The Spanish Developmental Contrastive Spelling Test: An Instrument for Investigating Intra-Linguistic and Crosslinguistic Influences on Spanish-Spelling Development

Igone Arteagoitia, Elizabeth R. Howard, Mohammed Loguit,
Valerie Malabonga, and Dorry M. Kenyon
Center for Applied Linguistics

Abstract

This article describes the development of a Spanish-spelling measure designed to assess the progress made by Spanish-English bilingual children from Grade 2 to Grade 5. Different stages of measure development are described, such as the selection of the focus features, the pilot phase of the assessment, and the finalizing of the operational version. Two underlying attributes characterize the spelling measure described here. First, it is *developmental*, meaning that it contains a wide variety of features and items that differ according to spelling difficulty, such that the assessment is able to measure the growth of Spanish-spelling ability over the full sequence of the upper elementary grades. Second, it is *contrastive*, as it was designed to detect some areas of potential crosslinguistic influence from English to Spanish. The combination of these two characteristics makes this spelling measure a unique tool for assessing the development of spelling ability by Spanish-English bilingual children.

Introduction

While there has been considerable research on bilingualism in the United States in the past 30 years, most studies have focused on the effects of bilingual education on the development of the English language only, and thus present a partial picture of the knowledge that bilingual children possess (see, for example, Oller & Eilers, 2002). This has led to conclusions about English language learners' (ELLs) linguistic ability being reached based solely

on their performance on English tests, without considering what has been called the "distributed characteristic of bilingual knowledge" (Oller & Eilers, p. 10), namely the fact that a bilingual speaker may have knowledge of some words (or other linguistic units) in one language, but not in the other and vice versa. Thus, lower performance of bilingual children on tests normed on monolinguals could imply a distribution of knowledge across the two languages rather than a general linguistic deficiency. In order to attain a more accurate picture of ELLs' overall linguistic ability, it is necessary to assess their bilingual knowledge.

The importance of assessing bilingual children's knowledge of their native language (L1) is supported by research that suggests that a strong foundation in the L1 may, in the long term, facilitate second-language (L2) development (Cummins, 1984; Medina & Escamilla, 1992; Oller & Eilers, 2002; Ramírez, Yuen, & Ramey, 1991; Thomas & Collier, 2002). For example, one of the main conclusions that can be drawn from Thomas and Collier's study, which documents the academic achievement of ELLs over the long term (4–12 years) and across content areas, is that the amount of formal L1 instruction that a student has received is the strongest predictor of second-language student achievement. That is, the greater the number of years of L1 grade-level schooling a student has received, the higher his or her English achievement. In fact, the only educational programs that were found to prepare ELLs to fully reach the 50th percentile in both their L1 and English in all subject areas and to maintain that level of high achievement were two-way immersion bilingual programs and one-way developmental bilingual programs. Both of them are additive bilingual programs that provide ongoing instruction through both the L1 and L2 for at least 5 or 6 years (Genesee, 1999). Given the relationship between strong L1 skills and L2 development found in the literature, there is a clear need to evaluate ELLs' L1 ability in addition to their L2 ability.

One area of literacy development that has been found to be important for predicting reading skills is the ability to spell words correctly (Adams, 1995; Ehri & Wilce, 1987; Ferroli & Krajenta, 1989; Schlagal, 1992). Research that has examined the relationship between early reading and spelling has generally found that spelling ability predicts reading ability even after phonological awareness has been controlled for, and this seems to be true in both English and Spanish (Cunningham & Stanovich, 1990; Ferroli & Krajenta, 1989; Stage & Wagner, 1992).

The Spanish language possesses a very transparent orthographic system with a high degree of correspondence between phonemes and graphemes. This contrasts with English where multiple representations of the same sound are the norm. Thus, for example, while there are at least three different ways to represent /i:/ in English (<ea> as in *neat*, <ee> as in *need*, and with the silent <e> as in *cede*), there is only one sound per vowel in Spanish. However, in spite of this transparency, some research suggests that even high school

Spanish-speaking students consistently make spelling mistakes in their L1, and thus have yet to master some aspects of the Spanish-spelling system (Carbonell de Grompone, Tuana, Piedra de Moratoria, Lluch de Pintos, & Corbo de Mandracho, 1980).

Carbonell de Grompone et al. (1980) identified three phases in the spelling development of native Spanish speakers. The first phase is characterized by unique or one-to-one phoneme-grapheme mappings, such as the vowels, except for /i/ which can be represented as <i> or <y> in some contexts, and many of the consonants in Spanish. The next phase involves the representation of rule-governed phonemes whose graphemic realization varies depending on the context in which they appear (e.g., the trill is represented as <r> at the beginning of a word or after a consonant, as in *ropa* or *honra*, but as <rr> between vowels, as in *carro*). The last phase includes the spelling of phonemes that have more than one representation, which, while not as common as in English, does happen with a few Spanish sounds, such as the bilabial voiced stop /b/, which has two graphemic representations, <b> and <v>, as in *boca* and *vaca*.

Very little research has been devoted to the development of spelling ability in Spanish by Spanish-English bilingual children. This is unfortunate, because when studied in conjunction with the development of English-spelling skills, the development of spelling ability in the L1 makes it possible to disentangle general spelling or literacy issues from issues of L2 proficiency. That is, if a Spanish-English bilingual student is found to have poor literacy skills in English but strong skills in similar domains in Spanish, then it can be determined that the student's difficulties are more reflective of limited L2 proficiency rather than a generalized difficulty with literacy development. If, however, the student struggles in both languages, then it is more likely the case that more general difficulties related to literacy development in either language are at play. Additionally, having L1 and L2 spelling data from the same children will make it possible to find out whether spelling skills transfer from one language to the other, or whether the different orthographic systems of the two languages involved interfere with this transfer of skills.

