

Learning English Bilingually: Age of Onset of Exposure and Rate of Acquisition Among English Language Learners in a Bilingual Education Program

Jeff MacSwan
Arizona State University

Lisa Pray
Utah State University

Abstract

This article asks whether children enrolled in a bilingual education program learn English in a reasonable amount of time, and whether older children learn English faster than younger children. Children ($N = 89$) were found to achieve parity with native English speakers in a range of 1 to 6.5 years and in an average of 3.31 years on measures of English language. Indirect comparisons with other data suggest that children in bilingual education programs learn English as fast as or faster than children in all-English programs, and an ANOVA analysis indicates that older school-age children in the sample learn English faster than younger children, $F(4, 84) = 9.037$, $p < .001$, adjusted $R^2 = .268$. The evidence supports the underlying rationale of bilingual education programs; in addition, the authors argue that English-only programs may inhibit successful learning of academic subject matter.

Introduction

Policy affecting English language learners (ELLs) has changed dramatically in recent years. California, Arizona, and Massachusetts have passed anti-bilingual education measures requiring that children “be taught English by being taught in English” (Arizona Revised Statutes, §15–752), using an all-English approach known as Structured English Immersion (SEI). An especially controversial aspect of English-only instruction measures is

the suggestion in the text of law that children would become proficient in English within a year's time: "Children who are English learners shall be educated through sheltered English immersion during a temporary transition period not normally intended to exceed one year" (Arizona Revised Statutes, §15-752). Indeed, during the campaign to pass Proposition 203 in Arizona, proponents made very strong claims about the promise of the SEI approach to teach children English very quickly. In one instance, Ron Unz, the California businessman who funded the initiative, told *The Arizona Republic* that within a few years following the passage of Proposition 203 "there will be no Arizona children in English acquisition classes." According to Unz, "Children will learn English in a couple months." Additionally, SEI advocates maintain that children younger than 10 are especially well suited to the approach, because—as the initiative claims in its preamble—"young immigrant children can easily acquire full fluency in a new language, such as English, if they are heavily exposed to that language in the classroom at an early age" (Proposition 203, Section 1).

The assumption that ELLs can learn English very quickly in an all-English instructional setting, especially if they begin when they are very young, plays a key role in the underlying rationale for SEI. In *Lau v. Nichols* (1973), the U.S. Supreme Court had found that "students who do not know English are effectively foreclosed from any meaningful education" because they cannot understand classroom instruction. SEI advocates contend that young children learn English so quickly that they can readily catch up to other students once classroom instruction has become understandable (Rossell, 2000, 2002). Proponents of bilingual education, on the other hand, have maintained that classroom instruction in the native language is necessary to help children keep up academically during the time it takes to learn English well enough to get by in an all-English instructional setting, which is understood to be a matter of years rather than months (Crawford, 2004). Opponents of the English-only measure warned that the negative effects of SEI are likely to show up most prominently in later years, when the accumulative effects of incomprehensible classroom instruction would begin to take a toll (Krashen, 1996; Crawford). Thus, the question of how much time immigrant children generally need to become proficient in English is a fundamental question underlying the current controversy.

Our study seeks to address two specific questions relevant to this policy debate. First, how long does it take school-age second-language learners (SLLs) to develop English proficiency in a bilingual education program? Second, since SEI advocates argue that younger children are especially quick to learn new languages, we ask whether younger school-age children develop English proficiency faster than older children do.

Although previous studies have been conducted to investigate these questions empirically, the results have often been problematic, generally due to the nature of the measures of language proficiency that have been employed.

In particular, some of the instruments used to assess children's growth in English are notoriously difficult, so that even native English speakers often do not pass them, as we discuss below. This suggests that some tests may be sensitive to achievement and other factors besides knowledge of English. Thus, a question of crucial first-order importance becomes, what is English proficiency and how does it differ from academic achievement?

We begin our paper with a discussion of this basic question, which forms a context for our review of previous research on the rate of acquisition of English. We then present our own study in detail, where we address our main research questions. In the concluding section, we address the implications of these findings for the policy question regarding the appropriateness of time limits and the efficacy of the underlying rationale for SEI programs, such as those mandated in California, Arizona, and Massachusetts.

What Is English-Language Proficiency?

Before one can meaningfully ask how much time is needed to learn English, the question of what English proficiency is must be addressed. While all normal children acquire the language of their speech community, it is widely known that children differ a great deal with respect to their success in school. Immigrant children, moreover, come to school with two important tasks to which they must be committed simultaneously: They must acquire a second language, English, while at the same time keeping pace academically.

Research in language acquisition has found cross-culturally that all normal children acquire the language of their respective speech communities, and do so effortlessly and without instruction (Pinker, 1994; Slobin & Bowerman, 1985). During the most active acquisition period in the preschool years, children learn approximately 10–12 new words per day, often on one exposure and in highly ambiguous contexts (Gleitman & Landau, 1994), and acquire knowledge of elementary aspects of sentence structure for which they have no evidence at all (Lightfoot, 1982). Moreover, as Tager-Flusberg (1997) has pointed out “by the time children begin school, they have acquired most of the morphological and syntactic rules of their language” (p. 188), and possess a grammar essentially indistinguishable from adults. These facts and others have led researchers to believe that language acquisition is inwardly directed by innate principles of Universal Grammar (UG) (Chomsky, 1981), or an internal “bioprogram,” as Bickerton (1981) has termed it.

UG is presumed to be a biological endowment common to the human species, and unique in essential respects. It defines a narrowly delineated hypothesis space for the language learner who uses primary linguistic data from the speech community to set options permitted by UG. Thus, UG begins in an initial state, S_0 , which successively approximates the language of the speech community through a series of intermediate states, $S_1 \dots S_n$, until it reaches a steady state, S_s , after which it appears to undergo only peripheral

changes (acquisition of new vocabulary, development of new speech styles, and so on) (Chomsky, 1986; Lightfoot, 1982; Pinker, 1994; Ritchie & Bhatia, 1999).

Of course, languages differ across communities of speakers, and across individuals as well (Fillmore, Kempler, & Wang, 1979). Thus, when we identify a speech community as “speakers of English” or “speakers of Tyrolian German,” we engage in an idealization, assuming homogeneity for the purpose of discussion, much in the way that natural scientists assume homogeneity of body organs and other objects of study. We might usefully think of “speakers of a language L” as those speakers whose internal and individual languages are each sufficiently alike as to permit intelligible intercommunication in L. In doing so, however, we recognize that in actual fact speech communities have considerable internal variation, even to the level of individual speakers (idiolects), and that speakers may be members of multiple speech communities.

