
L2 Proficiency Development in a Two-way and a Developmental Bilingual Program

**Ester de Jong
University of Florida**

Abstract

This longitudinal, quantitative study examined the English oral and literacy development of U.S.-born English language learners (ELLs) who had attended a two-way immersion (TWI) or developmental bilingual (DBE) program since Kindergarten, as measured by a five-point rating scale and the Language Assessment Scale Reading/Writing. Descriptive statistical analysis showed native-like academic oral proficiency developed within two (comprehension) to four (production) years. Basic literacy skills reached fluency by the end of third grade, though 30% of the students were still limited in writing. T-tests indicated significant differences between TWI students and DBE students' literacy skills. The findings of the study point to the importance of considering plateau effects in second language production skills.

Introduction

Recent proposals in California, Arizona, Massachusetts, and Colorado to replace bilingual education with one-year English-only structured immersion programs for English language learners (ELLs) have reinforced a long-standing debate on the expected rate of second language proficiency development of English language learners (ELLs). However, in absence of longitudinal studies that examine second language proficiency development there is a gap in our understanding of what students (can) learn in one year (as reflected in these one-year policy proposals) and ultimate attainment in a second language. Even less is understood about such second language development within the context of bilingual programs for minority language speakers. This study aimed to address this gap by examining the development of English oral proficiency and literacy skills of native Spanish speakers in maintenance bilingual programs from a longitudinal perspective. After providing a summary of the "how long" debate and the outcomes of different bilingual program types, the article presents the study, its outcomes, and implications.

The “How Long” Question

The question how long it takes learners to become proficiency in a second language (L2) has long occupied the field of English as a second language and bilingual education. Since the early 1980s, researchers have tried to establish the length of time needed to acquire academic second language proficiency. This line of research builds on Cummins’ distinction between social and academic English proficiency (Cummins, 1980, 2000) and finds that it may take ELLs between four and seven years to reach grade-level norms on formal language/reading tests (Collier, 1987; Cummins, 1981; Hakuta, Butler, & Witt, 2000; Klesmer, 1994). These studies are concerned with ultimate academic achievement in English and do not include program-related variables, such as program type or length of program attendance. One can therefore not infer that these findings support the need for specialized programs (i.e., bilingual education or English as a Second Language programs) for English language learners that last seven years or more. They do, however, emphasize the need for continued support for ELLs to learn to negotiate learning through their second language (Cummins, 2000).

To what extent such support must come through a specialized program or in a standard curriculum classroom is open for debate (e.g., Platt, Harper, & Mendoza, 2003). How long should ELLs remain in a specialized program? In contrast to the previous research, this alternate “how long” question deals with the identification of minimum criteria to determine ELL students’ readiness for reclassification as fluent English proficient and hence fulltime placement in a standard curriculum classroom (De Avila, 1990; De George, 1988; Gandara & Merino, 1993; Liguanti, 2001; Rossell & Baker, 1988). The policy proposals cited above assume that after one year “children will understand enough English so that they will be better off in a grade level mainstream classroom than in a remedial classroom” (Rossell, 2000, p. 17). However, it takes ELLs longer than one year to meet even minimal criteria for placement in a standard curriculum classroom, regardless of program type (e.g., Fillmore, 1998; Goldberg, 1997; Krashen, 2000; Ramirez, 1992). Moreover, a remedial approach to the schooling of ELLs has long been criticized as inappropriate if the goal is to provide ELLs with equal access to comprehensive education (Brisk, 1998). Rossell’s comment is therefore less an argument to quickly place ELLs in standard curriculum classrooms than it is a reminder that program goals, design, and implementation must reflect high expectations for ELLs.

These two perspectives on the “how long” question have raised important issues regarding English proficiency. However, few studies have considered the development of second language skills between a minimum of one year and ultimate attainment as it occurs over time in different programmatic contexts. The purpose of this study was to longitudinally examine these patterns for ELLs.

L2 Proficiency Development in Bilingual Programs

With few exceptions (Ramirez, 1992; Thomas & Collier, 2002), ELL program evaluation studies have typically compared English as a Second Language and bilingual program outcomes after 2-3 years of program implementation (e.g., Greene, 1998; Rossell & Baker, 1996; Willig, 1985). Besides methodological flaws (e.g., August & Hakuta, 1997; Willig & Ramirez, 1993), these studies have been unable to provide

insights regarding trends in second language development over time in various ELL program types. There continues to be a need for longitudinal research that follows the same students for an extended period of time (August & Hakuta, 1997; Thomas & Collier, 1997).

The program evaluations' exclusive emphasis on comparing bilingual versus English-only approaches has also prevented a more nuanced discussion about the effectiveness of different bilingual education models. Generally speaking, a distinction can be made between maintenance bilingual programs and transitional bilingual programs. Maintenance bilingual models are considered more consistent with an educational approach that maximizes the social, cognitive, and academic benefits of bilingualism documented in the literature (Christian, 1994; Cloud, Genesee, & Hamayan, 2000; Cummins, 2000; Lindholm-Leary, 2001; Thomas & Collier, 1997; Valdés, 1997). Maintenance bilingual programs typically use both languages (at least 50% for the minority language) for the duration of the program (minimally K-5). There are two types of maintenance bilingual programs: developmental (or late-exit) bilingual programs for minority students (DBE) and two-way immersion (TWI) programs that target majority and minority students (Brisk, 1998; Ovando & Collier, 1998). In contrast to maintenance bilingual programs, transitional (early-exit) bilingual education (TBE) programs are short-term programs (2-3 years), in which the native language is temporarily used as a bridge to learning English. They have been criticized for being assimilationist and subtractive (Ovando & Collier, 1998; Spener, 1988).

While the positive effects of attending a good bilingual program have been well documented (for overviews see for example Collier, 1992; Crawford, 1997; Cummins, 1996; Lindholm-Leary, 2001), little is known about the effectiveness of different bilingual program types. Studies that compared TWI and TBE programs (Cazabon, Nicoladis, & Lambert, 1998; Lindholm-Leary, 2001; Thomas & Collier, 2002) or DBE and TBE programs (Medina & Escamilla, 1992; Thomas & Collier, 2002) have consistently reported that MBE students outperform their TBE peers on measures of native language oral language and/or reading. This outcome is predictable and fully consistent with the goals of each program type. Regarding English proficiency development, some studies found no differences between MBE and TBE programs (Ramirez, 1992; Medina & Escamilla, 1992) or presented mixed results. For instance, Lindholm-Leary (2001) who examined cross-sectional data for English oral language development for TBE and TWI program students, reported different outcomes depending on the instrument used. TBE students outperformed TWI students on one oral assessment scale, but had a lower passing rate than TWI students on a different oral language assessment. Thomas & Collier (2002) provide the strongest data regarding the strength of TWI. They examined different bilingual program types in one district, the Houston Independent School District (HISD). The HISD bilingual programs included substantial native language instruction for the first three years (PK-2) with an increase in English in third grade. Differences among bilingual program types emerged in 4th grade when students in the TBE program shifted to an all English-medium classroom and students in the DBE and TWI programs maintained 50% of their instruction in Spanish. The only difference between the DBE and TWI program as described by the HISD guidelines was the presence of dominant English speakers in the TWI program. Thomas & Collier found that students in the TWI program consistently outperformed comparable groups of students attending the transitional bilingual education program on the Stanford English reading, math, and language. In short, while the superiority of maintenance bilingual education programs

for native language development and maintenance is clear, mixed results are reported for achievement in the second language, English.

The increased popularity of TWI programs across the nation (Christian, Howard, & Loeb, 2000) has caused an interest in comparisons between TWI and transitional or developmental bilingual programs. The inclusion of majority language speakers, systematic opportunities for native and non-native speaker interaction to build language and cross-cultural skills, and high program status create contexts for academic and language learning (Christian, 1994; Lindholm-Leary, 2001; Thomas & Collier, 1997) more difficult to obtain for DBE or TBE programs. The quasi-longitudinal data in Thomas & Collier (2002) indicated higher performance for TBE students than for the students in the DBE programs. Although the study did not directly compare DBE and TWI programs, combining the latter finding with the ones cited above (TWI students outperformed TBE students) leads to the speculation that the outcomes for TWI and DBE programs may indeed be different (see also Collier, 1992).

These findings confirm the importance of considering the effectiveness of different bilingual models in developing English language proficiency skills longitudinally over time (August & Hakuta, 1997; Lindholm-Leary, 2001). Whereas the superiority of native language proficiency appears similar across maintenance bilingual education programs for language minority students, the variation across models emerges particularly when considering English achievement patterns. The purpose of this study was to examine the development of second language proficiency by English language learners in the context of bilingual education programs. The study is unique for two reasons. First, it takes a longitudinal view to the “how long” question by following students from Kindergarten through fifth grade. Second, it focuses on two maintenance bilingual programs, a two-way immersion and a developmental bilingual education program. The first two research questions examined at which point ELLs reach native-like English oral proficiency and literacy, respectively. The third question explored whether there were any differences between the two programs.

Methodology

The study took place in medium-sized school district in the Northeastern United States. The district enrolls close to 8,000 students. At the time of the study, it had three types of programs for ELLs: two-way immersion, developmental bilingual education, and an English as a Second Language program. The district was chosen because of its long-standing experience with ELLs (the first bilingual classroom was established in the late 1960s), its consistency in program philosophy and implementation, and access to a TWI and a DBE program within the district. Specifically, the study was concerned with three research questions:

When do ELLs in maintenance bilingual programs reach fluency in English oral comprehension and production?

When do ELLs in maintenance bilingual programs reach fluency in English reading and writing?

Are there differences between the TWI and the DBE program regarding English oral proficiency and English literacy skills?

To answer these questions, a longitudinal, quantitative study was designed and implemented.

Sample

For the purpose of this study, a sample of U.S.-born native Spanish-speaking ELLs who entered the DBE or TWI program in Kindergarten and who remained in these programs through 3rd, 4th, or 5th grade without interruptions (e.g., moving away and returning) was selected to remove the role of length of residence and country of birth as confounding variables. Hence, students who entered at later grade levels (Grade 1 and up) were excluded from the sample. Also, students with special needs and those students for whom more than two consecutive data points were missing were taken out of the sample. The TWI program is located in a school with a higher percentage of students on free/reduced lunch. The two samples are similar in terms of entering students' country of birth and gender (Table 1).