As mentioned earlier, there is a lack of research on the development of L1 ability by bilingual children. Specifically, the number of studies that have investigated the development of L1 spelling ability by Spanish-English bilingual children to date is minimal. A few studies (e.g., Ferroli, 1991; Staczek & Aid, 1981) have focused on Spanish features that are indicative of phonological and orthographic differences between the two languages. However, these studies present a very constrained picture of bilingual children's spelling ability in Spanish by either focusing on a small set of features (Ferroli), or by providing a mere qualitative account of the types of spelling errors Spanish-English bilingual children make in Spanish (Staczek & Aid). Ferroli's study looked at how second- and third-grade bilingual children in a Spanish-

English transitional bilingual education program in a suburban Midwestern public elementary school spelled a small set of features in Spanish and English in order to find out whether Spanish-spelling knowledge interfered with learning English letter-sound associations. Findings from this study indicate that in general misspellings in English that can be attributed to Spanish influence were not produced by children whose Spanish-spelling knowledge was high but rather by those whose English-spelling knowledge was low, which suggests that Spanish-spelling knowledge did not interfere with English spelling. This finding, together with the finding that "children who progressed beyond a letter-name strategy in Spanish transferred that conceptual understanding of spelling to English" (Ferroli, p. 64), seems to provide some evidence for the transfer of spelling skills from Spanish to English.

Staczek and Aid (1981) conducted a qualitative analysis of the spelling errors found in writing samples by native Spanish-speaking students of varying ages (from elementary school to college) enrolled in a range of educational programs (including full and transitional bilingual programs) in the Miami area. They identified three types of errors: (a) intra-linguistic errors associated with the complexities of Spanish language, (b) crosslinguistic errors derived from English influence, and (c) others, which seem to be a result of ignoring the orthographic conventions in both languages. While their description of nonstandard orthographic patterns provides some useful information about the types of spelling errors produced by Spanish-English bilinguals, the fact that they do not include a control monolingual Spanish-speaking group to validate their classification of errors as crosslinguistic is problematic, as it is possible that monolingual Spanish speakers may make the same error. Additionally, findings are not presented by grade level, which makes it impossible to assess potential developmental trends.

Two studies document Spanish-spelling measures that have been created for use with Spanish-English bilingual students in the elementary grades (Estes & Richards, 2002; Ferroli & Krajenta, 1989). In both cases, the focus was on creating a measure designed to capture the Spanish-spelling development of Spanish-English bilingual children, so the measures only included spelling features that were considered difficult from an intra-linguistic standpoint. Contrastive features that would yield information about potential crosslinguistic influences in the Spanish spelling of Spanish-English bilinguals were not included. Moreover, there was no control group of monolingual Spanish speakers included in either study, so it is not possible to determine whether or not the features that were considered difficult from an intra-linguistic standpoint were in fact difficult for Spanish monolinguals. Finally, these measures focused on emergent spelling issues of students in the primary grades and did not allow for the tracking of student performance through the upper elementary grades.

Following this line of research and expanding upon it, the present paper attempts to make a contribution to the literature on the development of spelling ability by Spanish-English bilingual children by presenting a Spanish-spelling measure that is both developmental and contrastive. The measure is developmental in that it includes a variety of features that differ widely in terms of their spelling difficulty, as do the individual words within each feature category. Both of these factors make it possible to evaluate the progress children make from Grade 2 to Grade 5. Moreover, the measure is also contrastive, meaning that it was designed to detect some areas of potential crosslinguistic influence across English and Spanish for students who are bilingual in these two languages.

The rest of the article will be devoted exclusively to the development of the measure from its inception to its operational version. Future publications will present data collected through the use of this measure.

Development of the Measure

The Spanish Developmental Contrastive Spelling Test is a group-administered dictation assessment that was developed for a longitudinal spelling study of Spanish-English bilingual students in the upper elementary grades. The assessment was developed following Bachman and Palmer's (1996) framework, which provides the key issues to take into account when designing and developing language tests, such as the usefulness of the test, the construct being measured, the task being used, how the test is scored, and the reliability of the test.

Three different stages were involved in the development of Spanish Developmental Contrastive Spelling Test: (a) initial exploratory work, (b) piloting, and (c) developing the final measure. Each one of these stages will be discussed, so as to provide the reader with some background information on the development of the assessment instrument.

Exploratory Work

The purpose of the exploratory work was to generate focus features (e.g., the graphemes <r> and <rr> to represent the trill, <b> and <v> to represent the stop bilabial phoneme, <ñ>, etc.) for the spelling measures, which would later be used in the selection of items to include on the measures. Two sources were mined in the generation of focus features. First, preliminary analyses were conducted on the spelling errors found in the writing samples of a stratified random sample of 20 students in Grades 3–5 who participated in the Center for Applied Linguistics (CAL)/Center for Research on Education, Diversity & Excellence (CREDE) Study of Two-Way Immersion Education (Howard, Christian, & Genesee, 2004). Stratification was based on L1 (Spanish or English) as well as English-spelling. The fact that the sample in the CAL/CREDE study

was similar to that of the present study in terms of participants' age, L1, and educational program in which they were enrolled (in fact, some of the same schools participated in both studies) made these data a good starting point for the generation of focus features for the spelling measures. In addition to analyzing empirical data, information gathered from the review of relevant literature discussed earlier was taken into account in the selection of target features for the spelling test. The features that resulted from this exploratory work are presented in Table 1.