Thus, a particular language, such as English, German, or Swahili, is a set of expressions defined by a grammar, a psychological mechanism which maps sound to meaning and which is represented in the mind/brain of a speaker-hearer, and a vocabulary. The grammar of a particular language is a set of values over the range of variation permitted by UG once it has entered the steady (or “mature”) state (Chomsky, 1995). In the context of first-language acquisition, then, we take “language proficiency” to be a state of linguistic maturity in which a learner has acquired a grammar which is compatible with the language of the community of origin.

Although children’s acquisition of their native language is essentially complete by the time they reach school, school-age SLLs may exhibit linguistic errors of a sort which normally developing school-age children do not exhibit in their native language. Unlike school-age native speakers, SLLs have developed only partial knowledge of the structure of their target language, and exhibit substantial errors associated with tense, case, grammatical agreement, word order, pronunciation, and other aspects of language structure. In addition, while all normal human beings acquire a language effortlessly and without instruction, second-language acquisition often meets with only partial success, at times depends upon considerable effort, and may be facilitated by purposely structured input (Bley-Vroman, 1989; Coppieters, 1987). Thus, in the case of school-age ELLs, we expect “English proficiency” to reflect growing mastery of the structure and vocabulary of English over some range of time. In terms just outlined, SLLs move through successive linguistic states, $S_1 \dots S_n$, until they enter a steady state, S_s , which is approximately compatible with the language of the new (second-language) speech community.²

Literacy and other school subjects will no doubt make use of a child’s language ability, but these seem substantially different in character. Humans acquire language by instinct, the way birds acquire birdsong, but the learning of school subjects, such as literacy, physics and mathematics, do not follow a

biologically endowed program (Chomsky, 1986; Gee, 2001; MacSwan, 2000; MacSwan & Rolstad, 2003). Academic achievement denotes a domain of knowledge that is specific to a particular human context, namely, the world of formal schooling. While all normal children develop a vocabulary and a grammatical system, not all children will come to know that Mississippi entered the Union as the 20th state, or that the square of the hypotenuse of a right-angled triangle is equal to the sum of the squares of the other two sides.

In addition, a child's tacit understanding of the rules which govern language use is also sensitive to social and situational contexts, and the interpretation of particular linguistic expressions is tied to a language user's appreciation of relevance, coherence, and context (Kehler, 2002; Sperber & Wilson, 1986). As Gee (1999) has pointed out, language use has the effect of establishing a who and a what, a socially situated person engaged in a particular kind of craft or activity—a teacher, a doctor, a member of a club or street gang, a regular at a local bar, or a student at school. These roles enter into a speaker-hearer's perspective, and are part of what Gee calls "Discourses"—ways of acting, interacting, thinking, and valuing within a particular community of speakers. Gee uses the term "social language" to denote the role of language in Discourse, the set of conventions that results in an expression of personal and social identity, and of relationships among interlocutors and participants. Thus, as we each make meaning out of language, we do far more than compute an interpretation deriving from the interaction of syntax and word meaning. We make use of a wealth of knowledge and theories about the world, and of a particular set of cultural models, practices, and beliefs.

In sum, we define knowledge of language as a linguistic construct, reflecting a grammatical system which consists of the rules and principles that govern syntax (word order), morphology (word formation), and phonology (pronunciation), and that interface with principles of discourse, pragmatics, and semantic interpretation. Speakers and communities differ with regard to the particular form these principles might take, resulting in the formation of distinctive varieties and conventions on language use; but each community nonetheless has a language every bit as rich and complex as the next (Crystal, 1986; Milroy & Milroy, 1999; Newmeyer, 1986). School-age SLLs are engaged in a process of language growth, perhaps analogous to native language growth during the preschool years. Academic achievement, as used here, is another matter.³ It denotes the content and skill specific to the school environments. School, like any environment, will have effects on children's language, but the specific ways in which school is likely to alter children's language do not amount to qualitative differences from a linguistic point of view.

Previous studies that have investigated the length of time children need to learn English have sometimes taken a rather different view of language proficiency than the perspective developed here. In particular, Cummins's (1980, 2000) distinction between Basic Interpersonal Communication Skills

(BICS) and Cognitive Academic Language Proficiency (CALP) is often used to distinguish SLLs' early development, associated with "surface fluency," from their later development in school, associated with the special demands of academic language. Apart from its controversial nature (Edelsky et al. 1983; MacSwan, 2000; MacSwan & Rolstad, 2003; Martin-Jones & Romaine, 1986; Wiley, 1996), the distinction between BICS and CALP posits that indicators of academic achievement overlap with aspects of language proficiency (Cummins, 2000). In a study concerned with how long immigrant children need to learn English, this overlap in construct definitions makes it difficult to know whether a child's difficulty at school results from an inability to speak English or from other factors associated with low academic achievement, such as socioeconomic status (SES). Thus, in our discussion below, we will relate our findings to conclusions drawn in previous studies, but will conceptualize the question a bit differently. Rather than positing separate kinds of language proficiency, we assume that growth in English-language proficiency—a singular construct, purely linguistic in nature—will precede the achievement of parity with English-background children on English-medium measures of academic achievement.

We now turn to a review of some previous studies.

Previous Studies

An early study that addressed the question of whether age of onset of exposure to a second language affected rate of acquisition was Snow and Hoefnagel-Höhle's (1978). These researchers followed a group of English-background learners of Dutch longitudinally for a 12-month period, beginning within 6 months of their arrival in Holland, and compared a second group of advanced learners who had been in Holland for at least 18 months. The study participants were middle-class English-speaking children whose families had relocated to Holland for professional reasons and who were now learning Dutch as a second language. Snow and Hoefnagel-Höhle's longitudinal group included 10 children aged 3–5, 8 aged 6–7, 13 aged 8–10, 9 aged 12–15, and 11 adults. Their cross-sectional comparison group of advanced learners included 6 children aged 6–7, 6 aged 8–10, 8 aged 12–15, and 10 adults. The authors reported that children who began learning Dutch at 3–5 years of age consistently scored worse than learners who began at later ages on all measures, and that children who began at 12–15 years of age showed the most rapid acquisition. Differences were statistically significant.

Although they were not primarily concerned with rate of acquisition in relation to proficiency of native speakers, Snow and Hoefnagel-Höhle (1978) also provided comparisons with Dutch-speaking natives ($n = 8$) for two age cohorts, 6–7 years old ($n = 8$ for longitudinal group, $n = 6$ for cross-sectional group) and 12–15 years old ($n = 9$ for longitudinal group, $n = 8$ for cross-

sectional group). Comparisons were made by calculating the means of scores for SLLs expressed as percentages of native speakers' means. For instance, it was found that, after more than 12 months of exposure to Dutch, children aged 6–7 evidenced relatively weak performance on measures of auditory discrimination (19.8%) and morphology (57%). However, in two measures of syntax, Snow and Hoefnagel-Höhle's results indicated that the younger cohort (ages 6–7) performed at levels similar to native speakers after 12–18 months, with scores of 95.4% for a sentence repetition task and 100% for a sentence judgment task. The comparison group of advanced SLLs, who had been learning Dutch for at least 18 months, had a similar score for the sentence repetition task (97%), but surprisingly had a considerably lower score for the sentence judgment task (53.8%). Tests of significance and standard deviations are not reported for these comparisons.

