Long-Term Crosslinguistic Transfer of Skills From L1 to L2

Richard Sparks

College of Mt. St. Joseph

Jon Patton, Leonore Ganschow, and Nancy Humbach

Miami (Ohio) University

This study investigated the relationship of first language (L1) skills in elementary school and second language (L2) learning in high school. Students classified as high-, average-, and low-proficiency L2 learners were compared on L1 achievement measures of reading, spelling, vocabulary, phonological awareness, and listening comprehension administered in the first, second, third, and fifth grades. An L2 aptitude measure was administered in ninth grade and L2 word decoding and L2 spelling measures were administered at the end of the first- and second-year L2 courses. Outcome measures were oral and written L2 proficiency measures in Spanish, French, and German administered at the end of 2 years of L2 study. Results showed overall differences among the three proficiency groups on the L1 achievement measures from second through fifth grades, the L2 aptitude measure, and the L2 word decoding and spelling measures. High-proficiency L2 learners exhibited stronger L1 skills and L2 aptitude than the average and low-proficiency L2 learners. Findings showed that L1 skill differences emerged early in elementary school and are related to L2 proficiency and achievement several years later in high school. The findings provide support for long-term crosslinguistic transfer of L1 to L2 skills.

Keywords   L1 skills; L2 learning; L2 aptitude; crosslinguistic transfer; L2 proficiency

Do students who struggle to learn a foreign language (L2) in high school have weaker first language skills (L1) than students who perform well in an L2? If so, how early might one observe language differences between good and poor L2 learners? Are the differences in early L1 skills apparent when students attempt to learn an L2 many years after they have learned to speak and read their L1?

For a number of years now, educators have suspected a relationship between L1 and L2 learning, particularly for reading and spelling alphabetic languages.

Correspondence concerning this article should be addressed to Richard Sparks, College of Mount St. Joseph, 5701 Delhi Rd., Cincinnati, OH 45233. Internet: richard_sparks@mail.msj.edu

(For a review of this literature, see Geva & Verhoeven, 2000; Koda, 2005.) Some researchers have speculated that overall proficiency in the L1 plays the largest role in one's overall proficiency in the L2 (e.g., see Cummins, 1979; Sparks, 1995; Sparks & Ganschow, 1993). To date, however, there have been few research studies that have explored this relationship in a systematic manner in the United States by following students from their elementary school years into high school when they begin the study of an L2. There are several possible benefits from conducting longitudinal studies of L1 learners in their earliest school years who then study a L2 in secondary education. First, early recognition of the impact of L1 on L2 has implications for prediction of L2 proficiency. Second, appropriate teaching in the L1 in elementary school may benefit students when they learn an L2 several years later. Third, a model describing linkages between L1 and L2 provides a rich source for further exploration in countries with large bilingual and multilingual populations. Finally, investigating early L1 and later L2 relationships could provide insights into the extent to which there may be long-term crosslinguistic transfer of L1 skills to L2 learning.

In the present study the authors followed a population of L1 learners from the beginning of 1st grade through the end of the 10th grade, at which time these learners had completed 2 years of L2 study. They were interested in examining the L2 proficiency of these learners in relation to their L1 skills many years earlier. In particular, they were interested in whether there would be differences in the early L1 skills of students identified as high-, average-, and low-proficiency L2 learners according to their performance on measures of oral and written L2 proficiency.

L1-L2 Relationships

Research and discussion on linkages between L1 and L2 learning arose independently from L1 and L2 educators in the 1980s and early 1990s. However, as far back as the 1950s and 1960s, educators had designed tests that included select L1 measures to assess aptitude for learning an L2. (For a review of this literature, see Dörnyei, 2005; Ganschow & Sparks, 2001; Spolsky, 1995). The L1 measures thought to be particularly relevant included tests of vocabulary, sound-symbol relationships, grammatical concepts, and verbal memory (see, e.g., Carroll & Sapon, 1959/2000; Pimsleur, 1968). As a result of his research on the Modern Language Aptitude Test (MLAT; Carroll & Sapon), Carroll inferred that L2 aptitude is, in part, a remnant of L1 learning ability (Carroll, 1962, 1973). Later, Spolsky (1989) included "intact" language

skills as a "necessary condition" for learning an L2 in his model of language learning.

Speculation about the direct impact of L1 on L2 arose in the form of two hypotheses, one from the bilingual field and the other from special education. In the L2 bilingual literature, educator Jim Cummins proposed his linguistic interdependence hypothesis (Cummins, 1979, 1984). Here, Cummins (1979, p. 233) hypothesized that L2 language and literacy skills are dependent in part on L1 literacy competence at the time of critical exposure to the L2. In the threshold hypothesis, he speculated that if a student's L1 competence is low, the competence in the L2 will also be low. In a second hypothesis, the interdependence hypothesis, he contended that language skills will transfer from L1 to L2 if there is sufficient exposure to the L2 and motivation to learn the language. Cummins based his inferences on studies of reading skills in language minority and immigrant child populations in which findings showed a high correlation between L1 and L2 reading. He stressed the critical role of L1 in the home, which then "interacts" with L2 instruction in school (Cummins, 1979, p. 236). A number of studies by Cummins and others support the interdependence and threshold hypotheses; for example, Verhoeven (1994) tested the theory with a group of L1 Turkish children living in the Netherlands by asking, "To what extent do abilities in L1 predict similar abilities in L2?" His findings provided positive evidence for interdependence of L1 and L2 for phonological, literacy, and pragmatic skills.

Several years later, special educators Sparks and Ganschow proposed their linguistic coding differences hypothesis (LCDH; Sparks, 1995; Sparks & Ganschow, 1991, 1993, 1995). The LCDH proposed that both L1 and L2 learning depend on basic language learning mechanisms that are similar to both languages. Initially, they based their hypothesis on observations of college students who had demonstrated histories of difficulty with L2 learning (e.g., students classified as learning disabled). Their early findings indicated that weak L2 learners appeared to have particular difficulties in specific aspects of their L1 (e.g., the phonological /orthographic [sound/symbol] system). These findings supported both Carroll's and Pimsleur's inclusion of phonological measures in their tests of L2 aptitude and studies by L1 reading researchers in which poor readers exhibited problems primarily with phonological processing in learning to read their L1.

Since then, Sparks, Ganschow, and colleagues have conducted extensive research on L1-L2 relationships with L2 learners in the United States. Because language study in the United States generally begins in high school, their studies were limited to high school and college populations. The studies included,

for example, comparisons of strong versus weak L2 learners on a variety of cognitive, affective, and L1 and L2 language dimensions; teacher, parent, and student perceptions of L1/L2 learning; prediction studies over 1 and 2 years of L2 study; factor analysis of L1 and L2 test batteries to determine the best measures to assess L2 performance; and comparisons of instructional methods for "at-risk" L2 learners. The results of these studies have shown that L2 learners in high school and college who achieve higher levels of L2 proficiency and classroom achievement have significantly stronger L1 skills, especially phonological processing skills, than L2 learners who exhibit lower levels of L2 proficiency and achievement (see reviews by Ganschow & Sparks, 2001; Sparks, 2001; Sparks, Ganschow, & Patton, 2008). Their studies have also found that the best predictors of L2 proficiency and achievement are variables related to language skills (e.g., performance on measures of L1 reading and L2 aptitude; Sparks, Ganschow, & Patton, 1995; Sparks, Ganschow, Patton, Artzer, Siebenhar, et al., 1997; Sparks, Javorsky, Patton, & Ganschow, 1998; Sparks, Patton, Ganschow, Humbach, & Javorsky, 2006).

Evidence for L1-L2 linkages also has been generated by the proliferation of studies in the United States and other countries on the impact of L1 literacy and language skills on learning an L2; for example, in the United States, Meschyan and Hernandez (2002) examined the role of L1 nonword decoding in learning Spanish among a group of 80 college students. Their findings showed that L1 pseudoword decoding predicted L2 decoding, grammar, vocabulary, and reading comprehension skills. Durgunoglu, Nagy, and Hancin-Bhatt (1993) found that Spanish speakers who had strong L1 phonological awareness and decoding skills were also strong L2 word decoders. In a study with young Finnish children learning English, Dufva and Voeten (1999) found that 58% of the variance in L2 learning on three separate tasks (i.e., L2 vocabulary, writing, and listening comprehension) was attributed to L1 phonological memory, word decoding, and listening/reading comprehension skills. In a study of Hebrew students learning English, Olshtain, Shohamy, Kemp, and Chatow (1990) reported that L1 academic proficiency played the most important role in learning Engish. In two studies, Kahn-Horwitz, Shimron, and Sparks (2005, 2006) found that performance on L1 Hebrew measures (e.g., word reading, phonological awareness, vocabulary knowledge) predicted L2 reading performance in English. Other investigators have found that L1 skills, including word reading, distinguish between strong and weak L2 learners (see, e.g., Holm & Dodd, 1996; Hulstijn & Bossers, 1992; Humes-Bartlo, 1989; Service & Kohonen, 1995).

Recent studies also have indicated that L1 skills, particularly phonological awareness, are a good predictor of (or show strong correlations with) subsequent

reading and word decoding skills in another language (Durgunoglu, 2002). Investigations of crosslinguistic transfer in the development of literacy skills have shown that phonological awareness skills are correlated across languages, particularly with word recognition skills (see, e.g., Cisero & Royer, 1995; Comeau, Cormier, Grandmaison, & Lacroix, 1999; Geva, Yaghoub-Zadeh, & Schuster, 2000; Lindsey, Manis, & Bailey, 2003). Geva and Verhoeven (2000) concluded that "L1-L2 transfer at the level of underlying component skills such as phonological awareness relates to word recognition and spelling" (p. 265).

Other studies have found that L1 skills are strongly related to L2 reading comprehension. In one investigation with Spanish-speaking fourth-grade English language learners, Proctor, August, Carlo, and Snow (2006) explored the effects of L1 (Spanish) alphabetic knowledge, vocabulary knowledge, listening comprehension, and fluency on L2 (English) reading comprehension. The results showed a significant main effect for L1 vocabulary knowledge on L2 reading comprehension and an interaction between L1 vocabulary knowledge and L2 fluency; that is, faster L2 readers benefited more from L1 vocabulary knowledge than less fluent readers. Proctor et al. suggested that the findings demonstrated a "compelling relationship" between the L1 and L2 of the students. In another recent study with Dutch students in Grades 8 to 10 who were learning English, van Gelderen et al. found a strong relationship between L1 and L2 reading comprehension and a strong effect of L1 metacognitive knowledge on L2 reading comprehension (van Gelderen, Schoonen, Stoel, Glopper, & Hulstijn, 2007). Their findings supported the transfer hypothesis, although language-specific knowledge and fluency also played a role in L2 reading performance.