Using Waters, Bruck, and Malus-Abramowitz's (1988) classification of words in terms of their spelling difficulty and expanding upon it, features were categorized as: (a) regular words with various patterns or words that have more than one possible representation given the spelling patterns of the language, as in the bilabial stop, which can be represented by the letters *b* or *v* (e.g., *árboles* and *ventanilla*); (b) orthographic words, which adhere to the orthographic conventions of the language, such as the need to add a *u* after *g* to preserve the velar stop sound (e.g., *jugo* vs. *jugué*); and (c) contrastive words or words that contain phonemes that do not exist in English (e.g., the grapheme <ñ> in *sueño*).

Because the study focuses on both crosslinguistic and intra-linguistic sources of spelling errors, the features include both of the following: (a) those that are more likely to be susceptible to transfer from English, such as the vowel sounds, especially the representation of the sound /i/ with the grapheme <e>, or the sound /e/ with the grapheme <a>, which are supposedly "based on the association formed in the English alphabetical system" (Staczek & Aid, 1981, p. 153; see also, Serrano & Howard, 2003); and (b) those that seem to be difficult to learn in general, such as the regular words with various patterns (Carbonell de Grompone et al., 1980; Cuetos, 1993; Morais, 1998; Valle-Arroyo, 1990).

Table 1

Real-Word Features Resulting From Exploratory Work

Feature	Representation	Sample item
Regular	j/g; b/v; ll/y; h; s/c; s/z	gente, árboles, ayuda, hospital, cebolla, cabeza
Orthographic	r/rr; g/gu; c/qu; m before bilabial stop (b or p)	corrí, llegué, quince, ambos
Contrastive	ñ	sueño, compañía

Pilot Study

   The purpose of the pilot study was twofold: (a) to get a preliminary sense of the functionality of the features and items selected through the exploratory analyses, and (b) to determine the reliability and overall quality of the Spanish Developmental Contrastive Spelling Test. One hundred and ninety-six native Spanish-speaking and native English-speaking students in Spanish-English two-way immersion bilingual programs in Grades 2–5 participated in the pilot study. Participants were enrolled in two types of educational programs: a Spanish-dominant (90/10) two-way immersion program ($n = 107$), and a balanced (50/50) two-way immersion program ($n = 89$). In the 90/10 model, 90% of the instruction in the first year or two is in the minority language and 10% in English, and over the course of the primary grades, the percentage of instruction in the minority language decreases, while the percentage of instruction in English gradually increases, reaching a 50/50 instructional ratio by about fourth grade. In a 50/50 model, half of the instruction is in English and half is in the minority language at all grade levels. In each school, all of the students in Grades 2–5 who received parental consent and had been in the program since at least first grade participated in the pilot study (see Table 2 for the distribution of participants by grade level and educational program).

   The term *bilingual* will be used to refer to this group of children, namely, native Spanish-speaking and native English-speaking students enrolled in two-way immersion programs, even though together they represent a range of proficiency levels in both languages. The vast majority of the native English speakers were English-dominant, that is, they came from English monolingual families and their first exposure to Spanish took place in school, and the vast majority of the native Spanish speakers were already balanced bilinguals by

Table 2

Sample by Grade Level Across Educational Program

Grade	Educational program		Total
	90/10	50/50	
2	21	24	45
3	35	20	55
4	34	29	63
5	17	16	33
Total	107	89	196

Grade 2, if not earlier, since they were born in the United States and exposed to English from an earlier age through the community, media, daycare, and preschool, etc.

The two schools where the pilot study was conducted were chosen for the following reasons: (a) They were veteran two-way immersion programs and thus provided good model fidelity; (b) each represented one of the two models of two-way immersion education that would participate in the actual study (90/10 and 50/50); and (c) together they represented a wide range of national origins in their student body (at least as indicated by parental nationality, since most of the study participants were born in the United States). Additionally, students who participated in the study included gifted and talented and limited English proficiency students as well as special education students. Given all the conditions described above, it is reasonable to believe that the sample used was adequate for the purposes of the pilot, which was the development of a Spanish-spelling test that would be appropriate for the range of students enrolled in two-way immersion programs. The measure was also piloted in Mexico with a group of children from Grades 2 through 5, and was found to be adequate for showing a general developmental trend in Spanish-spelling ability for these Mexican children as well. Findings from the pilot test study have resulted in a slightly different operational version (see the section Finalizing the Assessment), which is being used with a larger, but comparable sample, to show how they perform over the course of 4 years.

The pilot version contained 60 items and was divided into two parts: a 40 item real-word component that used a traditional dictation format in which participants hear a word first in isolation, then in context, then a third time in isolation again, and are asked to write it down in its entirety; and a 20 item pseudo-word component that used a cloze format in which participants are provided with partial spellings of pseudo-words, which they then must complete with the appropriate letter or letters after hearing the word.

Because one of the main goals of the study was to investigate the ability of Spanish-English bilingual children to map phonemes and graphemes in Spanish, it was decided to use pseudo-words in addition to real-words, since the former would make it possible to tease out children's ability to map phonemes and graphemes in situations where they had no prior exposure and could not rely on visual memory for spelling. Thus, a separate component with 20 pseudo-words was developed. Of the three categories used in the real-word component of the test, namely, regular words with various patterns, orthographic words, and contrastive words, only the last two were included in the pseudo-word section (see Table 3 for the subset of features used in the pseudo-word component of the test). This was done because in the case of invented words, all of the possible spellings of an item with various patterns would have to be considered as valid spellings (e.g., *noyada* and *nollada*),

Table 3

Pseudo-Word Features Resulting From Exploratory Work

Feature	Representation	Sample item
Orthographic	r/rr; g/gu; c/qu; m before bilabial stop (b or p)	tirrapo, guensa
Contrastive	ñ	beñal, triño

and this would make these items very easy to spell, when in fact, regular words with various patterns are difficult to spell in Spanish (see, for example, Carbonell de Grompone et al., 1980).