It should be emphasized that Snow and Hoefnagel-Höhle (1978) were specifically interested in evaluating the Critical Period Hypothesis for SLLs, not in developing empirical results that could inform the policy questions raised here concerning length of time needed to learn a second language. As such, there are a number of limitations that prevent us from interpreting their results in relation to this question. First, because evidence suggests that factors such as acculturation and motivation affect second-language learning (Bialystok & Hakuta, 1994), it is reasonable to suspect that outcomes associated with the children of professionals whose native language has particularly high SES abroad may not be generalizable to Spanish-background immigrant children of relatively low SES in the United States. Furthermore, Snow and Hoefnagel-Höhle noted that considerable individual differences in second-language ability were observed within groups in their study (p. 1126). Given the small numbers of children studied and other demographic differences, we cannot safely conclude that second-language development among immigrants in the United States would pattern the way it did for Snow and Hoefnagel-Höhle's participants. Indeed, the fact that the advanced group of late beginning SLLs compared more favorably to the native speakers than the advanced learners did on the sentence judgment task may suggest that the observed pattern does not reflect a direct upward learning trajectory, as Snow and Hoefnagel-Höhle also note, or that the sample was too small to account for potentially confounding factors. In addition, very little detail is provided regarding the sentence judgment measure, making it difficult to evaluate its face validity, and, as the authors note, large differences among adult native speakers in their ability to render grammaticality judgments have been found in numerous studies (see Shültze, 1996, for review).

Cummins (1980) reanalyzed cross-sectional data in Wright and Ramsey (1970) and Ramsey and Wright (1974) with a primary interest in disentangling age of arrival from length of residence of immigrant children. Ramsey and Wright studied 1,200 students, a 25% random sample of the Toronto School

System's Grades 5, 7, and 9 classrooms, and found that children who had arrived at 6–7 years of age did not lag behind monolingual peers on grade-level norms, but later arrivals did. Cummins noted that, when students are grouped by length of residence rather than age of arrival, one sees that older learners acquired academic second-language skills more rapidly than younger learners. However, as Cummins noted, the measures used in Ramsey and Wright's original study tended to target academic rather than pure linguistic factors. Even the test designed to measure linguistic competence (English Competence Test or ECT) exclusively assessed pronunciation and a limited number of English vocabulary items.

In Cummins's analysis, children appeared to require 2 to 3 years to approach native-level ability on the ECT, but as long as 5 years to approach grade level on academic measures in Cummins's analysis of Ramsey and Wright's (1974) data. Cummins used the terms BICS and CALP to characterize these different "levels" of language proficiency observed in students.

Cummins's reanalysis usefully reports that ELLs reach grade-level norms on measures of academic achievement in an average of 5 years. The ECT data are suggestive of how long it may have taken children in the sample to achieve English proficiency. However, because the test did not provide a comprehensive assessment of morphosyntactic structure, it might have produced false positives, children whose acquisition of phonology and vocabulary was superior but had little or no knowledge of syntax and morphology; and because the test was in part a written test, it may have produced some false negatives, children who knew English but lacked procedural knowledge for taking tests.

Collier (1987) analyzed the length of time required for advantaged ELLs to learn English for academic purposes. The study design was cross-sectional, and involved 1,548 children who lived in affluent suburbs and had parents of middle-class background in the family of origin. Collier analyzed the scores of children by three age cohorts, and discovered that the middle group, who entered a U.S. school at ages 8–11, was the fastest achievers. These students took 2–5 years to reach the 50th percentile on national norms on a standardized achievement test. Younger children in Collier's study, who entered the program at ages 5–7, were 1–3 years behind the 8–11 age cohort when both groups had the same length of residency, again on academic achievement measures taken in English. Finally, children arriving at ages 12–15 experienced the greatest difficulty, and were projected to require 6–8 years to reach grade level in academic achievement when schooled in the second language.

Like Cummins (1981), Collier (1987) was concerned with English for academic purposes, and posited that a test of academic achievement in English would measure an SLL's ability in academic registers in the target language. However, while higher scores on English-medium tests of academic achievement might reflect mastery of some aspects of English-language

proficiency in some populations of students, it is not possible to know whether lower scores indicate that a child does not understand the language of the test, or simply that he or she does not know the correct answer (see Thompson, DeCerbo, Mahoney, & Macswan, 2002, for discussion). Thus, we see the primary value of Collier's research in its characterization of achievement trajectories of middle-class ELLs on English-medium achievement tests rather than in providing an unambiguous description of how much time is required to learn English.

Like Cummins (1981) and Collier (1987), Hakuta, Butler, and Witt (2000) adopted "the rough distinction between oral English proficiency and academic English proficiency," but emphasized that the dichotomy presents "a rather crude simplification of the theory of language proficiency" (p. 4). These authors found that oral proficiency takes 2 to 5 years to develop, while academic English proficiency can take 4 to 7 years. Hakuta's group reported original data from two different school districts to draw conclusions about how long it takes ELLs in the United States to acquire English in school settings, and from two other previous published reports, Wright and Ramsey (1970), reanalyzed by Cummins (1980) and discussed above, and Klesmer (1993). The original data reported by Hakuta and colleagues were collected from two districts in the San Francisco Bay area, identified as District A and District B. District A enrolled approximately 3,400 ELLs in its total student population of over 11,000, and evidenced relatively low poverty rates in comparison with other districts in the area. The ELL population was primarily Vietnamese- and Spanish-background. The district had never provided systematic instruction in their native language. The sample consisted of 1,872 students in Grades 1–6 in spring, 1998, who met two criteria: (a) they had been in the district since kindergarten, and (b) they were classified as ELLs when they were in kindergarten. In District B, approximately 7,000 of the district's 16,000 students were predominantly Spanish-background ELLs. The sample consisted of 122 randomly selected Spanish-background students who met the same criteria as in District A but who additionally attended high poverty schools with more than 70% of students participating in free and reduced-price lunch. Participants in District A had been administered the IDEA Oral Language Proficiency Test (IPT), and those in District B took the Woodcock-Muñoz Language Survey (revised) (WMLS).

Hakuta and colleagues (2000) interpreted the results of the IPT, in which 90% of District A students demonstrated proficiency in 2 to 5 years, as a measure of oral language proficiency. The WMLS was selected because it was developed for use with native speakers of English, and includes subparts for assessments of oral language, reading, and writing, and was therefore taken to reflect student's growth in academic language proficiency. In addition to the IPT, District A also gave its ELLs the MacMillan Informal Reading Inventory, used by district personnel to signal academic competitiveness

with native English speakers. These measures indicated that 4–7 years are required for students to develop academic language proficiency, according to Hakuta and colleagues.

