

The Crosslinguistic Role of Cognitive Academic Language Proficiency on Reading Growth in Spanish and English

Wilda Laija-Rodríguez
California State University, Northridge

Salvador Hector Ochoa
University of Texas–Pan American

Richard Parker
Texas A & M University

Abstract

The Office of Civil Rights (2001) and the No Child Left Behind Act of 2001 (Antunez, 2003) mandate that language proficiency and academic achievement be measured in order to provide equal educational opportunities to English language learners and have an accountability system for their language and academic growth. The Cognitive Academic Language Proficiency (CALP) construct is often cited in the literature as a milestone to second-language (L2) development and as having a significant relationship with academic achievement in the L2. Studies have typically explored the relationship between the primary language (L1) and the L2 separately with academic achievement. Language proficiency has often been viewed as a unitary construct without considering the interrelationship between L1 and L2 (Cummins, 2001). This study investigated the crosslinguistic relationship between the CALP in L1 and L2, as measured by the Woodcock–Muñoz Language Survey (WMLS) and reading growth, as measured by Curriculum Based Measurement Oral Reading Probes, with 77 second- and third-grade students in transitional bilingual classes. A significant, but weak relationship was found between Spanish CALP Broad Standard Score and English CALP Broad Standard Score with reading growth in Spanish and in English, respectively. The crosslinguistic relationship, as measured by the WMLS, and its relationship to reading growth is further discussed.

Introduction

The Office of Civil Rights (OCR) (2001) recommends that school districts evaluate English language learners' (ELLs) language proficiency to determine academic instruction based on the student's language proficiency in English in order to provide an equal educational opportunity. Moreover, the No Child Left Behind Act of 2001 (NCLB) has also included directives to ensure that ELLs are full participants in the learning process and gain academically from the educational system (Antunez, 2003). The NCLB delineates that ELLs' language proficiency and academic achievement should be assessed in order to have an accountability system for their language and academic growth (Antunez). The emphasis on ELLs improving second-language (L2) proficiency and academic achievement is due to the fact that ELLs have historically lagged behind in L2 development and academic achievement, especially in reading (Kindler, 2002; Laija & Ochoa, 1999; Sosa, 1990). The most popular reason given to explain the delay in reading performance is that ELLs do not have adequate language proficiency in the L2 or instruction in the primary language (L1) in order to understand and assimilate academic information (Cummins, 2001; Fitzgerald, 1995; Hudson & Smith, 2001; Thomas & Collier, 1997). It has been reported that it takes ELL students at least 5 to 10 years to attain grade-level norms in reading (Cummins, 2001; Thomas & Collier, 1997). The significant reading delays of ELLs make it imperative to understand the relationship between language proficiency and academic achievement, more specifically reading growth. The understanding of the relationship between language proficiency and reading growth is vital in providing educators with guidelines to assign appropriate instructional programs for ELLs, as mandated by NCLB and the OCR, early in their academic career in order to narrow the reading gap with their native English-speaking peers.

Various program evaluation studies in bilingual education and independent research studies on language proficiency and academic achievement have consistently showed a relationship between language and literacy skills gained in the L1 and literacy skills in a L2 (Cummins, 1978, 1984; Fitzgerald, 1995; Gottardo, 2002; Koda, 1994; Thomas & Collier, 1997; Yamashita, 2002). To help explain this relationship Cummins (1984) proposed that a common underlying proficiency (CUP) facilitates language transfer between students' L1 and L2 (Hudson & Smith, 2001). Cummins (1984) referred to this process as the developmental interdependence hypothesis, which proposes that the level of competence in English that an ELL attains is partially a function of the type of competence the child has developed in the L1. Thus reading ability in the L1 influences L2 reading competence.

Theoretical and Assessment Issues

While there are various theories in regard to L2 development, Cummins (1978) provides a framework to predict the academic achievement as a result of different forms of “bilingualism” or language proficiency. This framework is known as the *threshold hypothesis*. Cummins (1984) proposes that there are two thresholds referred to as Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP). Cummins (1984) describes BICS as the lower threshold, which involves having the ability to converse in peer-appropriate ways in everyday face-to-face situations. BICS requires a relatively low level of listening comprehension and expressive skills and it may be observed to develop between 1 to 2 years (Cummins, 1978; Thomas & Collier, 1997). CALP is described as the second threshold that involves having language necessary for cognitively demanding tasks and is required for accelerated cognitive as well as academic growth. CALP has been shown to develop in approximately 5 to 7 years (Cummins, 1978; Thomas & Collier, 1997). Cummins proposes that if children do not develop CALP in their L1, this will lead to cognitive and academic delays and possibly develop a state of “semilingualism,” where the student is limited in both L1 and L2 (Hudson & Smith, 2001).

With the threshold hypothesis, Cummins attempts to provide a framework to understand L2 development and proficiency and its effects on cognitive and academic growth. However, there is a lack of consensus in regards to a definition of language proficiency (Del Vecchio & Guerrero, 1995; Yamashita, 2002). There is also no common operational definition of what it means to be an ELL (Del Vecchio & Guerrero, 1995). This problem is evident when considering the threshold hypothesis. Del Vecchio and Guerrero indicate that the various “definitions of language proficiency share at least two critical features. First, each definition accommodates the four linguistic modalities: speaking, listening, reading, and writing. Second, each definition places language proficiency within a specific context” (p. 6). This coincides with the concept of CALP, where it includes skills in speaking, listening, reading, and writing. Cummins also adds that students are more apt to learn a L2 when it is more positively perceived, thus considering the context as well.