Table 1
Sample

	TWI	DBE
K-3 Cohort	55	51
K-4 Cohort	45	37
K-5 Cohort	26	19
Number of non-U.S. Born Students in sample	6	7
Gender	Male: 37%	Male: 37%
	Female: 63%	Female: 63%
School Percentage of students on free/reduced lunch	38%	27%
Number of Special education students (excluded from analysis)	11	9

The data were grouped into three longitudinal cohorts: a K-3 cohort, a K-4 cohort, and a K-5 cohort (Table 1; and below). The K-3 cohort includes scores of all students who entered in grade three and who stayed in the program through Grade 3 (i.e., it includes data from students who stayed longer in the program). There is therefore overlap between the K-3 and the K-4/5 cohorts. The K-4 cohort includes only the scores of students who entered in Kindergarten and who stayed through grade 4. It excludes students only attending K-3 but includes students who stayed through 5th grade.

Finally, the K-5 cohort includes only students who attended the program for their elementary school career and does not include data from students who attended only K-3 or K-4. Collapsing the cohorts increases the number per cohort. This step is justified because of the stability of program philosophy and implementation, and the stability of the teachers involved.

Instrumentation

Oral language proficiency. The [State] English Language Assessment- Oral [S]ELA-O is an English oral proficiency assessment instrument that was developed specifically for the state in which the district is located. It is a teacher-administered assessment that requires that teachers observe and rate students individually on a scale of 0 to 5 in oral comprehension and production. While the comprehension component consists of only one scale, the production component consists of the four subscales of grammar, pronunciation, vocabulary, and fluency. When it was developed in 1993-1995, inter-rater reliability for the [S]ELA-O was reported as between .74 and .77 for the comprehension component and between .71 and .85 for the production component. These levels of reliability are considered quite good for a teacher-administered assessment. All teachers in the district are trained to properly administer the [S]ELA-O. Teachers receive a day and a half of training their first year of administering the assessment; the following year they receive another half-day of follow-up training. The interpretation of the five levels is generally as follows. Students at Levels 0, 1 and 2 are in the process of mastering social language. Students at Level 3 are socially able to function in English and are beginning to function academically in English. Levels 4 and 5 indicate native-like academic English proficiency. The [S]ELA-O is administered in the Spring of each year to students in the bilingual program from Kindergarten through Grade 5 to all students who have been in the classroom for at least two months at the time of assessment.

Reading and writing proficiency. For the purpose of assessing English literacy skills, a widely used formal language tests, the Language Assessment Scale Reading/Writing (LAS; Duncan and DeAvila, 1988) was used. The LAS R/W is a standardized proficiency assessment that measures English reading and writing skills. It assesses vocabulary, fluency, reading comprehension, mechanics, and usage. Raw scores can be converted into a proficiency level (non-reader/writer, limited reader/writer, and fluent reader/writer). The LAS R/W uses different forms for different grade levels, e.g., Form 1 is used for Grades 2-3 and Form 2 is used for Grades 4-6. The test is widely used to document progress as well as to determine a student's readiness to exit specialized services. The LAS-R/W is administered in Grade 2 through 5 as a whole-class assessment. The district provides teachers with a scoring manual that includes detailed instructions and guidelines about the scoring of student responses.

Research Site

The TWI program aims to develop bilingualism and biliteracy, grade-level academic performance by 5th grade in Spanish and English, and cross-cultural competence. It enrolls 50% native English speakers and 50% native Spanish speakers. There are two teachers at each grade level, a Spanish and English classroom. The TWI program adheres to a one teacher-one language policy at each grade level, except for Kindergarten where the English teacher also teaches in Spanish to reinforce Spanish for all students. Initial literacy development occurs in the native language for both language groups. Formal English literacy instruction for ELLs does not start until second grade

through the English as a Second Language class and in third grade through the classroom teacher. Until second grade, the focus is on oral language development in English. In Kindergarten the native Spanish speakers receive all instruction in Spanish except for specials (music, art, and physical education). In grades 1 and 2 they spend about 70% of the time is spent in L1 (literacy and math) and 30% in English (social studies/science). As of third grade, all students receive 50% of their instruction in their native language and 50% in their second language.

The district's self-contained DBE program is also designed for native Spanish speakers with limited English proficiency and who are dominant in their native language. The ultimate goal of the program is to prepare its students to succeed in a standard curriculum classroom. The DBE program is a K-5 model in which students are not expected to exit until 4th or 5th grade (though some do). The distribution of the two languages in the program for students entering in Kindergarten is described in the district's program handbook approximately as follows: K-1 (L1: 85%, L2: 15%), Grade 2 (L1: 75%, L2: 25%), Grade 3 (L1: 60%-50%; L2: 40%-50%), Grades 4 and 5 (L1: 30%; L2: 70%). Initial literacy development starts in L1, Spanish. DBE teachers start formal English literacy towards the middle of second grade and increase the time spent in English Language Arts in third grade. Until second grade, the English as a Second Language component focuses on oral language development. The school in which the program is located has worked hard to create an inclusive, bilingual environment for all its students. Students are integrated for special areas (music, art, physical education) with standard curriculum students. Moreover, to avoid extended student segregation, the program started to integrate DBE and standard curriculum students for content area instruction (science, social studies). This project started in 1998 with 4th graders, continued in 1999 with 4th and 5th grade, and was further extended in 2000 to include third grade as well.

Both programs are designed as quality bilingual programs and are expected to follow the state and district guidelines for grade-level curriculum content in all subject areas. The two schools are also comparable in terms of access to bilingual special education services, Spanish Title I services, and a fulltime English as a Second Language teacher. In both programs, all teachers have bilingual or English as a Second Language certification and have comparable years of teaching experience across the grade levels. Main differences between the two programs include systematic access to and interaction with native English speakers for the TWI students in the content areas and the use of dual language models (one teacher) at all grade levels for DBE students. Additionally, the TWI program is clearly perceived as an enrichment program. The DBE program's status is more ambiguous as it is trying to distance itself from a transitional bilingual education philosophy, which dominated in the 1980s.

Program placement. The bilingual program student assignment process is part of a controlled choice program in the district. The TWI and the DBE program take students from the same "pool" of Spanish-speaking ELLs. Students are eligible for the TWI or the DBE program if they have been recommended for bilingual services. Such placement is determined by the score on the oral Language Assessment Scale in English and Spanish. Spanish-speaking students are eligible for enrollment in a bilingual program if they are Spanish dominant, i.e., their score is Level 3 or higher (out of a scale from 1 to 5, 5 being fluent native speaker) and they are limited proficient in English (their English score is below Level 4). After services have been determined and parents agree with a bilingual program placement, the parents indicate their preference of program

type (TWI or DBE). Due to the controlled choice process, they have placement guarantee for their child in the school/program attended by a sibling. Once the Kindergarten slots are filled through the sibling guarantee process, the remaining slots are assigned by lottery. In other words, both programs take in the same group of Spanish-speaking ELLs who have been assigned based on the same criteria. There is therefore no reason to assume significant differences in the two student populations entering each program in terms of their Spanish language proficiency.

Procedure

Data for this study were drawn from district-level data, which are annually collected to meet state reporting requirements. In the Spring of each year, the Office of Bilingual Education indicates a two-week timeframe to teachers during which they must administer the [S]ELA-O and the LAS, thus ensuring that these assessments are done at approximately the same time in all programs throughout the district. Each teacher submits individual student scores on a specific form to the Office of Bilingual Education. For the purpose of this study, copies of the teachers' reporting forms were obtained from the district for a period of six years (1995-2001) as well as information on each student's date of program entry and program type. This information was entered into a database. The database contained individual assessment data for each student for each year s/he remained in the program between 1995 and 2001 for the [S]ELA-O (K-5) and for the LAS R/W (Grades 2-5).

Based on these data, the following analyses were carried out for each research question. The first question focuses on the point at which English language learners who entered their program in Kindergarten reach native-like oral proficiency in English. To answer this question, mean scores were calculated for each cohort at each grade level for comprehension and production as well as the percentage of students at each of the oral assessment's five proficiency levels (0-5). T-test analysis was used to determine whether scores differed significantly from grade to grade for each cohort. To answer the second research question, when do ELLs reach fluency in reading and writing, the distribution of student standardized scores over the three literacy levels indicated in the LAS manual (non-reader/writer, limited reader/writer, and fluent reader/writer) was determined. T-test analysis was carried out to examine whether students' scores changed significantly from second to third and from fourth to fifth grade. Finally, the third research question regarding programmatic differences was examined using t-test analysis for oral comprehension and production scores and for standardized reading and writing scores. TWI and DBE cohorts were compared for each grade.

Results

At what point do TWI and DBE students reach native-like oral L2 fluency? The first research question considered the point at which ELLs reached fluency in oral English in the two bilingual programs. The results for comprehension and production are presented below. For each program, students' mean scores are presented first, followed by their distribution over the five levels distinguished by the [S]ELA-O (Level 0-5).

L2 Oral language: Comprehension

TWI Program Students

The mean scores for the oral language proficiency for the three cohorts show steady growth in oral language comprehension. A mean score of 4.0 or higher (indicating mastery of native-like academic English proficiency) is approached by the end of first grade, exceeded by the end of second grade (Table 2), and continues to increase through 5th grade. Growth from Kindergarten to first ($p < .01$) and from first to second grade is significant for each cohort ($p < .05$). In the K-5 cohort, scores from 4th to 5th grade also differed significantly ($p < .01$).

Table 2
[S]ELA-O L2 Oral Proficiency Mean Scores TWI Program

TWI PROGRAM						
Comprehension						
	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
TWI-K-3	2.4	3.8	4.2	4.1		
TWI-K-4	2.3	3.9	4.2	4.2	4.3	
TWI-K-5	2.2	3.8	4.3	4.2	4.3	5
Production						
	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
TWI-K-3	1.9	3.2	3.7	3.7		
TWI-K-4	1.7	3.2	3.7	3	4.3	
TWI-K-5	1.7	2.8	3.5	3.8	3.9	4.9

When considering the distribution of students over the [S]ELA-O's five proficiency levels (Level 0-5), it appears that the three cohorts show similar behaviors (Figures 1a-c). By the end of first grade, two-thirds to three-quarters of the TWI students are rated as native-like at a Level 4 or 5 in comprehension. Over 85% of the students have reached these levels by the end of second grade (K-3: 86%, K-4: 88%, K-5: 88%). By 4th grade, almost all students are rated at these levels for the K-4 and K-5 cohorts (96% and 93%, respectively).