However, like the WMLS, the IPT is designed in part to measure CALP, which overlaps substantially with academic achievement, as the test developers also acknowledge (Williams, Ballard, Tighe, Dalton, & Amori, 1998). Thus, while the WMLS and MacMillan are clear measures of academic achievement, as Hakuta's research group suggested, we worry that results of the IPT may likewise be highly sensitive to academic achievement, as López (2001) has suggested: "Caution should be used when using the IPT tests to assess immigrant children with little or no previous educational experiences because many of the items assess academic related experiences that may not be familiar to students with such backgrounds" (p. 559). Indeed, the IPT tryout study revealed that appreciably many monolingual English-speaking children were identified as non-fluent speakers of English (Dalton, Tighe, & Ballard, 1991, pp. 29–30), perhaps because some items reflect academic achievement factors rather than language ability, as López suggested. These results and conclusions were also found in Pray's (2003) study in which the English version of the IPT was administered to native English-speaking students. Hence, the IPT is probably sensitive in large measure to achievement factors, and may incorrectly identify children as less than proficient speakers of English.

Klesmer (1993) studied 285 English-as-second-language (ESL) students of various language backgrounds and nationalities and 43 comparison students, all 12-year-olds from North York, Ontario. Students were grouped according to length of residence in Canada into seven cohorts, ranging from 6 months to 6 years, and were administered tests of language proficiency, reading, writing, analogies, antonyms, and vocabulary. Hakuta and colleagues (2000) summarized Klesmer's results by transforming them into deficits with respect to the English-speaking controls, expressed in standard deviation units. The results show on measures of oral expression and listening comprehension that after 5 years of residence, students remain .75 standard deviation units below the performance of native-speaking students. On tests of literacy, the North York students were .5 standard deviation units behind the control group after 5 years.

In Table 1, we present a summary of the studies reviewed; Cummins's (1980) study is a reanalysis of Wright and Ramsey's (1970), and Hakuta and colleagues incorporate Wright and Ramsey and Klesmer as two of the four data sources they consider.

Finally, a secondary question of interest in our study asks whether younger or older children have a learning advantage in a second language. As mentioned earlier, Snow and Hoefnagel-Höhle (1978) found that children who began learning Dutch at 3–5 years of age consistently scored worse than learners who began at later ages on all linguistic measures, and that children who

Table 1

Summary of Previous Research

Source	Time to English proficiency benchmark	Time to academic achievement benchmark
Snow & Hoefnagel-Höhle (1978)	More than 1.5 years	Not investigated
Cummins (1980) Wright & Ramsey, 1970; Ramsey & Wright, 1974	2 to 3 years	3 to 5 years
Collier (1987)	Not investigated	2 to 8 years, depending on age of arrival
Hakuta, Butler, & Witt (2000) Wright & Ramsey, 1970 Klesmer, 1993	2 to 5 years	4 to 7 years

began at 12–15 years of age showed the most rapid acquisition. Indeed, as summarized in Krashen, Long, and Scarcella (1979), a frequent finding is that older children and adults initially acquire many aspects of a second language faster than younger children, but that younger children have a long-term advantage in terms of ultimate attainment. We will present an analysis of our data to address this issue as well as our main question: How long does it take immigrant children to learn English in a bilingual education program?

Current Study

Setting

Our study used specific information collected from students enrolled in an urban elementary school district, located in Central Arizona. This district was selected because it had a well-designed bilingual education program, independently determined by a review of the literature (Brisk, 1998; Faltis & Hudelson, 1998), and because it used the Bilingual Syntax Measure (BSM), discussed below, to assess language proficiency in its ELL population. The district is comprised of 14 elementary and middle schools servicing more than 11,000 students. Approximately 65% of students in the district qualified for free or reduced-price lunch. The district reported that about 96% of teachers servicing ELLs had appropriate certification to do so. The six schools in

which we were permitted to collect data collectively enrolled about 4,600 students in 1999, 689 (15%) of whom were classified as ELLs, with Spanish as the predominant home language.

Upon enrollment in the district, students who indicated that a language other than English was spoken at home were given the BSM language assessment; those who scored in the “non-English proficient” range were placed in a bilingual classroom. Students progressed through a series of four distinct classroom levels, depending upon indicators of appropriate growth in English-language proficiency and academic achievement. As students moved through the progression, teachers initially taught substantially in the native language, and increasingly used English using “sheltered techniques,” strategies that involve specifically designed academic instruction guided by the teacher’s sensitivity to students’ current state of proficiency in English. It is important to emphasize that levels did not correspond to grades; a student might have been in any given level for more than 1 year. In addition, each school in our study had an ESL and a bilingual classroom at every grade, so that students who entered school with limited English-language proficiency could begin in a bilingual classroom regardless of grade level.

In addition, the district included monolingual and proficient bilingual students at each level as a way of enhancing diversity within the classroom. The model employed was generally consistent with widely accepted criteria of well-designed bilingual education programs, and in particular strongly reflected the “Graduate Exit, Variable Threshold Model” outlined in Krashen (1996) and used in many California districts before the passage of Proposition 227 (Crawford, 2004). Students were exited from the bilingual program and placed in a regular mainstream classroom when a local school committee determined, based upon a portfolio of indicators, that the child could do well in a regular classroom without special language assistance services. The six schools at which we collected data were selected because district administrators responsible for the ELL programs had identified these sites as true to the district’s bilingual education model.

Instrumentation

Unlike the IPT and the WMLS, which were built around the distinction between BICS and CALP, the BSM purports to assess “the psycholinguistic construct of language development as indicated by syntactic proficiency” (Burt, Dulay, Hernández & Taleposos, 1980, p. 8). As such, it may be expected to measure English-language proficiency independently of academic achievement.

Although the BSM is available in both English and Spanish, our analysis only involves scores on the English version of the test. The test is individually administered in an oral interview format using colorful cartoon representations designed to prompt students to produce specific linguistic structures; the

BSM I is used with children in Grades K–2, and the BSM II with children in Grades 3–12. The test assesses a wide range of tacit knowledge of syntax, including simple and complex sentences, declaratives and interrogatives, sentential truncation, clausal subordination, conditionals, indirect questions, the use of subjects, negation, auxiliaries, adverbs, prepositions; internal constituent structure of noun phrases, verb phrases, prepositional phrases; auxiliary sequences involving progressives, perfect regular/irregular, present, past, future; subject-verb agreement; and wh-question constructions. The authors arranged syntactic structures in a hierarchical order of acquisition based on the results of a study involving 775 Spanish-background ELLs in Grades 3 through 12, a result consistent with other published research on the “natural order” in which language learners progress (Brown, 1973; Burt & Dulay, 1974; Long, 1990). These theoretical considerations, in addition to the extensive range of grammatical constructions assessed, give strong evidence for the content validity of the BSM.