Studies on L1-L2 Relationships Over Time

To the author's knowledge, there are few studies that have followed students from L1 to L2 learning over periods longer than 2 years. Skehan and Ducroquet (1988) conducted a study to examine long-term interrelationships among three areas of language: L1 skills, L2 aptitude, and L2 achievement. In that study, they followed students who had been participants in the Bristol Language Project (Wells, 1985). The results of those studies in which children from the ages of 15 to 60 months had been administered several measures of L1 development (e.g., rate of development, vocabulary size) showed that there was considerable variation in the rate at which children acquired their L1. When the children were 13–14 years of age and enrolled in secondary school, Skehan and

Ducroquet administered measures of L2 aptitude (i.e., a verbal intelligence test, subtests from the Elementary MLAT, York Language Aptitude Test, Pimsleur Language Aptitude Battery) and L2 achievement (i.e., the NFER/APU tests for spoken and written French and German). In particular, they examined whether success in L2 learning was related to speed of L1 learning and whether L2 aptitude was related to L1 capacity. The findings showed that L1 development prior to entering school, especially on measures of vocabulary and syntax, was significantly related to L2 aptitude many years later and that L2 aptitude was significantly related to L2 achievement. The findings also showed that there were significant relationships between measures of L1 comprehension and vocabulary and later L2 achievement. Studies have also found that L2 aptitude and performance on intelligence (IQ) measures are related; however, each makes separate contributions to L2 learning (Skehan, 1998); for example, Wesche, Edwards, and Wells (1982) and Sasaki (1996) found moderate to strong relationships between L2 aptitude and IQ. However, other studies have shown that exceptional learners, including those with low IQs, can exhibit talent for learning specific L2 skills because they have strong skills in one or more components of L2 aptitude (e.g., see Skehan, 1998; Sparks & Artzer, 2000).

Sparks et al. (2008) studied L1 and L2 in small populations of high school students in the United States over extended time periods. One data set involved 156 students in a suburban high school who had completed 1–2 years of Spanish and whose L1 records dated back to the fourth grade. Two studies were conducted with this population. In the first study, they examined testing measures that best discriminated membership in four groups of L2 learners enrolled in first-year high school Spanish classes: high- and low-achieving L2 learners, students classified as learning disabled (LD), and students with attention deficit hyperactivity disorder (ADHD). High-achieving L2 learners were chosen randomly from a list of students achieving grade A or B in the first semester of the first-year Spanish course and low-achieving L2 learners were those who had achieved grade C, D, or F in the first semester of the first-year Spanish course. Measures included tests of L1 literacy skills (reading, writing, language) administered from 2 to 5 years earlier, L1 cognitive ability, L2 motivation, L2 aptitude, L2 word decoding, and L2 spelling. Findings showed that measures of L1 literacy administered as early as the fourth grade and measures of L2 word decoding and L2 aptitude (MLAT) best discriminated the students in the four groups. The findings supported studies showing that skills related to L1 and L2 literacy, especially decoding, phonological awareness, spelling, and grammar, are related to success in L2 learning.

In a second study with this population, Sparks, Humbach, and Javorsky (2008) compared the aforementioned groups after 2 years of L2 study on a variety of measures that assessed L1 literacy and language, L2 aptitude (MLAT), and oral and written L2 proficiency. Findings showed significant differences between the high- and low-achieving students and the high-achieving and LD groups on all of the measures as early as the fourth grade. They inferred that students with weaker L2 skills in high school exhibited weaker L1 skills several years before their exposure to an L2.

The second data set involved 54 students from a rural school district whom the authors followed from 1st through 10th grades, at which time the students had completed 2 years of L2 study in Spanish, German, or French. The students' L1 skills had been measured in the 1st, 2nd, 3rd, and 5th grades. In the ninth grade, they were administered the MLAT; then, in the 10th grade, they were administered measures of L2 oral (listening, speaking) and written (writing, reading comprehension, decoding, spelling) proficiency. Two studies were published using this data set. In one study, Sparks and colleagues examined best L1 predictors of L2 aptitude on the MLAT and overall L2 proficiency (Sparks et al., 2006). The findings were similar to those of Skehan and Ducroquet (1988) and showed that performance on measures of L1 skill was highly predictive of L2 aptitude and L2 proficiency. The authors suggested that the findings demonstrated the likelihood of long-term relationships between L1 skills as early as primary school and students' L2 aptitude and L2 proficiency in high school.

In a second study with the same 54 students, the authors examined early L1 predictors of later L2 reading (word decoding, comprehension) and spelling (Sparks, Patton, et al., 2008). All L1 measures were administered in the 1st through 5th grades and all L2 measures were administered at the end of the 9th and 10th grades. Findings showed that the best predictor of L2 word decoding skill was a measure of L1 decoding, and the best predictors of L2 spelling were L1 spelling and L1 phonological awareness. The L1 word decoding measure accounted for 43–52% of the variance in L2 word decoding skill, and L1 spelling and L1 phonological awareness accounted for 54% of the variance in L2 spelling. The best predictor of L2 reading comprehension was a measure of L1 reading comprehension. When L2 word decoding skill replaced L1 word decoding as a predictor variable for L2 reading comprehension, results showed that L2 word decoding was also an important predictor of L2 reading comprehension, accounting for 39% of the variance. Findings suggested that even several years after students learned to read and spell their L1, word

decoding, spelling, and reading comprehension skills in the L2 are related to students' L1 skills.

Rationale and Hypotheses

Research evidence to date indicates that students with stronger L1 skills have stronger L2 aptitude and higher levels of L2 proficiency and L2 classroom achievement, demonstrating the possibility of crosslinguistic transfer of L1-L2 skills. Koda (2005) cited evidence that shows that aspects of L1 abilities are transferred during L2 comprehension and production in phonology (Hancin-Bhatt & Bhatt, 1997), communicative strategies (Irujo, 1986; Olshtain, 1983), morphosyntax (Hakuta, 1976; Hancin-Bhatt & Nagy, 1994; Sasaki, 1993), metalinguistic awareness (Durgunoglu et al., 1993; Koda, 1999), and pragmatics (Rutherford, 1983). A hypothesis that L1 skills in early elementary school transfer to L2 proficiency and achievement several years later in high school would entail long-term crosslinguistic transfer of L1 skills to the L2. The longitudinal studies conducted by Skehan and Ducroquet and Sparks, Ganschow, and colleagues described earlier provide some evidence that L1 skills prior to entry into school and L1 skills developed in elementary school are strongly related to later L2 aptitude and L2 oral and written proficiency.

In a previous study over 2 years, Sparks, Ganschow, et al. (1998) examined the extent to which there would be differences in L1 skills, L2 aptitude, and L2 classroom achievement with two different populations of high school students—one public and one private school—who had completed 2 years of L2 study in Spanish, French, or German. In these studies, the L1 achievement measures had been administered in high school at the time when the students began the study of an L2. In both populations, the students were divided into high-, average-, and low-proficiency groups based on their score on a measure of oral (listening and speaking) and written (reading and writing) L2 proficiency administered at the end of 2 years of foreign language (FL) study. The results showed that the high-proficiency L2 learners in both the public and private schools achieved significantly stronger scores on all L1 measures (reading, spelling, vocabulary, language, phonological awareness), L2 aptitude (MLAT), and L2 achievement (course grades) than the low-proficiency learners; on some of the L1 measures, the high-proficiency learners achieved stronger scores than the average-proficiency learners and the average-proficiency learners outscored the low-proficiency learners.

The present study replicates but also extends the aforementioned 1998 study by administering L1 measures to students when they began the first grade and then following the students over 10 years until they had completed 2 years of Spanish, French, or German in high school. The MLAT was administered when students began L2 study in the 9th grade and the oral and written L2 proficiency measures when they completed the second year of L2 study in the 10th grade. The authors then divided the students into high-, average-, and low-proficiency groups based on their L2 proficiency scores. Based on the work of Cummins (i.e., the interdependence and threshold hypotheses) and the LCDH (i.e., L1 and L2 learning depend on similar language learning mechanisms for both languages), the authors hypothesized that there would be long-term crosslinguistic transfer of L1 skills developed in elementary school to L2 skills several years later when students first encountered an L2 in high school. Thus, students with higher levels of L2 proficiency after 2 years of L2 study in high school were expected to exhibit stronger L1 skills early in their elementary school years and stronger L2 aptitude measured at the beginning of L2 study. In addition, students with higher levels of L2 proficiency were expected to exhibit stronger L2 word decoding and L2 spelling skills as well as stronger L2 classroom achievement (course grades). In line with previous results, such findings would provide support for assertions that (a) early L1 skills are related to later L2 learning, (b) students who demonstrate varying levels of L2 proficiency and L2 achievement exhibit differences in their L1 skills, and (c) potential differences in later L2 learning emerge early in elementary school prior to exposure to an L2. Such findings would also support the notion of long-term crosslinguistic transfer in alphabetic languages from L1 to L2.

The present study is part of a larger prospective investigation conducted with the participants. In their previously published studies described earlier in the review, Sparks and his colleagues examined the best L1 predictors of L2 aptitude and overall L2 proficiency (Sparks et al., 2006) and whether early L1 word decoding, reading comprehension, and spelling measures would predict later L2 word decoding, reading comprehension, and spelling in high school (Sparks, Patton, et al., 2008). The study reported here differs from the aforementioned studies by specifically examining whether students who have different levels of L2 proficiency after 2 years of L2 study in high school exhibited significantly different levels of L1 skills (i.e., reading, spelling, vocabulary, listening comprehension, phonological awareness) from 5–10 years earlier in elementary school, L2 aptitude in ninth grade, L2 word decoding and spelling after 2 years of L2 study, and L2 classroom achievement.

