Published in final edited form as:
J Speech Lang Hear Res. 2006 December ; 49(6): 1209–1223. doi:10.1044/1092-4388(2006/087).

Evaluating the Discriminant Accuracy of a Grammatical Measure With Spanish-Speaking Children

Vera F. Gutiérrez-Clellen,
San Diego State University, San Diego, CA

M. Adelaida Restrepo, and
Arizona State University, Tempe

Gabriela Simón-Cereijido
San Diego State University

Abstract

Purpose—The purpose of this study was to evaluate the discriminant accuracy of a grammatical measure for the identification of language impairment in Latino Spanish-speaking children. The authors hypothesized that if exposure to and use of English as a second language have an effect on the first language, bilingual children might exhibit lower rates of grammatical accuracy than their peers and be more likely to be misclassified.

Method—Eighty children with typical language development and 80 with language impairment were sampled from 4 different geographical regions and compared using linear discriminant function analysis.

Results—Results indicated fair-to-good sensitivity from 4;0 to 5;1 years, good sensitivity from 5;2 to 5;11 years, and poor sensitivity above age 6 years. The discriminant functions derived from the exploratory studies were able to predict group membership in confirmatory analyses with fair-to-excellent sensitivity up to age 6 years. Children who were bilingual did not show lower scores and were not more likely to be misclassified compared with their Spanish-only peers.

Conclusions—The measure seems to be appropriate for identifying language impairment in either Spanish-dominant or Spanish-only speakers between 4 and 6 years of age. However, for older children, supplemental testing is necessary.

Keywords

children; Spanish-speaking; discriminant accuracy

Accurate identification of language impairment in Spanish-speaking children who are learning English as a second language is critical for the design of appropriate treatments and for the prevention of academic delays later in school. For tests to have adequate classification, they should demonstrate high sensitivity (i.e., the proportion of true positives correctly identified by the test) and high specificity (i.e., the proportion of true negatives correctly identified by the test). Plante and Vance (1994) indicated that a sensitivity and specificity of 80% is considered fair and above 90% is considered good. However, it is difficult to find measures that identify language disorders with accuracies above 90%, even

Contact author: Vera F. Gutiérrez-Clellen, School of Speech, Language, and Hearing Sciences, and SDSU/UCSD Joint Doctoral Program in Language and Communicative Disorders, San Diego State University 5500 Campanile Drive, San Diego, CA 92182-1518. vclellen@mail.sdsu.edu.

in monolingual English speakers. Research in this area is limited, and most tests lack acceptable levels of diagnostic accuracy because of the language areas assessed, the tasks used, and the procedures utilized to sample participants in the development of test norms.

Attempts to adapt standardized English tests have not been successful. Restrepo and Silverman (2001), for example, found that the use of a Spanish adaptation of the Preschool Language Scale—3 (Zimmerman, Steiner, & Pond, 1993) with a group of Spanish-speaking children yielded scores below one standard deviation from the mean in the receptive and expressive language scales. They concluded that the use of such a test would lead to overidentification of typically developing children as having language impairment. In addition, item progression in the test did not demonstrate a developmental trend, and certain tasks seemed to pose great difficulty to the children, even when the concept was appropriate in the Spanish language. Similar results were reported for the use of Spanish adaptations of vocabulary and syntactic measures (Anderson, 1996; Restrepo, 1998).

The Spanish Structured Photographic Expressive Language Tests (preschool [SPELT-P] and school-age [SPELT-II] versions) do not appear to provide adequate results for Spanish-speaking children either (Anderson, 1996; Restrepo, 1998). Restrepo found that the sensitivity of the SPELT-II version was too low (65%), and Anderson showed that the SPELT-P did not give a good estimate of a child's grammar when compared with a structured elicitation task of the same target forms. These tests were adapted to Spanish using an English developmental model, and, as a result, they failed to assess structures such as articles and clitic pronouns, which would be more sensitive to Spanish language development (for a review, see Restrepo & Gutiérrez-Clellen, 2004).

New norm-referenced measures such as the Spanish Preschool Language Scale—4 (SPLS–4; Zimmerman, Steiner, & Pond, 2002) and the Spanish Clinical Evaluation of Language Functions (SCELF; Semel, Wiig, & Secord, 1997) continue to have significant validity problems. Both tests are based on social and academic tasks appropriate for middle-class English-speaking children in the United States, but are problematic for Spanish-speaking children with different language socialization experiences. For example, the items of the SPLS–4 test for body parts and clothes. However, these items are not typically taught in language interactions within some Latino speech communities. In addition, some items are especially problematic for children who live in Spanish-speaking homes and who have not been schooled in Spanish. For example, providing definitions in Spanish may not be a familiar task for children who are schooled in English.

Both the SPLS–4 and the SCELF are based on English measures that are not useful for identifying language disorders in Spanish. For example, the SPLS–4 evaluates Spanish prepositions, and the SCELF includes Spanish possessive pronouns and present tense third-person singular verbs as target forms. Yet, Bedore and Leonard (2001) found that children with specific language impairment (SLI) did not differ from any of the controls in their use of third-person singular or possessive pronouns. Furthermore, Spanish differs from English in the degree of specificity of location, thus preposition use differs across the two languages (Slobin & Bocaz, 1988). Because of these content validity problems, it is not surprising to find that neither the SPLS–4 nor the SCELF report acceptable diagnostic accuracy rates. The SPLS–4, for example, reports sensitivity values ranging from .68 to .95, depending on the age groups evaluated. Its specificity rate (typically developing children identified as "typical") is below 70% in most age groups, which indicates that at least 30% of these children would obtain a low score on the test and not have a language disorder. For the SCELF, the discrimination scores are at 65% sensitivity (56 out of 86 children with language impairment identified as "impaired"), again leaving the clinician with 35% of the sample misclassified. Specificity for the SCELF at about 78% is not adequate either (67 out

of 86 nonimpaired children identified as “nonimpaired”). Neither of these tests meets minimum standards for clinical identification of language disorder. Their sensitivity is considerably below the 80% acceptable rate for their use in diagnostic applications (Plante & Vance, 1994), and their specificity rates would likely result in an over-representation of typically developing children in clinical caseloads.

In the absence of valid standardized language measures, several alternative measures have been recommended, including the use of parent report and spontaneous language samples. Restrepo (1998) found that a parent interview in conjunction with language sample analyses discriminated with 90% accuracy between children with language disorders and those developing typically. Specifically, use of grammatical errors per terminable unit (T unit; Gutiérrez-Clellen & Hofstetter, 1994; Hunt, 1965), mean length of T-unit, family history of speech and language problems, and parental concerns of language problems were the most accurate measures, among a set of eight measures used. The same accuracy could also be maintained with only evidence of parent concern and grammatical errors per T unit as diagnostic indicators.

Dynamic assessment has also been recommended as an alternative to standardized and language sampling assessments (Gutiérrez-Clellen & Peña, 2001; Peña, Quinn, & Iglesias, 1992). These procedures are effective to differentiate children who are developing typically from those who have language disorders or are at risk of academic difficulties—in particular, for assessing semantic or lexical domains, narratives, and literacy. Yet, although these approaches can contribute to the identification of a language disorder, they do not provide a specific focus on a child’s grammatical skills, an area particularly vulnerable in children with language impairments.

Grammatical Characteristics of Spanish-Speaking Children With Language Impairment

Spanish-speaking children with language impairment exhibit specific grammatical deficits on spontaneous or structured language tasks. Restrepo and Gutiérrez-Clellen (2001), for example, found that Spanish-speaking children with SLI exposed to bilingual environments presented with significant definite article production difficulties, especially in the singular masculine form (*el*). The errors consisted of substitution and omission errors. These results have been corroborated in other research studies, although the specific types of errors seem to vary from study to study because of methodological differences and variability across samples in terms of ages and sociolinguistic environment (Bedore & Leonard, 2001; Bosch & Serra, 1997; Eng & O’Connor, 2000). Bedore and Leonard and Bosch and Serra found that the most common error was omission of the definite article, Eng and O’Connor found more substitutions for *el*, and Restrepo and Gutierrez-Clellen found both types of errors.