In addition to assessing a wide range of syntactic constructions, the BSM requires correct answers to be both syntactically and morphologically accurate, and further requires students’ responses to be “conversationally appropriate” so that answers which are pragmatically ill-formed are regarded as incorrect. Because students must understand and render intelligible answers, there is a sense in which phonological ability is also assessed by the BSM, although no items are constructed in relation to phonological theory. The test does not involve specialized vocabulary, and specifically is not intended as a measure of vocabulary. Vocabulary is notoriously difficult to measure (Bryson, 1994; Cooper, 1997), and varies tremendously as a function of accidental experience; syntax and morphology, on the other hand, reflect an internalized system of rules that is remarkably stable within a community of speakers. Thus, in terms of its content validity, the BSM may be regarded as a good measure of English-language proficiency, as defined earlier, a grammatical system that consists of the rules and principles which govern syntax (word order), morphology (word formation), and phonology (pronunciation), and which interface with principles of discourse, pragmatics, and semantic interpretation.

In addition, while we pointed out that even native English speakers often do not do well on the IPT and WMLS—an indication, we argued, that the tests assess academic achievement or other factors in addition to language proficiency—Burt and colleagues (1980) reported that 100% of native English speakers scored in the “English proficient” range (levels 5 and 6) on the BSM in field testing. The study involved 489 Spanish-background and 72 English-background (native) students. Spanish-background SLLs, however, evidenced much more variance in their performance. These results are as expected for a purely linguistic measure of English-language proficiency.

Additionally, in a study involving the BSM, Saville-Troike (1984) found that the test correlated only .258 with the reading subtest of the Comprehensive Tests of Basic Skills (a domain of academic achievement), and the correlation was not significant, whereas the test correlated very highly with other linguistically oriented measures of oral language (Northwest Syntax Screening Test, $r = .713$; Functional Language Survey, $r = .819$; both are significant at the .05 level, $N = 19$). Campbell (1960) noted that a test should correlate more highly with variables to which it should have a theoretical relationship (convergent validity) than it does with theoretically distinct variables (divergent validity). Saville-Troike's results therefore suggest, in addition to field testing reported by Burt and colleagues (1980), that the BSM would be unlikely to produce false negatives, proficient speakers of English assessed as limited English speakers due to the test's sensitivity to other factors. In addition, the strong correlation of the test with other linguistically-oriented measures discussed above, and the faithfulness of test items to a wide range of grammatical constructions reflective of the construct definition, further leads us to believe that the test would be unlikely to identify false positives, non-proficient speakers of English assessed as proficient in English.

García (1985) and Shellenberger (1985) reported that reliability on the BSM was established in a pilot study of 85 students. The test developers are reported to have provided two measures of reliability, internal consistency and test-retest reliability. Alpha coefficients based on internal consistency scores of clusters and totals for the English version of the test ranged from .80 to .90. Test-retest correlation ranged from .82 to .96. These reliability coefficients reach levels generally regarded as adequate for accurate discrimination among individual examinees for a test with this purpose. For further discussion, see García (1985) and Shellenberger (1985).

Sample

Our sample consisted of 89 identified Spanish-background ELLs from six schools in the district described above. Each student so identified was included in the study according to two criteria: (a) The student's file contained at least two BSM scores, and (b) the student's first BSM score was a 1 (*No English*) and a later score was a 5 or 6 (*Proficient English*). Students were assessed at least every 2 years, in accordance with state law in effect at the time.

Students who were tested more than 2 years apart were included in the study only if the score following the gap was less than 5, because we could not know if the student had attained English proficiency before the subsequent testing suggested. Because we were interested in studying progress over time, it was necessary to exclude students who did not have repeated measures of language proficiency on the BSM. Furthermore, because we sought to discover how long children needed to develop proficiency in English, we

excluded cases in which scores of 5 or 6, indicating English proficiency, were not available. In addition, we excluded children who had been classified as special education students. Otherwise, all students in our selected school sites who met these criteria were included in the study. In the end, 89 students (about 10.24% of the total 869 ELLs) were included in our study.

A natural worry is that our decision to exclude students whose academic files did not contain a BSM score of 5 or 6 would cause us to overlook students, possibly many, who took considerably longer to learn English than our 6- to 8-year window permitted us to detect. However, after applying our other selection criteria, only 10 students were excluded based upon this condition: 4 from a school with an ELL population of 177, and 6 from a school with an ELL population of 207. In all, our six sites enrolled 869 ELLs, so that only 1.15% of this group might have escaped our notice.

Findings

At the outset, we asked two questions:

1. How long does it take school-age SLLs to develop English proficiency in a bilingual program?
2. Do younger school-age children develop English proficiency faster than older children do?

To address the first question, we calculated the time elapsed between the first BSM score, always a 1 (*No English*) in our data, and a 5 or 6 (*Proficient English*) to define a rate of acquisition of English. Further analysis indicated that the average number of years required for children to achieve a score of 5 or 6 on the BSM was 3.31, with a standard deviation of 1.31 years. The range was 5.58, with a minimum of .92 ($n = 1$) and a maximum of 6.50 ($n = 1$). We found that 68.5% of students achieved English proficiency in 4 years, and after 5 years 92.13% had done so. We emphasize, too, that there is an impressive range in the rate of English acquisition ($SD = 1.31$ years), suggesting that children exhibit substantial variation in the time it takes them to learn English, as Snow and Hoefnagel-Höhle (1978) also found (see Table 2).

In response to the second question, we calculated the rate of language acquisition by child's grade level at the time the first BSM was administered. A one-way ANOVA, using unweighted means, was conducted to evaluate the relationship between the grade at which the first BSM was administered and the rate of English acquisition. The ANOVA was significant, $F(4, 84) = 9.037$, $p < .001$, adjusted $R^2 = .268$. The strength in the relationship between the grade at which the child was first tested and the rate of English acquisition was strong, with grade level of first testing accounting for 27% of the variance of the dependent variable. Table 3 displays the difference between means.

Table 2

Number of Students Within Specified Ranges of Rate Acquisition of English

Years	<i>N</i>	Relative frequencies (%)	Cumulative frequencies (%)
0.92–1.0	2	2.25	2.25
1.1–2.0	17	19.10	21.35
2.1–3.0	19	21.35	42.70
3.1–4.0	23	25.84	68.54
4.1–5.0	21	23.60	92.13
5.1–6.0	4	4.49	96.63
6.1–6.5	3	3.37	100

Table 3

Rate of Acquisition of English by Grade Level

Grade level	<i>M</i>	<i>SD</i>	<i>N</i>
K	3.7702	1.2057	61
1	2.1940	0.8743	10
2	2.6740	0.9706	12
3	1.9167	0.4291	5
Total	3.3100	1.3076	89

Consistent with previously published research, we observe a clear pattern indicating that younger children systematically require more time to achieve proficiency in English. This effect is unlikely to be related to differential treatment by grade level, since the schools we studied did not prescribe different treatments according to students' grade levels.