The dictation format was chosen for the real-word test because it made it possible to test more than one feature per test item, while maintaining a representative sample of each feature (a minimum of four items per feature) without having to create a test that was too long for elementary school students, especially those in the lower grade levels. For the pseudo-words, a cloze dictation format was thought to be more appropriate, since this format would reduce the cognitive load of the task and allow for greater precision in testing the targeted feature in each item.

The pilot items originated from the following sources: (a) spelling errors found in the writing samples used for the exploratory work, (b) age-appropriate children's literature books, (c) graded language arts textbooks, such as Hampton-Brown's Spanish Grammar Book series *¡Bien dicho!* (2001), and (d) University of Puerto Rico Superior Educational Council's (1952) word frequency Spanish book with a corpus of over 7 million words from 10 different sources that include ample samples of children and adult language. The words varied widely in terms of frequency of occurrence (range = 2–7,541 per million words), with a mean frequency of 1,188. Because the test was still in the pilot stage, a few extra items were included so as to allow for the possibility of eliminating any misfitting items that were discovered through the analyses of the pilot data. Randomly dividing the total pool of items into five sets made it possible to create four versions of the assessment (one per grade level) without compromising the length of the test. Thus, 50 real-words were selected and further divided into 5 sets of 10. A common core of 10 words was tested across all four grade levels, while varying combinations of 3 of the remaining 4 sets of 10 words were used to create the specific pilot assessment at each grade level (see Table 4 for a distribution of the 50 real-words across grade levels).

Similarly, 25 pseudo-words were randomly divided into 5 sets of 5 words, with a common core of 5 words tested across all grade levels, and different combinations of 3 of the remaining sets of 5 words used to create the particular test at each grade level. Constancy among versions was maintained by making

Table 4

Pilot Test Item Distribution Across Grade Levels (Real-Words)

Grade	Item	Item	Item	Item
2	1–10	11–20	21–30	31–40
3	1–10	11–20	21–30	41–50
4	1–10	11–20	31–40	41–50
5	1–10	21–30	31–40	41–50

Table 5

Pilot Test Item Distribution Across Grade Levels (Pseudo-Words)

Grade	Item	Item	Item	Item
2	1–5	6–10	11–15	16–20
3	1–5	6–10	11–15	21–25
4	1–5	6–10	16–20	21–25
5	1–5	11–15	16–20	21–25

sure that the composition of the four versions was equivalent in terms of relative hypothesized difficulty of the items and features being tested (see Table 5 for a distribution of the 25 pseudo-words across grade levels).

The test was group-administered to intact classes by a trained native Spanish-speaking researcher, who began the administration with four practice items. For the real-word component, she dictated each word to the whole class, read a sentence that contained the word, and finally repeated the word one more time (or as many times as needed), while the participants wrote down each word in its entirety on a blank space. For the second component, the pseudo-words were dictated (without accompanying sentences) and were repeated as many times as needed, while participants completed each item with the missing letter or letters (see Figures 1 and 2 for sample items of the real-word and pseudo-word components, respectively).

Real-word test items were scored in three different ways: accuracy at the whole-word level, accuracy at the feature level, and crosslinguistic influence at the whole-word level. Pseudo-words, on the other hand, were only scored at the feature level, both for accuracy and for crosslinguistic influence, since participants did not have to produce the whole item, but only the part that

This is what the Test Administrator says:

1. gente	Toda la gente necesita agua para vivir.	gente
2. hijos	Los padres deben hablar con sus hijos a diario.	hijos

This is what the Answer Sheet looks like:

1. _____	2. _____

Figure 1. Sample of real-word component of pilot test.

This is what the Test Administrator says:
    1. teñar
    2. tagué
This is what the Answer Sheet looks like:
    1. te _____
    2. ta _____

Figure 2. Sample of pseudo-word component of pilot test.

contained the feature. In this paper, we will present the analyses for accuracy at the whole-word level for real-words and accuracy at the feature level for pseudo-words.

As mentioned above, participants only took 40 out of the 50 real-words and 20 out of the 25 pseudo-words, and the particular items completed differed slightly across all test versions depending on the grade of the participants, with the exception of 10 common real-words and 5 common pseudo-words. Thus, since it was not possible to aggregate scores or compare the scores of participants in different grade levels using raw scores, it was necessary to create scaled scores based on the difficulty of the items and the ability of the participants.

