins [1981a] suggest that the effects of LOR tend to diminish after 5 years and thus, in terms of immigrant students' ability to approach grade norms in L2 academic skills, there may be a critical age on arrival at about age 12, after which it will become increasingly difficult for students to catch up. (p. 79)

Another confirmation of the significant amount of L1 transfer of content knowledge to L2 and the difficulty of losing time in academic development while acquiring L2 was evident in the scoring pattern among eighth-grade test takers. Twelve-year-old arrivals with only 1-2 years' LOR performed significantly better on the science, social studies, and mathematics tests than 11-year-old arrivals with 2-3 years' LOR. The loss of content-area instruction while 11-year-old arrivals were acquiring English appeared to lower their scores, whereas the 12-year-old arrivals had the advantage of an additional year of L1 content instruction to apply to their L2 content knowledge.

In a reexamination of the data that support Cummins's hypotheses, a number of the studies focusing on the development of CALP-related skills that contrast younger arrivals with older arrivals define older students as 8-12 years of age (see, for example, Burstall, 1975; Cummins et al., 1984; Ekstrand, 1976; Ervin-Tripp, 1974; Grinder, Otomo, & Toyota, 1962; Skutnabb-Kangas & Toukomaa, 1976). This study provides further confirmation for the hypothesis that the fastest attainment of the second language for academic purposes occurs among those whose age on arrival is 8-11 years, when these students are schooled only in the L2 after arrival.

Other studies comparing younger arrivals to teenage or adult arrivals generally focus on the earliest stages of language acquisition and on BICS-related skills (see, for example, Asher & Garcia, 1969; Asher & Price, 1967; Ekstrand, 1978; Olson & Samuels, 1973; Oyama, 1976, 1978; Patkowski, 1980; Seliger, Krashen, & Ladefoged, 1975; Snow & Hoefnagel-Höhle, 1977; Stern, 1967). Although these studies show *initial* faster gains in BICS among teenage and adult arrivals (as summarized by Krashen et al., 1979), few of these studies look at student gains across time.

Two studies examining CALP gains across time (Cummins, 1981a; Snow & Hoefnagle-Höhle, 1978) appear to support faster L2 development among 12- to 15-year-old students, but both studies used the same language measures across all ages and examined absolute gains. The present study, however, focused on tests that change with each grade level, with scaled standard scores used to

compare one age with another. The SRA test reflects appropriate age-grade cognitive and academic development across time. None of the studies cited above measured *academic* gains over time in language, social studies, science, and mathematics, as did this study. A more detailed analysis of the literature on age and time variables in second language acquisition is provided in Collier (1987).

This study, therefore, supports a new hypothesis that older students who arrive at ages 12-15 experience the greatest difficulty with acquisition of the L2 for academic purposes, combined with continuing content-area development, when these students are schooled only in the L2. Such students in Grades 7-12 cannot easily afford even 1 or 2 years' loss of cognitive and academic development in all subject areas while they are mastering English. The study data suggest that secondary-level students are most in need of content-area classes taught in the L1, in order for them to stay at grade level while they are mastering English. Another alternative might be the development of accelerated content-area classes for very advanced ESL students, covering 2-3 years' academic work in 1-2 years. Content-area ESL classes taught at students' level of English proficiency should also be developed for students in the early years of ESL.

Age on Arrival: 8-11

Given that the 5-year-old arrivals performed less well than their peers in 4th grade, that the 6- and 7-year-old arrivals did less well than their peers in 6th grade, and that the 12- to 15-year-old arrivals tested in 11th grade achieved significantly lower than the national average even after 4-5 years of all-English schooling, then it would appear that the 8- to 11-year-old arrivals experienced the shortest length of time for reaching the aspects of CALP development in the L2 measured by the SRA tests.

These students had reached at least the 50th NCE in language arts, social studies, and mathematics within their first 2 years of all-English schooling. In science, the 4th- and 6th-grade test takers made it to the 53rd and 54th NCEs in just 2 years, but it took 4-5 years for the 8th-grade test takers to reach the 51st NCE. In reading, the one test focused on a pragmatic measure of language proficiency, all the 4th- and 6th-grade test takers made it to the 51st NCE after 3-4 years' LOR, but those taking the 8th-grade reading test had only reached the 47th NCE after 4-5 years' LOR. Again, it must be remembered that the tests at each succeeding level become cognitively more complex.

Although this is a remarkable accomplishment, these advantaged LEP students with a middle-class background and adequate education in their L1 have high expectations of competing with native speakers for university admission and thus need to score higher than the 50th NCE on tests focused on language. Achievement attained by native speakers in the school district, at a mean of 64 NCEs across all subject areas, would require 3-4 more years of continuing CALP development and subject knowledge in the L2, projecting the present pattern of 8- to 11-year-old arrivals' increases made each year.