Figure 1a
L2 Oral Language Proficiency: Comprehension K-3 TWI Cohort

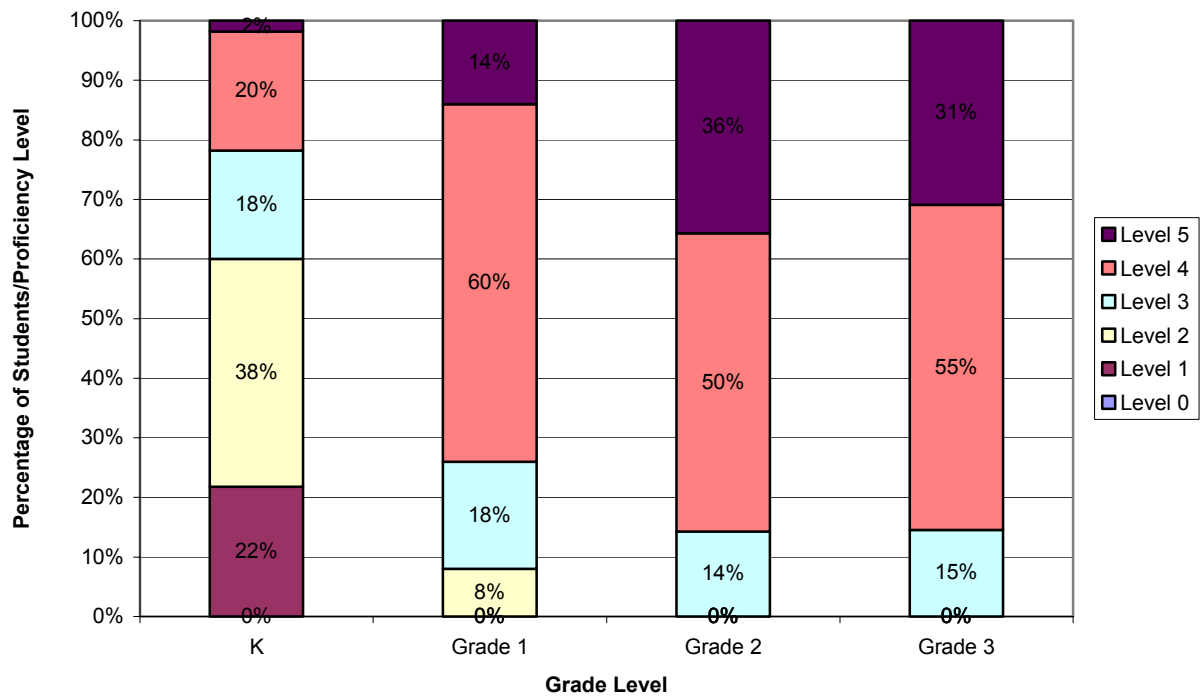


Figure 1b
L2 Oral Language Proficiency: Comprehension K-4 TWI Cohort

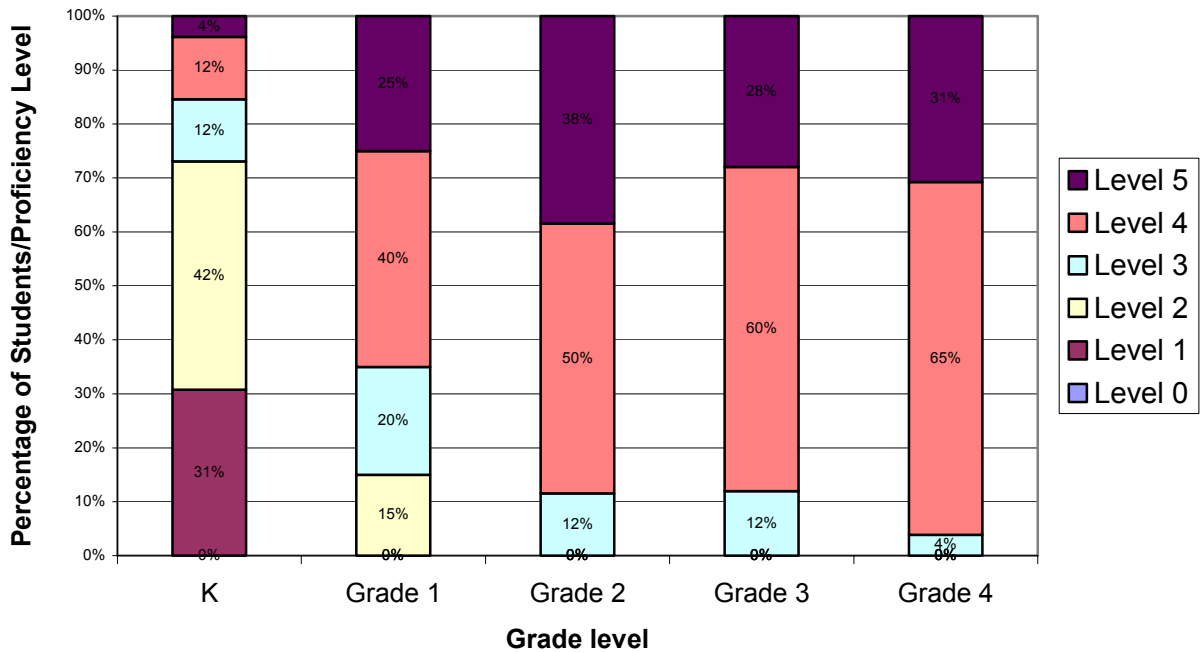
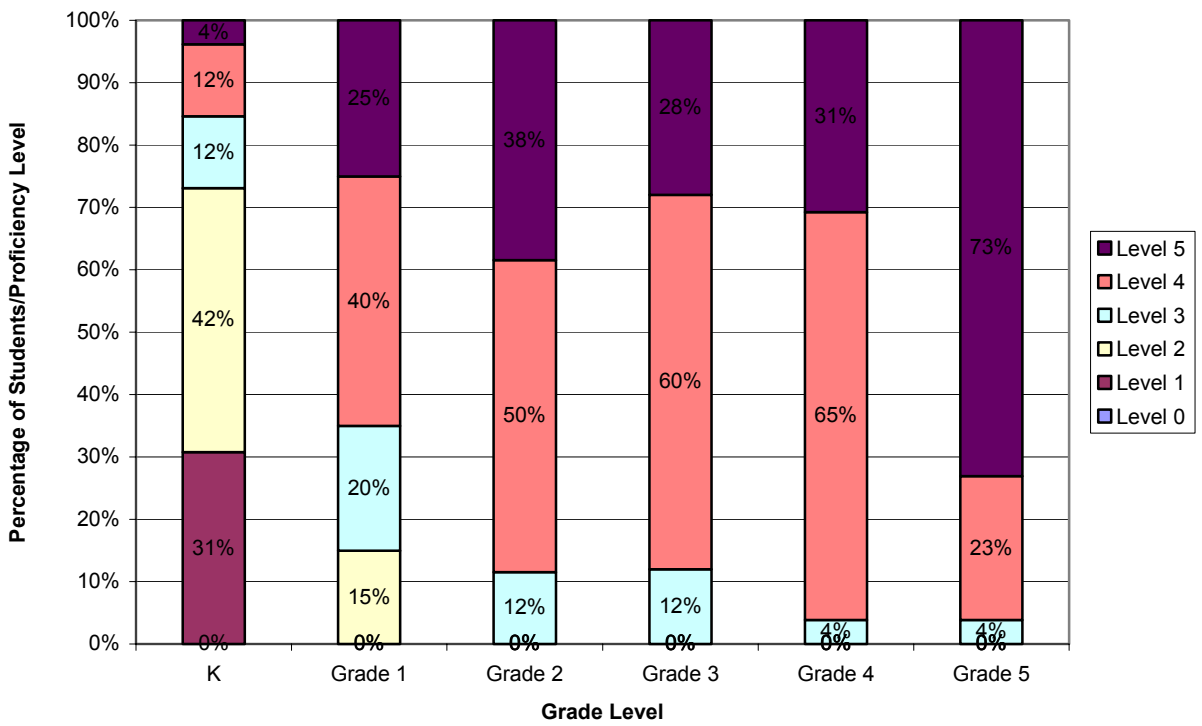


Figure 1c
L2 Oral Language Proficiency: Comprehension K-5 TWI Cohort



DBE Program Students

The DBE students' oral language development indicates beginning academic English skills in comprehension by the end of Kindergarten. By the end of first grade, their average score is 4.0 or higher (Table 3). For each cohort, growth from grade to grade is significant ($p < .01$). An exception is the K-5 cohort, where growth from first to second grade is significant at the .05 level and scores from 4th to 5th grade are not significantly different.

Table 3
[S]JELA-O L2 Oral Proficiency Mean Scores MBE Program

MBE PROGRAM

Comprehension

	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
MBE-K-3	3.3	4.1	4.6	4.2		
MBE-K-4	3.4	4	4.6	3.9	4.8	
MBE-K-5	3.4	3.8	4.6	4	4.8	4.8

Production

	K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5
MBE-K-3	2	3.3	3.7	3.6		
MBE-K-4	2.1	3.2	3.7	3.3	4	

MBE-K-5 2.2 3.2 3.8 3.4 3.8 4.1

Less than 20% of the students in the DBE program are rated as beginning English learners in comprehension at the end of Kindergarten (Figures 4a-c). About one quarter of the DBE students is at Level 3 and 70% of the students have reached native-like fluency by the end of first grade. As of second grade, almost all students are at Levels 4 and 5 for L2 oral language comprehension.