While Cummins’ threshold hypothesis involves the definition of BICS and CALP, Cummins’ definition of BICS and CALP appears to be broad and not well defined. Cummins rarely distinguishes between language and literacy, although he conceives language proficiency to include literacy skills (MacSwan, 2000). Other researchers have contributed to the concept of CALP. Wilen and van Maanen (1986) defined CALP as being composed of a broader vocabulary and proficiency in dealing with abstract linguistic messages, and Romaine (1995) indicated that CALP includes the development of literacy skills. Language proficiency, according to Cummins (1984), involves both BICS and CALP and BICS is a precursor to CALP.

Cummins' threshold hypothesis is not operationally defined or researched. It is such an abstract concept that it may be difficult to operationalize. As a result, many researchers have criticized his work in regards to the lack of definition and empirical support. In addition, not all researchers are in agreement with Cummins' threshold hypothesis. MacSwan (2000) argues that Cummins' threshold hypothesis should be abandoned due to lack of empirical and theoretical evidence. Other researchers, such as Baker (1998), indicate that developing the L1 does not necessarily promote higher academic achievement in the L2. Moreover, Baker indicates that instructional settings, such as the Structured English Immersion programs, produce better academic results for ELLs than bilingual programs that promote the development of L1.

While there are researchers who advocate against the threshold hypothesis for various reasons, much of the research exploring language proficiency and academic achievement in a L2 has found that the stronger the proficiency in the L1, the higher the academic achievement in the L2 (Fitzgerald, 1995; Thomas & Collier, 2002; Yamashita, 2002). More recently, Gottardo (2002) found that both reading skill and phonological awareness in the L1 were unique statistical predictors of reading in the child's L2. Other studies, such as García-Vázquez, Vázquez, López, and Ward (1997), directly relate high levels of CALP to higher achievement scores. "In spite of the growing recognition of the importance of academic language [CALP] in school achievement, little research has been conducted to define or describe academic language [CALP] in classroom contexts" (Fradd & Lee, 2001, p. 144) or in assessment measures.

While there are various language proficiency tests (e.g., Language Assessment Scales, Bilingual Syntax Measure, etc.), these tests report oral language proficiency as levels of proficiency ranging from *negligible* to *proficient*. In addition, these tests do not refer to language proficiency as levels of CALP nor do they assess literacy skills, which are believed to be a component of language proficiency (Del Vecchio & Guerrero, 1995). One of the few tests that has referred to language proficiency as levels of CALP is the Woodcock-Muñoz Language Survey (WMLS) (Woodcock & Muñoz-Sandoval, 1993); however, the WMLS has not operationalized each level of CALP besides just having a label ranging from *negligible* (level 1) to *proficient* (level 5).

There is a need to operationalize the different levels of language proficiency, as well as evaluate language proficiency measures. Different language proficiency tests have been shown to generate different language classifications (e.g., non-English speaking, limited English speaking and fully English proficient) for the same students (Ulibarri, Spencer, & Rivas, 1981) and classified native English speakers as limited or non-speakers (Pray, 2005). "If native English speakers do not receive scores in the 'native speaking ability' range on the assessment, the chances of an English language learner (ELL) achieving a score that accurately reflects his or her English proficiency is diminished" (Pray, p. 388). According to Pray in her validity study of three

language proficiency tests, the only test that resulted in classifying monolingual English students as proficient was the WMLS. While the WMLS is not perfect and still needs to further operationalize the levels of CALP, it appears to be an adequate test to measure language proficiency. According to Antunez (2003), 28 states use the WMLS to assess language proficiency of ELL students. Given the increased popularity of the WMLS, as a measure of language proficiency and as a measure of CALP, it is important that we understand how this particular instrument and its construct of CALP relates to student's reading growth as a precursor to academic achievement.

Studies on the relationship between language proficiency and academic achievement have typically explored the relationship of L1 and L2 separately with academic achievement. Koda (1994) indicates that a number of L2 acquisition studies have shown that various linguistic and metalinguistic elements are transferred from L1 to L2 production in both oral and written forms, thus implying that some reading skills acquired in one language can be applied to another language. Language proficiency has often been viewed as a unitary construct without considering the different development patterns of conversation and academic language acquisition in the L2 (Cummins, 2001). Therefore, when implementing L1 assessment, "the implicit assumption has frequently been that bilinguals can be assessed as though there were two separate monolingual proficiencies in their heads" (Cummins, 2001, p. 127). According to Cummins (2001), the strategy to view languages separately fails to take into account the close developmental relationship between L1 and L2 and underestimates the totality of an ELL's conceptual repertoire. Given that the WMLS is among the first to report language proficiency results as CALP and that it is the third most common language assessment measure in the United States (Antunez, 2003), it is important to explore the crosslinguistic factors that may exist when using this test. Moreover, Koda proposes that reading research should include language proficiency in both the L1 and L2 as "L2 reading is crosslinguistic in nature, involving at least two languages" (p. 5).