Item analysis at the whole-word level using the Rasch model was performed on the scores of all participants on all 50 items (real-words). Participants with perfect or zero scores were automatically removed from the Rasch analysis by the Winsteps program (Linacre & Wright, 2000), leaving a total sample of 195. The Rasch analysis provided information on the difficulty
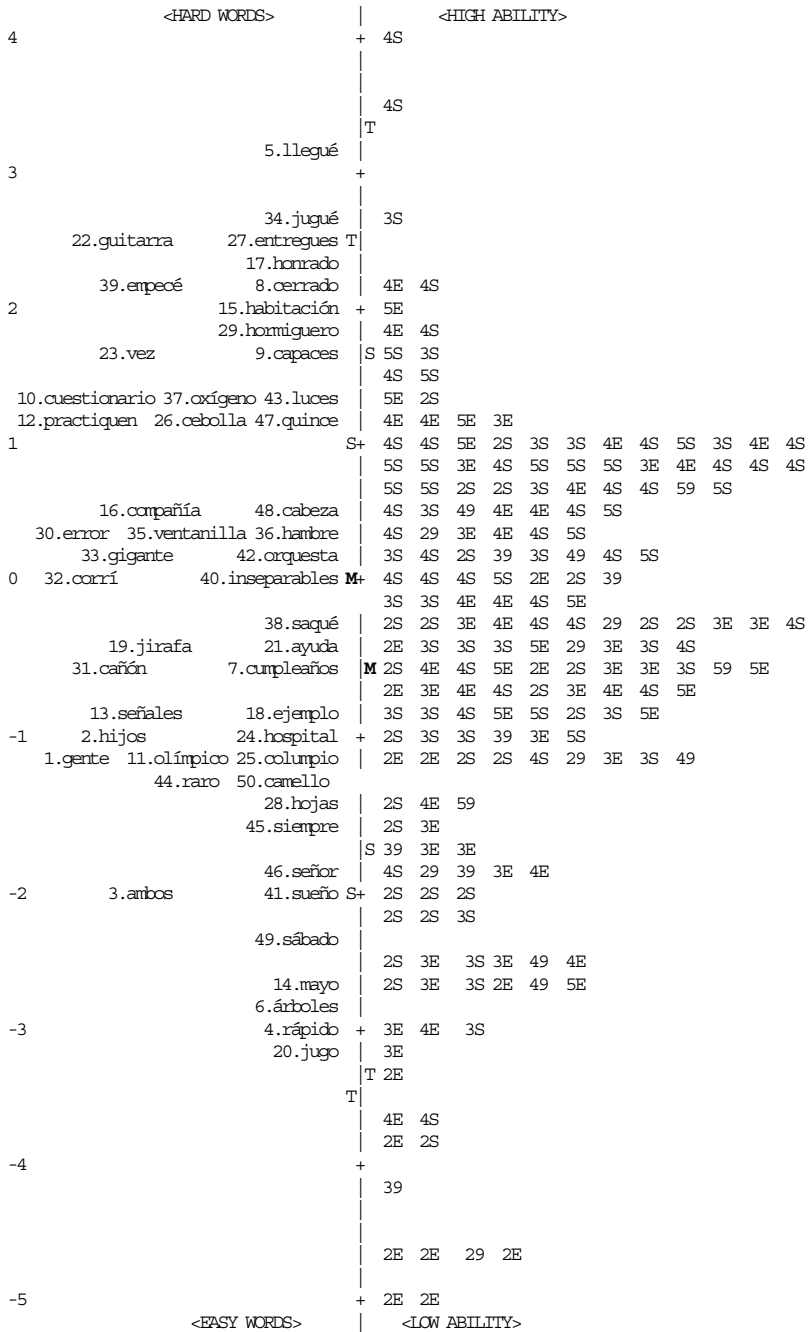
```
                    <HARD WORDS>            |     <HIGH ABILITY>
4                                           +  4S
                                            |
                                            |
                                            |  4S
                                            |T
                         5.llegué           |
3                                           +
                                            |
                        34.jugué            |  3S
        22.guitarra     27.entregues       T|
                        17.honrado          |
           39.empecé     8.cerrado          |  4E  4S
2                       15.habitación       +  5E
                        29.hormiguero        |  4E  4S
           23.vez        9.capaces          |S 5S  3S
                                            |  4S  5S
 10.cuestionario 37.oxígeno 43.luces        |  5E  2S
 12.practiquen  26.cebolla 47.quince        |  4E  4E  5E  3E
1                                          S+  4S  4S  5E  2S  3S  3S  4E  4S  5S  3S  4E  4S
                                            |  5S  5S  3E  4S  5S  5S  5S  3E  4E  4S  4S  4S
                                            |  5S  5S  2S  2S  3S  4E  4S  4S  59  5S
            16.compañía     48.cabeza       |  4S  3S  49  4E  4E  4S  5S
   30.error  35.ventanilla 36.hambre        |  4S  29  3E  4E  4S  5S
            33.gigante     42.orquesta      |  3S  4S  2S  39  3S  49  4S  5S
0    32.corrí         40.inseparables      M+  4S  4S  4S  5S  2E  2S  39
                                            |  3S  3S  4E  4E  4S  5E
                        38.saqué            |  2S  2S  3E  4E  4S  4S  29  2S  2S  3E  3E  4S
           19.jirafa       21.ayuda         |  2E  3S  3S  3S  5E  29  3E  3S  4S
         31.cañón          7.cumpleaños     |M 2S  4E  4S  5E  2E  2S  3E  3E  3S  59  5E
                                            |  2E  3E  4E  4S  2S  3E  4E  4S  5E
           13.señales      18.ejemplo       |  3S  3S  4S  5E  5S  2S  3S  5E
-1      2.hijos          24.hospital        +  2S  3S  3S  39  3E  5S
    1.gente  11.olímpico 25.columpio        |  2E  2E  2S  2S  4S  29  3E  3S  49
             44.raro   50.camello
                        28.hojas            |  2S  4E  59
                        45.siempre          |  2S  3E
                                            |S 39  3E  3E
                        46.señor            |  4S  29  39  3E  4E
-2         3.ambos         41.sueño        S+  2S  2S  2S
                                            |  2S  2S  3S
                        49.sábado           |
                                            |  2S  3E     3S  3E  49  4E
                        14.mayo             |  2S  3E     3S  2E  49  5E
                         6.árboles          |
-3                       4.rápido           +  3E  4E  3S
                        20.jugo             |  3E
                                            |T 2E
                                           T|
                                            |  4E  4S
                                            |  2E  2S
-4                                          +
                                            |  39
                                            |
                                            |
                                            |  2E  2E  29  2E
                                            |
-5                                          +  2E  2E
                    <EASY WORDS>            |    <LOW ABILITY>
```

Figure 3. Analysis of difficulty of Spanish-spelling real-words vis-à-vis children's grade and language.

level of each item on a scale from easiest to hardest vis-à-vis the participants' performance on the test, and also identified misfitting items (i.e., easy items that were misspelled by "high ability" participants or hard items that were spelled correctly by "low ability" participants).