Figure 2a
L2 Oral Language Proficiency: Comprehension MBE K-3 Cohort

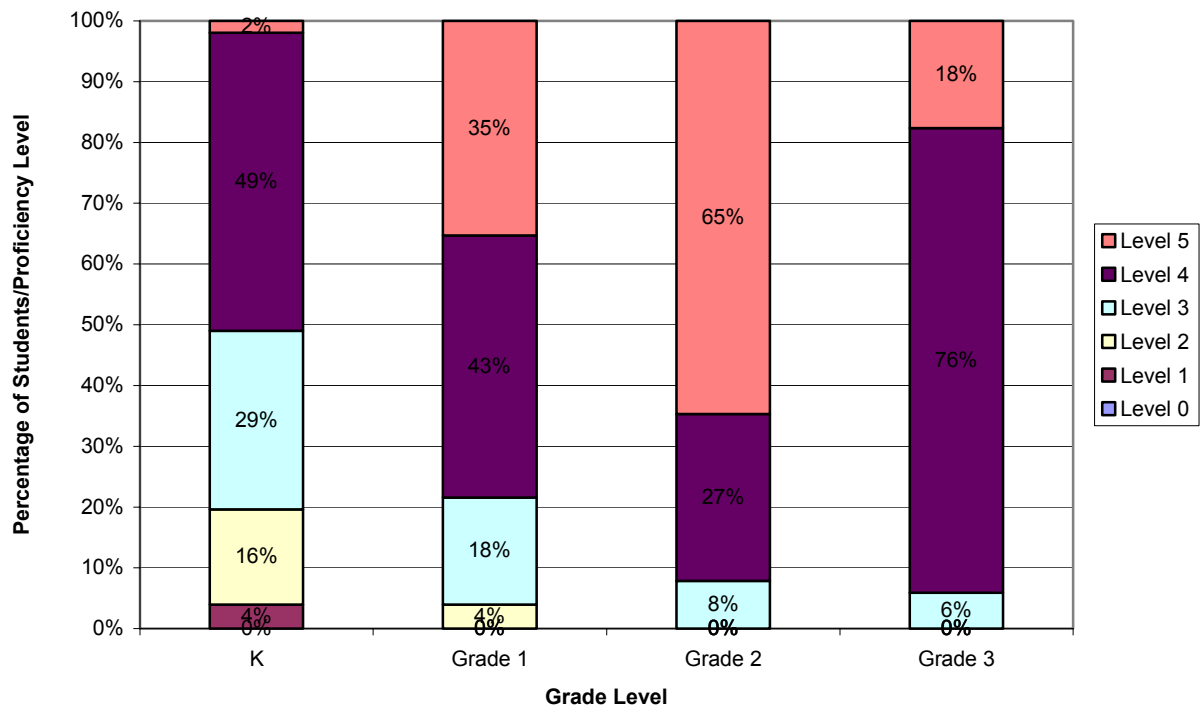


Figure 2b
L2 Oral Language Proficiency: Comprehension MBE K-4 Cohort

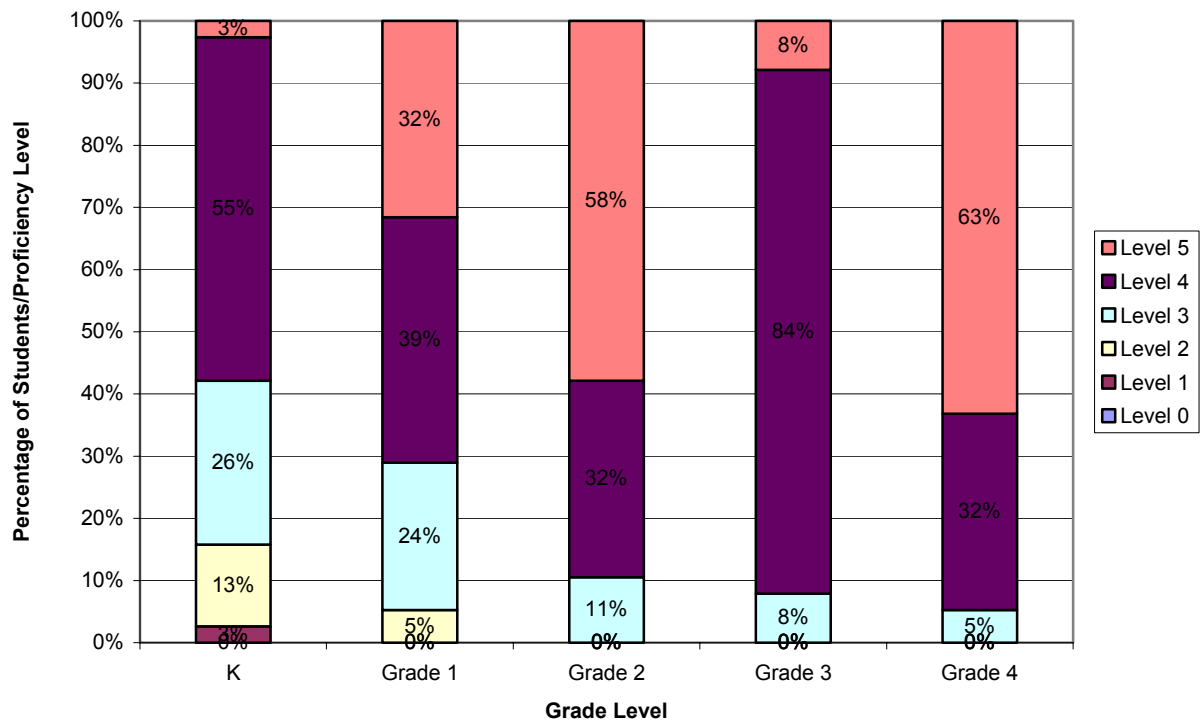
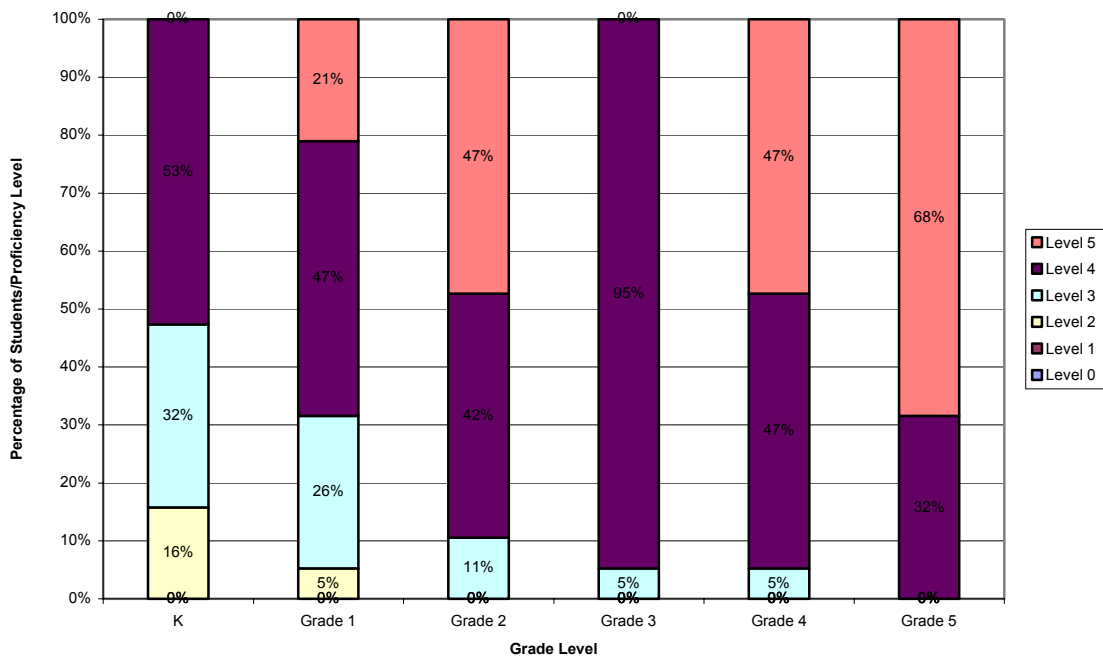


Figure 2c
L2 Oral Language Proficiency: Comprehension MBE K-5 Cohort



L2 Oral language: Production

TWI Program Students

It takes the TWI students until the end of 3rd grade to approach native-like proficiency in oral production (mean score of 4.0 and higher), although emergent academic proficiency (reflected by a mean score of 3.0 and higher) is reached or exceeded by second grade (Table 2). Significance was reached for growth from Kindergarten to first grade ($p < 0.01$) and from first to second grade ($p < .05$) for each cohort. Only the K-5 cohort showed further significant differences in scores for each subsequent grade level through fifth grade ($p < 0.01$).

The analysis of proficiency levels shows more variation in the early grades for English oral language production skills (Figure 3a-c). The majority of K-3 cohort and K-4 cohort students have reached Level 3 by the end of first grade. The K-5 cohort students take until third grade, probably because they started at much lower English production levels than the other two cohorts. Native-like academic English oral proficiency (Level 4 or 5) is reached by 71%-76% of the students in third grade and by almost all students (84% and 86% for the K-4 and K-5 cohorts, respectively) by the end of 4th grade.