Assessing Reading Growth

Given the fact that ELLs are often significantly delayed in reading and it takes them longer than their monolingual peers to reach grade-level equivalency, it is imperative to focus on reading growth with this population. Since standardized tests were not designed to measure growth, the use of informal tests, such as curriculum based assessments, are recommended. A successful method to measure reading growth has been shown to be Curriculum-Based Measurement (CBM) (Shinn, 1989), which measures oral reading fluency (ORF) and reading accuracy. ORF is determined by how fast or how many words per minute a student reads and is part of reading decoding,

a strong predictor of reading comprehension in the early grades (Floyd, Gregg, & Keith, 2004). Reading growth is measured by ascertaining a student's ORF, which has been found to have a high correlation with reading proficiency including reading comprehension (Shinn). In many studies, ORF, as a measure of reading, has been shown to have high reliability in both Spanish and English reading. An unpublished study of ORF using Spanish-reading probes with over 90 primary-grade low readers reported a test-retest reliability coefficient of .90 (Baker, Kameenui, Simmons, & Stahl, 1994; Parker & Hasbrouck, 1998). Other studies using CBM ORF English-reading measures have resulted in reliability coefficients ranging from .82 to .97 (Shinn).

Purpose

It is well documented that skills in L1 are transferred to L2, yet, there are many questions still unanswered in regards to ELLs' language and academic development. While studies have focused on the relationship between reading in the L1 and L2 separately (Cummins, 2001), they have not addressed the crosslinguistic nature of CALP in L1 and L2 and its relationship to reading growth in L1 and/or L2 with individual students. Research that supports the importance of L1 development for academic achievement has been based on group data (Cziko, 1992). Individual students' reading growth has not been the unit of measurement in most studies. Individual students' reading progress has not been measured, nor has this been related to the construct of CALP. In addition, the interrelationship between L1 and L2 with reading growth has not been addressed (Koda, 1994).

This study will examine the interrelationship between L1 and L2, as measured by combining both CALP Broad Standard Score (B-SS) Spanish and CALP B-SS English based on the WMLS, as predictors of reading growth in Spanish and English, respectively. It will compare whether combining L1 and L2, as measured by the WMLS, serves as a better predictor than using L1 or L2 as individual predictors. This study's hypothesis is that the combination of CALP in L1 and L2, as measured by CALP B-SS on the WMLS, will be a better predictor of reading growth in Spanish and English, respectively. It is assumed that a better understanding and proficiency in the L1 and L2 will aid in reading growth, as it provides a basic phonological awareness of each language.

Method

Setting and Participants

The study was conducted in a school district in the southwest of the United States. The school district had a student enrollment of approximately 14,000 students. Seven percent of the pupils received bilingual or English as

a second language (ESL) education. The ethnic breakdown of the school district was 44% White, 24% African American, 31% Hispanic, and 1% other. Approximately 53% of the students enrolled were of low socioeconomic status.

The district had a transitional bilingual education program. The six teachers, who participated in the study, volunteered to participate in the study after being recruited by their bilingual education director and they were the teachers of the students involved in the study. These teachers were all female, were bilingual certified with a range of 1–24 year ($M = 11.6$) teaching experience (Their years of teaching experience were 1, 7, 8, 10, 20, and 24, respectively), and were based on two separate schools. The six teachers were interviewed in order to obtain instructional and demographic information. Based on interviews, five of the six teachers who participated in the study reported using primarily Spanish to explain lessons and one third-grade teacher indicated using both Spanish and English. Four of the six teachers indicated that they expected the children to speak both Spanish and English in the classroom, while the remaining two teachers, both second-grade teachers, expected their students to speak primarily Spanish in the classroom.

Eighty-seven students from six transitional bilingual classes in a Southwestern school district participated in the study after obtaining parent permission and student assent.

Of the 87 students in the six classrooms, 77 (55 second-grade and 22 third-grade) students contributed sufficient data to be included in this study. Of the 10 students excluded, 7 moved during the course of the study, 2 were special education students, and 1 student was a recent arrival from South America who was illiterate in both Spanish and English. The age of the children ranged from 7 to 10 years old. All of the students were ELLs based on the Language Assessment Scales (LAS) and of Hispanic descent. Most of the students were first generation Mexican American.

Determination of their language status was based on the LAS, which had already been administered by school personnel as required by the school district. Students with a score of 1 to 3 on the LAS were included in this study. These three scores are described as: *negligible language skills* (1), *very limited in language skills* (2), and *limited language skills* (3). Five of the six teachers reported that 91% to 100% of the students in their classes were receiving free or reduced lunch. The remaining teacher reported that 70% to 90% of the students in her class were receiving free or reduced lunch.

Instrumentation

The instruments used in this study were: the WMLS (Woodcock & Muñoz-Sandoval, 1993) in both English and Spanish, CBM probes from Spanish and English basal readers, a teacher questionnaire, and interview.

Woodcock-Muñoz Language Survey

The WMLS “uses Cummins’ (1984) BICS and CALP distinction as the theoretical foundation for item selection and overall test design” (Del Vecchio & Guerrero, 1995, p. 18). The WMLS is composed of four subtests: (a) Picture Vocabulary, (b) Verbal Analogies, (c) Letter–Word Identification, and (d) Dictation, which attempt to measure oral vocabulary, synonyms, reading, and spelling. The WMLS (Woodcock & Muñoz-Sandoval, 1993) is individually administered in English and a parallel form is administered in Spanish. One distinction between the WMLS and other language proficiency test is that the WMLS is described to provide CALP levels and standard scores for three clusters: oral language, reading and writing, and broad language, which combines the oral language and reading and writing cluster scores. According to the manual, there are five CALP levels that range from 1 to 5 and are defined as follows: *negligible language skills* (level 1), *very limited language skills* (level 2), *limited language skills* (level 3), *fluent* (level 4), and *advanced language skills* (level 5). However, there is no clear explanation on how the WMLS measures each level of CALP. Neither the manual nor research studies using the WMLS have addressed this question as of yet.