Spanish-speaking children with SLI also present with decreased use of clitic pronouns and clitic gender agreement errors (Bedore & Leonard, 2001; Bosch & Serra, 1997; Jacobson & Schwartz, 2002). These deficits are specifically evident on elicited tasks because there may not be sufficient use of clitic pronouns in their spontaneous language samples.

Unlike their English-speaking peers, Spanish-speaking children with SLI do not exhibit significant difficulty marking verb tense (Bedore & Leonard, 2001. For example, Bedore and Leonard found that SLI children’s verbs were above 70% accuracy, similar to the mean length of utterance (MLU) controls, and somewhat below the 90% accuracy rates of age controls. Verb errors included person and number errors. Spanish-speaking SLI children make errors marking mood but, in general, verb morphology errors are less frequent than article or clitic errors in Spanish SLI. In Spanish, subject–verb agreement is learned early

because of the morphological saliency of the Spanish verb system. For example, between ages 2 and 3 years (or when children produce an MLU in words between 1.40–3.21 words), Spanish-speaking children exhibit appropriate use of verb person, number, and tense (Gathercole, Sebastián, & Soto, 2002). Verb morphology may be significantly impaired only in very young children with SLI. Torrens and Wexler (2001) examined the development of verb morphology in 4 young Spanish/Catalan-speaking children with SLI and compared them with MLU controls. On the basis of spontaneous language samples, they found that these children took longer to acquire subject–verb agreement than the MLU controls. Research conducted in Spain with monolingual children with SLI has indicated that Spanish-speaking children may have specific difficulty with complex verbs (e.g., verbs that require more than two arguments). Sanz Torrent (2002) analyzed the children's verb productions by the number of arguments to the verb (complements to the verb such as "give *the ball to me*"). The Spanish-speaking 5-year-old children with SLI ($n = 12$) produced significantly fewer verbs with three arguments compared with the language- and age-matched children. Furthermore, she found that the Spanish-speaking children with SLI produced fewer verbs per utterance than language- and age-matched control children. The majority of the errors were on verb inflections (20%), and a small percentage (4.18%) of the errors involved verb omissions. These findings are strikingly lower than the verb error rates of English speakers with SLI reported in the literature (e.g., Bedore & Leonard, 2001; Leonard, Eyer, Bedore, & Grela, 1997; Rice & Wexler, 1996). Errors marking mood (e.g., the subjunctive) have been reported, although the rate of impairment is not like that of verbs in English SLI (Bedore & Leonard, 2001; Gutiérrez-Clellen et al., 2000). These findings suggest important cross-linguistic differences in the manifestations of the disorder that need to be addressed if a grammatical measure is to be used for diagnostic purposes. Whereas in English, such measure should have a focus on verb use (e.g., Gavin, Klee, & Membrino, 1993; Redmond & Rice, 2001; Rice, Wexler, & Cleave, 1995; Watkins, Rice, & Moltz, 1993), the Spanish literature summarized earlier clearly indicates that articles and clitic pronouns across gender and number inflections may have greater clinical sensitivity. Verbs may prove to be clinically useful if tested in contexts requiring at least two arguments and across indicative and subjunctive mood (i.e., contexts with sufficient syntactic complexity). Research with English speakers has shown that grammatical accuracy may be significantly affected by increased utterance complexity (e.g., Masterson & Kamhi, 1991). Thus, a diagnostic grammatical measure should also provide for elicitation contexts that vary in syntactic complexity.

Grammatical Characteristics of Children Learning English as a Second Language

As Spanish-speaking children learn and use English, they may show differences related to crosslinguistic influence from English, in spite of the fact that they may not be fully proficient in that language. There is evidence that bilingual children (i.e., children who use both languages on a daily basis) may show grammatical patterns resulting from cross-linguistic influence that would not be found in the speech of their monolingual peers (e.g., Döpke, 2000). Within a competition model (MacWhinney, 1987, 1997), structural similarities and differences between the languages affect the strength of particular cues and may result in performance differences. Some of these grammatical patterns may be similar to those described in children who are undergoing language attrition (Restrepo & Kruth, 2000). Anderson (2001) reported a decrease in tense accuracy and a reduction of tense forms in two typically developing Spanish-speaking children as they started to use English at school and with peers. Children may also show less reliable use of gender agreement in their articles and clitic pronouns. Restrepo and Gutiérrez-Clellen (2001) and Anderson (2003) found that Spanish-speaking children with SLI exposed to English tended to make gender

agreement errors. Yet, these types of errors are not reported in children from Spain who are exposed only to Spanish or to Spanish and Catalan (Bosch & Serra, 1997). Both Spanish and Catalan mark gender in articles, thus gender becomes a reliable cue in speakers of these languages. For speakers of Spanish and English, however, gender is not reinforced in the input and as a result, its use may become less consistent in contexts of limited exposure to or use of the Spanish language. Although there are significant individual differences in the ultimate grammatical attainments of bilinguals (i.e., not all bilinguals exhibit language attrition), it is possible that the grammatical accuracy of Spanish-speaking children who are English speakers differs from that of monolingual children, especially if they are learning their second language in subtractive bilingual environments, as is the case in the United States. Thus, it is not yet known how well a grammatical measure would classify Spanish-speaking bilingual children compared with monolinguals. The need to understand if these two groups of children differ quantitatively and qualitatively from each other is essential for the development of appropriate assessment measures and norms. If these populations differ significantly from each other, standardized tests must provide separate norms for bilingual and monolingual children. However, if performance is similar between groups, it is possible that the normative samples can be combined. For example, the norms in the Spanish SPLS–4 are combined because the authors noted no differences between samples. Based on the test standardization research, there were only 1–3 standard score point differences between bilingual and monolingual children. However, in a separate study with children from other countries, the international cohort showed higher scores than the U.S. children above age 4.

The current study was based on the need for a standardized language measure to identify language impairment in Spanish-speaking children between the ages of 4 and 7 years. Specifically, it focuses on the validation of a Spanish Morphosyntax Test (S-MST), which is part of the Bilingual English Spanish Assessment, a battery of tests soon to be published in the United States (Peña, Gutiérrez-Clellen, & Iglesias, 2006). The main goal was to evaluate the discriminant accuracy of the S-MST to determine how well it classified Spanish-speaking children with typical development and Spanish-speaking children with language impairment. As discussed earlier, available measures are not based on grammatical features that address the characteristics of language impairment in Spanish-speaking children. They may also include forms that are not developmentally appropriate to identify language impairment because research with these children has varied in their age groups, tasks used, and specific grammatical forms tested. As a result, there is very limited information regarding the diagnostic accuracy of a grammatical measure to identify language impairment in Spanish-speaking children ages 4–7 years. The manifestations of the disorder may vary developmentally, and assessment measures should address these differences in order to have diagnostic utility. The sensitivity of a measure may vary by age because children with language impairment present with changing symptomatology over time. For example, 6-year-olds may not exhibit significant problems with inflections, but they may show difficulties in the semantic domains or in the use of complex syntax (Paul, 2001). Therefore, we examined this issue directly by evaluating the discriminant accuracy of the S-MST for Spanish-speaking children in each of three different age groups established a priori within the 4–7 years age range.

The second aspect of this research was to examine the performance of Spanish-speaking children who are bilingual. In this article, the term *bilingual* is used to refer to children who have high levels of English exposure and use, but are Spanish dominant (i.e., these children are not English dominant or “balanced” bilinguals). To validate a Spanish measure of language that is used for diagnostic purposes, it is critical that the samples are based on the child’s “best” language (i.e., Spanish). When dominance is not addressed in test norming, the likelihood of misdiagnosis is significantly increased. As discussed earlier, most Spanish tests are based on monolingual populations or on samples that do not consider the child’s

dominant language. We hypothesized that if exposure to and use of English as a second language have any effect on grammatical performance in the first language, the bilingual children would exhibit lower rates of grammatical accuracy and be more likely to be misclassified by the S-MST than their Spanish-only peers.