Figure 3a
L2 Oral Language Proficiency: Production K-3 TWI Cohort

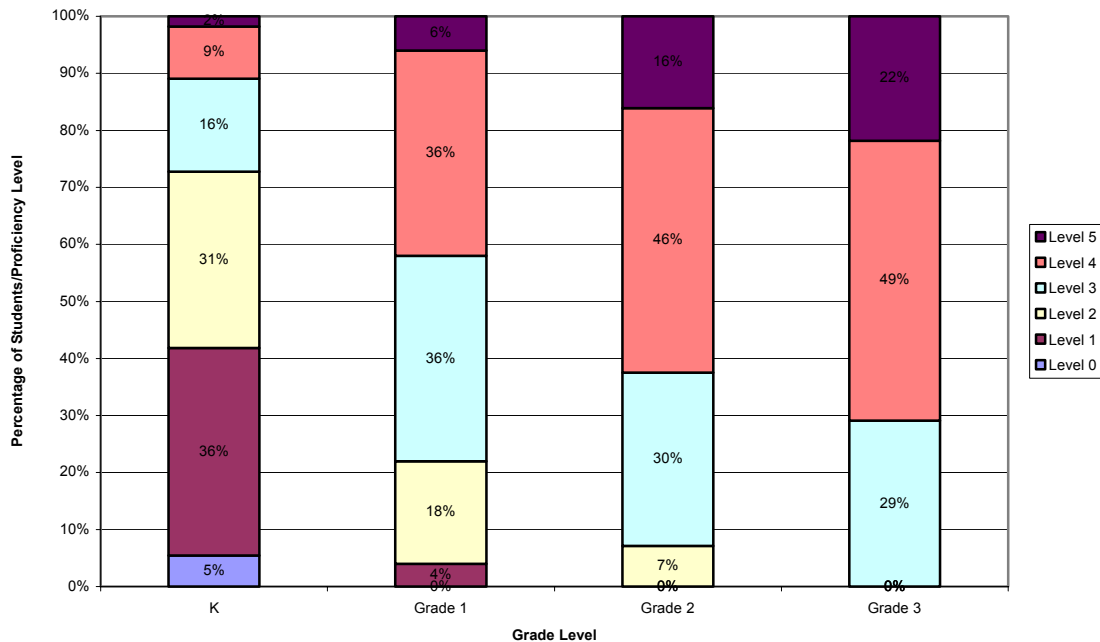


Table 3b
L2 Oral Language Proficiency: Production TWI K-4 Cohort

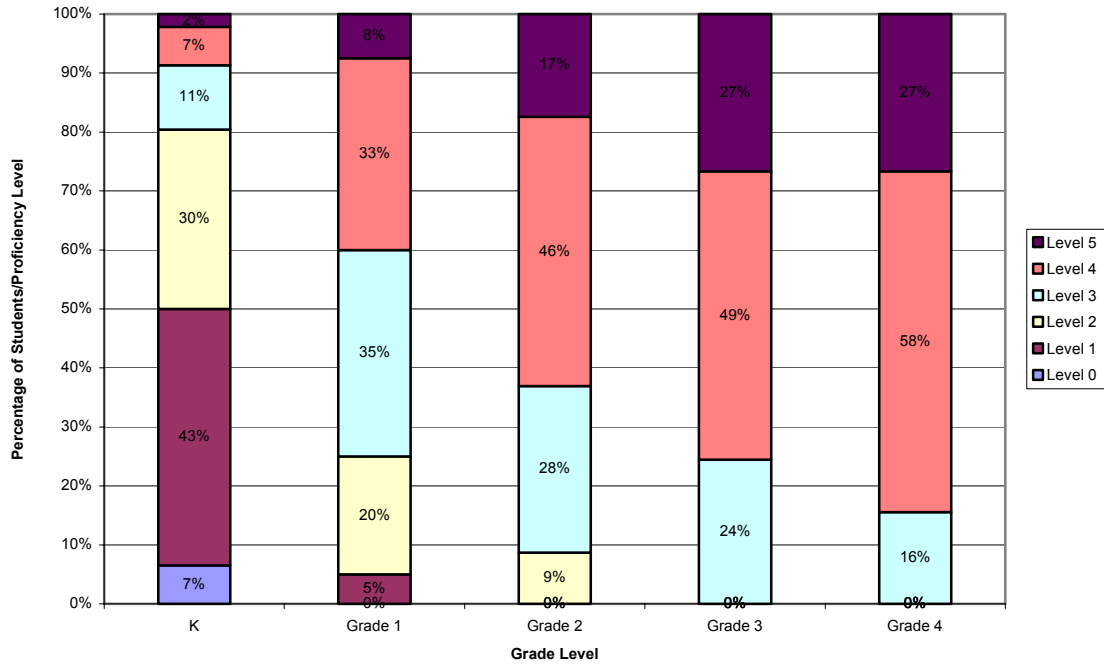
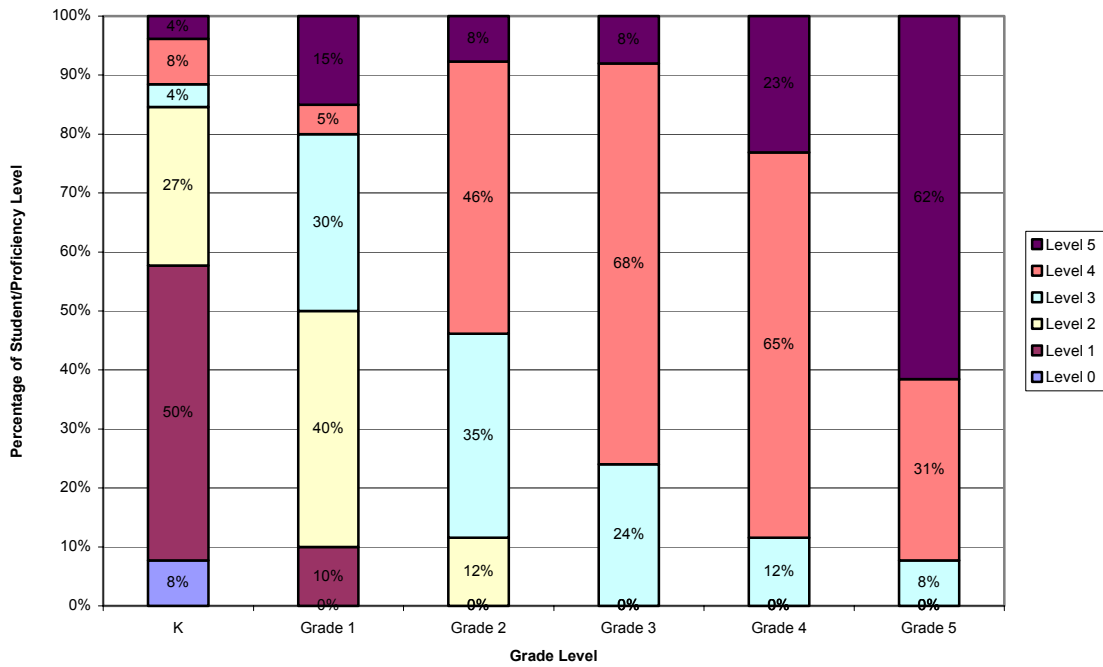


Figure 3c
L2 Oral Language Proficiency: Production K-5 TWI Cohort



DBE Program Students

The students in the DBE program reach an average score of 3 or higher by the end of first grade. Their mean score is between 3.0 and 4.0 between first grade and fourth grade. Native-like proficiency occurs by the end of 4th grade where their scores reach a mean of more than 4.0 (see Table 3). Growth is significant for all cohorts from Kindergarten to first grade ($p < 0.01$). Only for cohorts K-3 and K-4 is growth from first to second grade also significant ($p < .001$). No other significant differences between grade levels were found.

The distribution over the proficiency levels the DBE students' oral production remains just below native-like proficiency because many of them remain at level 3 in grades 2 and 3 (Figures 4a-c). Approximately 50% are at Level 3 by the end of first grade and 20%-35% of the students remain at this level through 4th and 5th grade. Seventy-two percent of the K-4 cohort and 58% of the K-5 cohort has reached Level 4 or higher by 4th grade. Almost all DBE program students are rated with grade-level academic English proficiency by the end of fifth grade (90%).