The WMLS CALP B-SS clusters in both English and Spanish were utilized to conduct statistical analyses. Norms for the WMLS test in Spanish and English are the same as those gathered from the 6,359 subjects for the standardization of the Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R ACH) (Woodcock & Mather, 1989). Norms for the Spanish WMLS were equated to the English WMLS norms. Internal consistency coefficients were reported and ranged from .87 for the Picture Vocabulary subtest to .96 for the broad English ability score (Del Vecchio & Guerrero, 1995).

Curriculum-Based Measurement

CBM ORF probes were used to assess reading fluency. The probes were selected and adapted from Spanish and English basal reading tests that were not used in the school district. These probes were designed based on basal series adopted in a Southwest state for use in second-grade classrooms. Narrative passages in Spanish and English were randomly selected and were reviewed for appropriateness of content, length, and difficulty. Administration of CBM probes was first done in Spanish then in English.

Teacher questionnaire and interview

A teacher questionnaire and interview were used to obtain information on student’s language, educational background, free lunch program status, and current classroom performance. This questionnaire and interview were conducted at the end of the study. The questionnaire was given to each teacher prior to the interview to assist with the interview in the attempt to

make it as short as possible due to teacher's limited time availability. Questions dealt with years of teaching experience and language usage for instruction, as well as to what language they expected students to use in the classroom.

Procedures

Training Procedures

Six doctoral school psychology students, one master-level bilingual education student, and two undergraduate students were recruited and trained to conduct the testing. Volunteers were trained using guidelines in the standardized WMLS test manuals to ensure standardization. Standardized procedures were also followed for the CBM ORF probe administration as suggested by Shinn (1989). All examiners passed "an administration proficiency check out" by the researchers in order to ensure that they were able to administer and record responses appropriately.

Testing Procedures

Permission to conduct this study was obtained from various sources. Consent forms were obtained from parents in order for the child to participate in the study. Thereafter, an assent form was read and explained to each student. Each student was asked to sign the assent form. The teachers also completed consent forms for the questionnaire and interview.

The WMLS in English and Spanish were administered at the beginning of the study. The examiners began with the WMLS in Spanish and administered the WMLS in English on a different day to minimize practice effects. Each administration of the WMLS lasted approximately 20 minutes.

To track reading growth, the ORF probes were administered once a month for 6 months. Spanish probes were administered first with English probes following. Each student was given instructions in the appropriate language prior to the administration. Each 2-minute probe was individually administered in Spanish and English to each participant once per month for 6 months. The examiner used standardized directions as described by Shinn (1989) and these directions were translated into Spanish. Directions were given in Spanish to ensure comprehension each time probe administration was conducted. Testing time for the two 2-minute probes was between 5 to 10 minutes for each monthly testing session. An average of the two 1-minute readings was taken from each set of reading probes in English and Spanish, which became the score used for ORF per month.

Data Analyses

To investigate whether the interrelationship between L1 and L2 had a stronger predictive relationship with reading growth than the individual language measured of CALP B-SS in L1 or CALP B-SS in L2, the combination of CALP B-SS in L1 and L2 based on the WMLS was used to serve as the predictor for two separate multiple regression analyses in which CBM ORF performance in Spanish and English, respectively, was the criterion variable. The criterion variables were the result of a time series design used with the CBM ORF probes that were the repeated measure and provided a simple linear regression representing reading growth. Reading growth per student, as measured by the standardized slope of CBM ORF performance in Spanish and English, respectively, across the 6-month period was the criterion variable. All tests were conducted using $\alpha = .05$.

To answer whether the combination of CALP B-SS Spanish and CALP B-SS English was a stronger predictor than individual levels of CALP B-SS Spanish or CALP B-SS English, two sets of simple linear regression analyses were also conducted. The two sets of simple linear regression analyses were performed to explore the predictive ability of the individual CALP B-SS in Spanish or English with the respective reading growth in English or Spanish.

Among the first steps in the data analysis was the calculation of ORF raw and standardized slopes, which represent the rate of reading growth across a 6-month period for each of the 77 students involved in the study. The raw and standardized slope coefficients for ORF in Spanish and English were computed using SPSS. This was done based on simple regression analyses using monthly ORF scores to represent scores on the regression line. The ORF score was the average number of the words read correctly per minute of the two 1-minute readings on each of the Spanish and English probe's oral reading passages. These scores were computed separately for the Spanish- and English-reading probes. The standardized slope, while not directly interpretable "as rate of growth or improvement," was used in the statistical analyses because of its standardized numerical value, which allows for direct comparison with other standardized scores. The linear regression derived per student revealed whether the slopes were significantly different than the mean, thus suggesting significant reading growth for each participating student. Results from individual linear regressions, indicating whether there was significant reading growth or not, were then used as aggregate data to conduct further analysis.

Results

Descriptive Information

The mean, standard deviation, kurtosis (peakedness of a distribution), and skewness (symmetry of a distribution) were calculated for predictor and criterion variables used in this study. As noted in Table 1, the mean score for CALP B-SS Spanish was higher ($M = 104$, $SD = 15$) than the one obtained in English ($M = 73$, $SD = 15$), indicating that most students were ELLs and therefore more proficient in their native language, Spanish.