Method

Participants

One hundred and sixty children and their families were sampled from schools in districts serving predominantly low-income families in California, Georgia, Texas, and Pennsylvania. Table 1 shows the combined educational and income profiles for the typical language development and language impairment groups. School lunch program status was used as a metric for income level. Each school independently determined lunch program qualification status, which was based on family income and the number of occupants in the household. The majority of the children from California, Georgia, and Texas were of Mexican American descent. The children from Pennsylvania were predominantly of Puerto Rican and Dominican descent (see Table 1 for a breakdown of the participants by dialect spoken). None of the participants had a history of hearing loss, sensorimotor delays, frank neurological deficits, or significant health problems, all of which were based on school records and parent report.

The following sections provide a description of the language characteristics of the participants. First, we indicate the procedures and criteria for assigning bilingual status based on parent and teacher questionnaire data. Second, we list the procedures and criteria used to determine language dominance based on spontaneous narrative samples. Third, we describe the criteria used to differentiate the language ability groups. Because the children from Georgia were sampled using a slightly different protocol, the procedures and criteria for establishing their language skills are described in a separate section.

Procedures and criteria for establishing bilingual status—The bilingual status of the children from California, Texas, and Pennsylvania (i.e., 141 children) was determined using parent and teacher reports based on previous research with these measures (Gutiérrez-Clellen & Kreiter, 2003). Parents or guardians were interviewed using a parent questionnaire that addressed the children's general language use across settings and interlocutors as well as history of developmental delay, hearing loss, or any concerns about speech or language skills. The parent questionnaire also served as a measure of exposure to and use of each language at home. Parents were asked to rate proficiency of each language spoken by each member of the household with whom the child had the opportunity to interact, and the child's language proficiency and use were rated using a 5-point scale for each measure: 0 (*no use or proficiency*) and 4 (*use all the time and nativelike proficiency*). They also reported the number of hours the child interacted with each member of the household and the language spoken during those interactions. The children's teachers were also given a questionnaire to rate the participants' use and proficiency of each language using the same 5-point scale. In addition, they provided an estimate of the percentage of time that the child was exposed to each language as a measure of input in school. To verify the accuracy of the questionnaire data (e.g., amount of input by language during the week, language exposure estimates across contexts at school), at least one half of all questionnaires were independently rescored by a second bilingual research assistant. Item-by-item reliability checks yielded above 90% agreement between the two judges.

Children were determined to be Spanish-only proficient if they had (a) minimal or no exposure to English, (b) a minimum parent rating of 3 for language use and proficiency in Spanish, and (c) a minimum teacher rating of 3 for language use and proficiency in Spanish.

Children were judged to be bilingual if they had (a) a minimum of 20% of time exposed to both Spanish and English (based on Pearson, Fernandez, Lewedeg, & Oller, 1997), (b) a minimum parent rating of 3 for language use and proficiency in Spanish, (c) a minimum teacher rating of 3 for language use and proficiency in Spanish, and (d) the ability to speak English based on an analysis of the child's spontaneous language (see details later in this section). There were no disagreements between parent and teacher reports in whether children were Spanish only or bilingual. The bilingual status of the children with language impairment was based on the same criteria, except that both parent and teacher ratings were based only on language use (proficiency was expected to be below a 3 in both languages).

In addition to the parent and teacher questionnaires, Spanish dominance for the California, Texas, and Pennsylvania samples was verified using a direct measure of grammatical proficiency based on spontaneous narrative samples. The narrative samples were obtained using wordless picture books: for Spanish, *Frog on His Own* (Mayer, 1973) and *Frog Goes to Dinner* (Mayer, 1974); for English, *Frog, Where Are You?* (Mayer, 1969) and *One Frog Too Many* (Mayer, 1975). The languages were tested in two separate administrations of the narrative procedure, one per language. Children were considered to be typically developing Spanish-dominant bilingual if they were able to produce Spanish narrative samples with minimal grammatical errors (i.e., below 20% ungrammatical utterances), but with a substantial amount of grammatical errors in English (i.e., more than 20% ungrammatical utterances). This criterion was a rough estimate of proficiency adapted from previously published guidelines for the spontaneous language assessment of Spanish-speaking children (Gutiérrez-Clellen et al., 2000; Restrepo, 1998).

Narratives were transcribed word by word and coded by a bilingual research assistant who used the Systemic Analysis of Language Transcripts computer program (Miller & Chapman, 1985) to find grammatical errors in each language. Lexical errors, phonological errors, or cohesion errors were not counted as grammatical errors. Spanish–English mixed utterances and utterances with unintelligible words were excluded from the analysis. For each narrative sample, the proportion of grammatical utterances was calculated as an external measure of proficiency in each language. This measure was selected for two reasons. First, it was assumed that first-language learners at these ages were capable of producing utterances with minimal grammatical errors. Second, guidelines for identifying language impairment in bilingual children indicated that the proportion of grammatical utterances in children's spontaneous language samples (i.e., a cutoff criterion of more than 18% ungrammatical utterances) were valid and reliable estimates of language development in bilingual children (Gutiérrez-Clellen et al., 2000). Transcription reliability was obtained with two bilingual transcribers who independently transcribed and coded 20% of the audiotapes in each language. Word transcription reliability was 93% for English and 90% for Spanish. Grammatical code reliability was 92% for English and 90% for Spanish. Any discrepancies were resolved by consensus with a third bilingual transcriber.

The majority of the children attended bilingual classrooms with Spanish-speaking aides. However, there was great variability in the use of Spanish across classrooms.

Criteria for identification of children with language impairment—Given the fact that there are no valid standardized tests to identify Spanish-speaking children with language impairment, the procedures used for identification could not be based on any available language tests. This situation presented a problem. Because the purpose of this investigation was to validate a grammatical measure for clinical use, it would have been inappropriate to use invalid tests as the standard for assigning children to ability groups. Given the lack of developmentally, culturally, and linguistically appropriate standardized measures for this population, the best research standard for identifying language impairment in Spanish-

speaking children should be based on measures that have been previously validated with this population. Restrepo (1998) found that evidence of parent concern and significant grammatical errors in the child's spontaneous language had a high agreement with the clinical judgment of trained bilingual speech-language pathologists, and differentiated children with and without SLI with high discriminant accuracy. Children with language impairment from California, Texas, and Pennsylvania met clinical (diagnostic) criteria based on evidence of parent or teacher concern and clinical judgment based on observations of trained bilingual speech-language pathologists (e.g., reported evidence of limited responsiveness in conversational samples, modifiability). In addition, these children had to demonstrate a significant percentage of ungrammatical utterances in their narrative samples (i.e., above 20% ungrammatical utterances; Gutiérrez-Clellen, Restrepo, Bedore, Peña, & Anderson, 2000; Restrepo, 1998) in both Spanish and English (if any spoken). None of the children evidenced hearing impairments, mental retardation, emotional disturbance, motor difficulties, or neurological deficits, according to parent report and school records. No direct measures of hearing functioning were administered, but the cognitive functioning of half of the children was verified using the non-verbal cognitive measures from the Differential Ability Scales (Elliott, 1990). About one third of these children were in the caseloads of trained bilingual clinicians. The remaining children had just been referred for a diagnostic evaluation by their teachers because of parent concerns, teacher concerns, or both. The children with language impairment and their typically developing peers were recruited from the same classrooms and schools. The children with typical language development were learning their language(s) without difficulty, according to parent and teacher reports as well as clinical observation. Upon verification that these children met the criteria, they were age matched (±3 months) with the children with language impairment. The groups were also matched by bilingual status for most cases. Further details on the language characteristics of the sample are described in Table 2.

Language characteristics of the Georgia sample—Nineteen children from Georgia were sampled. All of these children were Spanish-only proficient and had minimal exposure to English, on the basis of parents' ratings of their children's level of comprehension and expression of both English and Spanish using a 5-point rating scale for each measure: 0 (*no proficiency*) and 4 (*nativelike proficiency*). For these children, Spanish was the only language spoken at home, according to parent reports. The children were attending English-only prekindergarten, with English-speaking teachers, except for one classroom that had one bilingual aide. At the time of language testing, they had only attended their first 2 months of preschool; therefore, their contact with English was quite limited. There were no bilingual children in this sample. Children were not able to participate in English testing, and when they did, they could only respond to a receptive vocabulary measure, with a score that placed them two standard deviations below the mean of their age group. Children with language impairment had failed a Spanish language screening involving clinical observation during spontaneous language samples as well as elicited probes examining semantic and morphological skills. There were 8 typically developing children and 11 children with language impairment.