Reading and Language Arts

ESL graduates' better performance on the language arts test in comparison with the reading test was indicative of the different aspects of language that the two tests measured. The language arts test measured the more mechanical, easily taught aspects of language—grammar, spelling, and punctuation. These test items provided a measure of metalinguistic knowledge about the language, from a discrete-point, language-testing perspective (Lado, 1961).

The reading test, on the other hand, included a vocabulary test of synonyms and antonyms as well as questions following reading comprehension passages. The subject matter of these passages ranged from topics taken from content areas, to consumer skills such as newspaper editorials or advertisements, to literature. In contrast to the discrete-point focus of the language arts test, test items on the reading test were much more pragmatic (Oller, 1979), measuring a wide range of language domains through a reading passage that set an age-appropriate context for the questions that followed.

The reading test also assessed thinking skills and was more closely related to the acquisition end of the acquisition-learning continuum (Krashen, 1981). On the SRA reading test, the types of items used to measure vocabulary development and reading comprehension required the use of more complex cognitive processes, at the upper levels of Bloom's taxonomy (see Bloom & Krathwohl, 1977). Such processes appear to take longer to master in L2 CALP development than the mechanics of language measured in the very limited SRA language arts test. The content of the SRA reading test was also more closely related to the content presented in language subtests of standardized tests for university admission and thus might serve as a more appropriate predictor of these LEP students' future performance on standardized language tests.

Mathematics

It is encouraging that advantaged LEP students can perform so well in mathematics, even when portions of the exam include math concepts and problem solving, which rely more heavily on language skills. Although the remarkably high mathematics achievement was the exception to other content-area achievement, ESL graduates' scores still followed the same pattern as that found in the other four content-area tests. The 5-year-old arrivals tested in 4th grade and the 6- and 7-year-old arrivals tested in 6th grade achieved significantly below their peers who had been in the United States for a shorter time. Likewise, the ESL graduates' 11th-grade achievement in mathematics was still considerably below ESL graduates' 4th-, 6th-, and 8th-grade achievement in mathematics.

CONCLUSIONS AND IMPLICATIONS

Across all subject areas tested and all grade levels combined, LEP students in this study arriving between the ages of 8 and 11 were the fastest achievers. Seven-year-old arrivals were slightly below this performance, with an average of 3 NCEs below projected scores for their LOR. LEP students arriving at ages 5 and 6 were projected to require at least 2-3 more years' LOR to reach the 8- to 11-year-old arrivals' performance level. LEP students arriving between the ages of 12 and 15 were the lowest achievers, not having reached national average scores in any subject area except mathematics after 4-5 years' LOR. They were projected to need at least another 2-3 years to reach the 50th NCE on all subject-area tests.

The data imply that 5-, 6-, and 7-year-old arrivals might acquire English for academic purposes more rapidly if they were provided a minimum of 2 years of continuing cognitive academic development in the L1. Arrivals at ages 12 to 15 cannot afford to lose time in academic instruction in the content areas taught at grade level. These subjects might be taught either through the L1 or through intensive courses taught in the L2 when students are sufficiently proficient in English to be able to work at grade level. It clearly takes a long time to acquire CALP and attain appropriate levels of academic achievement in the L2 in all the subject areas. Depending on age of arrival, it may take these advantaged LEP students anywhere from 4-8 years or more to reach the 50th NCE on standardized tests across all the subject areas. It will take them even longer to reach native speakers' attainment in their own school district.

It should not be assumed from this study that standardized,

multiple-choice tests are an adequate measure of CALP. These SRA tests measure only very limited aspects of the whole range of language proficiency. For instance, they do not measure listening comprehension, oral production, writing skills, strategic reasoning, initiative, creativity, or many pragmatic aspects of language. Moreover, students' anxiety on a timed test can greatly limit their ability to demonstrate what they know. Standardized tests have many other limitations as well.

Nevertheless, as long as these tests are used in the mainstream as a significant measure of academic achievement for students to move from one level to the next or to be selected for special academic programs, we in ESL and bilingual education need to use such tests to assess our students' ability to achieve in the mainstream. These findings show that there is no shortcut to the development of cognitive academic second language proficiency and to academic achievement in the second language. It is a process that takes a long, long time.

We plan to continue analyzing this extensive data base for additional findings. Scores from future years of testing will be added, and a time-series study will be undertaken to track each ESL graduate. New independent variables will be added in order to analyze differences in educational background, sociocultural variables, and many other factors that may influence student achievement and second language acquisition.
■

ACKNOWLEDGMENTS

I would like to thank Wayne P. Thomas for his valuable assistance in construction of the computer files and data analysis; Eda Valero-Figueira for collegial support and suggestions; and Michael O'Malley for his helpful comments on an earlier version of this article, presented at the annual meeting of the American Educational Research Association in Washington, DC, in April 1987.