Method

Participants

Fifty-four students attending a large, middle-class, rural public school district in the midwestern United States participated in the study. The participants were followed over a 10-year period and had completed 2 years of Spanish, French, or German courses in the 9th and 10th grades. Thirty of the participants had taken Spanish, 14 had taken French, and 10 had taken German. There were 29 female and 25 male participants. The mean age of the participants at the beginning of the first grade was 6.9 years (age range: 6.3–7.9 years); at the end of the study 10 years later, the mean age of the participants was 16.4 years (age range: 15.9–16.9 years). The authors used a cohort model in which they selected a sample of first-grade students over 3 years and followed each cohort over 10 years. The study initially began with a total of 156 students when they entered the first grade in this school district. By the ninth grade, 101 of the 156 students still attended school in the district. Seventy-seven of the 101 students enrolled in an L2 course in high school and began their study of an L2 in the ninth grade. Fifty-four of the 77 students chose to continue their participation in the study when they began their first year in an FL course. None of the 54 participants had learned an L2 or had received L2 instruction prior to the ninth grade. All of the participants were monolingual and their home language background was English. Each participant was exposed to similar learning conditions and had the same number of hours of L2 instruction in the classroom over 2 years. None of the participants repeated an L2 course. To the authors' knowledge, the participants were not exposed to the L2 in which they were enrolled outside of the classroom. Parental permission was obtained for each participant.

Instruments

Measures of L1 Achievement, L2 Aptitude, L2 Word Decoding and Spelling, and L2 Classroom Achievement (Dependent Variables)

   Word decoding. The measure of L1 word recognition and decoding was the Woodcock Reading Mastery Test-Revised Basic Skills Cluster, Forms G and H (Woodcock, 1987). The Basic Skills Cluster is comprised of two subtests: Word Identification and Word Attack. On the Word Identification subtest, a student read aloud a list of increasingly difficult words. On the Word Attack subtest, a student read aloud a list of increasingly difficult pseudowords. For a response to be considered correct, the student had to produce a natural reading (pronunciation) of the word or pseudoword. A test-retest reliability of .96 was reported for the Basic Skills Cluster.

Spelling.   The measure of L1 spelling was the Test of Written Spelling-2 (Larsen & Hammill, 1986). On this dictated word test, the student wrote the words spoken by the examiner. The response was marked as correct or incorrect. A test-retest reliability of .95 was reported for the test.

Reading comprehension.   The measure of L1 reading comprehension was the Formal Reading Inventory, Forms A and B (Wiederholt, 1986). On this multiple-choice test, the student read silently a short paragraph and answered five questions about the paragraph. The response was marked as correct or incorrect. An internal consistency of .92–.97 was reported for the two forms of the test.

Phonological awareness.   The measure of L1 phonological awareness was the Lindamood Auditory Conceptualization Test, Forms A and B (Lindamood & Lindamood, 1979). On this test, the student manipulated blocks of various colors to indicate his/her conceptualization of speech sounds said aloud by the examiner. The student had to show the number, sameness, or difference of sounds (e.g., show me */b//b//z/*) and represent, with the blocks, changes in spoken patterns without associating the sounds with letters; for example, if that says */ups/*, show me */usp/*. This test yields raw scores with a range of 0–100. Pretest-posttest reliability of .96 was reported for the two forms of the test.

Vocabulary.   The measure of L1 vocabulary was the Peabody Picture Vocabulary Test-Revised, Forms L and M (Dunn & Dunn, 1981). The test measures receptive vocabulary for standard American English. On this test, the student was shown four pictures and asked to identify the picture for the word spoken by the examiner. The response was marked as correct or incorrect. A median test-retest reliability of .82 was reported for the two forms of the test.

Listening comprehension.   The measure of L1 listening comprehension was the Woodcock Reading Mastery Test-Revised Passage Comprehension subtest, Forms G and H (Woodcock, 1987). This cloze test consists of reading a short passage (one to two sentences) aloud to a student and asking him/her to identify aloud a key word missing from the passage. The student was not permitted to read (see) the passage, but the passage can be repeated. The response was marked as correct or incorrect. Generally, this subtest is used as a measure of reading comprehension; however, the aforementioned alternative procedure was recommended by Aaron (1989) as a diagnostic indicator in identifying problem readers and was used for this study. A test-retest reliability of .92 was reported when this subtest was used as a measure of reading comprehension.

The L1 achievement tests were chosen because they measured language skills that have been found in numerous studies to distinguish high-, average-, and low-achieving L2 learners (e.g., see Sparks, Humbach, & Javorsky, 2008).

In addition, the L1 tests measured skills similar to those on L2 aptitude tests such as the MLAT; for example, the word decoding test (Woodcock) assessed skills in the L1 similar to those measured by the MLAT Phonetic Script subtest (i.e., sound-symbol relationships). The vocabulary test (Peabody Picture) assessed vocabulary skills in the L1 similar to the MLAT Spelling Clues subtest, which measures vocabulary knowledge after the student has read a phonetically spelled word.

A description of each L1 measure is provided in Appendix A.

L2 aptitude. The Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959/2000) was used to measure students' L2 aptitude. This test is designed to provide an indication of a student's probable degree of success in learning an L2. The MLAT was chosen because it has been found to be a good predictor of L2 proficiency and classroom achievement. A description of the test is provided in Appendix A.

L2 word decoding. The L2 word decoding task for each of the three languages consisted of a list of 20 real words and a list of 20 pseudowords. The real words and pseudowords in the three L2s measured a specific decoding task that involved the use of vowel sounds in each L2 that were not consistent with the vowel sounds in English (i.e., vowels for which the phoneme-grapheme correspondence differs between English and the L2). The vowel sounds in the L2s used different letter combinations and contained multisyllabic words and words with diacritical markings. Students read each word list twice, once at the end of the first-year L2 course and again at the end of the second-year L2 course. For a response to be considered correct, the student had to produce a natural reading (pronunciation) of the word or pseudoword. The reliability of the L2 word decoding lists and pseudoword lists was checked by calculating Cronbach's alpha. For the Spanish words, the reliability coefficient for the word decoding list was .75, and for the pseudoword decoding list, it was .75. For the French words, the reliability coefficient for the word decoding list was .78, and for the pseudoword decoding list, it was .84. For the German words, the reliability coefficient for the word decoding list was .81, and for the pseudoword decoding list, it was .79. The word lists in the three L2s are presented in Appendix B.

The maximum score for each year was 40 words (20 real words and 20 pseudowords) for a potential maximum score of 80 words (over 2 years). The scores from the first- and second-year L2 courses were combined because there were not significant differences among the three L2 groups (Spanish, French, German) in first-year L2 word decoding, $F(2, 51) = 0.05$, $p = .95$, and second-year L2 word decoding, $F(2, 51) = 0.73$, $p = .49$. The L2 word

decoding measure was not included in the L2 total proficiency score because word decoding is not included as part of students' L2 proficiency in the ACTFL Guidelines.

L2 spelling.   The spelling task for each of the three L2s, which differ in terms of the transparency of their phoneme-grapheme correspondences, consisted of 20 words designed to measure specific encoding tasks that included vowel sounds in the L2s that were not consistent with the vowel sounds in English, used different letter combinations, and contained multisyllabic words and words with diacritical markings. The student wrote the words spoken by the examiner. A response was marked as correct or incorrect. The reliability of the L2 spelling list was checked by calculating Cronbach's alpha. For the Spanish words, the reliability coefficient was .78; for the French words, the reliability coefficient was .81; and for the German words, the reliability coefficient was .74. The spelling lists are presented in Appendix C. The L2 spelling measure was not included in the L2 total proficiency score because spelling is not included as part of the students' L2 proficiency in the ACTFL Guidelines.

L2 grades.   Each student's first- and second-year L2 grades (L2 Grade-Yr. 1, L2 Grade-Yr. 2) represented a compilation of their first- and second-semester grades of each school year. The grades represented scores on oral and written quizzes and tests, projects, in-class activities, and homework. All letter grades were transformed to a standard GPA scale (i.e., F = 0.00 to A+ = 4.33). The teachers for Spanish, French, and German taught both the first- and second-year L2 courses.

Measure of L2 Proficiency (Independent Variable)

The L2 proficiency measure used in this study was the combined scores of students' performance in the four areas identified by the American Council on Teaching Foreign Language (ACTFL) Proficiency Guidelines (1986, 1989) as essential for L2 acquisition: reading, writing, listening, and speaking. The justification for combining the scores on the three tests was twofold. First, the ACTFL guidelines define proficiency as reading comprehension, writing, and speaking/listening to an L2. Second, there were no significant differences among the three L2 groups (Spanish, French, German) on the L2 reading comprehension test, $F(2, 51) = 1.34, p = .27$, the L2 writing test, $F(2, 51) = 1.08$, $p = .35$, and the L2 listening/speaking test, $F(2, 51) = 1.86, p = .17$.

Three L2 educators, who were formally trained in ACTFL guidelines, designed the proficiency tests in their respective languages. The individual tests were designed to ensure uniformity as closely as possible across the three languages (i.e., Spanish, French, German); that is, reading, writing, speaking/

listening in German were measured by using the same items and prompts as those used for Spanish and French. The directions were the same for each of the three languages. All of these proficiency measures had been used by the authors in their previous studies (e.g., Sparks, Ganschow, et al., 1998; Sparks, Patton, et al., 2006, 2008). A description of each measure is provided here.

On the L2 reading comprehension test, students read a one-page letter written in the L2 and answered 10 multiple-choice questions in English about the letter. The questions were written in English to conform to the ACTFL proficiency guidelines. The letter was designed using criteria descriptive of the intermediate-high level of the ACTFL guidelines. Next, the students were given a passage to read from *Reader's Digest* in Spanish, French, or German and were asked to answer 10 multiple-choice questions in English about the passage. The students were given 15 min to read the letter and answer the questions and 15 min to read the passage and answer the questions (maximum score = 20).

On the L2 writing test, students wrote a response in the L2 to the letter that had been read for the reading comprehension task. The letter contained five questions to which the students were asked to respond. The students were given 15 min to write the letter. Each student's writing sample was scored for five criteria: vocabulary, cultural appropriateness, structure, comprehensibility, and spelling. The ACTFL guidelines for determining proficiency levels were used in assigning the scores (0–5) on each of the five criteria. Each student was evaluated in the five areas using the five-point scale: 0 = *no production*, 1 = *novice-low*, 2 = *novice-mid*, 3 = *novice-high*, 4 = *intermediate-low*, and 5 = *intermediate-high* and above (maximum score = 25). A score of 0 was used in the scoring because some students may have been unable to produce any response in the L2 after 2 years of L2 learning in high school.