The total sample (i.e., 160 children across sites) included 80 children with typical language development and 80 children with language impairment equally distributed in each of the following age groups: 60 children ages 4;1–5;1, 50 children ages 5;2–5;11, and 50 children ages 6;1–7;0. Because the goal was to have a sufficient number of children at each age, the age intervals varied from 13 months (ages 4;1–5;1), to 9 months (ages 5;2–5;11), to 11 months (ages 6;1 to 7;0).

Procedures

S-MST—Drawing from the available research on the grammatical characteristics of Spanish-speaking children with language impairment, the S-MSTwas designed to include morphosyntactic targets that were predicted to be sensitive to language impairment in Spanish speakers (i.e., articles, clitics, and verbs; Bosch & Serra, 1997; Gutiérrez-Clellen & Simón-Cereijido, 2004; Restrepo, 1997, 2003; Restrepo & Gutiérrez-Clellen, 2001; Restrepo & Kruth, 2000; Sanz Torrent, 2002). In addition, it was of interest to sample complex forms that would facilitate identification in children as old as 7 years. Previous research with Spanish speakers suggested that children with limited language skills or who are at risk would exhibit limited syntactic complexity as measured by the use of syntactic subordination (Gutiérrez-Clellen, 1998; Gutiérrez-Clellen & Hoffstetter, 1994). Thus, the S-MST was not designed to sample all grammatical features of the language, but only those that had clinical potential for children ages 4–7 years.

For a measure to accurately identify children with language impairment, it is critical that items address potential dialectal differences. Toward this end, a number of steps were taken. First, the grammatical forms selected excluded forms or grammatical contexts that are not consistently used across Spanish dialects based on the available literature (for detailed reviews on dialectal variation, see Lipski, 1990; Ramírez, 1992; Roca & Lipski, 1993; Silva-Corvalan, 1995; Varela, 1992). For example, present perfective forms (e.g., He comido/ [I] have eaten) is not frequently used in Puerto Rican Spanish, and complex subjunctive forms may not be used by speakers with limited literacy in Spanish (Stokes, Krashen, & Kartchner, 1998). Depending on the dialect, article gender may not be reliable for certain nouns. For example, in Castilian Spanish the feminine article *la* is obligatory in front of *radio* (when referring to the receiving set), but the masculine *el* would be required in Mexican Spanish.

Second, forms and items were reviewed by panels of speakers of Caribbean and Mexican American Spanish (graduate students from three different institutions) and colleagues with expertise in the target Spanish dialects. Questionable forms and items were then eliminated on the basis of evidence of inconsistent use across dialects. Specific dialect features consistently used by a given dialect, but deemed critical for the identification of the disorder in another dialect, were included. In these cases, specific scoring procedures were developed so as not to penalize the dialect. For example, omission of plural *–s* in articles and clitics by Puerto Rican Spanish speakers was not penalized but they were considered incorrect in Mexican Spanish. Similarly, pronoun case substitutions as in *leísmo* (i.e., *le/lo/him*), which is typical of some Mexican Spanish dialects, were not penalized but they were important for assessing clitics in other dialects.

Third, item analysis was used to determine item discrimination (i.e., which items worked best at eliciting target responses and which ones needed to be eliminated) and to verify which items elicited alternative responses that were dialectal in nature and grammatically acceptable. In this process, the percentage of Spanish speakers with language impairment who passed each item is subtracted from the percentage of Spanish speakers with typical development who passed each item, to obtain an index of discrimination or *D* value for each item. In general, a value of 2 indicates fair discrimination, a value of 3 indicates good discrimination, and a value above 3 indicates excellent discrimination. Items that are subject to dialectal differences yield unacceptable *D* values. Based on this preliminary work, the S-MST was refined to include only items appropriate for different Spanish dialects and that would maximize differentiation between children with different language abilities (i.e., only items that had a *D* value of at least 2 were retained).

The S-MST had a total of 73 items presented in either cloze format (e.g., Examiner: Los niños tienen unos carros. Y aquí, los niños tienen … Answer: un/el carro [Examiner: The

children have some cars. And here, the children have … Answer: a/the car]) or sentence repetition format (e.g., La niña que estaba jugando con la puerta se lastimó la mano [The girl who was playing with the door hurt her hand]: El gato no quería comer aunque tenía hambre [The cat did not want to eat even though it was hungry]).

For the cloze presentation format, there were 22 items assessing articles, verbs, clitic pronouns, and subjunctive mood. The cloze task had 7 article items (e.g., Examiner: María está jugando con unas muñecas. Y aquí, María está jugando con … Answer: una/la muñeca [Examiner: María is playing with [indefinite article feminine plural] dolls. And here, María is playing with … Answer: a/the doll]), 2 preterite items (e.g., Examiner: Los niños van a agarrar al gato. ¿Aquí, qué hicieron ya? Los niños ya … Answer: lo agarraron/cogieron [Examiner: The children are going to get the cat. Here, what did they do? The children … Answer: got/caught it]), 8 object clitic items (e.g., Examiner: Los niños van a abrir los regalos. Yaquí, ¿qué hacen los niños con los regalos? Los niños … Answer: los abren/están abriéndo los [Examiner: The children are going to open the presents. And here, what do the children do with the presents? The children … Answer: open/are opening them]), and 5 subjunctive items (e.g., Examiner: La mamá quiere que pongan la mesa. Y aquí, ¿qué quiere? La mamá quiere que … Answer: coman/tomen la sopa [Examiner: The mother wants them to set (subjunctive present third-person plural) the table. And here, what does she want? She wants them to … Answer: eat/drink (subjunctive present third-person singular) soup]).

The sentence repetition task had 51 items designed to examine use of complex syntax (e.g., relative pronouns, conditional verbs, and conjunctions; e.g., Si tuviera dinero me compraría un helado [If I had money, I would buy myself an ice-cream], La niña estaba triste porque se le rompió la muñeca [The girl was sad because her doll broke]). As with the cloze task items, the specific complex forms included in each sentence were based on preliminary item analyses. They were not determined a priori. That is, only forms that had fair-to-good item discrimination were included. No assumptions were made as to why a particular form differentiated the groups. Children were presented with two practice items before every cloze structure. When presenting the demonstration items, the target items were always stressed by exaggerated intonation and by pointing to the picture. For the sentence repetition task, children were presented with two practice items and were encouraged to repeat the full sentence without substitutions or comments.

The children were tested at their schools by trained bilingual research assistants. During testing, the child was addressed in only one language, Spanish. Children who were bilingual were tested in English on a different day by a different examiner. All responses were audio-taped and transcribed to obtain intra- and interrater reliability. About one half of all tests were independently scored by a second rater to achieve at least 90% item-by-item agreement. Disagreements between raters were related to audiotape intelligibility issues and were resolved by consensus with a third judge. Each item was then scored as correct, incorrect, or not applicable, to account for cases of unintelligible or code-switched responses to be excluded from the analyses. Semantic substitutions (e.g., use of a subjunctive mood with a different verb) were coded as correct as long as the response had the target inflection or form (e.g., a different verb was used that contained the target mood). Because the sentence repetition task was designed to target specific verb forms and subordinate clauses, it was not coded word by word. Thus, no credit was given for any type of substitution of the grammatical targets tested in each sentence.

Alternative responses based on different dialects were provided to account for dialectal differences. In addition, to avoid penalizing speakers of Caribbean Spanish, omissions of plurals in articles and clitic pronouns were scored as correct if there was evidence from the

parent questionnaire that the child was a speaker of that dialect. In Caribbean Spanish, these omissions are highly frequent and are not symptomatic of a grammatical deficit. All other items were scored using the same language standard. An S-MST score was derived based on the number of correct items divided by the total correct and incorrect items. To determine if there were any differences across sites, the test scores of the children with typical language development (i.e., 80 children) were compared using a one-way analysis of covariance (ANCOVA) with age as a covariate. Using a *t* test, dialectal differences between Caribbean (i.e., Dominican, Puerto Rican) and Mexican Spanish speakers were directly evaluated by examining the test performance of all available Caribbean children with typical language development (i.e., 9 children's mean age = 69.7 months, *SD* = 7.3 months) and the same number of Mexican Spanish speakers matched by age (9 children's mean age = 69.8 months, *SD* = 6.8 months).