Figure 3 shows the variable map generated from the Rasch analysis, following the dichotomous scoring system (i.e., correct vs. incorrect). Items are displayed on the left side of the line in order of spelling difficulty from the easiest on the bottom to the most difficult on the top. Participants are distributed along the right side of the line with "more able" participants on the top and "less able" ones on the bottom of the map. Participants are identified by L1 and grade level (e.g., 5S stands for a native Spanish-speaking participant in Grade 5), with the exception of a few participants whose L1 data were not obtained, which appear on the map as 9 (e.g., 49 stands for a participant in Grade 4 with no L1 information). The "M" to the left of the vertical line represents the mean item difficulty, located at an arbitrary value of 0. The "M" to the right of the line corresponds to the mean participant ability.

As Figure 3 shows, the mean ability of the participants is slightly below the mean difficulty of the words, which shows that the real-word segment of the pilot test was slightly hard for the sample being tested. Additionally, as illustrated in this figure, the test items are fairly well distributed in terms of their spelling difficulty, about half of them falling above the mean and the other half below. However, there is a small gap at the bottom of the item side of the map, which indicates that a few more easy items could be included in the test. Finally, the reliability of the real-word segment of the pilot test was found to be very high (.90).

As can be seen on the map, a large number of the "difficult to spell" words are orthographic words with the "<g> vs. <gu>" feature (e.g., *llegué*, *guitarra*, *jugué*, *entregues*) and the "<r> vs. <rr>" feature (e.g., *guitarra*). This finding does not seem to be consistent with studies of Spanish native speakers' spelling development, which show that native speakers tend to master orthographic words before regular words with various patterns, such as those containing the "<b> vs. <v>" or "<g> vs. <j>" features (e.g., *cebolla*, *gigante*) (Carbonell de Grompone et al., 1980; Morais, 1998; Valle-Arroyo, 1990). However, many of these orthographic words involve verbs (e.g., *llegué*, *jugué*, *entregues*) in which the focus feature undergoes an orthographic change when certain suffixes are added in order to preserve its original sound (e.g., <g> vs. <gu>, as in *llegar* vs. *llegué*). Thus, it is possible that these words are learned in their infinitive form, and this form is applied to every form of the verb, even if the inflection requires the use of a different grapheme. It is also worth noting that among these "difficult to spell" words, there are those, such as *llegué*, in which the <g> vs. <gu> feature appears in combination with a feature that has multiple graphemic representations (i.e., $/\lambda/$), and this may have contributed to the spelling difficulty of the word. Finally, in the case of

guitarra, the presence of not just one, but two orthographic features (<g> vs. <gu> and <r> vs. <rr>) may have increased the spelling difficulty of the word. Thus, these additional factors may help explain why Spanish-language learners have found orthographic words containing the <g> vs. <gu> and <r> vs. <rr> features very hard to spell.

On the other end of the scale, frequently used words, such as *árboles*, *mayo*, *sueño*, and *sábado*, were among the easiest ones to spell. This finding is consistent with studies that have looked at spelling development in Spanish native speakers (e.g., Carbonell de Grompone et al., 1980; Cuetos, 1993; Morais, 1998; Valle-Arroyo, 1990). The word *jugo* was found to be the easiest to spell for the sample being tested, and although it is not a high-frequency word, the two features it contains (the <j> and the <g>) both appear in their default form, thus presumably making this word very easy to spell. Additionally, it is likely that *jugo* is a popular word among participants this age, thus increasing chances that they may know how to spell it. Finally, contrastive words with the <ñ> feature (e.g., *sueño*, *señor*) were also among the easiest 10 words, and so were words that contained the <m> before bilabial stop feature (e.g., *siempre*, *ambos*).

As Figure 4 shows, the mean ability of the participants vis-à-vis the pseudo-words was close to (a little bit below) the mean difficulty of the pseudo-words, which means that this component of the test was neither too hard nor too easy for the sample being tested. Additionally, as illustrated in this figure, the test items are fairly well-distributed in terms of their spelling difficulty, with about half of them falling above the mean and the other half below. The reliability of the pseudo-word segment of the pilot test was found to be moderately high (.77).

When compared with the real-words, the pseudo-words were found to be easier to spell for the sample being tested. This may seem counterintuitive, given that participants could not have had any prior exposure to these made-up words, whereas they were likely to have been familiar with at least some of the real-words. However, the fact that the pseudo-word section only required participants to complete each item by filling in the blank with the corresponding feature considerably reduced the cognitive load of this section as compared with the real-word section in which the whole word had to be written, and this may be the reason for the relative ease of this component.