On the L2 speaking/listening test, students' proficiency was assessed through a 10–15-min individual oral proficiency interview. The interviewers had no previous information about the participants, including their L1 achievement and L2 aptitude test scores or their L2 classroom achievement. Interviewers used randomly selected topics about which the students conversed. Each student's interview was taped so that it could be scored at a later date. The oral interview was scored for five criteria: pronunciation, vocabulary, grammar, comprehensibility, and listening comprehension (maximum score = 25). The remainder of the scoring procedure and maximum score (25) was the same as for the L2 writing test.

The students' total proficiency score (L2 Proficiency) was the combined number of points on the L2 reading comprehension, writing, and listening/speaking tests (maximum score = 70). The reliability of the L2 proficiency

measure was checked by calculating Cronbach's alpha. For the total test, the reliability coefficient was .86.

Procedure

The students' scores on the L1 achievement measures were collected at five specific time points during the completion of this study. The time at which each analysis was conducted with its corresponding label is as follows: PRE (beginning of first grade), POST 1 (end of first grade), POST 2 (end of second grade), POST 3 (end of third grade), and POST 4 (end of fifth grade). The labels PRE–POST 4 for the time periods are used throughout the remainder of the article. The students were administered the Woodcock Reading Mastery Test, Test of Written Spelling, Formal Reading Inventory, and Peabody Picture Vocabulary Test at the following times: the beginning of first grade (PRE) and at the end of the first, second, third, and fifth grades (POST 1–POST 4). The Lindamood Auditory Conceptualization Test was administered at the beginning of the first grade (PRE) and at the end of the first, second, and third grades (POST 1–POST 3). The listening comprehension measure was administered at the end of the third (POST 3) and fifth (POST 4) grades. The aforementioned testing measures were administered within the first month and last month of each school year. All of the L1 measures were administered individually to the participants by the first two authors with assistance from graduate and undergraduate students trained by the authors.

The MLAT was administered at the beginning of the ninth grade in small groups by the first author. The L2 word decoding measure was administered at the end of the students' 9th- and 10th-grade years. The L2 proficiency (reading comprehension + writing + listening/speaking) and L2 spelling measures were administered at the end of the students' 10th-grade year. The L2 listening/speaking, word decoding, and spelling measures were administered individually by one of the authors of the L2 proficiency measures, who was fluent in all three L2s and had been trained to administer L2 proficiency tests according to ACTFL guidelines, and graduate students trained by the fourth author.

Data Analysis

To determine the L2 proficiency level (independent variable), the 54 students were classified into three groups—High Proficiency (HIGH PROF), Average Proficiency (AVE PROF), and Low Proficiency (LOW PROF)—based on their scores on the L2 proficiency test. The grouping procedure involved determining the extent to which a given student's score deviated from the total group's mean score on the L2 proficiency test. All of the students' scores on the L2

proficiency test were transformed to 2 scores. Next, students scoring more than two-thirds standard deviation above the overall sample mean were identified as HIGH PROF; those between $+.67$ and $-.67$ standard deviations from the sample mean were identified as AVE PROF; and those scoring more than two-thirds standard deviation below the sample mean were identified as LOW PROF. Two-thirds standard deviation was used as the cut point for the HIGH and LOW PROF groups to ensure that enough students were included in each group for comparison purposes.

To determine whether there would be a need to control for cognitive ability, an analysis of variance procedure (ANOVA) was conducted comparing the three groups' scores on a measure of academic aptitude and overall cognitive ability, the Test of Cognitive Skills (CTB/McGraw Hill, 1983). Students' scores on this test were obtained from school records at the end of the first grade. A description of this measure is provided in Appendix A. The results showed that there were no significant differences among the three proficiency groups on this measure, $F(2, 51) = 1.87$, $p = .16$. Thus, in all subsequent analyses, the authors did not control for students' level of cognitive ability.

A multiple analysis of variance procedure (MANOVA) was used to determine whether there were overall differences among the HIGH, AVE, and LOW PROF groups on the six L1 achievement measures (dependent variables). Separate MANOVA procedures were used to determine whether there were overall differences among the groups on the L1 word decoding and spelling measures and in first- and second-year L2 grades (dependent variables). In the event a MANOVA analysis was significant, a one-way analysis of variance (ANOVA) was used to compare the HIGH, AVE, and LOW PROF groups on the aforementioned measures. A separate ANOVA procedure was used to determine whether there would be group differences on the L2 aptitude measure, the MLAT (dependent variable). The criterion for significance was a level of $p \leq .05$. To reduce the possibility of Type I error, a Scheffé procedure was used in comparing individual group differences on each measure. Effect sizes (Cohen's d) were calculated for all group comparisons on which there were significant differences.

Appendix D reports the scores of the HIGH PROF, AVE PROF, and LOW PROF groups on the L1 achievement measures in the PRE through POST 4 time periods. Appendix E reports the scores of the three proficiency groups on the L2 proficiency, L2 aptitude, L2 word decoding, and L2 spelling measures, and first- and second-year L2 grades. Appendices F–J present intercorrelations in the PRE–POST 4 time periods among the L1 achievement, L1 cognitive ability,

L2 aptitude, L2 proficiency, L2 word decoding, and L2 spelling measures, and L2 grades.

Results

Results are reported under separate headings for the dependent variables: L1 achievement, L2 aptitude, L2 word decoding and spelling, and L2 grades. Table 1 profiles between-group differences on all of the testing measures as well as effect sizes for those measures on which between-group differences were found.

L1 Achievement

In the PRE analysis, results of the MANOVA procedure showed that the groups did not exhibit overall significant differences on the L1 measures, Wilks's lambda $= .799$, $F(10, 94) = 1.11$, $p = .36$.

In the POST 1 analysis, results of the MANOVA procedure showed that the groups did not exhibit overall differences on the L1 measures, Wilks's lambda $= .783$, $F(10, 94) = 1.22$, $p = .29$.

In the POST 2 analysis, results of the MANOVA procedure showed that the groups exhibited significant overall differences on the L1 measures, Wilks's lambda $= .595$, $F(10, 94) = 2.78$, $p = .005$. Individual ANOVAs showed between-group differences on the Woodcock Reading Mastery Test, the Test of Written Spelling, and the Peabody Picture Vocabulary Test favoring the HIGH PROF group over the AVE PROF and LOW PROF groups; and on the Formal Reading Inventory favoring the HIGH PROF and AVE PROF groups over the LOW PROF group. Effect sizes ranged from $d = .67$ to 1.44 for all between-group differences.

In the POST 3 analysis, results of the MANOVA procedure showed that the groups exhibited significant overall differences on the L1 measures, Wilks's lambda $= .595$, $F(12, 92) = 2.27$, $p = .01$. Individual ANOVAs showed between-group differences on the Woodcock Reading Mastery Test, the Test of Written Spelling, the Peabody Picture Vocabulary Test, and the Lindamood Auditory Conceptualization Test, favoring the HIGH PROF group over the AVE PROF and LOW PROG groups; and on the Formal Reading Inventory, favoring the HIGH PROF group over the LOW PROF group. Effect sizes ranged from $d = .70$ to .70 for all between-group differences.

In the POST 4 analysis, results of the MANOVA procedure showed that the groups exhibited significant overall differences on the L1 measures, Wilks's lambda $= .639$, $F(10, 94) = 2.36$, $p = .02$. Individual ANOVAs showed

Table 1 Between-group differences using Scheffé corrections and effect sizes on measures of L1 achievement, L2 aptitude, L2 word decoding and spelling, and L2 grades among the HIGH PROF, AVE PROF, and LOW PROF groups

Testing measures	HIGH PROF vs. LOW PROF			HIGH PROF vs. AVE PROF			AVE PROF vs. LOW PROF		
	Post 2	Post 3	Post 4	Post 2	Post 3	Post 4	Post 2	Post 3	Post 4
L1 achievement									
Woodcock Reading Mastery Test Basic Skills	*(1.35)	*(1.35)	*(1.15)	*(1.20)	*(0.70)	*(0.86)			
Test of Written Spelling	*(1.24)	*(1.50)	*(1.63)	*(0.67)	*(1.04)	*(0.86)			
Peabody Picture Vocabulary Test-Revised	*(1.27)	*(1.22)	*(1.18)	*(0.82)	*(0.93)	*(0.53)			
Formal Reading Inventory	*(1.44)	*(1.03)	*(1.08)			*(0.84)	*(0.85)		
Lindamood Auditory Conceptualization Test	—	*(0.95)	—		*(1.08)	—	—		—
Listening Comprehension	—		*(0.96)	—					
L2 aptitude		*(2.08)			*(1.13)			*(0.83)	
L2 word decoding		*(1.57)			*(1.08)				
L2 spelling		*(1.90)			*(1.41)				
L2 grades									
Year 1		*(2.39)			*(1.15)			*(0.81)	
Year 2		*(3.06)			*(1.18)			*(1.05)	

Note. PRE and POST 1 results are not included because the MANOVA showed no overall differences among the three groups. Effect sizes are listed in parentheses.

*Significant between-group differences.

between-group differences on the Woodcock Reading Mastery Test, the Test of Written Spelling, the Peabody Picture Vocabulary Test, and the Formal Reading Inventory, favoring the HIGH PROF group over the AVE PROF and LOW PROF groups; and on the listening comprehension measure, favoring the HIGH PROF group over the LOW PROF group. Effect sizes ranged from $d = .53$ to 1.63 for all between-group differences.

L2 Aptitude
Results of the ANOVA procedure showed that the groups exhibited significant differences on the MLAT, $F(2,51) = 12.57$, $p = .0001$. Individual ANOVAs showed significant between-group differences favoring the HIGH PROF group over AVE PROF and LOW PROF groups and the AVE group over the LOW PROF group. Effect sizes ranged from $d = .83$ to 2.08 for all between-group differences.