The skewness and kurtosis were assessed using the z -distribution as suggested by Tabachnick and Fidell (1989) in which the obtained skewness and kurtosis values are then compared to zero using the z -distribution. Both the values for skewness and kurtosis are divided by their respective standard errors and an α level of .01 was used to test their significance. This equation was applied to the predictor variables (i.e., CALP B-SS Spanish and CALP B-SS English). It was not applied to the criterion variables (ORF standardized slope [SS] Spanish and ORF SS English) because these criterion variables were already standardized as indicated previously. Tabachnick and Fidell report that when a distribution is normal, the skewness and kurtosis are equal or close to zero. In addition, in a normal distribution kurtosis can range from -3 to +3 (Tabachnick & Fidell). Two variables (CALP B-SS Spanish and ORF Spanish) met normality based on skewness and kurtosis. CALP B-SS English and ORF English did not meet normality and were negatively skewed. This supports the fact that the students in this study were ELLs. When non-normality is

Table 1

Descriptive Data for Predictor and Criterion Variables (N = 77)

| Instrument | Minimum score | Maximum score | <i>M</i> | <i>SD</i> | Skewness | Kurtosis |
|---------------------------|---------------|---------------|----------|-----------|----------|----------|
| Predictor variables | | | | | | |
| WMLS Spanish CALP B-SS | 53 | 139 | 104 | 15 | -0.30 | 0.71 |
| WMLS English CALP B-SS | 15 | 103 | 73 | 15 | -0.87 | 2.46 |
| Criterion variables | | | | | | |
| ORF SS Spanish | -.68 | .98 | .45 | .43 | -0.93 | 0.02 |
| ORF SS English | -.36 | .99 | .69 | .26 | -1.65 | 3.00 |

Note. CALP B-SS = Cognitive Academic Language Proficiency Broad Standard Score; ORF SS = Oral Reading Fluency Standardized Slope.

found, Tabachnick and Fidell recommend the transformation of variables; however, the interpretation of results becomes more complex. Moreover, transforming some variables and not others makes comparisons among those variables questionable. For clarity of interpretation, it was therefore decided not to transform the variable of CALP B-SS English or ORF English, as these clearly described the student population used in this study. In addition, Stevens (1996) and Keppel (1982) note that even when distributions depart from normality, using more than 50 observations will approximate normality.

It was of interest to understand the frequency of combinations of CALP B-SS Spanish and CALP B-SS English levels to better understand the combination of languages on reading growth. Results indicate that in general, students were more proficient in Spanish than in English. Most students scored at CALP level 4 (*fluent*) in Spanish, whereas in English more were evenly spread out over level 2 (*very limited*), level 3 (*limited*), and level 4 (*fluent*). Only one student scored at the level 2 (*very limited*) in Spanish, whereas 37 students scored at level 1 (*negligible*) and level 2 (*very limited*) in English. The most common Spanish/English combinations were levels 4/2 (*fluent/very limited*) and 4/3 (*fluent/limited*), which comprised nearly 62% of the entire group of participants.

Reliability of Variables

The inter-rater reliability for the CBM ORF for this study was $r = .91$ for both Spanish and English, which is a respectable level of internal consistency. The inter-rater reliability was computed using the pair of initial CBM ORF scores (each obtained for a 1-minute sample), which were entered for each student and correlated. The means for both halves were nearly identical ($M = 56.6$), which is further evidence supporting instrument validity.

Using Pearson correlation, reliabilities of CBM ORF SS in Spanish and English were computed, which refer to “the consistency of examinees’ relative performances over repeated administrations of the same test or parallel forms of the test” (Crocker & Algina, 1986, p. 127). Reliabilities for reading growth in Spanish (.83) and English (.79) were found to be adequate.

Two multiple regression analyses were conducted, each including both CALP B-SS Spanish and CALP B-SS English as predictors for reading growth in Spanish and in English. These analyses were conducted to explore the interrelationship between both L1 and L2 and reading growth, thus addressing the first question: Does combining CALP B-SS Spanish and CALP B-SS English serve as a better predictor for reading growth in either Spanish or English? Both multiple regression analyses shown in Table 2 were statistically significant, but had small effect sizes and therefore weak relationship among variables. The two multiple regression analyses will be discussed separately.

Relationship Between CALP B-SS Spanish and English and ORF SS Spanish

A standard multiple regression was performed between ORF Spanish as the criterion variable and CALP B-SS Spanish and CALP B-SS English as predictor variables (see Table 2). R for this multiple regression was significantly different from zero, $F(2,76) = 5.69$, $p < .05$.

The combined predictors, CALP B-SS Spanish and CALP B-SS English, predicted 13% of the variance of reading growth (ORF SS) in Spanish. When examining the contribution of each individual variable, the bivariate correlation (r^{a2}), as suggested by Thompson (1992), was examined to understand the proportion of reading growth explained by the predictors. When looking at individual predictors' contribution in this analysis, CALP B-SS English had more ($r^{a2} = .07$), although weak ($\beta = -.38$, $p < .001$), predictive relationship with reading growth (ORF SS) in Spanish compared to CALP B-SS Spanish ($r^{a2} = .01$). This relationship between CALP English and ORF Spanish was negative for this sample, which means that as one of these variables increased the other decreased.

Table 2

Multiple Regression Analyses With CALP B-SS Spanish and CALP B-SS English as Predictor Variables and ORF SS Spanish and ORF SS English as Criterion Variables

| Analysis | Predictor variable | Criterion variable | r^{a2} | β | p |
|----------|--------------------|--------------------|----------|---------|------|
| 1** | CALP B-SS Spanish | ORF SS Spanish | .0121 | .27 | .03* |
| | CALP B-SS English | | .0729 | -.38 | .00* |
| 2*** | CALP B-SS Spanish | ORF SS English | .0961 | .37 | .00 |
| | CALP B-SS English | | .0001 | -.16 | .18 |

Note. CALP B-SS = Cognitive Academic Language Proficiency Broad Standard Score; ORF SS = Oral Reading Fluency Standardized Slope; r^{a2} = Bivariate correlation; β = Standardized Beta coefficient.