We now turn to a discussion of these findings in the context of the current policy debate.

Discussion

We wish to frame our discussion in terms of specific questions that have arisen in relation to the current policy debate.

Does Bilingual Education Delay English-Language Acquisition?

Our study, which is situated in a well-designed bilingual education program, reveals that children acquire English very rapidly, requiring 3.31 years on average, with a range of 1 to 6.5 years. After 5 years, 92.13% of the students we studied had achieved native-level proficiency in English, as measured on the BSM.

It would be helpful to know how quickly children learn English in all-English environments, such as that mandated by Unz and recommended by Rossell (2000, 2002), as a context for judging whether bilingual education has the effect of “delaying” English acquisition. Although direct comparisons are not possible because different tests were used, among other factors, data reported in Wright and Ramsey (1970) and Klesmer (1993) suggests that rate of acquisition is not increased by an all-English instructional program. For instance, on a test of grammatical function words, students in Wright and Ramsey’s data set were still .5 standard deviations below native-speaking norms after 7 years. On a measure of oral expression, children in Klesmer’s data set were .75 standard deviations below native-speaking norms after 5 and 6 years. Additionally, Mahoney, MacSwan, and Thompson (2005) analyzed statewide language proficiency data in Arizona for the IPT and Language Assessment Scales-Oral (LAS-O) and found that concurrent with a strict implementation of the state’s English-only law 89% of ELLs who tested non-proficient in 2003 were still not proficient a year later, in 2004.

Hence, while no study has been conducted to determine the relative effectiveness of bilingual and all-English instructional programs in terms of rate of acquisition of English, the available evidence suggests that children in bilingual education programs learn English as fast as or faster than children in all-English programs.

Thus, we may conclude that, relatively speaking, students in bilingual education acquire English very quickly. Moreover, students’ involvement in a bilingual education program had no negative effects on the rate of acquisition of English, and we can reasonably assume that academic content instruction in the children’s native language helped them to keep up academically during the time needed to learn English well.

Do Younger Children Learn English Faster Than Older Children?

An important assumption of SEI is that younger children are especially fast to learn English when heavily exposed to it, so that total immersion in English is said to be especially effective at the younger ages. However, a very

clear pattern emerges in our data, which is consistent with others' findings (Krashen, Long, & Scarcella, 1979; Krashen, Scarcella, & Long, 1982; Snow & Hoefnagel-Höhle, 1978), that school-age children require more time to reach parity with native speakers at each successive grade level, K through 3. We conclude, then, that younger children do not generally learn English faster than older children, as is commonly believed.

Are Across-the-Board Time Limits Justified?

Every study that has been conducted to address the question of how long children need to learn English has found considerable variability. In our study, while children required an average of 3.31 years to attain native-level proficiency in English, there was considerable variability in rate, ranging from 1 to 6.5 years. The underlying cause of this variability is not understood. Children appear to acquire their first language on a fairly precise timetable, but once they enter the school years a great deal of variability is introduced, partly controlled by age of onset of the second language.

Whatever the cause, it would appear arbitrary and capricious to limit the number of years students are given to learn English. Thus, in our view, there would appear to be no empirical or ethical justification for placing across-the-board restrictions on how long children may receive language assistance services. Children who experience greater difficulty should receive an increase in services, not a decrease.

Is One Year Long Enough for Children to Acquire English?

Our data reveal that only 2.25% of children attained proficiency in 1 year's time. After 2 years, a little over 20% had reached parity with native speakers of English. One sees similar patterns in data pertaining to students in all-English programs. For instance, in Wright and Ramsey's (1970) data, children were 2.5 standard deviations below native-speaking norms on a measure of grammatical function words, and Klesmer's (1993) data showed that children in all-English programs were more than 1 standard deviation below native-speaking norms on a measure of listening comprehension. The data reviewed by Mahoney and colleagues (2005) consider only a 1-year span, but there we find only 11% of children tested on the IPT and the LAS-O achieving proficiency within a year's time. Clearly, except in highly unusual cases, children do not achieve parity with native speakers on measures of English-language proficiency after 1 or even 2 years of enrollment.

Klesmer's (1993) finding regarding listening comprehension is especially informative given Rossell's (2000, 2002) recent claims. She conjectures, based on inappropriate evidence, that after 1 year's time children know enough English to understand instruction, even though they may not have reached parity with native speakers. Considerably many factors will determine whether

a child can understand a teacher, not the least of which is the teacher's special training to communicate with ELLs. Based on no scientific evidence or theoretically defensible framework, Rossell (2002) recommends that students in immersion be limited to 1 year, and students in bilingual education to 2 years, even if data collected by the local school determines that they still do not know English.

What Kinds of Services Should Districts Offer to ELLs?

The National Research Council, primarily on the strength of Willig's (1985) meta-analysis program comparison studies, recognized "the elements of positive relationships that are consistent with empirical results from other studies and that support the theory underlying native language instruction" in both its 1992 and 1998 reports (August & Hakuta, 1998, p. 105; see also Meyer & Fienberg, 1992).⁴ However, August and Hakuta added a note of caution:

We see little value in conducting evaluations to determine which type of program is best. First, the key issue is not finding a program that works for all children and all localities, but rather finding a set of program components that works for the children in the community of interest, given the goals, demographics, and resources of that community. (p. 147)

Based on the available evidence, it makes no sense to restrict districts' options regarding the sort of program or programs they can make available to children. In particular, to the extent the current findings are indicative of other programs, bilingual education students appear to evidence no disadvantage over students in other programs in terms of the rate at which they acquire English. National or statewide mandates, made without considering local resources or the characteristics of particular students, cannot be justified.

When Should Children Be Exited From Language Assistance Services?

Rossell (2000, 2002) advocates that ELLs be exited from language assistance programs within 1 year, because such programs, she believes, may become a hindrance if provided to children who no longer need them. For Rossell (2000), children no longer need such programs once they can understand the teacher, which she believes is accomplished well within a year:

... all students understand enough English sometime during the first year to be able to comprehend English instruction. I base this conclusion on research conducted in Canada and the U.S. on immersion programs, research conducted in the U.S. and Europe on newcomer centers, my

conversations with LEP students in bilingual and ESL classrooms around the U.S., and my conversations with formerly LEP students in my classes at Boston University. (pp. 14–15)

Rossell (2000) cites two studies of Canadian immersion programs (Genesee, 1984; Swain & Lapkin, 1982), both of which relate the teachers' impression that children understood instruction within a year's time. However, these studies do not meet Rossell's criteria published elsewhere (Rossell & Baker, 1996) for scientific research, and involved privileged, self-selected, and highly motivated students. Furthermore, and very importantly, the students were said to understand instruction in the immersion classroom, that is, they may have understood instruction provided by teachers who were specially trained to make language input comprehensible to learners.