L2 Word Decoding and Spelling
Results of the MANOVA procedure showed that the groups exhibited significant overall differences on the L2 word decoding and L2 spelling measures, Wilks's lambda $= .657$, $F(5, 84) = 5.93$, $p = .0003$. Individual ANOVAs showed between-group differences on both the L2 word decoding and L2 spelling measures, favoring the HIGH PROF group over the AVE PROF and LOW PROF groups. Effect sizes ranged from $d = 1.08$ to 1.57 for L2 word decoding and from $d = 1.41$ to 1.90 for L2 spelling for all between-group differences.

L2 Grades
Results of the MANOVA procedure showed that the groups exhibited significant overall differences on the measures of classroom achievement, first- and second-year L2 grades, Wilks's lambda $= .601$, $F(4, 100) = 7.15$, $p = .0001$. Individual ANOVAs showed between-group differences in both first- and second-year L2 grades, favoring the HIGH PROF group over the AVE PROF and LOW PROF groups, and AVE PROF group over the LOW PROF group. Effect sizes ranged from $d = .81$ to 2.39 for first-year L2 grades and from $d = 1.05$ to 3.06 for second-year L2 grades for all between-group differences.

Discussion

High school students who had completed 2 years of L2 courses were grouped by their L2 proficiency level and compared on L1 and L2 achievement measures and a measure of L2 aptitude. The study examined the relationship among the groups' scores on the measure of L2 proficiency and their performance

on measures of (a) L1 skills administered in the first, second, third, and fifth grades, (b) L2 aptitude administered prior to the beginning of L2 study in the ninth grade, (c) L2 word decoding and L2 spelling administered at the end of first and second-year L2 courses, and (d) L2 classroom achievement (i.e., course grades). The results are discussed under the following headings: L1 Skills, L2 Aptitude and L2 Classroom Achievement, and L2 Word Decoding and Spelling.

L1 Skills

The results of the MANOVA analyses showed that the HIGH PROF, AVE PROF, and LOW PROF groups exhibited significant overall differences on the L1 measures (i.e., reading spelling, vocabulary, phonological awareness, listening comprehension) administered in elementary school several years prior to enrolling in an L2 course. Most group differences were found between the HIGH PROF and LOW PROF groups and the HIGH and AVE PROF groups. The findings are similar to Sparks et al.'s 1998 study in which overall group differences were found on measures of L1 skill administered at the beginning of L2 study in high school (Sparks, Ganschow, et al., 1998). The findings support research conducted by Sparks, Ganschow, and colleagues who have found in other studies that students with different levels of L2 proficiency exhibit significantly different levels of L1 skill (e.g., see Sparks, Artzer, Patton, Ganschow, Miller, & Hordubay, 1998; Sparks et al., 2008). A new finding in this study is that high school L2 learners with different levels of L2 proficiency exhibited significant differences on measures of L1 learning from 4 to 7 years prior to beginning L2 courses. This finding supports the premises of Cummins's linguistic threshold hypothesis; that is, if students' L1 competence is low, competence in the L2 will also be low.

An important finding is that students with different levels of L2 proficiency in high school exhibited significant differences in L1 skills as early as the second grade of elementary school. By the second grade, HIGH PROF students achieved significantly higher scores than the AVE PROF and LOW PROF students on L1 measures of word decoding, spelling, and vocabulary, and the HIGH PROF and AVE PROF students exhibited significantly higher scores than the LOW PROF students in L1 reading comprehension. These results are consistent with recent studies that have found that L1 differences among stronger and weaker L1 learners emerge early in elementary school (e.g., see Kahn-Horwitz, Shimron, & Sparks, 2005, 2006; Sparks, Humbach, et al., 2008; Sparks, et al., 2008). The finding that students with different levels of L2 proficiency in high school exhibited L1 differences in the early years of

primary school provides support for Cummins's and Sparks's and Ganschow's hypotheses about L1-L2 links and also suggests that differences in L1 skills in the primary school years are related to students' L2 aptitude and level of L2 learning several years later.

Despite the finding that the LOW PROF L2 learners achieved the lowest scores on the L1 measures, their scores on the L1 measures were still in the average range of achievement. Initially, the LOW PROF learners achieved in the below-average range at the beginning of first grade (PRE) on measures of L1 word decoding ($SS = 89$, 23rd percentile), L1 spelling ($SS = 74$, 4th percentile), and L1 reading comprehension ($SS = 80$, 9th percentile). However, their scores on these measures increased and were in the average range by the end of the first grade (L1 word decoding $SS = 96$, 39th percentile; L1 spelling $SS = 91$, 27th percentile; L1 reading comprehension $SS = 98$, 45th percentile). The LOW PROF L2 learners continued to achieve in the average range through the fifth grade. Nonetheless, their average performance on all of the oral and written L1 measures did not lead to L2 proficiency scores similar to the HIGH PROF and AVE PROF students, who exhibited even stronger scores on some of the L1 measures. In their earlier studies, Sparks and Ganschow found that L2 learners who achieved low scores on measures of L2 proficiency (and achieved low grades in L2 courses) scored in the average to low-average range on measures of L1 skill (see reviews by Ganschow & Sparks, 2001; Sparks, 1995, 2001). Other researchers have also found that weaker L2 learners exhibited average to low average L1 skills (e.g., see Humes-Bartlo, 1989; Kahn-Horwitz, Shimron, & Sparks, 2005). These findings support Cummins's threshold hypothesis (i.e., low L1 competence leads to low L2 competence) and reinforce Sparks's and Ganschow's findings that students with average to low-average L1 skills exhibit lower levels of L2 proficiency than students with stronger L1 skills.

The findings also revealed that the strength of the correlations on several L1 measures (Woodcock Basic Skills, Test of Written Spelling, Peabody Picture Vocabulary Test, Formal Reading Inventory) increased from the beginning of first grade (Appendix F) to the end of fifth grade (Appendix J). In most cases, the correlations at the end of first grade between L1 skills and L2 proficiency (.41–.60) were similar to or stronger than the correlation between L1 cognitive ability on the Test of Cognitive Skills and L2 proficiency (.42). This finding suggests that although L1 cognitive ability plays a role in L2 proficiency, as students progress through school, their L1 skills, especially L1 literacy, become more strongly related to their subsequent L2 proficiency even several years after they learned their L1. In a previous study with this population, Sparks et al. (2006) found that the Woodcock Reading Mastery Test, which measures both

L1 word decoding and L1 reading comprehension skills, explained 40% of the variance in L2 proficiency after 2 years of L2 study. These findings suggest that early L1 skills may play a strong role in the learning of an L2 and that language learning in L1 and L2 may rely on similar language learning mechanisms.

In L1 research, investigations have found that students with lower levels of L1 literacy generally are exposed to less complex language (e.g., read easier texts, read fewer books, read less frequently), resulting in lower levels of language ability over time; for example, Stanovich (1986, 2000) and others (e.g., see Cunningham & West, 2000) have shown that differences in reading skills result in differences in the acquisition of knowledge and skills taught in school. In the present study, students with lower L2 proficiency also exhibited the lowest levels of L1 literacy, L1 receptive vocabulary, and L1 listening comprehension. These findings lead to speculation that even students with lower but not deficient oral and written L1 skills may read less over time than students with higher L1 skills. The differential exposure to reading over time may explain differential bootstrapping of further vocabulary, knowledge, and cognitive structures that creates individual differences that manifest themselves in differential performance in school and in L2 learning, which depends in large part on language learning skills. Although acquisition of L1 differs from L2, MacWhinney (2005) has proposed that L2 learning is heavily influenced by transfer from L1 and described evidence for transfer in several language skills (e.g., lexicon, articulation, pragmatics, audition). The findings in this study suggest the possibility of long-term crosslinguistic transfer of L1 skills to L2 proficiency and lead to speculation that language learning may be driven, at least in part, by frequency of use and exposure to language.

L2 Aptitude and L2 Classroom Achievement

Differences were found in the groups' aptitude for L2 learning and their performance in L2 classes. On the MLAT, HIGH PROF students ($M = 113.8$) scored almost one standard deviation higher than the AVE PROF learners ($M = 100.7$) and over one and one-half standard deviations higher than the LOW PROF learners ($M = 89.5$). Likewise, there were large and significant differences among the three groups in their L2 classroom achievement (i.e., course grades). HIGH PROF learners achieved almost 1 point (0.8) higher than the AVE PROF group and 1.5 points higher than the LOW PROF group in first-year L1 grades. After another year of L2 study, the HIGH PROF learners achieved 1 full point higher than the AVE PROF learners in second-year L2 grades and 2.0 points higher than the LOW PROF students. These findings are supportive of numerous studies conducted over several years by Sparks and

Ganschow, who have found that students with higher scores on the MLAT and higher L2 course grades also achieved significantly higher levels of oral and written L2 proficiency than students who achieved lower scores on the MLAT (see reviews by Ganschow & Sparks, 2001; Sparks, 2001; Sparks, Humbach, et al., 2008).

Examination of the correlations between students' scores on the MLAT administered in the ninth grade and their L1 cognitive ability on the Test of Cognitive Skills administered in the first grade revealed a strong relationship between the two measures. The correlation between the MLAT and L1 cognitive ability (.64) was higher than the correlations between the students' MLAT score and their L1 skills at the beginning of first grade (.41–.61). However, the correlations between the MLAT and students' L1 skills by the end of the fifth grade had increased and, in most cases, were higher (.60–.76) than the correlation between the MLAT and L1 cognitive ability. In a previous study with this population, Sparks et al. (2006) found that L1 skills at the end of the fifth grade combined with L1 cognitive ability explained 73% of the variance in their MLAT scores, whereas the L1 skills alone accounted for 61% of the total variance. These findings suggest that the MLAT, an L2 aptitude test, is likely measuring, in large part, students' aptitude for language learning generally and that L2 aptitude, at least on the MLAT, is strongly related to their L1 skills, including L1 literacy skills. The results of this study and others may help to explain one reason why the MLAT has been found to be a good predictor of L2 proficiency and achievement: L1 skills are a strong predictor of L2 aptitude and, in turn, L2 aptitude is a good predictor of L2 proficiency and classroom achievement. The findings reinforce Carroll's (1962, 1973) speculation that L2 aptitude is a remnant of L1 learning ability as well as Skehan's (1998) contention that the skills measured by L2 aptitude tests are "qualitatively different" from other areas of knowledge and learning measured by intelligence tests (p. 233).