Linear discriminant analyses evaluating group classifications (i.e., children with typical language development and children with language impairment) were performed on the percentage of correct scores. This was accomplished in two phases. Children were randomly assigned to Phase 1 or Phase 2. In Phase 1 (exploratory discriminant analyses), the test scores of 116 children (i.e., 22 age-matched pairs from 4;0 to 5;1, 18 pairs from 5;2 to 5;11, and 18 pairs from 6;1 to 7;0) were entered into separate discriminant analyses for each age group. In Phase 2, confirmatory discriminant analyses were conducted with the remaining 44 children (i.e., 8 pairs from 4;0 to 5;1, 7 pairs from 5;2 to 5;11, and 7 pairs from 6;1 to 7;0). This was done to determine if the discriminant functions derived from the exploratory data sets were able to predict group membership with independent samples of children of the same ages. That is, the exploratory discriminant analyses generate canonical discriminant function coefficients that maximize differences between groups relative to differences within groups. Then, in the confirmatory analysis, a discriminant function score for each case was manually calculated by multiplying the original test score by the coefficient and by adding a constant (Tabachnick & Fidell, 1989). These discriminant function scores were used to classify cases into groups. A case was classified into the typical language development group if its discriminant function score was above 0 and into the language impairment group if the discriminant function score was below 0.

Chi-square tests were used to compare the classifications of children depending on their bilingual status or their dialect, because these tests are not affected by unequal or small sample sizes. Finally, a *t* test was used to evaluate score differences between typically developing Spanish-speaking monolingual children (i.e., Spanish-only proficient) and Spanish-dominant bilinguals. All comparisons were evaluated with an alpha error set at .05.

Results

Table 3 displays the means and standard deviations of the typical language development children sampled across California, Texas, Pennsylvania, and Georgia sites. Because the sites varied in sample size, homogeneity of variance was evaluated with the Levene's statistic: 1.476, $p = .228$. The sample variances were not significantly different. An ANCOVA with age as a covariate was used to compare site scores, given that the sites differed by age. There were no significant differences in test scores across sites, $F(3, 75) = 0.12$, $p = .94$, $\eta^2 = .10$.

There were no apparent differences between Caribbean and Mexican Spanish dialects either. The test performance of the Caribbean Spanish-speaking children did not differ from the Mexican Spanish-speaking children, $t(16) = 0.92$, $p = .36$, $d = .44$. Given these results, it was reasonable to assume that the groups were equivalent and that they could be combined for the remaining analyses. The next step was to examine the discriminant accuracy of the S-

MST, the primary goal of this study. First, it was important to examine the equality of variances for each sample to ensure that homogeneity of variance could be assumed. Although inference is usually robust with respect to heterogeneity of covariance matrices, classification is not. Cases tend to be over-classified into groups with greater dispersion. Box's test of equality of variance–covariance matrices indicated non-significant differences in equality of variance–covariance matrices for ages 4;0–5;1, but indicated differences for ages 5;2–5;11 and 6;1–7;0 (Box's $M = 0.141$, $p = .71$; Box's $M = 9.272$, $p = .003$; and Box's $M = 6.984$, $p = .009$, respectively). Given these results, the discriminant analyses for these two age groups were run using arcsine transformations of the S-MST scores because Box's test showed no violations of equality of variance–covariance matrices (for ages 5;2–5;11, Box's $M = 1.280$, $p = .265$; for ages 6;1–7;0, Box's $M = 0.563$, $p = .459$). Table 4 shows the mean test scores by the three age groups and the number of children correctly classified by the exploratory discriminant analyses. The results of the exploratory analyses yielded a significant squared canonical correlation of .75, $p = .000$; .83, $p = .000$; and .75, $p = .000$ for the three age groups, respectively. These results indicate large and significant associations between S-MST weighted scores and group membership. At ages 4;0–5;1, the test scores classified correctly 86.4% of the children with language impairment and 86.4% of the children with TLD. At ages 5;2–5;11, the measure classified 94.4% of the children with language impairment and 94.4% of the children with typical language development, but at ages 6;1–7;0, it only classified 72.2% of the children with language impairment and 88.9% of the children with typical language development.

Table 5 shows how the cutoff scores for each age group were derived. For each age group, and using the derived discriminant coefficients, the S-MST score that resulted in a discriminant function score of 0 was determined to be the cutoff score for the corresponding age group. The cutoff score was 50 for ages 4;0–5;1, 70 for ages 5;2–5;11, and 70 for ages 6;1 to 7;0. Table 6 shows the mean test scores and the number of children correctly classified in the validation phase (confirmatory discriminant analyses) based on these cutoff scores. At ages 4;0–5;1, 87.5% of the children with language impairment and 100% of the children with typical language development were classified correctly. At ages 5;2 to 5;11, both children with language impairment and typical language development were classified with 100% accuracy, but at ages 6;1 to 7;0, the measure classified only 42.8% of the language impairment children correctly and 57.2% of typical language development children. Although all the typical language development children in this age group were correctly classified, the S-MST score was not sufficiently accurate to identify children with language impairment above 6 years of age.

There appeared to be no specific forms responsible for the overall S-MST score in either group. Article and verb scores increased with age in both the language impairment and the typical language development groups. The mean percentage of correct article production for the language impairment group was 35 for age 1, 62 for age 2, and 63 for age 3. The typical language development group's mean percentage of correct article was 81 for age 1, 89 for age 2, and 91 for age 3. Regarding percentage of correct verbs, the language impairment group scored 36 for age 1, 41 for age 2, and 48 for age 3; the typical language development group's scores were 73 for age 1, 77 for age 2, and 79 for age 3. Clitic production was highly variable. The language impairment group reached a mean percentage score of 28 ($SD = 21$) for age 1, 48 ($SD = 27$) for age 2, and 49 ($SD = 29$) for age 3. The typical language development group's percentage of correct clitic scores were 70 ($SD = 27$) for age 1, 86 ($SD = 16$) for age 2, and 76 ($SD = 23$) for age 3. Given the fact that these forms were tested with a very limited number of items, it was not appropriate to evaluate these trends statistically. Future research examining clitic performance across age groups will be needed to determine if this variability is related to differences in English use or to the specific procedure used to evaluate clitics by the S-MST.

The second goal of this study was to examine the S-MST performance of the Spanish-dominant bilingual children (n = 38) compared with their Spanish-only proficient peers (n = 122). This was accomplished in two ways. First, we determined whether these children were more likely to be misclassified by the S-MST. A 2 (yes or no) × 2 (bilingual status) chi-square analysis showed no differences in the likelihood of being misclassified or correctly classified, $\chi^2(1, N = 160) = 1.22$, $p > .26$. Bilingual status did not appear to affect classification accuracy.

Next, we examined the possibility that the bilingual children had lower grammatical scores than their Spanish-only peers. Given the small number of children available, only 13 Spanish-dominant bilinguals and 13 Spanish-only proficient children matched by age (Spanish-dominant bilinguals' mean age = 6;3, SD = 6.96 months; Spanish-only proficients' mean age = 6;2, SD = 6.65 months) were identified and compared on their S-MST scores. All of these children had typical language development. The comparison was made only with these children because there were not enough children matched by age and bilingual status in each ability group. Table 7 shows the mean test scores for the two groups. There were no significant group differences in S-MST overall performance, $t(24) = 0.511$, $p = .614$, $d = .21$. Comparisons across specific grammatical forms showed no performance differences either. The Spanish-dominant bilinguals were not less grammatically proficient than their Spanish-only proficient peers in any of the grammatical forms: articles, $t(24) = -0.200$, $p = .843$, $d = .07$; clitics, $t(24) = 0.325$, $p = .748$, $d = .13$; or verbs, $t(24) = -0.528$, $p = .605$, $d = .19$. An error analysis did not result in any particular trends in types of errors (i.e., omissions vs. substitutions).