None of the other reports Rossell alludes to are studies at all, but rather simple descriptions of programs for newcomers offered in the United States and Europe. Rossell (2000) then provides two kinds of anecdotal evidence to support her claim that children can understand instruction within a year's time. First, she reports,

I have also had conversations with LEP students in public schools in California, Massachusetts, New York City, and St. Paul Minnesota. In most ESL classrooms I have been in, there are one or two students who are working independently because they already know what is being taught. I have taken the opportunity to talk to these students about how long it took them before they could understand what the teacher was saying in English when they entered the school. Those who started in September, having just come from a foreign country, believe they understood what the teacher was saying by the Christmas break. (p. 15)

Our data, too, reveal that in exceptional cases children develop English-language proficiency within the first year. Clearly, however, it is unreasonable to generalize on the basis of an informal chat with one or two students that no child should be allowed to remain in a language assistance program for more than a year.

Rossell's (2000) final source of evidence comes from informal conversations with young adult immigrants enrolled in her classes at Boston University. She asks students who successfully graduated from high school and entered college, and who were subject to her influence in class, to recollect how long it took them before they believed they could understand the teacher at the time they first enrolled in primary school. "[A]ll believed they could understand the teacher completely by the end of their first year in an English speaking classroom (p. 16)," she reported.

Based upon these transparently flawed data, Rossell (2002) recommended that Proposition 227, the anti-bilingual education proposition sponsored by Ron Unz in California, be amended to provide that children cannot be kept in a SEI classroom for more than 1 year (or for more than 2 years, for bilingual education students) “regardless of their test scores” (p. 101). However, Rossell’s position is based on a flawed analysis of the available evidence, one so flawed that any competent researcher, regardless of specialization, should be able to recognize it immediately.

We believe that ELLs can reasonably be expected to have fully benefitted from an ESL or bilingual education program once they have learned English well enough to manage without assistance in a mainstream classroom. In our research, we have recommended the BSM over other assessments in the context of a very narrow set of concerns, namely, identifying growth in English-language proficiency distinct from growth in academic achievement and other areas. However, at the interface between academic achievement and language proficiency is the cultural world of school, involving, among other things, the acquisition of a social language (a system of pragmatics and discourse appropriate to the school environment) that coheres with students’ developing identities as successful students in U.S. schools. Hence, schools should consider a broad range of issues before placing children in mainstream classrooms—not only should they have achieved a reasonable level of proficiency in English, but they should also be ready to engage texts and perform classroom assignments at grade level. Indeed, before being transitioned, schools should ascertain that a child has achieved second-language instructional competence maybe, SLIC consisting in an ability to understand instruction and perform grade-appropriate school activities using the second language alone, in the local educational setting (MacSwan & Rolstad, 2003).

References

- Arizona Revised Statutes*. (n.d.). §15–752. Retrieved November 22, 2005, from <http://www.azleg.state.az.us/FormatDocument.asp?inDoc=/ars/15/00752.htm&Title=15&DocType=ARS>
- August, D., & Hakuta, K. (Eds.). (1998). *Educating language-minority children*. Washington, DC: National Academy Press.
- Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second language acquisition*. New York: Basic Books.
- Bickerton, D. (1981). *The roots of language*. Ann Arbor, MI: Karoma Press.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41–67). New York: Cambridge University Press.

- Brisk, M. E. (1998). *Bilingual education: From compensatory to quality schooling*. Mahwah, NJ: Erlbaum.
- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Bryson, B. (1994). Order out of chaos. In P. Eschholz, A. Rosa, & V. Clark (Eds.), *Language awareness* (6th ed., pp.108–121). New York: St. Martin's Press.
- Burt, M., & Dulay, H. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24(1), 37–53.
- Burt, M., Dulay, H., Hernández, C. E., & Taleposos, E. (1980). *Bilingual syntax measure II: Technical manual*. New York: Harcourt Brace Jovanovich.
- Campbell, D. T. (1960). Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist*, 15, 546–553.
- Carroll, S. (2000). *Input and evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris Publications.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.
- Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21(4), 617–641.
- Cooper, T. (1997). Assessing vocabulary size: So, what's the problem? *Language Matters*, 28, 96–117.
- Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language*, 63, 544–573.
- Crawford, J. (2004). *Educating English learners: Language diversity in the classroom* (5th ed.). Los Angeles: Bilingual Education Services.
- Crystal, D. (1986). The prescriptive tradition. In D. Crystal (Ed.), *The Cambridge encyclopedia of language* (pp. 2–5). Cambridge, UK: Cambridge University Press.
- Cummins, J. (1980). The crosslinguistic dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14(3), 175–187.
- Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics*, 11(2), 132–149.

- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.
- Edelsky, C., Hudelson, S., Flores, B., Barkin, F., Altweger, J., & Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics*, 4, 1–22.
- Faltis, C. J., & Hudelson, S. (1998). *Bilingual education in elementary and secondary school communities: Toward understanding and caring*. Boston: Allyn & Bacon.
- Fillmore, C., Kempler, D., & Wang, S. Y. (1979). *Individual differences in language ability and language behavior*. New York: Academic Press.
- García, E. E. (1985). Review of the bilingual syntax measure II. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (pp.197–198). Lincoln: The University of Nebraska Press.
- Gee, J. P. (1999). *An introduction to discourse analysis: Theory and method*. New York: Routledge.
- Gee, J. P. (2001). Progressivism, critique, and socially situated minds. In C. Dudley-Marling & C. Edelsky (Eds.), *The fate of progressive language policies and practices* (pp. 31–58). Urbana, IL: National Council of Teachers of English.
- Genesee, F. (1984). French immersion programs. In S. Shapson & V. D’Oyley (Eds.), *Bilingual and multicultural education: Canadian perspectives*. (pp.33–54) Clevedon, UK: Multilingual Matters.
- Gleitman, L., & Landau, B. (1994). *The acquisition of the lexicon*. Cambridge: MIT Press.
- Grimshaw, J. (1981). Form, function, and the language acquisition device. In C. L. Baker & J. McCarthy (Eds.), *The logical problem of language acquisition* (pp. 165–182). Cambridge: MIT Press.
- Hakuta, K., Butler, Y., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara: University of California Linguistic Minority Research Institute.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: Center for the Study of Language and Information.
- Klesmer, H. (1993). *Development of ESL achievement criteria as a function of age and length of residence in Canada: ESL achievement project*. Ontario, Canada: North York Board of Education.
- Krashen, S. (1996). *Under attack: The case against bilingual education*. Culver City, CA: Language Education Associates.
- Krashen, S., Scarcella, R. C., & Long, M. H. (1982). Age, rate and eventual attainment in second language acquisition. In S. Krashen & R. Scarcella (Eds.), *Child, adult differences in second language acquisition* (pp.161–172). Rowley, MA: Newbury House Publishers, Inc.