In terms of item difficulty, the Rasch analysis performed on the pseudo-words revealed similar findings to those of the real-words. The six most difficult items were all orthographic words containing either the trill feature (e.g., *tirrapo*, *irrol*, *jotarra*), or the <g> vs. <gu> feature (e.g., *guensa*, *pregues*, *laguinte*), replicating the findings from the real-word test. The easiest items were those containing the infinitive counterpart of a feature (e.g., <g>, <r>, and <c> in *dugo*, *rótipa*, and *catafes* respectively), as well as words with the <ñ> and <m> before /p/or /b/ features (e.g., *triños*, *beñal*, *fumbro*, and *ambul*), a finding once again consistent with that of the real-word component of the test. Thus,

```
5                    +
                     |
                     |
                     |
                     |  4S
                     |
4                    +T
                     |
                     |
                     |  4S
                     |
          tirrapo T  |  3E
3                    +  4E
            irrol    |  5S
           guensa    |
           jotarra   |  49  4E  4E  4S  4S  4S  4S
   pregues  laguinte |  5E  5S
            tagué    |  3S
2                    +S 4E  4S  4S  4S  4S
            derrí    |  59  5E  5S
                   S |  3E  3S  4E  4E  4S  4S  4S  4S
                     |  2S  5E  5S  5S  5S
                     |  3E  3E  3E  3S  3S  3S  3S
        cuestario    |  29  29  2S  49  4E  4E  4S  4S  4S  4S  4S  4S
1                    +  5E  5E  5S  5S  5S  5S
                     |  3E  3S  3S  3S  4E  4S  4S  4S  4S  4S  4S  4S  4S
                     |  29  2E  2S  2S  5E  5S  5S
         telpañía    |  39  39  3S  3S  3S  3S  3S  4E  4E  4E  4S  4S  4S  4S
                     |  2S  2S  2S  5E  5E  5S
           uispo M   |
0          tiquen   +M 39  3S  3S  3S  4E  4E  5E  5S
          laumpre    |  29  2E  2E  2E  2S  2S  2S
                     |  3E  3E  3E  3E  3E  3E  3S  3S  3S  3S  3S  4E  4E  4S  4S
        arámpoco     |  2E  2S  2S  2S  2S
                     |  3E  3E  3E  3E  3S  3S  3S  4E  4E  4S  59  5E  5E  5E  5E  5S
          maqué      |
-1       etimpla    +  29  2E  2S  2S  2S  4E  59  5S
                     |  39  3E  3E  3S
     teñar friño S   |  29  2E  2S  2S  2S  4S
           beñal     |  39  3E  3E
   triños  fumbro    |  49  4E  4E  4S
           ambul     |  2E  2S  2S  2S  39  3E  3E
-2                   +S
                     |  4E
                     |  29  2E
                     |
          catafes    |
                   T |  2E  2S
-3                   +
                     |  49
                     |
     dugo   rótipa   |  2E  2E  2S
                     |
                     |
-4                   +T
                     |
                     |
                     |  2E  2E
                     |
                     |
-5                   +  49
```

Figure 4. Analysis of difficulty of Spanish-spelling pseudo-words vis-à-vis children's grade and language.

taken together, the results of the real-words and the pseudo-words seem to indicate that regardless of the context in which they appear, some features are hard to spell and thus acquired late (e.g., the velar stop /g/ before the /e/ or /i/ vowels, which has to be realized as <gu>, and the trill between vowels, as <rr>), while others are acquired earlier on (e.g., the default realization of a given feature, as well as the <ñ> and the <m> before bilabial stop).

Finalizing the Assessment

As mentioned earlier, 10 extra real-words and 5 extra pseudo-words were included in the pilot test, so that any misfitting items could be deleted from the final version while still retaining 60 piloted items on the final version (40 real-words and 20 pseudo-words). As per the Rasch analysis' results, the 10 real-words that were not included in the final version of the test were partly misfitting items, that is, hard words that "low ability" participants unexpectedly tended to spell correctly and easy words that "high ability" participants unexpectedly tended to spell incorrectly (e.g., *jirafa*, *raro*, *ambos*, *orquesta*). Another factor that played a role in the final decision to eliminate some of the pilot test items was the fact that there were a few cognate words that were spelled differently in both languages (e.g., *compañía*, *cañón*, *cuestionario*, *orquesta*). In these cases, it was hard to figure out whether the reason they were misspelled by some participants was because of the features themselves (e.g., the <ñ> in *compañía* or *cañón*), or because they were familiar with the English spelling and were using this instead of the Spanish spelling to write these words. Finally, there were a few items (i.e., *hijos*, *columpio*, *inseparables*) that were eliminated because there was no need for any extra features of the type those words represented (/x/, /mp/, and /s/, respectively). Similarly, of the five pseudo-words not included in the final test, three of them were misfits (*cuestario*, *laumpre*, *telpañía*), and the remaining two, *irrol* and *pregues*, were unnecessary to maintain the minimum requirement of four items per feature.

Conclusion

To date, the final version of the Spanish developmental contrastive spelling measure has been used to collect the first three waves (Grades 2, 3, and 4) of longitudinal data from 220 participants in four two-way immersion programs across the country. Data will be collected from the same participants through Grade 5 so that the development of their spelling ability can be evaluated over the full range of the upper elementary grades. Additionally, cross-sectional data have been collected from monolingual Spanish speakers in Grades 2 through 5 in Mexico and Puerto Rico so that it will be possible to confirm or disconfirm hypothesized crosslinguistic influence. The Spanish monolingual group will be the key in helping to tease out errors that are

influenced by English from those that are not, since these students will not have had any experience with the English language, and thus will serve as a comparison group for the Spanish-English bilingual groups in the United States.

An English Developmental Contrastive Spelling Test has been developed along with the Spanish one. It has the same number of items (40 real-words and 20 pseudo-words) and like the Spanish measure, it focuses on features that are interesting from both developmental and contrastive perspectives. Having longitudinal English- and Spanish-spelling data will make it possible to study and compare their spelling development in both languages. That is, it will be possible to find out, for example, whether good spellers in one language are also good spellers in the other language, whether spelling in one's L1 is easier than spelling in the L2, or whether, regardless of the child's L1, spelling in a language that relies on a rather complex phoneme-grapheme mapping system, like English, is more difficult than spelling in a language that has a shallow orthography, like Spanish.