Because the purpose of this study was to evaluate the discriminant accuracy of the S-MST with Spanish speakers, a population that may speak a variety of Spanish dialects, it was also important to determine whether Spanish dialect affected the S-MST classification accuracy. Speakers of Central American dialects were merged with Mexican Spanish speakers (n = 171) because of the similarities between these dialects for the specific grammatical forms targeted by the S-MST (Roca & Lipski, 1993). The Caribbean speakers were of Puerto Rican and Dominican descent (n = 27). A 2 (yes or no) × 2 (Mexican or Caribbean dialects) chi-square analysis showed no differences in the likelihood of being misclassified or correctly classified by the S-MST, $\chi^2(1, N = 198) = 0.205$, $p = .6505$. As with bilingual status, Spanish dialect did not appear to affect classification accuracy.

Discussion

The results showed that a grammatical measure may be useful to clinically identify language impairment in Spanish speakers. The S-MST was able to identify Spanish-speaking children with language impairment with good accuracy, especially from ages 4 to 6 years. These findings corroborate previous research that showed that a measure focused on articles, clitics, subjunctive verbs, and complex syntax may unveil the language difficulties of these children (e.g., Bedore & Leonard, 2001; Gutiérrez-Clellen, 1998; Restrepo & Gutiérrez-Clellen, 2001). The S-MST had overall good discrimination accuracy for the children between ages 4 and 6 years. Based on Plante and Vance's guidelines (1994), measures with accuracy above 90% are considered to have good discrimination and those with 80% to 89% are considered fair. At ages 4;0–5;1, the S-MST had fair-to-good discrimination (86%); at ages 5;2–5;11, it had good discrimination (94%). On the other hand, above 6 years, the sensitivity of the S-MST is considered to be poor (Plante & Vance, 1994). As described earlier in this article, when a test is first developed there is a need to determine which cutoff scores should be used. The usefulness of discriminant analysis is that it helps establish the appropriate cutoff scores for each age group. This is in contrast with the use of arbitrary cutoff scores (e.g., one or two standard deviations below the mean) that may not have

adequate classification accuracy because they are not derived by the database. The statistically derived cutoff scores maximally differentiate the ability groups for each age, and these are the scores that should be applied in making classification decisions. However, one of the limitations of discriminant analysis is that the cutoff scores may not work equally across samples. In the present study, the confirmatory analyses verified the adequate level of sensitivity and specificity of the S-MST for the younger groups, but showed a lower level of sensitivity than that found in the exploratory study for the older group. A closer look at the mean scores for the ability groups across exploratory and confirmatory studies suggests that the older language impairment group in the confirmatory study had a higher level of grammatical skill than the language impairment group in the exploratory study. These differences may have had an impact on the level of sensitivity obtained.

The low sensitivity of the S-MST with children above age 6 years may be related to developmental changes in the manifestation of the disorder in general (Paul, 2001), which makes accurate identification more difficult regardless of a child's second language proficiency. Table 8 illustrates the distributions of grammatical utterances from the children's spontaneous language samples across the three age groups from the large sample recruited at the San Diego, Philadelphia, and Austin sites. The older children with language impairment appeared to have less pronounced grammatical problems than the younger groups. It is important to note that in previous research in which accuracy results above these were found (Restrepo, 1998), the children in the study were identified as having moderate-to-severe language impairment and were receiving therapy by speech-language pathologists. There is some evidence that sensitivity and specificity may be enhanced when treatment status (as opposed to diagnostic status) is used to identify children with language impairment. Children who are receiving treatment are more likely to obtain poor test scores compared with children who are not (Ellis Weismer et al., 2000). Many of the children with language impairment in the present study had just been referred and, therefore, were not receiving any speech or language services at the time of the testing. Thus, these results, especially for ages between 4 and 6 years, represent the best discrimination rates to date based on morphosyntactic probes, and may indeed be reflective of the true diagnostic accuracy of these measures.

The developmental trends in classification accuracy across the two younger groups found in the present study are consistent with the characteristics of Spanish language impairment as well. There is evidence that Spanish-speaking children with language impairment can achieve a high level of morphological accuracy early compared with their English-speaking peers because of the frequency and saliency of Spanish morphological forms (for a review, see Restrepo & Gutiérrez-Clellen, 2004). The S-MST does not appear to be an appropriate measure to identify language impairment in older Spanish-speaking children. Future research with school-age children will be needed to examine the clinical sensitivity of other language features, such as verb arguments (Gutiérrez-Clellen & Simón-Cereijido, 2004), specific complex sentences (Gutiérrez-Clellen, 1998), or derivational morphemes (Auza, 2003).

The second aspect of this research focused on the performance of the bilingual children. These children were not more likely to be misclassified by the morphosyntactic measure used in the study. In addition, the comparisons between the typically developing Spanish-dominant bilinguals and their Spanish-only proficient peers indicated no differences between the groups based on the S-MST. Although these results need to be replicated with larger groups of bilingual children, the findings suggest that exposure to and use of a second language (i.e., English) do not necessarily affect overall morphosyntactic accuracy of the first language, at least for children with the linguistic profiles and age ranges studied. The findings are consistent with previous research showing no significant differences between

young bilingual children and those with proficiency in only one language on language processing tasks (Gutiérrez-Clellen, Calderón, & Ellis Weismer, 2004). It is important to note that the linguistic environments of the bilingual children in these comparisons included children from three different regions in the country growing up in different sociolinguistic contexts. Future research comparing the grammatical performance of children across specific language learning conditions is critically needed in order to better understand the effects of subtractive bilingualism on first-language development. Language dominance must be carefully considered in future bilingual studies with these children as well. Grammatical assessments should examine performance across the two languages to make decisions about a child's dominant language. S-MST differences between monolinguals and bilinguals are likely to exist if children are tested and compared in their weaker language (e.g., if the Spanish of English-dominant bilinguals is compared with the Spanish of Spanish-only or Spanish-dominant speakers).

The performance of children from different Spanish dialects is based on a small sample of speakers of Caribbean dialects compared with a sample of Mexican American children, yet the preliminary results suggest that the S-MST is applicable to Central American and Caribbean children and may be used clinically with both dialect groups. However, it is important to note that scoring adaptations consistent with the dialect are necessary, so that children are not misclassified because of dialectal differences. Therefore, it is recommended that clinicians be well acquainted with the dialectal characteristics of the children they assess and that they make these types of dialectal adaptations.

In summary, the present study is the first validation of a measure specifically designed to assess morphosyntactic skills in Spanish-speaking children living in the United States for the purpose of identifying language disorders. This measure was developed targeting language forms that have significance for Spanish learners rather than forms adapted from English. Although the discrimination ability of the test does not fall under the good category across all three age groups according to Plante and Vance's (1994) guidelines, these results constitute a significant improvement for the use of a standardized measure to identify language impairment in the bilingual Latino population in the United States. The findings of the study indicate that the S-MST is appropriate for children ages 4–6 years of age; however, for older children, supplemental testing is necessary. In addition, the results of this study indicate that there are no differences in performance on this measure between children who are bilingual (and for whom Spanish is their best language) and children who speak only Spanish. Moreover, the measure appears to be equally sensitive across Caribbean and Mexican Spanish dialects. This research is a preliminary step toward validating specific measures for the identification of bilingual children with language impairment. Studies with the S-MST, in combination with other language tasks (e.g., Gutiérrez-Clellen et al., 2004), may help improve the accuracy of language assessments with these children and ultimately prevent their misdiagnosis.

Acknowledgments

This project was partially supported by National Institute on Deafness and Other Communication Disorders Grant 1-DC-8-2100 and by the National Institutes of Health/National Institute of General Medical Sciences San Diego State University Minority Biomedical Research Support Program (Grant 5 R25 GM58906-06). We are grateful to Aquiles Iglesias, Elizabeth Peña, Lisa Bedore, Brian Goldstein, Janet Calderón, and Ellen Stubbe Kester for their invaluable assistance with many phases of the data collection. We would also like to extend our thanks to Darrell Sabers for his expert advice in test development and to Lorena Cataño for her help with data management and analysis.