Figure 4a
L2 Oral Language Proficiency: Production MBE K-3 Cohort

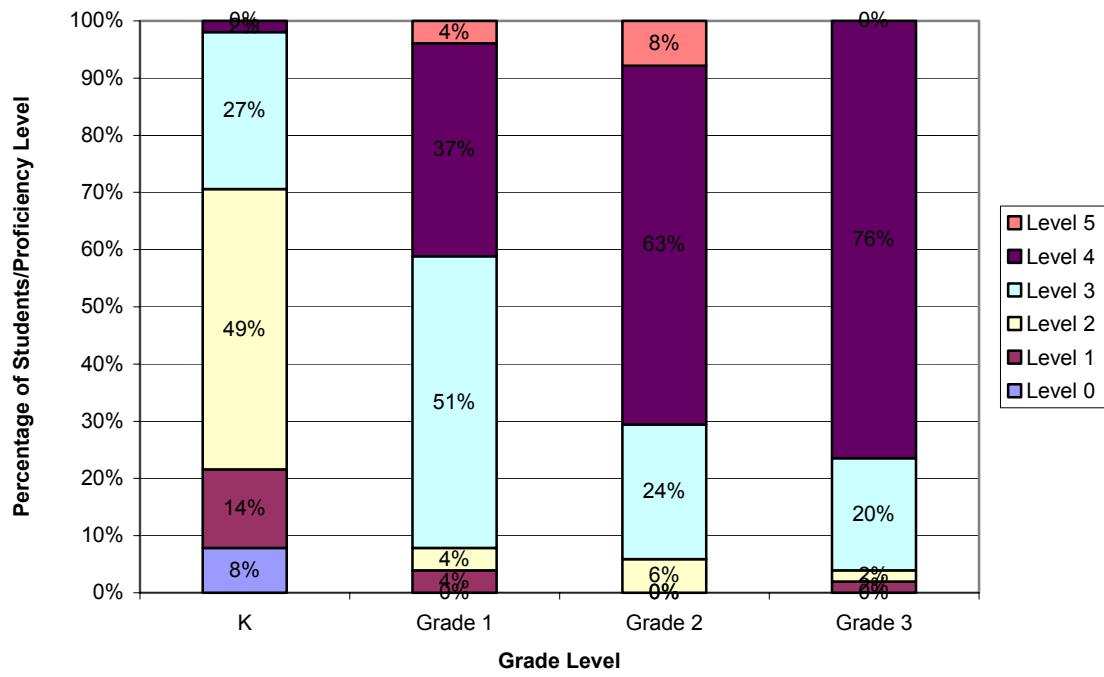


Figure 4b
L2 Oral Language Proficiency: Production MBE K-4 Cohort

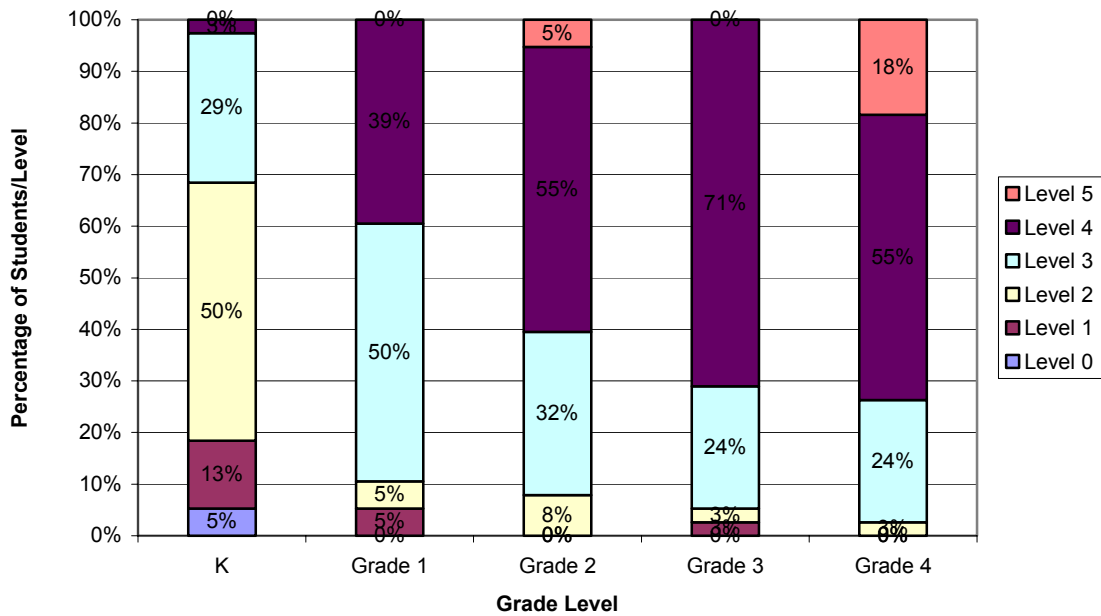
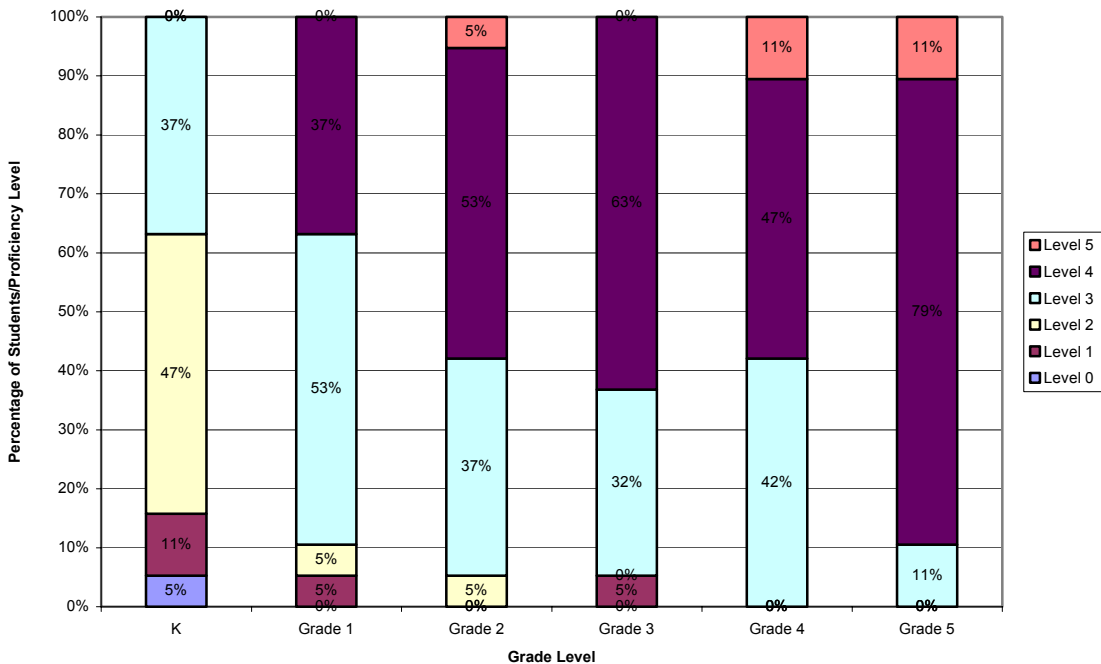


Figure 4c
Oral Language Proficiency: Production MBE K-5 Cohort



At what point do TWI and DBE students reach fluency in reading and writing?
 The second research question considered at which point ELLs reach fluency in English literacy in a TWI and DBE program. The results for reading and writing as measured by

the Language Assessment Scales are presented below for Grades 2-5. The LAS manual indicates three proficiency levels which were used for this analysis: non-reader or writer (Level 1), limited reader or writer (Level 2) and fluent reader or writer (Level 3) (Duncan & DeAvila, 1988).

L2 literacy: Reading

TWI Program

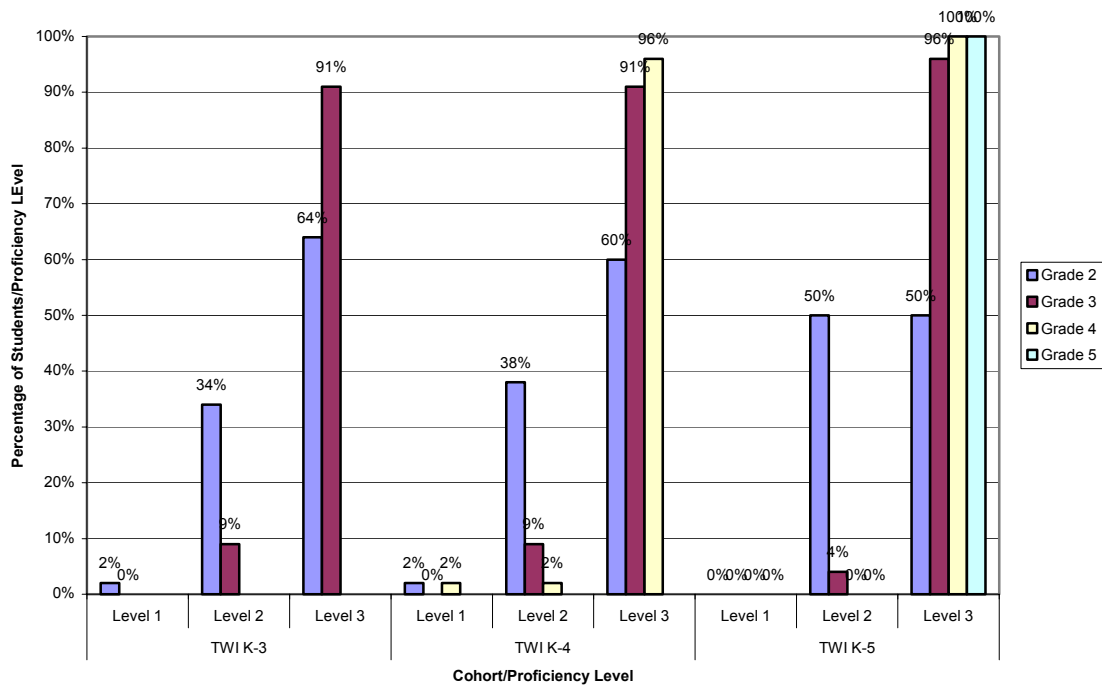
The TWI students do well on the reading portion of the Language Assessment Scale (Table 4). By the end of second grade, TWI students have reached a standardized score of 90 or more. T-test analysis shows that all scores from second to third grade represents significant growth for all three cohorts. Growth from 4th to 5th grade for the K-5 cohort is also significant ($p < .01$).

Table 4
LAS Reading and Writing: Average Standardized Scores TWI Program

TWI PROGRAM						
Reading			Writing			
	TWI K-3	TWI K-4	TWI K-5	TWI K-3	TWI K-4	TWI K-5
Grade 2	84	83	82	73	71	69
Grade 3	92	92	94	87	87	88
Grade 4	--	90	92	--	84	83
Grade 5	--	--	97	--	--	87

Analysis of the distribution of students across the three proficiency levels (non-reader, limited reader, and fluent reader) confirms that TWI students do well in reading. Several students have acquired limited reading fluency by the end of second grade (without formal English literacy instruction). For each cohort, a substantial shift towards fluency (Level 3) takes place from second to third grade. Ninety-one percent of the TWI students rates as a fluent reader as of the end of third grade and beyond (Figure 5).

Figure 5



DBE Program Students

The DBE program students are just below fluency by the end of second grade in reading, particularly the K-5 group. In Grades 3 and 4, these students' average standardized score is higher than 80. By 5th grade, the DBE students reach an average of a score of 95. The DBE students in each cohort make significant progress at each grade level ($p < .01$) (Table 5).

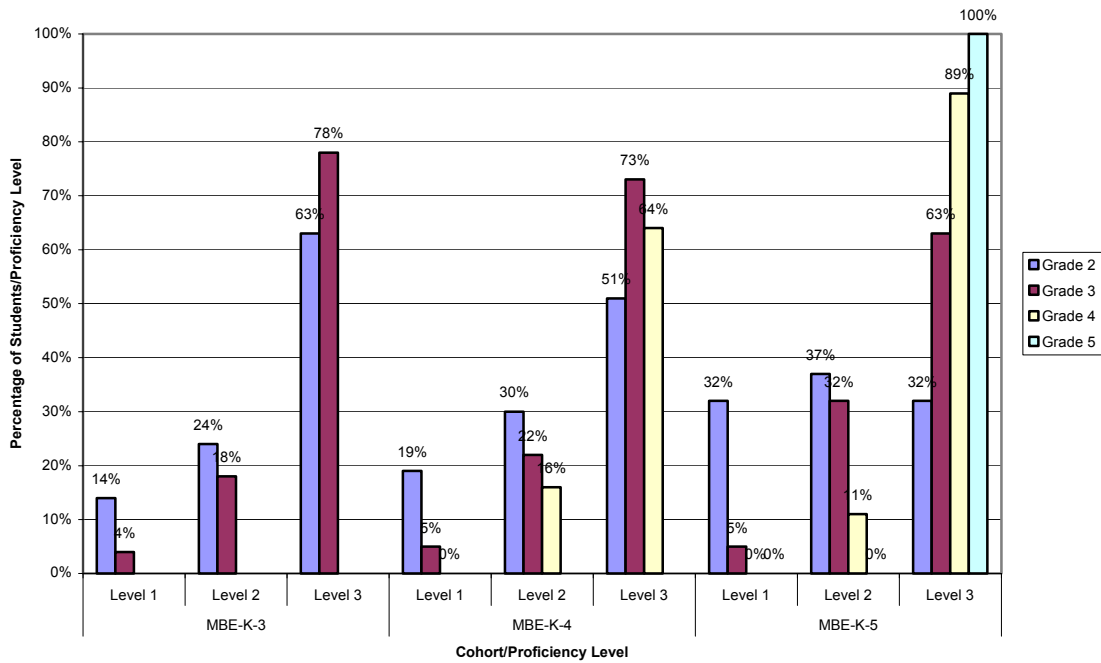
Table 5

MBE PROGRAM

	Reading			Writing		
	TWI K-3	TWI K-4	TWI K-5	TWI K-3	TWI K-4	TWI K-5
Grade 2	79	76	68	63	61	52
Grade 3	86	85	82	75	73	72
Grade 4	--	88	88	--	73	72
Grade 5	--	--	95	--	--	82

The analysis of the percentage of students scoring at the non-reader, limited reader, or fluent reader proficiency levels shows that the majority of students are score at a Level 3 (fluent reader) by the end of third grade (Figure 6). The K-4 cohort still shows close to a third of the students as limited readers by the end of 4th grade. The K-5 cohort shows 89% of the students at Level 3 by 4th grade and 100% of the students in grade 5.