* $p < .05$. ** $F(2, 76) = 5.69$, $p < .05$, $R^2 = .13$, adjusted $R^2 = .11$. *** $R = .34$, $F(2, 76) = 4.79$, $p < .01$, $R^2 = .12$, adjusted $R^2 = .09$.

To address whether the combination of CALP in L1 and L2 served as a better predictor of reading growth in L1 and L2 than individual CALP measures in L1 or L2, simple linear regression analyses were conducted exploring: (a) the relationship between CALP B-SS Spanish and reading growth in Spanish and English, and (b) CALP B-SS English and reading growth in Spanish and English. Results from the simple linear regression analyses addressing reading growth in Spanish are shown in Table 3.

Simple linear regression analysis between CALP B-SS English and reading growth (ORF SS) in Spanish was significantly different from zero, $F(1, 76) = 5.99, p < .05$ and had a negative relationship. The relationship between CALP B-SS Spanish and reading growth in Spanish was positive, but not significant $p = .32$, similar to the multiple regression results.

Relationship Between CALP B-SS Spanish and English and ORF SS English

R for this multiple regression was significantly different from zero, $F(2, 76) = 4.79, p < .01$ (see Table 2), where the combined predictors CALP B-SS Spanish and CALP B-SS English predicted 12% of the variance of ORF-SS English. The bivariate correlation (r^{a2}) was examined to understand the proportion of reading growth (ORF SS) in English explained by each predictor. When looking at individual predictors' contribution in this analysis, CALP B-SS Spanish had more ($r^{a2} = .10$), although weak ($\beta = .37, p < .001$), predictive relationship with reading growth (ORF SS) in English compared to CALP B-SS English ($r^{a2} = .00$). There was a positive relationship between CALP Spanish and ORF English.

Table 3

Simple Regression Analyses Between Predictor Variables, CALP B-SS Spanish and CALP B-SS English, and Criterion Variable, ORF SS Spanish

| Predictor | Criterion | r^a | B | p |
|----------------------|----------------|-------|---------|------|
| CALP B-SS Spanish** | ORF SS Spanish | .11 | .110000 | .32 |
| CALP B-SS English*** | | -.27 | .000781 | .02* |

Note. CALP B-SS = Cognitive Academic Language Proficiency Broad Standard Score; ORF SS = Oral Reading Fluency Standardized Slope.

* $p < .05$. ** $R = .11, F(1, 76) = 0.99, p = .32, R^2 = .01$. *** $R = .27, F(1, 76) = 5.99, p < .05, R^2 = .07$.

^aZero-order correlation.

Given individual predictor contribution displayed in Table 4, however, only CALP B-SS Spanish resulted in a significantly different than zero correlation ($p < .05$) but weak relationship ($R^2 = .09$) with reading growth (ORF SS) in English. The relationship between CALP B-SS English and the reading growth (ORF SS) in English was not different than zero.

Overall, both multiple regression analyses in which CALP B-SS Spanish and English were predictors for reading growth in Spanish and in English respectively showed significantly different from zero results; however, the relationships were weak. Multiple regression results were similar to the results of the simple linear regression analyses, although their values were different. Similar results on both the multiple regressions and simple linear regressions conducted separately indicate that the combination of CALP B-SS in Spanish and CALP B-SS in English did not serve as a better predictor of reading growth as was anticipated.

Discussion and Conclusion

This study focused on the interrelationship between CALP B-SS Spanish and CALP B-SS English as measured by the WMLS, and reading growth in the L1 and L2, respectively, based on CBM ORF probes. This study addressed whether the crosslinguistic nature of L2 acquisition, as measured by combining both CALP B-SS Spanish and CALP B-SS English, had a stronger predictive relationship with reading growth in Spanish and English, respectively, than individual language measures of CALP Spanish or CALP English. This study's

Table 4

Simple Regression Analyses Between Predictor Variables, CALP B-SS Spanish and CALP B-SS English, and Criterion Variable, ORF SS English

| Predictor | Criterion | r^a | B | p |
|----------------------|----------------|-------|--------|------|
| CALP B-SS Spanish** | ORF SS English | .31 | .00534 | .01* |
| CALP B-SS English*** | | -.01 | .00022 | .92 |

Note. CALP B-SS = Cognitive Academic Language Proficiency Broad Standard Score; ORF SS = Oral Reading Fluency Standardized Slope.

* $p < .05$. ** $R = .31$, $F(1, 76) = 7.70$, $p < .05$, $R^2 = .09$. *** $R = .01$, $F(1, 76) = 0.01$, $p = .92$, $R^2 = .00$.

^aZero-order correlation.

hypothesis was that the combination of CALP in L1 and L2 as measured by CALP B-SS in English and Spanish by the WMLS would be a better predictor of reading growth in Spanish and English, respectively.