References

Anderson R. Assessing the grammar of Spanish-speaking children: A comparison of two procedures. Language, Speech, and Hearing Services in Schools. 1996; 27:333–344.

Anderson R. Lexical morphology and verb use in child first language loss: A preliminary case study investigation. International Journal of Bilingualism. 2001; 5:377–401.

Anderson, R. Language loss in Spanish-speaking children with and without SLI; Presentation at the Annual Convention of the American Speech-Language-Hearing Association; Chicago. 2003 Nov.

Auza A. La transparencia morfológica y conceptual en la adquisición de términos centrals y periféricos. Revista Latina de Pensamiento y Lenguaje. 2003; 2:207–223. Morphological and conceptual transparency in the acquisition of central and peripheral terms.

Bedore L, Leonard LB. Grammatical morphology deficits in Spanish-speaking children with specific language impairment. Journal of Speech, Language, and Hearing Research. 2001; 44:905–924.

Bosch L, Serra M. Grammatical morphology deficits of Spanish-speaking children with specific language impairment. Amsterdam Series in Child Language Development. 1997; 6:33–45.

Döpke, S. The interplay between language-specific development and cross-linguistic influence. In: Döpke, S., editor. Cross-linguistic structures in simultaneous bilingualism. Philadelphia: John Benjamins; 2000. p. 79-103.

Elliott, CD. Differential Ability Scales (DAS). San Antonio, TX: Harcourt Assessment; 1990.

Ellis Weismer S, Tomblin JB, Zhang X, Buckwalter P, Chynoweth J, Jones M. Nonword repetition performance in school-age children with and without language impairment. Journal of Speech, Language, and Hearing Research. 2000; 43:865–878.

Eng N, O'Connor B. Acquisition of definite articles + noun agreement of Spanish–English bilingual children with specific language impairment. Communication Disorders Quarterly. 2000; 21:114–124.

Gathercole VCM, Sebastián E, Soto P. The emergence of linguistic person in Spanish-speaking children. Language Learning. 2002; 52:679–722.

Gavin WJ, Klee T, Membrino I. Differentiating specific language impairment from normal language development using grammatical analysis. Clinical Linguistics & Phonetics. 1993; 7:191–206.

Gutiérrez-Clellen VF. Syntactic skills of Spanish-speaking children at risk for academic underachievement. Language, Speech, and Hearing Services in Schools. 1998; 29:207–215.

GutiérrezClellen VF, Calderón J, Ellis Weismer WS. Verbal working memory in bilingual children. Journal of Speech, Language, and Hearing Research. 2004; 47:863–876.

Gutiérrez-Clellen VF, Hofstetter R. Syntactic complexity in Spanish narratives: A developmental study. Journal of Speech and Hearing Research. 1994; 37:645–654. [PubMed: 8084194]

Gutiérrez-Clellen VF, Peña E. Dynamic assessment of diverse children: A tutorial. Language, Speech, and Hearing Services in Schools. 2001; 32:212–224.

Gutiérrez-Clellen VF, Restrepo MA, Bedore L, Peña E, Anderson R. Language sample analysis in Spanish speaking-children: Methodological considerations. Language, Speech, and Hearing Services in Schools. 2000; 31:88–98.

Gutiérrez-Clellen, VF.; Simón-Cereijido, G. Language sample measures for Spanish-speaking children with specific language impairment; Paper presented at the Symposium on Research in Child Language Disorders; Madison, WI. 2004 Jun.

Hunt, K. Grammatical structures written at three grade levels (Research Report No 3). Champaign, IL: National Council of Teachers of English; 1965.

Jacobson PF, Schwartz R. Morphology in incipient bilingual Spanish-speaking preschool children with specific language impairment. Applied Psycholinguistics. 2002; 23:23–41.

Leonard L, Eyer J, Bedore L, Grela B. Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. Journal of Speech, Language, and Hearing Research. 1997; 40:741–753.

Lipski, JM. The language of the isleños. Baton Rouge: Louisiana State University Press; 1990.

MacWhinney, B. The competition model. In: MacWhinney, B., editor. The 20th Annual Carnegie Symposium on Cognition: Mechanisms of language acquisition. Hillsdale, NJ: Erlbaum; 1987. p. 249-308.

MacWhinney, B. Second language acquisition and the competition model. In: Groot, AD.; Kroll, J., editors. Tutorials in bilingualism: Psycholinguistic perspectives. Hillsdale, NJ: Erlbaum; 1997. p. 113-142.

Masterson JJ, Kamhi AG. The effects of sampling conditions on sentence production in normal, reading-disabled, and language-learning disabled children. Journal of Speech and Hearing Research. 1991; 34:549–558. [PubMed: 2072679]

Mayer, M. Frog, where are you?. New York: Puffin Pied Pipper; 1969.

Mayer, M. Frog on his own. New York: Puffin Pied Pipper; 1973.

Mayer, M. Frog goes to dinner. New York: Puffin Pied Pipper; 1974.

Mayer, M. One frog too many. New York: Puffin Pied Pipper; 1975.

Paul, R. Language disorders from infancy through adolescence. 2. St. Louis, MO: Mosby; 2001.

Pearson B, Fernandez S, Lewedeg V, Oller K. The relation of input factors to lexical learning by bilingual infants. Applied Psycholinguistics. 1997; 18:41–58.

Peña E, Gutiérrez-Clellen VF, Iglesias A. Bilingual English–Spanish Assessment (BESA). 2006 Manuscript in preparation.

Peña E, Quinn R, Iglesias A. The application of dynamic methods to language assessment: A nonbiased procedure. Journal of Special Education. 1992; 26:269–280.

Plante E, Vance R. Selection of preschool language tests: A database approach. Language, Speech, and Hearing Services in the Schools. 1994; 25:15–23.

Ramírez, A. El español de los Estados Unidos [Spanish spoken in the United States]. Madrid: Editorial MAPFRE; 1992.

Redmond SM, Rice ML. Detection of irregular verb violations by children with and without SLI. Journal of Speech, Language, and Hearing Research. 2001; 44:655–669.

Restrepo, MA. Grammatical errors in Spanish-speaking children with language impairment who are learning English as a second language; Paper presented at the First International Symposium on Bilingualism; Newcastle upon Tyne, England. 1997 Apr.

Restrepo MA. Identifiers of predominantly Spanish-speaking children with language impairment. Journal of Speech, Language, and Hearing Research. 1998; 41:1398–1411.

Restrepo, MA. Spanish language skills in bilingual children with specific language impairment. In: Montrul, S.; Ordóñez, F., editors. Papers from the Fifth Hispanic Linguistics Symposium and the 2001 Acquisition of Spanish and Portuguese Conference: Linguistic theory and language development in Hispanic languages. Somerville, MA: Cascadilla Press; 2003. p. 365-374.

Restrepo MA, Gutiérrez-Clellen VF. Article production in bilingual children with specific language impairment. Journal of Child Language. 2001; 28:433–452. [PubMed: 11449946]

Restrepo, MA.; Gutiérrez-Clellen, VF. Grammatical impairments in Spanish/English-speaking children. In: Goldstein, B., editor. Language development: A focus on Spanish–English speakers. Baltimore: Brookes; 2004. p. 213-256.

Restrepo MA, Kruth K. Grammatical characteristics of a bilingual student with specific language impairment. Journal of Children's Communication Development. 2000; 21:66–76.

Restrepo MA, Silverman S. The validity of the Spanish Preschool Language Scale for use with bilingual children. American Journal of Speech-Language Pathology. 2001; 10:382–393.

Rice M, Wexler K. Toward tense as a clinical marker of specific language impairment in English-speaking children. Journal of Speech and Hearing Research. 1996; 39:1236–1257.

Rice M, Wexler K, Cleave P. Specific language impairment as a period of extended optional infinitive. Journal of Speech and Hearing Research. 1995; 38:850–863. [PubMed: 7474978]

Roca, A.; Lipski, JM., editors. Spanish in the United States: Linguistic contact and diversity. New York: Mouton de Gruyter; 1993.