- Krashen, S., Long, M., & Scarcella, R. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly*, 13, 573–582.
- Lau v. Nichols*. (1973). No. 72–6520, Supreme Court of the United States, 414 U.S. 563; 94 S. Ct. 786; 39 L. Ed. 2d 1; 1974 U.S. LEXIS 151.
- Lightfoot, D. (1982). *The language lottery: Toward a biology of grammars*. Cambridge: MIT Press.
- Long, M. (1990). The least a second language acquisition theory needs to explain. *TESOL Quarterly*, 24, 649–666.
- López, E. C. (2001). Review of IDEA Oral Proficiency Test. In B.S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 558–560). Lincoln: The University of Nebraska Press.
- MacSwan, J. (2000). The threshold hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences*, 20(1), 3–45.
- MacSwan, J., & Rolstad, K. (2003). Linguistic diversity, schooling, and social class: Rethinking our conception of language proficiency in language minority education. In C. B. Paulston & R. Tucker (Eds.), *Essential readings in sociolinguistics* (pp. 329–340). Oxford: Blackwell.
- Mahoney, K., MacSwan, J., & Thompson, M. (2005). The condition of English language learners in Arizona: 2005. In D. García & A. Molnar (Eds.), *The condition of preK–12 education in Arizona, 2005* (pp. 3.1–3.24). Tempe, AZ: Education Policy Research Laboratory.
- Martin-Jones, M., & Romaine, S. (1986). Semilingualism: A half-baked theory of communicative competence. *Applied Linguistics*, 7(1), 26–38.
- McDaniel, D., McKee, C., & Cairns, H. S. (Eds.). (1996). *Methods for assessing children's syntax*. Cambridge: MIT Press.
- Meisel, J. M. (2000). On transfer at the initial state of L2 acquisition. *Working papers in multilingualism* (pp. 2–23). Hamburg, Germany: Hamburg University Centre for the Study of Multilingualism.
- Milroy, J., & Milroy, L. (1999). *Authority in language: Investigating standard English*. New York: Routledge.
- Newmeyer, F. J. (1986). *The politics of linguistics*. Chicago: The University of Chicago Press.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1994). *The language instinct: How the mind creates languages*. New York: William Morrow and Company.

- Pray, L. (2003). An analysis of language assessments used in the referral and placement of language minority students into special education. *Digital Dissertation*, DAI-A 64/03, 766.
- Proposition 203. (2000). Retrieved November 22, 2005, from <http://www.ade.state.az.us/asd/lep/PROPOSITION203.pdf>
- Ramsey, C. A., & Wright, E. N. (1974). Age and second language learning. *The Journal of Social Psychology*, 94, 115–121.
- Ritchie, W. C., & Bhatia, T. K. (1999). Child language acquisition: Introduction, foundations, and overview. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 3–28). San Diego, CA: Academic Press.
- Rolstad, K., Mahoney, K., & Glass, G. V. (2005). The big picture: A meta-analysis of program effectiveness research on English language learners. *Educational Policy*, 19, 572–594.
- Rossell, C. (2000). *Different questions, different answers: A critique of the Hakuta, Butler and Witt report, "How long does it take English learners to attain proficiency?"* Washington, DC: Read Institute.
- Rossell, C. (2002). *Dismantling bilingual education, implementing English immersion: The California initiative*. Manuscript, Boston University. Retrieved November 8, 2005, from <http://web.bu.edu/POLISCI/CROSSELL/CRossell.html>
- Rossell, C., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English*, 30(1), 7–74.
- Rossell, C., & Ross, J. M. (1986). The social science evidence on bilingual education. *Journal of Law and Education*, 15, 385–419.
- Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18(2), 199–219.
- Shellenberger, S. (1985). Review of bilingual syntax measure II. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (pp. 198–199). Lincoln: The University of Nebraska Press.
- Shültze, C. (1996). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- Slobin, D., & Bowerman, M. (1985). *Crosslinguistic evidence for the language making capacity: What shapes children's grammar?* Hillsdale, NJ: Erlbaum.
- Snow, S., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 49, 1114–1128.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, UK: Cambridge University Press.

- Swain, M., & Lapkin, S. (1982). *Evaluating bilingual education: A Canadian case study*. Clevedon, UK: Multilingual Matters.
- Tager-Flusberg, H. (1997). Putting words together: Morphology and syntax in the preschool years. In J. Berko-Gleason (Ed.), *The development of language* (pp. 159–209). Boston: Allyn & Bacon.
- Thompson, M. S., DiCerbo, K., Mahoney, K. S., & MacSwan, J. (2002). ¿Éxito en California? A validity critique of language program evaluations and analysis of English learner test scores. *Education Policy Analysis Archives*, 10(7), entire issue. Retrieved November 8, 2005, from <http://epaa.asu.edu/epaa/v10n7/>
- Wiley, T. (1996). *Literacy and language diversity in the United States*. McHenry, IL: Center for Applied Linguistics and Delta Systems, Co.
- Williams, C., Ballard, W., Tighe, P., Dalton, E., & Amori, B. (1998). *IDEA Oral Language Proficiency Test*. Brea, CA: Ballard and Tighe Publishers.
- Willig, A. C. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 56(3), 269–317.
- Wright, E. N., & Ramsey, C. (1970). *Students of non-Canadian origin: Age on arrival, academic achievement and ability*. Research Report No. 88. Toronto Board of Education.

Endnotes

- ¹ The Arizona Republic. (2000, October 29). *Few kids affected by Proposition 203*.
- ² Researchers in second language acquisition dispute the nature of the mapping to S_s, but this debate will not be relevant to the present discussion. For recent reviews, see Meisel (2000) and Carroll (2000).
- ³ Although we see academic achievement as distinct from English-language proficiency, it might nonetheless be true that increases in academic achievement have a positive effect on the acquisition of English, as Krashen (1996) has argued, because knowledge of formal school settings will form part of the conceptual structure that learners have available to guide interpretations of linguistic expressions used in turn to construct semantic models that correspond to particular structural analyses of utterances. Thus, “semantic bootstrapping” (Grimshaw, 1981; Pinker, 1984) of language structure may take place in the second-language environment, facilitated by growth in students’ understanding of the context in which the second language is acquired.
- ⁴ For an excellent update on Willig’s research, see Rolstad, Mahoney, and Glass (2005).