CALP B-SS Spanish and CALP B-SS English as Predictors of ORF Spanish

A significant ($p < .01$), but weak ($R^2 = .13$) relationship was found for the combined predictors CALP B-SS Spanish and CALP B-SS English with reading growth in Spanish. In this weak multiple regression, CALP B-SS English had the more ($r^2 = .07$) predictive contribution, although negative, to reading growth in Spanish, compared to CALP B-SS Spanish ($r^2 = .01$). This indicates that there was an inverse relationship between CALP B-SS English and reading growth in Spanish for this sample. That is, as one increases, the other decreases. In this case, it appeared that when CALP B-SS English was higher, reading growth in Spanish was lower. This may be explained due to the practice in transitional bilingual programs where instruction in Spanish is reduced, thus affecting the growth in Spanish reading as student's language proficiency in English is increased. This produces a lack of growth or "language loss" in Spanish.

Contrary to expectations, CALP B-SS Spanish failed to significantly predict reading growth in Spanish. Again, this was similar to the multiple regression results, which showed that CALP Spanish did not have a significant predictive relationship with reading growth in Spanish. The most likely reason for these results is that the lack of continued reading growth in Spanish limited the range of scores obtained, thus affecting this variable's ability to capture differences in reading growth in Spanish. Limited growth in Spanish is typical with transitional bilingual education programs, given that the goal of these types of programs is to develop the L2 and not to continue developing the L1. For this sample, it appears that this is a feasible interpretation.

CALP B-SS Spanish and CALP B-SS English as Predictors of ORF English

A significant ($p < .01$), but weak ($R^2 = .12$) statistical relationship was found for the multiple regression for CALP B-SS Spanish and CALP B-SS English with reading growth in English. CALP B-SS Spanish was a better predictor of reading growth in English than CALP B-SS English, and CALP B-SS English did not play a significant role as a predictor of reading growth in English. Combining both CALP B-SS Spanish and CALP B-SS English to investigate whether this combination would serve as a better predictor of reading growth in English did not prove fruitful. It may be because CALP B-SS Spanish and CALP B-SS English shared much variance and thus did not add much to the prediction when combined. Given that the WMLS Spanish is an equated translation of the WMLS English version (Woodcock & Muñoz-

Sandoval, 1993), this appears plausible. Another possible reason is that in fact there is a crosslinguistic relationship between L1 and L2 and therefore the shared variance is to be expected, given Cummins' explanation of CUP.

The simple regression analyses also indicated that CALP B-SS Spanish had a significant ($p < .01$), but weak relationship ($R^2 = .09$) with reading growth in English. This was consistent with the multiple regression results. Although weak, the significant positive relationship between CALP B-SS Spanish and reading growth in English is consistent with Cummins' (1978) developmental interdependence hypothesis as well as with Thomas and Collier's (1997, 2002) research. This is also consistent with other research on L2 development and bilingual education, which shows that there is a positive relationship between L1 development and literacy skills in English (Calero-Breckheimer & Goetz, 1993; Fitzgerald, 1995; Goldenberg, 1996; Gottardo, 2002; Koda, 1994; Ramírez, 1992).

For this sample, no significant relationship was found between CALP B-SS English and reading growth in English. A possible reason for this is the fact that only 11 students scored at a moderate (CALP level 3) to proficient (CALP level 5) levels of CALP English, thus, this lack of a significant relationship could be due to poor development of CALP in English for the majority of the sample, which results in a limited variance and limited predictive ability for this variable. A larger sample of English proficient students would make it easier to evaluate whether the limited variance is the reason for these results.

To summarize, combining CALP B-SS Spanish and CALP B-SS English as measured by the WMLS, to investigate the interrelationship between L1 and L2 and reading in Spanish and English respectively did not prove fruitful for this sample of second and third grade ELL students. Results did not show that there is benefit in combining WMLS CALP B-SS results in L1 and L2 to predict reading growth. It may be that in fact there is a crosslinguistic relationship between L1 and L2 and therefore the shared variance reduces predictive ability. If this is so, this adds to the literature supporting Cummins' CUP concept of the threshold hypothesis. For this sample, the best predictor of L2 reading growth was primary language development. This result is consistent with other research, which supports the theory that students with higher language proficiency in their L1 will achieve higher L2 proficiency. This adds to the literature and research supporting language and literacy skills in L1 as contributors to L2 language and reading development. The weak relationship, however, makes it necessary for researchers to expand on this theory and explore what other variables besides L1 language proficiency and literacy better predict reading in L2. It is important to begin to incorporate reading development theory into research and how this may or may not be similar to L2 readers. In addition, research needs to be conducted to further investigate the crosslinguistic nature of language acquisition and its relationship to academic achievement and reading growth. Future studies

need to address limitations in regards to consistency of instruction, robustness of variables, adequacy of language measures to address the crosslinguistic nature of second language acquisition and CALP, and further define and operationalize the construct of CALP. Research should also expand the sample to different ages and grades, where the crosslinguistic nature of language acquisition may be more important.