In their study, Skehan and Ducroquet (1988) found that early L1 development was significantly related to L2 aptitude many years later and that L2 aptitude was significantly related to L2 achievement. The results of the present study are supportive of their findings because performance on measures of early L1 literacy were not only strongly related to L2 aptitude but that L2 aptitude was also strongly related to subsequent L2 proficiency and classroom achievement. The finding that early L1 skills are strongly related to and predictive of L2 aptitude many years later suggests that L1 skills, especially reading, have important consequences for students' cognitive and language development (e.g., see Cunningham & Stanovich, 1991; Stanovich, 1993; Stanovich & Cunningham, 1993) and that their L1 language development may have

important consequences for later L2 proficiency; that is, more frequent expo-
sure to language through reading contributes to, for example, general language
skills such as vocabulary, background knowledge, and familiarity with com-
plex syntactic structures. The findings here suggest the possibility of long-term
crosslinguistic transfer between L1 learning, especially L1 literacy, and, later,
L2 aptitude, L2 classroom achievement, and L2 proficiency.

L2 Word Decoding and L2 Spelling
On measures of L2 word decoding and L2 spelling, the HIGH PROF learners
achieved significantly stronger scores than both the AVE PROF and LOW
PROF learners. These findings mirrored those with similar L1 measures
used in this study; that is, the HIGH PROF learners achieved significantly
higher scores than the AVE PROF and LOW PROF groups on the L1 word
decoding measure, the Woodcock Basic Skills Cluster, and the L1 spelling
measure, the Test of Written Spelling. In numerous studies conducted over
several years, Sparks and Ganschow have found that strong and weak L2
learners exhibit significant differences on measures of L1 and L2 phonological
(sound) and phonological/orthographic (sound/symbol) measures (see reviews
by Ganschow & Sparks, 2001; Sparks, 1995, 2001). In a recent study Sparks,
Humbach, and Javorsky (2008) found that high- and low-achieving L2 learners
exhibited significant differences on measures of L1 reading as early as the
fourth grade. In another study, Sparks, Ganschow, and Patton (2008) found
that measures of L1 reading administered in elementary school and a measure
of L2 word decoding skill were the best discriminators among groups of
high- and low -achieving L2 learners enrolled in high school L2 classes. The
findings in the present study and the other studies cited here support long-term
crosslinguistic transfer of L1 decoding and spelling skills to L2 decoding and
spelling, at least for alphabetic languages.

Examination of the correlations between students' performance on mea-
sures of L1 decoding (Woodcock) and L2 decoding and measures of L1 spelling
(Test of Written Spelling) and L2 spelling increased from the beginning of the
first grade to the end of the fifth grade. By the fifth grade, L1 and L2 decoding
were strongly correlated (.77), as were L1 and L2 spelling (.68). The findings
provide support for the idea that language learning mechanisms—in this case,
for phonology—may be similar for L1 and L2.

Conclusions

In sum, the findings from this study suggest that students' early L1 skills are
strongly related to their L2 learning several years later and that L1 skills may be

an important source of individual differences among L2 learners. The present study is unique in several ways. In particular, the participants were followed over 10 years, so the possibility of long-term crosslinguistic transfer of L1 skills to L2 proficiency could be examined. However, the study also has limitations that restrict generalization of the findings. The small sample size limits the power of the statistical analyses. Additionally, the use of nonvalidated measures for assessing L2 oral and written proficiency, L2 word decoding, and L2 spelling was necessary because standardized instruments for measuring these skills in an L2 are unavailable. Attempts were made to construct L2 assessments that closely resembled the L1 measures in this study (e.g., multiple-choice formats for the L1 and L2 reading comprehension tests, written spelling for the L1 and L2 spelling tests, real words and pseudowords for the L1 and L2 decoding tests). Although the internal consistency of the L2 measures was checked, their construct or criterion-related validity was not determined. In addition, the participants studied three different L2s, each of which has a different orthography, and their scores were combined in the analyses. Combining the three different L2 groups may have served to blur unique distinctions among L1-L2 proficiency, decoding, and spelling in the three L2s.

There are several implications that can be drawn from this study. First, students' early L1 skills appear to play a role in individual differences for L2 learning even several years after students have mastered their L1. In recent years, L2 educators have proposed several hypotheses to account for students' L2 learning differences (e.g., high anxiety, low motivation, poor use of language learning strategies). However, the results of this study suggest that a large part of the differences on L2 proficiency measures displayed by L2 learners is likely to be explained by their L1 skills. L2 educators should consider the role that L1 plays in learning a L2.

Second, L1 and L2 learning may depend on basic language learning mechanisms that are similar to both languages, as predicted by the LCDH. The findings of this study suggest that theories of cognition that propose connections between basic L1-L2 language learning skills (e.g., syntax, morphology, phonology) may be beneficial in explaining individual differences in L2 learning and how students acquire L2 skills. L2 educators should consider theories of cognition that propose crosslinguistic interactions between L1 and L2 when seeking explanations for more and less successful L2 learning.

Third, there appear to be strong connections between students' L1 and L2s that have less orthographic distance. In this study, the students' L1 and the three L2s were alphabetic and displayed orthographic similarities. Thus, it may have been more likely that the participants, especially those with stronger

L1 skills, could make stronger connections between their L1 and the L2. L2 researchers could conduct studies to determine whether specific skills in the L1 (e.g., phonology) better predict performance in some L2s as opposed to others.

Fourth, students who have lower levels of L1 skills will likely have lower levels of L2 aptitude. Because L1 skills, including L1 literacy, and L2 aptitude (on the MLAT) have been found to be strong predictors of L2 proficiency, students with average L1 skills may find it more challenging to develop oral and written L2 proficiency several years after learning their L1. If so, L2 educators need to be aware of studies that show long term L1-L2 connections and develop effective teaching methods that take into account language-related variables in L2 learning.

Fifth, because both L2 aptitude and subsequent L2 proficiency appear to be strongly related to early L1 skills, it is important for all educators to know that early L1 language development prior to entering school is important for later L2 learning several years after students have mastered their L1 (e.g., see Sénéchal, 2006). Likewise, all L1 and L2 educators should understand that mastery of L1 literacy skills in the primary school years is important for students attempting to learn a L2 several years after learning to read their L1.

<div align="right">Revised version accepted 10 June 2008</div>

References

Aaron, P. G. (1989). Can reading disabilities be diagnosed without using intelligence tests? *Journal of Learning Disabilities*, *24*, 178–191.

American Council on the Teaching of Foreign Languages. (1986). *ACTFL proficiency guidelines*. Yonkers, NY: Author.

American Council on the Teaching of Foreign Languages (ACTFL). (1989). *American Council on the Teaching of Foreign Languages proficiency guidelines*. Hastings-on-Hudson, NY: Author.

Carroll, J. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training and research in education* (pp. 87–136). Englewood Cliffs, NJ: Prentice-Hall.

Carroll, J. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language learning. *International Journal of Psycholinguistics*, *2*, 5–14.

Carroll, J., & Sapon, S. (2000). *Modern Language Aptitude Test (MLAT): Manual*. Rockville, MD: Second Language Testing, Inc. (Original work published 1959)

Cisero, C., & Royer, J. (1995). The development of cross-language transfer of phonological awareness. *Contemporary Educational Psychology*, *20*, 275–303.

Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read a second language. *Journal of Educational Psychology*, *91*, 29–43.

CTB/McGraw Hill (1983). *Test of cognitive skills*. Monterey, CA: Author.

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

Cummins, J. (1984). Implications of bilingual proficiency for the education of minority language students. In P. Allen, M. Swain, & C. Brumfit (Eds.), *Language issues and education policies: Exploring Canada's multilingual resources*. Oxford: Pergamon Press.

Cunningham, A., & Stanovich, K. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*, 264–274.

Cunningham, A., & West, R. (2000). Literacy experiences and the shaping of cognition. In K. Stanovich (Ed.), *Progress in understanding reading: Scientific foundations and new frontiers* (pp. 309–319). New York: Guilford Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dufva, M., & Voeten, M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholinguistics*, *20*, 329–348.

Dunn, L., & Dunn, L. (1981). *Peabody picture vocabulary test-revised*. Circle Pines, MN: American Guidance.

Durgunoglu, A. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, *52*, 189–204.

Durgunoglu, A., Nagy, W., & Hancin-Bhatt, B. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology*, *85*, 453–465.

Ganschow, L., & Sparks, R. (2001). Learning difficulties and foreign language learning: A review of research and instruction. *Language Teaching*, *34*, 79–98.

Geva, E., & Verhoeven, L. (2000). Introduction: The development of second language reading in primary children: Research issues and trends. *Scientific Studies of Reading*, *4*, 261–266.

Geva, E., Yaghoub-Zadeh, Z., & Schuster, B. (2000). Understanding individual differences in word recognition skills of ESL children. *Annals of Dyslexia*, *50*, 121–154.

Hakuta, K. (1976). A case study of a Japanese child learning English as a second language. *Language Learning*, *26*, 321–351.

Hancin-Bhatt, B., & Bhatt, R. (1997). Optimal L2 syllables: Interactions of transfer and developmental effects. *Studies in Second Language Acquisition*, *19*, 331–378.

Hancin-Bhatt, B., & Nagy, W. (1994). Lexical transfer and second language morphological development. *Applied Psycholinguistics*, *15*, 289–310.

Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition*, *59*, 119–147.

Hulstijn, J., & Bossers, B. (1992). Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology*, *4*, 341–353.

Humes-Bartlo, M. (1989). Variation in children's ability to learn second languages. In K. Hyltenstam & L. Obler (Eds.), *Bilingualism across the life span* (pp. 41–54). Cambridge: Cambridge University Press.

Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms on a second language. *TESOL Quarterly*, *20*, 287–304.

Kahn-Horwitz, J., Shimron, J., & Sparks, R. (2005). Predicting foreign language reading achievement in elementary school students. *Reading and Writing: An Interdisciplinary Journal*, *18*, 527–558.

Kahn-Horwitz, J., Shimron, J., & Sparks, R. (2006). Weak and strong novice readers of English as a foreign language: Effects of first language and socioeconomic status. *Annals of Dyslexia*, *56*, 161–185.

Koda, K. (1999). Development of L2 intraword structural sensitivity and decoding skills. *Modern Language Journal*, *83*, 51–64.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.

Larsen, S., & Hammill, D. (1986). *Test of written spelling-2 (TWS-2)*. Austin, TX: PRO-ED.