In addition, other measures of Spanish-English language and literacy development are being administered to the participating students each year, and this will make it possible to not only examine the development of spelling ability over time for native Spanish speakers and native English speakers, but also to look at potential relationships between spelling and other literacy skills such as decoding, reading comprehension, and writing. In other words, this measure was designed to be used both to investigate the development of Spanish-spelling ability as an important topic in and of itself, and to determine the extent to which Spanish spelling is related to English spelling and other literacy abilities in both Spanish and English. In this way, both the measure and the research findings produced with it will make contributions to the field of biliteracy development.

References

Adams, M. J. (1995). *Beginning to read*. Cambridge, MA: MIT Press.

Bachman, L., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.

Carbonell de Grompone, M. A., Tuana, E. J., Piedra de Moratorio, M., Lluch de Pintos, E., & Corbo de Mandracho, H. (1980). Evolución de la ortografía según la clasificación estructural de los errores ortográficos [Spelling development according to the structural classification of spelling errors]. *Lectura y Vida*, *1*(4), 11–17.

Cuetos, F. (1993). Writing processes in a shallow orthography. *Reading and Writing: An Interdisciplinary Journal*, *5*, 17–28.

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

Cunningham, A. C., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing in children: A quick measure of reading experience. *Journal of Educational Psychology*, *88*, 733–740.

Ehri, L. C., & Wilce, L. (1987). Does learning to spell help beginners learn to read words? *Reading Research Quarterly*, *20*, 163–179.

Estes, T., & Richards, H. (2002). Knowledge of orthographic features in Spanish among bilingual children. *Bilingual Research Journal*, *26*(2), 295–307.

Ferroli, L. J. (1991). Developmental spelling and the transfer of literacy skills among primary grade bilinguals. *Dissertation Abstracts International*, *52*, 91A. (UMI No. 9132282).

Ferroli, L., & Krajenta, M. (1989). Validating a Spanish developmental spelling test. *National Association for Bilingual Education Journal*, *14*, 41–61.

Linacre, J. M., & Wright B. D. (2000). *WINSTEPS: Multiple-choice, rating scale, and partial credit Rasch analysis* [Computer software]. Chicago: MESA Press.

Genesee, F. (Ed.). (1999). *Program alternatives for linguistically diverse students.* Educational Practice Report 1. Santa Cruz, CA and Washington, DC: Center for Research on Education, Diversity and Excellence.

Gomes de Morais, A. (1998). La ortografía en la escuela: representaciones del aprendiz y acción didáctica [Spelling in school: Learner's representations and pedagogical action]. In M. E. Rodríguez (Ed.), *Textos en contexto: La escuela y la formación de lectores y escritores* [Texts in context: School and the development of readers and writers], *Vol. 4* (pp. 35–58). Buenos Aires: Asociación Internacional de Lectura.

Howard, E. R., Christian, D., & Genesee, F. (2004). *The development of bilingualism and biliteracy from Grades 3 to 5: A summary of findings from the CAL/CREDE study of two-way immersion education*. Research Report. Santa Cruz, CA and Washington, DC: Center for Research on Education, Diversity & Excellence and Center for Applied Linguistics.

Medina, M., & Escamilla, K. (1992). English acquisition by fluent- and limited-Spanish-proficient Mexican Americans in a 3-year maintenance bilingual program. *Hispanic Journal of Behavioral Sciences*, *14*, 252–267.

Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children*. Clevendon, UK: Multilingual Matters.

Ramírez, D. J., Yuen, S. D., & Ramey, D. R. (1991). *Executive summary, final Report: Longitudinal study of structured English immersion strategy, early-exit and late-exit transitional bilingual education programs for language-minority children* (Contract No. 300-87-0156). San Mateo, CA: Aguirre International.

Rueda, R., Saldivar, T., Shapiro, L., Templeton, S., Terry, C. A., Valentino, S., & Wolf, S. A. (2001). *¡Bien dicho! Gramática, estilo y uso para expresarte mejor* [Well-said! Grammar, style, and use to express yourself better]. Grades 2–5. Carmel, CA: Hampton-Brown.

Schlagal, R. (1992). Patterns of orthographic development into the intermediate grades. In S. Templeton & D. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy* (pp. 31–52). Hillsdale, NJ: Erlbaum Associates.

Serrano, R., & Howard, E. R. (2003). Maintaining Spanish proficiency in the United States: The influence of English on the Spanish writing of native Spanish speakers in two-way immersion programs. In L. Sayahi (Ed.), *Selected proceedings of the first workshop on Spanish sociolinguistics* (pp. 77–88). Sommerville, MA: Cascadilla Proceedings Project.

Staczek, J. J., & Aid, F. M. (1981). Hortografía Himortal: Spelling problems among bilingual students. In G. Valdés, A. G. Lozano, & R. García-Moya (Eds.), *Teaching Spanish to the Hispanic bilingual: Issues, aims, and methods* (pp. 146–55). New York: Teachers College Press.

Stage, S. A., & Wagner, R. K. (1992). Development of young children's phonological and orthographic knowledge as revealed by spellings. *Developmental Psychology*, *28*, 287–296.

Thomas, W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement* (Final Paper No. 1.1). Santa Cruz, CA: University of California, Center for Research on Education, Diversity & Excellence.

University of Puerto Rico. Superior Educational Council. (1952). *Recuento de Vocabulario Español, editado por la OEA y la UNESCO* [Recount of Spanish vocabulary edited by the OEA and UNESCO]. Río Piedras, Puerto Rico: Author.

Valle-Arroyo, F. (1990). Spelling errors in Spanish. *Reading and Writing: An Interdisciplinary Journal*, *2*, 83–98.

Waters, G. S., Bruck, M., & Malus-Abramowitz, M. (1988). The role of linguistic and visual information in spelling: A developmental study. *Journal of Experimental Child Psychology*, *45*, 400–421.

Acknowledgment