Figure 6



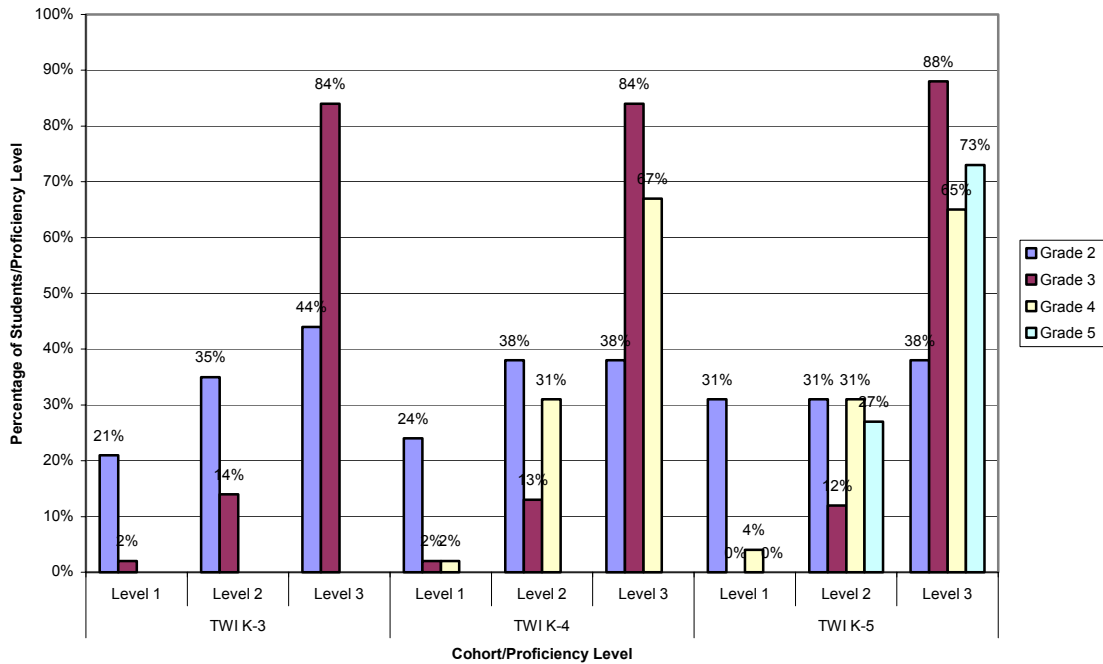
L2 literacy: Writing

TWI Program

Table 4 also illustrates that, in the absence of formal second language reading or writing instruction, TWI students are clearly not non-writers in English by the end of second grade nor are they at the bottom of the limited writing proficiency level (the average standardized score ranges from 60 to 79 across the three cohorts). Their average standardized score approaches the lower end of the fluent writer interval (which starts at 80). Fluency in writing is reached by the three cohorts by the end of third grade. Progress in writing is significant from second to third (all cohorts) and from 4th to 5th grade (K-3 and K-4 cohort). For K-5 cohort, the 4th to 5th growth approaches significance ($p=.078$).

Considering the three proficiency levels, it appears that TWI students are evenly distributed over the three levels by the end of second grade (Figure 7). Third grade TWI students score in the fluent writer proficiency level (84%, 84%, and 88% respectively for the K-3, K-4 and K-5 cohorts). The increased writing demand on the test can be seen at the 4th and 5th grade level where one-third of the students score at Level 2 and 67% (K-4 cohort) and 65% (K-5 cohort) score at Level 3. By 5th grade, 73% the K-5 cohort are considered fluent writers.

Figure 7
TWI L2 Writing by Cohort, Proficiency Level, and Grade Level

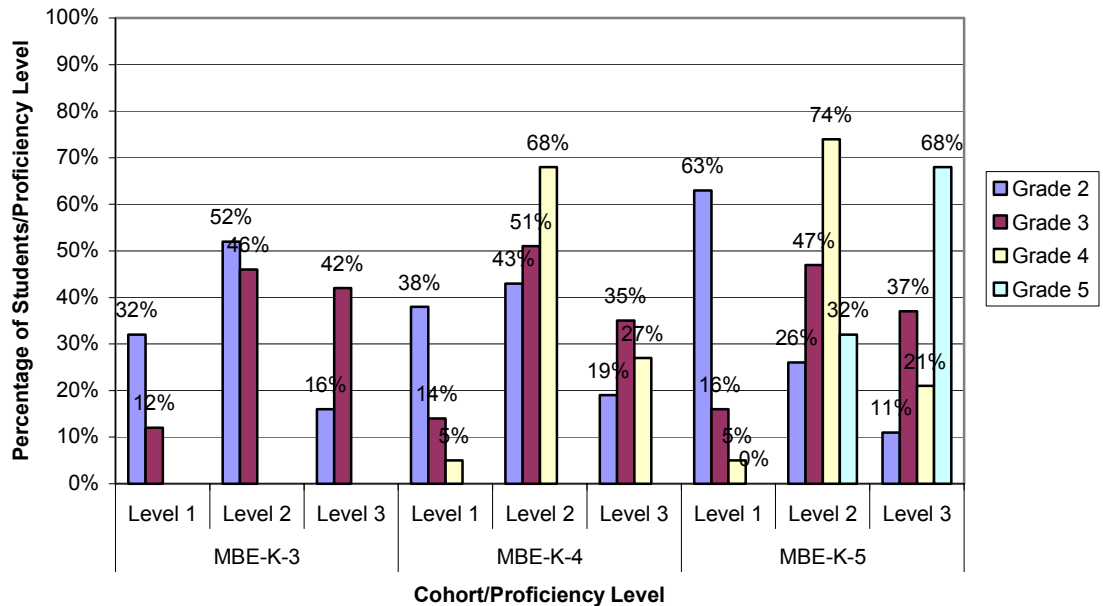


DBE Program

The DBE program students start at the lower end of the limited writing proficiency level in second grade (average standardized score is 60-79). By 5th grade, they reach an average standardized score of more than 80, indicating fluent writing proficiency (see Table 5). Tests of significance show that the mean scores are significant from second to third grade ($p < .01$). For the K-5 cohort, there is a significant difference between the 4th and the 5th grade scores ($p < .01$).

The distribution of the students across the three proficiency levels shows that a substantial number of students (about 50%) remain at Level 2 through third grade (Figure 8). By 4th grade, 27% of the K-4 cohort and none of the K-5 cohort have reached fluency. The DBE K-5 cohort students are rated as fluent writers by the end of 5th grade (68%).

Figure 8
MBE L2 Writing by Cohort, Grade, and Proficiency Level



Are there significant differences between the two programs? The third research question focused on whether there were statistically significant differences between the two programs. T-test analysis showed significant differences for oral language proficiency. Only for oral language comprehension at the Kindergarten level were DBE students' scores were significantly higher for each cohort ($p < .01$). This finding confirms the differences observed in Figures 2a-c and Figures 4a-c where the TWI program had many more students scoring at the beginning levels of oral proficiency (almost 60%-70% for TWI versus 20% or less for MBE).

The results are different for reading and writing. For reading, third grade TWI reading scores were significantly higher for all three cohorts ($p < .01$). For writing, the second and third grade scores for the TWI were significantly higher than those of the DBE students (all three cohorts), as were the fourth grade writing scores for the K-4 and K-5 cohorts. No statistically significant differences were found for the 5th grade writing scores.

Discussion

This quantitative study took a longitudinal view of the English proficiency development for cohorts of native Spanish speakers in two maintenance bilingual programs, a TWI and a DBE program. The three cohorts in each program showed similar patterns, which strengthens conclusions regarding the second language

development patterns and also shows that these bilingual programs have stable outcomes regardless of the student population entering the program. Students in both programs demonstrated significant growth from grade to grade in oral language development (particularly in the early grades, K-2) and in reading and writing (all grades). In short, participation in a bilingual program does not keep students from learning the English language (Cummins, 2000).

The first research question examined when ELLs in these two programs reached native-like oral language fluency (as measured by a rating scale which ranged from Level 0 to 5). Data analysis indicated that oral academic language output takes about four years to develop for students entering in Kindergarten, confirming the findings from the “how long” L2 attainment studies cited earlier. Second, although most students move towards native-like oral academic proficiency, several students remain at the intermediate proficiency level (Level 3) and students often remain at this level longer (Grades 1-3) before moving to native-like proficiency. This may point to a possible “plateau-effect” in oral language development in which students do not move beyond an advanced intermediate fluency in academic English (Scarcella, 2002). This phenomenon has been noted in French immersion program lasting as long as four years (Harley and Swain, 1984; cited in Long, 2003). While this plateau does not appear to persist for younger second language learners, it raises important questions regarding the impact it might have on student achievement. If teachers fail to stretch ELLs’ comprehensible output (Swain, 1995) and allow students to “get by” on basic oral fluency, it may ultimately hinder their access to dense, language-mediated curriculum and instructional practices (Corson, 2001).

The second research question considered when ELLs in these two programs reached the level of “fluent reader” or “fluent writer” on a formal language test, the Language Assessment Scale. The analysis showed students reaching fluency more quickly in reading than in writing. Ninety percent of the TWI students reach the test ceiling in reading by the end of third grade; 60%-80% of the DBE students do so. At the same time, it takes most students (65%-70% of TWI and 70% of DBE students) until 4th or 5th grade to reach fluency in writing. Second, the data analysis showed that 25% of the students were still rated as a limited writer by 5th grade. Contrary to claims that students will acquire sufficient second language proficiency within one year, this result makes it clear that it takes more than one year to become proficient in English, particularly in writing. While undoubtedly partially due to the test (students lose points if they make simple punctuation or spelling mistakes), this finding may also reflect the beginnings of a plateau effect similar to the one observed for oral language proficiency. The “intermediate” writing level allows many ELLs to survive elementary school as elementary teachers typically use many additional strategies to support the learning process. However, chances are that ELLs will become “long-term English learners,” if teachers do not effectively address their academic writing needs, i.e., students who receive passing grades but fail “when they try to pass high-stakes exit exams or when they take standardized test, their scores are low” (Freeman, Freeman and Mercuri, 2002, p. 5).