Lindamood, L., & Lindamood, P. (1979). *Lindamood auditory conceptualization test*. Allen, TX: PRO-ED.

Lindsey, K., Manis, F., & Bailey, C. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology*, *95*, 482–494.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Meschyan, G., & Hernandez, A. (2002). Is native-language decoding skill related to second-language learning? *Journal of Educational Psychology*, *94*, 14–22.

Olshtain, E. (1983). Sociocultural competence and language transfer: The case of apology. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 232–249). Rowley, MA: Newbury House.

Olshtain, E., Shohamy, E., Kemp, J., & Chatow, R. (1990). Factors predicting success in EFL among culturally different learners. *Language Learning*, *40*, 23–44.

Pimsleur, P. (1968). Language aptitude testing. In A. Davies (Ed.), *Language testing symposium: A linguistic approach* (pp. 98–106). New York: Oxford University Press.

Proctor, C., August, D., Carlo, M., & Snow, C. (2006). The intriguing role of Spanish language vocabulary in predicting English reading comprehension. *Journal of Educational Psychology*, *98*, 159–169.

Rutherford, W. (1983). Language typology and language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 358–370). Rowley, MA: Newbury House.

Sasaki, Y. (1993). Paths of processing strategy transfer in learning Japanese and English as foreign languages: A competition model approach. *Studies in Second Language Acquisition*, *16*, 43–72.

Sasaki, Y. (1996). *Second language proficiency, foreign language aptitude, and intelligence: Quantitative and qualitative analyses*. New York: Peter Lang.

Sénéchal, M. (2006). Testing the Home Literacy Model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading*, *10*, 59–87.

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, *16*, 155–172.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P., & Ducroquet, L. (1988). *A comparison of first and foreign language ability* (Working Documents No. 8). London: ESOL Department, Institute of Education, London University.

Sparks, R. (1995). Examining the linguistic coding differences hypothesis to explain individual differences in foreign language learning. *Annals of Dyslexia*, *45*, 187–214.

Sparks, R. (2001). Foreign language learning problems of students classified as learning disabled and non-learning disabled: Is there a difference? *Topics in Language Disorders*, *21*, 38–54.

Sparks, R., & Artzer, M. (2000). Foreign language learning, hyperlexia, and early word recognition. *Annals of Dyslexia*, *50*, 189–211.

Sparks, R., Artzer, M., Patton, J., Ganschow, L., Miller, K., Hordubay, D., *et al*. (1998). Benefits of multisensory language instruction in Spanish for at-risk learners: A comparison study of high school Spanish students. *Annals of Dyslexia*, *48*, 239–270.

Sparks, R., & Ganschow, L. (1991). Foreign language learning difficulties: Affective or native language aptitude differences? *Modern Language Journal*, *75*, 3–16.

Sparks, R., & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning problems: Linking first and second language learning. *Modern Language Journal*, *77*, 289–302.

Sparks, R., & Ganschow, L. (1995). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *Modern Language Journal*, *79*, 235–244.

Sparks, R., Ganschow, L., Artzer, M., Siebenhar, D., Plageman, M., & Patton, J. (1998). Differences in native language skills, foreign language aptitude, and foreign language grades among high, average, and low proficiency learners: Two studies. *Language Testing*, *15*, 181–216.

Sparks, R., Ganschow, L., & Patton, J. (1995). Prediction of performance in first-year foreign language courses: Connections between native and foreign language learning. *Journal of Educational Psychology*, *87*, 638–655.

Sparks, R., Ganschow, L., & Patton, J. (2008). L1 and L2 literacy, L1 and L2 aptitude, and L2 affective variables as discriminators among high and low-achieving, LD, and ADHD L2 learners. In J. Kormos & E. Kontra (Eds.), *Language learners with special needs: An international perspective* (pp. 11–35). London: Multilingual Matters.

Sparks, R., Humbach, N., & Javorsky, J. (2008). Individual and longitudinal differences among high- and low-achieving, LD, and ADHD foreign language learners. *Learning and Individual Differences*, *18*, 29–43.

Sparks, R., Javorsky, J., Patton, J., & Ganschow, L. (1998). Principal components analysis of a test battery to predict proficiency in a foreign language. *Applied Language Learning*, *9*, 71–106.

Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, *56*, 129–160.

Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2008). Early first-language reading and spelling skills predict later second-language reading and spelling skills. *Journal of Educational Psychology*, *100*, 162–174.

Spolsky, B. (1989). *Conditions for second language learning*. Oxford: Oxford University Press.

Spolsky, B. (1995). Prognostication and language aptitude testing: 1925–1962. *Language Testing*, *12*, 321–340.

Stanovich, K. (1986). Mathew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.

Stanovich, K. (1993). Does reading make you smarter? Literacy and the development of verbal intelligence. In H. Reese (Ed.), *Advances in child development and behavior* (pp. 133–180). San Diego, CA: Academic Press.

Stanovich, K. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford.

Stanovich, K., & Cunningham, A. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, *85*, 211–229.

van Gelderen, A., Schoonen, R., Stoel, R., Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*, 477–491.

Verhoeven, L. (1994). Transfer in bilingual development: The Linguistic Interdependence Hypothesis revisited. *Language Learning*, *44*, 381–415.

Wells, C. (1985). *Language development in the pre-school years*. Cambridge: Cambridge University Press.

Wesche, M., Edwards, H., & Wells, W. (1982). Foreign language aptitude and
   intelligence. *Applied Psycholinguistics*, *3*, 127–140.

Wiederholt, L. (1986). *Formal reading inventory*. Austin, TX: PRO-ED.

Woodcock, R. (1987). *Woodcock reading mastery test-revised*. Circle Pines, MN:
   American Guidance.

Appendix A: Lists and Descriptions of the L1 Achievement Measures and the L2 Aptitude Test

Formal Reading Inventory, Forms A and B (FRI)

This standardized test consists of a series of graded stories that measure the ability to read and answer multiple-choice comprehension questions in an untimed format.

Listening Comprehension

This standardized test consists of reading aloud to the student passages from the Woodcock Reading Mastery Test-Revised, Passage Comprehension subtest, and asking the student to identify a key word missing from the passage. The student is not permitted to see (read) the passage, but the passages can be repeated. To measure listening comprehension, the "alternate" form of the Woodcock was administered each year (i.e., when a student was administered Form G of the Woodcock to measure reading skill, s/he was administered Form H of the Woodcock to measure listening comprehension, and vice versa).

Modern Language Aptitude Test–Long Form (MLAT)

This standardized test measures foreign language aptitude with a simulated format to provide an indication of the probable degree of success in learning an L2. The Long Form includes all five subtests. The subtests are as follows:

> Part I–Number Learning. Student learns numbers of a made-up language and then transcribes spoken number words into written digits on hearing them presented rapidly.
>
> Part II–Phonetic Script. Student listens to a sequence of syllables (many with no meaning in English) while looking at their graphemic transcriptions and is asked to quickly learn how the sounds (phonemes) correspond to the letters (graphemes).
>
> Part III–Spelling Clues. Student reads English words presented as abbreviated spellings (e.g., luv) and then chooses the one word (out of five) that corresponds most nearly in meaning (e.g., carry, exist, affection, wash, spy).

Part IV–Words in Sentences. Student reads a key sentence in which a word is underlined, reads another sentence in which five words and phrases are marked as possible choices, and chooses the word or phrase in the second sentence that has the same grammatical function that the marked word or phrase has in the key sentence.

Part V–Paired Associates. Student studies a list of nonsense words with assigned English meanings and chooses the correct English word from a multiple choice format to match the nonsense word.

Peabody Picture Vocabulary Test-Revised, Forms L and M

This standardized test measures receptive vocabulary for standard American English.

Test of Cognitive Skills

This standardized test is designed to assess a student's academic aptitude and predict the student's level of success in school. The test consists of four subtests: *Sequences* tests the ability to recognize a rule or principle in a pattern or sequence of figures, letters, or numbers; *Analogies* tests the ability to recognize relationships that may be literal or symbolic; *Memory* tests the ability to recall previously presented material; and *Verbal Reasoning* tests the ability to reason logically and discern relationships between pictures or words. All test items are presented orally and no reading or writing is necessary to complete the items.

Test of Written Spelling-2 (TWS)

This standardized test measures performance on writing single words from dictation.

Woodcock Reading Mastery Test-Revised, Forms G and H (WRMT-R)

This standardized test consists of four subtests: *Word Identification* tests the ability to read isolated words (e.g., urgent, hysterical, causation, heterogeneous); *Word Attack* tests the ability to read (pseudo) nonsense words (e.g., dee, poe, vunhip, mancingful); *Word Comprehension* tests the ability to read words and provide an antonym or synonym and to read and compete an analogy; and *Passage Comprehension* tests the ability to read a short passage and then identify a key word missing from the passage. The Total Test score is a combination of the four subtests.