References

- Antunez, B. (2003). *Assessing English language learners in the great city schools*. Washington, DC: Council of the Great City Schools.
- Baker, K. (1998). Structured English immersion. *Phi Delta Kappan*, 80(3), 199–204.
- Baker, S. K., Kameenui, E. J., Simmons, D. C., & Stahl, S. A. (1994). Beginning reading: Educational tools for diverse learners. *School Psychology Review*, 23(3), 372–391.
- Calero-Breckheimer, G. S., & Goetz, E. T. (1993). Reading strategies of biliterate children for English and Spanish texts. *Reading Psychology: An International Quarterly*, 14, 177–204.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Cummins, J. (1978). Cognitive factors associated with attainment of intermediate levels of bilingual skills. *The Modern Language Journal*, 31 (1–2), 3–12.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. San Diego, CA: College Hill Press, Inc.
- Cummins, J. (2001). Assessment and intervention with culturally and linguistically diverse learners. In S. R. Hurley & J. V. Tinajero (Eds.), *Literacy assessment of second language learners*. Boston: Allyn and Bacon.
- Cziko, G. A. (1992). The evaluation of bilingual education: From necessity and probability to possibility. *Educational Researcher*, 21(2), 10–15.
- Del Vecchio, A., & Guerrero, M. (1995). *Handbook of English language proficiency tests*. Retrieved April 5, 2006, from <http://www.ncela.gwu.edu/pubs/eacwest/elptests.htm>
- Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research*, 65(2), 145–190.
- Floyd, R. G., Gregg, N., & Keith, T. (2004). *Explaining reading comprehension across childhood adolescence and early adulthood is somewhat simple*. Retrieved March 10, 2006, from <http://www.woodcock-munoz-foundation.org/floyd2004b.pdf>

- Fradd, S. H., & Lee, O. (2001). Needed: A framework for integrating standardized and informal assessment for students developing academic language proficiency in English. In S. R. Hurley & J. V. Tinajero (Eds.), *Literacy assessment of second language learners*. Boston: Allyn and Bacon.
- García-Vázquez, E., Vázquez, L. A., López, I. C., & Ward, W. (1997). Language proficiency and academia success: Relationships between proficiency in two languages and achievement among Mexican American students. *Bilingual Research Journal*, 21(4), 395–408.
- Goldenberg, C. (1996). The education of language-minority students: Where are we, and where do we need to go? *The Elementary School Journal*, 96(3), 353–362.
- Gottardo, A. (2002). The relationship between language and reading skills in bilingual Spanish–English speakers. *Topics in Language Disorders*, 22(5), 46–77.
- Hudson, R., & Smith, S. (2001). Effective reading instruction for struggling Spanish-speaking readers: A combination of two literatures. *Intervention in School and Clinic*, 37(1), 36–40.
- Kindler, A. L. (2002). *Survey of states' limited English proficient students and available educational programs and services 1999–2000 summary report*. Washington, DC: National Clearinghouse for English Acquisition and Language Instruction Educational Programs.
- Koda, K. (1994). Second language reading research: Problems and possibilities. *Applied Psycholinguistics*, 15, 1–28.
- Laija, W., & Ochoa, S. H. (1999). Historical roots of politics and policies regarding the education of Hispanics: The Mexican American experience. In A. Tashakkori & S. H. Ochoa (Eds.), *Readings on equal education: Education of Hispanics in the United States: Politics, policies, and outcomes* (Vol. 6, pp. 1–23). New York: AMS Press, Inc.
- MacSwan, J. (2000). The threshold hypothesis, semilingualism, and other contributions to a deficit view of linguistic minorities. *Hispanic Journal of Behavioral Sciences*, 22(1), 3–45.
- Office of Civil Rights. (2001). *Title VI prohibition against national origin discrimination as it affects persons with limited English proficiency*. Retrieved February 15, 2006, from <http://www.hhs.gov/ocr/lep/guide.html>
- Parker, R., & Hasbrouck, J. (1998). *The reliability of the CBM measure, oral reading fluency, in Spanish reading with LEP grade 2 students*. Manuscript in preparation.
- Pray, L. (2005). How Well Do Commonly Used Language Instruments Measure English Oral-Language Proficiency? *Bilingual Research Journal*, 29(2), 387–410.

- Ramírez, J. D. (1992). Executive summary. *Bilingual Research Journal*, 16 (1–2), 1–62.
- Romaine, S. (1995). *Bilingualism* (2nd ed.). Cambridge, MA: Blackwell Publishers.
- Shinn, M. R. (1989). *Curriculum-based measurement: Assessing special children*. New York: The Guilford Press.
- Sosa, A. (1990). *Making education work for Mexican Americans: Promising community practices* (Report No. EDO-RC-90-2). KY: ERIC Clearinghouse on Rural Education and Small Schools Institute for Urban and Minority Education. (ERIC Document Reproduction Service No. ED319580)
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics*. New York: Harper & Row, Publishers.
- Thomas, W. P., & Collier, V. (1997). *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. Retrieved March 8, 2006, from http://www.crede.org/research/llaa/1.1_final.html
- Thompson, B. (1992, April). *Interpreting regression results: Beta weights and structure coefficients are both important*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED344897)
- Ulibarri, D., Spencer, M., & Rivas, G. (1981). Language proficiency and academic achievement: A study of language proficiency tests and their relationship to school ratings as predictors of academic achievement. *Journal of the National Association for Bilingual Education*, 5(3), 47–80.
- Wilen, D. K., & van Maanen, C. (1986). Assessment of limited English proficient Hispanic students. *School Psychology Review*, 15(1), 59–79.
- Woodcock, R. W., & Mather, N. (1989). WJ-R ACH Tests of Achievement: Examiner's Manual. In R. W. Woodcock & M. B. Johnson (Eds.), *Woodcock-Johnson Psycho-Educational Battery-Revised*. Chicago, IL: Riverside Publishing Company.
- Woodcock, R. W., & Muñoz-Sandoval, A. F. (1993). *Woodcock-Muñoz Language Survey: Comprehensive manual*. Chicago, IL: Riverside Publishing Company.
- Yamashita, J. (2002). Mutual compensation between L1 reading ability and L2 language proficiency in L2 reading comprehension. *Journal of Research in Reading*, 25(1), 81–95.