The third research question focused on whether there were significant differences between TWI and DBE program students. Significant differences were found in reading and writing. These differences emerged early on (Grade 2) in favor of the TWI program. A possible explanation for this finding is that the TWI students are more systematically and frequently exposed to academic literacy through their integration

content component. Although they are not introduced to formal English literacy until second grade (like DBE students), TWI students are engaged in science and social studies in English from first grade onwards in the “English” side classroom with native English speakers (unlike DBE students). As a result, ELLs in TWI programs regularly spend time in print-rich and literature-rich English language environments. For DBE teachers, it is more difficult to create such environments within one classroom (while also maintaining a rich Spanish language environment). Furthermore, TWI teachers must challenge the L1 (English) readers and writers in the integrated classroom. Reading and writing activities are naturally integrated into the science/social studies, while allowing ELLs to participate orally. ELLs in TWI programs have therefore more opportunities to take advantage of access to literacy in their second language and can develop their English literacy skills informally early on.

Limitations of the Study

A weakness of this study is its reliance on a standardized test to measure literacy. These tests do not adequately reflect the increasing academic reading and writing demands that ELLs encounter in the classroom. Moreover, it does not allow for the documentation of quantitative and qualitative growth that students make over time. While more authentic, the classroom-based rating scale applied by the teachers introduces a subjective component, which might have affected results.

Conclusions

This quantitative, longitudinal study confirms existing studies on how long it takes English language learners to acquire academic English proficiency. In response to the first two research questions, when do ELLs become proficient orally and in reading/writing, the study found a minimum of four years for academic English language development, at least for productive skills (speaking and writing). Moreover, the analysis indicated a possible “plateau” effect for productive language skills. This finding has important implications for ELL programs. Long-term persistence of an intermediate plateau may eventually prevent ELLs from having full access to academically complex learning environments that make more sophisticated demands on their language skills. Teachers must therefore learn to pay conscious attention to academic English proficiency, especially after initial fluency has been reached. As Gravelle (1996) points out, “Once pupils become more fluent in their use of English the immediate need for extra support is less obvious. It is at this stage that learners are often left to manage as best as they can and their achievement is affected” (p. 9). Bilingual and ESL teachers must provide ELLs with models of and practice with a varied, in-depth repertoire of linguistic resources to carry out a range of academic language functions (Gibbons, 2002). More detailed analyses of written and oral language samples will be necessary to develop a better understanding of this plateau effect, ways that teachers can scaffold productive academic language skills, as well as the consequences of such an effect when ELLs enter secondary school with its high reliance on language and literacy to mediate learning (Corson, 2001; Cummins, 2000).

The finding that there were significant differences between the TWI and the DBE program students’ performance in reading and writing points to the potential of making native/non-native speaker integrated (content-based) learning environment an integral part of any bilingual program. While an integrated program design need not be limited to

TWI programs (De Jong, 1996, 2001), such student integration rarely occurs systematically and for academic purposes within the context of developmental or even transitional bilingual education programs. This study suggests that, at least within the context of an additive bilingual program design, early (informal) exposure to English literacy may accelerate students' L2 development.

Exploring second language development patterns in different ELL program types links second language learning with effective programmatic contexts for ELLs. The issue of L2 expectancies plays a central role in discussions about the length of ELL program attendance as well as issues of program effectiveness. With increased inclusion of ELLs in accountability systems, an understanding of when English assessments can be interpreted with validity for ELLs who are still developing their second language becomes crucial. Finally, longitudinal studies that document second language proficiency development can inform bilingual practices, for example, by building links between bilingual and standard curriculum programs to support second language development.

References

- August, D., & Hakuta, K. (1997) (Eds.). *Improving schooling for language-minority children. A research agenda*. Washington, DC: National Academy Press.
- Brisk, M. E. (1998). *Bilingual education: From compensatory to quality schooling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cazabon, M. T., Nicoladis, E., & Lambert, W. E. (1998). *Becoming bilingual in the Amigos two-way immersion program*. Santa Cruz, CA: Center for Research on Education, Diversity, & Excellence.
- Christian, D. (1994). *Two-way immersion education: Students learning through two languages*. Santa Cruz, CA: National Center for Research on Cultural Diversity and Second Language Learning.
- Christian, D., Howard, E. R., & Loeb, M. (2000). Bilingualism for all: Two-way immersion education in the United States. *Theory into Practice*, 39 (4), 258-266.
- Cloud, N., Genesee, F., & Hamayan, E. (2000). *Dual language instruction. A handbook for enriched education*. Boston, MA: Heinle and Heinle.
- Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21, 4, pp. 617-641.
- Collier, V. P. (1992). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Research Journal*, 16 (1&2), pp.187-212.
- Corson, D. (2001). *Language diversity and education*. Mahwah, NJ: Lawrence Erlbaum.
- Crawford, J. (1997). *Best Evidence. Research foundations of the Bilingual Education Act*. Washington, DC: *National Clearinghouse on Bilingual Education*.

- Cummins, J. (1980). The entry and exit fallacy in bilingual education. *NABE Journal*, 4, 25-60.
- Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics*, 11, 2, pp. 132-149.
- Cummins, J. (1996). *Negotiating Identities. Education for empowerment in a diverse society*. Los Angeles, CA: *California Association for Bilingual Education*.
- Cummins, J. (2000). *Language, power and pedagogy. Bilingual children in the crossfire*. Clevedon: *Multilingual Matters*.
- De Avila, E. (1990). Assessment of language minority students: Political, technical, practical and moral imperatives. *Proceedings of the first research symposium on Limited English Proficient students*, OBEMLA.
- De George, G. P. (1988). Assessment and placement of language minority students: Procedures for mainstreaming. *FOCUS*, 3, Winter. NCBE Occasional Papers in Bilingual Education.
- De Jong, E.J. (1996). *Integrating language minority education in elementary schools*. Unpublished dissertation. Boston University.
- De Jong, E. J. (2001). From compensatory to quality. An example of change. *NABE News*, 24(3), p. 22-24.
- Duncan, S. E., & De Avila, E. A. (1988). *LAS: Language Assessment Scales. Reading/Writing*. Monterey, CA: MacMillan/McGraw-Hill.
- Fillmore, L. W. (1998). Supplemental declaration.
<http://www.humnet.ucla.edu/humnet/linguistics/people/grads/macswan/fillmor2.htm>
- Freeman, Y. S., Freeman, D. E., & Mercuri, S. (2001). *Closing the achievement gap. How to reach limited-formal-schooling and long-term English learners*. Portsmouth, NH: Heinemann.
- Gandara, P., & Merino, B. (1993). Measuring the outcomes of LEP programs: Test scores, exit rates, and other mythological data. *Educational Evaluation and Policy Analysis*, 15(3), 320-338.
- Gibbons, P. (2002). *Scaffolding language, scaffolding learning. Teaching second language learners in the mainstream classroom*. Portsmouth, NH: Heinemann.
- Goldberg, A. (1997). Follow-up study on the Bethlehem, PA, school district's english acquisition program. *Read Perspectives*, 4(1), 59-94.
- Gravelle, M. (1996). *Supporting bilingual learners in school*. Stoke-On-Trent, UK: Trentham Books.

- Greene, J. (1998). *A meta-analysis of the effectiveness of bilingual education*. Claremont, CA: Tomas Rivera Policy Institute.
- Hakuta, K., Butler, Y.G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* University of California Minority Research Institute. Policy Report 2000-1.
- Klesmer, H. (1994). Assessment and teacher perceptions in ESL student achievement. *English Quarterly*, 26, 3, 8-11.
- Krashen, S. (2000). Is one year/180 days enough? Retrieved April 5, 2002, from <http://www.languagebooks.com/2.0/articles/IsOneYear.180DaysEnough.html>
- Lindholm-Leary, K. J. (2001). *Dual language education*. Tonawanda, NY: Multilingual Matters.
- Linquanti, R. (2001). *The redesignation dilemma: Challenges and choices in fostering meaningful accountability for English language learners*. The University of California Linguistic Minority Research Institute Policy Report 2001-1. WestEd.
- Long, M.H. (2003). Stabilization and fossilization in interlanguage development. In: C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition*. (pp. 487-536). Malden, MA: Blackwell Publishers.
- Medina, M. Jr., & Escamilla, K. (1992). Evaluation of transitional and maintenance bilingual programs. *Urban Education*, 27(3), 263-290.
- Ovando, C. J., & Collier, V. P. (1998). *Bilingual and ESL classrooms. Teaching in multicultural contexts*. 2nd edition. Boston, MA: McGraw Hill.
- Platt, E., Harper, C., & Mendoza, M. B. (2003). Dueling philosophies: Inclusion or separation for Florida's English language learners? *TESOL Quarterly*, 37(1), 105-134.
- Ramirez, J. D. (1992). Executive summary. *Bilingual Research Journal*, 16, 1-62.
- Rossell, C. H. (2000). Different questions, different answers: A critique of the Hakuta, Butler, and Witt report, "How long does it take English learners to attain proficiency?" READ Abstract. Retrieved from the internet <http://www.ceousa.org/READ/hakuta.html>.
- Rossell, C. H., & Baker, K. (1988). Selecting and exiting students in bilingual education programs. *Journal of Law and Education*, 17, 4, 589-623.
- Rossell, C. H., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the teaching of English*, 30 (1), 7-74.
- Spener, D. (1988). Transitional bilingual education and the socialization of immigrants. *Harvard Education Review*, 58, 133-153.

- Scarcella, R. (2002). *What teachers need to know to be successful*. Paper presented at the OELA Summit on Language Acquisition, Washington, DC. Retrieved on November 19, 2003, at <http://www.ncela.gwu.edu/oela/summit/index.html>
- Swain, M. (1995). Three functions of output in second language learning. In: G. Cook, and B. Seidhofer (eds.). *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson*. (pp. 125-144). Oxford: Oxford University Press.
- Thomas, W. P., & Collier, V. (1997). *School effectiveness for language minority students*. Washington, DC. National Clearinghouse for Bilingual Education.
- Thomas, W. P., & Collier, V. (2002). *A national study on the school effectiveness for language minority students' long-term academic achievement. Final Report: Project 1.1*. Santa Cruz, CA: Center for Excellence on Education, Diversity, & Excellence. Retrieved on September 19, 2002, http://www.crede.uscs.edu/research/llaa/1.1_sc.html
- Valdés, G. (1997). Dual-language immersion programs: A cautionary note concerning the education of language-minority students. *Harvard Education Review*, 67, 391-429.
- Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. *Review of Educational Research*, 55 (3), 269-317.
- Willig, A., & Ramirez, D. J. (1993). The evaluation of bilingual education. In: M. Beatriz Arias & U. Casanova (Eds.). *Bilingual education: Politics, practice, and research*. (pp. 65-87). Ninety-second yearbook of the National Society for the Study of Education. The University of Chicago Press: Chicago, IL.