Appendix B: Lists of Real Words and Pseudowords for Spanish, French, and German

Spanish		French		German	
Real Words	Pseudowords	Real Words	Pseudowords	Real Words	Pseudowords
anoche	loche	adorable	elars	frei	zogen
enero	regua	midi	trousante	schier	angesammen
isla	traceo	robe	etez	nachbar	biedel
orilla	placeta	chaise	pentants	bitte	leiner
usted	sucrete	difficile	tonseur	bücher	nochtbar
mesa	popeta	croix	pateau	einwand	auwinnern
señora	porrosca	ouest	gaissons	todesurteile	sotürlich
jefa	asurge	soeur	pretons	pflicht	zeiben
entretenamiento	hastilla	magnifique	els	jawohl	speulen
salón	movadiza	aéroport	lace	deshalb	möffel
inventado	vestuto	soleil	desentrez	ausgezeichnet	besuchgekammen
mural	cantón	Joël	tadelle	fussboden	kössen
la amada	calahoria	longtemps	ponte	leute	weschen
agencia	meradario	pluie	exploter	flugzeug	hilgescher
filólogo	zebajo	médicament	d'assoint	besonders	hols
repentinamente	pantaora	rançon	soulangerie	walstatt	wangst
antena	cebaduría	hier	trençons	geschlecht	vertaunen
corriente	grallanado	décidé	boignon	räuber	mautern
alto	llenosidad	nationalité	ramposé	rauchen	leisling
preocupación	yagüe	réveillon	touvertes	unbezähmbar	dömpferbehn

Appendix C: List of Spelling Words for Spanish, French, and German

Spanish	French	German
bien	faire	zeit
arpa	comme	handschuh
dulce	maison	natürlich
gafas	cher	lederhosen
litro	vraiment	kuchen
maíz	voudrais	bücher
pago	étoile	verbinden
después	pendant	hoffentlich
secreto	visiter	wechsel
paisaje	poivre	bahnhof
placita	déjeuner	bienenstich
tranquilo	choisir	pferderei
mirando	pourrais	gefallen
debajo	adorable	schmeckt
horarios	printemps	frucht
tagarote	chanteuse	gewöhnlich
periódico	raconteur	besuchen
felicidad	magnifique	schule
zanahoria	chaise	obst
abecedario	ouvrir	mittagessen

Appendix D: Means and Standard Deviations of High-, Average-, and Low-Proficiency Groups on the L1 Achievement and Cognitive Ability Measures

Testing measure	PRE						POST 1						POST 2						POST 3						POST 4					
	High		Ave		Low		High		Ave		Low		High		Ave		Low		High		Ave		Low		High		Ave		Low	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Woodcock Reading Mastery Test Basic Skills[a]	102.9	15.4	97.4	14.2	89.1	13.3	109.5	13.5	104.0	13.2	95.6	18.3	114.1	9.2	101.3	11.9	95.2	17.5	108.7	7.9	101.4	12.6	94.5	12.6	112.3	9.4	103.4	11.4	95.3	13.2
Test of Written Spelling[a]	83.9	9.6	80.0	12.6	73.6	8.5	102.4	11.0	97.6	14.9	91.0	10.6	107.7	7.9	101.0	11.8	94.9	12.3	111.5	8.5	100.4	12.4	93.0	15.3	112.5	7.1	98.4	12.0	94.4	14.1
Peabody Picture Vocabulary Test-Revised[a]	109.2	19.1	106.4	11.3	98.9	12.1	113.8	9.9	103.4	14.1	99.6	13.3	114.6	15.6	103.4	11.3	97.9	10.2	117.9	12.9	107.6	8.9	101.1	14.7	117.5	11.1	111.2	12.5	102.6	14.0
Formal Reading Inventory[a]	93.5	19.8	86.4	15.8	80.4	10.9	110.3	11.9	100.2	11.9	98.1	17.5	113.5	15.6	104.4	17.4	89.6	17.5	109.5	16.6	106.3	15.9	90.8	19.6	119.0	13.9	108.0	12.3	94.3	13.5
Lindamood Auditory Conceptualization Test[b]	0.5	0.9	−.07	1.0	−.19	1.2	0.4	0.9	−0.0	1.0	−.06	1.2	0.5	0.8	−.12	0.9	−.07	1.3	0.6	0.7	−.15	.09	−.15	1.2	—	—	—	—	—	—
Listening Comprehension[a]	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	105.8	11.5	104.0	11.0	96.9	8.5	110.8	10.0	106.1	10.4	98.3	15.5
Test of Cognitive Skills[a]	110.7	10.8	104.4	13.7	99.6	15.7	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

[a]Standard Scores, M = 100, SD = 15.
[b]z-Scores.

Appendix E: Means and Standard Deviations of High-, Average-, and Low-Proficiency Groups on the L2 Proficiency, L2 Aptitude, L2 Word Decoding, and L2 Spelling Measures and L2 Course Grades

Testing measure	HIGH PROF (n = 13)		AVE PROF (n = 27)		LOW PROF (n = 14)	
	M	SD	M	SD	M	SD
L2 Proficiency[a] (Read Comprehension, Writing, Listening/Speaking)	41.8	8.4	23.9	4.8	12.2	5.0
Modern Language Aptitude Test[b]	113.8	9.5	100.7	13.3	89.5	13.5
L2 word decoding[c]	56.3	9.3	44.9	11.7	39.9	12.0
L2 spelling[d]	12.3	2.8	8.0	3.3	7.1	2.6
L2 grade–Year 1[e]	3.9	0.5	3.1	0.9	2.4	0.8
L2 grade–Year 2[e]	3.8	0.5	2.8	1.1	1.8	0.8

[a]Raw score, maximum score = 70.
[b]Standard scores, M = 100, SD = 15.
[c]Raw score, maximum score = 80.
[d]Raw score, maximum score = 20.
[e]Based on a scale ranging (in 0.33 increments) from F = 0.00 to A+ = 4.33.

Appendix F: Intercorrelations Among PRE L1 Variables, MLAT, L2 Proficiency, L2 Grades, and L2 Word Decoding and Spelling

Testing measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Woodcock Reading Mastery Test Basic Skills		.80**	.55**	.28*	.44**	.46**	.57**	.37**	.28*	.22	.54**	.48**
2. Test of Written Spelling			.65**	.34**	.49**	.49**	.61**	.40**	.41**	.38**	.53**	.59**
3. Formal Reading Inventory				.35**	.33**	.37**	.41**	.38**	.19	.26	.30*	.44**
4. Peabody Picture Vocabulary Test					.46**	.46**	.46**	.26	.18	.21	.24	.33**
5. Lindamood Auditory Conceptualization Test						.60**	.56**	.40**	.34**	.23	.52**	.60**
6. Test of Cognitive Skills							.64**	.42**	.51**	.42**	.30*	.49**
7. Modern Language Aptitude Test								.68**	.61**	.62**	.66**	.72**
8. L2 proficiency									.65**	.70**	.55**	.67**
9. L2 grade–Year 1										.86**	.59**	.55**
10. L2 grade–Year 2											.47**	.56**
11. L2 word decoding												.66**
12. L2 spelling												

*$p < .05$.
**$p < .01$.

Appendix G: Intercorrelations Among POST 1 L1 Variables, MLAT, L2 Proficiency, L2 Grades, and L2 Word Decoding and Spelling

Testing measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Woodcock Reading Mastery Test Basic Skills		.77**	.61**	.34*	.45**	.53**	.60**	.24	.27*	.20	.57**	.50**
2. Test of Written Spelling			.57**	.37**	.50**	.54**	.63**	.40**	.38**	.39**	.57**	.57**
3. Formal Reading Inventory				.55**	.57**	.40**	.48**	.30*	.41**	.30*	.41**	.33*
4. Peabody Picture Vocabulary Test					.45**	.11	.53**	.38**	.46**	.36**	.40**	.28*
5. Lindamood Auditory Conceptualization Test						.44**	.47**	.28*	.24	.24	.46**	.49**
6. Test of Cognitive Skills												
7. Modern Language Aptitude Test												
8. L2 proficiency												
9. L2 grade–Year 1												
10. L2 grade–Year 2												
11. L2 word decoding												
12. L2 spelling												

*p < .05.
**p < .01.

Appendix H: Intercorrelations Among POST 2 L1 Variables, MLAT, L2 Proficiency, L2 Grades, and L2 Word Decoding and Spelling

Testing measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Woodcock Reading Mastery Test Basic Skills												
2. Test of Written Spelling	.83**											
3. Formal Reading Inventory	.55**	.53**										
4. Peabody Picture Vocabulary Test	.44**	.45**	.29*									
5. Lindamood Auditory Conceptualization Test	.54**	.53**	.39**	.43**								
6. Test of Cognitive Skills	.58**	.45**	.42**	.28*	.49**							
7. Modern Language Aptitude Test	.73**	.74**	.64**	.53**	.51**							
8. L2 proficiency	.56**	.49**	.55**	.45**	.37**							
9. L2 grade–Year 1	.44**	.37**	.49**	.22	.28*							
10. L2 grade–Year 2	.38**	.35**	.41**	.27*	.35**							
11. L2 word decoding	.74**	.62**	.51**	.39**	.39**							
12. L2 spelling	.45**	.61**	.50**	.38**	.48**							

*$p < .05$.
**$p < .01$.

Appendix I: Intercorrelations Among POST 3 L1 Variables, MLAT, L2 Proficiency, L2 Grades, and L2 Word Decoding and Spelling

Testing measure	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Woodcock Reading Mastery Test Basic Skills		.91**	.50**	.55*	.43**	.47**	.56**	.77**	.51**	.38**	.31*	.76**	.60**
2. Test of Written Spelling			.45**	.57**	.32*	.56**	.59**	.81**	.58**	.49**	.44**	.70**	.68**
3. Formal Reading Inventory				.49**	.28*	.32*	.19	.45**	.20	.31*	.33*	.44**	.31*
4. Peabody Picture Vocabulary Test					.48**	.46**	.30*	.62**	.46**	.37**	.34**	.44**	.35**
5. L1 Listening Comprehension						.45**	.21	.43**	.33*	.25	.16	.34**	.22
6. Lindamood Auditory Conceptualization Test							.49**	.58**	.39**	.21	.30*	.44**	.53**
7. Test of Cognitive Skills													
8. Modern Language Aptitude Test													
9. L2 proficiency													
10. L2 grade–Year 1													
11. L2 grade–Year 2													
12. L2 word decoding													
13. L2 spelling													

*p < .05.
**p < .01.

Appendix J: Intercorrelations Among POST 4 L1 Variables, MLAT, L2 Proficiency, L2 Grades, and L2 Word Decoding and Spelling

Testing measure	1	2	3	4	5	6	7	8	9	10	11	12
1. Woodcock Reading Mastery Test Basic Skills												
2. Test of Written Spelling	.85**											
3. Formal Reading Inventory	.51*	.59**										
4. Peabody Picture Vocabulary Test	.59**	.56**	.50**									
5. L1 Listening Comprehension	.62**	.62**	.60**	.56**								
6. Test of Cognitive Skills	.47**	.45**	.44**	.58**	.53**							
7. Modern Language Aptitude Test	.72**	.76**	.60**	.67**	.65**							
8. L2 proficiency	.53**	.60**	.46**	.46**	.41**							
9. L2 grade–Year 1	.49**	.45**	.44**	.35**	.48**							
10. L2 grade–Year 2	.44*	.41**	.45**	.40**	.48**							
11. L2 word decoding	.77**	.73**	.45**	.45**	.55**							
12. L2 spelling	.69**	.68**	.53**	.52**	.48**							

*$p < .05$.
**$p < .01$.