THE AUTHOR

Virginia P. Collier is Assistant Professor of Education and Associate Director of the Center for Bilingual/Multicultural/ESL Teacher Preparation at George Mason University. Her research interests are focused on applications of second language acquisition research to the classroom. She is co-author of the teacher training textbook *Bilingual and ESL Classrooms: Teaching in Multicultural Contexts* (McGraw-Hill, 1985).

REFERENCES

Asher, J., & Garcia, G. (1969). The optimal age to learn a foreign language. *Modern Language Journal, 38*, 334-341.

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

Bloom, B.S., & Krathwohl, D.R. (1977). *Taxonomy of educational objectives: Handbook I. Cognitive domain*. New York: Longman.

Burstall, C. (1975). Primary French in the balance. *Educational Research, 17*, 193-198.

Collier, V.P. (1987). *Age and time variables in acquisition of second language for academic purposes* (New Focus No. 2). Wheaton, MD: National Clearinghouse for Bilingual Education.

CTB/McGraw-Hill. (1986). *California achievement tests*. Monterey, CA: Author.

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism, 9*, 1-43.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimal age question, and some other matters. *Working Papers on Bilingualism, 19*, 197-205.

Cummins, J. (1980). The entry and exit fallacy in bilingual education. *NABE Journal, 4*(3), 25-59.

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

Cummins, J. (1981b). The role of primary language development in promoting educational success for language minority students. In *Schooling and language minority students: A theoretical framework* (pp. 3-49). Los Angeles: California State University, National Evaluation, Dissemination and Assessment Center.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon, England: Multilingual Matters.

Cummins, J., & Swain, M. (1983). Analysis-by-rhetoric: Reading the text or the reader's own projections? A reply to Edelsky et al. *Applied Linguistics, 4*, 23-41.

Cummins, J., Swain, M., Nakajima, K., Handscombe, J., Green, D., & Tran, C. (1984). Linguistic interdependence among Japanese and Vietnamese immigrant students. In C. Rivera (Ed.), *Communicative competence approaches to language proficiency assessment: Research and application* (pp. 60-81). Clevedon, England: Multilingual Matters.

Dulay, H., & Burt, M. (1978). From research to method in bilingual education. In J.E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics 1978* (pp. 551-575). Washington, DC: Georgetown University Press.

Edelsky, C., Hudelson, S., Flores, B., Barkin, F., Altwerger, B., & Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics, 4*, 1-22.

Ekstrand, L. (1976). Age and length of residence as variables related to the adjustment of migrant children, with special reference to second language learning. In G. Nickel (Ed.), *Proceedings of the Fourth International Congress of Applied Linguistics* (Vol. 3, pp. 179-197). Stuttgart: Hochschulverlag.

Ekstrand, L. (1978). English without a book revisited: The effect of age on second language acquisition in a formal setting. *Didakometry*, No. 60. Malmö, Sweden: School of Education, Department of Educational and Psychological Research.

Ervin-Tripp, S.M. (1974). Is second language learning like the first? *TESOL Quarterly*, 8, 111-127.

Grinder, R., Otomo, A., & Toyota, W. (1962). Comparisons between second, third, and fourth grade children in the audio-lingual learning of Japanese as a second language. *Journal of Educational Research*, 56, 463-469.

Krashen, S.D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Krashen, S.D. (1982). Accounting for child-adult differences in second language rate and attainment. In S.D. Krashen, R.C. Scarcella, & M.A. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 202-226). Rowley, MA: Newbury House.

Krashen, S.D., Long, M.A., & Scarcella, R.C. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly*, *13*, 573-582.

Krashen, S.D., Scarcella, R.C., & Long, M.A. (Eds.). (1982). *Child-adult differences in second language acquisition*. Rowley, MA: Newbury House.

Lado, R. (1961). *Language testing*. London: Longmans, Green.

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

McLaughlin, B. (1984). *Second-language acquisition in childhood: Vol. 1. Preschool children* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Oller, J.W., Jr. (1979). *Language tests at school*. London: Longman.

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

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

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

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

Rivera, C. (Ed.). (1984). *Language proficiency and academic achievement*. Clevedon, England: Multilingual Matters.

Science Research Associates. (1978). *Science Research Associates achievement series*. Chicago: Author.

Seliger, H., Krashen, S., & Ladefoged, P. (1975). Maturational constraints in the acquisition of second languages. *Language Sciences, 38*, 20-22.

Skutnabb-Kangas, T., & Toukomaa, P. (1976). *Teaching migrant children's mother tongue and learning the language of the host country in the context of the socio-cultural situation of the migrant family.* Helsinki: The Finnish National Commission for UNESCO.

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

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

Stern, H.H. (Ed.). (1967). *Foreign languages in primary education: The teaching of foreign or second languages to younger children.* London: Oxford University Press.

Tallmadge, G.K. (1976). *Interpreting NCEs.* Mountain View, CA: RMC Research Corporation.

de Villiers, P.A., & de Villiers, J.G. (1979). *Early language.* Cambridge, MA: Harvard University Press.