Sanz Torrent M. Los verbos en niños con trastorno de lenguaje [Verbs in children with language disorders]. Revista de Logopedia, Foniatría y Audiología. 2002; XXI:100–110.

Semel, E.; Wiig, EH.; Secord, W. Clinical Evaluation of Language Fundamentals. 3. San Antonio, TX: The Psychological Corporation; 1997.

Silva-Corvalan, C., editor. Spanish in four continents: Studies in language contact and bilingualism. Washington, DC: Georgetown University Press; 1995.

Slobin DI, Bocaz A. Learning to talk about movement through time and space: The development of narrative abilities in Spanish and English. Lenguas Modernas. 1988; 15:5–24.

Stokes J, Krashen SD, Kartchner J. Factors in the acquisition of the present subjunctive in Spanish: The role of reading and study. I T L Review of Applied Linguistics. 1998; 121–122:19–25.

Tabachnick, BG.; Fidell, LS. Using multivariate statistics. New York: HarperCollins; 1989.

Torrens V, Wexler K. El retraso del lenguaje en la adquisición del Castellano y el Catalán [Language delay in the acquisition of Castillian and Catalán]. Aloma. 2001; 9:131–148.

Varela, B. El español Cubano–Americano [Cuban– American Spanish]. New York: Senda Nueva Ediciones; 1992.

Watkins RV, Rice ML, Moltz CC. Verb use by language impaired and normally developing children. First Language. 1993; 37:133–143.

Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scale:Third Edition Spanish. San Antonio, TX: The Psychological Corporation; 1993.

Zimmerman, IL.; Steiner, VG.; Pond, RE. Preschool Language Scale: Fourth Edition Spanish. San Antonio, TX: The Psychological Corporation; 2002.

Table 1

Percentage of families of the children with typical language development (TLD) and language impairment (LI) in each category of educational level, eligibility to lunch program, and Spanish dialect spoken.

Characteristic	TLD (*n* = 80)	LI (*n* = 80)
Educational level in the home		
Primary education	25	18
Some secondary education	33	40
High school graduate	32	20
Some college experience	7	20
College graduate	3	1
Some graduate school experience	0	0
Eligibility to lunch program		
Free	50	32
Reduced	4	12
Regular	46	56
Spanish dialect		
Mexican American	84	87
Puerto Rican and Dominican	14	13
Other	1	1

Table 2

Means and standard deviations (in parentheses) of amount of language input, and ratings of use and proficiency for the participants from San Diego, Austin, and Philadelphia ($N = 124$).

	Language ability			
Characteristic	**Typical**	**Impaired**	***p***	***d***
Proportion of Spanish input at home	.77 (.20)	.78 (.21)	.751	.06
Proportion of English input at home	.22 (.20)	.21 (.21)	.751	.06
Parents' rating of use of Spanish	3.90 (0.35)	3.82 (0.42)	.278	.19
Parents' rating of use of English	2.13 (0.97)	2.37 (1.20)	.227	.22
Parents' rating of proficiency in Spanish	3.88 (0.32)	3.60 (0.66)	.003	.54
Parents' rating of proficiency in English	1.95 (1.12)	1.95 (1.26)	.991	.00
Teachers' rating of use of Spanish	3.52 (0.97)	2.95 (1.13)	.008	.54
Teachers' rating of use of English	1.78 (1.25)	1.81 (1.13)	.890	.03
Teachers' rating of proficiency in Spanish	3.15 (1.44)	2.94 (1.16)	.422	.16
Teachers' rating of proficiency in English	1.57 (1.26)	1.96 (1.28)	.113	.31
Proportion of input at school in Spanish	.48 (.33)	.56 (.30)	.221	.26
Proportion of input at school in English	.37 (.31)	.37 (.28)	.986	.00

Note. Seventeen children were excluded because of incomplete questionnaire data.

Table 3

Means and standard deviations of S-MST scores of all children with TLD and LI across California, Texas, Pennsylvania, and Georgia sites.

	Language ability					
	TLD			LI		
Site	***n***	***M***	***SD***	***n***	***M***	***SD***
California	34	79	15	42	45	22
Texas	27	80	12	15	44	18
Pennsylvania	11	81	9	12	49	24
Georgia	8	71	15	11	34	13

Note. S-MST (Spanish Morphosyntax Test) scores represent percentage of correct items produced.

Table 4

Means and standard deviations of S-MST scores and percentages of children correctly classified as with TLD or LI based on exploratory discriminant analyses by age.

	Ability					
	TLD		LI			
Age group	***M***	***SD***	***M***	***SD***	**Sensitivity**	**Specificity**
Age Group 1 4;0–5;1 (*n* = 44)	67	15	32	16	86.4% (19/22)	86.4% (19/22)
Age Group 2 5;2–5;11 (*n* = 36)	85	8	50	17	94.4% (17/18)	94.4% (17/18)
Age Group 3 6;1–7;0 (*n* = 36)	84	11	48	21	72.2% (13/18)	83.3% (15/18)

Note. S-MST scores represent percentage of correct items produced.

Table 5

Canonical discriminant function coefficients across age groups.

Age group	Constant	Coefficient × (Score)	Cutoff score
Age Group 1 4;0–5;1	−3.236	6.499 × (*x*)	50
Age Group 2 5;2–5;11	−4.500	5.734 × [arcsine (*x*)]	70
Age Group 3 6;1–7;0	−3.393	4.392 × [arcsine (*x*)]	70

Note. The formula used to derive the cutoff score for age 1 is the following:

$$0=\text{constant}+\text{coefficient}\ (x),$$

where x is the value of the cutoff score. Given the significant differences in equality of variance–covariance matrices for Age Groups 2 and 3, the cutoff scores were derived by using the following formula:

$$0=\text{constant}+\text{coefficient}\ [\text{arcsine}\ (x)]$$

The value of x was solved by using the following formula:

$$x=\sin\ [\text{arcsine}\ (x)].$$

Table 6

Means and standard deviations of S-MST scores and percentage of children correctly classified as with TLD or LI based on confirmatory discriminant analyses by age.

	Ability					
	TLD		LI			
Age group	***M***	***SD***	***M***	***SD***	**Sensitivity**	**Sensitivity**
Age Group 1 4;0 to 5;1 (*n* = 16)	77	15	33	22	87.5% (7/8)	100% (8/8)
Age Group 2 5;2 to 5;11 (*n* = 14)	84	6	44	13	100% (7/7)	100% (7/7)
Age Group 3 6;1 to 7;0 (*n* = 14)	88	6	65	25	42.8% (3/7)	100% (7/7)

Note. S-MST scores represent percentage of correct items produced.

Table 7

Means and standard deviations of S-MST scores by age-matched Spanish-only proficient and Spanish-dominant bilingual children with TLD.

	Spanish-only proficient (*n* = 13)	Spanish-dominant bilingual (*n* = 13)		
Characteristic	***M (SD)***	***M (SD)***	***p***	***d***
Age in months	74.23 (6.65)	75.08 (6.96)		
S-MST score	84 (10)	82 (09)	.614	.21
Articles	89 (14)	90 (13)	.843	.07
Clitics	73 (24)	76 (21)	.748	.13
Verbs	76 (20)	80 (23)	.605	.19

Note. S-MST scores represent percentage of correct items produced.

Table 8

Distribution of percentage of grammatical utterances by age, bilingual status, and language ability of the California, Texas, and Pennsylvania samples.

	Spanish language sample				English language sample			
	Language ability				Language ability			
	Typical		Impaired		Typical		Impaired	
Age group	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
Age Group 1 4;0–5;1								
Monolingual (*n* = 31)	90	5	64	16	—	—	—	—
Bilingual (*n* = 4)	88	4	70	5	65	10	57	—
Age Group 2 5;2–5;11								
Monolingual (*n* = 30)	92	5	78	8	—	—	—	—
Bilingual (*n* = 12)	88	5	78	16	63	10	55	15
Age Group 3 6;1–7;0								
Monolingual (*n* = 25)	92	4	80	8	—	—	—	—
Bilingual (*n* = 20)	89	5	79	18	61	13	60	19

Note. Dashes indicate data are not applicable.